Pacific Journal of Mathematics

Volume 302 No. 1 September 2019

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 2019 is US $490/year for the electronic version, and $665/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 1945-5844 electronic, 0030-8730 printed) 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. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.1

ON MASAS IN q-DEFORMED VON NEUMANN ALGEBRAS

MARTIJN CASPERS, ADAM SKALSKI AND MATEUSZ WASILEWSKI

We study certain q-deformed analogues of the maximal abelian subalgebras of the group von Neumann algebras of free groups. The radial subalgebra is defined for Hecke deformed von Neumann algebras of the Coxeter group $(\mathbb{Z}/2\mathbb{Z})^{\star k}$ and shown to be a maximal abelian subalgebra which is singular and with Pukánszky invariant $\{\infty\}$. Further all nonequal generator masas in the q-deformed Gaussian von Neumann algebras are shown to be mutually nonintertwinable.

1. Introduction

Our aim is to investigate maximal abelian subalgebras in certain II_1-factors that can be viewed as deformations of $\mathrm{VN}(\mathbb{F}_n)$. Our particular interest lies in the analysis of counterparts of the radial masa A_r in $\mathrm{VN}(\mathbb{F}_n)$, studied for example in [Boca and Rădulescu 1992] and in [Cameron et al. 2010] (see also [Trenholme 1988]). The main open problem concerning the radial masa in $\mathrm{VN}(\mathbb{F}_n)$ is the question whether it is isomorphic to the generator masa(s); so far they share all the known properties, such as maximal injectivity, the same Pukánszky invariant, etc. They are also known not to be unitarily conjugate (see Proposition 3.1 of [Cameron et al. 2010]). More generally, radial masas have been studied for von Neumann algebras of groups of the type $(\mathbb{Z}/_{n\mathbb{Z}})^{\star k}$ in [Trenholme 1988] and [Boca and Rădulescu 1992].

Here we want to analyse the behaviour of counterparts of the radial/generator masa in some deformed versions of $\mathrm{VN}(\mathbb{F}_n)$ or $\mathrm{VN}((\mathbb{Z}/_{n\mathbb{Z}})^{\star k})$; more specifically in Hecke deformed von Neumann algebras of right-angled Coxeter groups $\mathrm{VN}_q(W)$ of Dymara [2006] (see also [Garncarek 2016] and [Caspers 2016a]) and in q-deformed Gaussian von Neumann algebras $\Gamma_q(\mathcal{H}_\mathbb{R})$ of Bożejko, Kümmerer and Speicher [Bożejko et al. 1997]. In the former case we can naturally define the radial subalgebra (and not the generator one), and in the latter the object that intuitively corresponds to the radial subalgebra is in fact obviously isomorphic to the generator one (as studied by Ricard [2005] and further by Wen [2017] and Parekh, Shimada

MSC2010: primary 46L10; secondary 46L65.
Keywords: maximal abelian subalgebras, singular masas, Hecke von Neumann algebra, q-Gaussian algebras.

and Wen [Parekh et al. 2018]). We show in Section 4 however that the different generator masas inside the $\Gamma_q(\mathcal{H}_{\mathbb{R}})$ are not unitarily conjugate.

Note that another example of a counterpart of the radial subalgebra in $\mathrm{VN}(\mathbb{F}_n)$ was studied and shown to be maximal abelian and singular in [Freslon and Vergnioux 2016]. It was a von Neumann subalgebra of the algebra $L^\infty(O_N^+)$, which shares many properties with $\mathrm{VN}(\mathbb{F}_n)$, although very recently the latter two were shown to be nonisomorphic [Brannan and Vergnioux 2018].

The plan of the paper is as follows: after finishing this section introducing certain notation, in Section 2 we define the radial subalgebra of the Hecke deformed von Neumann algebra $\mathrm{VN}_q(W)$ and show it to be maximal abelian. In Section 3, we compute its Pukánszky invariant and deduce its singularity. Finally Section 4 discusses the nonintertwinability of (a continuous family of) different generator masas in the q-deformed Gaussian von Neumann algebras.

Notation. Throughout this paper, by a *masa* we mean a maximal abelian von Neumann subalgebra of a given von Neumann algebra M. Let $U(\mathsf{M})$ be the group of unitaries in M. For a (unital) subalgebra $\mathsf{A} \subseteq \mathsf{M}$ we define the *normalizer* of A in M as

$$N_{\mathsf{M}}(\mathsf{A}) = \{u \in U(\mathsf{M}) \mid u\mathsf{A}u^* \subseteq \mathsf{A}\}.$$

A subalgebra $\mathsf{A} \subseteq \mathsf{M}$ is called *singular* if $N_{\mathsf{M}}(\mathsf{A}) \subseteq \mathsf{A}$. These notions were first introduced by Dixmier [1954].

$\mathbb{N}_0$ denotes the natural numbers including 0.

2. The radial Hecke masa

In this section we show that right-angled Hecke von Neumann algebras admit a radial algebra and prove that it is in fact a masa.

Let W denote a *right-angled Coxeter group*. Recall that this is the universal group generated by a finite set S of elements of order 2, with the relations forcing some of the distinct elements of S to commute, and some other to be free. This is formally encoded by a function

$$m : S \times S \setminus \{(s, s) : s \in S\} \to \{2, \infty\}$$

such that for all $s, t \in S$, $s \neq t$ we have

$$(st)^{m(s,t)} = e$$

(and $(st)^\infty = e$ means that s and t are free; necessarily $m(s, t) = m(t, s)$). We will always associate to W the length function $|\cdot| : W \to \mathbb{N}_0$ given by the generating set S. All the information about W is encoded by a graph Γ with a vertex set $V\Gamma = S$ and the edge set $E\Gamma = \{(s, t) \in S \times S : m(s, t) = 2\}$. Let $q \in (0, 1]$ and put $p = (q - 1)/q^{1/2}$ (note that our convention on q means that $p \leq 0$). The algebra $\mathbb{C}_q[W]$ is a *-algebra with a linear basis $\{T_w : w \in W\}$ satisfying the conditions ($s \in S$, $w \in W$)

$$T_s T_w = \begin{cases} T_{sw} & \text{if } |sw| > |w|, \\ T_{sw} + pT_w & \text{if } |sw| < |w|. \end{cases}$$

The algebra $\mathbb{C}_q[W]$ acts in a natural way (via bounded operators) on the space $\ell^2(W)$ and its von Neumann algebraic closure in $B(\ell^2(W))$ will be denoted by $\mathrm{VN}_q(W)$. The vector $\delta_e \in \ell^2(W)$ will sometimes be denoted by Ω; the corresponding vector state $\tau := \omega_\Omega$ on $\mathrm{VN}_q(W)$ is a faithful trace. More generally to any element $T \in \mathrm{VN}_q(W)$ we can associate its *symbol* $T\Omega$, and as Ω is a separating vector for $\mathrm{VN}_q(W)$ this correspondence is injective. Finally note that using the right action of the Hecke algebra on itself, we can define another von Neumann algebra acting on $\ell^2(W)$, say $\mathrm{VN}_q(W)^r$. It is obviously contained in the commutant of $\mathrm{VN}_q(W)$; in fact Proposition 19.2.1 of [Davis 2008] identifies it with $\mathrm{VN}_q(W)'$.

In what follows, we will write L to denote the cardinality of S.

Hecke von Neumann algebras were first considered in [Dymara 2006] and [Davis et al. 2007] in order to study weighted L^2-cohomology of Coxeter groups. In [Davis et al. 2007], the authors raised a natural question: how large is the centre of $\mathrm{VN}_q(W)$? A precise answer for the right-angled case was found in [Garncarek 2016], where the following result was shown.

Theorem 2.1. *Let $|S| \geq 3$ and assume that Γ is irreducible. Then for $q \in [\rho, 1]$ the right-angled Hecke von Neumann algebra $\mathbb{C}_q[W]$ is a II_1-factor and for $(0, \rho)$ we have that $\mathbb{C}_q[W]$ is a direct sum of a II_1-factor and $\mathbb{C}$. Here ρ is the radius of convergence of the fundamental power series $\sum_{k=0}^{\infty} \big|\{w \in W \mid |w| = k\}\big| z^k$.*

In particular $\mathrm{VN}_q(W)$ is diffuse if and only if $q \in [\rho, 1]$. Further structural results were obtained in [Caspers 2016a; 2016b; Caspers and Fima 2017] where for example noninjectivity, approximation properties, absence of Cartan subalgebras, the Connes embedding property and the existence of graph product decompositions were established for $\mathrm{VN}_q(W)$.

In this paper we consider the special case $W = (\mathbb{Z}_2)^{*L}$, i.e., the case where m is constantly equal to infinity. We assume also that $L \geq 3$. Here the main result of [Garncarek 2016] (see Theorem 2.1) says that $\mathrm{VN}_q(W)$ is a factor if and only if $q \in \left[\frac{1}{L-1}, 1\right]$, and results of [Dykema 1993] together with a calculation in Section 5 of [Garncarek 2016] show that for that range of q we have

$$\mathrm{VN}_q(W) \approx \mathrm{VN}(\mathbb{F}_{2Lq/(1+q)^2}),$$

where $\mathrm{VN}(\mathbb{F}_s)$ for $s \geq 1$ denote the interpolated free group factors of Dykema and Radulescu.

Definition 2.2. An element $T \in \mathrm{VN}_q(W)$ is said to be radial if for its symbol decomposition $T\Omega = \sum_{w\in W} c_w\delta_w$, where $c_w \in \mathbb{C}$, we have $c_w = c_v$ for every $v, w \in W$ with $l(v) = l(w)$. We say that T has radius (at most) n if the frequency support (i.e., the set of those $w \in W$ for which $c_w \neq 0$) of T_w is contained in the ball $\{w \in W : |w| \leq n\}$.

Define $h \in \mathbb{C}_q[W] \subset \mathrm{VN}_q(W)$ by the formula $h = \sum_{s\in S} T_s$ and put $\mathsf{B} := \{h\}''$.

Proposition 2.3. *The von Neumann algebra* B *coincides with the collection of all radial operators in* $\mathrm{VN}_q(W)$*. In particular the set of all radial operators forms an algebra.*

Proof. For each $n \in \mathbb{N}$ consider the radial operator $h_n := \sum_{w\in W,\, |w|=n} T_w \in \mathbb{C}_q[W]$ and put $h_0 := I$.

For each $n \in \mathbb{N},\ n \geq 2$, we have

$$hh_n = \sum_{s\in S} \sum_{\substack{|w|=n,\\ |sw|>|w|}} T_sT_w + \sum_{s\in S} \sum_{\substack{|w|=n,\\ |sw|<|w|}} T_sT_w \tag{2-1}$$
$$= \sum_{s\in S} \sum_{\substack{|w|=n,\\ |sw|>|w|}} T_{sw} + \sum_{s\in S} \sum_{\substack{|w|=n,\\ |sw|<|w|}} T_{sw} + \sum_{s\in S} \sum_{\substack{|w|=n,\\ |sw|<|w|}} pT_w$$
$$= h_{n+1} + (L-1)h_{n-1} + ph_n.$$

We also have $h^2 = h_2 + ph + Lh_0$. This shows in particular that the algebra generated by h consists of radial operators. Moreover viewing the above as a recurrence formula we see that each h_n can be expressed as a polynomial in h and I, so that the subspace A generated by $\{h_n : n \in \mathbb{N}\}$ coincides with the unital *-algebra generated by h.

Further define the radial subspace

$$\ell^2(W)_r := \{(c_w)_{w\in W} \in \ell^2(W) : c_v = c_w \text{ for all } w, v \in W, |w| = |v|\}$$

and denote the orthogonal projection from $\ell^2(W)$ onto $\ell^2(W)_r$ by P_r. It is easy to see that $A\Omega$ is norm dense in $\ell^2(W)_r$. Thus the unique trace-preserving conditional expectation $\mathbb{E}$ onto $A'' \subset \mathrm{VN}_q(W)$ is given by the formula

$$\mathbb{E}(T)\Omega = P_rT\Omega, \quad T \in \mathrm{VN}_q(W).$$

This shows that the set of radial operators in $\mathrm{VN}_q(W)$ coincides with A'' and passing now to ultraweak closures we see that h generates the von Neumann algebra of all radial operators. □

Note that the above fact is not true (even for $p = 0$) for a general right-angled Coxeter group. Also note that formulae such as (2-1) (and the subsequent line in the proof) play a very relevant role in our proof of singularity in Section 3.

The first main theorem of this paper is based on the idea of Pytlik for the radial algebra in $\mathrm{VN}(\mathbb{F}_n)$ [1981]; see also [Sinclair and Smith 2008]. By $R_h \in \mathrm{VN}_q(W)^r$, we understand the operator on $\ell^2(W)$ given by the *right* action of $\sum_{s\in S} T_s$.

Lemma 2.4. *For every $v, w \in W$ with $|v| = |w|$ and for every $\varepsilon > 0$ there exists a vector $\eta \in \ell^2(W)$ such that*

$$\|e_v - e_w - (h\eta - R_h\eta)\|_2 < \varepsilon.$$

Proof. We first assume that $w = az$ and $v = zb$ for some word $z \in W$ with $|z| = |v| - 1$ and some letters $a, b \in S$. In the proof x and y will always be words in W and summations are always over x and y. Put for $k \in \mathbb{N}$

$$\psi_k = \sum_{\substack{|x|=|y|=k,\\ |xa|=|by|=k+1}} e_{xazby} \in \ell^2(W),$$

and define also $\psi_0 = e_{azb}$. Let $\delta > 0$. As for each $k \in \mathbb{N}$ there are $L(L-1)^{k-1}$ reduced words in W of length k,

$$\left\|\left(\frac{1-\delta}{L-1}\right)^k \psi_k\right\|_2^2 \le \left(\frac{1-\delta}{L-1}\right)^{2k} (L-1)^{2k-2} L^2 \le 4(1-\delta)^{2k}. \tag{2-2}$$

This means that we can define

$$\eta_\delta = \sum_{k=0}^{\infty} \left(\frac{1-\delta}{L-1}\right)^k \psi_k \in \ell^2(W).$$

We claim that the vector η_δ, for δ small enough (dependent on ε), satisfies the condition of the lemma. To show that we need to analyse the actions of h and R_h on ψ_k. For $k \ge 1$ we have (the bracket term included; the brackets are there in order to define further vectors in the remainder of the proof)

$$h\psi_k = \sum_{s\in S} \sum_{\substack{|x|=|y|=k,\ |sx|=k+1\\ |xa|=|by|=k+1}} e_{sxazby} + \sum_{s\in S} \sum_{\substack{|x|=|y|=k,\ |sx|=k-1\\ |xa|=|by|=k+1}} e_{sxazby} (+p e_{xazby}). \tag{2-3}$$

and similarly, for $k \ge 1$,

$$R_h\psi_k = \sum_{s\in S} \sum_{\substack{|x|=|y|=k,\ |ys|=k+1\\ |xa|=|by|=k+1}} e_{xazbys} + \sum_{s\in S} \sum_{\substack{|x|=|y|=k,\ |ys|=k-1\\ |xa|=|by|=k+1}} e_{xazbys} (+p e_{xazby}). \tag{2-4}$$

Finally

$$h\psi_0 = e_{zb} + pe_{azb} + \sum_{s\in S\setminus\{a\}} e_{sazb}, \qquad R_h\psi_0 = e_{az} + pe_{azb} + \sum_{s\in S\setminus\{b\}} e_{azbs}. \tag{2-5}$$

We now analyse the "commutators" $h\psi_k - R_h\psi_k$ and their sum. Note first that for each $k \in \mathbb{N}_0$ the summand in $h\psi_k$ given by pe_{xazby} also occurs in $R_h\psi_k$.

We define (compare to (2-5))

$$\phi_{1,0} = \sum_{s\in S\setminus\{a\}} e_{sazb}, \quad \phi_{2,0} = e_{zb}, \quad \chi_{1,0} = \sum_{s\in S\setminus\{b\}} e_{azbs}, \chi_{2,0} = e_{az}.$$

For $k \geq 1$ we set the following notation: let $\phi_{1,k}$ and $\phi_{2,k}$ be the two large sums on, respectively, the first and second line of (2-3), without the vectors between brackets. Similarly we define $\chi_{1,k}$ and $\chi_{2,k}$ to be the two large sums on, respectively, the first and second line of (2-4), without the vectors between brackets.

Then we have for all $k \in \mathbb{N}_0$

$$\phi_{1,k} = \frac{1}{L-1}\chi_{2,k+1}, \qquad \chi_{1,k} = \frac{1}{L-1}\phi_{2,k+1},$$

so that

$$\phi_{1,k} - \frac{1-\delta}{L-1}\chi_{2,k+1} = \delta\phi_{1,k}, \qquad \chi_{1,k} - \frac{1-\delta}{L-1}\phi_{2,k+1} = \delta\chi_{1,k}.$$

Thus a version of the telescopic argument yields the equality

$$\begin{aligned} h\eta_\delta - R_h\eta_\delta &= \sum_{k=0}^{\infty}\left(\frac{1-\delta}{L-1}\right)^k (\phi_{1,k} + \phi_{2,k} - \chi_{1,k} - \chi_{2,k}) \\ &= e_{zb} - e_{az} + \delta\left(\sum_{k=1}^{\infty}\left(\frac{1-\delta}{L-1}\right)^k (\phi_{1,k} - \chi_{1,k})\right). \end{aligned}$$

As $\delta \searrow 0$ this can be shown via a similar ℓ^2-counting estimate to that above to converge in norm to $e_{zb} - e_{az}$. From this we conclude the claim.

For general $v = v_1 \dots v_n$ and $w = w_1 \dots w_n$ with $v_n \neq w_1$ the proposition follows from a triangle inequality and an application of the argument in the first part of the proof to each pair $w_k \dots w_n v_1 \dots v_{k-1}$ and $w_{k+1} \dots w_n v_1 \dots v_k$. In the case where $v_n = w_1$ one can apply the above to the pairs $v_k \dots v_n b w_1 \dots w_{k-2}$ and $v_{k+1} \dots v_n b w_1 \dots w_{k-1}$ for some letter $b \neq v_n$. □

We are ready to formulate the first main result in this section.

Theorem 2.5. *The radial algebra* B *is a masa in* $\mathrm{VN}_q(W)$.

Proof. Suppose that $T \in \mathsf{B}' \cap \mathrm{VN}_q(W)$ and write $T\Omega = \sum_{u\in W} c_u e_u$. Let $v, w \in W$ with $|v| = |w|$, let $\varepsilon > 0$ and let η be as in Lemma 2.4. Note that as T commutes with h we have $\langle T\Omega, h\eta - R_h\eta\rangle = \langle (hT - R_hT)\Omega, \eta\rangle = \langle T(h - R_h)\Omega, \eta\rangle = 0$. Then we get

$$|\langle T\Omega, e_v - e_w\rangle| \leq |\langle T\Omega, e_v - e_w + h\eta - R_h\eta\rangle| \leq \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, we see that $c_w = c_v$. Thus T is radial, which is equivalent to the fact that $T \in \mathsf{B}$ by Proposition 2.3. □

Remark 2.6. The recurrence formula (2-1) allows us to compute explicitly the distribution of h with respect to the canonical trace. As the formula (2-1) is valid only from $n = 2$ we first define "new" h_0 as $L/\widetilde{L}$, where $\widetilde{L} := L - 1$, so that with respect to the new variables it holds for all $n \in \mathbb{N}$. For simplicity assume that $q \in [1/\widetilde{L}, 1]$, so that $\mathrm{VN}_q(W)$ is a (finite) factor. Then the distribution of h is continuous (as B is diffuse) and the main result of [Cohen and Trenholme 1984] implies that the corresponding density is given (up to a normalising factor) by

$$\frac{\widetilde{L}\sqrt{4\widetilde{L} - (x-p)^2}}{\pi[-(x-p)^2 - p(2-L)(x-p) + p^2(L-1) + L^2]}dx.$$

Note that for $p = 0$ we obtain, as expected, the distribution of the radial element in the group $(\mathbb{Z}_2)^{*L}$ as computed in Theorem 4 of [Cohen and Trenholme 1984].

3. The Pukánszky invariant and singularity of the Hecke MASA

The Pukánszky invariant $\mathcal{P}(\mathsf{A})$ of a masa $\mathsf{A} \subseteq \mathsf{M}$ is determined by the von Neumann algebra generated by all A-A bimodule homomorphisms of $L^2(\mathsf{M})$. We refer to [Sinclair and Smith 2008] for further discussion of $\mathcal{P}(\mathsf{A})$. Popa [1985] showed that the Pukánszky invariant can be used to prove singularity of certain masas (and indeed this was successfully applied by Radulescu [1991] in order to obtain singularity of the radial masa in $\mathrm{VN}(\mathbb{F}_n)$). We will use this strategy in this section, following very closely the proof of [Rădulescu 1991], to show that the Hecke radial masa discussed in Section 2 is singular. In particular we determine its Pukánszky invariant.

We need some terminology. Let again $L \geq 3$, $W = (\mathbb{Z}_2)^{*L}$, $q \in \left[\frac{1}{L-1}, 1\right]$ and let B be the radial subalgebra of the factor $\mathrm{VN}_q(W)$ (shown to be a masa in Theorem 2.5).

Definition 3.1. The Pukánszky invariant of $\mathsf{B} \subseteq \mathrm{VN}_q(W)$ is defined as the type of the von Neumann algebra $\langle h, R_h\rangle' \subseteq \mathsf{B}(\ell^2(W))$, where h and R_h were defined in Section 2.

Next we introduce the necessary notation in order to determine the Pukánszky invariant of $\mathsf{B} \subseteq \mathrm{VN}_q(W)$. We need to construct certain bases, which are inspired by Radulescu's bases in free group factors (see [Rădulescu 1991]). For $l \in \mathbb{N}_0$

let $q_l : \mathbb{C}_q[W] \to \mathbb{C}_q[W]$ be the natural projection onto the span of $\{T_w, |w| = l\}$. Write $\mathbb{C}_q^l[W] = q_l(\mathbb{C}_q[W])$. As before set $h_l = \sum_{|w|=l} T_w$. We have for $m \geq 1$ (see (2-1) and its subsequent line)

$$h_1 h_m = h_m h_1 = h_{m+1} + p h_m + (L_m - 1) h_{m-1}, \tag{3-1}$$

where $L_m = L$ if $m \geq 2$ and $L_m = L+1$ if $m = 1$. Let

$$S_l = \operatorname{span}\{q_l(h_1 x), q_l(x h_1) \mid x \in q_{l-1}(\mathbb{C}_q[W])\};$$

in particular $S_1 = \mathbb{C} h_1$. Further for $l \in \mathbb{N},\ \gamma \in \mathbb{C}_q^l[W]$, set

$$\gamma_{m,n} = q_{m+n+l}(h_m \gamma h_n), \qquad m, n \in \mathbb{N}_0.$$

We also set $\gamma_{m,n} = 0$ in case $m < 0$ or $n < 0$. Finally for $l \in \mathbb{N}$ and $\gamma \in \mathbb{C}_q^l[W] \ominus S_l$ set

$$X_\gamma = \overline{\operatorname{span}}^{\|\ \|_2}\{\gamma_{m,n} \mid m, n \in \mathbb{N}_0\} \subset \ell^2(W).$$

Lemma 3.2 collects all computational results we need in what follows. As all the (rather easy) arguments are basically contained in [Rădulescu 1991, Lemma 1] we merely sketch the proof; all other proofs we give in this section will then be self-contained.

Lemma 3.2. (1) *For $\gamma \in \mathbb{C}_q^l[W],\ l \geq 1,\ m \geq 1,\ n \geq 0$, we have*

$$h_1 \gamma_{m,n} = \gamma_{m+1,n} + p\gamma_{m,n} + (L-1)\gamma_{m-1,n}.$$

(2) *For $\gamma \in \mathbb{C}_q^l[W] \ominus S_l,\ l \geq 2,\ m \geq 0,\ n \geq 0$, we have*

$$h_1 \gamma_{m,n} = \gamma_{m+1,n} + p\gamma_{m,n} + (L-1)\gamma_{m-1,n}.$$

(*Note that only the case $m = 0$ was not already covered by* (1)).

(3) *For $\beta \in \mathbb{C}_q^1[W] \ominus S_1,\ n \geq 0$, we have*

$$h_1 \beta_{0,n} = \beta_{1,n} + p\beta_{0,n} - \beta_{0,n-1}.$$

(4) *For $\gamma \in \mathbb{C}_q^l[W],\ l \geq 1$ we have*

$$q_{l+m+n+1}(h_1 h_m \gamma h_n) = q_{l+n+m+1}(h_1 q_{l+m+n}(h_m \gamma h_n)), \quad m, n \in \mathbb{N},$$
$$q_{l-m-n-1}(h_1 h_m \gamma h_n) = q_{l-m-n-1}(h_1 q_{l-m-n}(h_m \gamma h_n)), \quad 0 \leq m+n \leq l.$$

(5) *For $\gamma \in \mathbb{C}_q^l[W],\ l \geq 1$, we have $q_l(h_1 q_{l+1}(h_1 \gamma)) = (L-1)\gamma$.*

(6) *For $\beta \in \mathbb{C}_q^1[W] \ominus S_1$, we have $q_n(h_1 q_{n+1}(\beta h_n)) = -q_n(\beta h_{n-1})$.*

(7) *For all $\gamma \in \mathbb{C}_q^l[W] \ominus S_l,\ l \geq 2, n \in \mathbb{N},\ m \geq 1$, we have*

$$q_l(q_{m+n+l}(h_m \gamma h_n) h_{m+n}) = 0.$$

Proof. The proofs of (1)–(2) are easy consequences of (3-1); see also [Rădulescu 1991, Lemma 1 (a) and (b)]. The proof of (3) is essentially the same as [Rădulescu 1991, Lemma 1 (c)]. Statement (4) is a direct consequence of (3-1), and (5) and (6) follow from (1) and (3), respectively. Statement (7) follows from (1) and (2). □

The following theorem gives the cornerstone in our computation of the Pukánszky invariant. The idea is based on first showing that for suitable β and γ the mapping $T : X_\beta \to X_\gamma$ defined by the formula (3-2) is bounded and invertible. Then one uses a basis transition to the respective bases $\{h_m \beta h_n\}_{m,n\in\mathbb{N}}$ and $\{h_m \gamma h_n\}_{m,n\in\mathbb{N}}$ to show that T is actually a B-B bimodule map.

Theorem 3.3. *Let $l \in \mathbb{N}$, $l \geq 2$, let $\beta \in \mathbb{C}^1_q[W] \ominus S_1$ and let $\gamma \in \mathbb{C}^l_q[W](W) \ominus S_l$. Then the following hold*:

(1) *There exists a bounded invertible linear map $T : X_\beta \to X_\gamma$ determined by*

$$T : \beta_{m,n} \mapsto \gamma_{m,n} + \gamma_{m-1,n-1}, \qquad m, n \in \mathbb{N}_0. \tag{3-2}$$

(2) *We have $X_\beta = \overline{\mathsf{B}\beta\mathsf{B}}^{\|\ \|_2}$ and $X_\gamma = \overline{\mathsf{B}\gamma\mathsf{B}}^{\|\ \|_2}$. Moreover the map T defined by* (3-2) *agrees with the linear map*

$$T : h_m \beta h_n \mapsto h_m \gamma h_n, \qquad m, n \in \mathbb{N}_0. \tag{3-3}$$

The proof of Theorem 3.3 proceeds through a couple of lemmas, which we prove in two separate subsections.

Proof of Theorem 3.3 (1). The first statement of Theorem 3.3 is essentially a consequence of the following orthogonality property.

Lemma 3.4. *Let $l \in \mathbb{N}$, $l \geq 2$, and let $\beta, \beta' \in \mathbb{C}^1_q[W] \ominus S_1$, $\gamma \in \mathbb{C}^l_q[W] \ominus S_l$, $\gamma' \in \mathbb{C}^l_q[W]$, $l \geq 2$. We have then for each $m, n, m', n' \in \mathbb{N}_0$*

$$\langle \beta_{m,n}, \beta'_{m',n'} \rangle = \delta_{m+n,n'+m'} (L-1)^{m+n-|n-n'|} (-1)^{|n-n'|} \langle \beta, \beta' \rangle; \tag{3-4}$$

similarly,

$$\langle \gamma_{m,n}, \gamma'_{m',n'} \rangle = \delta_{m,m'} \delta_{n,n'} (L-1)^{m+n} \langle \gamma, \gamma' \rangle. \tag{3-5}$$

Proof. Let us first prove (3-5). Firstly, as $\gamma_{m,n}$ (resp. $\gamma'_{m',n'}$) is in the range of q_{m+n+l} (resp. $q_{m'+n'+l}$), we must have $m+n = m'+n'$ or else both sides of (3-5) are nonzero. We claim that

$$q_l (h_{m'} q_{m+n+l} (h_m \gamma h_n) h_{n'}) = \delta_{m,m'} \delta_{n,n'} (L-1)^{m+n} \gamma. \tag{3-6}$$

For $k := m+n = 0$ this is obvious. We proceed by induction on k and assume the assertion for $k-1$. For $k \geq 1$ one of m and n is nonzero and we may assume without loss of generality that $m \neq 0$ (the proof for n can be done in the same way, or one can consider the adjoint of (3-6) which interchanges the roles of m and n).

If the left-hand side of (3-6) is nonzero, then we must have that m' is nonzero, because otherwise this expression reads $q_l(q_{m+n+l}(h_m\gamma h_n)h_{n+m})$ which is zero by Lemma 3.2 (7).

Using (3-1) together with the fact that $q_l(h_r q_{m+n+l}(x)h_n)=0$ for every $r<m$ and $x\in\mathbb{C}_q[W]$ and $q_{m+n+l}(h_s\gamma h_n)=0$ for $s<m$, we get

$$q_l(h_{m'}q_{m+n+l}(h_m\gamma h_n)h_{n'})=q_l(h_{m'-1}h_1q_{m+n+l}(h_1h_{m-1}\gamma h_n)h_{n'}).$$

Using Lemma 3.2 (4) and (5) for the first two of the following equalities and then the induction hypothesis yields

$$\begin{aligned}(3\text{-}7)\quad q_l(h_{m'}q_{m+n+l}(h_m\gamma h_n)h_{n'})&\\ &=q_l(h_{m'-1}q_{m+n+l-1}(h_1q_{m+n+l}(h_1q_{m+n+l-1}(h_{m-1}\gamma h_n))h_{n'})\\ &=(L-1)q_l(h_{m'-1}q_{m+n+l-1}(h_{m-1}\gamma h_n)h_{n'})\\ &=(L-1)(L-1)^{m+n-1}\delta_{m,m'}\delta_{n,n'}\gamma.\end{aligned}$$

This completes the proof of (3-6). Then using the fact that $h_{m'}$ and $h_{n'}$ are self-adjoint we get

$$\begin{aligned}(3\text{-}8)\qquad \langle\gamma_{m,n},\gamma'_{m',n'}\rangle&=\langle q_{m+n+l}(h_m\gamma h_n),q_{m'+n'+l}(h_{m'}\gamma' h_{n'})\rangle\\ &=\langle h_{m'}q_{m+n+l}(h_m\gamma h_n)h_{n'},\gamma'\rangle\\ &=\langle q_l(h_{m'}q_{m+n+l}(h_m\gamma h_n)h_{n'}),\gamma'\rangle\\ &=(L-1)^{m+n}\delta_{m,m'}\delta_{n,n'}\langle\gamma,\gamma'\rangle.\end{aligned}$$

Next we sketch the proof of (3-4); it is largely the same as (3-5). The claim (3-6) gets replaced by

$$(3\text{-}9)\quad q_l(h_{m'}q_{m+n+l}(h_m\beta h_n)h_{n'})=(L-1)^{|m+n|-|n-n'|}(-1)^{|n-n'|}\delta_{m+n,m'+n'}\beta.$$

Again the proof proceeds by induction with respect to $k:=m+n=m'+n'$. The case $k=0$ is obvious so assume $k\geq 1$. First assume that both $m,m'\geq 1$. Similar to (3-7) and using the same results from Lemma 3.2 we find that

$$\begin{aligned}(3\text{-}10)\quad q_l(h_{m'}q_{m+n+l}(h_m\beta h_n)h_{n'})&=q_l(h_{m'-1}h_1q_{m+n+l}(h_1h_{m-1}\beta h_n)h_{n'})\\ &=(L-1)q_l(h_{m'-1}q_{m+n+l-1}(h_{m-1}\beta h_n)h_{n'-1})\\ &=(L-1)^{m+n-|n-n'|}(-1)^{|n-n'|}\delta_{m+n,m'+n'}\langle\beta,\beta'\rangle.\end{aligned}$$

The proof of (3-10) (disregarding the intermediate steps) for the case $n,n'\geq 1$ proceeds in the same manner (or follows by taking adjoints of (3-10) which swaps the roles of m,m' and n,n'). The only case that remains is then $m=0$ and $n'=0$ (again the case $m'=0$ and $n=0$ follows by taking adjoints, or by symmetry).

Then $n \geq 1$ and $m' \geq 1$ and using Lemma 3.2 (6) for the second equality and then applying the induction hypothesis we obtain

$$\begin{aligned} q_1(h_{m'} q_{n+1}(\beta h_n)) &= q_1(h_{m'-1} q_n(h_1 q_{n+1}(\beta h_{n-1} h_1))) \\ &= -q_1(h_{m'-1} q_n(\beta h_{n-1})) \\ &= (L-1)^{m+n-|n-n'|} \delta_{m+n,m'+n'} (-1)^{|n-n'|} \langle \beta, \beta' \rangle. \end{aligned}$$

Then the lemma follows by replacing γ by β in (3-8). □

Recall the elementary fact (see [Rădulescu 1991, Lemma 5] for a proof) that for a real number a, $|a| < 1$, there exist constants $B_a > 0$ and $C_a > 0$ such that for any $k \in \mathbb{N}$, $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$, we have

(3-11) $$B_a \sum_{i=1}^{k} |\lambda_i|^2 \leq \sum_{i=1}^{k} \lambda_i \overline{\lambda}_j a^{|i-j|} \leq C_a \sum_{i=1}^{k} |\lambda_i|^2.$$

Proof of Theorem 3.3 (1). By Lemma 3.4 and (3-11) we see that the assignment $\beta_{m,n} \mapsto \gamma_{m,n}$ extends to a bounded invertible linear mapping $T_0 : X_\beta \to X_\gamma$. By Lemma 3.4 we see that $S : X_\gamma \mapsto X_\gamma : \gamma_{m,n} \mapsto \gamma_{m-1,n-1}$ is bounded with norm $\|S\| \leq (L-1)^{-2}$. Therefore $\mathrm{Id}_{X_\gamma} + S$ is bounded and invertible. As the composition $(I+S) \circ T_0$ is bounded and invertible and agrees with (3-2) we are done. □

Proof of Theorem 3.3 (2). The following Lemma 3.5 is the crucial part of the proof of Theorem 3.3 (2).

Lemma 3.5. *Let $l \geq 2$, $\beta \in \mathbb{C}_q^1[W] \ominus S_1$, and let $\gamma \in \mathbb{C}_q^l[W] \ominus S_l$. For every $m, n \in \mathbb{N}_0$ there exist certain constants $b_{k,j}^{m,n}, c_{k,j}^{m,n} \in \mathbb{R}$, $k = 0, \ldots, m$, $j = 0, \ldots, n$, such that we have the expansions*

(3-12) $$h_m \beta h_n = \sum_{k \leq m, j \leq n} b_{k,j}^{m,n} \beta_{k,j}, \qquad h_m \gamma h_n = \sum_{k \leq m, j \leq n} c_{k,j}^{m,n} \gamma_{k,j}.$$

Moreover, these constants satisfy

(3-13) $$c_{k,j}^{m,n} = b_{k,j}^{m,n} + b_{k+1,j+1}^{m,n}, \quad m, n \in \mathbb{N},\ k = 0, \ldots, m,\ j = 0, \ldots, n,$$

where $b_{m+1,n+1}^{m,n} = 0$.

Proof. If $m = 0$ and $n \in \mathbb{N}$ arbitrary, then the existence of decompositions (3-12) is a consequence of Lemma 3.2. The relation (3-13) for $m = 0$ becomes $c_{k,j}^{0,n} = b_{k,j}^{0,n}$ which is a rather direct consequence of Lemma 3.2 as well.

The proof proceeds by induction on m. Let $L_k = L$ if $k > 1$ and let $L_1 = L + 1$. We have by (3-1) and then Lemma 3.2 (1) and (3),

$$\begin{aligned}
h_m \beta h_n &= (h_1 - p) h_{m-1} \beta h_n - (L_{m-1} - 1) h_{m-2} \beta h_n \\
&= (h_1 - p) \sum_{k=0}^{m-1} \sum_{j=0}^{n} b_{k,j}^{m-1,n} \beta_{k,j} - (L_{m-1} - 1) \sum_{k=0}^{m-2} \sum_{j=0}^{n} b_{k,j}^{m-2,n} \beta_{k,j} \\
&= \sum_{k=0}^{m-1} \sum_{j=0}^{n} b_{k,j}^{m-1,n} (\beta_{k+1,j} + (L-1)\beta_{k-1,j}) \\
&\qquad - \sum_{j=0}^{n} b_{0,j}^{m-1,n} \beta_{0,j-1} - (L_{m-1} - 1) \sum_{k=0}^{m-2} \sum_{j=0}^{n} b_{k,j}^{m-2,n} \beta_{k,j} \\
&= \sum_{k=0}^{m} \sum_{j=0}^{n} (b_{k-1,j}^{m-1,n} + (L-1) b_{k+1,j}^{m-1,n}) \beta_{k,j} \\
&\qquad - \sum_{j=0}^{n-1} b_{0,j+1}^{m-1,n} \beta_{0,j} - (L_{m-1} - 1) \sum_{k=0}^{m-2} \sum_{j=0}^{n} b_{k,j}^{m-2,n} \beta_{k,j}.
\end{aligned} \tag{3-14}$$

This shows that for all $0 \leq k \leq m,\ 0 \leq j \leq n$, we obtain

$$b_{k,j}^{m,n} = b_{k-1,j}^{m-1,n} + (L-1) b_{k+1,j}^{m-1,n} - (L_{m-1} - 1) b_{k,j}^{m-2,n} - \delta_{k,0} b_{0,j+1}^{m-1,n}.$$

Let $\delta_{k \geq 1}$ be 1 if $k \geq 1$ and 0 otherwise. We get then

$$\begin{aligned}
b_{k,j}^{m,n} + b_{k+1,j+1}^{m,n} = \delta_{k \geq 1} (b_{k-1,j}^{m-1,n} + b_{k,j+1}^{m-1,n}) + (L-1)(b_{k+1,j}^{m-1,n} + b_{k+2,j+1}^{m,n+1}) \\
- (L_{m-1} - 1)(b_{k,j}^{m-2,n} + b_{k+1,j+1}^{m-2,n}).
\end{aligned}$$

So by induction

$$\begin{aligned}
b_{k,j}^{m,n} + b_{k+1,j+1}^{m,n} &= \delta_{k \geq 1} c_{k-1,j}^{m-1,n} + (L-1) c_{k+1,j}^{m-1,n} - (L_{m-1} - 1) c_{k,j}^{m-2,n} \\
&= c_{k-1,j}^{m-1,n} + (L-1) c_{k+1,j}^{m-1,n} - (L_{m-1} - 1) c_{k,j}^{m-2,n}.
\end{aligned} \tag{3-15}$$

Exactly as we computed (3-14) (with the difference that Lemma 3.2 (3) is replaced by Lemma 3.2 (2)) we get

$$h_m \gamma h_n = \sum_{k=0}^{m+1} \sum_{j=0}^{n} (c_{k-1,j}^{m-1,n} + (L-1) c_{k+1,j}^{m-1,n}) \gamma_{k,j} - (L_{m-1} - 1) \sum_{k=0}^{m-2} \sum_{j=0}^{n} c_{k,j}^{m-2,n} \gamma_{k,j}.$$

Thus

$$c_{k,j}^{m,n} = c_{k-1,j}^{m-1,n} + (L-1) c_{k+1,j}^{m-1,n} - (L_m - 1) c_{k,j}^{m-2,n}.$$

Combining the above with (3-15) gives $c_{k,j}^{m,n} = b_{k,j}^{m,n} + b_{k+1,j+1}^{m,n}$ for all $0 \leq k \leq m$, $0 \leq j \leq n$. □

Proof of Theorem 3.3 (2). Lemma 3.5 shows that $\mathsf{B}\gamma\mathsf{B} \subseteq X_\gamma$ and $\mathsf{B}\beta\mathsf{B} \subseteq X_\beta$ and hence the inclusions hold also for the $\| \|_2$-closures. For the converse inclusion proceed by induction: take $h_n\gamma h_m \in \mathsf{B}\gamma\mathsf{B}$ and assume that all vectors $h_r\beta h_s$ with $r < n, s \leq m$ are contained in X_γ (if $n = 0$ then assume that $r \leq n, s < m$ and consider adjoints, or use a similar induction argument on m). By (3-1) we have

$$h_n\gamma h_m = (h_1 - p)h_{n-1}\gamma h_m - (L_n - 1)h_{n-2}\gamma h_m \in h_1X_\gamma + X_\gamma.$$

Here again $L_n = L$ if $n \geq 2$ and $L_1 = L + 1$. So it suffices to show that $h_1X_\gamma \subseteq X_\gamma$, but this is a consequence of Lemma 3.2 (2). The proof for β instead of γ is the same but uses Lemma 3.2 (1) and (3) for the latter argument.

The fact that (3-3) agrees with (3-2) is now a direct consequence of Lemma 3.5. Indeed,

$$\begin{aligned} T(h_m\beta h_n) = T\Bigg(\sum_{k\leq m, j\leq n} b^{m,n}_{k,j}\beta_{k,j}\Bigg) &= \sum_{k\leq m, j\leq n} b^{m,n}_{k,j}(\gamma_{k,j} + \gamma_{k-1,j-1}) \\ &= \sum_{k\leq m, j\leq n} (b^{m,n}_{k,j} + b^{m,n}_{k+1,j+1})\gamma_{k,j} = \sum_{k\leq m, j\leq n} c^{m,n}_{k,j}\gamma_{k,j} = h_m\gamma h_n. \qquad \square \end{aligned}$$

Consequences of Theorem 3.3. Let $\mathsf{B}_r = \langle R_h\rangle''$ (note that as $\mathrm{VN}_q(W)$ is in the standard form on $\ell^2(W)$, it is also equal to $J\mathsf{B}J$, where J is the antilinear Tomita–Takesaki modular conjugation $\delta_x \mapsto \delta_{x^{-1}}$). For a vector $\gamma \in \bigcup_{l\in\mathbb{N}_0} \mathbb{C}^l_q[W]$ we let p_γ be the central support in $(\mathsf{B}\cup\mathsf{B}_r)''$ of the vector state $\omega_{\gamma,\gamma}$. The operator p_γ is then given by the projection onto the closure of $\mathsf{B}\gamma\mathsf{B}$.

Lemma 3.6. *If vectors $\xi, \xi' \in \cup_{l\geq 1}\mathbb{C}^l_q[W] \ominus S_l$ are orthogonal then p_ξ and $p_{\xi'}$ are orthogonal projections.*

Proof. Let $\xi \in \mathbb{C}^l_q[W] \ominus S_l$ and let $\xi' \in \mathbb{C}^{l'}_q[W] \ominus S_{l'}$ with $l, l' \geq 1$. If $l = l'$ then the lemma follows directly from Lemma 3.4. So assume that $l \neq l'$ and say that $l' \leq l$. It suffices to show that

$$\xi'_{r,s} \perp \xi_{m,n} \quad \text{for every } r, s, m, n \in \mathbb{N}_0. \tag{3-16}$$

If $m + n + l \neq r + s + l'$ this is obvious as then the images of q_{m+n+l} and $q_{r+s+l'}$ are mutually orthogonal. We may then assume $m + n + l = r + s + l'$, so that $r + s \geq m + n$. If $m + n = 0$ then (3-16) is obvious, as $\xi \perp S_l$ whereas $\xi'_{r,s} \in S_l$. But then note that $\xi'_{r,s} = (\xi'_{a,b})_{r-a,s-b}$ for any $a = 0, \ldots, r,\ b = 0, \ldots s$ such that $l' + a + b = l$. As $\xi'_{a,b} \in S_l$ we see from Lemma 3.4 that $(\xi'_{a,b})_{r-a,s-b} \perp \xi_{m,n}$. $\square$

We can now state and prove the main result of this section.

Theorem 3.7. *The von Neumann algebra $(\mathsf{B}\cup\mathsf{B}_r)'(1 - p_\Omega)$ is homogeneous of type I_∞.*

Proof. Because $(\mathsf{B}\cup\mathsf{B}_r)''$ is abelian, the commutant $(\mathsf{B}\cup\mathsf{B}_r)'$ decomposes as a direct sum $\bigoplus_{n=1}^{\infty}\mathsf{A}_n\overline{\otimes}B(\mathcal{H}_n)$, where $\dim(\mathcal{H}_n)=n$ and the algebras A_n are abelian (see [Dixmier 1969]). Let $(\xi_i)_{i\in\mathbb{N}}$ be an orthonormal basis in $\bigcup_{l\geq 1}\mathbb{C}_q^l[W]\ominus S_l$. By Lemma 3.6 the projections $(p_{\xi_i})_{i\in\mathbb{N}}$ are mutually orthogonal and by Theorem 3.3 they have the same central support in $(\mathsf{B}\cup\mathsf{B}_r)'$. As by Lemma 3.6 we have $\sum_{i\in\mathbb{N}}p_{\xi_i}=1-p_\Omega$ and $1-p_\Omega$ is central in $(\mathsf{B}\cup\mathsf{B}_r)'$ (see [Popa 1985, Lemma 3.1]), we see that the central support of each p_{ξ_i} in $(\mathsf{B}\cup\mathsf{B}_r)'$ is $1-p_\Omega$, which is the unit in $(\mathsf{B}\cup\mathsf{B}_r)'(1-p_\Omega)$. Since we have a partition of unity formed by projections with the same central support, the above decomposition of $(\mathsf{B}\cup\mathsf{B}_r)'$ must in fact consist of only one element. As there are infinitely many orthogonal projections, this summand must correspond to $n=\infty$, so that we have $(\mathsf{B}\cup\mathsf{B}_r)'(1-p_\Omega)=A_\infty\overline{\otimes}B(\ell^2)$ □

Remark 3.8. Theorem 3.7 is phrased in the literature as follows: the Pukánszky invariant of B is $\{\infty\}$. This is because in the B-B-bimodule $(1-p_\Omega)L^2(\mathsf{M})$, the only factors occurring in the direct integral decomposition of the commutant of $\mathsf{B}\cup\mathsf{B}_r$ are infinite (and necessarily of type I).

Corollary 3.9. *The radial subalgebra* B *is a singular masa of* $\mathrm{VN}_q(W)$.

Proof. This follows from Theorem 3.7 by [Popa 1985, Remark 3.4]. □

4. Generator masas in q-deformed Gaussian von Neumann algebras

In this section we consider masas in a different deformation of the free group factors, i.e., so-called q-Gaussian algebras.

The starting point of the construction of q-Gaussian algebras is a real Hilbert space $\mathcal{H}_\mathbb{R}$. We complexify it, obtaining a complex Hilbert space $\mathcal{H}$, and form an algebraic direct sum $\bigoplus_{n\geqslant 0}\mathcal{H}^{\otimes n}$, where $\mathcal{H}^{\otimes 0}=\mathbb{C}$. Following [Bożejko et al. 1997] (see that paper for all facts stated below without proofs), we will define an inner product on this space using the parameter $q\in(-1,1)$. For each $n\in\mathbb{N}$ we define an operator $P_q^n:\mathcal{H}^{\otimes n}\to\mathcal{H}^{\otimes n}$ by the formula

$$P_q^n(e_1\otimes\cdots\otimes e_n)=\sum_{\pi\in S_n}q^{i(\pi)}e_{\pi(1)}\otimes\cdots\otimes e_{\pi(n)},$$

where $e_1,\ldots,e_n\in\mathcal{H}$, S_n is the permutation group on n letters and $i(\pi)$ denotes the number of inversions in the permutation π. These operators are strictly positive, so they define an inner product on $\bigoplus_{n\geqslant 0}\mathcal{H}^{\otimes n}$ — the Hilbert space that we get after completion is called the q-Fock space and is denoted by $\mathcal{F}_q(\mathcal{H})$. The direct sum decomposition of the q-Fock space allows us to define shift-like operators.

Definition 4.1. Let $\xi\in\mathcal{H}$. We define the *creation operator* $a_q^*(\xi):\mathcal{F}_q(\mathcal{H})\to\mathcal{F}_q(\mathcal{H})$ by $a_q^*(\xi)(e_1\otimes\cdots\otimes e_n)=\xi\otimes e_1\otimes\cdots\otimes\cdots e_n$. The *annihilation operator* $a_q(\xi):\mathcal{F}_q(\mathcal{H})\to\mathcal{F}_q(\mathcal{H})$ is defined as the adjoint of $a_q^*(\xi)$. Using the definition of the q-deformed inner product we can find the formula for $a_q(\xi)$:

$$a_q(\xi)(e_1 \otimes \cdots \otimes e_n) = \sum_{i=1}^{n} q^{i-1} \langle \xi, e_i \rangle e_1 \otimes \cdots \widehat{e_i} \cdots \otimes e_n,$$

where $\widehat{e_i}$ means that the factor e_i is omitted. All the above operators extend to bounded operators on $\mathcal{F}_q(\mathcal{H})$.

Creation and annihilation operators will allow us to define q-Gaussian algebras.

Definition 4.2. Let $\mathcal{H}_{\mathbb{R}}$ be a real Hilbert space and let $\mathcal{H}$ be its complexification. The von Neumann subalgebra of $\mathrm{B}(\mathcal{F}_q(\mathcal{H}))$ generated by the set $\{a_q^*(\xi)+a_q(\xi) : \xi \in \mathcal{H}_{\mathbb{R}}\}$ is called the *q-Gaussian algebra* associated with $\mathcal{H}_{\mathbb{R}}$ and is denoted by $\Gamma_q(\mathcal{H}_{\mathbb{R}})$.

The vector $\Omega = 1 \in \mathbb{C} \subset \mathcal{H}^{\otimes 0} \subset \mathcal{F}_q(\mathcal{H})$ is called the *vacuum vector*. It is a cyclic and separating vector for $\Gamma_q(\mathcal{H}_{\mathbb{R}})$ and the associated vector state $\omega(x) := \langle \Omega, x\Omega \rangle$ is a normal faithful trace on $\Gamma_q(\mathcal{H}_{\mathbb{R}})$.

Remark 4.3. For $q = 0$ the assignment $\mathcal{H}_{\mathbb{R}} \mapsto \Gamma_q(\mathcal{H}_{\mathbb{R}})$ is precisely Voiculescu's free Gaussian functor. In particular $\Gamma_0(\mathcal{H}_{\mathbb{R}}) \simeq \mathrm{L}(\mathbb{F}_{\dim(\mathcal{H}_{\mathbb{R}})})$.

We will study problems pertaining to conjugacy of masas in the q-Gaussian algebras. It is a nice feature of these objects that the orthogonal operators on $\mathcal{H}_{\mathbb{R}}$ give rise to automorphisms of $\Gamma_q(\mathcal{H}_{\mathbb{R}})$. To introduce these automorphisms, we need to present the *first quantisation*.

Definition 4.4. Let $T : \mathcal{H} \to \mathcal{H}$ be a contraction. The assignment

$$\bigoplus_{k\geqslant 0} \mathcal{H}^{\otimes k} \ni e_1 \otimes \cdots \otimes e_n \mapsto Te_1 \otimes \cdots \otimes Te_n \in \bigoplus_{k\geqslant 0} \mathcal{H}^{\otimes k}$$

extends to a contraction $\mathcal{F}_q(T) : \mathcal{F}_q(\mathcal{H}) \to \mathcal{F}_q(\mathcal{H})$ and is called the first quantisation of T.

Remark 4.5. If $U : \mathcal{H} \to \mathcal{H}$ is a unitary then $\mathcal{F}_q(U)$ is also a unitary.

To work with $\Gamma_q(\mathcal{H}_{\mathbb{R}})$ we need a convenient notation for its generators. For any $\xi \in \mathcal{H}_{\mathbb{R}}$ we put $W(\xi) := a_q^*(\xi) + a_q(\xi)$. If $\eta = \xi_1 + i\xi_2 \in \mathcal{H}$ then we denote $W(\eta) = W(\xi_1) + iW(\xi_2)$; therefore $W(\eta)$ is complex-linear in η. Recall that the vacuum vector Ω is cyclic and separating. One can check that for any vectors $\eta_1, \dots, \eta_n \in \mathcal{H}$ we have $\eta_1 \otimes \cdots \otimes \eta_n \in \Gamma_q(\mathcal{H}_{\mathbb{R}})\Omega$; the unique operator $W(\eta_1 \otimes \cdots \otimes \eta_n) \in \Gamma_q(\mathcal{H}_{\mathbb{R}})$ such that $W(\eta_1 \otimes \cdots \otimes \eta_n)\Omega = \eta_1 \otimes \cdots \otimes \eta_n$ is called a *Wick word*. The span of all such operators associated with finite simple tensors forms a strongly dense *-subalgebra of $\Gamma_q(\mathcal{H}_{\mathbb{R}})$, which we call the *algebra of Wick words*. Finally note that much as in Section 2 we can also consider the "right" version of $\Gamma_q(\mathcal{H}_{\mathbb{R}})$, generated by the combinations of right creation and annihilation operators, in particular containing the right Wick words, to be denoted $W_r(\xi)$. We are ready to introduce the *second quantisation*.

Definition 4.6. Let $\mathcal{H}_{\mathbb{R}}$ be a real Hilbert space and let $\mathcal{H}$ be its complexification. Suppose that $T : \mathcal{H} \to \mathcal{H}$ is a contraction such that $T(\mathcal{H}_{\mathbb{R}}) \subset \mathcal{H}_{\mathbb{R}}$. Then the assignment $\Gamma_q(\mathcal{H}_{\mathbb{R}}) \ni W(\eta_1 \otimes \cdots \otimes \eta_n) \mapsto W(T\eta_1 \otimes \cdots \otimes T\eta_n) \in \Gamma_q(\mathcal{H}_{\mathbb{R}})$, where $\eta_1, \ldots, \eta_n \in \mathcal{H}$, may be extended to a normal, unital, completely positive map on $\Gamma_q(\mathcal{H}_{\mathbb{R}})$, denoted by $\Gamma_q(T)$.

Remark 4.7. Note that the condition $T(\mathcal{H}_{\mathbb{R}}) \subset \mathcal{H}_{\mathbb{R}}$ is essential, otherwise $\Gamma_q(T)$ would not even preserve the adjoint, let alone be completely positive.

We will only deal with automorphisms and, in this construction, they come from orthogonal operators on $\mathcal{H}_{\mathbb{R}}$. If $U : \mathcal{H}_{\mathbb{R}} \to \mathcal{H}_{\mathbb{R}}$ is orthogonal then $\Gamma_q(U)(x) = \mathcal{F}_q(U)x\mathcal{F}_q(U)^*$, where we still denote by U its canonical unitary extension to $\mathcal{H}$. It is easy to check that $\Gamma_q(U)W(\xi) = W(U\xi)$.

To find candidates for masas, we draw inspiration from the case $q = 0$, in which the most basic masas are the so-called generator masas. In our picture they correspond to subalgebras generated by a single element $W(\xi)$, where $\xi \in \mathcal{H}_{\mathbb{R}}$. Ricard [2005] proved they are also masas in the case of q-Gaussian algebras. As an application, he established factoriality of all q-Gaussian algebras $\Gamma_q(\mathcal{H}_{\mathbb{R}})$ with $\dim(\mathcal{H}_{\mathbb{R}}) \geqslant 2$. Recently these generator masas were also shown to be singular [Wen 2017] and maximally injective [Parekh et al. 2018] (the latter for sufficiently small $|q|$).

Using the automorphisms produced by the second quantisation procedure, we can easily show that all these masas are conjugate by an outer automorphism. Indeed, consider masas generated by $W(\xi)$ and $W(\eta)$, where $\xi, \eta \in \mathcal{H}_{\mathbb{R}}$. By rescaling, we may assume that $\|\xi\| = \|\eta\| = 1$. Therefore one can find an orthogonal operator U such that $U\xi = \eta$; then $\Gamma_q(U)((W(\xi))'') = (W(\eta))''$. Our aim now is to show that they are never conjugate by a unitary.

Case of orthogonal vectors. We first want to deal with the case when $\mathsf{A} := (W(e_1))''$ and $\mathsf{B} := (W(e_2))''$ are masas in $\mathsf{M} := \Gamma_q(\mathcal{H}_{\mathbb{R}})$ coming from two orthogonal vectors. In the case $q = 0$ these masas correspond to two different generator masas of the free group factor. One can prove that these are not unitarily conjugate using Popa's notion of orthogonal pairs of subalgebras (see [Popa 1983, Corollary 4.3]). We will use another technique due to Popa giving a criterion for embedding A into B inside M (in a certain technical sense). We will actually only state the part of the theorem that is useful for us; for the full statement consult [Popa 2006, Theorem 2.1 and Corollary 2.3] or [Popa 2019, Theorem 1.3.1]. We call A and B *intertwinable* (inside M) if the intertwiner space $\mathcal{I}_{\mathsf{M}}(\mathsf{A}, \mathsf{B})$, defined in [Popa 2019, Subsection 1.3] is nontrivial.

Proposition 4.8 (Popa). *Let* A *and* B *be von Neumann subalgebras of a finite von Neumann algebra* (M, τ). *Suppose that there exists a sequence of unitaries* $(u_k)_{k\in\mathbb{N}} \subset \mathcal{U}(\mathsf{A})$ *such that for any* $x, y \in \mathsf{M}$ *we have* $\lim_{k\to\infty} \|\mathbb{E}_{\mathsf{B}}(xu_k y)\|_2 = 0$, *where* $\mathbb{E}_{\mathsf{B}}$ *is the unique* τ*-preserving conditional expectation from* M *onto* B*. Then* A *and* B *are nonintertwinable*; *in particular there does not exist a unitary* $u \in \mathsf{M}$ *such that* $u\mathsf{A}u^* = \mathsf{B}$.

Remark 4.9. Note that it suffices to check that $\lim_{k\to\infty} \|\mathbb{E}_{\mathsf{B}}(xu_k y)\|_2 = 0$ only for $x, y \in \widetilde{\mathsf{M}}$, where $\widetilde{\mathsf{M}}$ is a strongly dense *-subalgebra. It follows from Kaplansky's density theorem, because we can approximate in the strong operator topology (in particular in L^2) and control the norm of the approximants at the same time.

Proposition 4.10. *Let* $e_1, e_2 \in \mathcal{H}_{\mathbb{R}}$, $\|e_1\| = \|e_2\| = 1$, $e_1 \perp e_2$*. Set* $\mathsf{A} = (W(e_1))''$, $\mathsf{B} = (W(e_2))''$, *and* $\mathsf{M} = \Gamma_q(\mathcal{H}_{\mathbb{R}})$*. There exists a sequence of unitaries* $(u_k)_{k\in\mathbb{N}} \subset \mathcal{U}(\mathsf{A})$ *such that we have* $\lim_{k\to\infty} \|\mathbb{E}_{\mathsf{B}}(xu_k y)\|_2 = 0$ *for all* $x, y \in \widetilde{\mathsf{M}}$, *where* $\widetilde{\mathsf{M}}$ *is the algebra of Wick words.*

Proof. Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{H}_{\mathbb{R}}$. Assume $x = W(e_{i_1} \otimes \cdots \otimes e_{i_n})$ and $y = W(e_{j_1} \otimes \cdots \otimes e_{j_m})$; it clearly suffices because the span of such elements is equal to $\widetilde{\mathsf{M}}$. By definition of the trace on $\Gamma_q(\mathcal{H}_{\mathbb{R}})$ we have $\|\mathbb{E}_{\mathsf{B}}(xu_k y)\|_2 = \|(\mathbb{E}_{\mathsf{B}}(xu_k y))\Omega\|$. Since the conditional expectation on the level of the Fock space is just the orthogonal projection (denoted P) onto the closed linear span of the set $\{e_2^{\otimes n} : n \in \mathbb{N}\}$, we get $\|(\mathbb{E}_{\mathsf{B}}(xu_k y))\Omega\| = \|P(xu_k y\Omega)\|$. Note now that as the left and right actions of y on Ω produce the same result, $e_{j_1} \otimes \cdots \otimes e_{j_m}$, we can change y to its right version, $W_r(e_{j_1} \otimes \cdots \otimes e_{j_m})$, denoted now by $\tilde{y}$. Since $\tilde{y} \in \mathsf{M}'$, we get $\|P(xu_k y\Omega)\| = \|P(x\tilde{y}u_k\Omega)\|$. We now choose the sequence $(u_k)_{k\in\mathbb{N}}$ — it is an arbitrary sequence of unitaries in A such that the corresponding vectors $\eta_k := u_k\Omega$ converge weakly to zero (such a sequence exists, because A is diffuse). Let Q_l be the orthogonal projection from $\mathcal{F}_q(\mathbb{C}e_1)$ onto $\mathrm{span}\{e_1^{\otimes j} : j \leqslant l\}$. Then for any l the sequence $(Q_l\eta_k)_{k\in\mathbb{N}}$ converges to zero in norm. Therefore to check that $\lim_{k\to\infty} \|P(x\tilde{y}\eta_k)\| = 0$, it suffices to do so for η_k replaced by $(\mathbb{1} - Q_l)\eta_k$. We now choose $l = n+m$. Therefore any η_k consists solely of tensors $e_1^{\otimes d}$, where $d \geqslant n+m+1$. Since x can be written as a sum of products of n (in total) creation and annihilation operators and y can be decomposed similarly into products of m creation and annihilation operators, any simple tensor appearing in $x\tilde{y}(\mathbb{1} - Q_{n+m})\eta_k$ will contain at least one e_1. But all such simple tensors are orthogonal to $\mathcal{F}_q(\mathbb{C}e_2)$, so they are killed by P. □

Corollary 4.11. *If the vectors* e_1 *and* e_2 *in* $\mathcal{H}_{\mathbb{R}}$ *are orthogonal, then masas* $(W(e_1))''$ *and* $(W(e_2))''$ *are not intertwinable inside* $\Gamma_q(\mathcal{H}_{\mathbb{R}})$.

General case. Let us check now if the method used for a pair of orthogonal vectors can be applied in a more general setting. Assume now that e_1 and v are two unit vectors and write $v = \alpha e_1 + \beta e_2$, where $e_2 \perp e_1$, $\alpha^2 + \beta^2 = 1$, and $\beta \neq 0$. We fix now an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of $\mathcal{H}_{\mathbb{R}}$ (if $\mathcal{H}_{\mathbb{R}}$ is finite-dimensional then this should be a finite sequence).

Proposition 4.12. *The masas* $\mathsf{A} := W(v)''$ *and* $\mathsf{B} := (W(e_1))''$ *are not intertwinable* (*so in particular are not unitarily conjugate*).

Proof. We proceed exactly as in the proof of Proposition 4.10 and also use the same notation; note however that this time P will be the orthogonal projection onto $\overline{\operatorname{span}}\{e_1^{\otimes n} : n \geqslant 0\}$. The only problem is that now we do not have orthogonality. Write $\eta_k = \sum_{j\in\mathbb{N}} a_j^{(k)} v^{\otimes j}$. We have $\|v^{\otimes j}\| \simeq (1/\sqrt{1-q})^j$ (see the third displayed formula on page 660 of [Ricard 2005]). Let us compute $v^{\otimes j}$:

$$v^{\otimes j} = \sum_{k=0}^{j} \alpha^{j-k}\beta^k R_{j,k}(e_1^{\otimes(j-k)} \otimes e_2^{\otimes k}),$$

where $R_{j,k}(e_1^{\otimes(j-k)} \otimes e_2^{\otimes k})$ is equal to the sum of all simple tensors such that $j-k$ factors are equal to e_1 and k factors are equal to e_2; there are $\binom{j}{k}$ such simple tensors. Note now that if $k \geqslant n+m+1$ then after applying $x\tilde{y}$ at least one e_2 remains as a factor, so the orthogonal projection P kills it. We conclude that it suffices to perform the summation in the displayed formula above only up to $j \wedge (n+m)$; we call the resulting tensors $\tilde{v}^{\otimes j}$ and the corresponding η_k is dubbed $\widetilde{\eta}_k$. Since k is bounded, the number $\binom{j}{k}$ is polynomial in j, so if we get exponential decay of the norm of the individual factors in the sum, the factor $\binom{j}{k}$ does not affect the overall convergence. After neglecting the terms with $k > n+m$, we use the trivial estimate $\|P(x\tilde{y}\widetilde{\eta}_k)\| \leqslant C\|\widetilde{\eta}_k\|$. The proof will be completed if we show that $\|\widetilde{\eta}_k\|$ converges to 0. Note now that the square of the norm of $\widetilde{\eta}_k$ is equal to $\sum_{j\in\mathbb{N}} |a_j^{(k)}|^2 \cdot \|\tilde{v}^{\otimes j}\|^2$. Recall that $\|\eta_k\| \leqslant 1$ and $\|v^{\otimes j}\| \simeq (1/\sqrt{1-q})^j$, so the coefficients $a_j^{(k)}$ satisfy $\sum_{j\in\mathbb{N}} |a_j^{(k)}|^2 \big(\frac{1}{1-q}\big)^j \lesssim 1$. It therefore suffices to show that $\lim_{j\to\infty}(1-q)^j\|\tilde{v}^{\otimes j}\|^2 = 0$, remembering that the vectors η_k converge weakly to 0, so we only care about large j. We estimate the norm of $\tilde{v}^{\otimes j}$ by the triangle inequality:

$$\|\tilde{v}^{\otimes j}\| \leqslant \sum_{k=0}^{j\wedge(n+m)} |\alpha|^{j-k}|\beta|^k \binom{j}{k} \|e_1^{\otimes k} \otimes e_2^{j-k}\|.$$

Since k is bounded, one can easily get an estimate of the form

$$\|e_1^{\otimes k} \otimes e_2^{\otimes(j-k)}\| \leqslant C(1/\sqrt{1-q})^j$$

(see [Ricard 2005, Remark 2]). This yields $\|\tilde{v}^{\otimes j}\| \leqslant C(1/\sqrt{1-q})^j|\alpha|^j \cdot j^k$. This is the inequality that we wanted, i.e., we find out that $(1-q)^j\|\tilde{v}^{\otimes j}\|^2$ is bounded by $Cj^k|\alpha|^j$, which converges to zero very fast, as we assumed that $|\alpha| < 1$. This finishes the proof of the proposition. □

We can now use the result to prove that the second quantisation automorphisms are never inner, unless trivial; it extends a result of Houdayer and Shlyakhtenko in the free case [2011, Theorem 5.1].

Corollary 4.13. *Let $\mathcal{H}_\mathbb{R}$ be a real Hilbert space and let $U : \mathcal{H}_\mathbb{R} \to \mathcal{H}_\mathbb{R}$ be an orthogonal transformation. If $\Gamma_q(U) : \Gamma_q(\mathcal{H}_\mathbb{R}) \to \Gamma_q(\mathcal{H}_\mathbb{R})$ is an inner automorphism then $U = \mathbb{1}$.*

Proof. If U is not a multiple of identity then there exists a vector $v \in \mathcal{H}_\mathbb{R}$ such that Uv is not a multiple of v. The masas $\mathsf{A} := W(v)''$ and $\mathsf{B} := (W(Uv))''$ are conjugate by the automorphism $\Gamma_q(U)$, but by Proposition 4.12 they are not conjugate by an inner automorphism.

The only remaining case is now $U = -\mathbb{1}$. We may assume that the dimension of $\mathcal{H}_\mathbb{R}$ is at least 2, because otherwise $\Gamma_q(\mathcal{H}_\mathbb{R})$ is commutative and any nontrivial automorphism is outer. Pick two orthogonal vectors e_1 and e_2 and consider the masas $\mathsf{A} = (W(e_1))''$ and $\mathsf{B} = (W(e_2))''$. Assume now that the automorphism $x \mapsto \mathcal{F}_q(-\mathbb{1})x\mathcal{F}_q(-\mathbb{1})$ is inner, so there is a unitary $u \in \Gamma_q(\mathcal{H}_\mathbb{R})$ implementing it. Since $\mathcal{F}_q(-\mathbb{1})W(e_1)\mathcal{F}_q(-\mathbb{1}) = -W(e_1)$, this automorphism preserves A; it also preserves B. But the masas in question are singular, so $u \in \mathsf{A} \cap \mathsf{B}$. It follows that $u\Omega \in L^2(\mathsf{A}) \cap L^2(\mathsf{B}) = \mathbb{C}\Omega$, so u has to be a multiple of identity, which is a contradiction, because this would yield the trivial automorphism. □

Remark 4.14. The results above exhibit in particular explicitly a continuum of nonmutually intertwinable singular masas in $\Gamma_q(\mathcal{H}_\mathbb{R})$. Very recently Popa [2019] showed the existence of such uncountable families in every separable II_1-factor (see Corollary 2.2 of that paper).

Remark 4.15. Generator masas can be also studied for the so-called mixed q-Gaussians (see [Speicher 1993]). They are known to be masas by [Skalski and Wang 2018], and in fact an application of methods of that paper and general results of [Bikram and Mukherjee 2017] show that they are singular, as noted by Simeng Wang. There seems to be however nothing known about the "radial" subalgebra in this more general context. Is it a masa? Is it isomorphic to a generator one?

Acknowledgements. Caspers received funding from the European Union's Horizon 2020 research and innovation programme through grant agreement No. 702139. Skalski was partially supported by the National Science Centre (NCN) grant No. 2014/14/E/ST1/00525. Wasilewski was partially supported by the National Science Centre (NCN) grant No. 2016/21/N/ST1/02499. The work on the paper was started during Caspers's visit to IMPAN in January 2017, partially funded by the Warsaw Centre of Mathematics and Computer Science. We thank the referee for helpful comments.

References

[Bikram and Mukherjee 2017] P. Bikram and K. Mukherjee, "Generator masas in q-deformed Araki–Woods von Neumann algebras and factoriality", *J. Funct. Anal.* **273**:4 (2017), 1443–1478. MR Zbl

[Boca and Rădulescu 1992] F. Boca and F. Rădulescu, "Singularity of radial subalgebras in II_1 factors associated with free products of groups", *J. Funct. Anal.* **103**:1 (1992), 138–159. MR Zbl

[Bożejko et al. 1997] M. Bożejko, B. Kümmerer, and R. Speicher, "q-Gaussian processes: non-commutative and classical aspects", *Comm. Math. Phys.* **185**:1 (1997), 129–154. MR Zbl

[Brannan and Vergnioux 2018] M. Brannan and R. Vergnioux, "Orthogonal free quantum group factors are strongly 1-bounded", *Adv. Math.* **329** (2018), 133–156. MR Zbl

[Cameron et al. 2010] J. Cameron, J. Fang, M. Ravichandran, and S. White, "The radial masa in a free group factor is maximal injective", *J. Lond. Math. Soc.* (2) **82**:3 (2010), 787–809. MR Zbl

[Caspers 2016a] M. Caspers, "Absence of Cartan subalgebras for right-angled Hecke von Neumann algebras", preprint, 2016. To appear in *Anal. PDE*. arXiv

[Caspers 2016b] M. Caspers, "Connes embeddability of graph products", *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **19**:1 (2016), art. id. 1650004. MR Zbl

[Caspers and Fima 2017] M. Caspers and P. Fima, "Graph products of operator algebras", *J. Noncommut. Geom.* **11**:1 (2017), 367–411. MR Zbl

[Cohen and Trenholme 1984] J. M. Cohen and A. R. Trenholme, "Orthogonal polynomials with a constant recursion formula and an application to harmonic analysis", *J. Funct. Anal.* **59**:2 (1984), 175–184. MR Zbl

[Davis 2008] M. W. Davis, *The geometry and topology of Coxeter groups*, London Mathematical Society Monographs Series **32**, Princeton University Press, 2008. MR Zbl

[Davis et al. 2007] M. W. Davis, J. Dymara, T. Januszkiewicz, and B. Okun, "Weighted L^2-cohomology of Coxeter groups", *Geom. Topol.* **11** (2007), 47–138. MR Zbl

[Dixmier 1954] J. Dixmier, "Sous-anneaux abéliens maximaux dans les facteurs de type fini", *Ann. of Math.* (2) **59** (1954), 279–286. MR Zbl

[Dixmier 1969] J. Dixmier, *Les algèbres d'opérateurs dans l'espace hilbertien* (*algèbres de von Neumann*), 2nd ed., Gauthier-Villars, Paris, 1969. MR Zbl

[Dykema 1993] K. Dykema, "Free products of hyperfinite von Neumann algebras and free dimension", *Duke Math. J.* **69**:1 (1993), 97–119. MR Zbl

[Dymara 2006] J. Dymara, "Thin buildings", *Geom. Topol.* **10** (2006), 667–694. MR Zbl

[Freslon and Vergnioux 2016] A. Freslon and R. Vergnioux, "The radial MASA in free orthogonal quantum groups", *J. Funct. Anal.* **271**:10 (2016), 2776–2807. MR Zbl

[Garncarek 2016] Ł. Garncarek, "Factoriality of Hecke–von Neumann algebras of right-angled Coxeter groups", *J. Funct. Anal.* **270**:3 (2016), 1202–1219. MR Zbl

[Houdayer and Shlyakhtenko 2011] C. Houdayer and D. Shlyakhtenko, "Strongly solid II_1 factors with an exotic MASA", *Int. Math. Res. Not.* **2011**:6 (2011), 1352–1380. MR Zbl

[Parekh et al. 2018] S. Parekh, K. Shimada, and C. Wen, "Maximal amenability of the generator subalgebra in q-Gaussian von Neumann algebras", *J. Operator Theory* **80**:1 (2018), 125–152. MR Zbl

[Popa 1983] S. Popa, "Orthogonal pairs of $*$-subalgebras in finite von Neumann algebras", *J. Operator Theory* **9**:2 (1983), 253–268. MR Zbl

[Popa 1985] S. Popa, "Notes on Cartan subalgebras in type II_1 factors", *Math. Scand.* **57**:1 (1985), 171–188. MR Zbl

[Popa 2006] S. Popa, "Strong rigidity of II_1 factors arising from malleable actions of w-rigid groups, I", *Invent. Math.* **165**:2 (2006), 369–408. MR Zbl

[Popa 2019] S. Popa, "Constructing MASAs with prescribed properties", *Kyoto J. Math.* **59**:2 (2019), 367–397.

[Pytlik 1981] T. Pytlik, "Radial functions on free groups and a decomposition of the regular representation into irreducible components", *J. Reine Angew. Math.* **326** (1981), 124–135. MR Zbl

[Ricard 2005] E. Ricard, "Factoriality of q-Gaussian von Neumann algebras", *Comm. Math. Phys.* **257**:3 (2005), 659–665. MR Zbl

[Rădulescu 1991] F. Rădulescu, "Singularity of the radial subalgebra of $\mathscr{L}(F_N)$ and the Pukánszky invariant", *Pacific J. Math.* **151**:2 (1991), 297–306. MR Zbl

[Sinclair and Smith 2008] A. M. Sinclair and R. R. Smith, *Finite von Neumann algebras and masas*, London Mathematical Society Lecture Note Series **351**, Cambridge University Press, 2008. MR Zbl

[Skalski and Wang 2018] A. Skalski and S. Wang, "Remarks on factoriality and q-deformations", *Proc. Amer. Math. Soc.* **146**:9 (2018), 3813–3823. MR Zbl

[Speicher 1993] R. Speicher, "Generalized statistics of macroscopic fields", *Lett. Math. Phys.* **27**:2 (1993), 97–104. MR Zbl

[Trenholme 1988] A. R. Trenholme, "Maximal abelian subalgebras of function algebras associated with free products", *J. Funct. Anal.* **79**:2 (1988), 342–350. MR Zbl

[Wen 2017] C. Wen, "Singularity of the generator subalgebra in q-Gaussian algebras", *Proc. Amer. Math. Soc.* **145**:8 (2017), 3493–3500. MR Zbl

Received April 10, 2017. Revised February 17, 2019.

MARTIJN CASPERS
MATHEMATISCH INSTITUUT
UNIVERSITEIT UTRECHT
UTRECHT
THE NETHERLANDS
Current address:
DELFT UNIVERSITY OF TECHNOLOGY
DELFT
NETHERLANDS
m.p.t.caspers@tudelft.nl

ADAM SKALSKI
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAW
POLAND
a.skalski@impan.pl

MATEUSZ WASILEWSKI
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSAW
POLAND
Current address:
DEPARTMENT OF MATHEMATICS
KATHOLIEKE UNIVERSITEIT LEUVEN
LEUVEN
BELGIUM
mateusz.wasilewski@kuleuven.be

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.23

THE COMPACT PICTURE OF SYMMETRY-BREAKING OPERATORS FOR RANK-ONE ORTHOGONAL AND UNITARY GROUPS

JAN FRAHM AND BENT ØRSTED

We present a method to calculate intertwining operators between the underlying Harish-Chandra modules of degenerate principal series representations of a reductive Lie group G and a reductive subgroup G', and between their composition factors. Our method describes the restriction of these operators to the K'-isotypic components, $K' \subseteq G'$ a maximal compact subgroup, and reduces the representation-theoretic problem to an infinite system of scalar equations of a combinatorial nature. For rank-one orthogonal and unitary groups and spherical principal series representations we calculate these relations explicitly and use them to classify intertwining operators. We further show that in these cases automatic continuity holds; i.e., every intertwiner between the Harish-Chandra modules extends to an intertwiner between the Casselman–Wallach completions, verifying a conjecture by Kobayashi. Altogether, this establishes the compact picture of the recently studied symmetry-breaking operators for orthogonal groups by Kobayashi and Speh, gives new proofs of their main results, and extends them to unitary groups.

1. Introduction

Representation theory of reductive Lie groups consists to a large extent in the study of the structure of standard families of representations, for example principal series

MSC2010: primary 22E46; secondary 05E10, 17B15.

Keywords: symmetry-breaking operators, intertwining operators, Harish-Chandra modules, principal series, spectrum-generating operator.

representations. Here intertwining operators, such as the classical Knapp–Stein operators, play an important role, and they also provide important examples of integral kernel operators appearing in classical harmonic analysis. Recently similar operators have been introduced in [Kobayashi 2015] in connection with branching laws, i.e., the study of how representations behave when restricted to a closed subgroup of the original group; see also [Kobayashi and Speh 2015; Möllers et al. 2016a]. Again these are integral kernel operators, now intertwining with respect to the subgroup, and they appear to be very natural objects, not only for the problem of restricting representations, see [Möllers and Oshima 2015], but also for questions in classical harmonic analysis and automorphic forms, see [Möllers and Ørsted 2017; Möllers et al. 2016c].

In this paper we shall give an alternative approach to this new class of symmetry-breaking operators, namely one based on the Harish-Chandra module, i.e., the K-finite vectors in the representation, in analogy with the idea of spectrum-generating operators [Branson et al. 1996]. This gives new proofs of the main results of [Kobayashi and Speh 2015] and generalizes these results to unitary groups. Moreover, our more algebraic framework provides an alternative proof of the discrete spectrum in certain unitary representations.

The approach is quite general and discussed in the first part of the paper, while in the second part we carry out all details for the real conformal case and the CR case.

1A. ***Symmetry-breaking operators.*** Let G be a reductive Lie group with compact center and $G' \subseteq G$ a reductive subgroup also with compact center. For irreducible smooth representations π of G and τ of G' the space

$$\mathrm{Hom}_{G'}(\pi|_{G'}, \tau)$$

of continuous G'-intertwining operators between π and τ and its dimension $m(\pi, \tau)$ have received considerable attention recently, in particular in connection with multiplicity-1 statements asserting that $m(\pi, \tau) \leq 1$ for certain pairs (G, G') of classical groups such as $(\mathrm{GL}(n, \mathbb{R}), \mathrm{GL}(n-1, \mathbb{R}))$, $(\mathrm{O}(p, q), \mathrm{O}(p, q-1))$ or $(\mathrm{U}(p, q), \mathrm{U}(p, q-1))$; see [Sun and Zhu 2012]. A more refined problem is to determine whether for given representations π and τ there exist nontrivial G'-intertwining operators $\pi|_{G'} \to \tau$, also called *symmetry-breaking operators* in [Kobayashi 2015], and to classify them. For the pair $(G, G') = (\mathrm{O}(1, n), \mathrm{O}(1, n-1))$ this question was completely answered in [Kobayashi and Speh 2015] in the case where π and τ are spherical principal series representations, and in joint work with Y. Oshima we generalized in [Möllers et al. 2016a] their construction of symmetry-breaking operators to a large class of symmetric pairs.

Instead of studying this problem in the smooth category we attempt to apply the "spectrum-generating method" by Branson, Ólafsson, and Ørsted [Branson et al. 1996] in the study of intertwining operators in the category of $(\mathfrak{g}', K')$-modules, and verify a conjecture by Kobayashi on the automatic continuity of symmetry-breaking operators between Harish-Chandra modules. To given smooth admissible representations π of G and τ of G' one can associate the underlying Harish-Chandra modules π_{HC} and τ_{HC}. These are admissible $(\mathfrak{g}, K)$-modules, resp. $(\mathfrak{g}', K')$-modules, realized on the spaces of K-finite, resp. K'-finite, vectors of π, resp. τ, where $K \subseteq G$ and $K' \subseteq G'$ are maximal compact subgroups. We consider the space

$$\mathrm{Hom}_{(\mathfrak{g}',K')}(\pi_{\mathrm{HC}}|_{(\mathfrak{g}',K')}, \tau_{\mathrm{HC}})$$

of intertwining operators in the category of Harish-Chandra modules. The natural restriction map

$$\mathrm{Hom}_{G'}(\pi|_{G'}, \tau) \to \mathrm{Hom}_{(\mathfrak{g}',K')}(\pi_{\mathrm{HC}}|_{(\mathfrak{g}',K')}, \tau_{\mathrm{HC}}) \tag{1-1}$$

is injective but in general not surjective and hence there might be more intertwining operators in the category of Harish-Chandra modules than in the smooth category. According to [Kobayashi 2014, Remark 10.2 (4)] it is plausible that this map is surjective if the space $(G \times G')/\operatorname{diag}(G')$ is real spherical. (Note that for $G' = G$ the map is surjective by the Casselman–Wallach theorem.)

We remark that for $(G, G') = (\mathrm{GL}(2, \mathbb{F}) \times \mathrm{GL}(2, \mathbb{F}), \mathrm{GL}(2, \mathbb{F}))$, $\mathbb{F} = \mathbb{R}, \mathbb{C}$, and $(G, G') = (\mathrm{GL}(2, \mathbb{C}), \mathrm{GL}(2, \mathbb{R}))$ intertwining operators between Harish-Chandra modules were previously studied in [Loke 2001] using explicit computations.

1B. ***Symmetry breaking of principal series.*** In this paper we outline a method to classify symmetry-breaking operators between the Harish-Chandra modules of principal series representations induced from maximal parabolic subgroups, and their composition factors. Let $P = MAN \subseteq G$ be a maximal parabolic subgroup of G such that $P' = P \cap G' = M'A'N'$ is maximal parabolic in G' and write $\mathfrak{a}$ and $\mathfrak{a}'$ for the Lie algebras of A and A'. Fix $\nu \in \mathfrak{a}^*$ such that the roots of (P, A) are given by $\{\nu, 2\nu, \ldots, q\nu\}$ and do similarly for $\nu' \in (\mathfrak{a}')^*$. Consider the principal series representations (smooth normalized parabolic induction)

$$\pi_{\xi,r} = \mathrm{Ind}_P^G(\xi \otimes e^{r\nu} \otimes \mathbf{1}), \qquad \tau_{\xi',r} = \mathrm{Ind}_{P'}^{G'}(\xi' \otimes e^{r'\nu'} \otimes \mathbf{1}),$$

where ξ and ξ' are finite-dimensional representations of M and M' and $r, r' \in \mathbb{C}$. Let $\xi' = \xi|_{M'}$ and assume that for all K-types α of $\pi_{\xi,r}$ and all K'-types α' of $\tau_{\xi',r'}$ the multiplicity-free properties

$$\dim \mathrm{Hom}_K(\alpha, \pi_{\xi,r}|_K) \le 1,$$
$$\dim \mathrm{Hom}_{K'}(\alpha', \tau_{\xi',r'}|_{K'}) \le 1,$$
$$\dim \mathrm{Hom}_{K'}(\alpha|_{K'}, \alpha') \le 1.$$

hold; i.e., $\pi_{\xi,r}$ is K-multiplicity-free, $\tau_{\xi',r'}$ is K'-multiplicity-free, and every K-type in $\pi_{\xi,r}$ is K'-multiplicity-free.

Let $T : (\pi_{\xi,r})_{\mathrm{HC}} \to (\tau_{\xi',r'})_{\mathrm{HC}}$ be a $(\mathfrak{g}', K')$-intertwining operator; then T is in particular K'-intertwining. Consider a pair $(\alpha; \alpha')$ consisting of a K-type α in $\pi_{\xi,r}$ and a K'-type α' in $\tau_{\xi',r'}$ which also occurs in $\alpha|_{K'}$. By the multiplicity-free assumptions the restriction of T to the K'-type α' inside the K-type α in $\pi_{\xi,r}$ maps to the K'-type α' in $\tau_{\xi',r'}$ and is unique up to a scalar $t_{\alpha,\alpha'} \in \mathbb{C}$ (see Section 3A for the precise definition). This encodes every K'-intertwining operator $T : (\pi_{\xi,r})_{\mathrm{HC}} \to (\tau_{\xi',r'})_{\mathrm{HC}}$ into scalars $t_{\alpha,\alpha'}$. Using the method of spectrum-generating operators by Branson, Ólafsson, and Ørsted [Branson et al. 1996] we prove:

Theorem A (see Theorem 3.4 and Corollary 3.6). *Let $T : (\pi_{\xi,r})_{\mathrm{HC}} \to (\tau_{\xi',r'})_{\mathrm{HC}}$ be a K'-intertwining operator given by scalars $t_{\alpha,\alpha'}$. Then T is $(\mathfrak{g}', K')$-intertwining if and only if for all pairs $(\alpha; \alpha')$ and every K'-type β' the following relation holds*:

$$\sum_{\substack{\beta \\ (\alpha;\alpha')\leftrightarrow(\beta;\beta')}} \lambda^{\beta,\beta'}_{\alpha,\alpha'}(\sigma_\beta - \sigma_\alpha + 2r)t_{\beta,\beta'} = (\sigma'_{\beta'} - \sigma'_{\alpha'} + 2r')t_{\alpha,\alpha'}. \tag{1-2}$$

Here we write $(\alpha; \alpha') \leftrightarrow (\beta; \beta')$ if the K'-type β' inside the K-type β in $\pi_{\xi,r}$ can be reached from α' inside α by a single application of $\pi_{\xi,r}(\mathfrak{g}')$ for generic $r \in \mathbb{C}$ (see Section 3B for details). Further, σ_α and $\sigma'_{\alpha'}$ as well as $\lambda^{\beta,\beta'}_{\alpha,\alpha'}$ are certain constants depending only on the representations ξ and ξ' (see Sections 2C and 3C for their definition).

We note that the relations characterizing intertwining operators depend linearly on the induction parameters r and r' and turn the representation-theoretic problem of classifying symmetry-breaking operators into a combinatorial problem. We also remark that Theorem A admits a slight modification characterizing also intertwining operators between any subquotients of $\pi_{\xi,r}$ and $\tau_{\xi',r'}$ (see Remark 3.5).

1C. *Examples.* For the two pairs $(G, G') = (\mathrm{O}(1, n), \mathrm{O}(1, n-1))$, $n \geq 3$, and $(\mathrm{U}(1, n), \mathrm{U}(1, n-1))$, $n \geq 2$, we explicitly write down the linear relations for the scalars $t_{\alpha,\alpha'}$ characterizing intertwining operators in the case where $\xi = \mathbf{1}$ is the trivial representation (see Theorems 4.1 and 5.1), and use these relations to compute multiplicities. For the statements we abbreviate $\pi_r = \pi_{\mathbf{1},r}$ and $\tau_{r'} = \tau_{\mathbf{1},r'}$. If $\mathcal{V}$ is a $(\mathfrak{g}, K)$-module and $\mathcal{W}$ a $(\mathfrak{g}', K')$-module we write

$$m(\mathcal{V}, \mathcal{W}) = \dim \mathrm{Hom}_{(\mathfrak{g}',K')}(\mathcal{V}|_{(\mathfrak{g}',K')}, \mathcal{W}).$$

We note that much of the notation used here follows [Kobayashi and Speh 2015].

Theorem B (see Theorems 4.2(1) and 5.2(1)).

(1) *For $(G, G') = (\mathrm{O}(1, n), \mathrm{O}(1, n-1))$ we have*

$$m((\pi_r)_{\mathrm{HC}}, (\tau_{r'})_{\mathrm{HC}}) = \begin{cases} 1 & \text{for } (r, r') \in \mathbb{C}^2 \setminus L_{\mathrm{even}}, \\ 2 & \text{for } (r, r') \in L_{\mathrm{even}}, \end{cases}$$

where $L_{\mathrm{even}} = \left\{\left(-\frac{n-1}{2} - i, -\frac{n-2}{2} - j\right) : i, j \in \mathbb{N},\ i - j \in 2\mathbb{N}\right\}$.

(2) *For* $(G, G') = (\mathrm{U}(1, n), \mathrm{U}(1, n-1))$ *we have*

$$m((\pi_r)_{\mathrm{HC}}, (\tau_{r'})_{\mathrm{HC}}) = \begin{cases} 1 & \text{for } (r, r') \in \mathbb{C}^2 \setminus L, \\ 2 & \text{for } (r, r') \in L, \end{cases}$$

where $L = \{(-n-2i, -(n-1)-2j) : i, j \in \mathbb{N},\ j \leq i\}$.

Multiplicity 2 does not contradict the multiplicity-1 statements for the above pairs (G, G'), because for $(r, r') \in L_{\mathrm{even}}$, resp. L, both representations π_r and $\tau_{r'}$ are reducible. In the case $(G, G') = (\mathrm{O}(1, n), \mathrm{O}(1, n-1))$ the representation $(\pi_r)_{\mathrm{HC}}$ is reducible if and only if $r = \pm\left(\frac{n-1}{2} + i\right)$, $i \in \mathbb{N}$, its composition factors consisting of a finite-dimensional subrepresentation $\mathcal{F}(i)$ and an infinite-dimensional unitarizable quotient $\mathcal{T}(i)$. Similarly, in the case $(G, G') = (\mathrm{U}(1, n), \mathrm{U}(1, n-1))$ the representation $(\pi_r)_{\mathrm{HC}}$ is reducible if and only if $r = \pm(n + 2i)$, $i \in \mathbb{N}$, and its composition factors consist of a finite-dimensional subrepresentation $\mathcal{F}(i)$, two proper subquotients $\mathcal{T}_{\pm}(i)$, and a unitarizable quotient $\mathcal{T}(i)$. Write $\mathcal{F}'(j)$, $\mathcal{T}'_{\pm}(j)$ and $\mathcal{T}'(j)$ for the corresponding composition factors of $(\tau_{r'})_{\mathrm{HC}}$ at $r' = -\frac{n-2}{2} - j$, resp. $r' = -(n-1) - 2j$.

Theorem C (see Theorems 4.2(2) and 5.2(2)).
(1) *For* $(G, G') = (\mathrm{O}(1, n), \mathrm{O}(1, n-1))$, *the multiplicities* $m(\mathcal{V}, \mathcal{W})$ *are given by*

$\mathcal{V}\downarrow$ $\mathcal{W}\rightarrow$	$\mathcal{F}'(j)$	$\mathcal{T}'(j)$
$\mathcal{F}(i)$	1	0
$\mathcal{T}(i)$	0	1

for $i - j \in 2\mathbb{N}$,

$\mathcal{V}\downarrow$ $\mathcal{W}\rightarrow$	$\mathcal{F}'(j)$	$\mathcal{T}'(j)$
$\mathcal{F}(i)$	0	0
$\mathcal{T}(i)$	1	0

otherwise.

(2) *For* $(G, G') = (\mathrm{U}(1, n), \mathrm{U}(1, n-1))$, *the multiplicities* $m(\mathcal{V}, \mathcal{W})$ *are given by*

$\mathcal{V}\downarrow$ $\mathcal{W}\rightarrow$	$\mathcal{F}'(j)$	$\mathcal{T}'_+(j)$	$\mathcal{T}'_-(j)$	$\mathcal{T}'(j)$
$\mathcal{F}(i)$	1	0	0	0
$\mathcal{T}_+(i)$	0	1	0	0
$\mathcal{T}_-(i)$	0	0	1	0
$\mathcal{T}(i)$	0	0	0	1

for $j \leq i$,

$\mathcal{V}\downarrow$ $\mathcal{W}\rightarrow$	$\mathcal{F}'(j)$	$\mathcal{T}'_+(j)$	$\mathcal{T}'_-(j)$	$\mathcal{T}'(j)$
$\mathcal{F}(i)$	0	0	0	0
$\mathcal{T}_+(i)$	0	0	0	0
$\mathcal{T}_-(i)$	0	0	0	0
$\mathcal{T}(i)$	1	0	0	0

otherwise.

We further construct a basis of $\mathrm{Hom}_{(\mathfrak{g}',K')}((\pi_r)_{\mathrm{HC}}|_{(\mathfrak{g}',K')},(\tau_{r'})_{\mathrm{HC}})$ for all $r,r'\in\mathbb{C}$ by solving the relations (1-2) explicitly. More precisely, we find a family $(t_{\alpha,\alpha'}(r,r'))_{\alpha,\alpha'}$ consisting of rational functions in $r,r'\in\mathbb{C}$ that solve the relations (1-2). Renormalizing the functions $t_{\alpha,\alpha'}(r,r')$ gives a family $(t^{(1)}_{\alpha,\alpha'}(r,r'))_{\alpha,\alpha'}$ of holomorphic functions in $r,r'\in\mathbb{C}$ satisfying the relations (1-2). By Theorem A this constructs intertwining operators $T^{(1)}(r,r')$ depending holomorphically on $r,r'\in\mathbb{C}$. We show that

$$T^{(1)}(r,r')=0 \quad \text{if and only if} \quad (r,r')\in L_{\mathrm{even}},\ \text{resp. } L.$$

For each $(r,r')\in L_{\mathrm{even}}$, resp. L, the holomorphic function $T^{(1)}(r,r')$ can be renormalized along two different affine complex lines through (r,r'), and one obtains two different nontrivial operators $T^{(2)}(r,r')$, $T^{(3)}(r,r')$ for every $(r,r')\in L_{\mathrm{even}}$, resp. L (see Propositions 4.6 and 5.6 for details).

Theorem D (see Theorems 4.9 and 5.8 and Remarks 4.10 and 5.9). *For the pair* $(G,G')=(\mathrm{O}(1,n),\mathrm{O}(1,n-1))$, *resp.* $(\mathrm{U}(1,n),\mathrm{U}(1,n-1))$, *we have*

$$\mathrm{Hom}_{(\mathfrak{g}',K')}((\pi_r)_{\mathrm{HC}}|_{(\mathfrak{g}',K')},(\tau_{r'})_{\mathrm{HC}}) = \begin{cases}\mathbb{C}T^{(1)}(r,r') & \text{for } (r,r')\in\mathbb{C}^2\setminus\mathcal{L},\\ \mathbb{C}T^{(2)}(r,r')\oplus\mathbb{C}T^{(3)}(r,r') & \text{for } (r,r')\in\mathcal{L},\end{cases}$$

where $\mathcal{L}=L_{\mathrm{even}}$, *resp. L. Moreover, by composing* $T^{(1)}(r,r')$ *with embeddings and quotient maps for the composition factors of* π_r *and* $\tau_{r'}$, *and renormalizing along certain affine complex lines, one can obtain every intertwining operator between arbitrary composition factors of* $(\pi_r)_{\mathrm{HC}}$ *and* $(\tau_{r'})_{\mathrm{HC}}$.

The previous theorem shows that basically all the information about intertwining operators between spherical principal series of G and G' and their composition factors is contained in the single holomorphic family $T^{(1)}(r,r')$ of intertwiners.

Finally we turn to the question of whether every intertwining operator between the Harish-Chandra modules $(\pi_r)_{\mathrm{HC}}$ and $(\tau_{r'})_{\mathrm{HC}}$ lifts to an intertwining operator between the smooth globalizations π_r and $\tau_{r'}$, i.e., the question of whether (1-1) is an isomorphism.

Theorem E (see Corollaries 4.12 and 5.11). *For the pairs* $(G,G')=(\mathrm{O}(1,n),\mathrm{O}(1,n-1))$ *and* $(\mathrm{U}(1,n),\mathrm{U}(1,n-1))$ *every intertwining operator between* $(\pi_r)_{\mathrm{HC}}$ *and* $(\tau_{r'})_{\mathrm{HC}}$ *(resp. any of their subquotients) extends to a continuous intertwining operator between* π_r *and* $\tau_{r'}$ *(resp. the Casselman–Wallach completions of the subquotients). In particular, the injective map* (1-1) *is surjective for all spherical principal series representations and their subquotients.*

This verifies Kobayashi's conjecture [2014, Remark 10.2 (4)] in the above cases.

For $(G,G')=(\mathrm{O}(1,n),\mathrm{O}(1,n-1))$ the analogues of Theorems B, C and D in the smooth category, i.e., for π_r and $\tau_{r'}$ instead of $(\pi_r)_{\mathrm{HC}}$ and $(\tau_{r'})_{\mathrm{HC}}$, were established in [Kobayashi and Speh 2015] using analytic techniques. With Theorem E we obtain a new proof of their results as well as the corresponding results for $(G, G') = (\mathrm{U}(1,n), \mathrm{U}(1,n-1))$.

1D. *Application.* For $(G, G') = (\mathrm{O}(1,n), \mathrm{O}(1,n-1))$ we further present an application of the classification of symmetry-breaking operators. In Theorem 4.14 we use the explicit formula for the numbers $t_{\alpha,\alpha'}$ to construct discrete components in the restriction of certain unitary representations of G to G'. The representations in question are either spherical complementary series representations (i.e., those π_r which are unitarizable) or the unitarizable quotients $\mathcal{T}(i)$. This extends and gives new proofs of previous results of [Speh and Venkataramana 2011; Zhang 2015; Kobayashi and Speh 2015; Möllers and Oshima 2015] (see Remark 4.15). Analogous results hold for $(G, G') = (\mathrm{U}(1,n), \mathrm{U}(1,n-1))$.

1E. *Structure of the paper.* In Section 2 we fix the notation for principal series representations and recall the method of spectrum-generating operators [Branson et al. 1996]. This method is applied in Section 3 to obtain an equivalent characterization of intertwining operators in the category of $(\mathfrak{g}', K')$-modules by means of scalar identities. After this quite general approach, we study in Section 4 the special case $(G, G') = (\mathrm{O}(1,n), \mathrm{O}(1,n-1))$ in detail and give some applications. Finally, in Section 5 we repeat the same procedure for $(G, G') = (\mathrm{U}(1,n), \mathrm{U}(1,n-1))$ providing a new classification of symmetry-breaking operators in this example. Appendix A contains some basic properties of Gegenbauer and Jacobi polynomials which are used in Appendix B to describe explicit branching laws for real and complex spherical harmonics.

Throughout we will use the notation $\mathbb{N} = \{0, 1, 2, \dots\}$.

2. Preliminaries

We fix the necessary notation, discuss induced representations and the method of the spectrum-generating operator by Branson, Ólafsson, and Ørsted [Branson et al. 1996].

2A. *Compatible maximal parabolic subgroups.* Let G be a reductive Lie group with compact center and $G' \subseteq G$ a reductive subgroup also with compact center. Denote by $\mathfrak{g}$ and $\mathfrak{g}'$ the Lie algebras of G and G'. Choose a maximal parabolic subgroup $P \subseteq G$ with the property that $P' = P \cap G'$ is maximal parabolic in G' and write $P = MAN$ and $P' = M'A'N'$ for the Langlands decompositions of P and P'. We fix a Cartan involution θ of G which leaves G' and the Levi subgroups MA and $M'A'$ invariant. Write $K = G^\theta$ and $K' = (G')^\theta$ for the corresponding fixed point subgroups of G and G' which are maximal compact and denote by $\mathfrak{k}$ and $\mathfrak{k}'$ their Lie algebras. Let $\mathfrak{s}$ and $\mathfrak{s}'$ be the (-1)-eigenspaces of θ on $\mathfrak{g}$ and $\mathfrak{g}'$ so that

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{s}, \quad \mathfrak{g}' = \mathfrak{k}' \oplus \mathfrak{s}'.$$

Example 2.1. (1) Let (G, G') be one of the pairs

$$(\mathrm{O}(1,n), \mathrm{O}(1,n-1)), \quad (\mathrm{U}(1,n), \mathrm{U}(1,n-1)),$$
$$(\mathrm{Sp}(1,n), \mathrm{Sp}(1,n-1)), \quad (F_{4(-20)}, \mathrm{Spin}(8,1)).$$

Then one can choose the minimal parabolic P such that $P' = P \cap G'$ is minimal parabolic in G'. Since G and G' are of rank 1, minimal parabolics are maximal and hence satisfy our assumptions.

(2) Let

$$(G, G') = (\mathrm{SL}(n,\mathbb{R}), \mathrm{SL}(n-1,\mathbb{R})),$$

with G' embedded in G as the upper-left block. Then all standard maximal parabolics $P = P_{p,q} = (S(\mathrm{GL}(p,\mathbb{R}) \times \mathrm{GL}(q,\mathbb{R}))) \ltimes \mathbb{R}^{p\times q}$ corresponding to the partition $n = p + q$ with $q > 1$ satisfy the assumptions. In this case $P' = P \cap G'$ is the standard maximal parabolic of G' corresponding to the partition $n-1 = p+(q-1)$.

2B. ***Principal series representations.*** For any finite-dimensional representation (ξ, V_ξ) of M and any $\nu \in \mathfrak{a}_\mathbb{C}^*$, where $\mathfrak{a}$ denotes the Lie algebra of A, consider the induced representation $\mathrm{Ind}_P^G(\xi \otimes e^\nu \otimes \mathbf{1})$ (normalized smooth parabolic induction). This representation is realized on the space

$$\mathcal{E}(G;\xi,\nu) = \{F \in C^\infty(G, V_\xi) : F(gman) = a^{-\nu-\rho}\xi(m)^{-1}F(g) \\ \text{for all } g \in G,\ man \in MAN\},$$

where $\rho = \frac{1}{2}\operatorname{tr}\operatorname{ad}|_\mathfrak{n} \in \mathfrak{a}^*$. The group G acts on $\mathcal{E}(G;\xi,\nu)$ by the left-regular action. Since $G = KP$, restriction to K is an isomorphism $\mathcal{E}(G;\xi,\nu) \to \mathcal{E}(K;\xi|_{M\cap K})$, where

$$\mathcal{E}(K;\xi|_{M\cap K}) = \{F \in C^\infty(K, V_\xi) : F(km) = \xi(m)^{-1}F(k) \text{ for all } k \in K, m \in M\cap K\}.$$

Let $\pi_{\xi,\nu}$ denote the action of G on $\mathcal{E}(K;\xi|_{M\cap K})$ which makes this isomorphism G-equivariant. Then $(\pi_{\xi,\nu}, \mathcal{E}(K;\xi|_{M\cap K}))$ is a smooth admissible representation of G. The restriction of $\pi_{\xi,\nu}$ to K is simply the left-regular representation of K on $\mathcal{E}(K;\xi|_{M\cap K})$.

Corresponding to the smooth representation $\pi_{\xi,\nu}$ we consider its underlying $(\mathfrak{g}, K)$-module $(\pi_{\xi,\nu})_{\mathrm{HC}}$ realized on the space $\mathcal{E} = \mathcal{E}(K;\xi|_{M\cap K})_K$ of K-finite vectors. Abusing notation we denote the action of the Lie algebra $\mathfrak{g}$ on $\mathcal{E}$ also by $\pi_{\xi,\nu}$. Then the restriction of $(\pi_{\xi,\nu})_{\mathrm{HC}}$ to K decomposes as

$$\mathcal{E} = \bigoplus_{\alpha \in \widehat{K}} \mathcal{E}(\alpha),$$

with $\mathcal{E}(\alpha)$ being the α-isotypic component in $\mathcal{E}$. Note that $\mathcal{E}$ and hence its decomposition into K-isotypic components is independent of $\nu \in \mathfrak{a}_\mathbb{C}^*$ and only depends on ξ.

Similarly we consider $\tau_{\xi',\nu'} = \mathrm{Ind}_{P'}^{G'}(\xi' \otimes e^{\nu'} \otimes \mathbf{1})$ for a finite-dimensional representation $(\xi', V_{\xi'})$ of M' and an element $\nu' \in (\mathfrak{a}')^*_{\mathbb{C}}$, and its underlying $(\mathfrak{g}', K')$-module $(\tau_{\xi',\nu'})_{\mathrm{HC}}$ realized on the space $\mathcal{E}' = \mathcal{E}(K'; \xi'|_{K'\cap M'})_{K'}$. As above we decompose the restriction of $(\tau_{\xi',\nu'})_{\mathrm{HC}}$ to K'

$$\mathcal{E}' = \bigoplus_{\alpha' \in \widehat{K}'} \mathcal{E}'(\alpha'),$$

with $\mathcal{E}'(\alpha')$ being the α'-isotypic component.

2C. ***The spectrum-generating operator.*** Since P is a maximal parabolic subgroup we have $\dim \mathfrak{a} = 1$ and we can choose $H \in \mathfrak{a}$ such that the eigenvalues of $\mathrm{ad}(H)$ on the Lie algebra $\mathfrak{n}$ of N are $1, \dots, q$. Define $\nu \in \mathfrak{a}^*$ by $\nu(H) = 1$; then $\Sigma(\mathfrak{g}, \mathfrak{a}) = \{\pm\nu, \dots, \pm q\nu\}$. We abbreviate $\pi_{\xi,r} = \pi_{\xi,r\nu}$ for $r \in \mathbb{C}$.

Let B be an invariant nondegenerate symmetric bilinear form on $\mathfrak{g}$ normalized by $B(H, H) = 1$. For $1 \le j \le q$ let

$$\mathfrak{k}_j = \mathfrak{k} \cap (\mathfrak{g}_{j\nu} + \mathfrak{g}_{-j\nu}).$$

Choose a basis $(X_{j,k})_k$ of $\mathfrak{k}_j$, denote by $(X'_{j,k})_k$ the corresponding dual basis with respect to B and put

$$\mathrm{Cas}_j = \sum_k X_{j,k} X'_{j,k}.$$

Then Cas_j is an element of $\mathcal{U}(\mathfrak{k})$, the universal enveloping algebra of $\mathfrak{k}$. Clearly the elements $\mathrm{Cas}_j \in \mathcal{U}(\mathfrak{k})$ do not depend on the choice of the corresponding bases. Following [Branson et al. 1996] we define the *spectrum-generating operator* as the second-order element in $\mathcal{U}(\mathfrak{k})$ given by

$$\mathcal{P} = \sum_{j=1}^{q} j^{-1} \mathrm{Cas}_j.$$

We remark that even though the spaces $\mathfrak{k}_j$ do not form subalgebras the operator $\mathcal{P}$ can be written as a rational linear combination of Casimir elements of subalgebras of $\mathfrak{k}$; see [Branson et al. 1996, Remark 2.4]. Since the left-regular representation of K on $\mathcal{E}$ commutes with the right-action $\mathcal{R}_{\mathcal{P}}$ of $\mathcal{P}$ the restriction of $\mathcal{R}_{\mathcal{P}}$ to each isotypic component $\mathcal{E}(\alpha)$ is a linear transformation

$$\sigma_\alpha = \sigma_{\alpha,\xi|_{M\cap K}} \in \mathrm{End}\,\mathcal{E}(\alpha)$$

which only depends on ξ but not on ν.

Similarly we define $H' \in \mathfrak{a}'$, $\nu' \in (\mathfrak{a}')^*$ and choose an invariant nondegenerate symmetric bilinear form B' on $\mathfrak{g}'$ with $B'(H', H') = 1$. Let $\mathcal{P}'$ denote the spectrum-generating operator for G' and write $\sigma'_{\alpha'} \in \mathrm{End}\,\mathcal{E}'(\alpha')$ for the restriction of $\mathcal{R}_{\mathcal{P}'}$ to the isotypic component $\mathcal{E}'(\alpha')$.

2D. ***Reduction to the cocycle.*** For each $X \in \mathfrak{g}_{\mathbb{C}}$ we define a scalar-valued function $\omega(X)$ on K by

$$\omega(X)(k) = B(\mathrm{Ad}(k^{-1})X, H), \quad k \in K,$$

where we extend B to a symmetric $\mathbb{C}$-bilinear form on $\mathfrak{g}_{\mathbb{C}}$. This defines a K-equivariant map

$$\omega : \mathfrak{g}_{\mathbb{C}} \to \mathcal{E}(K; \mathbf{1}) \cong C^\infty(K/(M \cap K)),$$

where $\mathbf{1}$ is the trivial $M \cap K$-representation. The map ω is called a *cocycle*. Note that ω vanishes on $\mathfrak{k}_{\mathbb{C}}$. Let $m(\omega(X))$ denote the multiplication operator

$$\mathcal{E} \to \mathcal{E}, \quad \varphi \mapsto \omega(X)\varphi.$$

For $\alpha, \beta \in \widehat{K}$ with $\mathcal{E}(\alpha), \mathcal{E}(\beta) \neq 0$ we let

$$\omega_\alpha^\beta(X) = \mathrm{proj}_{\mathcal{E}(\beta)} \circ m(\omega(X))|_{\mathcal{E}(\alpha)}, \quad X \in \mathfrak{g}_{\mathbb{C}},$$

where $\mathrm{proj}_{\mathcal{E}(\beta)}$ denotes the projection from $\mathcal{E}$ onto $\mathcal{E}(\beta)$. We can now express the differential representation $\pi_{\xi,r}$ of $\mathfrak{g}$ on $\mathcal{E}$ in terms of the cocycle ω and the maps σ_α:

Theorem 2.2 [Branson et al. 1996, Corollary 2.6]. *For $X \in \mathfrak{s}_{\mathbb{C}}$ and any $\alpha, \beta \in \widehat{K}$ with $\mathcal{E}(\alpha), \mathcal{E}(\beta) \neq 0$ we have*

$$\mathrm{proj}_{\mathcal{E}(\beta)} \circ \pi_{\xi,r}(X)|_{\mathcal{E}(\alpha)} = \tfrac{1}{2}(\sigma_\beta \omega_\alpha^\beta(X) - \omega_\alpha^\beta(X)\sigma_\alpha + 2r\omega_\alpha^\beta(X)). \tag{2-1}$$

Similarly we denote by $\omega'(X)$ the corresponding cocycle for G' and by $\omega_{\alpha'}^{\beta'}(X)$ the corresponding map from $\mathcal{E}'(\alpha')$ to $\mathcal{E}'(\beta')$. Then we obtain for $X \in \mathfrak{s}'_{\mathbb{C}}$ and any $\alpha', \beta' \in \widehat{K}'$ with $\mathcal{E}'(\alpha'), \mathcal{E}'(\beta') \neq 0$ the analogous identity

$$\mathrm{proj}_{\mathcal{E}'(\beta')} \circ \tau_{\xi',r'}(X)|_{\mathcal{E}'(\alpha')} = \tfrac{1}{2}(\sigma'_{\beta'} \omega_{\alpha'}^{\beta'}(X) - \omega_{\alpha'}^{\beta'}(X)\sigma'_{\alpha'} + 2r'\omega_{\alpha'}^{\beta'}(X)). \tag{2-2}$$

3. The compact picture of symmetry-breaking operators

Consider the admissible $(\mathfrak{g}, K)$-module $(\pi_{\xi,r})_{\mathrm{HC}}$. Then its restriction $(\pi_{\xi,r})_{\mathrm{HC}}|_{(\mathfrak{g}',K')}$ is a $(\mathfrak{g}', K')$-module which is in general not admissible anymore. However, we can still study the space

$$\mathrm{Hom}_{(\mathfrak{g}',K')}((\pi_{\xi,r})_{\mathrm{HC}}|_{(\mathfrak{g}',K')}, (\tau_{\xi',r'})_{\mathrm{HC}})$$

of intertwining operators between the $(\mathfrak{g}', K')$-modules. In this section we use Theorem 2.2 to characterize these intertwining operators in terms of their action on the K'-isotypic components in the K-types $\mathcal{E}(\alpha)$.

3A. ***Relating K-types and K'-types.*** From now on we assume that both $\mathcal{E}$ and $\mathcal{E}'$ are multiplicity-free; i.e.,

$$\text{(MF1)} \qquad \dim \operatorname{Hom}_K(\alpha,\ \mathcal{E}), \dim \operatorname{Hom}_{K'}(\alpha', \mathcal{E}') \leq 1 \quad \text{for all } \alpha \in \widehat{K},\ \alpha' \in \widehat{K}'.$$

This implies by Schur's lemma that the maps σ_α and $\sigma'_{\alpha'}$ are scalars which we denote by the same symbols. We further assume that each K'-type $\mathcal{E}'(\alpha') \neq 0$ occurs at most once in each K-type $\mathcal{E}(\alpha) \neq 0$; i.e.,

$$\text{(MF2)} \qquad \dim \operatorname{Hom}_{K'}(\mathcal{E}(\alpha), \mathcal{E}'(\alpha')) \leq 1 \quad \text{for all } \mathcal{E}(\alpha), \mathcal{E}'(\alpha') \neq 0.$$

Each K-isotypic component $\mathcal{E}(\alpha)$ decomposes under the action of $K' \subseteq K$ into

$$\mathcal{E}(\alpha) = \bigoplus_{\alpha' \in \widehat{K}'} \mathcal{E}(\alpha; \alpha'),$$

where $\mathcal{E}(\alpha; \alpha')$ is the α'-isotypic component in $\mathcal{E}(\alpha)$. Then our assumptions imply that if $\mathcal{E}(\alpha; \alpha'), \mathcal{E}'(\alpha') \neq 0$ then $\mathcal{E}(\alpha; \alpha') \cong \mathcal{E}'(\alpha')$. In all such cases we fix an isomorphism

$$R_{\alpha,\alpha'} : \mathcal{E}(\alpha; \alpha') \xrightarrow{\sim} \mathcal{E}'(\alpha').$$

To simplify notation, let $R_{\alpha,\alpha'} = 0$ whenever $\mathcal{E}'(\alpha') = 0$, so that we have surjective K'-equivariant maps $R_{\alpha,\alpha'} : \mathcal{E}(\alpha; \alpha') \to \mathcal{E}'(\alpha')$ for all $\mathcal{E}(\alpha; \alpha') \neq 0$.

In applications, it is often useful to choose a natural isomorphism $\mathcal{E}(\alpha; \alpha') \cong \mathcal{E}'(\alpha')$ relating K-types and K'-types. For this we study the restriction of functions from K to K'. Assume for simplicity that $\xi' = \xi|_{M'}$. In this case we can consider the restriction operator

$$\operatorname{rest} : \mathcal{E} \to \mathcal{E}', \quad \varphi \mapsto \varphi|_{K'}.$$

This operator is K'-equivariant and hence, if rest is nonzero on some K'-type $\mathcal{E}(\alpha; \alpha')$ in $\mathcal{E}$ then $\operatorname{rest}|_{\mathcal{E}(\alpha;\alpha')}$ is an isomorphism onto $\mathcal{E}'(\alpha')$ by Schur's lemma. However, rest might also vanish on some $\mathcal{E}(\alpha; \alpha')$ and therefore we need to combine the restriction with differentiation in the normal direction.

For this we write

$$\mathfrak{k} = (\mathfrak{m} \cap \mathfrak{k}) \oplus \mathfrak{q},$$

where $\mathfrak{q}$ is the orthogonal complement of $(\mathfrak{m} \cap \mathfrak{k})$ in $\mathfrak{k}$ with respect to the invariant form B. Note that $M \cap K$ acts on $\mathfrak{q}$. Similarly

$$\mathfrak{k}' = (\mathfrak{m}' \cap \mathfrak{k}') \oplus \mathfrak{q}'.$$

Let $\mathfrak{q}''$ denote the orthogonal complement of $\mathfrak{q}'$ in $\mathfrak{q}$; then

$$\mathfrak{q} = \mathfrak{q}' \oplus \mathfrak{q}''.$$

We note that since $M \cap K$ acts on $\mathfrak{q}$ and $M' \cap K'$ acts on $\mathfrak{q}'$, the group $M' \cap K'$ also acts on $\mathfrak{q}''$. We then have

$$\mathfrak{k}/(\mathfrak{m} \cap \mathfrak{k}) \cong \mathfrak{k}'/(\mathfrak{m}' \cap \mathfrak{k}') \oplus \mathfrak{q}'';$$

i.e., $\mathfrak{q}''$ identifies with the normal space of $K'/(M' \cap K')$ in $K/(M \cap K)$ at the base point. Denote by $S(\mathfrak{q}'')$ the symmetric algebra over $\mathfrak{q}''$ and by $S(\mathfrak{q}'')^{M' \cap K'}$ its $(M' \cap K')$-invariants. Note that $S(\mathfrak{q}'')^{M' \cap K'}$ acts naturally from the right by differential operators on functions defined on a small neighborhood of $K'/(M' \cap K')$ in $K/(M \cap K)$.

Lemma 3.1. *Let* $(\alpha, \alpha') \in \widehat{K} \times \widehat{K}'$ *with* $\mathcal{E}(\alpha; \alpha') \neq 0$ *and* $D \in S(\mathfrak{q}'')^{M' \cap K'}$. *Then the map* $\operatorname{rest} \circ D : \mathcal{E} \to \mathcal{E}'$ *is* K'*-equivariant. In particular,*

$$(\operatorname{rest} \circ D)|_{\mathcal{E}(\alpha;\alpha')} : \mathcal{E}(\alpha; \alpha') \to \mathcal{E}'(\alpha')$$

is an isomorphism whenever it is nonzero.

Remark 3.2. Of course one could as well consider other irreducible $M' \cap K'$-subrepresentations of $S(\mathfrak{q}'')$ than the trivial one. In fact, using an idea of [Ørsted and Vargas 2004] one can construct an injective K'-equivariant map

$$\mathcal{E} = C^\infty(K \times_{M \cap K} \xi)_K \to \bigoplus_{m=0}^{\infty} C^\infty(K' \times_{M' \cap K'} (\xi \otimes S^m(\mathfrak{q}'')))_{K'}$$

and use it to relate K-types and K'-types of the induced representations $\pi_{\xi,r}$ and $\tau_{\xi',r'}$ for $\xi'|_{M' \cap K'}$ any subrepresentation of $\xi|_{M' \cap K'} \otimes S(\mathfrak{q}'')$. Lemma 3.1 can then be viewed as the special case $\xi' = \xi \otimes \mathbb{C}D$, where $D \in S(\mathfrak{q}'')^{M' \cap K'}$ and hence $\mathbb{C}D$ is the trivial $M' \cap K'$-representation.

3B. ***Intertwining operators between Harish-Chandra modules.*** Let $\mathcal{V} \subseteq \mathcal{E}$ be a $(\mathfrak{g}', K')$-submodule of $(\pi_{\xi,r})_{\mathrm{HC}}$; i.e., $\mathcal{V}$ is stable under $\pi_{\xi,r}(\mathfrak{g}')$ and stable under $\pi_{\xi,r}(K')$. A linear map $T : \mathcal{V} \to \mathcal{E}'$ is called an *intertwining operator for* $\pi_{\xi,r}$ *and* $\tau_{\xi',r'}$ if for every $v \in \mathcal{V}$ we have

$$(T \circ \pi_{\xi,r}(X))v = (\tau_{\xi',r'}(X) \circ T)v \quad \text{for all } X \in \mathfrak{g}', \tag{3-1}$$

$$(T \circ \pi_{\xi,r}(k))v = (\tau_{\xi',r'}(k) \circ T)v \quad \text{for all } k \in K'. \tag{3-2}$$

In particular an intertwining operator commutes by (3-2) with the action of K' and hence restricts to a map $T_{\alpha,\alpha'} = T|_{\mathcal{E}(\alpha;\alpha')} : \mathcal{E}(\alpha; \alpha') \to \mathcal{E}'(\alpha')$ for all $\mathcal{E}(\alpha; \alpha') \subseteq \mathcal{V}$. If $\mathcal{E}'(\alpha') = 0$ then clearly $T_{\alpha,\alpha'} = 0$. Recall that we fixed in Section 3A K'-equivariant maps $R_{\alpha,\alpha'} : \mathcal{E}(\alpha; \alpha') \to \mathcal{E}'(\alpha')$; then by Schur's lemma $T_{\alpha,\alpha'}$ is a scalar multiple of $R_{\alpha,\alpha'}$. We write

$$T_{\alpha,\alpha'} = t_{\alpha,\alpha'} \cdot R_{\alpha,\alpha'} \quad \text{for all } 0 \neq \mathcal{E}(\alpha; \alpha') \subseteq \mathcal{V}, \tag{3-3}$$

with $t_{\alpha,\alpha'} \in \mathbb{C}$.

Restricting (3-1) to $\mathcal{E}(\alpha;\alpha')$ and composing with the projection $\mathrm{proj}_{\mathcal{E}'(\beta')}$ we obtain

$$\mathrm{proj}_{\mathcal{E}'(\beta')}\circ T\circ\pi_{\xi,r}(X)|_{\mathcal{E}(\alpha;\alpha')}=\mathrm{proj}_{\mathcal{E}'(\beta')}\circ\tau_{\xi',r'}(X)\circ T|_{\mathcal{E}(\alpha;\alpha')}. \tag{3-4}$$

To simplify both sides we let

$$\omega_{\alpha,\alpha'}^{\beta,\beta'}:\mathfrak{s}'_{\mathbb{C}}\otimes\mathcal{E}(\alpha;\alpha')\to\mathcal{E}(\beta;\beta'),\quad \omega_{\alpha,\alpha'}^{\beta,\beta'}(X)=\mathrm{proj}_{\mathcal{E}(\beta;\beta')}\circ m(\omega(X))|_{\mathcal{E}(\alpha;\alpha')},$$

where we view $\omega_{\alpha,\alpha'}^{\beta,\beta'}(X)$, $X\in\mathfrak{s}'$, as a linear map $\mathcal{E}(\alpha;\alpha')\to\mathcal{E}(\beta;\beta')$. Write $(\alpha;\alpha')\to(\beta;\beta')$ if $\omega_{\alpha,\alpha'}^{\beta,\beta'}\neq 0$. The following lemma is proved along the same lines as [Branson et al. 1996, Lemma 4.4 (c)] and justifies the use of the notation $(\alpha;\alpha')\leftrightarrow(\beta;\beta')$ instead of $(\alpha;\alpha')\to(\beta;\beta')$:

Lemma 3.3. *For an orthonormal basis $(X_k)_k\subseteq\mathfrak{s}'$ put*

$$s_{\alpha,\alpha'}^{\beta,\beta'}=\sum_k\omega_{\beta,\beta'}^{\alpha,\alpha'}(X_k)\circ\omega_{\alpha,\alpha'}^{\beta,\beta}(X_k).$$

Then $s_{\alpha,\alpha'}^{\beta,\beta'}$ is independent of the choice of $(X_k)_k$ and

$$(\alpha;\alpha')\to(\beta;\beta')\iff s_{\alpha,\alpha'}^{\beta,\beta'}\neq 0\iff(\beta;\beta')\to(\alpha;\alpha').$$

Now, on the left-hand side of the identity (3-4) we can express $\pi_{\xi,r}(X)|_{\mathcal{E}(\alpha;\alpha')}$ in terms of the cocycle using (2-1):

$$\begin{aligned}\mathrm{proj}_{\mathcal{E}'(\beta')}\circ T\circ\pi_{\xi,r}(X)|_{\mathcal{E}(\alpha;\alpha')}&=\sum_{\substack{\beta\\(\alpha;\alpha')\leftrightarrow(\beta;\beta')}}T\circ\mathrm{proj}_{\mathcal{E}'(\beta;\beta')}\circ\pi_{\xi,r}(X)|_{\mathcal{E}(\alpha;\alpha')}\\&=\frac{1}{2}\sum_{\substack{\beta\\(\alpha;\alpha')\leftrightarrow(\beta;\beta')}}(\sigma_\beta-\sigma_\alpha+2r)\cdot(T\circ\omega_{\alpha,\alpha'}^{\beta,\beta'}(X))\\&=\frac{1}{2}\sum_{\substack{\beta\\(\alpha;\alpha')\leftrightarrow(\beta;\beta')}}(\sigma_\beta-\sigma_\alpha+2r)t_{\beta,\beta'}\cdot(R_{\beta,\beta'}\circ\omega_{\alpha,\alpha'}^{\beta,\beta'}(X)).\end{aligned}$$

Similarly we use (2-2) to obtain for the right-hand side

$$\mathrm{proj}_{\mathcal{E}'(\beta')}\circ\tau_{\xi',r'}(X)\circ T|_{\mathcal{E}(\alpha;\alpha')}=\tfrac{1}{2}(\sigma'_{\beta'}-\sigma'_{\alpha'}+2r')t_{\alpha,\alpha'}\cdot(\omega_{\alpha'}^{\beta'}(X)\circ R_{\alpha,\alpha'}).$$

Inserting both expressions into the initial equation (3-4) we obtain:

Theorem 3.4. *Assume (MF1) and (MF2) and fix $R_{\alpha,\alpha'}:\mathcal{E}(\alpha;\alpha')\to\mathcal{E}'(\alpha')$ as in Section 3A. Let $\mathcal{V}\subseteq\mathcal{E}$ be a $(\mathfrak{g}',K')$-submodule of $(\pi_{\xi,r})_{\mathrm{HC}}$. A linear map $T:\mathcal{V}\to\mathcal{E}'$ is an intertwining operator for $\pi_{\xi,r}$ and $\tau_{\xi',r'}$ if and only if*

$$T|_{\mathcal{E}(\alpha;\alpha')}=t_{\alpha,\alpha'}\cdot R_{\alpha,\alpha'}\quad\textit{for all } 0\neq\mathcal{E}(\alpha;\alpha')\subseteq\mathcal{V},$$

and for all $0 \neq \mathcal{E}(\alpha;\alpha') \subseteq \mathcal{V}$ *and* $\mathcal{E}'(\beta') \neq 0$ *we have*

$$\sum_{\substack{\beta \\ (\alpha;\alpha')\leftrightarrow(\beta;\beta')}} (\sigma_\beta - \sigma_\alpha + 2r)t_{\beta,\beta'} \cdot (R_{\beta,\beta'} \circ \omega_{\alpha,\alpha'}^{\beta,\beta'}) = (\sigma'_{\beta'} - \sigma'_{\alpha'} + 2r')t_{\alpha,\alpha'} \cdot (\omega_{\alpha'}^{\beta'} \circ R_{\alpha,\alpha'}). \tag{3-5}$$

Remark 3.5. Through the formulation of Theorem 3.4 for any submodule $\mathcal{V}$ of $(\pi_{\xi,r})_{\mathrm{HC}}$ one can also use (3-5) to describe intertwining operators from subquotients of $(\pi_{\xi,r})_{\mathrm{HC}}$ to $(\tau_{\xi',r'})_{\mathrm{HC}}$. In fact, if $\mathcal{V}' \subseteq \mathcal{V} \subseteq \mathcal{E}$ are $(\mathfrak{g}, K)$-submodules of $(\pi_{\xi,r})_{\mathrm{HC}}$ then any intertwining operator $\mathcal{V}/\mathcal{V}' \to \mathcal{E}'$ for the actions $\pi_{\xi,r}$ and $\tau_{\xi',r'}$ is given by an intertwining operator $\mathcal{V} \to \mathcal{E}'$ which vanishes on $\mathcal{V}'$.

A little more complicated is the description of intertwining operators into subquotients of $(\tau_{\xi',r'})_{\mathrm{HC}}$. Let $\mathcal{W}' \subseteq \mathcal{W} \subseteq \mathcal{E}'$ be $(\mathfrak{g}', K')$-submodules of $(\tau_{\xi',r'})_{\mathrm{HC}}$ and decompose $\mathcal{W} = \mathcal{W}' \oplus \mathcal{W}''$ as K'-modules. Then a close examination of the arguments above shows that any operator $\mathcal{V} \to \mathcal{W}/\mathcal{W}'$ which intertwines the actions of $\pi_{\xi,r}$ and $\tau_{\xi',r'}$ is given by a K'-intertwining linear map $T : \mathcal{V} \to \mathcal{W}''$ with $T|_{\mathcal{E}(\alpha;\alpha')} = t_{\alpha,\alpha'} \cdot R_{\alpha,\alpha'}$ such that the relations (3-5) hold for any $0 \neq \mathcal{E}(\alpha;\alpha') \subseteq \mathcal{V}$ and $0 \neq \mathcal{E}'(\beta') \subseteq \mathcal{W}''$. Note that $t_{\alpha,\alpha'} = 0$ whenever $\mathcal{E}'(\alpha') \nsubseteq \mathcal{W}''$.

3C. *Scalar identities.* To extract from (3-5) information on the constants $t_{\alpha,\alpha'}$ we have to transform it into a scalar identity. For this we assume additionally that

(MF3) $\quad \dim \mathrm{Hom}_{K'}(\mathfrak{s}'_{\mathbb{C}} \otimes \alpha', \beta') \leq 1 \quad$ for all $0 \neq \mathcal{E}(\alpha;\alpha') \subseteq \mathcal{V},\ \mathcal{E}(\beta') \neq 0.$

This implies that the maps

$$\eta_{\alpha,\alpha'}^{\beta,\beta'} = R_{\beta,\beta'} \circ \omega_{\alpha,\alpha'}^{\beta,\beta'} : \mathfrak{s}'_{\mathbb{C}} \otimes \mathcal{E}'(\alpha,\alpha') \to \mathcal{E}'(\beta')$$

are proportional to each other. If further the map

$$\eta_{\alpha,\alpha'}^{\beta'} = \omega_{\alpha'}^{\beta'} \circ R_{\alpha,\alpha'} : \mathfrak{s}'_{\mathbb{C}} \otimes \mathcal{E}'(\alpha;\alpha') \to \mathcal{E}'(\beta')$$

is nonzero then there exist constants $\lambda_{\alpha,\alpha'}^{\beta,\beta'} \neq 0$ such that

$$\eta_{\alpha,\alpha'}^{\beta,\beta'} = \lambda_{\alpha,\alpha'}^{\beta,\beta'} \eta_{\alpha,\alpha'}^{\beta'}.$$

We call $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$ the *proportionality constants*. In this case (3-5) simplifies:

Corollary 3.6. *Under the multiplicity-freeness assumption* (MF3) *the identity* (3-5) *is equivalent to*

$$\sum_{\substack{\beta \\ (\alpha;\alpha')\leftrightarrow(\beta;\beta')}} \lambda_{\alpha,\alpha'}^{\beta,\beta'} (\sigma_\beta - \sigma_\alpha + 2r)t_{\beta,\beta'} = (\sigma'_{\beta'} - \sigma'_{\alpha'} + 2r')t_{\alpha,\alpha'}. \tag{3-6}$$

Whereas the constants σ_α and $\sigma'_{\alpha'}$ are easy to calculate using the highest weights of α and α', see [Branson et al. 1996], we do not have a general method to find the constants $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$. Of course one can always try to compute the action of the cocycle on explicit K-finite vectors and decompose the result, but this turns out to be quite involved already in low-rank cases. However, in some special cases the following information is enough to determine $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$:

Lemma 3.7. *Assume that the elements $H \in \mathfrak{a}$ and $H' \in \mathfrak{a}'$ coincide. Let $\mathcal{E}(\alpha;\alpha') \neq 0$ and $\mathcal{E}'(\beta') \neq 0$ and assume that $R_{\alpha,\alpha'} = R_{\beta,\beta'} = \mathrm{rest}$ for all β with $(\alpha;\alpha') \leftrightarrow (\beta;\beta')$. Then*

$$\sum_{\substack{\beta \\ (\alpha;\alpha')\leftrightarrow(\beta;\beta')}} \lambda_{\alpha,\alpha'}^{\beta,\beta'} = 1, \qquad \sum_{\substack{\beta \\ (\alpha;\alpha')\leftrightarrow(\beta;\beta')}} (\sigma_\beta - \sigma_\alpha)\lambda_{\alpha,\alpha'}^{\beta,\beta'} = \sigma'_{\beta'} - \sigma'_{\alpha'} + 2(\rho - \rho').$$

Here ρ and ρ' are identified with the numbers $\rho(H)$ and $\rho'(H')$.

Proof. For the first identity we note that $H = H'$ implies $\omega(X)|_{K'} = \omega'(X)$ for all $X \in \mathfrak{s}'$. Hence

$$R_{\beta,\beta'} \circ \omega(X) = \omega'(X) \circ R_{\alpha,\alpha'} \quad \text{for all } X \in \mathfrak{g}',$$

which implies

$$\eta_{\alpha,\alpha'}^{\beta'} = \sum_{\substack{\beta \\ (\alpha;\alpha')\leftrightarrow(\beta;\beta')}} \eta_{\alpha,\alpha'}^{\beta,\beta'}$$

and the claimed identity follows. For the second identity note that for $r + \rho = r' + \rho'$ the restriction operator $\mathrm{rest} : \mathcal{E} \to \mathcal{E}'$ is intertwining for $\pi_{\xi,r}$ and $\tau_{\xi',r'}$. Hence the identity (3-6) is satisfied with $t_{\alpha,\alpha'} = 1$ for all $\mathcal{E}(\alpha;\alpha') \neq 0$. Eliminating r and r' gives the desired formula. □

Remark 3.8. The knowledge of any intertwining operator $T : (\pi_{\xi,r})_{\mathrm{HC}} \to (\tau_{\xi',r'})_{\mathrm{HC}}$ and the corresponding numbers $t_{\alpha,\alpha'}$ provides an additional identity for the constants $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$ just as in the proof of Lemma 3.7 for the restriction operator $T = \mathrm{rest}$ with $r + \rho = r' + \rho'$ and $t_{\alpha,\alpha'} = 1$.

3D. ***Automatic continuity.*** In this section we study the question of whether $(\mathfrak{g}', K')$-intertwining operators $(\pi_{\xi,r})_{\mathrm{HC}} \to (\tau_{\xi',r'})_{\mathrm{HC}}$ between Harish-Chandra modules extend to G'-intertwining operators $\pi_{\xi,r} \to \tau_{\xi',r'}$ between the smooth representations, i.e., whether the natural injective map

$$\mathrm{Hom}_{G'}(\pi_{\xi,r}|_{G'}, \tau_{\xi',r'}) \to \mathrm{Hom}_{(\mathfrak{g}',K')}((\pi_{\xi,r})_{\mathrm{HC}}|_{(\mathfrak{g}',K')}, (\tau_{\xi',r'})_{\mathrm{HC}})$$

is an isomorphism. It is expected, see [Kobayashi 2014, Remark 10.2 (4)], that this is true if the space $(G \times G')/\operatorname{diag}(G')$ is real spherical. Statements of this type are also known as "automatic continuity theorems" since they imply continuity with respect to the smooth topologies of every intertwining operator between the algebraic Harish-Chandra modules. We provide a criterion to show automatic continuity in the context of this paper.

Fix a Haar measure dk' on K'. Then the nondegenerate bilinear pairing

$$\mathcal{E}(K';\xi'|_{M'\cap K'})\times\mathcal{E}(K';\xi'^{\vee}|_{M'\cap K'})\to\mathbb{C},\ (f_1,f_2)\mapsto\int_{K'}\langle f_1(k'),f_2(k')\rangle\,dk'$$

is invariant under $\tau_{\xi',r'}\otimes\tau_{\xi'^{\vee},-r'}$ for any $r'\in\mathbb{C}$, where $\xi'^{\vee}$ denotes the contragredient representation of ξ' on the dual space $V_{\xi'}^{\vee}$. Using this pairing we identify $\tau_{\xi',r'}$ with a subrepresentation of the contragredient representation $\tau_{\xi'^{\vee},-r'}^{\vee}$ of $\tau_{\xi'^{\vee},-r'}$, which is realized on the topological dual space $\mathcal{E}(K';\xi'^{\vee}|_{M'\cap K'})^{\vee}$ carrying the weak-$\star$ topology.

Lemma 3.9. *Every continuous G'-intertwining operator $T:\pi_{\xi,r}\to\tau_{\xi'^{\vee},-r'}^{\vee}$ maps into $\tau_{\xi',r'}$ and defines a continuous G'-intertwining operator $T:\pi_{\xi,r}\to\tau_{\xi',r'}$.*

Proof. Let $T:\mathcal{E}(K;\xi|_{M\cap K})\to\mathcal{E}(K';\xi'^{\vee}|_{M'\cap K'})^{\vee}$ be a continuous linear operator which is G'-intertwining for $\pi_{\xi,r}$ and $\tau_{\xi'^{\vee},-r'}^{\vee}$. Then T induces a continuous linear functional

$$\overline{T}:\mathcal{E}(K;\xi|_{M\cap K})\,\widehat{\otimes}\,\mathcal{E}(K';\xi'^{\vee}|_{M'\cap K'})\to\mathbb{C},$$

which is invariant under the diagonal action of K'. The left-hand side is naturally isomorphic to $\mathcal{E}(K\times K';(\xi\otimes\xi'^{\vee})|_{(M\cap K)\times(M'\cap K')})$. Composing with the surjective continuous linear operator

$$\flat:C^{\infty}(K\times K';V_\xi\otimes V_{\xi'}^{\vee})\to\mathcal{E}(K\times K';(\xi\otimes\xi'^{\vee})|_{(M\cap K)\times(M'\cap K')}),$$
$$\flat F(k,k')=\int_{M\cap K}\int_{M'\cap K'}(\xi(m)\otimes\xi'(m')^{\vee})F(km,k'm')\,dm'\,dm,$$

we obtain a functional

$$K_T:=\overline{T}\circ\flat:C^{\infty}(K\times K';V_\xi\otimes V_{\xi'}^{\vee})\to\mathbb{C},$$

i.e., a distribution on $K\times K'$ with values in $V_\xi\otimes V_{\xi'}^{\vee}$. (This is basically the Schwartz kernel of the operator T, avoiding distribution sections of vector bundles.) The distribution K_T is invariant under the diagonal action of K' from the left and equivariant under the action of $(M\cap K)\times(M'\cap K')$ from the right. We define a distribution $\widetilde{K}_T$ on K with values in $V_\xi\otimes V_{\xi'}^{\vee}$, i.e., a continuous linear functional on $C^{\infty}(K;V_\xi\otimes V_{\xi'}^{\vee})$, by

$$\langle\widetilde{K}_T,\phi\rangle:=\langle K_T(x,x'),\phi(x'^{-1}x)\rangle.$$

Then for $\phi \in \mathcal{E}(K;\xi|_{M\cap K})$ and $\psi \in \mathcal{E}(K';\xi'^{\vee}|_{M'\cap K'})$ we have

$$\begin{aligned}\langle T\phi,\psi\rangle = \langle K_T,\phi\otimes\psi\rangle &= \int_{K'} \langle K_T(x,x'),\phi(k'x)\otimes\psi(k'x')\rangle\,dk' \\ &= \left\langle K_T(x,x'), \int_{K'} \phi(k'x)\otimes\psi(k'x')\,dk' \right\rangle \\ &= \left\langle K_T(x,x'), \int_{K'} \phi(k'x'^{-1}x)\otimes\psi(k')\,dk' \right\rangle \\ &= \int_{K'} \langle K_T(x,x'),\phi(k'x'^{-1}x)\otimes\psi(k')\rangle\,dk' = \int_{K'} \langle \widetilde{K}_T,\phi(k'\cdot)\otimes\psi(k')\rangle\,dk'.\end{aligned}$$

This implies that for any $\lambda \in V_{\xi'}^{\vee}$

$$\langle \lambda, T\phi(k')\rangle = \langle \widetilde{K}_T, \phi(k'\cdot)\otimes\lambda\rangle,$$

which shows that $T\phi \in C^\infty(K';V_{\xi'})$. That $T\phi \in \mathcal{E}(K';\xi'|_{M'\cap K'})$ easily follows from the equivariance property of $\widetilde{K}_T$ with respect to $M\cap K$ and $M'\cap K'$. Finally, continuity of the thus defined operator $T:\pi_{\xi,r}\to\tau_{\xi',r'}$ follows from the continuity of the functional $\widetilde{K}_T$ on $C^\infty(K;V_\xi\otimes V_{\xi'}^{\vee})$ and the proof is complete. □

Fix invariant inner products on the representation $\xi|_{M\cap K}$, resp. $\xi'|_{M'\cap K'}$, and let $\|\cdot\|$, resp. $\|\cdot\|'$, denote the corresponding L^2-norm on $L^2(K\times_{M\cap K}\xi)$, resp. $L^2(K'\times_{M'\cap K'}\xi')$. These norms induce norms on each K'-type $\mathcal{E}(\alpha;\alpha')$ resp. $\mathcal{E}'(\alpha')$. Write $\|R_{\alpha,\alpha'}\|_{L^2\to L^2}$ for the operator norm of $R_{\alpha,\alpha'}:\mathcal{E}(\alpha;\alpha')\to\mathcal{E}'(\alpha')$ with respect to the L^2-norms.

For any $F\in L^2(K\times_{M\cap K}\xi)$ write

$$F = \sum_{\alpha\in\widehat{K}} F_\alpha,$$

with $F_\alpha\in\mathcal{E}(\alpha)$. Then the sequence $\{\|F_\alpha\|\}_\alpha$ belongs to $\ell^2(\widehat{K})$, the space of square-summable sequences. We identify the set $\widehat{K}$, resp. $\widehat{K}'$, with the corresponding weight lattice so that it becomes a subset of a finite-dimensional vector space. Denote by $|\cdot|$, resp. $|\cdot|'$, a norm on this finite-dimensional vector space. It is known that $F\in\mathcal{E}(K;\xi_{M\cap K})$ if and only if the sequence $\{\|F_\alpha\|\}_\alpha$ belongs to $s(\widehat{K})$, the space of rapidly decreasing sequences, i.e., those that are still bounded if multiplied with any power $|\alpha|^N$. Moreover, $\mathcal{E}(K;\xi^\vee|_{M\cap K})^\vee$ is identified with all formal expansions $F=\sum_\alpha F_\alpha$, where $\{\|F_\alpha\|\}_\alpha$ belongs to $s'(\widehat{K})$, the space of tempered sequences, i.e., those that grow at most at the rate of $|\alpha|^N$ for some $N\in\mathbb{N}$.

Proposition 3.10. *A $(\mathfrak{g}',K')$-intertwining operator $T:(\pi_{\xi,r})_{\mathrm{HC}}\to(\tau_{\xi',r'})_{\mathrm{HC}}$ with $T|_{\mathcal{E}(\alpha;\alpha')}=t_{\alpha,\alpha'}\cdot R_{\alpha,\alpha'}$ extends to a continuous G'-intertwining operator $\pi_{\xi,r}\to\tau_{\xi',r'}$ if both $t_{\alpha,\alpha'}$ and $\|R_{\alpha,\alpha'}\|_{L^2\to L^2}$ are of at most polynomial growth in α and α'.*

Proof. By Lemma 3.9 it suffices to show that T extends to a continuous G'-intertwining operator $\pi_{\xi,r} \to (\tau_{\xi'^\vee,-r'})'$. Let $F \in \pi_{\xi,r}$; then $F = \sum_\alpha F_\alpha$ with $\{\|F_\alpha\|\}_\alpha$ a sequence in $s(\widehat{K})$. We have $TF = \sum_{\alpha'}(TF)_{\alpha'}$ with

$$(TF)_{\alpha'} = \sum_\alpha t_{\alpha,\alpha'} \cdot R_{\alpha,\alpha'} F_\alpha.$$

By the assumptions

$$|t_{\alpha,\alpha'}| \le C_1(1+|\alpha|+|\alpha'|)^{N_1},$$
$$\|R_{\alpha,\alpha'}\|_{L^2\to L^2} \le C_2(1+|\alpha|+|\alpha'|)^{N_2}$$

for some $C_1, C_2 > 0$ and $N_1, N_2 \in \mathbb{N}$. Further, since $\|F_\alpha\| \in s(\widehat{K})$, for every $N \in \mathbb{N}$ there exists $C > 0$ such that $\|F_\alpha\| \le C(1+|\alpha|)^{-N}$. Hence, we have for any α'

$$\|(TF)_{\alpha'}\|' \le CC_1C_2 \sum_\alpha (1+|\alpha|+|\alpha'|)^{N_1+N_2}(1+|\alpha|)^{-N}.$$

Choosing N large enough, this is uniformly bounded by a constant times $(1+|\alpha'|)^{N_1+N_2}$, and hence $\|(TF)_{\alpha'}\|' \in s'(\widehat{K}')$ so that $TF \in \mathcal{E}(K';\xi'^\vee|_{M'\cap K'})^\vee$. This shows that T extends to a G'-intertwining operator $\pi_{\xi,r} \to \tau^\vee_{\xi'^\vee,-r'}$. Continuity of this operator also follows by the above estimates. □

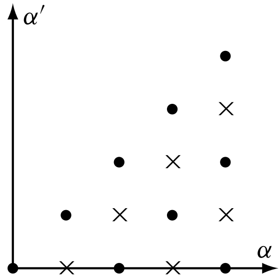
4. Rank-one orthogonal groups

In this section we apply our method to classify symmetry-breaking operators for rank-one orthogonal groups. Let $n \ge 3$ and consider the indefinite orthogonal group $G = \mathrm{O}(1,n)$ of $(n+1)\times(n+1)$ real matrices leaving the standard bilinear form on $\mathbb{R}^{n+1}$ of signature $(1,n)$ invariant. The subgroup $G' \subseteq G$ of matrices fixing the last standard basis vector e_{n+1} is isomorphic to $\mathrm{O}(1,n-1)$.

4A. ***K-types.*** We fix $K = \mathrm{O}(1)\times\mathrm{O}(n)$ and choose

$$H = \begin{pmatrix} 0 & 1 & \\ 1 & 0 & \\ & & \mathbf{0}_{n-1} \end{pmatrix}$$

so that $P = MAN$, with $M = \Delta\,\mathrm{O}(1)\times\mathrm{O}(n-1)$, where $\Delta\,\mathrm{O}(1) = \{\mathrm{diag}(x,x) : x \in \mathrm{O}(1)\}$. Note that $\rho = \frac{n-1}{2}$. Then K acts transitively on S^{n-1} via $\mathrm{diag}(\varepsilon,k)\cdot x = \varepsilon kx$, $\varepsilon \in \mathrm{O}(1)$, $k \in \mathrm{O}(n)$, $x \in S^{n-1}$, and M is the stabilizer subgroup of the first standard basis vector $e_1 \in S^{n-1}$, whence $K/M \cong S^{n-1}$. The subgroup $G' = \mathrm{O}(1,n-1)$ is embedded into G such that $K' = \mathrm{O}(1)\times\mathrm{O}(n-1)$ and $P' = G'\cap P = M'A'N'$, with $A' = A$ and $M' = \Delta\,\mathrm{O}(1)\times\mathrm{O}(n-2)$. Then $K'/M' = S^{n-2}$, viewed as the equator in $K/M = S^{n-1} \subseteq \mathbb{R}^n$ given by $x_n = 0$. Further we have $\nu = \nu'$ and $\rho' = \frac{n-2}{2}$.



Legend: $\bullet$ K'-types $\mathcal{E}(\alpha;\alpha')$ with $\alpha-\alpha'\in 2\mathbb{Z}$
$\times$ K'-types $\mathcal{E}(\alpha;\alpha')$ with $\alpha-\alpha'\in 2\mathbb{Z}+1$

Figure 1

Let $\xi = \mathbf{1}$, $\xi' = \mathbf{1}$ be the trivial representations of M and M' and abbreviate $\pi_r = \pi_{\xi,r}$ and $\tau_{r'} = \tau_{\xi',r'}$. As K-modules, resp. K'-modules, we have

$$\mathcal{E} = \bigoplus_{\alpha=0}^{\infty} \underbrace{\mathrm{sgn}^{\alpha} \boxtimes \mathcal{H}^{\alpha}(\mathbb{R}^n)}_{\mathcal{E}(\alpha)}, \qquad \mathcal{E}' = \bigoplus_{\alpha'=0}^{\infty} \underbrace{\mathrm{sgn}^{\alpha'} \boxtimes \mathcal{H}^{\alpha'}(\mathbb{R}^{n-1})}_{\mathcal{E}'(\alpha')},$$

so that (MF1) is satisfied. Further, each K-type decomposes by (B-1) into K'-types as

$$(\mathrm{sgn}^{\alpha} \boxtimes \mathcal{H}^{\alpha}(\mathbb{R}^n))|_{K'} \simeq \bigoplus_{0\le\alpha'\le\alpha} (\mathrm{sgn}^{\alpha} \boxtimes \mathcal{H}^{\alpha'}(\mathbb{R}^{n-1})),$$

and hence (MF2) holds. Comparing the sign representations of the $\mathrm{O}(1)$-factor of K' we find that $\mathrm{Hom}_{K'}(\mathcal{E}(\alpha)|_{K'}, \mathcal{E}'(\alpha')) \neq 0$ if and only if $\alpha-\alpha' \in 2\mathbb{Z}$. In this case formulas (B-2) and (A-2) show that the restriction operator

$$R_{\alpha,\alpha'} = \mathrm{rest}\,|_{\mathcal{E}(\alpha;\alpha')} : \mathcal{E}(\alpha;\alpha') \to \mathcal{E}'(\alpha')$$

is an isomorphism. Hence the restriction $T_{\alpha,\alpha'} = T|_{\mathcal{E}(\alpha;\alpha')}$ of a K'-intertwining operator $T : \mathcal{E} \to \mathcal{E}'$ is given by $T_{\alpha,\alpha'} = t_{\alpha,\alpha'} R_{\alpha,\alpha'}$ for $\alpha-\alpha' \in 2\mathbb{N}$ and $T_{\alpha,\alpha'} = 0$ else. The K- and K'-types are illustrated in Figure 1.

4B. ***Proportionality constants.*** The eigenvalues of the spectrum-generating operator on the K-types are simply the eigenvalues of the Laplacian on S^{n-1} and given by, see [Branson et al. 1996, Section 3.a],

$$\sigma_{\alpha} = \alpha(\alpha+n-2), \qquad \sigma'_{\alpha'} = \alpha'(\alpha'+n-3).$$

We identify $\mathfrak{s} \cong \mathbb{R}^n$ via

$$\mathbb{R}^n \to \mathfrak{s}, \qquad y \mapsto X_y = \begin{pmatrix} 0 & y^t \\ y & \mathbf{0}_n \end{pmatrix}.$$

Then $\mathfrak{s}' \cong \mathbb{R}^{n-1}$, embedded in $\mathbb{R}^n$ as the first $n-1$ coordinates. Since $\mathfrak{s}'_{\mathbb{C}} \simeq \mathbb{C}^{n-1}$ is a weight multiplicity-free K'-module, (MF3) holds and we can use Corollary 3.6. To compute the proportionality constants $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$ we use Lemma 3.7, which applies to this situation because $H = H'$ and $R_{\alpha,\alpha'} = \text{rest}$. The cocycle ω is given by

$$\omega(X_y)(x) = y^t x, \quad x \in S^{n-1},\ y \in \mathbb{R}^n.$$

Using (B-4) it is easy to see that for fixed $0 \le \alpha' \le \alpha$

$$(\alpha;\alpha') \leftrightarrow (\beta;\beta') \quad \Longleftrightarrow \quad |\alpha-\beta| = |\alpha'-\beta'| = 1.$$

By Lemma 3.7 we have the following equations for $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$: for $\beta' = \alpha'+1$ we obtain

$$\begin{aligned} \lambda_{\alpha,\alpha'}^{\alpha+1,\alpha'+1} + \lambda_{\alpha,\alpha'}^{\alpha-1,\alpha'+1} &= 1,\\ (2\alpha+n-1)\lambda_{\alpha,\alpha'}^{\alpha+1,\alpha'+1} - (2\alpha+n-3)\lambda_{\alpha,\alpha'}^{\alpha-1,\alpha'+1} &= 2\alpha'+n-1, \end{aligned}$$

which gives

$$\lambda_{\alpha,\alpha'}^{\alpha+1,\alpha'+1} = \frac{\alpha+\alpha'+n-2}{2\alpha+n-2}, \quad \lambda_{\alpha,\alpha'}^{\alpha-1,\alpha'+1} = \frac{\alpha-\alpha'}{2\alpha+n-2},$$

and for $\beta' = \alpha'-1$ we get

$$\begin{aligned} \lambda_{\alpha,\alpha'}^{\alpha+1,\alpha'-1} + \lambda_{\alpha,\alpha'}^{\alpha-1,\alpha'-1} &= 1,\\ (2\alpha+n-1)\lambda_{\alpha,\alpha'}^{\alpha+1,\alpha'-1} - (2\alpha+n-3)\lambda_{\alpha,\alpha'}^{\alpha-1,\alpha'-1} &= -2\alpha'-n+5, \end{aligned}$$

implying

$$\lambda_{\alpha,\alpha'}^{\alpha+1,\alpha'-1} = \frac{\alpha-\alpha'+1}{2\alpha+n-2}, \quad \lambda_{\alpha,\alpha'}^{\alpha-1,\alpha'-1} = \frac{\alpha+\alpha'+n-3}{2\alpha+n-2}.$$

We remark that the constants $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$ can in this case also be obtained by computing the action of $\omega(X)$ on explicit K-finite vectors using (B-2) and recurrence relations for the Gegenbauer polynomials. With the explicit form of the constants $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$ Corollary 3.6 now provides the following characterization of symmetry-breaking operators:

Theorem 4.1. *An operator $T : \mathcal{E} \to \mathcal{E}'$ is intertwining for π_r and $\tau_{r'}$ if and only if*

$$T|_{\mathcal{E}(\alpha;\alpha')} = \begin{cases} t_{\alpha,\alpha'} \cdot \text{rest}\,|_{\mathcal{E}(\alpha;\alpha')} & \text{for } \alpha-\alpha' \in 2\mathbb{Z},\\ 0 & \text{for } \alpha-\alpha' \in 2\mathbb{Z}+1, \end{cases}$$

with numbers $t_{\alpha,\alpha'}$ satisfying

$$(2\alpha+n-2)(2r'+2\alpha'+n-2)t_{\alpha,\alpha'} = (\alpha+\alpha'+n-2)(2r+2\alpha+n-1)t_{\alpha+1,\alpha'+1} + (\alpha-\alpha')(2r-2\alpha-n+3)t_{\alpha-1,\alpha'+1} \tag{4-1}$$

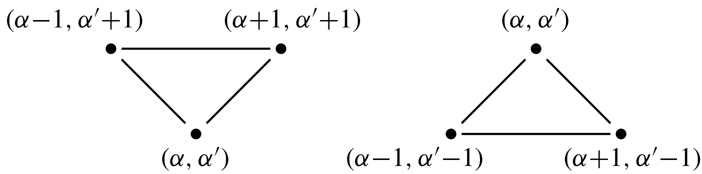


Figure 2. The relations (4-1) and (4-2).

and

$$(2\alpha+n-2)(2r'-2\alpha'-n+4)t_{\alpha,\alpha'} = (\alpha-\alpha'+1)(2r+2\alpha+n-1)t_{\alpha+1,\alpha'-1} + (\alpha+\alpha'+n-3)(2r-2\alpha-n+3)t_{\alpha-1,\alpha'-1}. \tag{4-2}$$

We view these two relations as triangles connecting three vertices in the K-type picture (see Figure 2).

Note that if $r \notin -\rho-\mathbb{N}$ then $2r+2\alpha+n-1 \neq 0$ for all α and hence one can define $t_{\alpha+1,\alpha'+1}$ in terms of $t_{\alpha,\alpha'}$ and $t_{\alpha-1,\alpha'+1}$ using (4-1) and do similarly for $t_{\alpha+1,\alpha'-1}$ using (4-2). If $r = -\rho-i \in -\rho-\mathbb{N}$ and $\alpha = i$ the coefficient $(2r+2\alpha+n-1)$ vanishes and (4-1) and (4-2) reduce to identities involving only two terms. We indicate this by drawing a vertical line between i and $i+1$ indicating that one cannot "step" from the left-hand side to the right-hand side (see Figure 3). Similarly we have that if $r' \notin -\rho'-\mathbb{N}$ then $2r'+2\alpha'+n-2 \neq 0$ for all α' and we can define $t_{\alpha,\alpha'}$ in terms of $t_{\alpha\pm1,\alpha'+1}$ using (4-1). If $r' = -\rho'-j \in -\rho'-\mathbb{N}$ and $\alpha' = j$ we obtain a horizontal line between j and $j+1$ as barrier, indicating that we cannot step from the part above this line to the part below. Note that if there is a vertical, resp. horizontal, barrier like this the coefficient $(2r-2\alpha-n+3)$, resp. $(2r'-2\alpha'-n+4)$, never vanishes and one can step in the other direction, namely from right to left, resp. from the part below the line to the part above.

4C. *Multiplicities.* The $(\mathfrak{g}, K)$-module $(\pi_r)_{\mathrm{HC}}$ is reducible if and only if $r \in \pm(\rho+\mathbb{N})$. More precisely, for $r = -\rho-i$ the module $(\pi_r)_{\mathrm{HC}}$ contains a unique nontrivial finite-dimensional $(\mathfrak{g}, K)$-submodule $\mathcal{F}(i) \subseteq \mathcal{E}$ with K-types $\mathcal{E}(\alpha)$, $0 \leq \alpha \leq i$. Its quotient $\mathcal{T}(i) = \mathcal{E}/\mathcal{F}(i)$ is irreducible and can be identified with the unique nontrivial $(\mathfrak{g}, K)$-submodule of $(\pi_{-r})_{\mathrm{HC}}$. Similarly we denote for $r' = -\rho'-j$ by

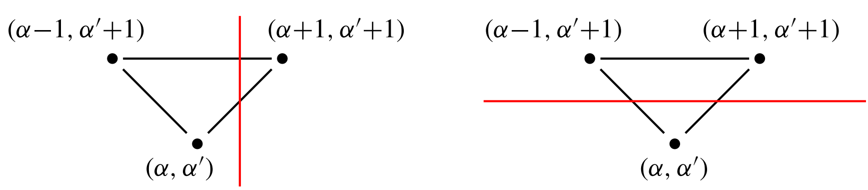


Figure 3. Barriers for $r = -\rho-i$ and $r' = -\rho'-j$.

$\mathcal{F}'(j)$ the unique finite-dimensional $(\mathfrak{g}', K')$-submodule of $(\tau_{r'})_{\mathrm{HC}}$ and by $\mathcal{T}'(j)$ its irreducible quotient. Let

$$L_{\mathrm{even}} = \{(r, r') : r = -\rho - i,\ r' = -\rho' - j,\ i - j \in 2\mathbb{N}\},$$
$$L_{\mathrm{odd}} = \{(r, r') : r = -\rho - i,\ r' = -\rho' - j,\ i - j \in 2\mathbb{N} + 1\}.$$

This notation agrees with the notation used in [Kobayashi and Speh 2015].

Theorem 4.2. (1) *The multiplicities between spherical principal series of G and G' are given by*

$$m((\pi_r)_{\mathrm{HC}}, (\tau_{r'})_{\mathrm{HC}}) = \begin{cases} 1 & \text{for } (r, r') \in \mathbb{C}^2 \setminus L_{\mathrm{even}}, \\ 2 & \text{for } (r, r') \in L_{\mathrm{even}}. \end{cases}$$

(2) *For $i, j \in \mathbb{N}$ the multiplicities $m(\mathcal{V}, \mathcal{W})$ between subquotients are given by*

$\mathcal{V}\downarrow$ $\mathcal{W}\rightarrow$	$\mathcal{F}'(j)$	$\mathcal{T}'(j)$
$\mathcal{F}(i)$	1	0
$\mathcal{T}(i)$	0	1

for $i - j \in 2\mathbb{N}$,

$\mathcal{V}\downarrow$ $\mathcal{W}\rightarrow$	$\mathcal{F}'(j)$	$\mathcal{T}'(j)$
$\mathcal{F}(i)$	0	0
$\mathcal{T}(i)$	1	0

otherwise.

To prove Theorem 4.2 we study how the relations (4-1) and (4-2) determine the numbers $t_{\alpha,\alpha'}$. We first consider the diagonal $\alpha = \alpha'$. Relation (4-1) then simplifies to

$$(2r' + 2\alpha + n - 2)t_{\alpha,\alpha} = (2r + 2\alpha + n - 1)t_{\alpha+1,\alpha+1}. \tag{4-3}$$

This immediately yields:

Lemma 4.3. (1) *For $(r, r') \in \mathbb{C}^2 \setminus (L_{\mathrm{even}} \cup L_{\mathrm{odd}})$ the space of diagonal sequences $(t_{\alpha,\alpha})_\alpha$ satisfying (4-3) has dimension* 1. *Any generator $(t_{\alpha,\alpha})_\alpha$ satisfies*:

(a) *For $r \notin -\rho - \mathbb{N}$, $r' \notin -\rho' - \mathbb{N}$,*

$$t_{\alpha,\alpha} \neq 0 \quad \textit{for all } \alpha \in \mathbb{N}.$$

(b) *For $r = -\rho - i \in -\rho - \mathbb{N}$, $r' \notin -\rho' - \mathbb{N}$,*

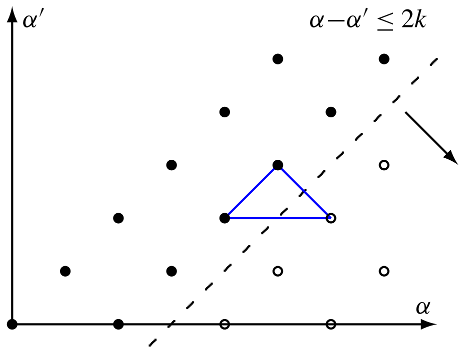
$$t_{\alpha,\alpha} = 0 \quad \textit{for all } \alpha \leq i \qquad \textit{and} \qquad t_{\alpha,\alpha} \neq 0 \quad \textit{for all } \alpha > i.$$

(c) *For $r \notin -\rho - \mathbb{N}$, $r' = -\rho' - j \in -\rho' - \mathbb{N}$,*

$$t_{\alpha,\alpha} \neq 0 \quad \textit{for all } \alpha \leq j \qquad \textit{and} \qquad t_{\alpha,\alpha} = 0 \quad \textit{for all } \alpha > j.$$

(d) *For $r = -\rho - i \in -\rho - \mathbb{N}$, $r' = -\rho' - j \in -\rho' - \mathbb{N}$, with $i < j$,*

$$t_{\alpha,\alpha} \neq 0 \quad \textit{for all } i < \alpha \leq j \qquad \textit{and} \qquad t_{\alpha,\alpha} = 0 \quad \textit{else}.$$



Legend: • K'-types $\mathcal{E}(\alpha;\alpha')$ with $\alpha-\alpha' \leq 2k$ ($t_{\alpha,\alpha'}$ already defined)
∘ K'-types $\mathcal{E}(\alpha;\alpha')$ with $\alpha-\alpha' > 2k$ ($t_{\alpha,\alpha'}$ yet to define)

Figure 4

(2) *For* $(r,r') = (-\rho - i, -\rho' - j) \in (L_{\text{even}} \cup L_{\text{odd}})$, *the space of diagonal sequences* $(t_{\alpha,\alpha})_\alpha$ *satisfying* (4-3) *has dimension* 2. *It has a basis* $(t'_{\alpha,\alpha})_\alpha$, $(t''_{\alpha,\alpha})_\alpha$ *with the properties*

$$t'_{\alpha,\alpha} \neq 0 \quad \text{for all } \alpha \leq j, \qquad t'_{\alpha,\alpha} = 0 \quad \text{for all } \alpha > j,$$
$$t''_{\alpha,\alpha} = 0 \quad \text{for all } \alpha \leq i, \qquad t''_{\alpha,\alpha} \neq 0 \quad \text{for all } \alpha > i.$$

Next we investigate how a diagonal sequence $(t_{\alpha,\alpha})_\alpha$ satisfying (4-3) can be extended to a sequence $(t_{\alpha,\alpha'})_{(\alpha,\alpha')}$ satisfying (4-1) and (4-2).

Lemma 4.4. *Let* $(r,r') \in \mathbb{C}^2 \setminus (L_{\text{even}} \cup L_{\text{odd}})$. *Then every diagonal sequence* $(t_{\alpha,\alpha})_\alpha$ *satisfying* (4-3) *has a unique extension to a sequence* $(t_{\alpha,\alpha'})_{(\alpha,\alpha')}$ *satisfying* (4-1) *and* (4-2).

Proof. Step 1. We first treat the case $r \notin -\rho - \mathbb{N}$. In this case the coefficients $(2r + 2\alpha + n - 1)$ in (4-1) and (4-2) never vanish. We now extend the diagonal sequence $(t_{\alpha,\alpha})_\alpha$ inductively to a sequence $(t_{\alpha,\alpha'})_{\alpha-\alpha' \leq 2k}$ with $k \in \mathbb{N}$ which satisfies (4-1) for (α,α') with $\alpha - \alpha' \leq 2k$ and (4-2) for (α,α') with $\alpha - \alpha' \leq 2k-2$ as visualized in Figure 4 (i.e., the two relations hold whenever the corresponding triangles in Figure 2 are contained in the region $\alpha - \alpha' \leq 2k$). For $k = 0$ the diagonal sequence we start with satisfies these assumptions. For the induction step $k \to k+1$ let $\alpha - \alpha' = 2k$ and define $t_{\alpha+1,\alpha'-1}$ and $t_{\alpha+2,\alpha'}$ using (4-2) (the blue triangles in Figure 5) in terms of $t_{\alpha-1,\alpha'-1}$, $t_{\alpha,\alpha'}$ and $t_{\alpha+1,\alpha'+1}$. This is possible, because $2r + 2\alpha + n - 1 \neq 0$ for all α and hence the corresponding coefficients in (4-2) are nonzero. Now we have to prove that (4-1) holds for $(\alpha+1, \alpha'-1)$ (the red triangle). This can be done by an elementary calculation using the blue triangles that are by definition valid as well as the green triangles that are valid by the induction assumption. Hence this extends the diagonal sequence $(t_{\alpha,\alpha})_\alpha$ to a

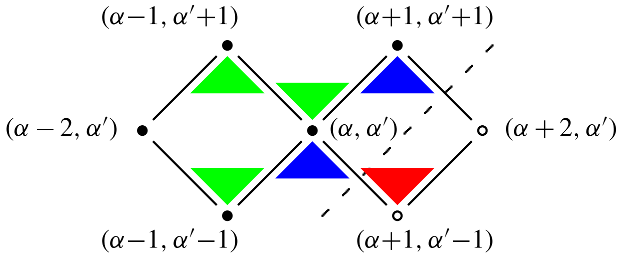
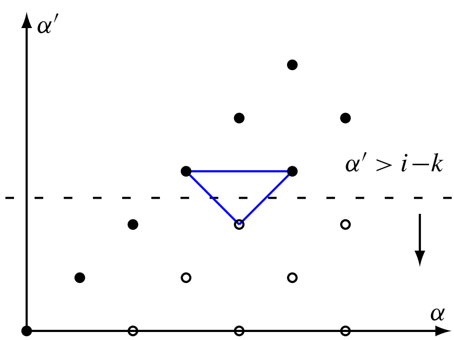


Figure 5

sequence $(t_{\alpha,\alpha'})_{0\le\alpha'\le\alpha}$ satisfying (4-1) and (4-2). Since the relations were used to extend the diagonal sequence this extension is unique.

Step 2. Next assume $r=-\rho-i\in-\rho-\mathbb{N}$ and $r'\notin-\rho'-\mathbb{N}$. Then the coefficient $(2r+2\alpha+n-1)$ vanishes if and only if $\alpha=i$. We can therefore use the technique in Step 1 to extend the upper part $(t_{\alpha,\alpha})_{\alpha>i}$ of the diagonal sequence to a sequence $(t_{\alpha,\alpha'})_{i<\alpha'\le\alpha}$ in the region $\alpha'>i$. Next we extend the sequence $(t_{\alpha,\alpha})_{\alpha'>i}$ inductively to a sequence $(t_{\alpha,\alpha'})_{\alpha'>i-k}$ with $k=0,\dots,i+1$ which satisfies (4-1) for (α,α') with $\alpha'>i-k$ and (4-2) for (α,α') with $\alpha'>i-k+1$ as visualized in Figure 6 (i.e., the two relations hold whenever the corresponding triangles in Figure 2 are contained in the region $\alpha'>i-k$). For $k=0$ the sequence we obtained using Step 1 satisfies these assumptions by Step 1. For the induction step $k\to k+1$ let $\alpha'=i-k+1$ and define $t_{\alpha-1,\alpha'-1}$ and $t_{\alpha+1,\alpha'-1}$ using (4-1) (the blue triangles in Figure 7) in terms of $t_{\alpha-2,\alpha'}$, $t_{\alpha,\alpha'}$ and $t_{\alpha+2,\alpha'}$. This is possible, because $r'\notin-\rho'-\mathbb{N}$ and hence the corresponding coefficient $(2r'+2\alpha'+n-2)$ in (4-1) never vanishes. Now we have to prove that (4-2) holds for (α,α') (the red triangle) which is done in a similar fashion as in Step 1 using the green triangle. This finishes Step 2.



Legend: $\bullet$ K'-types $\mathcal{E}(\alpha;\alpha')$ with $\alpha'>i-k$ ($t_{\alpha,\alpha'}$ already defined)
$\circ$ K'-types $\mathcal{E}(\alpha;\alpha')$ with $\alpha'\le i-k$ ($t_{\alpha,\alpha'}$ yet to define)

Figure 6

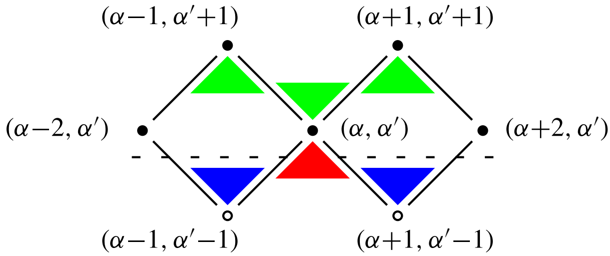


Figure 7

Step 3. Now let $r = -\rho - i \in -\rho - \mathbb{N}$ and $r' = -\rho' - j \in -\rho' - \mathbb{N}$ with $i, j \in \mathbb{N}$, $j > i$. Note that to carry out Step 2 we only need that $(2r' + 2\alpha' + n - 2) \neq 0$ for $\alpha' \leq i$. This is satisfied since

$$2r' + 2\alpha' + n - 2 = 2(\alpha' - j) < 2(\alpha' - i) \leq 0$$

by assumption. Hence the technique in Step 2 carries over to this case. □

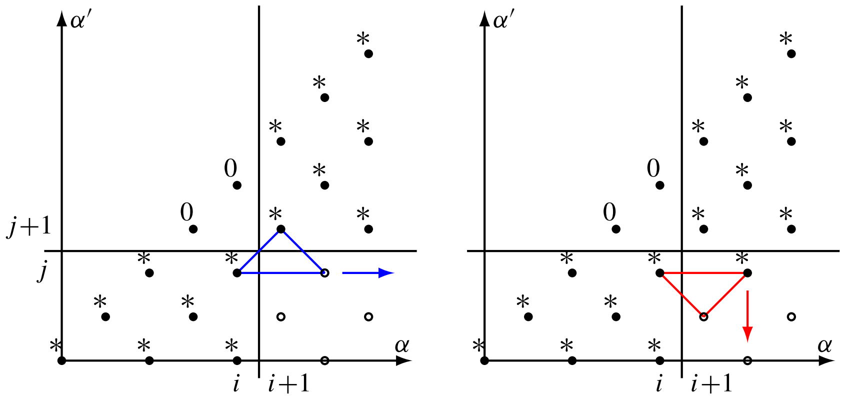
Lemma 4.5. *Let* $(r, r') = (-\rho - i, -\rho' - j)$, $i, j \in \mathbb{N}$:

(1) *For* $(r, r') \in L_{\text{even}}$ *every diagonal sequence* $(t_{\alpha,\alpha})_\alpha$ *satisfying* (4-3) *has a unique extension to a sequence* $(t_{\alpha,\alpha'})_{(\alpha,\alpha')}$ *satisfying* (4-1) *and* (4-2).

(2) *For* $(r, r') \in L_{\text{odd}}$ *any sequence* $(t_{\alpha,\alpha'})_{\alpha,\alpha'}$ *satisfying* (4-1) *and* (4-2) *has the property* $t_{\alpha,\alpha'} = 0$ *for* $\alpha \leq i$ *or* $\alpha' > j$. *Conversely, for any choice of* $t_{i+1,j} \in \mathbb{C}$ *there exists a unique extension to a sequence* $(t_{\alpha,\alpha'})_{(\alpha,\alpha')}$ *satisfying* (4-1) *and* (4-2).

Proof. (1) First Steps 1 and 2 in the proof of Lemma 4.4 extend a diagonal sequence $(t_{\alpha,\alpha})_\alpha$ uniquely to the range $\{(\alpha, \alpha') : \alpha \leq i \text{ or } \alpha' > j\}$. This extension satisfies $t_{\alpha,\alpha'} = 0$ whenever $j < \alpha' \leq \alpha \leq i$. Next one can use (4-2) for $(\alpha, \alpha') = (i+1, j+1)$ to define $t_{i+2,j}$ in terms of $t_{i,j}$ and $t_{i+1,j+1}$ (the blue triangle in Figure 8). Inductively, using (4-2) for $(\alpha, \alpha') = (i + 2k + 1, j + 1)$, $k = 0, 1, 2, \ldots$, the values of $t_{i+2k+2,j}$ are determined for all k. In the next step the technique from Step 2 in the proof of Lemma 4.4 is used to inductively define $t_{\alpha,\alpha'}$ for $\alpha > i$ and $\alpha' = j - k$, $k = 0, \ldots, j$ (the red triangle). That all relations (4-1) and (4-2) are satisfied within the four quadrants in Figure 8 is clear from the arguments in Steps 1 and 2 in the proof of Lemma 4.4. That these relations are also satisfied at the edges between the quadrants holds either by definition or since all terms vanish.

(2) Let $(t_{\alpha,\alpha'})_{\alpha,\alpha'}$ be a sequence satisfying (4-1) and (4-2). Note that Lemma 4.3(2) already implies $t_{\alpha,\alpha} = 0$ for $j < \alpha \leq i$. Then by Step 1 in the proof of Lemma 4.4 we have $t_{\alpha,\alpha'} = 0$ whenever $j < \alpha' \leq \alpha \leq i$ (the black zeroes in Figure 9). We first show inductively that $t_{i-2k-1,j} = 0$ for $k = 0, \ldots, \frac{i-j-1}{2}$ (the red zeroes). To show the statement for $k = 0$ consider the relation (4-2) for $(\alpha, \alpha') = (i, j+1)$. By the previous considerations $t_{\alpha,\alpha'} = 0$ and further the coefficient $(2r + 2\alpha + n - 1)$ of $t_{\alpha+1,\alpha'-1}$ vanishes. Hence $t_{\alpha-1,\alpha'-1} = t_{i-1,j} = 0$. For the induction step assume $t_{i-2k-1,j} = 0$ and



Legend: • K'-types $\mathcal{E}(\alpha;\alpha')$ with $t_{\alpha,\alpha'}$ already defined
∘ K'-types $\mathcal{E}(\alpha;\alpha')$ with $t_{\alpha,\alpha'}$ yet to define

Figure 8

consider the relation (4-2) for $(\alpha,\alpha')=(i-2k-2,\,j+1)$. Then $t_{\alpha,\alpha'}=t_{\alpha+1,\alpha'-1}=0$ and therefore $t_{\alpha-1,\alpha'-1}=t_{i-2(k+1)-1,j}=0$. Thus we have showed $t_{j,j}=0$. But in view of (4-3) this yields $t_{\alpha,\alpha}=0$ for $\alpha\le j$. In a similar way one uses (4-1) and (4-2) for $(\alpha,\alpha')=(i+1,\,j+2k)$, $k=0,\dots,\frac{i-j+1}{2}$, to show that $t_{i+1,i+1}=0$ and hence $t_{\alpha,\alpha}=0$ for all $\alpha>i$. From the vanishing of the diagonal the techniques in Steps 1 and 2 in the proof of Lemma 4.4 yield $t_{\alpha,\alpha'}=0$ whenever $\alpha\le i$ or $\alpha'>j$.

Now let $t_{i+1,j}\in\mathbb{C}$ be given and put $t_{\alpha,\alpha'}=0$ whenever $\alpha\le i$ or $\alpha'>j$. Then (4-1) and (4-2) are trivially satisfied whenever all three terms are defined. Further, using Steps 1 and 2 it is again easy to see that this sequence has a unique extension $(t_{\alpha,\alpha'})_{\alpha,\alpha'}$ with the required properties. □

Proof of Theorem 4.2. (1) Let first $(r,r')\in\mathbb{C}^2\setminus(L_{\mathrm{even}}\cup L_{\mathrm{odd}})$. Then by Lemma 4.3 the space of diagonal sequences satisfying (4-3) is one-dimensional and each such

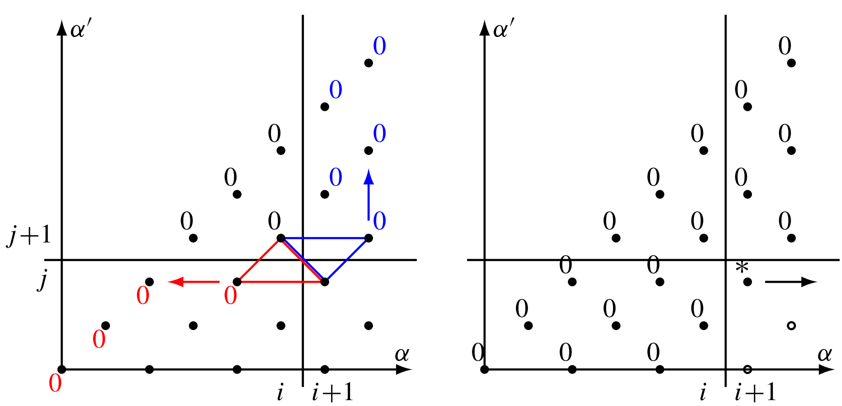


Figure 9

sequence gives by Lemma 4.4 rise to a unique extension $(t_{\alpha,\alpha'})_{(\alpha,\alpha')}$ satisfying (4-1) and (4-2). Hence, by Theorem 4.1 the multiplicity is 1. Similarly we obtain multiplicity 2 for $(r,r')\in L_{\text{even}}$ using Lemma 4.5(1). For $(r,r')\in L_{\text{odd}}$ the multiplicity statement is contained in Lemma 4.5(2).

(2) We first consider the case $\mathcal{V}=\mathcal{F}(i)$ and $\mathcal{W}=\mathcal{F}(j)$. Then any intertwining operator in $\operatorname{Hom}_{(\mathfrak{g}',K')}(\mathcal{V}|_{(\mathfrak{g}',K')},\mathcal{W})$ corresponds to an intertwining operator $T:(\pi_r)_{\text{HC}}\to(\tau_{r'})_{\text{HC}}$ for $r=\rho+i$ and $r'=-\rho'-j$ such that $T|_{\mathcal{E}(\alpha)}=0$ for all $\alpha>i$ and $T(\mathcal{E})\subseteq\mathcal{F}'(j)$. This implies that T is given by a sequence $(t_{\alpha,\alpha'})_{\alpha,\alpha'}$ with $t_{\alpha,\alpha'}=0$ if either $\alpha>i$ or $\alpha'>j$. By part (1) the space of intertwining operators $T:(\pi_r)_{\text{HC}}\to(\tau_{r'})_{\text{HC}}$ is one-dimensional, and using Lemma 4.3(1c) and Step 1 in the proof of Lemma 4.4 it is easy to see that this operator satisfies the conditions on $t_{\alpha,\alpha'}$ if and only if $i-j\in 2\mathbb{N}$. Hence $m(\mathcal{F}(i),\mathcal{F}'(j))=1$ for $i-j\in 2\mathbb{N}$ and $=0$ else. Similar considerations for $r=-\rho-i$ and $r'=\rho'+j$ show that $m(\mathcal{T}(i),\mathcal{T}'(j))=1$ for $i-j\in 2\mathbb{N}$ and $=0$ else.

Now let $\mathcal{V}=\mathcal{T}(i)$ and $\mathcal{W}=\mathcal{F}'(j)$. Then $m(\mathcal{V},\mathcal{W})\neq 0$ if and only if there exists a nontrivial sequence $(t_{\alpha,\alpha'})_{\alpha,\alpha'}$ satisfying (4-1) and (4-2) for $r=-\rho-i$ and $r'=-\rho'-j$ such that $t_{\alpha,\alpha'}=0$ whenever $\alpha\le i$ or $\alpha'>j$. First assume $j>i$, then by part (1) there exists a unique sequence $(t_{\alpha,\alpha'})_{\alpha,\alpha'}$, and by Lemma 4.3(1d) and Step 3 in the proof of Lemma 4.4 it is easy to see that for this sequence $t_{\alpha,\alpha'}=0$ if either $\alpha\le i$ or $\alpha'>j$. Hence $m(\mathcal{T}(i),\mathcal{F}'(j))=1$ in this case. Next assume $j\le i$; then by Lemmas 4.3(2) and 4.5 there can only exist a sequence $(t_{\alpha,\alpha'})_{\alpha,\alpha'}$ with the above properties if $i-j\in 2\mathbb{N}+1$. This shows the claimed formulas for $m(\mathcal{T}(i),\mathcal{F}'(j))$. That $m(\mathcal{F}(i),\mathcal{T}'(j))=0$ for any i,j follows easily by similar considerations. □

4D. ***Explicit formula for the spectral function.*** From the relations (4-1) and (4-2) one can deduce an explicit *spectral function* $(t_{\alpha,\alpha'}(r,r'))_{0\le\alpha'\le\alpha}$, i.e., a set of solutions to the relations for all $r,r'\in\mathbb{C}$ depending meromorphically on r and r':

Proposition 4.6. *For $(\alpha,\alpha')\in\mathbb{N}$ with $\alpha-\alpha'\in 2\mathbb{Z}$ the numbers*

$$t_{\alpha,\alpha'}(r,r')=\sum_{k=0}^{\infty}\frac{2^{4k}\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}+k\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}\big)}{(2k)!\,\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}-k\big)}\times\frac{\Gamma(r+\rho)\Gamma(r'+\rho'+\alpha')\Gamma\big(\frac{2r'+2r+1}{4}+k\big)\Gamma\big(\frac{2r'-2r+3}{4}\big)}{\Gamma(r+\rho+\alpha'+2k)\Gamma(r'+\rho')\Gamma\big(\frac{2r'+2r+1}{4}\big)\Gamma\big(\frac{2r'-2r+3}{4}-k\big)}\tag{4-4}$$

are rational functions in r and r' satisfying (4-1) and (4-2). They are normalized to $t_{0,0}\equiv 1$.

Proof. First note that since $\alpha-\alpha'\in 2\mathbb{Z}$ the number $\frac{\alpha-\alpha'+2}{2}-k$ is a negative integer for $k\gg 0$ and hence the sum is actually finite for each fixed pair (α,α'). It is also easy to see that each summand is a rational function in r and r'. A short calculation shows that for each $k \in \mathbb{N}$ the term

$$\frac{\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}+k\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}\big)\Gamma(r'+\rho'+\alpha')}{\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}-k\big)\Gamma(r+\rho+\alpha'+2k)}$$

solves (4-1). If we further make the ansatz

$$t_{\alpha,\alpha'} = \sum_{k=0}^{\infty} b_k \frac{\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}+k\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}\big)\Gamma(r'+\rho'+\alpha')}{\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}-k\big)\Gamma(r+\rho+\alpha'+2k)},$$

with $b_k = b_k(r, r')$ not depending on α and α' then we find that (4-2) holds if and only if

$$\sum_{k=0}^{\infty} b_k \frac{\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}+k-1\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}\big)\Gamma(r'+\rho'+\alpha'-1)}{\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}-k+1\big)\Gamma(r+\rho+\alpha'+2k)}$$
$$\times\Big[(2r'+2r+4k+1)(2r'-2r-4k-1)\big(\tfrac{\alpha-\alpha'+2}{2}-k\big)\big(\tfrac{\alpha+\alpha'+n-2}{2}+k-1\big)$$
$$-2k(2k-1)(r+\rho+\alpha'+2k-1)(r+\rho+\alpha'+2k-2)\Big] = 0.$$

Substituting $k-1$ for k in the first summand in the brackets gives the condition

$$\sum_{k=1}^{\infty} \frac{\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}+k-1\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}\big)\Gamma(r'+\rho'+\alpha'-1)}{\Gamma\big(\frac{\alpha+\alpha'+n-2}{2}\big)\Gamma\big(\frac{\alpha-\alpha'+2}{2}-k+1\big)\Gamma(r+\rho+\alpha'+2k-2)}$$
$$\times\,[(2r'+2r+4k-3)(2r'-2r-4k+3)b_{k-1} - 2k(2k-1)b_k] = 0,$$

which holds if

$$2k(2k-1)b_k = (2r'+2r+4k-3)(2r'-2r-4k+3)b_{k-1}.$$

This recurrence relation has the solution

$$b_k = c\cdot\frac{2^{4k}\Gamma\big(\frac{2r'+2r+1}{4}+k\big)}{(2k)!\,\Gamma\big(\frac{2r'-2r+3}{4}-k\big)},$$

with $c = c(r, r')$ not depending on k. Finally $t_{0,0} \equiv 1$ implies

$$c = \frac{\Gamma(r+\rho)\Gamma\big(\frac{2r'-2r+3}{4}\big)}{\Gamma(r'+\rho')\Gamma\big(\frac{2r'+2r+1}{4}\big)}. \qquad \square$$

Corollary 4.7. (1) *The renormalized numbers*

$$t^{(1)}_{\alpha,\alpha'}(r, r') = \frac{1}{\Gamma(r+\rho)} t_{\alpha,\alpha'}(r, r')$$

are holomorphic in $(r, r') \in \mathbb{C}^2$ *for every* $\alpha, \alpha' \in \mathbb{N}$, $\alpha - \alpha' \in 2\mathbb{N}$. *Further,* $t^{(1)}_{\alpha,\alpha'}(r, r') = 0$ *for all* α, α' *if and only if* $(r, r') \in L_{\text{even}}$.

(2) *Fix $r'=-\rho'-j,\ j\in\mathbb{N}$; then the renormalized numbers*

$$t^{(2)}_{\alpha,\alpha'}(r,r')=\frac{\Gamma\big(\frac{(r+\rho)-(r'+\rho')}{2}\big)}{\Gamma(r+\rho)}t_{\alpha,\alpha'}(r,r')$$

are holomorphic in $r\in\mathbb{C}$ for every $\alpha,\alpha'\in\mathbb{N},\ \alpha-\alpha'\in 2\mathbb{N}$. We have $t^{(2)}_{\alpha,\alpha'}(r,r')\equiv 0$ for $\alpha'>j$. Further, for every $r\in\mathbb{C}$ there exists a pair (α,α') with $t^{(2)}_{\alpha,\alpha'}(r,r')\neq 0$.

(3) *Fix $N\in\mathbb{N}$ and let $r'+\rho'=r+\rho+2N$; then the renormalized numbers*

$$t^{(3)}_{\alpha,\alpha'}(r,r')=\frac{\Gamma(r'+\rho')}{\Gamma(r+\rho)}t_{\alpha,\alpha'}(r,r')$$

are holomorphic in $r\in\mathbb{C}$ for every $\alpha,\alpha'\in\mathbb{N},\ \alpha-\alpha'\in 2\mathbb{N}$. Further, for every $r\in\mathbb{C}$ there exists $\alpha_0\in\mathbb{N}$ such that $t^{(3)}_{\alpha,\alpha}(r,r')\neq 0$ for $\alpha\geq\alpha_0$.

Proof. (1) We can write

$$t^{(1)}_{\alpha,\alpha'}(r,r')=(r'+\rho')_{\alpha'}\sum_{k=0}^{\frac{\alpha-\alpha'}{2}}\frac{2^{4k}\big(\frac{\alpha+\alpha'+n-2}{2}\big)_k\big(-\frac{\alpha-\alpha'}{2}\big)_k\big(\frac{2r'+2r+1}{4}\big)_k\big(\frac{2r-2r'+1}{4}\big)_k}{(2k)!\,\Gamma(r+\rho+\alpha'+2k)},$$

where $(\lambda)_n=\lambda(\lambda+1)\cdots(\lambda+n-1)$ denotes the Pochhammer symbol. This expression is obviously holomorphic in $(r,r')\in\mathbb{C}^2$. Now assume $t^{(1)}_{\alpha,\alpha'}(r,r')=0$ for all α,α'. For $\alpha=\alpha'$ we have $\big(-\frac{\alpha-\alpha'}{2}\big)_k=0$ for $k>0$ and hence

$$t^{(1)}_{\alpha,\alpha}(r,r')=\frac{(r'+\rho')_\alpha}{\Gamma(r+\rho+\alpha)},$$

which vanishes for all $\alpha\in\mathbb{N}$ if and only if $r+\rho=-i$ and $r'+\rho'=-j$ with $j\leq i$. We claim that $i-j\in 2\mathbb{N}$. In fact, if $i-j\in 2\mathbb{N}+1$ then for $(\alpha,\alpha')=(i+1,j)$ only the summand for $k=\frac{i-j+1}{2}$ is nonzero and hence $t^{(1)}_{\alpha,\alpha'}(r,r')\neq 0$, a contradiction. Therefore $i-j\in 2\mathbb{N}$ which means $(r,r')\in L_{\text{even}}$.

Conversely assume $r+\rho=-i,\ r'+\rho'=-j$, with $i-j\in 2\mathbb{N}$. Then in each summand at least one of the three factors

$$\big(\tfrac{2r-2r'+1}{4}\big)_k=\big(-\tfrac{i-j}{2}\big)_k,\quad (r'+\rho')_{\alpha'}=(-j)_{\alpha'},\quad \tfrac{1}{\Gamma(r+\rho+\alpha'+2k)}=\tfrac{1}{\Gamma(-i+\alpha'+2k)}$$

vanishes and hence $t^{(1)}_{\alpha,\alpha'}(r,r')=0$ for all α,α'.

(2) We can write

$$t^{(2)}_{\alpha,\alpha'}(r,r')=(-j)_{\alpha'}\sum_{k=0}^{\frac{\alpha-\alpha'}{2}}\frac{2^{4k}\big(\frac{\alpha+\alpha'+n-2}{2}\big)_k\big(-\frac{\alpha-\alpha'}{2}\big)_k\big(\frac{2r-2j-n+3}{4}\big)_k\Gamma\big(\frac{2r+2j+n-1}{4}+k\big)}{(2k)!\,\Gamma(r+\rho+\alpha'+2k)}$$

as a meromorphic function of r. Then $t^{(2)}_{\alpha,\alpha'}(r,r')\equiv 0$ for $\alpha'>j$. Further, for $\alpha'\leq j$ each pole r of the factor $\Gamma\big(\frac{2r+2j+n-1}{4}+k\big)$ is simple and also a pole of the denominator $\Gamma(r+\rho+\alpha'+2k)$, whence $t^{(2)}_{\alpha,\alpha'}(r,r')$ is holomorphic in $r\in\mathbb{C}$. Now assume $t^{(2)}_{\alpha,\alpha'}(r,r')=0$ for all α,α'. Then

$$0=t^{(2)}_{j,j}(r,r')=(-j)_j\frac{\Gamma\big(\frac{2r+2j+n-1}{4}\big)}{\Gamma(r+\rho+j)}$$

and hence r has to be a pole of the denominator while it is a regular point for the numerator. This means $r+\rho=-i\in-\mathbb{N}$ with $i\geq j$ and $\frac{2r+2j+n-1}{4}=\frac{j-i}{2}\notin-\mathbb{N}$, i.e., $i-j\in 2\mathbb{N}+1$. But for $(\alpha,\alpha')=(i+1,j)$ only the summand for $k=\frac{i-j+1}{2}$ is nonzero and hence $t^{(2)}_{\alpha,\alpha'}(r,r')\neq 0$, a contradiction.

(3) Note that $\big(\frac{2r-2r'+1}{4}\big)_k=(-N)_k=0$ for $k>N$ and hence we can write

$$t^{(3)}_{\alpha,\alpha'}(r,r')=\sum_{k=0}^{N}\frac{2^{4k}(-N)_k\big(\frac{\alpha+\alpha'+n-2}{2}\big)_k\big(-\frac{\alpha-\alpha'}{2}\big)_k}{(2k)!}\times\big(r+N+\tfrac{1}{2}\big)_k(r+\rho+\alpha'+2k)_{2N-2k},\tag{4-5}$$

which is clearly holomorphic in $r\in\mathbb{C}$. Further, $t^{(3)}_{\alpha,\alpha}(r,r')=(r+\rho+\alpha)_{2N}$, which is nonzero for $\alpha>-(r+\rho)$. □

Remark 4.8. After a few modifications we find that

$$t_{\alpha,\alpha'}(r,r')=\frac{(r'+\rho')_{\alpha'}}{(r+\rho)_{\alpha'}}{}_4F_3\Big(-\tfrac{\alpha-\alpha'}{2},\tfrac{\alpha+\alpha'+n-2}{2},\tfrac{2r+2r'+1}{4},\tfrac{2r-2r'+1}{4};\ \tfrac{1}{2},\tfrac{r+\rho+\alpha'}{2},\tfrac{r+\rho+\alpha'+1}{2};1\Big).$$

Note that the generalized hypergeometric function ${}_4F_3(a_1,a_2,a_3,a_4;b_1,b_2,b_3;z)$ occurring here is balanced, i.e., $a_1+a_2+a_3+a_4+1=b_1+b_2+b_3$. However, there does not exist an explicit formula for its special value at $z=1$ in the literature. Also, we could not find estimates for special values of such hypergeometric functions for large/small parameters, and therefore were not able to show that $t_{\alpha,\alpha'}(r,r')$ grows at most polynomially in $\alpha,\alpha'\geq 0$ for fixed $(r,r')\in\mathbb{C}^2$. This is what is needed to apply Proposition 3.10 in order to show automatic continuity of intertwining operators. We will therefore first describe all intertwining operators in terms of the holomorphic family $T^{(1)}(r,r')$ (see Theorem 4.9) and then show automatic continuity using the corresponding holomorphic family in the smooth category obtained in joint work with Y. Oshima [Möllers et al. 2016a]. This is done in Corollary 4.12.

Theorem 4.9. *For $i=1,2,3$ we let $T^{(i)}(r,r')$ be the intertwining operators $(\pi_r)_{\mathrm{HC}}\to(\tau_{r'})_{\mathrm{HC}}$ corresponding to the numbers $t^{(i)}_{\alpha,\alpha'}(r,r')$ in Corollary 4.7. Then the operator $T^{(1)}(r,r')$ is defined for $(r,r')\in\mathbb{C}^2$, the operator $T^{(2)}(r,r')$ is defined for $r'\in-\rho'-\mathbb{N}$ and the operator $T^{(3)}(r,r')$ is defined for $(r+\rho)-(r'+\rho')\in-2\mathbb{N}$.*

We have

$$\operatorname{Hom}_{(\mathfrak{g}',K')}((\pi_r)_{\mathrm{HC}}|_{(\mathfrak{g}',K')},(\tau_{r'})_{\mathrm{HC}}) = \begin{cases} \mathbb{C}T^{(1)}(r,r') & \text{for } (r,r')\in\mathbb{C}^2\setminus L_{\mathrm{even}}, \\ \mathbb{C}T^{(2)}(r,r')\oplus\mathbb{C}T^{(3)}(r,r') & \text{for } (r,r')\in L_{\mathrm{even}}. \end{cases}$$

Remark 4.10. By the proof of Theorem 4.2(2) every intertwining operator between the subquotients $\mathcal{F}(i),\mathcal{T}(i)$ and $\mathcal{F}'(j),\mathcal{T}'(j)$ can be constructed by composing an intertwining operator $(\pi_r)_{\mathrm{HC}}\to(\tau_{r'})_{\mathrm{HC}}$ for particular r,r' with embeddings and/or quotient maps for the subquotients. Hence, also every intertwining operator between subquotients is given by an operator in one of the three families $T^{(i)}(r,r')$. Therefore, all information about intertwining operators between $(\pi_r)_{\mathrm{HC}}$ and $(\tau_{r'})_{\mathrm{HC}}$ and any of their subquotients is contained in the holomorphic family $T^{(1)}(r,r')$.

Remark 4.11. The family of operators $T^{(3)}$ is (up to a constant) equal to Juhl's family of conformally invariant differential restriction operators $D_{2N}(r):C^\infty(S^{n-1})\to C^\infty(S^{n-2})$; see [Juhl 2009; Kobayashi and Speh 2015]. The constants $t^{(3)}_{\alpha,\alpha'}$ then give the "spectrum" of Juhl's operators in the sense that they describe how the operators are acting on explicit K-finite vectors. Note that by (4-5) the number of summands for $t^{(3)}_{\alpha,\alpha'}(r,r')$ is at most $N+1$.

Corollary 4.12. *For* $(G,G')=(\mathrm{O}(1,n),\mathrm{O}(1,n-1))$ *the natural injective map*

$$\operatorname{Hom}_{G'}(\pi|_{G'},\tau)\to\operatorname{Hom}_{(\mathfrak{g}',K')}(\pi_{\mathrm{HC}}|_{(\mathfrak{g}',K')},\tau_{\mathrm{HC}}) \tag{4-6}$$

is an isomorphism for all spherical principal series π *of* G *and* τ *of* G' *and their subquotients.*

Proof. By Remark 4.10 all intertwining operators between subquotients arise by composing with quotient maps and embeddings. It therefore suffices to show that (4-6) is an isomorphism for $\pi=\pi_r$ and $\tau=\tau_{r'}$ for all $(r,r')\in\mathbb{C}^2$. In [Möllers et al. 2016a] a holomorphic family $A(r,r')\in\operatorname{Hom}_{G'}(\pi_r|_{G'},\tau_{r'})$ was constructed in the smooth category using singular integral operators (see Section 4F for details). Denote by $\bar{A}(r,r')\in\operatorname{Hom}_{(\mathfrak{g}',K')}((\pi_r)_{\mathrm{HC}}|_{(\mathfrak{g}',K')},(\tau_{r'})_{\mathrm{HC}})$ its image under the map (4-6). By Theorem 4.9 this space is generically spanned by $T^{(1)}(r,r')$, and since both $\bar{A}(r,r')$ and $T^{(1)}(r,r')$ depend holomorphically on $(r,r')\in\mathbb{C}^2$ there exists a meromorphic function $\phi(r,r')$ such that

$$\bar{A}(r,r')=\phi(r,r')\cdot T^{(1)}(r,r').$$

Replacing $A(r,r')$ and $\bar{A}(r,r')$ by $\phi(r,r')^{-1}A(r,r')$ and $\phi(r,r')^{-1}\bar{A}(r,r')$ we may assume that

$$\bar{A}(r,r')=T^{(1)}(r,r').$$

This already implies that for $(r, r') \in \mathbb{C}^2 \setminus L_{\text{even}}$ every intertwining operator in the space $\operatorname{Hom}_{(\mathfrak{g}',K')}((\pi_r)_{\text{HC}}|_{(\mathfrak{g}',K')}, (\tau_{r'})_{\text{HC}})$ extends to the smooth globalization. Further, for $(r, r') \in L_{\text{even}}$ we may restrict $(r, r') \mapsto T^{(1)}(r, r')$ to an affine complex line and renormalize to obtain all intertwining operators in $\operatorname{Hom}_{(\mathfrak{g}',K')}((\pi_r)_{\text{HC}}|_{(\mathfrak{g}',K')}, (\tau_{r'})_{\text{HC}})$ by Theorem 4.9. The same restriction and renormalization can be applied to $(r, r') \mapsto A(r, r')$, and in this way one obtains extensions of all operators in $\operatorname{Hom}_{(\mathfrak{g}',K')}((\pi_r)_{\text{HC}}|_{(\mathfrak{g}',K')}, (\tau_{r'})_{\text{HC}})$ to the smooth globalization. Note that renormalization of $A(r, r')$ preserves continuity of the operators. This shows that the map (4-6) is surjective, hence an isomorphism for all $(r, r') \in \mathbb{C}^2$. □

Remark 4.13. The operators $T^{(i)}(r, r')$ are related to the operators $\widetilde{\mathbb{A}}_{\lambda,\nu}$, $\widetilde{\widetilde{\mathbb{A}}}_{\lambda,\nu}$ and $\widetilde{\mathbb{C}}_{\lambda,\nu}$ studied in [Kobayashi and Speh 2015] for $\lambda = r + \rho$, $\nu = r' + \rho'$. In fact, using their notation we have

$$T^{(1)}(r, r') = \pi^{-\frac{n-2}{2}} \widetilde{\mathbb{A}}_{\lambda,\nu}, \quad T^{(2)}(r, r') = \pi^{-\frac{n-2}{2}} \widetilde{\widetilde{\mathbb{A}}}_{\lambda,\nu}, \quad T^{(3)}(r, r') = \frac{(-1)^N N!}{2^{2N}} \widetilde{\mathbb{C}}_{\lambda,\nu},$$

where for $i = 3$ we write $r' + \rho' = r + \rho + 2N$ with $N \in \mathbb{N}$.

4E. *Discrete components in the restriction of unitary representations.* We apply our results to branching problems for unitary representations. The $(\mathfrak{g}, K)$-modules $(\pi_r)_{\text{HC}}$ are unitarizable if and only if $r \in i\mathbb{R} \cup (-\rho, \rho)$ and we denote by $\hat{\pi}_r$ their unitary completions. For $r \in i\mathbb{R}$ these representations form the unitary principal series and for $r \in (-\rho, \rho)$ they belong to the complementary series. Further, all irreducible quotients $\mathcal{T}(i)$ are unitarizable and their unitary completions will be denoted by $\hat{\pi}_{-\rho-i}$. We note that for $r \in -(\rho + \mathbb{Z})$, $r < 0$, each representation $\hat{\pi}_r$ is isomorphic to some Zuckerman derived functor module $A_{\mathfrak{q}}(\lambda)$ and occurs discretely in the decomposition of the regular representation on $L^2(G/G')$.

Similarly we denote by $\hat{\tau}_{r'}$, $r' \in i\mathbb{R} \cup (-\rho', \rho')$, the unitary completions of $\tau_{r'}$ and by $\hat{\tau}_{-\rho'-j}$, $j \in \mathbb{N}$, the unitary completions of $\mathcal{T}'(j)$.

For $r \in \mathbb{R}$ we define the finite set

$$D(r) = \left(r + \tfrac{1}{2} + 2\mathbb{N}\right) \cap (-\infty, 0)$$

and note that for $r \in (-\rho, 0) \cup (-\rho - \mathbb{N})$ and $r' \in D(r)$ we have $r' \in (-\rho', 0) \cup (-\rho' - \mathbb{N})$; i.e., $\hat{\tau}_{r'}$ is a unitary representation.

Theorem 4.14. *Let $r \in (-\rho, 0) \cup (-\rho - \mathbb{N})$. Then for every $r' \in D(r)$ the representation $\hat{\tau}_{r'}$ occurs discretely with multiplicity* 1 *in the restriction of $\hat{\pi}_r$ to G'.*

We note that for a complementary series representation $\hat{\pi}_r$, $r \in (-\rho, 0)$, all representations $\hat{\tau}_{r'}$, $r' \in D(r)$, are complementary series representations. If $\hat{\pi}_r$ is an $A_{\mathfrak{q}}(\lambda)$-module, $r \in -\rho + \mathbb{Z}$, $r < 0$, then so are the representations $\hat{\tau}_{r'}$, $r' \in D(r)$.

The restriction of the $A_{\mathfrak{q}}(\lambda)$-modules $\hat{\pi}_r$ to G' decomposes with both discrete and continuous spectrum and is therefore hard to study by purely algebraic methods.

Remark 4.15. For the special case $r' = r + \frac{1}{2}$, i.e., $N = 0$, the occurrence of $\hat{\tau}_{r'}$ in $\hat{\pi}_r|_{G'}$ was first proved in [Speh and Venkataramana 2011] for $r \in \left[-\rho, -\frac{1}{2}\right)$ and generalized in [Zhang 2015] to the case $r \in \left(-\rho, -\frac{1}{2}\right) \cup (-\rho - \mathbb{N})$. Later Kobayashi and Speh [2015, Theorem 1.4] proved Theorem 4.14 for the case $r \in (-\rho, 0)$. The full decomposition of $\hat{\pi}_r|_{G'}$ for $r \in (-\rho, 0) \cup (-\rho - \mathbb{N})$ including the continuous spectrum was given in [Möllers and Oshima 2015].

We first describe the invariant norms on the unitarizable constituents for $r \in \mathbb{R}$. For this we fix the L^2-norm $\|\cdot\|_{L^2(S^{n-1})}$ on $L^2(K/M) = L^2(S^{n-1})$ corresponding to the standard Euclidean measure on S^{n-1}. For $r \in (-\rho, \rho)$ the norm $\|\cdot\|_r$ on $\mathcal{E}$ given by

$$\|v\|_r^2 = \sum_{\alpha=0}^{\infty} b_\alpha(r) \|v_\alpha\|_{L^2(S^{n-1})}^2 \quad \text{for } v = \sum_{\alpha=0}^{\infty} v_\alpha \in \bigoplus_{\alpha=0}^{\infty} \mathcal{E}(\alpha),$$

with

$$b_\alpha = \frac{\Gamma(\rho - r + \alpha)}{\Gamma(\rho + r + \alpha)} \sim (1+\alpha)^{-2r}$$

turns $(\pi_r)_{\mathrm{HC}}$ into a unitary $(\mathfrak{g}, K)$-module. Further, for $r = -\rho - i$ the seminorm $\|\cdot\|_r$ on $\mathcal{E}$ has kernel $\mathcal{F}(i)$ and turns the quotient $\mathcal{T}(i) = \mathcal{E}/\mathcal{F}(i)$ into a unitary $(\mathfrak{g}, K)$-module.

Similarly we denote by $\|\cdot\|'_{r'}$ the $\tau_{r'}$-invariant norm on $\mathcal{E}'$, respectively $\mathcal{T}'(j)$, given by

$$\|w\|_{r'}^{\prime 2} = \sum_{\alpha=0}^{\infty} b'_{\alpha'}(r') \|w_{\alpha'}\|_{L^2(S^{n-2})}^2 \quad \text{for } w = \sum_{\alpha'=0}^{\infty} w_{\alpha'} \in \bigoplus_{\alpha'=0}^{\infty} \mathcal{E}'(\alpha'),$$

with

$$b'_{\alpha'} = \frac{\Gamma(\rho' - r' + \alpha')}{\Gamma(\rho' + r' + \alpha')} \sim (1+\alpha')^{-2r'}.$$

We need the following two basic results; see, e.g., [Zhang 2015, Lemmas 3.2 and 3.5]:

Lemma 4.16. *Let $\mathcal{V} \subseteq \mathcal{E}$ be a K-invariant subspace and $\mathcal{W} \subseteq \mathcal{E}'$ a K'-invariant subspace and assume that $\mathcal{V}$ and $\mathcal{W}$ are endowed with pre-Hilbert space structures with respect to which the groups K and K' act unitarily. A linear map $T : \mathcal{V} \to \mathcal{W}$ is bounded if and only if there exists a constant $C > 0$ such that*

$$\sum_{\substack{\alpha \\ \mathcal{E}(\alpha;\alpha') \subseteq \mathcal{V}}} \|T|_{\mathcal{E}(\alpha;\alpha')}\|_{\mathcal{V}\to\mathcal{W}}^2 \leq C \quad \textit{for all } \alpha',$$

where $\|\cdot\|_{\mathcal{V}\to\mathcal{W}}$ denotes the operator norm with respect to the given pre-Hilbert space structures.

Lemma 4.17. *Suppose that* $\alpha > -1$, $\beta \geq 0$, *and* $\beta - \alpha > 1$. *Then there exists a constant* $C > 0$ *such that*

$$\sum_{p=0}^{\infty} \frac{(1+p)^\alpha}{(1+p+q)^\beta} \leq \frac{C}{(1+q)^{\beta-\alpha-1}} \quad \textit{for all } q \geq 0.$$

Proof of Theorem 4.14. For $r \in (-\rho, 0)$ let $\mathcal{V} = \mathcal{E}$ and for $r = -\rho - i \in -\rho - \mathbb{N}$ let $\mathcal{V} = \bigoplus_{\alpha=i+1}^{\infty} \mathcal{E}(\alpha)$. Let $r' \in D(r)$; then similarly we put $\mathcal{W} = \mathcal{E}'$ for $r' \in (-\rho', 0)$ and $\mathcal{W} = \bigoplus_{\alpha'=j+1}^{\infty} \mathcal{E}'(\alpha')$ for $r' = -\rho' - j \in -\rho' - \mathbb{N}$. By Theorem 4.2 there exists (up to scalar) a unique nonzero intertwining operator $T : (\pi_r)_{\mathrm{HC}} \to (\tau_{r'})_{\mathrm{HC}}$ with $T(\mathcal{V}) \subseteq \mathcal{W}$ and if $r = -\rho - i$ additionally $T|_{\mathcal{F}(i)} = 0$. In our notation

$$T|_{\mathcal{E}(\alpha;\alpha')} = t_{\alpha,\alpha'} \cdot \mathrm{rest}|_{\mathcal{E}(\alpha;\alpha')},$$

with $t_{\alpha,\alpha'} = t^{(3)}_{\alpha,\alpha'}$ for $\alpha' > j$ and $t_{\alpha,\alpha'} = 0$ else (see Corollary 4.7 for the definition of $t^{(3)}_{\alpha,\alpha'}$). We show that T is bounded if we endow $\mathcal{V}$ with the norm $\|\cdot\|_r$ and $\mathcal{W}$ with the norm $\|\cdot\|_{r'}$. To apply Lemma 4.16 we calculate

$$\|T|_{\mathcal{E}(\alpha;\alpha')}\|^2_{\mathcal{V}\to\mathcal{W}} = t^2_{\alpha,\alpha'} \|\,\mathrm{rest}|_{\mathcal{E}(\alpha;\alpha')}\|^2_{\mathcal{E}(\alpha;\alpha')\to\mathcal{E}'(\alpha')} \frac{b'_{\alpha'}(r')}{b_\alpha(r)},$$

where $\|\cdot\|_{\mathcal{E}(\alpha;\alpha')\to\mathcal{E}'(\alpha')}$ denotes the operator norm with respect to the L^2-inner products on $\mathcal{E}(\alpha;\alpha') \subseteq L^2(S^{n-1})$ and $\mathcal{E}'(\alpha') \subseteq L^2(S^{n-2})$. Using (A-2), (B-2) and (B-3) it is easy to see that for $\alpha = \alpha' + 2\ell$ we have

$$\begin{aligned}\|\,\mathrm{rest}|_{\mathcal{E}(\alpha;\alpha')}\|^2_{\mathcal{E}(\alpha;\alpha')\to\mathcal{E}'(\alpha')} &= \frac{2^{2\alpha'+n-3}\big(\alpha'+2\ell+\frac{n-2}{2}\big)(2\ell)!\,\Gamma\big(\alpha'+\ell+\frac{n-2}{2}\big)^2}{\pi(\ell!)^2\Gamma(2\alpha'+2\ell+n-2)} \\ &= \frac{\big(\alpha'+2\ell+\frac{n-2}{2}\big)\Gamma\big(\ell+\frac{1}{2}\big)\Gamma\big(\alpha'+\ell+\frac{n-2}{2}\big)}{\pi\Gamma(\ell+1)\Gamma\big(\alpha'+\ell+\frac{n-1}{2}\big)} \sim \frac{(1+\alpha'+\ell)^{\frac{1}{2}}}{(1+\ell)^{\frac{1}{2}}}.\end{aligned}$$

Then Lemma 4.16 translates into

$$\sum_{\ell=0}^{\infty} t^2_{\alpha'+2\ell,\alpha'} \frac{(1+\alpha'+\ell)^{\frac{1}{2}+2r}}{(1+\ell)^{\frac{1}{2}}} \leq C(1+\alpha')^{2r'}.$$

It is enough to check this for each of the $N+1$ summands of $t_{\alpha'+2\ell,\alpha'}$ in (4-5) where $r' + \rho' = r + \rho + 2N$. The k-th summand grows of order

$$\sim (1+\alpha')^{(r'+\rho')-(r+\rho+2k)}(1+\ell)^k(1+\alpha'+\ell)^k$$

and hence the claim follows by Lemma 4.17. Altogether this shows that T induces a bounded G'-intertwining operator $\widetilde{T} : \hat{\pi}_r|_{G'} \to \hat{\tau}_{r'}$ whose adjoint $\widetilde{T}^* : \hat{\tau}_{r'} \to \hat{\pi}_r|_{G'}$ embeds $\hat{\tau}_{r'}$ isometrically as a subrepresentation of $\hat{\pi}_r$ by Schur's lemma. Multiplicity 1 follows from the fact that any G'-equivariant embedding $S : \hat{\tau}_{r'} \to \hat{\pi}_r|_{G'}$ induces an intertwiner $S^* : (\pi_r)_{\mathrm{HC}} \to (\tau_{r'})_{\mathrm{HC}}$ between the Harish-Chandra modules by taking the adjoint operator and then passing to K-finite vectors. Such an operator is unique (up to scalars) by Theorem 4.2 and since K-finite vectors are dense in $\hat{\pi}_r$ the embedding S is unique (up to scalars). □

4F. *Comparison with singular integral operators.* In [Kobayashi and Speh 2015; Möllers et al. 2016a] a meromorphic family of intertwining operators $A(r,r'):u_{\mathbf{1},r\nu}|_{G'}\to u'_{\mathbf{1},r'\nu}$ in the smooth category is constructed as family of singular integral operators. In the compact picture this family is (up to scalars) given by

$$A(r,r'):C^\infty(S^{n-1})\to C^\infty(S^{n-2}),$$
$$A(r,r')f(y)=\int_{S^{n-1}}(|x'-y|^2+x_n^2)^{-(r'+\rho')}|x_n|^{(r-\rho)+(r'+\rho')}f(x)\,dx,$$

where dx denotes the Euclidean measure on S^{n-1}.

Theorem 4.18. *Let $T(r,r'):C^\infty(S^{n-1})\to C^\infty(S^{n-2})$ denote the intertwining operator with spectrum given by the numbers $t_{\alpha,\alpha'}(r,r')$ in (4-4). Then*

$$A(r,r')=\frac{2^{r-r'+\frac{1}{2}}\pi^{\frac{n-2}{2}}\Gamma\left(\frac{2r+2r'+1}{4}\right)\Gamma\left(\frac{2r-2r'+1}{4}\right)}{\Gamma\left(r+\frac{n-1}{2}\right)}\cdot T(r,r').$$

Proof. Since by Theorem 4.2(1) and Corollary 4.12 we generically have

$$\dim\operatorname{Hom}_{G'}(\pi_r|_{G'},\tau_{r'})=1$$

and both $A(r,r')$ and $T(r,r')$ are meromorphic in $r,r'\in\mathbb{C}$ there exists a scalar meromorphic function $c(r,r')$ with $A(r,r')=c(r,r')T(r,r')$. To determine $c(r,r')$ we put $f\equiv 1$:

$$c(r,r')=\int_{S^{n-1}}(|x'-y|^2+x_n^2)^{-\left(r'+\frac{n-2}{2}\right)}|x_n|^{r+r'-\frac{1}{2}}\,dx.$$

Using the stereographic projection

$$x=\left(\frac{1-|z|^2}{1+|z|^2},\frac{2z}{1+|z|^2}\right),\qquad z\in\mathbb{R}^{n-1},$$

the measure transforms by $dx=2^{n-1}(1+|z|^2)^{-(n-1)}\,dz$, where dz is the standard Lebesgue measure on $\mathbb{R}^{n-1}$. Writing

$$y=\left(\frac{1-|w|^2}{1+|w|^2},\frac{2w}{1+|w|^2}\right),\qquad w\in\mathbb{R}^{n-2},$$

we find

$$c(r,r')=2^{r-r'+\frac{1}{2}}(1+|w|^2)^{r'+\frac{n-2}{2}}\int_{\mathbb{R}^{n-1}}(|z'-w|^2+z_{n-1}^2)^{-\left(r'+\frac{n-2}{2}\right)}\times|z_{n-1}|^{r+r'-\frac{1}{2}}(1+|z|^2)^{-\left(r+\frac{n-1}{2}\right)}\,dz,$$

where we have written $z = (z', z_{n-1})$. This integral is evaluated in [Kobayashi and Speh 2015, Proposition 7.4] and we obtain

$$c(r,r') = \frac{2^{r-r'+\frac{1}{2}}\pi^{\frac{n-2}{2}}\Gamma\big(\frac{2r+2r'+1}{4}\big)\Gamma\big(\frac{2r-2r'+1}{4}\big)}{\Gamma\big(r+\frac{n-1}{2}\big)},$$

which shows the claim. □

Remark 4.19. The special value of the intertwiners $A(r,r')$ at the spherical vector $f \equiv 1$ was also calculated in [Möllers and Ørsted 2017] by a different method.

Remark 4.20. The action of $A(r,r')$ on K-finite vectors was also computed in [Kobayashi and Speh 2015, Lemma 7.7]. However, their parametrization of K-finite vectors differs from our parametrization by (α,α'), and therefore it is nontrivial to see the equivalence of their identity and our identity (4-4).

5. Rank-one unitary groups

We indicate in this section how the calculations in Section 4 can be generalized to rank-one unitary groups and state the corresponding results. Let $n \geq 2$ and consider the indefinite unitary group $G = \mathrm{U}(1,n)$ of $(n+1)\times(n+1)$ complex matrices leaving the standard Hermitian form on $\mathbb{C}^{n+1}$ of signature $(1,n)$ invariant. The subgroup $G' \subseteq G$ of matrices fixing the last standard basis vector e_{n+1} is isomorphic to $\mathrm{U}(1,n-1)$.

5A. ***K-types.*** We fix $K = \mathrm{U}(1)\times\mathrm{U}(n)$ and choose

$$H = \begin{pmatrix} 0 & 1 & \\ 1 & 0 & \\ & & \mathbf{0}_{n-1} \end{pmatrix}$$

so that $P = MAN$ with $M = \Delta\,\mathrm{U}(1)\times\mathrm{U}(n-1)$, where $\Delta\,\mathrm{U}(1) = \{\mathrm{diag}(x,x) : x \in \mathrm{U}(1)\}$. Note that $\rho = n$. Then K acts transitively on the unit sphere $S^{2n-1} \subseteq \mathbb{C}^n$ via $\mathrm{diag}(\lambda,k)\cdot z = \lambda^{-1}kz,\ \lambda \in \mathrm{U}(1),\ k \in \mathrm{U}(n),\ z \in S^{2n-1}$, and M is the stabilizer subgroup of the first standard basis vector e_1, whence $K/M \cong S^{2n-1}$. The subgroup $G' = \mathrm{U}(1,n-1)$ is embedded into G such that $K' = \mathrm{U}(1)\times\mathrm{U}(n-1)$ and $P' = G' \cap P = M'A'N'$ with $A' = A$ and $M' = \Delta\,\mathrm{U}(1)\times\mathrm{U}(n-2)$. Then $K'/M' = S^{2n-3} \subseteq \mathbb{C}^{n-1}$, viewed as the codimension 2 submanifold in $K/M = S^{2n-1} \subseteq \mathbb{C}^n$ given by $z_n = 0$. Further we have $\rho' = n-1$.

Let $\xi = \mathbf{1},\ \xi' = \mathbf{1}$ be the trivial representations of M and M' and abbreviate $\pi_r = \pi_{\xi,r}$ and $\tau_{r'} = \tau_{\xi',r'}$. Then as K-modules, resp. K'-modules, we have

$$\mathcal{E} = \bigoplus_{\alpha_1,\alpha_2=0}^{\infty} \underbrace{e^{i(\alpha_1-\alpha_2)\theta} \boxtimes \mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n)}_{\mathcal{E}(\alpha)}, \qquad \mathcal{E}' = \bigoplus_{\alpha_1',\alpha_2'=0}^{\infty} \underbrace{e^{i(\alpha_1'-\alpha_2')\theta} \boxtimes \mathcal{H}^{\alpha_1',\alpha_2'}(\mathbb{C}^{n-1})}_{\mathcal{E}'(\alpha')},$$

where we abbreviate $\alpha=(\alpha_1,\alpha_2)$ and $\alpha'=(\alpha_1',\alpha_2')$. Hence, (MF1) is satisfied. Further, each K-type decomposes by (B-5) into K'-types as

$$(e^{i(\alpha_1-\alpha_2)\theta}\boxtimes\mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n))|_{K'}=\bigoplus_{\substack{0\le\alpha_1'\le\alpha_1\\0\le\alpha_2'\le\alpha_2}}(e^{i(\alpha_1-\alpha_2)\theta}\boxtimes\mathcal{H}^{\alpha_1',\alpha_2'}(\mathbb{C}^{n-1})),$$

so that (MF2) holds. Comparing the characters of the U(1)-factor of K' we find that $\mathrm{Hom}_{K'}(\mathcal{E}(\alpha)|_{K'},\mathcal{E}'(\alpha'))\neq 0$ if and only if $\alpha_1-\alpha_2=\alpha_1'-\alpha_2'$. In this case formulas (B-6) and (A-3) show that the restriction operator

$$R_{\alpha,\alpha'}=\mathrm{rest}\,|_{\mathcal{E}(\alpha;\alpha')}:\mathcal{E}(\alpha;\alpha')\to\mathcal{E}'(\alpha')$$

is an isomorphism. Hence the restriction $T_{\alpha,\alpha'}=T|_{\mathcal{E}(\alpha;\alpha')}$ of a K'-intertwining operator $T:\mathcal{E}\to\mathcal{E}'$ is given by $T_{\alpha,\alpha'}=t_{\alpha,\alpha'}R_{\alpha,\alpha'}$ for $\alpha_1-\alpha_2=\alpha_1'-\alpha_2'$ and $T_{\alpha,\alpha'}=0$ else.

5B. ***Proportionality constants.*** The eigenvalues of the spectrum-generating operator on the K-types are given by (see [Branson et al. 1996, Section 3.b])

$$\sigma_{(\alpha_1,\alpha_2)}=2\alpha_1(\alpha_1+n-1)+2\alpha_2(\alpha_2+n-1),$$
$$\sigma'_{(\alpha_1',\alpha_2')}=2\alpha_1'(\alpha_1'+n-2)+2\alpha_2'(\alpha_2'+n-2).$$

We write $\mathfrak{s}_{\mathbb{C}}=\mathfrak{s}+J\mathfrak{s}=\mathfrak{s}_++\mathfrak{s}_-$ and identify $\mathfrak{s}_\pm\cong\mathbb{C}^n$ via

$$\mathbb{C}^n\to\mathfrak{s}_\pm,\quad w\mapsto X_{w,\pm}=\begin{pmatrix}0 & w^*\mp Jiw^*\\ w\pm Jiw & \mathbf{0}_n\end{pmatrix}.$$

Then $\mathfrak{s}_\pm'\simeq\mathbb{C}^{n-1}$, embedded in $\mathbb{C}^n$ as the first $n-1$ coordinates. Since both $\mathfrak{s}_\pm'$ are multiplicity-free K'-modules, (MF3) holds (with $\mathfrak{s}_{\mathbb{C}}'$ replaced by $\mathfrak{s}_\pm'$) and we can use Corollary 3.6. The cocycle ω is given by

$$\omega(X_{w,+})(z)=w^*z,\quad w\in\mathfrak{s}_+,\qquad \omega(X_{w,-})(z)=z^*w,\quad w\in\mathfrak{s}_-,$$

where $z\in S^{2n-1}\subseteq\mathbb{C}^n$.

We note by (B-7) that if $X\in\mathfrak{s}_+$ then the multiplication map $m(\omega(X))$ maps the K-type $\mathcal{E}(\alpha_1,\alpha_2)$ into the K-types $\mathcal{E}(\alpha_1+1,\alpha_2)$ and $\mathcal{E}(\alpha_1,\alpha_2-1)$ and if $X\in\mathfrak{s}_-$ into the K-types $\mathcal{E}(\alpha_1,\alpha_2+1)$ and $\mathcal{E}(\alpha_1-1,\alpha_2)$. Because of similar considerations for $\mathfrak{s}_+'$ and $\mathfrak{s}_-'$ the equivalence relation $(\alpha,\alpha')\leftrightarrow(\beta,\beta')$ is given by

$$((\alpha_1,\alpha_2);(\alpha_1',\alpha_2'))\leftrightarrow(\beta;(\alpha_1'+1,\alpha_2'))\iff\beta\in\{(\alpha_1+1,\alpha_2),(\alpha_1,\alpha_2-1)\},$$
$$((\alpha_1,\alpha_2);(\alpha_1',\alpha_2'))\leftrightarrow(\beta;(\alpha_1'-1,\alpha_2'))\iff\beta\in\{(\alpha_1-1,\alpha_2),(\alpha_1,\alpha_2+1)\},$$
$$((\alpha_1,\alpha_2);(\alpha_1',\alpha_2'))\leftrightarrow(\beta;(\alpha_1',\alpha_2'+1))\iff\beta\in\{(\alpha_1-1,\alpha_2),(\alpha_1,\alpha_2+1)\},$$
$$((\alpha_1,\alpha_2);(\alpha_1',\alpha_2'))\leftrightarrow(\beta;(\alpha_1',\alpha_2'-1))\iff\beta\in\{(\alpha_1+1,\alpha_2),(\alpha_1,\alpha_2-1)\}.$$

Now, Lemma 3.7 yields the following equations for $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$: for $\beta'=(\alpha_1'+1,\alpha_2')$ we obtain

$$\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1+1,\alpha_2),(\alpha_1'+1,\alpha_2')} + \lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1,\alpha_2-1),(\alpha_1'+1,\alpha_2')} = 1,$$

$$(2\alpha_1+n)\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1+1,\alpha_2),(\alpha_1'+1,\alpha_2')} - (2\alpha_2+n-2)\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1,\alpha_2-1),(\alpha_1'+1,\alpha_2')} = 2\alpha_1'+n,$$

which gives

$$\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1+1,\alpha_2),(\alpha_1'+1,\alpha_2')} = \frac{\alpha_1'+\alpha_2+n-1}{\alpha_1+\alpha_2+n-1}, \quad \lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1,\alpha_2-1),(\alpha_1'+1,\alpha_2')} = \frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2+n-1},$$

for $\beta'=(\alpha_1'-1,\alpha_2')$ we get

$$\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1-1,\alpha_2),(\alpha_1'-1,\alpha_2')} + \lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1,\alpha_2+1),(\alpha_1'-1,\alpha_2')} = 1,$$

$$(2\alpha_1+n-2)\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1-1,\alpha_2),(\alpha_1'-1,\alpha_2')} - (2\alpha_2+n)\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1,\alpha_2+1),(\alpha_1'-1,\alpha_2')} = 2\alpha_1'+n-4,$$

implying

$$\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1-1,\alpha_2),(\alpha_1'-1,\alpha_2')} = \frac{\alpha_1'+\alpha_2+n-2}{\alpha_1+\alpha_2+n-1}, \quad \lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1,\alpha_2+1),(\alpha_1'-1,\alpha_2')} = \frac{\alpha_1-\alpha_1'+1}{\alpha_1+\alpha_2+n-1},$$

and similarly we find

$$\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1,\alpha_2+1),(\alpha_1',\alpha_2'+1)} = \frac{\alpha_1+\alpha_2'+n-1}{\alpha_1+\alpha_2+n-1}, \quad \lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1-1,\alpha_2),(\alpha_1',\alpha_2'+1)} = \frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2+n-1},$$

$$\lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1,\alpha_2-1),(\alpha_1',\alpha_2'-1)} = \frac{\alpha_1+\alpha_2'+n-2}{\alpha_1+\alpha_2+n-1}, \quad \lambda_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}^{(\alpha_1+1,\alpha_2),(\alpha_1',\alpha_2'-1)} = \frac{\alpha_2-\alpha_2'+1}{\alpha_1+\alpha_2+n-1}.$$

We remark that the constants $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$ can in this case also be obtained by computing the action of $\omega(X)$ on explicit K-finite vectors using (B-6) and recurrence relations for the Jacobi polynomials. With the explicit form of the constants $\lambda_{\alpha,\alpha'}^{\beta,\beta'}$ Corollary 3.6 now provides the following characterization of symmetry-breaking operators:

Theorem 5.1. *An operator $T:\mathcal{E}\to\mathcal{E}'$ is intertwining for π_r and $\tau_{r'}$ if and only if*

$$T|_{\mathcal{E}(\alpha;\alpha')} = \begin{cases} t_{\alpha,\alpha'}\cdot \operatorname{rest}|_{\mathcal{E}(\alpha;\alpha')} & \text{for } \alpha_1-\alpha_2=\alpha_1'-\alpha_2', \\ 0 & \text{else}, \end{cases}$$

with numbers $t_{\alpha,\alpha'}$ satisfying the following four relations:

$$\begin{aligned}(5\text{-}1)\quad (\alpha_1+\alpha_2+n-1)(r'+2\alpha_1'+n-1)t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')} \\ = (\alpha_1'+\alpha_2+n-1)(r+2\alpha_1+n)t_{(\alpha_1+1,\alpha_2),(\alpha_1'+1,\alpha_2')} \\ + (\alpha_1-\alpha_1')(r-2\alpha_2-n+2)t_{(\alpha_1,\alpha_2-1),(\alpha_1'+1,\alpha_2')},\end{aligned}$$

$$\begin{aligned}(5\text{-}2)\quad &(\alpha_1+\alpha_2+n-1)(r'-2\alpha_1'-n+3)t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}\\ &\quad=(\alpha_1'+\alpha_2+n-2)(r-2\alpha_1-n+2)t_{(\alpha_1-1,\alpha_2),(\alpha_1'-1,\alpha_2')}\\ &\qquad\qquad+(\alpha_1-\alpha_1'+1)(r+2\alpha_2+n)t_{(\alpha_1,\alpha_2+1),(\alpha_1'-1,\alpha_2')},\end{aligned}$$

$$\begin{aligned}(5\text{-}3)\quad &(\alpha_1+\alpha_2+n-1)(r'+2\alpha_2'+n-1)t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}\\ &\quad=(\alpha_1+\alpha_2'+n-1)(r+2\alpha_2+n)t_{(\alpha_1,\alpha_2+1),(\alpha_1',\alpha_2'+1)}\\ &\qquad\qquad+(\alpha_2-\alpha_2')(r-2\alpha_1-n+2)t_{(\alpha_1-1,\alpha_2),(\alpha_1',\alpha_2'+1)},\end{aligned}$$

$$\begin{aligned}(5\text{-}4)\quad &(\alpha_1+\alpha_2+n-1)(r'-2\alpha_2'-n+3)t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}\\ &\quad=(\alpha_1+\alpha_2'+n-2)(r-2\alpha_2-n+2)t_{(\alpha_1,\alpha_2-1),(\alpha_1',\alpha_2'-1)}\\ &\qquad\qquad+(\alpha_2-\alpha_2'+1)(r+2\alpha_1+n)t_{(\alpha_1+1,\alpha_2),(\alpha_1',\alpha_2'-1)}.\end{aligned}$$

5C. *Multiplicities.* The $(\mathfrak{g},K)$-module $(\pi_r)_{\mathrm{HC}}$ is reducible if and only if $r\in\pm(\rho+2\mathbb{N})$. More precisely, for $r=-\rho-2i$ the module $(\pi_r)_{\mathrm{HC}}$ contains a unique nontrivial finite-dimensional $(\mathfrak{g},K)$-submodule

$$\mathcal{F}(i)=\bigoplus_{\alpha_1,\alpha_2=0}^{i}\mathcal{E}(\alpha_1,\alpha_2)$$

as well as the two nontrivial infinite-dimensional submodules

$$\mathcal{F}_+(i)=\bigoplus_{\alpha_1=0}^{\infty}\bigoplus_{\alpha_2=0}^{i}\mathcal{E}(\alpha_1,\alpha_2),\qquad \mathcal{F}_-(i)=\bigoplus_{\alpha_1=0}^{i}\bigoplus_{\alpha_2=0}^{\infty}\mathcal{E}(\alpha_1,\alpha_2).$$

Then the composition series of $(\pi_r)_{\mathrm{HC}}$ is given by

$$\{0\}\subseteq\mathcal{F}(i)\subseteq\mathcal{F}_+(i)\subseteq(\mathcal{F}_+(i)+\mathcal{F}_-(i))\subseteq\mathcal{E}$$

(or equivalently with $\mathcal{F}_+$ and $\mathcal{F}_-$ switched). Hence the quotients

$$\mathcal{T}(i)=\mathcal{E}/(\mathcal{F}_+(i)+\mathcal{F}_-(i))\quad\text{and}\quad\mathcal{T}_\pm(i)=\mathcal{F}_\pm(i)/\mathcal{F}(i)$$

are irreducible and infinite-dimensional. Similarly we denote by $\mathcal{F}'(j)$, $\mathcal{F}'_\pm(j)$ and $\mathcal{T}'(j)$, $\mathcal{T}'_\pm(j)$ the corresponding composition factors of $(\tau_{r'})_{\mathrm{HC}}$ for $r'=-\rho'-2j$, $j\in\mathbb{N}$.

Define

$$L=\{(r,r')\in\mathbb{C}^2:r=-\rho-2i,\ r'=-\rho'-2j,\ 0\le j\le i\}.$$

Theorem 5.2. (1) *The multiplicities between spherical principal series of G and G' are given by*

$$m((\pi_r)_{\mathrm{HC}},(\tau_{r'})_{\mathrm{HC}})=\begin{cases}1 & \text{for } (r,r')\in\mathbb{C}^2\setminus L,\\ 2 & \text{for } (r,r')\in L.\end{cases}$$

(2) *For $i,j\in\mathbb{N}$ the multiplicities $m(\mathcal{V},\mathcal{W})$ between subquotients are given by*

$\mathcal{V}\downarrow\ \mathcal{W}\rightarrow$	$\mathcal{F}'(j)$	$\mathcal{T}'_+(j)$	$\mathcal{T}'_-(j)$	$\mathcal{T}'(j)$
$\mathcal{F}(i)$	1	0	0	0
$\mathcal{T}_+(i)$	0	1	0	0
$\mathcal{T}_-(i)$	0	0	1	0
$\mathcal{T}(i)$	0	0	0	1

for $j\le i$,

$\mathcal{V}\downarrow\ \mathcal{W}\rightarrow$	$\mathcal{F}'(j)$	$\mathcal{T}'_+(j)$	$\mathcal{T}'_-(j)$	$\mathcal{T}'(j)$
$\mathcal{F}(i)$	0	0	0	0
$\mathcal{T}_+(i)$	0	0	0	0
$\mathcal{T}_-(i)$	0	0	0	0
$\mathcal{T}(i)$	1	0	0	0

otherwise.

To prove Theorem 5.2 we proceed approximately as in Section 4C. For this we first reduce the four relations (5-1)–(5-4) in the four parameters $\alpha_1,\alpha_2,\alpha_1',\alpha_2'$ with $\alpha_1-\alpha_2=\alpha_1'-\alpha_2'$ to two pairs of two relations with only two parameters.

Put

$$p=\alpha_1+\alpha_2,\qquad q_1=\alpha_1',\qquad q_2=\alpha_2'.$$

Then

$$\alpha_1=\tfrac{p+q_1-q_2}{2},\qquad \alpha_2=\tfrac{p-q_1+q_2}{2},\qquad \alpha_1'=q_1,\qquad \alpha_2'=q_2.$$

Then $0\le\alpha_1'\le\alpha_1$, $0\le\alpha_2'\le\alpha_2$, and $\alpha_1-\alpha_2=\alpha_1'-\alpha_2'$ if and only if $p,q_1,q_2\in\mathbb{N}$ with $p-q_1-q_2\in 2\mathbb{N}$. With this reparametrization, the parameter q_2 is constant in the identities (5-1) and (5-2) and the parameter q_1 is constant in (5-3) and (5-4). Abusing notation and writing t_{p,q_1,q_2} for $t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}$ the relations (5-1)–(5-4) become

$$\begin{aligned}(p+n-1)(r'+2q_1+n-1)t_{p,q_1,q_2}&\\ =\big(\tfrac{p+q_1+q_2}{2}+n-1\big)&(r+p+q_1-q_2+n)t_{p+1,q_1+1,q_2}\\ &+\big(\tfrac{p-q_1-q_2}{2}\big)(r-p+q_1-q_2-n+2)t_{p-1,q_1+1,q_2},\end{aligned}\tag{5-5}$$

$$\begin{aligned}(p+n-1)(r'-2q_1-n+3)t_{p,q_1,q_2}&\\ =\big(\tfrac{p+q_1+q_2}{2}+n-2\big)&(r-p-q_1+q_2-n+2)t_{p-1,q_1-1,q_2}\\ &+\big(\tfrac{p-q_1-q_2}{2}+1\big)(r+p-q_1+q_2+n)t_{p+1,q_1-1,q_2},\end{aligned}\tag{5-6}$$

$$\begin{aligned}(p+n-1)(r'+2q_2+n-1)t_{p,q_1,q_2}&\\ =\big(\tfrac{p+q_1+q_2}{2}+n-1\big)&(r+p-q_1+q_2+n)t_{p+1,q_1,q_2+1}\\ &+\big(\tfrac{p-q_1-q_2}{2}\big)(r-p-q_1+q_2-n+2)t_{p-1,q_1,q_2+1},\end{aligned}\tag{5-7}$$

$$\begin{aligned}(p+n-1)(r'-2q_2-n+3)t_{p,q_1,q_2}&\\ =\big(\tfrac{p+q_1+q_2}{2}+n-2\big)&(r-p+q_1-q_2-n+2)t_{p-1,q_1,q_2-1}\\ &+\big(\tfrac{p-q_1-q_2}{2}+1\big)(r+p+q_1-q_2+n)t_{p+1,q_1,q_2-1}.\end{aligned}\tag{5-8}$$

Note that q_2 is fixed in (5-5) and (5-6), and these relations hold for $p, q_1 \in \mathbb{N}$ with $p - q_1 \in q_2 + 2\mathbb{N}$. The obvious similar statement holds for (5-7) and (5-8).

We first consider the diagonal $p = q_1 + q_2$; then relations (5-5) and (5-7) simplify to

$$(r' + 2q_1 + n - 1)t_{q_1+q_2,q_1,q_2} = (r + 2q_1 + n)t_{q_1+q_2+1,q_1+1,q_2}, \tag{5-9}$$

$$(r' + 2q_2 + n - 1)t_{q_1+q_2,q_1,q_2} = (r + 2q_2 + n)t_{q_1+q_2+1,q_1,q_2+1}. \tag{5-10}$$

This immediately yields:

Lemma 5.3. (1) *For* $(r, r') \in \mathbb{C}^2 \setminus L$ *the space of diagonal sequences* $(t_{q_1+q_2,q_1,q_2})_{q_1,q_2}$ *satisfying* (5-9) *and* (5-10) *has dimension* 1. *Any generator* $(t_{q_1+q_2,q_1,q_2})_{q_1,q_2}$ *satisfies*:

(a) *For* $r \notin -\rho - 2\mathbb{N},\ r' \notin -\rho' - 2\mathbb{N}$,

$$t_{q_1+q_2,q_1,q_2} \neq 0 \quad \textit{for all } q_1, q_2 \in \mathbb{N}.$$

(b) *For* $r = -\rho - 2i \in -\rho - 2\mathbb{N},\ r' \notin -\rho' - 2\mathbb{N}$,

$$t_{q_1+q_2,q_1,q_2} = 0 \quad \textit{for all } q_1 \le i \textit{ or } q_2 \le i \qquad \textit{and} \qquad t_{q_1+q_2,q_1,q_2} \neq 0 \quad \textit{for all } q_1, q_2 > i.$$

(c) *For* $r \notin -\rho - 2\mathbb{N},\ r' = -\rho' - 2j \in -\rho' - 2\mathbb{N}$,

$$t_{q_1+q_2,q_1,q_2} \neq 0 \quad \textit{for all } q_1, q_2 \le j \qquad \textit{and} \qquad t_{q_1+q_2,q_1,q_2} = 0 \quad \textit{for all } q_1 > j \textit{ or } q_2 > j.$$

(d) *For* $r = -\rho - 2i \in -\rho - 2\mathbb{N},\ r' = -\rho' - 2j \in -\rho' - 2\mathbb{N}$ *with* $i < j$,

$$t_{q_1+q_2,q_1,q_2} \neq 0 \quad \textit{for all } i < q_1, q_2 \le j \qquad \textit{and} \qquad t_{q_1+q_2,q_1,q_2} = 0 \quad \textit{else}.$$

(2) *For* $(r, r') = (-\rho - 2i, -\rho' - 2j) \in L$ *the space of diagonal sequences* $(t_{q_1+q_2,q_1,q_2})_{q_1,q_2}$ *satisfying* (5-9) *and* (5-10) *has dimension* 4.

Next we investigate how a diagonal sequence $(t_{q_1+q_2,q_1,q_2})_{q_1,q_2}$ satisfying (5-9) and (5-10) can be extended to a sequence $(t_{p,q_1,q_2})_{p,q_1,q_2}$ satisfying (5-5) and (5-6) and the corresponding relations in q_2. For this note that if we fix, say, q_2, and put $p' = p - q_2$, then the relations (5-5) and (5-6) read

$$\begin{aligned}(p' + q_2 + n - 1)(r' + 2q_1 + n - 1)t_{p',q_1} &= \left(\tfrac{p'+q_1}{2} + q_2 + n - 1\right)(r + p' + q_1 + n)t_{p'+1,q_1+1} \\ &\quad + \left(\tfrac{p'-q_1}{2}\right)(r - p' + q_1 - 2q_2 - n + 2)t_{p'-1,q_1+1},\end{aligned} \tag{5-11}$$

$$\begin{aligned}(p' + q_2 + n - 1)(r' - 2q_1 - n + 3)t_{p',q_1} &= \left(\tfrac{p'+q_1}{2} + q_2 + n - 2\right)(r - p' - q_1 - n + 2)t_{p'-1,q_1-1} \\ &\quad + \left(\tfrac{p'-q_1}{2} + 1\right)(r + p' - q_1 + 2q_2 + n)t_{p'+1,q_1-1},\end{aligned} \tag{5-12}$$

where we again abuse notation and write t_{p',q_1} for t_{p,q_1,q_2}. Similar relations hold if q_1 is fixed. We note that (5-11) and (5-12) have to be satisfied for all $p', q_1 \in \mathbb{N}$ with

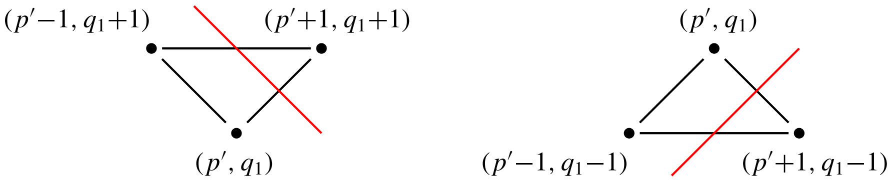


Figure 10. Barriers for $r = -\rho - 2i$.

$p' - q_1 \in 2\mathbb{N}$, just as in the case of orthogonal groups, see Figure 1. Thus, many arguments used in the orthogonal situation can be translated to this context. There are, however, differences to the orthogonal situation. If $r = -\rho - 2i \in -\rho - 2\mathbb{N}$ then the coefficient $(r + p' + q_1 + n)$ in (5-11) vanishes for $p' + q_1 = 2i$ and the coefficient $(r + p' - q_1 + 2q_2 + n)$ in (5-12) vanishes for $p' - q_1 = 2(i - q_2)$, which we indicate by diagonal lines as in Figure 10. Further, if $r' = -\rho' - 2j \in -\rho' - 2\mathbb{N}$ then the coefficient $(r' + 2q_1 + n - 1)$ in (5-11) vanishes for $q_1 = j$, which we indicate by a vertical line as in Figure 11.

Lemma 5.4. *Let $(r, r') \in \mathbb{C}^2 \setminus L$. Then every diagonal sequence $(t_{q_1+q_2,q_1,q_2})_{q_1,q_2}$ satisfying (5-9) and (5-10) has a unique extension to a sequence $(t_{p,q_1,q_2})_{p,q_1,q_2}$ satisfying (5-5)–(5-8).*

Proof. The proof is similar to the proof of Lemma 4.4 and we only indicate the relevant steps.

Step 1. We first treat the case $r \notin -\rho - 2\mathbb{N}$. We fix q_2; then the diagonal sequence determines t_{p',q_1} for $p' = q_1$. Since $r \notin -\rho - 2\mathbb{N}$ the coefficient $(r + p' - q_1 + 2q_2 + n)$ in (5-12) never vanishes. Hence, (5-12) can be used to express $t_{p'+1,q_1-1}$ in terms of t_{p',q_1} and $t_{p'-1,q_1-1}$. As in the proof of Lemma 4.4, Step 1, this uniquely determines all numbers t_{p',q_1}. Since q_2 was arbitrary this determines all numbers t_{p,q_1,q_2}.

Step 2. Next assume $r = -\rho - 2i \in -\rho - 2\mathbb{N}$ and $r' \notin -\rho' - 2\mathbb{N}$. Then the coefficient $(r + p' - q_1 + 2q_2 + n)$ vanishes if and only if $p' - q_1 = 2(i - q_2)$. In particular, it does not vanish for $q_2 > i$. We can therefore use the technique in Step 1 to extend the diagonal sequence to t_{p,q_1,q_2} for $q_2 > i$ and all p, q_1. Fixing q_1 instead of q_2 we are in the situation that t_{p',q_2} is given on the diagonal $p' = q_2$ and in the region $q_2 > i$. Since $r' \notin -\rho' - 2\mathbb{N}$ the coefficient $(r' + 2q_2 + n - 1)$ in (5-11) (with q_1 and

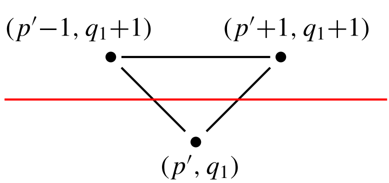


Figure 11. Barrier for $r' = -\rho' - 2j$.

q_2 interchanged) never vanishes, so we can use (5-11) (with q_1 and q_2 interchanged) to extend t_{p',q_2} to all p', q_2 as in the proof of Lemma 4.4 Step 2. Since q_1 was arbitrary this determines all numbers t_{p,q_1,q_2}.

Step 3. Now let $r = -\rho - 2i \in -\rho - 2\mathbb{N}$ and $r' = -\rho' - 2j \in -\rho' - 2\mathbb{N}$, with $i, j \in \mathbb{N}$, $j > i$. Note that to carry out Step 2 we only need that $r' + 2q_2 + n - 1 \neq 0$ for $q_2 \leq i$. This is satisfied since

$$r' + 2q_2 + n - 1 = 2(q_2 - j) < 2(q_2 - i) \leq 0$$

by assumption. Hence the technique in Step 2 carries over to this case. □

The case $(r, r') \in L$ has to be handled a little differently from the orthogonal situation.

Lemma 5.5. *Let $(r, r') = (-\rho - 2i, -\rho' - 2j) \in L$. Then every choice of $t_{0,0,0}$ and $t_{2i+2,0,0}$ determines a unique sequence $(t_{p,q_1,q_2})_{p,q_1,q_2}$ satisfying (5-5)–(5-8).*

Proof. Fix $q_2 = 0$, $p' = p - q_2 = p$; then by the assumption t_{p',q_1} is known for $(p', q_1) = (0, 0)$ and $(2i+2, 0)$. This is illustrated in Figure 12, where the barriers are as in Figures 10 and 11. Then the techniques from the proof of Lemma 5.4 extend $t_{0,0}$ uniquely to the region $p' + q_1 \leq 2i$; see also Figure 12. To overcome the barrier given by $p' + q_1 = 2i$ we use (5-12) for $p' - q_1 = 2i$ in which the coefficient $(r + p' - q_1 + 2q_2 + n)$ vanishes. Hence, this relation can be applied to extend along the diagonal line $p' - q_1 = 2i$ as indicated in Figure 12. It may also be applied anywhere above the diagonal $p' - q_1 = 2i$ so that we actually extend to the area $p' - q_1 \leq 2i$; see Figure 13. Next we need to overcome the barrier $p' - q_1 = 2i$, which we do by using (5-11) for $q_1 = j$. In this relation the coefficient $(r' + 2q_1 + n - 1)$ vanishes, and hence we can extend along the line $q_1 = j + 1$. Using again (5-12) even extends to the whole region

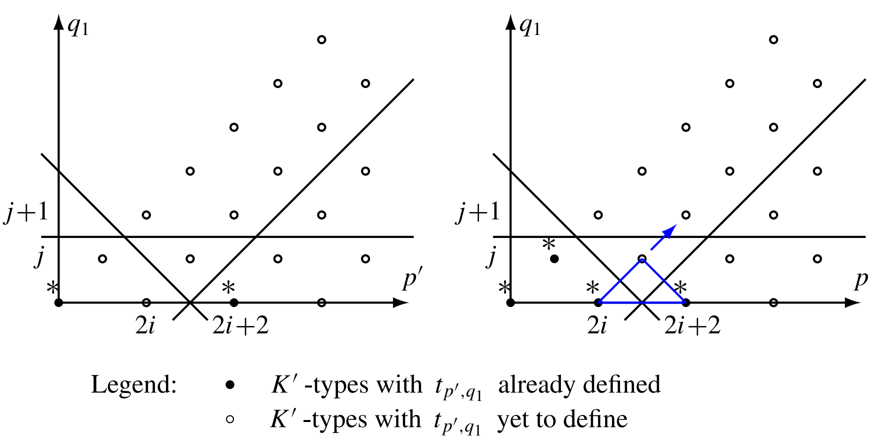
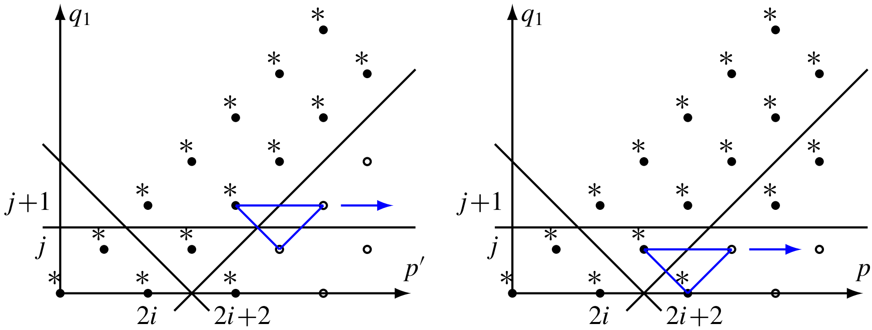


Figure 12



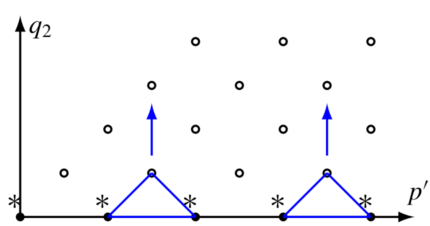
Legend: • K'-types with t_{p',q_1} already defined
∘ K'-types with t_{p',q_1} yet to define

Figure 13

$q_1 > j$; see Figure 13. We note that up to this point we have not yet made use of $t_{2i+2,0}$. This is needed now to extend into the region $\{(p', q_1) : p' - q_1 > 2i,\ q_1 \leq j\}$; see Figure 13. Here both relations (5-11) and (5-12) are needed. Summarizing, we have extended $t_{0,0,0}$ and $t_{2i+2,0,0}$ uniquely to a sequence $(t_{p,q_1,0})_{p,q_1}$. Next fix q_1 and let $p' = p - q_1$. Then t_{p',q_2} is already determined for $(p', q_2) = (p', 0)$ with p' arbitrary; see Figure 14. Note that in relation (5-12) (with q_1 and q_2 interchanged) the coefficient $(r' - 2q_2 - n + 3)$ never vanishes, and hence this relation can be used to extend $(t_{p',0})_{p'}$ uniquely to $(t_{p',q_2})_{p',q_2}$; see Figure 14. Since q_1 was arbitrary this finally yields t_{p,q_1,q_2} for any p, q_1, q_2 and finishes the proof. □

Proof of Theorem 5.2. (1) This statement is contained in Lemmas 5.3, 5.4, and 5.5.

(2) Composing with embeddings and quotient maps most of the multiplicity statements can be reduced to statements about the (non-)existence of intertwining operators $T : (\pi_r)_{\mathrm{HC}} \to (\tau_{r'})_{\mathrm{HC}}$ for particular r and r' such that the numbers



Legend: • K'-types with t_{p',q_1} already defined
∘ K'-types with t_{p',q_1} yet to define

Figure 14

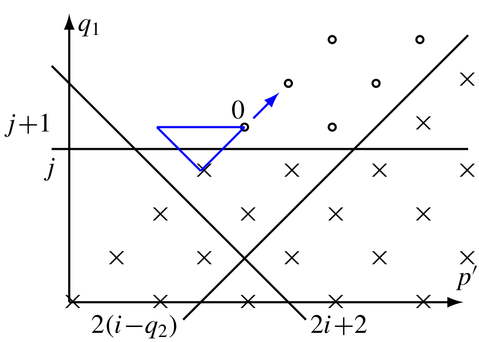
$t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}$ vanish in certain regions. These statements can be checked using the techniques used in Lemmas 5.3, 5.4 and 5.5. This does not work if either $\mathcal{V}=\mathcal{T}_\pm(i)$ or $\mathcal{W}=\mathcal{T}_\pm'(j)$. We therefore show the multiplicity statements for $m(\mathcal{T}_+(i),\mathcal{T}_+(j))$ in detail, using Remark 3.5. Similar considerations can then be applied to the remaining cases.

Let first $\mathcal{V}=\mathcal{T}_+(i)$ and $\mathcal{W}=\mathcal{T}_+'(j)$. Then, due to Remark 3.5, an intertwining operator $\mathcal{T}_+(i)\to\mathcal{T}_+(j)$ is given by an operator

$$T:\mathcal{F}_+(i)\to\bigoplus_{\alpha_1'=j+1}^{\infty}\bigoplus_{\alpha_2'=0}^{j}\mathcal{E}'(\alpha_1',\alpha_2'),\qquad T|_{\mathcal{E}(\alpha;\alpha')}=t_{\alpha,\alpha'}\cdot R_{\alpha,\alpha'}$$

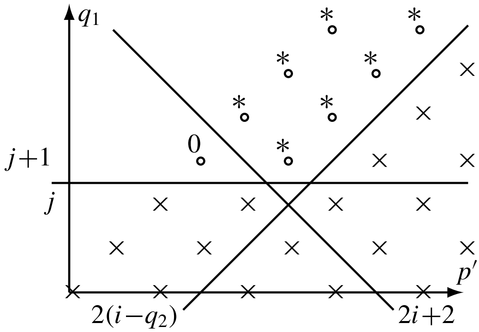
such that $T|_{\mathcal{F}(i)}=0$, and the numbers $t_{\alpha,\alpha'}$ solve the relations (5-1)–(5-4) whenever the two terms $t_{(\beta_1,\beta_2),(\beta_1',\beta_2')}$ on the right-hand sides of (5-1)–(5-4) satisfy $\beta_1'>j$, $\beta_2'\le j$ (i.e., the two upper, resp. lower, vertices of the corresponding triangles are contained in the region $\{(\beta_1',\beta_2'):\beta_1'>j,\ \beta_2'\le j\}$).

Assume first that $j>i$. Then for any fixed $q_2\le j$ and $p'=p-q_2$ we are looking for numbers t_{p',q_1} which vanish if either $q_1\le j$ (i.e., $\alpha_1'\le j$, the region below the horizontal line in Figure 15) or $p'-q_1>2(i-q_2)$ (i.e., $\alpha_2>i$, the region below the diagonal line going into the upper right corner in Figure 15). As indicated in Figure 15, relation (5-11) can be used along the diagonal to obtain $t_{q_1,q_1}=0$ for $q_1>j$. Then using (5-12) yields $t_{p',q_1}=0$ for all p', q_1, so that $m(\mathcal{T}_+(i),\mathcal{T}_+'(j))=0$. Next assume $j\le i$. Then for fixed $q_2\le j$ and $p'=p-q_2$ we have to find numbers as indicated in Figure 16. Here the relations (5-11) and (5-12) don't force any of the numbers in the region $\{(p',q_1):p'-q_1\le 2(i-q_2),\ p'+q_1>2i\}$ to vanish and hence the choice of one t_{p',q_1} determines the remaining numbers. We note that in this case $t_{p',q_1}=0$ for $p'+q_1\le 2i$ and $q_1>j$ as desired. Similarly, if we fix



Legend: ∘ K'-types with t_{p',q_1} to be determined
× K'-types with $t_{p',q_1}=0$ by formal reasons

Figure 15

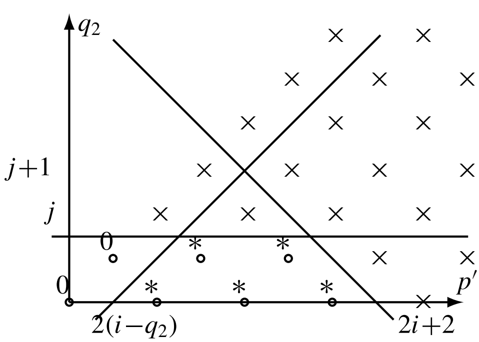


Legend: $\circ$ K'-types with t_{p',q_1} to be determined
$\times$ K'-types with $t_{p',q_1}=0$ by formal reasons

Figure 16

$q_1 > j$ and let $p' = p - q_1$ we are in the situation of Figure 17. More precisely, we need to find numbers t_{p',q_2} satisfying the relations (5-11) and (5-12) (with q_1 and q_2 interchanged) in the region $\{(p', q_2) : q_2 \le j,\ p' + q_2 \le 2i\}$ such that $t_{p',q_2} = 0$ for $p' - q_2 \le 2i$. Again the relations do not force any number in the nontrivial region to vanish (indicated by stars in Figure 17). Within this region, the choice of one of the numbers uniquely determines the rest. Together with the previous observation for the case of $q_2 \le j$ fixed we obtain $m(\mathcal{T}_+(i), \mathcal{T}_+(j)) = 1$. □

5D. ***Explicit formula for the spectral function.*** As in Section 4D we also find the generic solution to the relations (5-1)–(5-4) as a meromorphic function in $r, r' \in \mathbb{C}$.



Legend: $\circ$ K'-types with t_{p',q_1} to be determined
$\times$ K'-types with $t_{p',q_1}=0$ by formal reasons

Figure 17

Proposition 5.6. *For $\alpha_1, \alpha_2 \in \mathbb{N}$ and $0 \le \alpha_1' \le \alpha_1,\ 0 \le \alpha_2' \le \alpha_2$, with $\alpha_1 - \alpha_2 = \alpha_1' - \alpha_2'$, the numbers*

$$t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r, r') = \sum_{k=0}^{\infty} \frac{2^k \Gamma\bigl(\frac{\alpha_1+\alpha_2-\alpha_1'-\alpha_2'+2}{2}\bigr)\Gamma\bigl(\frac{\alpha_1+\alpha_2+\alpha_1'+\alpha_2'}{2}+n-1+k\bigr)}{(k!)^2\Gamma\bigl(\frac{\alpha_1+\alpha_2-\alpha_1'-\alpha_2'+2}{2}-k\bigr)\Gamma\bigl(\frac{\alpha_1+\alpha_2+\alpha_1'+\alpha_2'}{2}+n-1\bigr)} \times \frac{\Gamma\bigl(\frac{r+n}{2}\bigr)^2\Gamma\bigl(\frac{r'+n-1}{2}+\alpha_1'\bigr)\Gamma\bigl(\frac{r'+n-1}{2}+\alpha_2'\bigr)\Gamma\bigl(\frac{r'-r+1}{2}\bigr)\Gamma\bigl(\frac{r'+r+1}{2}+k\bigr)}{\Gamma\bigl(\frac{r'+n-1}{2}\bigr)^2\Gamma\bigl(\frac{r+n}{2}+\alpha_1'+k\bigr)\Gamma\bigl(\frac{r+n}{2}+\alpha_2'+k\bigr)\Gamma\bigl(\frac{r'+r+1}{2}\bigr)\Gamma\bigl(\frac{r'-r+1}{2}-k\bigr)}$$

are rational functions in r and r' satisfying the relations (5-1)–(5-4). They are normalized to $t_{(0,0),(0,0)} \equiv 1$.

Proof. The proof is similar to the proof of Proposition 4.6 and we omit some of the details. For simplicity we use the reparametrization (p, q_1, q_2) instead of (α_1, α_2) and (α_1', α_2'). Fix q_2 and let $p' = p - q_2$; then it is easy to see that for every $k \in \mathbb{N}$ the expression

$$\frac{\Gamma\bigl(\frac{p'-q_1+2}{2}\bigr)\Gamma\bigl(\frac{p'+q_1}{2}+q_2+n-1+k\bigr)\Gamma\bigl(\frac{r'+2q_1+n-1}{2}\bigr)}{\Gamma\bigl(\frac{p'-q_1+2}{2}-k\bigr)\Gamma\bigl(\frac{p'+q_1}{2}+q_2+n-1\bigr)\Gamma\bigl(\frac{r+2q_1+n}{2}+k\bigr)}$$

satisfies (5-11). Further, the series

$$\sum_{k=0}^{\infty} b_k \frac{\Gamma\bigl(\frac{p'-q_1+2}{2}\bigr)\Gamma\bigl(\frac{p'+q_1}{2}+q_2+n-1+k\bigr)\Gamma\bigl(\frac{r'+2q_1+n-1}{2}\bigr)}{\Gamma\bigl(\frac{p'-q_1+2}{2}-k\bigr)\Gamma\bigl(\frac{p'+q_1}{2}+q_2+n-1\bigr)\Gamma\bigl(\frac{r+2q_1+n}{2}+k\bigr)}$$

satisfies (5-12) if and only if

$$b_k = c\frac{2^k\Gamma\bigl(\frac{r'+r+1}{2}+k\bigr)}{(k!)^2\Gamma\bigl(\frac{r+2q_2+n}{2}+k\bigr)\Gamma\bigl(\frac{r'-r+1}{2}-k\bigr)}$$

for some constant $c = c(r, r', q_2)$ which does not depend on p', q_1 and k. Plugging in $p' = p - q_2$, using the symmetry of the relations (5-5)–(5-8) in q_1 and q_2, and normalizing to $t_{0,0,0} \equiv 0$ yields

$$t_{p,q_1,q_2}(r,r') = \sum_{k=0}^{\infty} \frac{2^k\Gamma\bigl(\frac{p-q_1-q_2+2}{2}\bigr)\Gamma\bigl(\frac{p+q_1+q_2}{2}+n-1+k\bigr)}{(k!)^2\Gamma\bigl(\frac{p-q_1-q_2+2}{2}-k\bigr)\Gamma\bigl(\frac{p+q_1+q_2}{2}+n-1\bigr)} \times \frac{\Gamma\bigl(\frac{r+n}{2}\bigr)^2\Gamma\bigl(\frac{r'+n-1}{2}+q_1\bigr)\Gamma\bigl(\frac{r'+n-1}{2}+q_2\bigr)\Gamma\bigl(\frac{r'-r+1}{2}\bigr)\Gamma\bigl(\frac{r'+r+1}{2}+k\bigr)}{\Gamma\bigl(\frac{r'+n-1}{2}\bigr)^2\Gamma\bigl(\frac{r+n}{2}+q_1+k\bigr)\Gamma\bigl(\frac{r+n}{2}+q_2+k\bigr)\Gamma\bigl(\frac{r'+r+1}{2}\bigr)\Gamma\bigl(\frac{r'-r+1}{2}-k\bigr)} \tag{5-13}$$

Reparametrizing p, q_1, q_2 to $\alpha_1, \alpha_2, \alpha_1', \alpha_2'$ shows the claimed formula. Rewriting (5-13) as

(5-14) $$t_{p,q_1,q_2}(r,r') = \sum_{k=0}^{\frac{p-q_1-q_2}{2}} \frac{1}{(k!)^2\left(\frac{r+n}{2}\right)_{q_1+k}\left(\frac{r+n}{2}\right)_{q_2+k}} 2^k\left(-\tfrac{p-q_1-q_2}{2}\right)_k\left(\tfrac{p+q_1+q_2}{2}+n-1\right)_k \times \left(\tfrac{r'+n-1}{2}\right)_{q_1}\left(\tfrac{r'+n-1}{2}\right)_{q_2}\left(\tfrac{r-r'+1}{2}\right)_k\left(\tfrac{r'+r+1}{2}\right)_k$$

further shows that this is a rational function in r and r'. □

Also the next two results are proven along the same lines as Corollary 4.7 and Theorem 4.9.

Corollary 5.7. (1) *The renormalized numbers*

$$t^{(1)}_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r') = \frac{1}{\Gamma\left(\frac{r+\rho}{2}\right)^2} t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r')$$

are holomorphic in $(r,r') \in \mathbb{C}^2$ *for all* (α_1,α_2), (α_1',α_2'). *Further,*

$$t^{(1)}_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r') = 0$$

for all (α_1,α_2), (α_1',α_2') *if and only if* $(r,r') \in L$.

(2) *Fix* $r' = -\rho' - 2j,\ j \in \mathbb{N}$; *then the renormalized numbers*

$$t^{(2)}_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r') = \frac{\Gamma\left(\frac{(r+\rho)-(r'+\rho')}{2}\right)}{\Gamma\left(\frac{r+\rho}{2}\right)^2} t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r')$$

are holomorphic in $r \in \mathbb{C}$ *for all* (α_1,α_2), (α_1',α_2'). *We have* $t^{(2)}_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r') \equiv 0$ *whenever* $\alpha_1' > j$ *or* $\alpha_2' > j$. *Further, for every* $r \in \mathbb{C}$ *there exist* (α_1,α_2), (α_1',α_2') *with* $t^{(2)}_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r') \neq 0$.

(3) *Fix* $N \in \mathbb{N}$ *and let* $r'+\rho' = r+\rho+2N$; *then the renormalized numbers*

$$t^{(3)}_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r') = \frac{\Gamma\left(\frac{r'+\rho'}{2}\right)^2}{\Gamma\left(\frac{r+\rho}{2}\right)^2} t_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r')$$

are holomorphic in $r \in \mathbb{C}$ *for all* (α_1,α_2), (α_1',α_2'). *Further, for every* $r \in \mathbb{C}$ *there exists* $\alpha_0 \in \mathbb{N}$ *such that* $t^{(3)}_{(\alpha_1,\alpha_2),(\alpha_1,\alpha_2)}(r,r') \neq 0$ *for* $\alpha_1, \alpha_2 \geq \alpha_0$.

Theorem 5.8. *For* $i = 1,2,3$ *we let* $T^{(i)}(r,r')$ *be the intertwining operators* $(\pi_r)_{\mathrm{HC}} \to (\tau_{r'})_{\mathrm{HC}}$ *corresponding to the numbers* $t^{(i)}_{(\alpha_1,\alpha_2),(\alpha_1',\alpha_2')}(r,r')$ *in Corollary 5.7. Then the operator* $T^{(1)}(r,r')$ *is defined for* $(r,r') \in \mathbb{C}^2$, *the operator* $T^{(2)}(r,r')$ *is* *defined for* $r' \in -\rho' - 2\mathbb{N}$ *and the operator* $T^{(3)}(r,r')$ *is defined for* $(r+\rho) - (r'+\rho') \in -2\mathbb{N}$. *We have*

$$\operatorname{Hom}_{(\mathfrak{g}',K')}((\pi_r)_{\mathrm{HC}}|_{(\mathfrak{g}',K')}, (\tau_{r'})_{\mathrm{HC}}) = \begin{cases} \mathbb{C}T^{(1)}(r,r') & \text{for } (r,r') \in \mathbb{C}^2 \setminus L, \\ \mathbb{C}T^{(2)}(r,r') \oplus \mathbb{C}T^{(3)}(r,r') & \text{for } (r,r') \in L. \end{cases}$$

Remark 5.9. We remark that also every intertwining operator between subquotients $\mathcal{V} = \mathcal{F}(i), \mathcal{T}_\pm(i), \mathcal{T}(i)$ and $\mathcal{W} = \mathcal{F}'(j), \mathcal{T}'_\pm(j), \mathcal{T}'(j)$ can be obtained from the holomorphic family $T^{(i)}(r,r')$ by restricting and renormalizing. More precisely, if $\mathcal{V}$ is a quotient of $(\pi_r)_{\mathrm{HC}}$ and $\mathcal{W}$ is a subrepresentation of $(\tau_{r'})_{\mathrm{HC}}$ then any intertwining operator $T : \mathcal{V} \to \mathcal{W}$ gives rise to an intertwining operator $(\pi_r)_{\mathrm{HC}} \to (\tau_{r'})_{\mathrm{HC}}$ and is hence of the form $T^{(i)}(r,r')$ for some $i = 1, 2, 3$. This constructs all except the intertwiners $\mathcal{T}_\pm(i) \to \mathcal{T}'_\pm(j)$ for $0 \le j \le i$. These can be obtained from $T^{(1)}(r,r')$ as follows:

We first construct an intertwining operator $T^+ : \mathcal{F}_+(i) \to (\tau_{r'})_{\mathrm{HC}}$ for $r' = -\rho' - 2j$ such that $T^+(\mathcal{F}_+(i)) \subseteq \mathcal{F}_+(j)$. Since $\mathcal{F}_+(i)$ consists of all K-type $\mathcal{E}(\alpha_1, \alpha_2)$ with $\alpha_2 \le i$ it is given by a sequence $(t^+_{(\alpha_1,\alpha_2),(\alpha'_1,\alpha'_2)})_{\alpha_2 \le i}$. Reparametrizing to p, q_1, q_2 this means that we have to find a sequence $(t^+_{p,q_1,q_2})_{p-q_1+q_2 \le 2i}$ satisfying the necessary relations. Let $r' + \rho' = r + \rho + 2N$, with $N = i - j \in \mathbb{N}$, and define

$$t^+_{p,q_1,q_2}(r,r') := \frac{\Gamma\left(\frac{r'+\rho'}{2}\right)}{\Gamma\left(\frac{r+\rho}{2}\right)} t_{p,q_1,q_2}(r,r'), \qquad p - q_1 + q_2 \le 2i.$$

Then by (5-14) we have

$$t^+_{p,q_1,q_2}(r,r') = \sum_k \frac{1}{(k!)^2 \left(\frac{r+n}{2}\right)_{q_2+k}} 2^k \left(-\tfrac{p-q_1-q_2}{2}\right)_k \left(\tfrac{p+q_1+q_2}{2} + n - 1\right)_k \times \left(\tfrac{r+n}{2} + k + q_1\right)_{N-k} \left(\tfrac{r+n}{2} + N\right)_{q_2} (-N)_k (r+N+1)_k.$$

In the sum all terms for $k > \frac{p-q_1-q_2}{2}$ vanish, so that $k \le \frac{p-q_1-q_2}{2} = \frac{p-q_1+q_2}{2} - q_2 \le i - q_2$. This implies that the denominator does not vanish at $r = -\rho - 2i$. Therefore $t^+_{p,q_1,q_2}(r,r')$ is holomorphic in $r = -\rho - 2i$ and evaluation there yields

$$\begin{aligned} t^+_{p,q_1,q_2} &= t^+_{p,q_1,q_2}(-\rho - 2i, -\rho' - 2j) \\ &= \sum_{k=0}^{i-j} \frac{1}{(k!)^2 (-i)_{q_2+k}} 2^k \left(-\tfrac{p-q_1-q_2}{2}\right)_k \\ &\qquad \times \left(\tfrac{p+q_1+q_2}{2} + n - 1\right)_k (k + q_1 - i)_{i-j-k} (-j)_{q_2} (j-i)_k (1-n-i-j)_k. \end{aligned}$$

The sequence t^+_{p,q_1,q_2} clearly satisfies the necessary relations since it is simply a renormalization of the sequence t_{p,q_1,q_2}, and hence it defines an intertwining operator $T^+ : \mathcal{F}_+(i) \to (\tau_{r'})_{\mathrm{HC}}$. We note that for $q_2 > j$ the term $(-j)_{q_2}$ vanishes so that $t^+_{p,q_1,q_2}=0$. Therefore $T^+(\mathcal{F}_+(i))\subseteq\mathcal{F}'_+(j)$. Composing with the quotient map $\mathcal{F}'_+(j)\to\mathcal{T}'_+(j)$ yields an intertwiner $T^+:\mathcal{F}_+(i)\to\mathcal{T}'_+(j)$. We claim that this intertwiner vanishes on $\mathcal{F}(i)$ and hence factorizes through $\mathcal{T}_+(i)$. In fact, for $\alpha_2=\frac{p+q_1-q_2}{2}\le i$ and $q_1>j$ we have $\frac{p-q_1-q_2}{2}\le i-q_1$ so that we may take the sum over all $k\le i-q_1$. But then $q_1+k-i\le 0$ and therefore $(q_1+k-i)_{i-j-k}=0$, whence $t^+_{p,q_1,q_2}=0$. This implies that $T^+:\mathcal{F}_+(i)\to\mathcal{T}'_+(j)$ factorizes to an intertwiner $T^+:\mathcal{T}_+(i)\to\mathcal{T}_+(j)$. To finally see that this intertwiner is nontrivial we note that for all $q_1>i,\ q_2\le j$, and $p=q_1+q_2$ we have

$$t^+_{p,q_1,q_2}=\frac{(q_1-i)_{i-j}(-j)_{q_2}}{(-i)_{q_2}}\neq 0.$$

Remark 5.10. The operators $T^{(1)}(r,r')$ are related to the meromorphic family of singular integral operators constructed in [Möllers et al. 2016a]. Further, the family $T^{(3)}$ is (up to a constant) equal to the differential restriction operators on the Heisenberg group constructed in [Möllers et al. 2016b]. They can be viewed as a generalization of Juhl's conformally invariant operators (see Remark 4.11). It would be interesting to carry out a detailed investigation of all operators $T^{(i)}(r,r')$, $i=1,2,3$, in the noncompact picture as in [Kobayashi and Speh 2015].

As in the real case, we can prove automatic continuity using the full classification in Theorem 5.8 in terms of the holomorphic family $T^{(1)}(r,r')$. Note that the corresponding holomorphic family of intertwining operators in the smooth category was also constructed in [Möllers et al. 2016a].

Corollary 5.11. *For $(G,G')=(\mathrm{U}(1,n),\mathrm{U}(1,n-1))$ the natural injective map*

$$\mathrm{Hom}_{G'}(\pi|_{G'},\tau)\to\mathrm{Hom}_{(\mathfrak{g}',K')}(\pi_{\mathrm{HC}}|_{(\mathfrak{g}',K')},\tau_{\mathrm{HC}})$$

is an isomorphism for all spherical principal series π of G and τ of G' and their subquotients.

Appendix A: Orthogonal polynomials

Gegenbauer polynomials. The classical Gegenbauer polynomials $C_n^\lambda(z)$ can be defined by, see [Erdélyi et al. 1953, 10.9, equation (18)],

$$C_n^\lambda(z)=\sum_{m=0}^{\lfloor\frac{n}{2}\rfloor}\frac{(-1)^m(\lambda)_{n-m}}{m!\,(n-2m)!}(2z)^{n-2m}.$$

They obviously satisfy the parity condition, see [Erdélyi et al. 1953, 10.9, equation (16)],

$$C_n^\lambda(-z)=(-1)^nC_n^\lambda(z). \tag{A-1}$$

The special value at $z=0$ can be written as

(A-2) $$C_n^\lambda(0)=\frac{2^n\sqrt{\pi}\,\Gamma\left(\lambda+\frac{n}{2}\right)}{n!\,\Gamma\left(\frac{1-n}{2}\right)\Gamma(\lambda)} \overset{(n=2k)}{=} \frac{(-1)^k\Gamma(\lambda+k)}{k!\,\Gamma(\lambda)}.$$

Jacobi polynomials. The classical Jacobi polynomials $P_n^{(\alpha,\beta)}(z)$ can be defined by, see [Erdélyi et al. 1953, 10.8, equation (12)],

$$P_n^{(\alpha,\beta)}(z)=2^{-n}\sum_{m=0}^{n}\binom{n+\alpha}{m}\binom{n+\beta}{n-m}(x-1)^{n-m}(x+1)^m.$$

The special value at $z=1$ is given by

(A-3) $$P_n^{(\alpha,\beta)}(1)=\binom{n+\alpha}{n}.$$

Appendix B: Spherical harmonics

Real spherical harmonics. Let $\mathcal{H}^\alpha(\mathbb{R}^n)$ denote the space of harmonic homogeneous polynomials of degree α on $\mathbb{R}^n$. Endowed with the natural action of $\mathrm{O}(n)$, the space $\mathcal{H}^\alpha(\mathbb{R}^n)$ is an irreducible representation. It is unitary with respect to the norm on $\mathcal{H}^\alpha(\mathbb{R}^n)$ given by

$$\|\phi\|^2_{L^2(S^{n-1})}=\int_{S^{n-1}}|\phi(x)|^2\,dx,$$

where dx denotes the Euclidean measure on S^{n-1}. Upon restriction to the subgroup $\mathrm{O}(n-1)$ the representation $\mathcal{H}^\alpha(\mathbb{R}^n)$ decomposes into

(B-1) $$\mathcal{H}^\alpha(\mathbb{R}^n)\simeq\bigoplus_{0\le\alpha'\le\alpha}\mathcal{H}^{\alpha'}(\mathbb{R}^{n-1}).$$

Explicit $\mathrm{O}(n-1)$-equivariant embeddings of the direct summands are given by, see [Kobayashi and Mano 2011, Fact 7.5.1],

(B-2) $$I^n_{\alpha'\to\alpha}:\mathcal{H}^{\alpha'}(\mathbb{R}^{n-1})\to\mathcal{H}^\alpha(\mathbb{R}^n),\qquad I^n_{\alpha'\to\alpha}(\phi)(x',x_n)=\phi(x')C^{\frac{n-2}{2}+\alpha'}_{\alpha-\alpha'}(x_n),$$

where $x=(x',x_n)\in S^{n-1}$. The following Plancherel formula holds for $\phi\in\mathcal{H}^{\alpha'}(\mathbb{R}^{n-1})$ (see [Kobayashi and Mano 2011, Fact 7.5.1(3)], note the different normalization of the Gegenbauer polynomials):

(B-3) $$\|I^n_{\alpha'\to\alpha}(\phi)\|^2_{L^2(S^{n-1})}=\frac{2^{3-n-2\alpha'}\pi\,\Gamma(n-2+\alpha+\alpha')}{(\alpha-\alpha')!\,\left(\alpha+\frac{n-2}{2}\right)\Gamma\left(\alpha'+\frac{n-2}{2}\right)^2}\|\phi\|^2_{L^2(S^{n-2})}.$$

For $\phi\in\mathcal{H}^\alpha(\mathbb{R}^n)$ we have

(B-4) $$x_j\phi=\phi_j^+ + |x|^2\phi_j^-,$$

with $\phi_j^\pm \in \mathcal{H}^{\alpha\pm1}(\mathbb{R}^n)$ given by

$$\phi_j^+ = x_j\phi - \frac{|x|^2}{n+2\alpha-2}\frac{\partial\phi}{\partial x_j}, \quad \phi_j^- = \frac{1}{n+2\alpha-2}\frac{\partial\phi}{\partial x_j}.$$

Complex spherical harmonics. Identifying $\mathbb{R}^{2n} \simeq \mathbb{C}^n$ we embed $\mathrm{U}(n)$ into $\mathrm{O}(2n)$. Then the restriction of the irreducible representation $\mathcal{H}^\alpha(\mathbb{R}^{2n})$ of $\mathrm{O}(2n)$ to the subgroup $\mathrm{U}(n)$ decomposes into

$$\mathcal{H}^\alpha(\mathbb{R}^{2n}) = \bigoplus_{\alpha_1+\alpha_2=\alpha} \mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n),$$

where $\mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n)$ denotes the space of harmonic polynomials on $\mathbb{C}^n$ which are holomorphic of degree α_1 and antiholomorphic of degree α_2. Endowed with the natural action of $\mathrm{U}(n)$ the space $\mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n)$ is an irreducible representation. It is unitary with respect to the norm $\|\cdot\|_{L^2(S^{2n-1})}$, where we view S^{2n-1} as the unit sphere in $\mathbb{C}^n$. Upon restriction to the subgroup $\mathrm{U}(n-1)$ the representation $\mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n)$ decomposes into

$$\mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n) = \bigoplus_{\substack{0\le\alpha_1'\le\alpha_1\\ 0\le\alpha_2'\le\alpha_2}} \mathcal{H}^{\alpha_1',\alpha_2'}(\mathbb{C}^{n-1}). \tag{B-5}$$

Explicit $\mathrm{U}(n-1)$-equivariant embeddings

$$I^n_{(\alpha_1',\alpha_2')\to(\alpha_1,\alpha_2)} : \mathcal{H}^{\alpha_1',\alpha_2'}(\mathbb{C}^{n-1}) \to \mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n)$$

are given by

$$\begin{aligned} &I^n_{(\alpha_1',\alpha_2')\to(\alpha_1,\alpha_2)}(\phi)(z',z_n) \\ &=\phi(z')\begin{cases} z_n^{(\alpha_1-\alpha_2)-(\alpha_1'-\alpha_2')} P_{\alpha_2-\alpha_2'}^{((\alpha_1-\alpha_2)-(\alpha_1'-\alpha_2'),\alpha_1'+\alpha_2'+n-2)}(1-2|z_n|^2) \\ \qquad\qquad\qquad \text{for } \alpha_1-\alpha_2 \ge \alpha_1'-\alpha_2', \\ \bar{z}_n^{(\alpha_1'-\alpha_2')-(\alpha_1-\alpha_2)} P_{\alpha_1-\alpha_1'}^{((\alpha_1'-\alpha_2')-(\alpha_1-\alpha_2),\alpha_1'+\alpha_2'+n-2)}(1-2|z_n|^2) \\ \qquad\qquad\qquad \text{for } \alpha_1-\alpha_2 \le \alpha_1'-\alpha_2', \end{cases} \end{aligned} \tag{B-6}$$

where $z=(z',z_n)\in S^{2n-1}$. For $\phi \in \mathcal{H}^{\alpha_1,\alpha_2}(\mathbb{C}^n)$ we have

$$z_j\phi = \phi_j^{+,\mathrm{hol}} + |z|^2\phi_j^{-,\mathrm{ahol}}, \quad \bar{z}_j\phi = \phi_j^{+,\mathrm{ahol}} + |z|^2\phi_j^{-,\mathrm{hol}}, \tag{B-7}$$

with $\phi_j^{\pm,\mathrm{hol}} \in \mathcal{H}^{\alpha_1\pm1,\alpha_2}(\mathbb{C}^n)$ and $\phi_j^{\pm,\mathrm{ahol}} \in \mathcal{H}^{\alpha_1,\alpha_2\pm1}(\mathbb{C}^n)$ given by

$$\begin{aligned} \phi_j^{+,\mathrm{hol}} &= z_j\phi - \frac{|z|^2}{\alpha_1+\alpha_2+n-1}\frac{\partial\phi}{\partial\bar{z}_j}, & \phi_j^{-,\mathrm{hol}} &= \frac{1}{\alpha_1+\alpha_2+n-1}\frac{\partial\phi}{\partial z_j}, \\ \phi_j^{+,\mathrm{ahol}} &= \bar{z}_j\phi - \frac{|z|^2}{\alpha_1+\alpha_2+n-1}\frac{\partial\phi}{\partial z_j}, & \phi_j^{-,\mathrm{ahol}} &= \frac{1}{\alpha_1+\alpha_2+n-1}\frac{\partial\phi}{\partial\bar{z}_j}. \end{aligned}$$

References

[Branson et al. 1996] T. Branson, G. Ólafsson, and B. Ørsted, "Spectrum generating operators and intertwining operators for representations induced from a maximal parabolic subgroup", *J. Funct. Anal.* **135**:1 (1996), 163–205. MR Zbl

[Erdélyi et al. 1953] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher transcendental functions, II*, McGraw-Hill, New York, 1953. MR Zbl

[Juhl 2009] A. Juhl, *Families of conformally covariant differential operators, Q-curvature and holography*, Prog. Math. **275**, Birkhäuser, Basel, 2009. MR Zbl

[Kobayashi 2014] T. Kobayashi, "Shintani functions, real spherical manifolds, and symmetry breaking operators", pp. 127–159 in *Developments and retrospectives in Lie theory*, edited by G. Mason et al., Dev. Math. **37**, Springer, 2014. MR Zbl

[Kobayashi 2015] T. Kobayashi, "A program for branching problems in the representation theory of real reductive groups", pp. 277–322 in *Representations of reductive groups*, edited by M. Nevins and P. E. Trapa, Progr. Math. **312**, Birkhäuser, Cham, 2015. MR Zbl

[Kobayashi and Mano 2011] T. Kobayashi and G. Mano, *The Schrödinger model for the minimal representation of the indefinite orthogonal group* $\mathrm{O}(p,q)$, Mem. Amer. Math. Soc. **1000**, 2011. MR Zbl

[Kobayashi and Speh 2015] T. Kobayashi and B. Speh, *Symmetry breaking for representations of rank one orthogonal groups*, Mem. Amer. Math. Soc. **1126**, 2015. MR Zbl

[Loke 2001] H. Y. Loke, "Trilinear forms of $\mathfrak{gl}_2$", *Pacific J. Math.* **197**:1 (2001), 119–144. MR Zbl

[Möllers and Ørsted 2017] J. Möllers and B. Ørsted, "Estimates for the restriction of automorphic forms on hyperbolic manifolds to compact geodesic cycles", *Int. Math. Res. Not.* **2017**:11 (2017), 3209–3236. MR Zbl

[Möllers and Oshima 2015] J. Möllers and Y. Oshima, "Restriction of most degenerate representations of $O(1,N)$ with respect to symmetric pairs", *J. Math. Sci. Univ. Tokyo* **22**:1 (2015), 279–338. MR Zbl

[Möllers et al. 2016a] J. Möllers, B. Ørsted, and Y. Oshima, "Knapp–Stein type intertwining operators for symmetric pairs", *Adv. Math.* **294** (2016), 256–306. MR Zbl

[Möllers et al. 2016b] J. Möllers, B. Ørsted, and G. Zhang, "Invariant differential operators on H-type groups and discrete components in restrictions of complementary series of rank one semisimple groups", *J. Geom. Anal.* **26**:1 (2016), 118–142. MR Zbl

[Möllers et al. 2016c] J. Möllers, B. Ørsted, and G. Zhang, "On boundary value problems for some conformally invariant differential operators", *Comm. Partial Differential Equations* **41**:4 (2016), 609–643. MR Zbl

[Ørsted and Vargas 2004] B. Ørsted and J. Vargas, "Restriction of square integrable representations: discrete spectrum", *Duke Math. J.* **123**:3 (2004), 609–633. MR Zbl

[Speh and Venkataramana 2011] B. Speh and T. N. Venkataramana, "Discrete components of some complementary series", *Forum Math.* **23**:6 (2011), 1159–1187. MR Zbl

[Sun and Zhu 2012] B. Sun and C.-B. Zhu, "Multiplicity one theorems: the Archimedean case", *Ann. of Math.* (2) **175**:1 (2012), 23–44. MR Zbl

[Zhang 2015] G. Zhang, "Discrete components in restriction of unitary representations of rank one semisimple Lie groups", *J. Funct. Anal.* **269**:12 (2015), 3689–3713. MR Zbl

Received February 14, 2017.

JAN FRAHM
DEPARTMENT OF MATHEMATICS
AARHUS UNIVERSITY
AARHUS
DENMARK
frahm@math.au.dk

BENT ØRSTED
DEPARTMENT OF MATHEMATICS
AARHUS UNIVERSITY
AARHUS
DENMARK
orsted@math.au.dk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.77

ON THE LANDSBERG CURVATURE OF A CLASS OF FINSLER METRICS GENERATED FROM THE NAVIGATION PROBLEM

LIBING HUANG, HUAIFU LIU AND XIAOHUAN MO

In this paper, we study the Landsberg curvature of a Finsler metric via conformal navigation problem. We show that the Landsberg curvature of F is proportional to its Cartan torsion where F is the Finsler metric produced from a Landsberg metric and its closed vector field in terms of the conformal navigation problem generalizing results previously known in the cases when F is a Randers metric or the Funk metric on a strongly convex domain. We also prove that the Killing navigation problem has the Landsberg curvature preserving property for a closed vector field.

1. Introduction

The flag curvature of a Finsler metric produced from a Riemann–Finsler metric and its conformal field in terms of the navigation problem has been determined [Huang and Mo 2015; Chern and Shen 2005; Cheng and Shen 2009; Mo and Huang 2007]. The flag curvature is an important Riemannian quantity in Finsler geometry because it takes the place of the sectional curvature in the Riemannian case and lies in the second variation formula of arc length.

Finsler geometry is more colorful than Riemannian geometry because there are several non-Riemannian quantities on a Finsler manifold besides the Riemannian quantities, such as the Cartan torsion $\boldsymbol{A}$ and the Landsberg curvature $\boldsymbol{L}$. The Cartan torsion $\boldsymbol{A}$ gives a measure of the failure of a Finsler metric to be a Riemannian metric. The Landsberg curvature $\boldsymbol{L}$ measures the rate of changes of the Cartan torsion along geodesics in a Finsler manifold. They all vanish for Riemannian metrics, hence they said to be *non-Riemannian*. For a Randers metrics F with its

Xiaohuan Mo is the corresponding author. Huang was supported by the National Natural Science Foundation of China 11301283 and 11571185. Liu was supported by the Beijing Natural Science Foundation 1182006. Mo was supported by the National Natural Science Foundation of China 11771020 .
MSC2010: 53B40, 58E20.
Keywords: Finsler metric, Landsberg curvature, conformal navigation problem.

navigation data (h, V), Shen [2004] proves its Landsberg curvature $\boldsymbol{L}$ and Cartan torsion $\boldsymbol{A}$ satisfy

$$\boldsymbol{L} + c(x)\boldsymbol{A} = 0, \tag{1-1}$$

i.e., $\boldsymbol{L}$ is proportional to $\boldsymbol{A}$ where V is a conformal field of h with dilation $c(x)$ and $V^\flat$ is closed. This interesting result prompts us to establish the relation between the Landsberg curvatures of a Finsler metric via a conformal navigation problem. Precisely we show the following:

Theorem 1.1. *Let $F = F(x, y)$ be a Finsler metric on a manifold M with its Landsberg curvature $\mathbf{L}$ and V a closed vector field on (M, F) with $F(x, V_x) < 1$. Let $\widetilde{F} = \widetilde{F}(x, y)$ denote the Finsler metric on M defined by*

$$F\left(x, \frac{y}{\widetilde{F}(x, y)} + V_x\right) = 1, \quad \text{for all } (x, y) \in TM. \tag{1-2}$$

Suppose that V is conformal with dilation $c(x)$. Then the Landsberg curvature $\widetilde{\boldsymbol{L}}$ and the Cartan torsion $\widetilde{\boldsymbol{A}}$ of $\widetilde{F}$ satisfy

$$\widetilde{\boldsymbol{L}}_y + c(x)\widetilde{\boldsymbol{A}}_y = \boldsymbol{L}_{\widetilde{y}},$$

where $\widetilde{y} = y + F(x, \widetilde{y})V$.

Theorem 1.1 tells us that the Killing navigation problem (i.e., $c(x) \equiv 0$) has the Landsberg curvature preserving property for a closed vector field.

For the definition of a closed vector field on a Finsler manifold see Section 4. In the case of a vector field V on a Riemannian manifold, our notion is reduced to $dV^\flat = 0$, where $\flat : TM \to T^*M$ denotes the musical isomorphism. It follows that if F is Riemannian, then Theorem 1.1 reduces to Shen's result (1-1).

Our method of proving Theorem 1.1 is partially in the contact geometry. It follows that our method is quite different from that of Shen [2004].

Theorem 1.2. *Let $F = F(x, y)$ be a Landsberg metric on a manifold M and V a closed vector field on (M, F) with $F(x, V_x) < 1$. Let $\widetilde{F} = \widetilde{F}(x, y)$ denote the Finsler metric on M defined in (1-2). Suppose that V is a conformal field of F. Then $\widetilde{F}$ has relatively isotropic Landsberg curvature, i.e., the Landsberg curvature of $\widetilde{F}$ is proportional to its Cartan torsion.*

A Finsler metric F is said to be *Landsberg type* if it has vanishing Landsberg curvature. It is known that on a Landsberg manifold (M, F), all $(T_xM\backslash\{0\}, \hat{g}_x)$ are isometric as Riemannian manifolds, where $\hat{g}_x := g_{ij}(x, y)dy^i \otimes dy^j$.

In $c =$ constant, we have another nontrivial example satisfying the conditions and conclusions in Theorem 1.2. Given a Minkowski norm $\varphi : \mathbb{R}^n \to \mathbb{R}$, a constant vector b and a constant c, one can construct a domain $\Omega := \{v \in \mathbb{R}^n \mid \varphi(2cv+b) < 1\}$. For each $x \in \Omega$, identify $T_x\Omega$ with $\mathbb{R}^n$. This $F(x, y))$ is a Minkowski metric on the domain Ω, where $F(x, y) = \varphi(y)$ and $V_x := 2cx + b$ is a vector field on Ω satisfying $F(x, V_x) = \varphi(2cx + b) < 1$. It can be shown that V is conformal with constant dilation c and V is a closed vector on (F, Ω). The proof will be given in Section 5. Define a new Finsler metric $\widetilde{F}$ by (1-2). Note that any Minkowski metric must be Landsberg type. By Theorem 1.2 we have

$$\widetilde{\boldsymbol{L}} + c\widetilde{\boldsymbol{A}} = 0. \tag{1-3}$$

We also present explicitly the geodesics of the Finsler metric $\widetilde{F}$ on the domain Ω. When $c = \frac{1}{2}$ and $b = 0$, $\widetilde{F}$ is the Funk metric on a strongly convex domain. Our result (1-3) has been obtained in [Shen 2001]. Again, the technique and method used in this paper is quite different from that of Shen.

2. Preliminaries

A *Finsler metric* on a manifold is a family of Minkowski norms on the tangent spaces. To characterize Riemannian metrics among Finsler metric, we define the *Cartan torsion* $\boldsymbol{A} = \{\boldsymbol{A}_y\}_{y \in T_xM \setminus \{0\}}$ by

$$\boldsymbol{A}_y(u, v) = A^i_{jk} u^j v^k \frac{\partial}{\partial x^i}, \qquad A^i_{jk} := \frac{F}{4} g^{il} \frac{\partial^2 F^2}{\partial y^j \partial y^k \partial y^l},$$

where $u = u^j(\partial/\partial x^j)$, $v = v^k(\partial/\partial x^k) \in T_xM$. Besides the Cartan torsion, there are other quantities which always vanish on Riemannian manifold. For instance, the following Landsberg curvature, it gives the rate of change of the Cartan torsion along geodesics. For a vector $y \in T_xM \setminus \{0\}$, define

$$\boldsymbol{L}_y(u, v) = \dot{A}^i_{jk} u^j v^k \frac{\partial}{\partial x^i},$$

where $u = u^j(\partial/\partial x^j)$, $v = v^k(\partial/\partial x^k) \in T_xM$ and "$\,\cdot\,$" denotes the covariant derivative along geodesics. $\boldsymbol{L} = \{\boldsymbol{L}_y\}_{y \in T_xM \setminus \{0\}}$ is called the *Landsberg curvature*. We say that F has *relatively isotropic Landsberg curvature* if $\boldsymbol{L} + c\boldsymbol{A} = 0$, where $c = c(x)$ is a scalar function on M [Shen 2004]. We say that F is a *Landsberg metric* if $\boldsymbol{L} = 0$.

Now we are going to give some lemmas. For related notions, such as the Reeb field, the Hilbert form and the angular metric, see [Huang and Mo 2011; 2015; Mo and Huang 2007].

Lemma 2.1. *Given a Finsler metric F and a vector field V with $F(x, V_x) < 1$, define a new Finsler metric $\widetilde{F}$ in (1-2). Denote the Cartan metrics of F and $\widetilde{F}$ by H and $\widetilde{H}$. Then the Reeb vector fields of H and $\widetilde{H}$ satisfy $\widetilde{\xi} = \xi - X_f$, where $f := p(V)/H$ and X_f is the Reeb vector field associated with f (or the infinitesimal contact transformation determined by f, see* [*Blair 2002*]).

Proof. We denote the corresponding objects with respect to $\widetilde{H}$ by adding a tilde "$\sim$". By (2.6) and (2.9) in [Huang and Mo 2015], we have

$$\widetilde{H}(x,p) = H(x,p) - p(V), \quad \omega^\flat = \frac{p}{H}, \tag{2-1}$$

where $\omega^\flat$ is the Hilbert form of H. It follows that

$$\phi := \frac{\widetilde{H}}{H} = 1 - \frac{p(V)}{H} = 1 - f. \tag{2-2}$$

By the second equation of (2-1), we obtain

$$\widetilde{\omega}^\flat = \phi^{-1}\omega^\flat. \tag{2-3}$$

We claim that

$$\widetilde{\xi} = X_\phi. \tag{2-4}$$

In fact,

$$\widetilde{\omega}^\flat(X_\phi) = \phi^{-1}\omega^\flat(X_\phi) = \phi^{-1}\phi = 1. \tag{2-5}$$

On the other hand, $0 = (X_\phi \lrcorner d\omega^\flat)(X_\phi) = -d\phi(X_\phi) + \xi(\phi)\omega^\flat(X_\phi) = \xi(\phi)\phi - X_\phi(\phi)$. Thus we have

$$\begin{aligned}\mathcal{L}_{X_\phi}\widetilde{\omega}^\flat = \mathcal{L}_{X_\phi}(\phi^{-1}\omega^\flat) &= X_\phi(\phi^{-1})\omega^\flat + \phi^{-1}\mathcal{L}_{X_\phi}\omega^\flat \\ &= -\phi^{-1}\xi(\phi)\omega^\flat + \phi^{-1}\xi(\phi)\omega^\flat = 0,\end{aligned}$$

where we have made use of (2-3). Together with (2-5) we obtain (2-4). It is easy to verify $X_{h+g} = X_h + X_g$ for functions h and g. In particular, we have $\widetilde{\xi} = X_\phi = X_{1-f} = X_1 - X_f = \xi - X_f$, where we have used (2-2). □

Lemma 2.2. *The vertical endomorphisms of H and $\widetilde{H}$ satisfy*

$$\widetilde{\mathcal{V}}^\flat = \mathcal{V}^\flat + \psi\mathcal{V}^\flat X_f \otimes \omega^\flat, \quad \psi := \phi^{-1}. \tag{2-6}$$

Proof. Let φ_t be the flow of X_f. The lift of φ_t is a flow $\widehat{\varphi_t}$ on S^*M, the co-sphere bundle of M, $\widehat{\varphi_t}(x,[p]) := (\varphi_t(x), [(\varphi_t^*)^{-1}(p)])$. It follows that $\iota \circ \widehat{\varphi_t} = \varphi_t \circ \iota$, where $\iota : S^*M \to M$ is the natural projection. Thus

$$\iota_* \circ (\widehat{\varphi_t})_* = \varphi_{t*} \circ \iota_*. \tag{2-7}$$

Now we assume that v is a vertical field, that is, $v \in VS^*M = \operatorname{Ker}\iota_*$. Then $(\widehat{\varphi_t})_* v$ is vertical from (2-7). Hence

$$[X_f, v] := \lim_{t\to 0} \frac{v - (\widehat{\varphi_t})_* v}{t} \tag{2-8}$$

is also vertical. It follows that

$$\omega^\flat[X_f, v] = 0, \quad \mathcal{V}^\flat[X_f, v] = 0. \tag{2-9}$$

Note that $\omega^\flat[\xi, v] = 0$. Then we obtain

$$(\mathcal{V}^\flat + \psi \mathcal{V}^\flat X_f \otimes \omega^\flat)[\widetilde{\xi}, v] = \mathcal{V}^\flat[\xi, v] + \psi \mathcal{V}^\flat X_f \otimes \omega^\flat[\xi, v] - \mathcal{V}^\flat[X_f, v] - \psi \mathcal{V}^\flat X_f \otimes \omega^\flat[X_f, v] = -v, \tag{2-10}$$

where $\widetilde{\xi} = \xi - X_f$. Observe that $\mathcal{V}^\flat(v) = \omega^\flat(v) = 0$. Then

$$(\mathcal{V}^\flat + \psi \mathcal{V}^\flat X_f \otimes \omega^\flat)(v) = 0. \tag{2-11}$$

Finally, we have

$$\begin{aligned}(\mathcal{V}^\flat + \psi \mathcal{V}^\flat X_f \otimes \omega^\flat)(\widetilde{\xi}) &= \mathcal{V}^\flat(\xi) - \mathcal{V}^\flat(X_f) + \psi \mathcal{V}^\flat X_f \omega^\flat(\xi) - \psi \mathcal{V}^\flat X_f \omega^\flat(X_f) \\ &= \mathcal{V}^\flat(X_f)(-1 + \psi - \psi f) \\ &= \mathcal{V}^\flat(X_f) \frac{-\widetilde{H} + H - p(V)}{\widetilde{H}} = 0.\end{aligned}$$

Together with (2-10), (2-11) and Proposition 4.6 in [Mo and Huang 2007], we have (2-6). □

Lemma 2.3. *The angular metric $h^\flat$ on VS^*M satisfies*

$$h^\flat(u, v) = -\omega^\flat([v, \mathcal{H}^\flat(u)]). \tag{2-12}$$

Proof. The angular metric $h^\flat$ is given by

$$h^\flat(u, v) = h^\flat(v, u) = d\omega^\flat(v, \mathcal{H}^\flat(u)) \quad \text{for all } u, v \in VS^*M, \tag{2-13}$$

where $\mathcal{H}^\flat$ is the horizontal endomorphism. Note that $\omega^\flat(u) = \omega^\flat(\mathcal{H}^\flat(v)) = 0$. Together with (2-13), we obtain

$$h^\flat(u, v) = v(\omega^\flat(\mathcal{H}^\flat(u))) - \mathcal{H}^\flat(u)(\omega^\flat(v)) - \omega^\flat([v, \mathcal{H}^\flat(u)]) = -\omega^\flat([v, \mathcal{H}^\flat(u)]). \quad \square$$

Lemma 2.4. *Assume that V is a conformal field on the Cartan manifold (M, H) with dilation $c(x)$. Then, for $u, v \in VS^*M$, the affine connections of H and $\widetilde{H}$ satisfy $\widetilde{\nabla}_v u = \nabla_v u - \psi h^\flat(u, v) \mathcal{V}^\flat X_f$.*

Proof. By using Lemmas 2.3 and 4.1 in [Huang and Mo 2015], we have

$$\widetilde{\nabla}_v u = \widetilde{\mathcal{V}}^\flat[v, \widetilde{\mathcal{H}}^\flat(u)], \quad \mathcal{H}^\flat(u) = \widetilde{\mathcal{H}}^\flat(u) - cu. \tag{2-14}$$

It follows that

$$\widetilde{\nabla}_v u = \widetilde{\mathcal{V}}^\flat[v, \mathcal{H}^\flat(u) + cu] = \widetilde{\mathcal{V}}^\flat[v, \mathcal{H}^\flat(u)] + \widetilde{\mathcal{V}}^\flat[v, cu] = \widetilde{\mathcal{V}}^\flat[v, \mathcal{H}^\flat(u)].$$

Together with Lemmas 2.2 and 2.3 and the first equation of (2-14), we have

$$\begin{aligned}\widetilde{\nabla}_v u &= (\mathcal{V}^\flat + \psi \mathcal{V}^\flat X_f \otimes \omega^\flat)[v, \widetilde{\mathcal{H}}^\flat(u)] \\ &= \mathcal{V}^\flat[v, \widetilde{\mathcal{H}^\flat}(u)] + \psi \mathcal{V}^\flat X_f \omega^\flat[v, \widetilde{\mathcal{H}^\flat}(u)] = \nabla_v u - \psi h^\flat(u, v) \mathcal{V}^\flat X_f. \quad \square\end{aligned}$$

Proposition 2.5. *The angular metric $h^\flat$ is given by*

$$h^\flat(\mathcal{V}^\flat X_f, v) = \frac{v(\psi)}{\psi^2}, \tag{2-15}$$

where ψ is given in the second equation of (2-6).

Proof. By the definition of the Reeb vector field associated with f, we have

$$X_f = \mathcal{P}^\flat_{\mathcal{H}} X_f + f\xi + \mathcal{P}^\flat_{\mathcal{V}} X_f, \tag{2-16}$$

where $\mathcal{P}^\flat_{\mathcal{H}}$ (resp. $\mathcal{P}^\flat_{\mathcal{V}}$) is the horizontal (resp. vertical) projection. The integrability of VS^*M tells us that

$$d\omega^\flat(\mathcal{P}^\flat_{\mathcal{V}} X_f, v) = -\omega^\flat([\mathcal{P}^\flat_{\mathcal{V}} X_f, v]) = 0. \tag{2-17}$$

On the other hand, $d\omega^\flat(f\xi, v) = f d\omega^\flat(\xi, v)$. Together with (2-16) and (2-17), we obtain

$$d\omega^\flat(X_f, v) = d\omega^\flat(\mathcal{P}^\flat_{\mathcal{H}} X_f, v). \tag{2-18}$$

By using (2-2) and (2-6), we have $1/\psi = 1 - f$. It follows that $df = \psi^{-2} d\psi$. Taking this together with (2-18) and (2-12), we obtain

$$\begin{aligned} h^\flat(\mathcal{V}^\flat X_f, v) &= d\omega^\flat(v, \mathcal{H}^\flat \mathcal{V}^\flat X_f) \\ &= d\omega^\flat(v, \mathcal{P}_{\mathcal{H}} X_f) \\ &= -d\omega^\flat(X_f, v) \\ &= [df - \xi(f)\omega](v) = df(v) = \psi^{-2} d\psi(v). \end{aligned}$$ □

Corollary 2.6. *For $u, v, w \in VS^*M$, the affine connections of H and $\widetilde{H}$ satisfy $\widetilde{h}^\flat(\widetilde{\nabla}_u v, w) = \psi h^\flat(\nabla_u v, w) - h^\flat(u, v) w(\psi)$.*

Proof. By Proposition 2.5 and Lemma 2.4, we have

$$\begin{aligned} \widetilde{h}^\flat(\widetilde{\nabla}_u v, w) &= \psi h^\flat(\nabla_u v - \psi h^\flat(u, v)\mathcal{V} X_f, w) \\ &= \psi h^\flat(\nabla_u v, w) - \psi^2 h^\flat(u, v) h^\flat(\mathcal{V} X_f, w) \\ &= \psi h^\flat(\nabla_u v, w) - h^\flat(u, v) w(\psi). \end{aligned}$$ □

Lemma 2.7. *The Cartan torsion of H and $\widetilde{H}$ satisfy*

$$\begin{aligned} 2\psi A^\flat(u, v, w) + u(\psi) h^\flat(v, w) + v(\psi) h^\flat(u, w) + w(\psi) h^\flat(u, v) \\ = 2\widetilde{A}^\flat(u, v, w), \end{aligned} \tag{2-19}$$

for $u, v, w \in VS^*M$.

Proof. According to Lemma 2.3 in [Huang and Mo 2015] and Lemma 6.7 in [Mo and Huang 2007],

$$2A^\flat(u,v,w) = (\nabla_u h^\flat)(v,w), \quad \widetilde{h}^\flat(u,v) = \psi h^\flat(u,v) \tag{2-20}$$

for $u, v, w \in VS^*M$. It follows that

$$\begin{aligned} 2\widetilde{A}^\flat(u,v,w) &= (\widetilde{\nabla}_u \widetilde{h}^\flat)(v,w) \\ &= u(\widetilde{h}^\flat(v,w)) - \widetilde{h}^\flat(\widetilde{\nabla}_u v, w) - \widetilde{h}^\flat(v, \widetilde{\nabla}_u w) \\ &= u(\psi)h^\flat(v,w) + v(\psi)h^\flat(u,w) + w(\psi)h^\flat(u,v) + (I), \end{aligned} \tag{2-21}$$

where

$$\begin{aligned} (I) := &\ \psi[u(h^\flat(v,w)) - h^\flat(\nabla_u v, w) - h^\flat(\nabla_u w, v)] \\ = &\ \psi(\nabla_u h^\flat)(v,w) = 2\psi A^\flat(u,v,w), \end{aligned} \tag{2-22}$$

where we have used Corollary 2.6. Plugging (2-22) into (2-21) yields (2-19). □

Define *(2,1)-Cartan torsion* $A^\flat$ by

$$h(A^\flat(u,v),w) = A^\flat(u,v,w) \tag{2-23}$$

for $u, v, w \in VS^*M$.

Corollary 2.8. *The* (2,1)*-Cartan torsion* $A^\flat$ *of* H *and* $\widetilde{H}$ *satisfy*

$$\widetilde{A^\flat} = A^\flat + \tfrac{1}{2} d(\ln\psi) \otimes \mathrm{Id} + \tfrac{1}{2}\,\mathrm{Id} \otimes d(\ln\psi) + \tfrac{1}{2}\psi h^\flat \otimes \mathcal{V}^\flat X_f. \tag{2-24}$$

Proof. From (2-23), (2-15), (2-19) and the second equation of (2-20), we obtain

$$\begin{aligned} \psi h^\flat(2\widetilde{A}^\flat(u,v),w) &= 2\widetilde{h}^\flat(\widetilde{A}^\flat(u,v),w) = 2\widetilde{A}^\flat(u,v,w) \\ &= 2\psi A^\flat(u,v,w) + u(\psi)h^\flat(v,w) \\ &\qquad + v(\psi)h^\flat(u,w) + w(\psi)h^\flat(u,v) \\ &= 2\psi h^\flat(A^\flat(u,v),w) + u(\psi)h^\flat(v,w) \\ &\qquad + v(\psi)h^\flat(u,w) + h^\flat(u,v)\psi^2 h^\flat(\mathcal{V}^\flat X_f, w) \\ &= \psi h^\flat\Big(2A^\flat(u,v) + \frac{1}{\psi}u(\psi)v + \frac{1}{\psi}v(\psi)u + \psi h^\flat(u,v)\mathcal{V}^\flat X_f, w\Big) \end{aligned}$$

for $u, v, w \in VS^*M$. This gives (2-24). □

3. Landsberg curvature

We say the navigation problem (1-2) is *conformal* if V is a conformal field [Huang and Mo 2015]. In this section we are going to give the relation between the Landsberg curvatures of F and $\widetilde{F}$, where $\widetilde{F}$ is the Finsler metric produced by conformal navigation problem (1-2).

We define the covariant derivative along the Reeb vector field ξ by

$$\nabla_\xi v := \mathcal{P}^\flat_{\mathcal{V}}[\xi, v] = \mathcal{V}^\flat \circ \mathcal{H}^\flat[\xi, v] \tag{3-1}$$

for $v \in VS^*M$.

Let V be a conformal field on a Cartan manifold with dilation $c(x)$. Let $\widetilde{H}$ be the Cartan metric given in the first equation of (2-1).

Lemma 3.1. *The covariant derivatives along ξ and $\widetilde{\xi}$ satisfy*

$$\widetilde{\nabla}_{\widetilde{\xi}} = \nabla_\xi - \mathcal{L}_{X_f} + c\,\mathrm{Id}\,. \tag{3-2}$$

Proof. By Lemma 4.2 in [Huang and Mo 2015], we have

$$\widetilde{\mathcal{P}}^\flat_{\widetilde{\mathcal{V}}} = \mathcal{P}^\flat_{\mathcal{V}} - c(x)\mathcal{V}^\flat \tag{3-3}$$

on $HS^*M \oplus VS^*M$. By the definition of $\widetilde{\xi}$, we have $\widetilde{\omega}^\flat([\widetilde{\xi}, v]) = -d\widetilde{\omega}^\flat(\widetilde{\xi}, v) = 0$. It follows that $[\widetilde{\xi}, v] \in \operatorname{Ker}\widetilde{\omega}^\flat = \operatorname{Ker}\omega^\flat = HS^*M \oplus VS^*M$. Together with (3-1) and (3-3) we obtain

$$\begin{aligned}
\widetilde{\nabla}_{\widetilde{\xi}} v &= \widetilde{\mathcal{P}^\flat}_{\mathcal{V}}[\widetilde{\xi}, v] \\
&= (\mathcal{P}^\flat_{\mathcal{V}} - c(x)\mathcal{V}^\flat)[\xi - X_f, v] \\
&= \mathcal{P}^\flat_{\mathcal{V}}[\xi, v] - \mathcal{P}^\flat_{\mathcal{V}}[X_f, v] - c(x)\mathcal{V}^\flat[\xi, v] + c(x)\mathcal{V}^\flat[X_f, v] \\
&= \nabla_\xi v - [X_f, v] + c(x)v = (\nabla_\xi - \mathcal{L}_{X_f} + c(x)\,\mathrm{Id})(v),
\end{aligned}$$

where we have used (2-8) and (2-9). This gives (3-2). □

The Landsberg curvature $L^\flat$ is given by

$$L^\flat = \nabla_\xi A^\flat, \tag{3-4}$$

where ξ is the Reeb vector field and $A^\flat$ is the Cartan torsion.

Proposition 3.2. *The Landsberg curvature $\widetilde{L}^\flat$ and the Cartan torsion $\widetilde{A}^\flat$ of $\widetilde{H}$ satisfy*

$$\widetilde{L}^\flat = T\widetilde{A}^\flat - c\widetilde{A}^\flat, \tag{3-5}$$

where

$$T := \nabla_\xi - \mathcal{L}_{X_f}. \tag{3-6}$$

Proof. By virtue of (3-2), (3-4) and (3-5),

$$\begin{aligned}
\widetilde{L}^\flat(u, v) = (\widetilde{\nabla}_{\widetilde{\xi}}\widetilde{A}^\flat)(u, v) &= \widetilde{\nabla}_{\widetilde{\xi}}(\widetilde{A}^\flat(u, v)) - \widetilde{A}^\flat(\widetilde{\nabla}_{\widetilde{\xi}}u, v) - \widetilde{A}^\flat(u, \widetilde{\nabla}_{\widetilde{\xi}}v) \\
&= (T + c\,\mathrm{Id})(\widetilde{A}^\flat(u, v)) - \widetilde{A}^\flat((T + c\,\mathrm{Id})u, v) - \widetilde{A}^\flat(u, (T + c\,\mathrm{Id})v) \\
&= T(\widetilde{A}^\flat(u, v)) - \widetilde{A}^\flat(Tu, v) - \widetilde{A}^\flat(u, Tv) - c\widetilde{A}^\flat(u, v) \\
&= (T\widetilde{A}^\flat)(u, v) - c\widetilde{A}^\flat(u, v)
\end{aligned}$$

for $u, v \in VS^*M$. This proves the proposition. □

Corollary 3.3. *For the conformal navigation problem (1-2),*

$$\widetilde{L}^\flat + c(x)\widetilde{A}^\flat = T[A^\flat + \tfrac{1}{2}d(\ln\psi)\otimes \mathrm{Id} + \tfrac{1}{2}\mathrm{Id}\otimes d(\ln\psi) + \tfrac{1}{2}\psi h^\flat \otimes \mathcal{V}^\flat X_f], \tag{3-7}$$

where $c(x)$ is the dilation of V.

Proof. Plugging (2-24) into (3-5) yields (3-7). □

Recall that a *flow* on a manifold M is a map $\phi : (-\varepsilon, \varepsilon)\times M \to M$, also denoted by $\phi_t := \phi(t, \cdot)$, satisfying (i) $\phi_0 = \mathrm{Id} : M \to M$, (ii) $\phi_s \circ \phi_t = \phi_{s+t}$ for any $s, t \in (-\varepsilon, \varepsilon)$ with $s+t \in (-\varepsilon, \varepsilon)$. Hence the lift of a flow ϕ_t on M is a flow on T^*M,

$$\widehat{\phi_t}(x, p) := (\phi_t(x), (\phi_t^*)^{-1}(p)). \tag{3-8}$$

By the relationship between vector fields and flows, (3-8) induces by a natural way a lift of a vector field V on M to a vector field X_V^* on T^*M. Now we assume that V is a conformal field of Cartan metric H with dilation $c(x)$ and ϕ_t its flow.

Lemma 3.4. *X_V^* is the Reeb vector field associated with $f := p(V)/H$, i.e.,*

$$X_V^* = X_f. \tag{3-9}$$

Therefore, $\widehat{\phi_t}$ is the flow of X_f.

Proof. Simple calculations give the following

$$\omega^\flat(X_V^*) = f, \quad X_V^* \lrcorner (d\omega^\flat) = -df + \xi(f)\omega^\flat.$$

Now our conclusion can be obtained from the uniqueness of the Reeb vector field associated with f. □

Lemma 3.5. *The flow of X_f satisfies*

$$\text{(i)} \quad (\widehat{\phi}_t^* h^\flat)(u, v) = [h^\flat(u, v)] \circ \phi_t, \tag{3-10}$$

$$\text{(ii)} \quad \widehat{\phi}_{t*}(\nabla_u v) = \nabla_{\widehat{\phi}_{t*}(u)} \widehat{\phi}_{t*}(v). \tag{3-11}$$

for $u, v \in VS^*M$.

Proof. By the definition of the vertical endomorphism $\mathcal{V}^\flat$, we have [Huang and Mo 2015]

$$\mathcal{V}^\flat(u) = 0, \quad \mathcal{V}^\flat(\xi) = 0, \quad \mathcal{V}^\flat[\xi, u] = -u, \qquad \text{for all } u \in VS^*M. \tag{3-12}$$

It follows that $\mathcal{V}^\flat[\xi, [\xi, v]] \in VS^*M$ for all $v \in VS^*M$.

Note that $VS^*M = \mathrm{Ker}\,\iota_*$ is integrable. We obtain

$$\omega^\flat[u, \mathcal{V}^\flat[\xi, [\xi, v]]] = 0 \tag{3-13}$$

for $u, v \in VS^*M$. The horizontal endomorphism $\mathcal{H}$ is given by

$$\mathcal{H}^\flat(v) = -[\xi, v] - \tfrac{1}{2}\mathcal{V}^\flat[\xi, [\xi, v]]. \tag{3-14}$$

Taking this together with (2-12) and (3-13), we obtain

$$h^\flat(u,v) = -\omega^\flat[u, \mathcal{H}^\flat(v)] = -\omega^\flat[u, -[\xi, v] - \tfrac{1}{2}\mathcal{V}^\flat[\xi,[\xi,v]]] = \omega^\flat[u,[\xi,v]]. \tag{3-15}$$

By using (3.1) in [Huang and Mo 2015],

$$\widehat{\phi}_t^*\omega^\flat = e^{2\sigma_t}\omega^\flat. \tag{3-16}$$

It follows that

$$\widehat{\phi}_{t*}\xi = e^{2\sigma_t}\xi. \tag{3-17}$$

Because ξ is the dual vector field of $\omega^\flat$ with respect to the Riemannian metric on S^*M, from which together with (3-15) we obtain

$$(\widehat{\phi}_t^* h^\flat)(u,v) = h^\flat(\widehat{\phi}_{t*}(u), \widehat{\phi}_{t*}(v)) = \omega^\flat[\widehat{\phi}_{t*}(u), [\xi, \widehat{\phi}_{t*}(v)]] = \omega^\flat[\widehat{\phi}_{t*}(u), (I)], \tag{3-18}$$

where

$$(I) := [e^{-2\sigma_t}\widehat{\phi}_{t*}(\xi), \widehat{\phi}_{t*}(v)] = -\widehat{\phi}_{t*}(v)(e^{-2\sigma_t})\widehat{\phi}_{t*}(\xi) + e^{-2\sigma_t}[\widehat{\phi}_{t*}(\xi), \widehat{\phi}_{t*}(v)] = e^{-2\sigma_t}\widehat{\phi}_{t*}[\xi, v], \tag{3-19}$$

where we have used the fact that

$$\widehat{\phi}_{t*}(v) \in VS^*M, \qquad \text{for all } v \in VS^*M. \tag{3-20}$$

Plugging (3-19) into (3-18) yields

$$\begin{aligned}(\widehat{\phi}_t^* h^\flat)(u,v) &= \omega^\flat[\widehat{\phi}_{t*}(u), e^{-2\sigma_t}\widehat{\phi}_{t*}[\xi, v]] \\ &= e^{-2\sigma_t}\omega^\flat[\widehat{\phi}_{t*}(u), \widehat{\phi}_{t*}[\xi, v]] \\ &= e^{-2\sigma_t}\omega^\flat(\widehat{\phi}_{t*}[u,[\xi,v]]) \\ &= e^{-2\sigma_t}(\widehat{\phi}_t^*\omega^\flat)([u,[\xi,v]]) \\ &= \omega^\flat([u,[\xi,v]])\circ\phi_t = h^\flat(u,v)\circ\phi_t,\end{aligned}$$

where we have made use of (3-20), (3-15) and (3-16). This gives (3-10).

We show that (3-11) holds. By Lemma 3.3 in [Huang and Mo 2015], we have

$$\widehat{\phi}_{t*}\circ\mathcal{V}^\flat = e^{-2\sigma_t}\mathcal{V}^\flat\circ\widehat{\phi}_{t*}. \tag{3-21}$$

Combining this with the first equation of (2-14), we have

$$\widehat{\phi}_{t*}(\nabla_u v) = \widehat{\phi}_{t*}(\mathcal{V}^\flat[u, \mathcal{H}^\flat(v)]) = e^{-2\sigma_t}\mathcal{V}^\flat(\widehat{\phi}_{t*}[u, \mathcal{H}^\flat(v)]) = e^{-2\sigma_t}\mathcal{V}^\flat([\widehat{\phi}_{t*}(u), (II)]), \tag{3-22}$$

where

$$(II) := \widehat{\phi}_{t*} \circ \mathcal{H}^\flat(v) = \widehat{\phi}_{t*}(-[\xi, v] - \tfrac{1}{2}\mathcal{V}^\flat[\xi, [\xi, v]]) \tag{3-23}$$
$$= - [\widehat{\phi}_{t*}\xi, \widehat{\phi}_{t*}v] - \tfrac{1}{2}e^{-2\sigma_t}\mathcal{V}^\flat \circ \widehat{\phi}_{t*}[\xi, [\xi, v]]$$
$$= - [e^{2\sigma_t}\xi, \widehat{\phi}_{t*}v] - \tfrac{1}{2}e^{-2\sigma_t}\mathcal{V}^\flat[\widehat{\phi}_{t*}\xi, \widehat{\phi}_{t*}[\xi, v]]$$
$$= - e^{2\sigma_t}[\xi, \widehat{\phi}_{t*}v] - \tfrac{1}{2}e^{-2\sigma_t}\mathcal{V}^\flat(III)$$

where we have used (3-17), (3-20) and (3-21) and

$$(III) := [e^{2\sigma_t}\xi, [e^{2\sigma_t}\xi, \widehat{\phi}_{t*}(v)]] = [e^{2\sigma_t}\xi, e^{2\sigma_t}[\xi, \widehat{\phi}_{t*}(v)]]$$
$$= e^{2\sigma_t}\xi(e^{2\sigma_t})[\xi, \widehat{\phi}_{t*}(v)] - e^{2\sigma_t}([\xi, \widehat{\phi}_{t*}(v)](e^{2\sigma_t}))\xi + e^{4\sigma_t}[\xi, [\xi, \widehat{\phi}_{t*}(v)]].$$

Plugging this into (3-23) yields

$$(II) = -e^{2\sigma_t}[\xi, \widehat{\phi}_{t*}v]$$
$$-\tfrac{1}{2}e^{-2\sigma_t}\mathcal{V}^\flat\big\{e^{2\sigma_t}\xi(e^{2\sigma_t})[\xi, \widehat{\phi}_{t*}(v)] - e^{2\sigma_t}([\xi, \widehat{\phi}_{t*}(v)](e^{2\sigma_t}))\xi + e^{4\sigma_t}[\xi, [\xi, \widehat{\phi}_{t*}(v)]]\big\}.$$

Together with (3-12) and (3-14) we obtain

$$(II) = -e^{2\sigma_t}[\xi, \widehat{\phi}_{t*}v] - \tfrac{1}{2}e^{-2\sigma_t}\Big\{ - e^{2\sigma_t}\xi(e^{2\sigma_t})\widehat{\phi}_{t*}(v) + e^{4\sigma_t}\mathcal{V}^\flat[\xi, [\xi, \widehat{\phi}_{t*}(v)]]\Big\}$$
$$= e^{2\sigma_t}\mathcal{H}^\flat \circ \widehat{\phi}_{t*}(v) + \tfrac{1}{2}\xi(e^{2\sigma_t})\widehat{\phi}_{t*}(v).$$

Substituting this into (3-22) yields

$$\widehat{\phi}_{t*}(\nabla_u v) = e^{-2\sigma_t}\mathcal{V}^\flat\big([\widehat{\phi}_{t*}(u), e^{2\sigma_t}\mathcal{H}^\flat \circ \widehat{\phi}_{t*}(v) + \tfrac{1}{2}\xi(e^{2\sigma_t})\widehat{\phi}_{t*}(v)]\big)$$
$$= \mathcal{V}^\flat([\widehat{\phi}_{t*}(u), \mathcal{H}^\flat \circ \widehat{\phi}_{t*}(v)]) = \nabla_{\widehat{\phi}_{t*}(u)}\widehat{\phi}_{t*}(v)$$

where we have used (3-12) and (3-20). □

Proposition 3.6. *The Cartan torsion $A^\flat$ of H satisfies $\mathcal{L}_{X_f}A^\flat = 0$.*

Proof. We can consider the flow of X_f, the lift of conformal transformation ϕ_t, by Lemma 3.4. Thus $\widehat{\phi}_t$ satisfies Lemma 3.5. Together with the first equation of (2-20), we have

$$2(\widehat{\phi}_t^* A^\flat)(u, v, w) = 2A^\flat(\widehat{\phi}_{t*}(u), \widehat{\phi}_{t*}(v), \widehat{\phi}_{t*}(w)) = (\nabla_{\widehat{\phi}_{t*}(u)}h^\flat)(\widehat{\phi}_{t*}(v), \widehat{\phi}_{t*}(w))$$
$$= (\widehat{\phi}_{t*}(u))(h(\widehat{\phi}_{t*}(v), \widehat{\phi}_{t*}(w))) - h^\flat(\nabla_{\widehat{\phi}_{t*}(u)}\widehat{\phi}_{t*}(v), \widehat{\phi}_{t*}(w))$$
$$- h^\flat(\widehat{\phi}_{t*}(v), \nabla_{\widehat{\phi}_{t*}(u)}\widehat{\phi}_{t*}(w))$$
$$= [u(h^\flat(u, v)) - h^\flat(\nabla_u v, w) - h^\flat(v, \nabla_u w)] \circ \phi_t$$
$$= [(\nabla_u h^\flat)(v, w)] \circ \phi_t = 2A^\flat(u, v, w) \circ \phi_t.$$

It follows that the Cartan torsion $A^\flat$ of H is invariant under the flow $\widehat{\phi}_t^*$. Thus we have proved Proposition 3.6. □

We define the covariant derivative of the angular metric along the Reeb field ξ by

$$(\nabla_\xi h^\flat)(u,v) := \nabla_\xi(h^\flat(u,v)) - h^\flat(\nabla_\xi u, v) - h^\flat(u, \nabla_\xi v) \tag{3-24}$$

for $u, v \in VS^*M$.

Lemma 3.7. $\nabla_\xi h^\flat = 0$.

Proof. Denote the orthonormal frame on VS^*M by $\{\widehat{e_{\bar\alpha}}\}$. Then we have that $h^\flat = (\omega^{\bar 1})^2 + \cdots + (\omega^{\overline{n-1}})^2$, where $\{\omega^{\bar\alpha}\}$ is the dual frame of $\{\widehat{e_{\bar\alpha}}\}$ where $n = \dim M$. It follows that

$$h^\flat(\widehat{e_{\bar\alpha}}, \widehat{e_{\bar\beta}}) = \delta_{\alpha\beta}, \quad \nabla_\xi(h^\flat(\widehat{e_{\bar\alpha}}, \widehat{e_{\bar\beta}})) = \xi(\delta_{\alpha\beta}) = 0. \tag{3-25}$$

By using (3-1), we have

$$\nabla_\xi \widehat{e_{\bar\alpha}} = \mathcal{P}^\flat_{\mathcal{V}}[\xi, \widehat{e_{\bar\alpha}}]. \tag{3-26}$$

Lemma 3.1 in [Mo and Huang 2007] tells us $[\xi, \widehat{e_{\bar\alpha}}] = -\widehat{e_\alpha} + \omega_\alpha^\beta(\xi)\widehat{e_{\bar\beta}}$, where $\{\omega_\alpha^\beta\}$ is Chern connection 1-form. Combining this with (3-26) we obtain $\nabla_\xi \widehat{e_{\bar\alpha}} = \omega_\alpha^\beta(\xi)\widehat{e_{\bar\beta}}$. It follows that

$$h^\flat(\nabla_\xi \widehat{e_{\bar\alpha}}, \widehat{e_{\bar\beta}}) = h(\omega_\alpha^\gamma(\xi)\widehat{e_{\bar\gamma}}, \widehat{e_{\bar\beta}}) = \omega_\alpha^\gamma(\xi)\delta_{\gamma\beta} = \omega_\alpha^\beta(\xi), \tag{3-27}$$

where we have used the first equation of (3-25). Similarly, we get

$$h^\flat(\widehat{e_{\bar\alpha}}, \nabla_\xi \widehat{e_{\bar\beta}}) = h^\flat(\nabla_\xi \widehat{e_{\bar\beta}}, \widehat{e_{\bar\alpha}}) = \omega_\beta^\alpha(\xi). \tag{3-28}$$

From the structure equation (2-4) in [Mo and Huang 2007], we have

$$\omega_\alpha^\beta + \omega_\beta^\alpha = -2H_{\alpha\beta\gamma}\omega^{\bar\gamma}. \tag{3-29}$$

By using (3-25), (3-27), (3-28) and (3-29), we have

$$\begin{aligned}(\nabla_\xi h^\flat)(\widehat{e_{\bar\alpha}}, \widehat{e_{\bar\beta}}) &= \nabla_\xi(h^\flat(\widehat{e_{\bar\alpha}}, \widehat{e_{\bar\beta}})) - h^\flat(\nabla_\xi \widehat{e_{\bar\alpha}}, \widehat{e_{\bar\beta}}) - h^\flat(\widehat{e_{\bar\alpha}}, \nabla_\xi \widehat{e_{\bar\beta}}) \\ &= \nabla_\xi(\delta_{\alpha\beta}) - (\omega_\alpha^\beta + \omega_\beta^\alpha)(\xi) = 2H_{\alpha\beta\gamma}\omega^{\bar\gamma}(\xi) = 0.\end{aligned}$$

Note that $\nabla_\xi h^\flat$ is a tensor. Thus we have proved Lemma 3.7. □

In the second equation of (3-25), we used the following definition: $\nabla_\xi g := \xi(g)$, for all $g \in C^\infty(S^*M)$. Together with Lemma 2.1 we obtain

$$\widetilde{\nabla}_{\tilde\xi} g = \tilde\xi(g) = (\xi - X_f)(g) = \nabla_\xi g - \mathcal{L}_{X_f} g = Tg, \tag{3-30}$$

where T is given in (3-6).

Lemma 3.8. *The angular metric $\widetilde{h}^\flat$ of $\widetilde{H}$ satisfies*

$$T\widetilde{h}^\flat = 2c(x)\widetilde{h}^\flat. \tag{3-31}$$

Proof. By (3-30), (3-2), (3-6) and Lemma 3.7, we see that

$$\begin{aligned} 0 = (\widetilde{\nabla}_{\widetilde{\xi}}\widetilde{h}^{\flat})(u,v) &= \widetilde{\nabla}_{\widetilde{\xi}}(\widetilde{h}^{\flat}(u,v)) - \widetilde{h}^{\flat}(\widetilde{\nabla}_{\widetilde{\xi}}u, v) - \widetilde{h}^{\flat}(u, \widetilde{\nabla}_{\widetilde{\xi}}v) \\ &= T(\widetilde{h}^{\flat}(u,v)) - \widetilde{h}^{\flat}(Tu + cu, v) - \widetilde{h}^{\flat}(u, Tv + cv) \\ &= T(\widetilde{h}^{\flat}(u,v)) - \widetilde{h}^{\flat}(Tu, v) - \widetilde{h}^{\flat}(u, Tv) - 2c\widetilde{h}^{\flat}(u,v) \\ &= (T\widetilde{h}^{\flat})(u,v) - 2c\widetilde{h}^{\flat}(u,v) \end{aligned}$$

for $u, v \in VS^*M$. Then (3-31) holds. □

Proposition 3.9. *The derivation T satisfies*

$$T(\mathcal{V}^{\flat}X_f) = \mathcal{P}^{\flat}_{\mathcal{V}}X_f - 2c\mathcal{V}^{\flat}X_f, \tag{3-32}$$

$$T(d\ln\psi) = \widetilde{h}^{\flat}(\mathcal{P}^{\flat}_{\mathcal{V}}X_f, \cdot), \tag{3-33}$$

where ψ is given in the second equation of (2-6).

Proof. By Lemma 3.4 in [Huang and Mo 2015], we have

$$[\xi, X_{\phi}] = X_{\xi(\phi)}, \quad \phi \in C^{\infty}(S^*M).$$

Together with (3-1), (2-16) and (3-12), we have

$$\begin{aligned} \nabla_{\xi}\mathcal{V}^{\flat}X_f &= \mathcal{V}^{\flat}[\xi, \mathcal{H}^{\flat}\circ\mathcal{V}^{\flat}X_f] = \mathcal{V}^{\flat}[\xi, \mathcal{P}^{\flat}_{\mathcal{H}}X_f] = \mathcal{V}^{\flat}[\xi, X_f - \mathcal{P}^{\flat}_{\mathcal{V}}X_f - f\xi] \\ &= \mathcal{V}^{\flat}[\xi, X_f] - \mathcal{V}^{\flat}[\xi, \mathcal{P}^{\flat}_{\mathcal{V}}X_f] - \mathcal{V}^{\flat}[\xi, f\xi] = \mathcal{V}^{\flat}X_{\xi(f)} + \mathcal{P}^{\flat}_{\mathcal{V}}X_f. \end{aligned} \tag{3-34}$$

By using (3-9) and (3-16), we get

$$\begin{aligned} \mathcal{L}_{X_f}\omega^{\flat} &= \lim_{t\to 0}\frac{\widehat{\phi}_t^*\omega^{\flat} - \omega^{\flat}}{t} = \frac{d}{dt}(\widehat{\phi}_t^*\omega^{\flat})\,|_{t=0} \\ &= \frac{d}{dt}(e^{2\sigma_t}\omega^{\flat})\,|_{t=0} = 2\frac{d\sigma_t^{\flat}}{dt}\,|_{t=0}\,\omega^{\flat} = 2c(x)\omega^{\flat}. \end{aligned} \tag{3-35}$$

By the definition of the Reeb vector field associated with f, we have

$$\mathcal{L}_{X_f}\omega^{\flat} = d(\omega^{\flat}(X_f)) + \iota_{X_f}(d\omega^{\flat}) = df + d\omega^{\flat}(X_f, \cdot) = \xi(f)\omega^{\flat}.$$

Combining this with (3-35) we get

$$\xi(f) = 2c(x). \tag{3-36}$$

Recall that $VS^*M = \{v \in TSM \mid v(g) = 0, \text{ for all } g \in C^{\infty}(M) \subset C^{\infty}(S^*M)\}$. Together with (3-36) and the proof of Proposition 2.5 we have $0 = v(2c(x)) = v(\xi(f)) = h^{\flat}(\mathcal{V}^{\flat}X_{\xi(f)}, v)$. It follows that

$$\mathcal{V}^{\flat}X_{\xi(f)} = 0. \tag{3-37}$$

Plugging this into (3-34) yields

$$\nabla_\xi \mathcal{V}^\flat X_f = \mathcal{P}^\flat_{\mathcal{V}} X_f. \tag{3-38}$$

By using Lemma 3.4 we have $\widehat{\phi}_{t*}X_f = X_f$. Taking this together with (3-21) yields

$$\begin{aligned}\mathcal{L}_{X_f}\mathcal{V}^\flat X_f &= \lim_{t\to 0}\frac{\mathcal{V}^\flat X_f - \widehat{\phi}_{t*}\mathcal{V}^\flat X_f}{t} = -\frac{d}{dt}\widehat{\phi}_{t*}\mathcal{V}^\flat X_f\,|_{t=0}\\ &= -\frac{d}{dt}(e^{-2\sigma_t}\mathcal{V}^\flat\circ\widehat{\phi}_{t*}X_f)\,|_{t=0}\\ &= -\frac{d}{dt}(e^{-2\sigma_t}\mathcal{V}^\flat X_f)\,|_{t=0} = 2\frac{d\sigma_t}{dt}\,|_{t=0}\,\mathcal{V}^\flat X_f = 2c(x)\mathcal{V}X_f.\end{aligned}$$

Together with (3-38) and (3-6), we get (3-32).

Now we show that (3-33) holds. By (2-15) and the second equation of (2-20) we have $d(\ln\psi)(v) = v(\psi)/\psi = \psi h^\flat(\mathcal{V}^\flat X_f, v) = \widetilde{h}^\flat(\mathcal{V}^\flat X_f, v)$. Together with (3-31) and (3-32) we obtain

$$\begin{aligned}[T(d(\ln\psi))](v) &= T(d(\ln\psi)(v)) - (d(\ln\psi))(T(v))\\ &= T(\widetilde{h}^\flat(\mathcal{V}^\flat X_f, v)) - \widetilde{h}^\flat(\mathcal{V}^\flat X_f, T(v))\\ &= (T\widetilde{h}^\flat)(\mathcal{V}^\flat X_f, v) + \widetilde{h}^\flat(T(\mathcal{V}^\flat X_f), v)\\ &= 2c\widetilde{h}^\flat(\mathcal{V}^\flat X_f, v) + \widetilde{h}^\flat(\mathcal{P}^\flat_{\mathcal{V}}X_f - 2c\mathcal{V}^\flat X_f, v) = \widetilde{h}^\flat(\mathcal{P}^\flat_{\mathcal{V}}X_f, v)\end{aligned}$$

for $v \in VS^*M$. This gives (3-33). □

Proposition 3.10. *For the conformal navigation problem* (1-2),

$$\widetilde{L}^\flat + c(x)\widetilde{A}^\flat = L^\flat + (I), \tag{3-39}$$

where $c(x)$ is the dilation of V and

$$(I) := \tfrac{1}{2}\widetilde{h}^\flat(\mathcal{P}^\flat_{\mathcal{V}}X_f,\cdot)\otimes Id + \tfrac{1}{2}Id\otimes\widetilde{h}^\flat(\mathcal{P}^\flat_{\mathcal{V}}X_f,\cdot) + \tfrac{1}{2}\widetilde{h}^\flat\otimes\mathcal{P}^\flat_{\mathcal{V}}X_f. \tag{3-40}$$

Proof. By using (3-7), (3-31), (3-32), (3-33), (2-20) and Proposition 3.6, we have

$$\begin{aligned}&T[A^\flat + \tfrac{1}{2}d(\ln\psi)\otimes\mathrm{Id} + \tfrac{1}{2}\mathrm{Id}\otimes d(\ln\psi) + \tfrac{1}{2}\psi h^\flat\otimes\mathcal{V}^\flat X_f]\\ &\quad= TA^\flat + \tfrac{1}{2}T[d(\ln\psi)\otimes\mathrm{Id}] + \tfrac{1}{2}T[\mathrm{Id}\otimes d(\ln\psi)] + \tfrac{1}{2}T(\widetilde{h}^\flat\otimes\mathcal{V}^\flat X_f)\\ &\quad= \nabla_\xi A^\flat - \mathcal{L}_{X_f}A^\flat + \tfrac{1}{2}T(d\ln\psi)\otimes\mathrm{Id}\\ &\qquad\quad + \tfrac{1}{2}\mathrm{Id}\otimes T(d\ln\psi) + \tfrac{1}{2}(T\widetilde{h}^\flat)\otimes\mathcal{V}^\flat X_f + \tfrac{1}{2}\widetilde{h}^\flat\otimes T(\mathcal{V}^\flat X_f)\\ &\quad= L^\flat + \tfrac{1}{2}\widetilde{h}^\flat(\mathcal{P}^\flat_{\mathcal{V}}X_f,\cdot)\otimes\mathrm{Id} + \tfrac{1}{2}\mathrm{Id}\otimes\widetilde{h}^\flat(\mathcal{P}^\flat_{\mathcal{V}}X_f,\cdot)\\ &\qquad\quad + c\widetilde{h}^\flat\otimes\mathcal{V}^\flat X_f + \tfrac{1}{2}\widetilde{h}^\flat\otimes(\mathcal{P}^\flat_{\mathcal{V}}X_f - 2c\mathcal{V}^\flat X_f)\\ &\quad= L^\flat + (I).\end{aligned}$$

Plugging this into (3-7) yields (3-39). □

4. Closed vector field

Let (M, F) be a Finsler manifold. A vector field V on (M, F) is said to be *closed* if $dV^\flat \equiv 0 \pmod{\delta y^i}$, where

$$V^\flat = V^j g_{ij} dx^i, \qquad V = V^j \frac{\partial}{\partial x^j}, \tag{4-1}$$

and δy^i are defined in (4-2).

Proposition 4.1. *Let V be a vector field on a Finsler manifold (M, F). Then the following assertions are equivalent*:

(i) *V is closed*;

(ii) $\delta V_i/\delta x^j = \delta V_j/\delta x^i$;

(iii) $V_{i|j} = V_{j|i}$,

where $V_i = V^\flat(\partial/\partial x^i)$, $\delta/\delta x^j := \partial/\partial x^j - N^i_j \partial/\partial y^i$ and "$|$" denotes the horizontally covariant derivative with respect to Chern connection or Berwald connection.

Proof. Recall that two natural (local) bases that are dual to each other:

- $\{\delta/\delta x^i, F(\partial/\partial y^i)\}$ for the tangent bundle of $TM\backslash\{0\}$,
- $\{dx^i, \delta y^i/F\}$ for the cotangent bundle of $TM\backslash\{0\}$,

where

$$\delta y^i := dy^i + N^i_j dx^j. \tag{4-2}$$

It follows that $dV^\flat = d(V_i dx^i) = dV_i \wedge dx^i = \big((\delta V_i/\delta x^j)dx^j + (\partial V_i/\partial y^j)\delta y^i\big) \wedge dx^i$. Thus we have $dV^\flat|_{HTM\times HTM} = (\delta V_i/\delta x^j)dx^j \wedge dx^i$. We obtain that V is closed if and only if (ii) holds. By the definition of covariant derivative, we have $V_{i|j} = \partial V_i/\partial x^j - V_k\Gamma^k_{ij} - (\partial V_i/\partial y^k)N^k_j = \delta V_i/\delta x^j - V_k\Gamma^k_{ij}$, where Γ^k_{ij} satisfies $\Gamma^k_{ij} = \Gamma^k_{ji}$. It follows that $V_{i|j} - V_{j|i} = \delta V_i/\delta x^j - \delta V_j/\delta x^i - V_k(\Gamma^k_{ij} - \Gamma^k_{ji}) = \delta V_i/\delta x^j - \delta V_j/\delta x^i$. Thus (ii) is equivalent to (iii). □

Remark. By Proposition 4.1 our notion of closed vector field V reduces to $dV^\flat = 0$ if (M, F) is a Riemannian manifold where $\flat : TM \to T^*M$ is a musical isomorphism. For a Randers metric we have the following result, the proof of which is omitted.

Proposition 4.2. *Let $F = \sqrt{h_{ij}(x)y^i y^j}$ be a Riemannian metric on a manifold M and V a vector field on M with $F(x, V_x) < 1$. Let $\widetilde{F} = \alpha + \beta$ denote the Randers metric on M defined in (1-2). Then*:

(i) *If V is conformal with respect to F, then V is closed if and only if $d\beta = 0$.*

(ii) *V is conformal and closed if and only if $\widetilde{F}$ is of relatively isotropic Landsberg curvature.*

Lemma 4.3. *Let $F = F(x, y)$ be a Finsler metric on a manifold M and V a closed vector field on M with $F(x, V_x) < 1$. Suppose that V is conformal with dilation $c(x)$. Let $\widetilde{F} = \widetilde{F}(x, y)$ denote the Finsler metric on M defined in (1-2). Then the lift of V satisfies*

$$X_V = V^i \frac{\delta}{\delta x^i} + 2c(x) y^i \frac{\partial}{\partial y^i}. \tag{4-3}$$

In particular, $X_V|_{SM} \subset HSM$.

Proof. By Lemma 3.1 in [Huang and Mo 2013], $X_V(F) = 2c(x)F$. It follows that $X_V(L) = 4c(x)L$, where $L := F^2/2$ is the Lagrangian function. In natural coordinates

$$X_V = V^i \frac{\partial}{\partial x^i} + y^j \frac{\partial V^i}{\partial x^j} \frac{\partial}{\partial y^i}, \tag{4-4}$$

where $V = V^i(\partial/\partial x^i)$ [Mo 2011; 2012]. Hence $X_V(L) = y_i(\partial V^i/\partial x^j)y^j + V^i L_{x^i}$, where we have used the fact $L_{y^i} = \partial L/\partial y^i = g_{ij} y^j := y_i$. Together with (4-4) we get

$$V^k L_{x^k} + y^k V^i_k y_i = 4c(x)L. \tag{4-5}$$

Differentiating (4-5) with respect to y^j, we obtain

$$V^k L_{x^k y^j} + V^i_j y_i + y^k V^i_k g_{ij} = 4c y_j. \tag{4-6}$$

We know (see (2.14) and (2.18) in [Chern and Shen 2005])

$$G^i = \tfrac{1}{2} g^{il}(L_{x^j y^l} y^j - L_{x^l}), \qquad N^i_k = \frac{\partial G^i}{\partial y^k}.$$

It follows that

$$y_i N^i_k = \tfrac{1}{2}\left[y_i \frac{\partial g^{il}}{\partial y^k}(L_{x^j y^l} y^j - L_{x^l}) + y^l (L_{x^j y^l y^k} y^j + L_{x^j y^l} \delta^j_k - L_{x^l y^k}) \right] = L_{x^k}. \tag{4-7}$$

Differentiating (4-7) with respect to y^j, we get $L_{x^k y^j} = g_{ij} N^i_k + y_i {}^b\Gamma^i_{kj}$, where ${}^b\Gamma^i_{kj}$ are Berwald connection coefficients [Bao et al. 2000]. Plugging this into (4-7) yields

$$g_{ij}(V^k N^i_k + y^k V^i_k) + y_i(V^{k\,b}\Gamma^i_{kj} + V^i_j) = 4c y_j. \tag{4-8}$$

Note that

$$V_{j|0} = (V^i g_{ij})_{|0} = g_{ij}\left(\frac{\partial V^i}{\partial x^l} + V^k \Gamma^i_{lk} \right) y^l = g_{ij}(y^k V^i_k + V^k N^i_k). \tag{4-9}$$

On the other hand, $y_{i|j} = (y^k g_{ki})_{|j} = y^k_{|j} g_{ki} + y^k g_{ki|j} = 0$. It follows that

$$\begin{aligned} V_{0|j} &= (y_i V^i)_{|j} = y_i(V^i_j + V^k \Gamma^i_{kj}) \\ &= y_i[V^i_j + V^k({}^b\Gamma^i_{kj} - L^i_{kj})] = y_i(V^i_j + V^{k\,b}\Gamma^i_{kj}). \end{aligned} \tag{4-10}$$

Plugging (4-9) and (4-10) into (4-8) yields

$$V_{j|0} + V_{0|j} = 4c y_j. \tag{4-11}$$

Now we assume that V is closed. From Proposition 4.1(iii), we have

$$V_{i|0} = V_{i|j} y^j = V_{j|i} y^j = (V_j yj)_i = V_{0|i}. \tag{4-12}$$

Substituting (4-12) into (4-11) yields

$$V_{i|0} = V_{0|i} = 2c(x) y_i. \tag{4-13}$$

It follows that $y^k V^i_k + V^k N^i_k = V^i_{|0} = (V_j g^{ij})_{|0} = V_{j|0} g^{ij} = 2c(x) y^i$, where we have used (4-9). Plugging this into (4-4) yields

$$X_V = V^i \left(\frac{\partial}{\partial x^i} - N^k_i \frac{\partial}{\partial y^k} \right) + 2cy^i \frac{\partial}{\partial y^i} = V^i \frac{\delta}{\delta x^i} + 2cy^i \frac{\partial}{\partial y^i}. \qquad \square$$

Let

$$\theta := L_{y^i} dx^i. \tag{4-14}$$

The 1-form θ is global. In fact, $\theta = F\omega$, where ω is the Hilbert form. Define $u : TM \setminus \{0\} \to \mathbb{R}$ by

$$u := \theta(V). \tag{4-15}$$

Consider a unique vector field X_u on $TM \setminus \{0\}$ satisfying

$$\theta(X_u) = u, \qquad X_u \lrcorner (d\theta) = -du + 2c(x)\theta. \tag{4-16}$$

Lemma 4.4. *Let "$\cdot$" denote the covariant derivative along the Reeb vector field. Then*

$$X_u = V^i \frac{\delta}{\delta x^i}, \tag{4-17}$$

$$\dot{u} = 4c(x) L. \tag{4-18}$$

Proof. To prove (4-17), put $X_u := a^k(\delta/\delta x^k) + b^k(\partial/\partial y^k)$. Together with (4-15) and the first equation of (4-16) we have

$$y_i V^i = u = \theta(X_u) = y_i dx^i (a^k(\delta/\delta x^k) + b^k(\partial/\partial y^k)) = y_i a^i.$$

By using (4-14) we have $d\theta = g_{ij} \delta y^i \wedge dx^j$, where δy^j are given by (4-2). Therefore

$$X_u \lrcorner (d\theta) = \left(a^k \frac{\delta}{\delta x^k} + b^k \frac{\partial}{\partial y^k} \right) \lrcorner (g_{ij} \delta y^i \wedge dx^j) = g_{ij} b^i dx^j - g_{ij} a^i \delta y^j. \tag{4-19}$$

On the other hand, $V_0 = V_i y^i = y_i V^i = u$. Hence we get

$$u_{|j} = V_{0|j} = 2cy_j,$$

using (4-13). It follows that $du = u_{|j} dx^j + u_{y^j} \delta y^j = 2cy_j dx^j + g_{ij} V^i \delta y^j$. Thus $-du + 2c(x)\theta = -g_{ij} V^i \delta y^j$. Together with (4-19) and the second equation of (4-16) we have $g_{ij} b^i = 0$, $g_{ij} a^i = g_{ij} V^i$. We get $b^i = 0$, $a^i = V^i$. This gives (4-17). By using (4-18) we have $\dot{u} = u_{|j} y^j = 2cF^2 = 4c(x) L$. $\square$

Let L^F be the Legendre transformation of F [Huang and Mo 2011]. Note that $\theta_{(x,y)} = F(x, y)\omega_{(x,y)} = L^F_x(y) = p$. If we view θ as a 1-form on TM, then $\theta^\flat$, its dual quantity on T^*M, satisfies

$$\theta^\flat = ((L^F)^{-1})^*\theta = p. \tag{4-20}$$

Define $v := T^*M \setminus \{0\} \to \mathbb{R}$ by

$$v := \theta^\flat(V) = p(V). \tag{4-21}$$

By (4-18) we get

$$\dot{v} = 4cL^\flat = 2cH^2. \tag{4-22}$$

Consider a unique vector field X_v on $T^*M \setminus \{0\}$ satisfying

$$\theta^\flat(X_v) = v, \quad X_v \lrcorner (d\theta^\flat) = -dv + 2c(x)\theta^\flat. \tag{4-23}$$

Together with (4-20), (4-15), (4-21) and (4-22), we get the following:

Lemma 4.5. $$L^F_* X_u = X_v.$$

Let $\iota : S^*M \to M$ be the natural projection and $\pi : SM \to M$ the natural projection. It is clear that $\iota \circ L^F = \pi$. Thus we have

$$L^F_*(VSM) \subset VS^*M. \tag{4-24}$$

On the other hand, $(L^F)_*\mathcal{H} = \mathcal{H}^\flat$. Hence we have the following:

Lemma 4.6. $$L^F_*(HSM) = HS^*M.$$

Proposition 4.7. *Let $F = F(x, y)$ be a Finsler metric on a manifold M and V a closed field on M with $F(x, V_x) < 1$. Suppose that V is conformal with respect to F. Let $\widetilde{F} = \widetilde{F}(x, y)$ denote the Finsler metric on M defined by* (1-2). *Then*

$$\mathcal{P}^\flat_{\mathcal{V}} X_f = 0. \tag{4-25}$$

Proof. Plugging (4-17) into (4-3) yields $X_V = X_u + 2c(x)y^i(\partial/\partial y^i)$. It follows that $X_V|_{SM} = X_u$. Together with Lemma 4.6 we get

$$L^F_* X_u = L^F_* X_V|_{SM} \subset HS^*M. \tag{4-26}$$

On the other hand, $v|_{S^*M} = f$. It follows that, on S^*M, $\dot{f} = \dot{v} = 2c(x)$, where we have used (4-22). From which together with (4-23) we have, on S^*M, $X_f = X_v$ is determined uniquely by $\omega^\flat(X_f) = f$ and $X_f \lrcorner d\omega^\flat = -df + \dot{f}\omega^\flat$, where we have used the fact $\theta^\flat = \omega^\flat$ on S^*M. Combining this with (4-26) yields $X_f = X_v = L^F_* X_u \subset HS^*M$. This gives Proposition 4.7. □

Proposition 4.8. *For the conformal navigation problem* (1-2), *if V is closed, then*

$$\widetilde{L}^\flat + c(x)\widetilde{A}^\flat = L^\flat. \tag{4-27}$$

Proof. Plugging (4-25) into (3-40) and combining with (3-39) yields (4-27). □

5. Proof of Theorem 1.1 and examples

Proof. By Proposition 4.8, we have that $\widetilde{L}^\flat(u,v)+c(x)\widetilde{A}^\flat(u,v)=L^\flat(u,v)$ for $u,v\in VS^*M$. Hence $[\widetilde{L}^\flat(u,v)]_{(x,[p])}+c(x)[\widetilde{A}^\flat(u,v)]_{(x,[p])}=[L^\flat(u,v)]_{(x,[p])}$. Pulling back to the sphere bundle, we have $[\widetilde{L}(w,z)]_{(x,[y])}+c(x)[\widetilde{A}(w,z)]_{(x,[y])}=[L(w,z)]_{(x,[y])}$, where $w:=(L_x^F)_*u,\ z:=(L_x^F)_*v$ and we have used $\partial\widetilde{H}/\partial p_i=y^i/\widetilde{F}$. As we know, there is a globally defined one-to-one map between $\pi^*TM\backslash\{\mathcal{Y}\}$ and VSM, where π^*TM the pull-back bundle over SM, $\pi:SM\to M$ is the natural projection and $\mathcal{Y}$ is the canonical section. Both the Landsberg tensor and the Cartan tensor are vanishing along the canonical section $\mathcal{Y}$. Thus we get the desired result. □

Example. A Randers metric can be expressed in the navigation form

$$F=\frac{\sqrt{(1-b^2)\alpha^2+\beta^2}}{1-b^2}+\frac{\beta}{1-b^2},$$

where $(\alpha,\beta^{\flat^{-1}})$ is the navigation data of F and $b:=\|\beta\|_\alpha$ is the length of β. Suppose that $\beta^{\flat^{-1}}$ is closed and conformal with dilation $c(x)$. Then F satisfies $\boldsymbol{L}+c(x)\boldsymbol{A}=0$ [Shen 2004].

Proposition 5.1. *Let $\varphi:\mathbb{R}^n\to\mathbb{R}$ be a Minkowski norm, $c=$ constant and $b\in\mathbb{R}^n$. Let $\Omega:=\{v\in\mathbb{R}^n\mid\varphi(2cv+b)<1\}$. Assume $\widetilde{F}$ is the Finsler metric on Ω defined by*

$$\varphi(y/\widetilde{F}(x,y)+2cx+b)=1$$

Then its Landsberg curvature $\widetilde{\boldsymbol{L}}$ and the Cartan torsion $\widetilde{\boldsymbol{A}}$ satisfy (1-3). Moreover, the geodesics of $\widetilde{F}$ are given by $e^{-2ct}[x+y((e^{2ct}-1)/(2c\varphi(y)))]-tb$ (resp. $x+(t/\varphi(y))y-tb$) for $c\neq0$ (resp. $c=0$).

Proof. For each $x\in\Omega$, identify $T_x\Omega$ with $\mathbb{R}^n$. It is easy to see that $(\Omega,F(x,y))$ is a Minkowski manifold, where $F(x,y)=\varphi(y)$. It follows that $V_x:=2cx+b$ is a vector field on Ω satisfying $F(x,V_x)=\varphi(2cx+b)<1$. By a straightforward computation one obtains $X_V(F)=2cF$, where X_V is the lift of V given in (4-4). It follows that V is a conformal field of F with constant dilation c [Huang and Mo 2011; Mo and Huang 2007]. Now we show that V is closed. In fact,

$$G^i=0,\qquad N^i_j=0$$

for Minkowski manifold $(\Omega,F(x,y))$. Plugging this into (4-2) yields

$$\delta y^i=dy^i. \tag{5-1}$$

On the other hand, $g_{ij}=g_{ij}(y)$, $V^\flat=(2cx^i-b^i)g_{ij}dx^j$. Together with (5-1) we have

$$dV^\flat=(2cx^i-b^i)\frac{\partial g_{ij}}{\partial y^k}\delta y^k\wedge dx^j.$$

It follows that $dV^\flat|_{HTM\times HTM}=0$. Thus V is closed. Define a new Finsler metric $\widetilde{F}$ by (1-2). Note that any Minkowski manifold is of Landsberg type. By Theorem 1.1, we have that (1-3) holds. Define a 1-parameter transformation ψ_t on Ω by $\psi_t(x):=e^{-2ct}x-tb$. Hence $(d\psi_t(x)/dt)\,|_{t=0}=-2cx-b=-V$. Since Φ is Minkowskian, it is locally projectively flat. This implies that the unit speed geodesic through x with tangent vector $y(\neq 0)$ is $\gamma(t) := x + (t/\varphi(y))y$. Theorem 1.1 in [Huang and Mo 2011] implies that a geodesic of $\widetilde{F}$ is given by $e^{-2ct}[x+y((e^{2ct}-1)/(2c\varphi(y)))]-tb$ (resp. $x+(t/\varphi(y))y-tb$) for $c \neq 0$ (resp. $c = 0$). □

References

[Bao et al. 2000] D. Bao, S.-S. Chern, and Z. Shen, *An introduction to Riemann–Finsler geometry*, Graduate Texts in Math. **200**, Springer, 2000. MR Zbl

[Blair 2002] D. E. Blair, *Riemannian geometry of contact and symplectic manifolds*, Progr. Math. **203**, Birkhäuser, Boston, 2002. MR Zbl

[Cheng and Shen 2009] X. Cheng and Z. Shen, "Randers metrics of scalar flag curvature", *J. Aust. Math. Soc.* **87**:3 (2009), 359–370. MR Zbl

[Chern and Shen 2005] S.-S. Chern and Z. Shen, *Riemann–Finsler geometry*, Nankai Tracts in Math. **6**, World Sci., Hackensack, NJ, 2005. MR Zbl

[Huang and Mo 2011] L. Huang and X. Mo, "On geodesics of Finsler metrics via navigation problem", *Proc. Amer. Math. Soc.* **139**:8 (2011), 3015–3024. MR Zbl

[Huang and Mo 2013] L. Huang and X. Mo, "On conformal fields of a Randers metric with isotropic S-curvature", *Illinois J. Math.* **57**:3 (2013), 685–696. MR Zbl

[Huang and Mo 2015] L. Huang and X. Mo, "On the flag curvature of a class of Finsler metrics produced by the navigation problem", *Pacific J. Math.* **277**:1 (2015), 149–168. MR Zbl

[Mo 2011] X. Mo, "On some Finsler metrics of non-positive constant flag curvature", *Results Math.* **60**:1-4 (2011), 475–485. MR Zbl

[Mo 2012] X. Mo, "On some Finsler metrics of constant (or scalar) flag curvature", *Houston J. Math.* **38**:1 (2012), 41–54. MR Zbl

[Mo and Huang 2007] X. Mo and L. Hang, "On curvature decreasing property of a class of navigation problems", *Publ. Math. Debrecen* **71**:1-2 (2007), 141–163. MR Zbl

[Shen 2001] Z. Shen, *Lectures on Finsler geometry*, World Sci., Singapore, 2001. MR Zbl

[Shen 2004] Z. Shen, "Landsberg curvature, S-curvature and Riemann curvature", pp. 303–355 in *A sampler of Riemann–Finsler geometry*, edited by D. Bao et al., Math. Sci. Res. Inst. Publ. **50**, Cambridge Univ. Press, 2004. MR Zbl

Received February 9, 2017. Revised June 30, 2018.

LIBING HUANG
NANKAI UNIVERSITY
CHINA
huanglb@nankai.edu.cn

HUAIFU LIU
BEIJING UNIVERSITY OF TECHNOLOGY
CHINA
liuhf@bjut.edu.cn

XIAOHUAN MO
PEKING UNIVERSITY
CHINA
moxh@pku.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.97

SYMPLECTIC AND ODD ORTHOGONAL PFAFFIAN FORMULAS FOR ALGEBRAIC COBORDISM

THOMAS HUDSON AND TOMOO MATSUMURA

In the Chow ring of symplectic/odd orthogonal Grassmann bundles the degeneracy loci classes can be expressed as a sum of Schur–Pfaffians. An analogous Schur–Pfaffian formula was obtained for K-theory by the authors together with T. Ikeda and M. Naruse. Here we generalize this explicit formula of degeneracy loci classes to algebraic cobordism, which is universal among all oriented cohomology theories.

1. Introduction

The r-th *degeneracy locus* for a morphism of vector bundles $\varphi : E \to F$ over a smooth quasi-projective scheme M is the subvariety X_r of M consisting of all the points at which the rank of φ is at most r. Assuming φ to be sufficiently general, the classical Giambelli–Thom–Porteous formula describes the Chow ring fundamental class $[X_r]$ as a Schur-determinant in the Chern classes of E and F. Similarly, one can consider more restrictive settings in which φ is either *skewsymmetric* or *symmetric*. In both cases $[X_r]$ is given as a Schur-Pfaffian instead of a Schur-determinant. A more general family of degeneracy loci can be constructed by considering flags of subbundles of E and F and imposing multiple rank conditions.

Fundamental examples of these loci are the Schubert varieties of *isotropic* Grassmannians. The *isotropic* Grassmannian consists of subspaces on which a given symplectic or odd orthogonal form vanishes identically. Inside this ambient space, the degeneracy loci correspond to the Schubert varieties indexed by the combinatorial objects known as *k-strict* partitions.

Pragacz [1991] considered the maximal isotropic case and showed that the Chow ring fundamental classes of Schubert varieties can be expressed through a *Schur-Pfaffian* formula. Kazarian [2000] generalised Pragacz's formula to general degeneracy loci (compare [Ikeda 2007]). Buch, Kresh and Tamvakis [Buch et al. 2017] obtained a *theta polynomial* formula for the non-maximal isotropic Grassmannians, which can also be written as a sum of Schur–Pfaffian. Wilson [2010] conjectured an analogous formula for general degeneracy loci, which was proved

MSC2010: primary 14M15, 55N22; secondary 05E05, 14C17.
Keywords: generalised Schubert calculus, algebraic cobordism, pfaffian, isotropic grassmannian.

in [Ikeda and Matsumura 2015] (compare [Tamvakis and Wilson 2016; Anderson and Fulton 2018]).

In recent years, following the trend of generalised Schubert calculus, there has been an attempt to lift results as the ones above from the Chow ring CH^* to other functors like *connective K-theory* CK^* and *algebraic cobordism* Ω^*, highlighting the role played in the formulas by the associated formal group law F and formal inverse χ. In [Hudson et al. 2017], together with T. Ikeda and H. Naruse, we generalized aforementioned results for CH^* to CK^*, and established a Pfaffian-sum formula describing the degeneracy loci classes of symplectic and odd orthogonal Grassmann bundles in CK^*. The goal of this paper is to further extend these formulas to Ω^*.

We begin by explaining our results in the symplectic case. Let $E \to X$ be a vector bundle of rank $2n$ with a nowhere vanishing skewsymmetric form and fix a non-negative integer $k \le n$. Consider the *symplectic Grassmann bundle* $SG^k(E)$ whose fiber at $x \in X$ is the Grassmannian of $(n-k)$-dimensional isotropic subspaces of E_x. For each k-strict partition λ, there is the degeneracy locus $X_\lambda \subset SG^k(E)$. Following [Kazarian 2000], we can construct a resolution of singularities $\varpi : Y_\lambda \to X_\lambda$ inside of a certain flag bundle over $SG^k(E)$. In CH^* or CK^*, the fundamental class of X_λ is well-defined and it coincides with the pushforward $\varpi_*[Y_\lambda]$ of the fundamental class of Y_λ along ϖ. *However, in algebraic cobordism, not all degeneracy loci have a well-defined notion of fundamental class. Hence we consider* $\varpi_*[Y_\lambda]$ *as a replacement of* $[X_\lambda]$. As in [Hudson et al. 2017], the fundamental class $[Y_\lambda]$ can be expressed as a product of top Chern classes of certain bundles. In our previous paper [Hudson and Matsumura 2019], we developed a technique to compute the pushforward of such classes along a flag bundle in terms of *relative Segre classes* of vector bundles. With that method at our disposal, we are able to obtain the following description of the class $\varpi_*[Y_\lambda]$ as our main result. The tautological isotropic subbundle of $SG^k(E)$ is denoted by U and the subbundles

$$0 = F^n \subset F^{n-1} \subset \cdots \subset F^1 \subset F^0 \subset F^{-1} \subset \cdots \subset F^{-n} = E,$$

form the reference flag used to define the degeneracy loci. In $\Omega^*(SG^k(E))$, we consider the relative Segre classes

$$\mathscr{C}_m^{(\ell)} := \mathscr{S}_m\left(U^\vee - (E/F^\ell)^\vee)\right) \qquad (\forall m \in \mathbb{Z}, -n \le \forall \ell \le n).$$

Main Theorem (Theorem 3.9). *Let* $\lambda = (\lambda_1, \dots, \lambda_r)$ *be a k-strict partition such that* $r \le n-k$ *and* $\lambda_1 \le n+k$, *and let* $\boldsymbol{\chi} = (\chi_1, \dots, \chi_r)$ *be its characteristic index* (see (3-1)). *In* $\Omega^*(SG^k(E))$, *we have*

$$(1\text{-}1) \qquad [Y_\lambda \to SG^k(E)] := \varpi_*[Y_\lambda] = \sum_{\mathbf{s}=(s_1,\dots,s_r)\in\mathbb{Z}^r} f_{\mathbf{s}}^\lambda \mathscr{C}_{\lambda_1+s_1}^{(\chi_1)} \cdots \mathscr{C}_{\lambda_r+s_r}^{(\chi_r)}.$$

Here $c^{\lambda}_{\mathbf{s}} \in \mathbb{L}$ are the coefficients of the Laurent series expansion

$$\frac{\prod_{1\le i<j\le r}(1-t_i/t_j)P(t_j,t_i)}{\prod_{(i,j)\in C(\lambda)}(1-\chi(t_i)/t_j)P(t_j,\chi(t_i))} = \sum_{\mathbf{s}=(s_1,\dots,s_r)\in\mathbb{Z}^r} f^{\lambda}_{\mathbf{s}}\cdot t_1^{s_1}\cdots t_r^{s_r}, \tag{1-2}$$

where $C(\lambda) := \{(i,j) \mid 1\le i<j\le r,\ \ \chi_i+\chi_j\ge 0\}$ and $P(u,v)$ is the unique power series satisfying $F(u,\chi(v)) = (u-v)P(u,v)$.

Now consider the odd orthogonal Grassmann bundle $OG^k(E)$, with E of rank $2n+1$ and each fiber being an orthogonal Grassmannian of $(n-k)$-dimensional isotropic subspaces. The essential difference with the previous situation is that it is far more complex to deal with the case of quadric bundles $OG^{n-1}(E) = Q(E)$, the orthogonal analogue of projective bundles. Let the reference flag be denoted by

$$0 = F^n \subset F^{n-1}\subset\cdots\subset F^1\subset F^0\subset (F^0)^{\perp}\subset F^{-1}\subset\cdots\subset F^{-n} = E.$$

The fundamental classes of the Schubert varieties $X_{(\lambda_1)}$ are actually well-defined in $\Omega^*(Q(E))$ and, as elements of $\Omega^*(Q(E))\otimes_{\mathbb{Z}}\mathbb{Z}[1/2]$, they are given by

$$[X_{\lambda_1}\to Q(E)] = \mathscr{B}^{(\chi_1)}_{\lambda_1} := \begin{cases} \mathscr{S}_{\lambda_1}\big(U^{\vee}-(E/F^{\chi_1})^{\vee}\big) & (0\le\lambda_1<n),\\ \dfrac{1}{F^{(2)}\big(c_1(U^{\vee})\big)}\mathscr{S}_{\lambda_1}\big(U^{\vee}-(E/F^{\chi_1})^{\vee}\big) & (n\le\lambda_1<2n),\end{cases}$$

where $F^{(2)}(u)$ is the power series defined by the equation $F(u,u) = u\cdot F^{(2)}(u)$. More generally, the pushforward classes $[Y_\lambda\to OG^k(E)]$ are obtained from (1-1) by replacing $\mathscr{C}^{(i)}_m$ with $\mathscr{B}^{(i)}_m$ (see Theorem 4.12).

A key aspect of algebraic cobordism, which was established in [Levine and Morel 2007], is its universality. In particular, this means that formulas which hold for Ω^* can be specialised to any other oriented cohomology theory. An easy example of this phenomenon is illustrated by the behaviour of the first Chern class of line bundles. In CH^* one has

$$c_1(L\otimes M) = c_1(L)+c_1(M)\ \text{ and }\ c_1(L^{\vee}) = -c_1(L),$$

while in CK^* these equalities become

$$c_1(L\otimes M) = c_1(L)+c_1(M)-\beta c_1(L)c_1(M)\ \text{ and }\ c_1(L^{\vee}) = \frac{-c_1(L)}{1-\beta c_1(L)},$$

where $\beta\in CK^*(\operatorname{Spec}\mathbf{k})$ is the pushforward of the fundamental class of $\mathbb{P}^1$ to the point. If we set $\beta = 0$, we recover the identities for CH^*. In algebraic cobordism Ω^*, the expressions describing $c_1(L\otimes M)$ and $c_1(L^{\vee})$ are respectively given by the *universal formal group law* $F(u,v)$ and the *universal formal inverse* $\chi(u)$ which are certain power series with coefficients in $\Omega(\operatorname{Spec}\mathbf{k})$. The universality of Ω^* implies that in any other oriented cohomology theory A^*, $c_1(L\otimes M)$ and $c_1(L^{\vee})$ can be obtained by specializing the coefficients of $F(u,v)$ and $\chi(u)$ to $A^*(\operatorname{Spec}\mathbf{k})$.

In particular, in CK^* we have $P(u,v)=\frac{1}{1-\beta v}$ and $\chi(u)=\frac{-u}{1-\beta u}$ and the Laurent series (1-2) can be expressed as a sum of Pfaffians ([Hudson et al. 2017, Lemma 5.18]). As a consequence (1-1) reduces to the Pfaffian sum formula describing the K-theoretic degeneracy loci classes [Hudson et al. 2017, Theorem 5.20].

Our choice of resolutions Y_λ has the advantage of being stable: the class $\varpi_*[Y_\lambda]$ doesn't change when $n\to\infty$. On the other hand, there are different resolutions for X_λ, such as *Bott–Samelson resolutions*. These resolutions are well-studied in the context of generalized Schubert calculus. On the one hand the advantage of Bott–Samelson classes is represented by their compatibility with divided difference operators, however this comes at the cost of not being stable along the limit $n\to\infty$. See, for example, [Hornbostel and Kiritchenko 2011; Kiritchenko and Krishna 2013; Hornbostel and Perrin 2018]. The classes related to other resolutions are also studied in [Nakagawa and Naruse 2016; 2018], a Hall–Littlewood type formula in Ω^* is derived. All of these resolution classes coincide with honest Schubert classes if one works in CK^*, while they form different families of classes in Ω^*. As an application of our explicit formulas, it would be interesting to compare those different classes which replace Schubert classes in algebraic cobordism. To this aim it would be advisable to first consider functors more suitable for computations like the infinitesimal theories used in [Hudson and Matsumura 2018].

Anderson [2019] extended the results of [Hudson et al. 2017] to more general degeneracy loci including those arising from even orthogonal Grassmann bundles. His work is based on the approach he and Fulton employed in their study of the Chow ring fundamental classes of degeneracy loci for all types [Anderson and Fulton 2012; 2018]. In our future work we would like to lift Anderson's results to Ω^* so to cover the even orthogonal case as well.

The organisation of the paper is as follows. In section 2 we recall some basic facts about Borel–Moore homology theories and we translate into this setting the results on Segre classes presented in [Hudson and Matsumura 2019]. This becomes necessary because the resolutions are not smooth in general. In section 3 we prove the main theorem for symplectic Grassmann bundles, while in section 4 we first deal with the special case of quadric bundles and then establish the main theorem for odd orthogonal Grassmann bundles.

Notations and conventions. Throughout this paper $\mathbf{k}$ will be a field of characteristic 0. By $\mathbf{Sch_k}$ we will denote the category of separated schemes of finite type over $\mathbf{k}$ and $\mathbf{Lci_k}$ will stand for its full subcategory constituted by the objects whose structural morphism is a local complete intersection. For a given category C we will write C' to refer to its subcategory given by allowing only projective morphisms. $\mathbf{Ab}_*$ represents the category of graded abelian groups.

2. Preliminaries

The goal of this section is to collect some basic properties of Borel–Moore homology theories and to translate in this context some of the results on generalised Segre classes presented in [Hudson and Matsumura 2019].

Borel–Moore homology theories. An oriented Borel–Moore (BM) homology theory on $\mathbf{Sch_k}$ (or *mutatis mutandis* on $\mathbf{Lci_k}$) is given by a covariant functor $A_* : \mathbf{Sch}'_{\mathbf{k}} \to \mathbf{Ab}_*$, by a family of pullback maps $\{f^* : A_*(Y) \to A_*(X)\}$ associated to *l.c.i.* morphism and by an external product $A_*(X) \otimes A_*(Y) \to A_*(X \times_{\mathrm{Spec}\ k} Y)$. Let us remind the reader that a morphism is a *local complete intersection* if and only if it can be factored as the composition of a regular embedding and a smooth morphism. A detailed description of the properties that these three components have to satisfy would force us to take a significant detour, so we will focus only on the aspects that are more relevant to our work and refer the reader to [Levine and Morel 2007, Definition 5.1.3] for the precise definition.

For us the most relevant feature of oriented BM homology theories is that they satisfy the projective bundle formula. Roughly speaking it states that for every vector bundle E of rank e with $X \in \mathbf{Sch_k}$, the evaluation of A_* on the associated dual projective bundle $\mathbb{P}^*(E) \xrightarrow{q} X$ can be described in terms of $A_*(X)$. More precisely for $i \in \{0, 1, \dots, e-1\}$ one has operations

$$\xi^{(i)} : A_{*+i-e+1}(X) \longrightarrow A_*(\mathbb{P}^*(E))$$

given by $\xi^{(i)} := \tilde{c}_1(\mathcal{Q})^i \circ q^*$, where $\mathcal{Q} \to \mathbb{P}^*(E)$ is the tautological line bundle and $\tilde{c}_1(\mathcal{Q}) := s^* \circ s_*$, for any section $s : \mathbb{P}^*(E) \to \mathcal{Q}$. Altogether these yield the following isomorphism

$$\Psi : \bigoplus_{i=0}^{e-1} A_{*+i-e+1}(X) \xrightarrow{\sum_{i=0}^{e-1} \xi^{(i)}} A_*(\mathbb{P}^*(E)).$$

A very important consequence of this is that every oriented BM homology theory admits a theory of Chern class operators: to E one associates $\{\tilde{c}_i^A(E) : A_*(X) \to A_{*-i}(X)\}_{0 \le i \le e}$. These are defined by setting $\tilde{c}_0^A(E) = \mathrm{id}_{A_*(X)}$ and, up to a sign, by considering the different components of $\Psi^{-1} \circ \xi^{(e)}$, so that one obtains the relation

$$\sum_{i=0}^{e} (-1)^i \xi^{(e-i)} \circ \tilde{c}_i^A(E) = 0.$$

These operators can be collected in the so-called *Chern polynomial* $\tilde{c}^A(E; u) := \sum_{i=0}^{e} \tilde{c}_i^A(E) u^i$ and it is worth mentioning that, in view of the Whitney formula, its definition can be extended to the Grothendieck group of vector bundles by setting

$$\tilde{c}^A(E - F; u) := \frac{\tilde{c}^A(E; u)}{\tilde{c}^A(F; u)}.$$

Beside being extremely useful for computations, Chern classes allow one to get some insight on how a general oriented BM homology theory A_* differs from the Chow group CH_*, probably the most commonly known example. Let us consider, as an example, the behaviour of the first Chern class with respect to the tensor product of two line bundles L and M. While in CH_* one has

$$\tilde{c}_1^{CH}(L \otimes M) = \tilde{c}_1^{CH}(L) + \tilde{c}_1^{CH}(M),$$

in general the relation between the three Chern class operators is described by a *formal group law* $\big(A_*(\operatorname{Spec} k), F_A\big)$, where $F_A(u, v)$ is a special power series with coefficients in the coefficient ring of the theory $A_*(\operatorname{Spec} k)$. The precise relation is given by

$$\tilde{c}_1^A(L \otimes M) = F_A(\tilde{c}_1^A(L), \tilde{c}_1^A(M)).$$

In a similar fashion, whereas in CH_* one simply has $\tilde{c}_1^{CH}(L^\vee) = -\tilde{c}_1^{CH}(L)$, in general one needs to introduce the *formal inverse* χ_A, a power series in one variable satisfying both

$$\tilde{c}_1^A(L^\vee) = \chi_A(\tilde{c}_1(L)) \quad \text{and} \quad F_A(u, \chi_A(u)) = 0.$$

In some case we will denote the formal inverse $\chi_A(u)$ simply by $\bar{u}$.

All our computations will take place in the algebraic cobordism of Levine–Morel Ω_* and our choice is motivated by the following fundamental result.

Theorem 2.1 [Levine and Morel 2007, Theorems 7.1.3 and 4.3.7]. *The algebraic cobordism Ω_* is universal among oriented BM homology theories on* $\mathbf{Lci_k}$. *That is, for any other oriented BM homology theory A_* there exists a unique morphism*

$$\vartheta_A : \Omega_* \to A_*$$

of oriented BM homology theories. Furthermore, its associated formal group law $\big(\Omega_*(\operatorname{Spec} \mathbf{k}), F_\Omega\big)$ *is isomorphic to the universal one defined on the Lazard ring* $(\mathbb{L}, F)$.

One consequence of this universality is that all the formulas obtained for Ω_* can be specialised to every other oriented BM homology theory A_*. In other words, algebraic cobordism allows one to work with all theories at once. Since we will only work with algebraic cobordism, in the remainder of the paper we will remove the subscript $_\Omega$ from the notation.

Let us conclude our general discussion by briefly mentioning the construction of fundamental classes and some results which can be used to compute them. To every $X \in \mathbf{Sch_k}$ whose structural morphism π_X is l.c.i. we associate its fundamental class by setting $1_X := \pi_X^*(1)$. Notice that here 1 stands for the multiplicative unit in $A_*(\operatorname{Spec}\mathbf{k})$. In the special case of the zero scheme of a bundle, the fundamental class can be computed via the following lemma.

Lemma 2.2 [Levine and Morel 2007, Lemma 6.6.7]. *Let E be a vector bundle of rank e over $X \in \mathbf{Sch_k}$. Suppose that E has a section $s : X \to E$ such that the zero scheme of s, $i : Z \to X$ is a regularly embedded closed subscheme of codimension e. Then we have*

$$\tilde{c}_e(E) = i_* \circ i^*.$$

In particular, if X is an l.c.i. scheme, we have

$$\tilde{c}_e(E)(1_X) = i_*(1_Z).$$

Finally, as it will play an important role in our computations, we would like to make more explicit the case of the fundamental class of a nonreduced divisor. For this we will require a bit of notation. For every integer $n \geq 2$, let $n \cdot_{F_A} u$ be the *formal multiplication* by n, that is, the power series obtained by adding n times the variable u using the formal group law F_A. Since F_A is a formal group law, one has

$$n \cdot_{F_A} u = u \cdot F_A^{(n)}(u) \tag{2-1}$$

for some degree 0 power series $F_A^{(n)}(u)$ whose costant term is n. We are now able to restate [Levine and Morel 2007, Proposition 7.2.2] for the particular case we will need.

Lemma 2.3. *Let W be a smooth scheme and D a smooth prime divisor of W. For any integer $n \geq 2$, let $|E|$ be the closed subscheme associated to the divisor $E = nD$. If L is the line bundle corresponding to D and $\iota : D \to |E|$ is the natural morphism, then in $A_*(|E|)$ we have*

$$1_{|E|} = \iota_*\big(F_A^{(n)}(\tilde{c}_1^A(L_{|D}))(1_D)\big),$$

where $L_{|D}$ is the restriction of L to D.

Segre class operators. In [Hudson and Matsumura 2019], in order to be able to describe the pushforwards along projective bundles over a smooth scheme, we generalised to algebraic cobordism the classical definition of Segre classes given in [Fulton 1998]. As in this paper we deal with the resolutions of symplectic or orthogonal degeneracy loci, it becomes necessary to extend such description to the case of projective bundles over non-smooth schemes. Therefore, we will now introduce Segre class operators for oriented BM homology theories, since these can be defined for more general schemes.

Following [Hudson et al. 2017, §4], we define the relative Segre operators in terms of generating functions. Let $X \in \mathbf{Sch_k}$.

Definition 2.4. Let $\tilde{x}_1, \dots, \tilde{x}_e$ be Chern root operators of a vector bundle E over X so that $\tilde{c}(E; u) = \prod_{i=1}^{e}(1 + \tilde{x}_i u)$. We define

$$\widetilde{w}(E; u) = \sum_{s \geq 0}^{\infty} \widetilde{w}_{-s}(E)u^{-s} = \prod_{i=1}^{e} P(u^{-1}, \tilde{x}_i),$$

where $P(u, v)$ is defined by $F(u, \chi(v)) = (u - v)P(u, v)$ (compare [Hudson and Matsumura 2019, Lemma 4.1]). Since the right-hand side is symmetric in the $\tilde{x}_i$, this definition of $\widetilde{w}_{-s}(E)$ is independent of the choice of Chern root operators of E. It should be noticed that $\widetilde{w}_0(E)$ has constant term 1 and as a consequence $\widetilde{w}(E; u)$ is an invertible power series in u^{-1}. One can also define $\widetilde{w}(E - F; u)$ for a virtual bundle $[E - F]$, where E and F are vector bundles over X, by setting

$$\widetilde{w}(E - F; u) = \sum_{s \geq 0}^{\infty} \widetilde{w}_{-s}(E - F)u^{-s} = \frac{\widetilde{w}(E; u)}{\widetilde{w}(F; u)}.$$

Definition 2.5. Let E be a vector bundle of rank e over X and n a nonnegative integer. Consider the dual projective bundle $\pi : \mathbb{P}^*(E \oplus O_X^{\oplus n}) \to X$ where O_X is the trivial line bundle over X. For every integer $m \geq -e - n + 1$, define the degree *m Segre class operator* $\widetilde{\mathscr{S}}_m(E)$ of E by setting

$$\widetilde{\mathscr{S}}_m(E) = \pi_* \circ \tilde{c}_1(\mathcal{Q})^{m+e+n-1} \circ \pi^*, \tag{2-2}$$

where $\mathcal{Q}$ is the tautological quotient line bundle of $\mathbb{P}^*(E \oplus O_X^{\oplus n})$. It is easy to verify (see [Hudson and Matsumura 2019, Remark 4.4]) that this assignment is independent of n. Finally, we set

$$\widetilde{\mathscr{S}}(E; u) := \sum_{m \in \mathbb{Z}} \widetilde{\mathscr{S}}_m(E)u^m.$$

Proposition 2.6. *Let $E \to X$ be a vector bundle of rank e over $X \in \mathbf{Sch_k}$. Then we have the following equality of power series*:

$$\widetilde{\mathscr{S}}(E; u) = \frac{\widetilde{\mathscr{P}}(u)}{\tilde{c}(E; -u)\widetilde{w}(E; u)}.$$

Here $\widetilde{\mathscr{P}}(u) := \sum_{i=0}^{\infty} \widetilde{[\mathbb{P}^i]} u^{-i}$ is the power series collecting the operators given by external multiplication with the pushforwards classes of projective spaces $[\mathbb{P}^i] := [\mathbb{P}^i \to \operatorname{Spec} k] \in \mathbb{L}^{-i}$.

Proof. Once one has translated in the language of operators the proof given in [Hudson and Matsumura 2019, Theorem 4.6], the only thing left to check is that for every trivial dual projective bundle $(\mathbb{P}^n_X)^* \xrightarrow{\pi} X$ the composition $\pi_* \circ \pi^*$ coincides with external multiplication by $[\mathbb{P}^n]$. This can be verified directly at the level of cobordism cycles by making use of the definitions of pushforward and pullback morphisms and of the external product. □

Remark 2.7. It is worth mentioning that, provided one restricts to the case $X \in \mathbf{Sm_k}$, Proposition 2.6 can be derived from the analogue of Quillen's formula for algebraic cobordism established in [Vishik 2007, Theorem 5.35]. The same formula can be used to express the classes $[\mathbb{P}^i]$ in terms of the generators of the Lazard ring and, as a consequence, of the coefficients of the formal group law. On the other hand, an easy computation shows that Quillen's formula can be recovered from Proposition 2.6, provided one knows the expression for the classes of projective spaces. In this sense our approach allows us to extend the validity of Vishik's result from $\mathbf{Sm_k}$ to $\mathbf{Sch_k}$.

In view of the last proposition, we are now able to extend to virtual bundles the definition of Segre classes.

Definition 2.8. For vector bundles E and F over X, define the *relative Segre class operators* $\widetilde{\mathscr{S}}_m(E-F)$ on $\Omega_*(X)$ as

$$\widetilde{\mathscr{S}}(E-F;u) := \sum_{m\in\mathbb{Z}} \widetilde{\mathscr{S}}_m(E-F)u^m = \widetilde{\mathscr{P}}(u)\frac{\tilde{c}(F;-u)\widetilde{w}(F;u)}{\tilde{c}(E;-u)\widetilde{w}(E;u)}. \tag{2-3}$$

Remark 2.9. If the rank of F is f, then we have

$$\widetilde{\mathscr{S}}_m(E-F) = \sum_{p=0}^{f}\sum_{q=0}^{\infty}(-1)^p\tilde{c}_p(F)\circ\widetilde{w}_{-q}(F)\circ\widetilde{\mathscr{S}}_{m-p+q}(E).$$

Even if F itself is a virtual bundle, this equation holds by replacing f with ∞.

We conclude this section by providing a description of relative Segre classes in terms of pushforwards of Chern classes. This should be seen as an analogue of [Hudson and Matsumura 2019, Theorem 4.9].

Theorem 2.10. *Let $X \in \mathbf{Sch_k}$ and let E and F be two vector bundles over X, respectively of rank e and f. Let $\pi : \mathbb{P}^*(E) \to X$ be the dual projective bundle of E and $\mathcal{Q}$ its universal quotient line bundle. As operators over $\Omega_*(X)$, we have*

$$\pi_* \circ \tilde{c}_1(\mathcal{Q})^s \circ \tilde{c}_f(\mathcal{Q}\otimes F^\vee)\circ\pi^* = \widetilde{\mathscr{S}}_{s+f-e+1}(E-F). \tag{2-4}$$

In particular if $X \in \mathbf{Lci_k}$, then one has

$$\pi_* \circ \tilde{c}_1(\mathcal{Q})^s \circ \tilde{c}_f(\mathcal{Q}\otimes F^\vee)(1_{\mathbb{P}^*(E)}) = \widetilde{\mathscr{S}}_{s+f-e+1}(E-F)(1_X).$$

Proof. Let us begin by observing that an easy Chern roots computation analogue to [Hudson and Matsumura 2019, formula (4.1)] gives us

$$\tilde{c}_f(\mathcal{Q}\otimes F^\vee)=\sum_{p=0}^{f}\sum_{q=0}^{\infty}(-1)^p\tilde{c}_p(F)\circ\tilde{w}_{-q}(F)\circ\tilde{c}_1(\mathcal{Q})^{f-p+q}.$$

Thus the left-hand side of (2-4) can be rewritten as

$$\sum_{p=0}^{f}\sum_{q=0}^{\infty}(-1)^p\tilde{c}_p(F)\circ\tilde{w}_{-q}(F)\circ\pi_*\circ\tilde{c}_1(\mathcal{Q})^{s+f-p+q}\circ\pi^*. \tag{2-5}$$

By (2-2), we find that (2-5) equals to

$$\sum_{p=0}^{f}\sum_{q=0}^{\infty}(-1)^p\tilde{c}_p(F)\circ\tilde{w}_{-q}(F)\circ\widetilde{\mathscr{S}}_{s+f-e+1-p+q}(E),$$

which coincides with the right-hand side of (2-4) in view of Remark 2.9. The second statement follows immediately by applying both sides of (2-4) to the fundamental class 1_X. □

Remark 2.11. If E is a line bundle, then one has $\pi=\mathrm{id}_X$ and $\mathcal{Q}=E$. As a consequence we have

$$\tilde{c}_1(\mathcal{Q})^s\circ\tilde{c}_f(\mathcal{Q}\otimes F^\vee)(1_X)=\widetilde{\mathscr{S}}_{f-e+1+s}(E-F)(1_X).$$

3. Symplectic degeneracy loci

For this section we fix a nonnegative integer k.

k-strict partitions and characteristic indices. A k-strict partition λ is a weakly decreasing infinite sequence $(\lambda_1,\lambda_2,\dots)$ of nonnegative integers such that the number of nonzero parts is finite, and if $\lambda_i>k$, then $\lambda_i>\lambda_{i+1}$. The length of λ is the number of nonzero parts of λ. Let $\mathcal{SP}^k$ be the set of all k-strict partitions. Let $\mathcal{SP}^k_r$ be the set of all k-strict partitions with the length at most r. If $\lambda\in\mathcal{SP}^k_r$, then we often write $\lambda=(\lambda_1,\dots,\lambda_r)$. Let $\mathcal{SP}^k(n)$ be the set of all k-strict partitions such that $\lambda_1\le n+k$ and the length of λ is at most $n-k$.

Let W_∞ be the infinite hyperoctahedral group which can be identified with the group of all signed permutations (permutations w of $\mathbb{Z}\backslash\{0\}$ such that $w(i)\neq i$ for only finitely many $i\in\mathbb{Z}\backslash\{0\}$, and $\overline{w(i)}=w(\bar{i})$ for all i where $\bar{i}:=-i$). A signed permutation w is determined by the sequence $(w(1),w(2),\dots)$ which we call one line notation. An element $w\in W_\infty$ is called k-Grassmannian if

$$0<w(1)<\cdots<w(k),\quad w(k+1)<w(k+2)<\cdots.$$

The set of all k-Grassmannian elements in W_∞ is denoted by $W_\infty^{(k)}$.

Between $W_\infty^{(k)}$ and $\mathcal{SP}^k$, there is a bijection defined as follows. For each $w \in W_\infty^{(k)}$, the corresponding k-strict partition is given by

$$\lambda_i := \begin{cases} w(k+i) & \text{if } w(k+i) < 0, \\ \#\{j \le k \mid w(j) > w(k+i)\} & \text{if } w(k+i) > 0. \end{cases}$$

For each $\lambda \in \mathcal{SP}^k$ (with the corresponding $w \in W_\infty^{(k)}$), we define its characteristic index $\boldsymbol{\chi} = (\chi_1, \chi_2, \dots)$ by

$$\chi_i := \begin{cases} -w(k+i) - 1 & \text{if } w(k+i) < 0, \\ -w(k+i) & \text{if } w(k+i) > 0. \end{cases} \tag{3-1}$$

Moreover, the following notations are necessary for our formulas of Grassmannian degeneracy loci in type C and B: for each $\lambda \in \mathcal{SP}^k$ and the corresponding characteristic index $\boldsymbol{\chi}$, define

$$C(\lambda) := \{(i, j) \mid 1 \le i < j, \quad \chi_i + \chi_j \ge 0\},$$
$$\gamma_j := \sharp\{i \mid 1 \le i < j, \quad \chi_i + \chi_j \ge 0\} \quad \text{for each } j > 0.$$

Symplectic degeneracy loci and the class κ_λ^C. Let E be a symplectic vector bundle over a smooth scheme X of rank $2n$, i.e., we are given a nowhere degenerating section of $\bigwedge^2 E$. For a subbundle F of E, we denote by $F^\perp$ the orthogonal complement of F with respect to the symplectic form. Fix a reference flag $F^\bullet$ of subbundles of E,

$$0 = F^n \subset F^{n-1} \subset \cdots \subset F^1 \subset F^0 \subset F^{-1} \subset \cdots \subset F^{-n} = E,$$

where $\operatorname{rk} F^i = n - i$ and $(F^i)^\perp = F^{-i}$ for all i with $-n \le i \le n$. Let $SG^k(E) \to X$ be the Grassmannian bundle over X consisting of pairs (x, U_x) where $x \in X$ and U_x is an $n-k$ dimensional isotropic subspace of E_x. Let U be the tautological bundle of $SG^k(E)$.

For each $\lambda \in \mathcal{SP}^k(n)$ of length r, let X_λ^C be the symplectic degeneracy locus in $SG^k(E)$ defined by

$$X_\lambda^C = \left\{(x, U_x) \in SG^k(E) \mid \dim(U_x \cap F_x^{\chi_i}) \ge i, \quad i = 1, \dots, r\right\},$$

where $\boldsymbol{\chi} = (\chi_1, \chi_2, \dots)$ is the characteristic index for λ.

Let $Fl_r(U) \to SG^k(E)$ be the r-step flag bundle of U where the fiber at $(x, U_x) \in SG^k(E)$ consists of the flag $(D_\bullet)_x = \{(D_1)_x \subset \cdots \subset (D_r)_x\}$ of subspaces of U_x with $\dim(D_i)_x = i$. Let $D_1 \subset \cdots \subset D_r$ be the flag of tautological bundles of $Fl_r(U)$. We set $D_0 = 0$. The bundle $Fl_r(U)$ can be constructed as a tower of projective bundles

$$\begin{aligned} \pi : Fl_r(U) = \mathbb{P}(U/D_{r-1}) &\xrightarrow{\pi_r} \mathbb{P}(U/D_{r-2}) \xrightarrow{\pi_{r-1}} \cdots \\ &\xrightarrow{\pi_3} \mathbb{P}(U/D_1) \xrightarrow{\pi_2} \mathbb{P}(U) \xrightarrow{\pi_1} SG^k(E). \end{aligned} \tag{3-2}$$

The quotient line bundle D_j/D_{j-1} is regarded as the tautological line bundle of $\mathbb{P}(U/D_{j-1})$ and we set $\tilde{\tau}_j := \tilde{c}_1((D_j/D_{j-1})^\vee)$.

We are now able to define the resolution of singularities of the degeneracy loci.

Definition 3.1. For each $j = 1, \ldots, r$, we define a subvariety Y_j of $\mathbb{P}(U/D_{j-1})$ by

$$Y_j := \big\{(x, U_x, (D_1)_x, \ldots, (D_j)_x) \in \mathbb{P}(U/D_{j-1}) \mid (D_i)_x \subset F_x^{\chi_i},\ i = 1, \ldots, j\big\}.$$

We set $Y_0 := SG_r^k(U)$ and $Y_\lambda^C := Y_r$. Let $P_{j-1} := \pi_j^{-1}(Y_{j-1})$, $\pi'_j : P_{j-1} \to Y_{j-1}$ the projection and $\iota_j : Y_j \to P_{j-1}$ the obvious inclusion. Let $\varpi_j := \pi'_j \circ \iota_j$. We have the commutative diagram

$$\begin{array}{ccc}
\mathbb{P}(U/D_{j-1}) & \xrightarrow[\pi_j]{} & \mathbb{P}(U/D_{j-2}) \\
\uparrow & & \uparrow \\
P_{j-1} & \xrightarrow[\pi'_j]{} & Y_{j-1} \\
\uparrow{\scriptstyle \iota_j} & \nearrow_{\varpi_j} & \\
Y_j & &
\end{array}$$

Definition 3.2. Let $\varpi := \varpi_1 \circ \cdots \circ \varpi_r : Y_\lambda^C \to SG^k(E)$. Define the class $\kappa_\lambda^C \in \Omega_*(SG^k(E))$ by

$$\kappa_\lambda^C = [Y_\lambda^C \to SG^k(E)] := \varpi_*(1_{Y_\lambda^C}).$$

Remark 3.3. It is also known that Y_λ^C is irreducible and has at worst rational singularities. Furthermore Y_λ^C is birational to X_λ^C through the projection π (see [Hudson et al. 2017], for example). Therefore in K-theory and Chow ring of $SG^k(E)$ the class κ_λ^C coincides with the fundamental class of the degeneracy loci X_λ^C. Note that in a general oriented cohomology theory, the fundamental class of X_λ^C is not defined since X_λ^C may not be an l.c.i. scheme.

Computing κ_λ^C. In this section, we establish an explicit formula of the class κ_λ^C in $\Omega_*(SG^k(E))$ in terms of a power series in relative Segre classes. The key ingredients for the computation are twofold: one is the formula that computes pushforwards along each ϖ_j and the other is so-called *umbral calculus* which is a computational technique to combine the pushforwards along all the ϖ_j.

We begin by the following lemma which was proved in [Hudson et al. 2017] for CK_*. One can easily check that the proof works for an arbitrary oriented BM homology and in particular for Ω_*.

Lemma 3.4. *For each $j = 1, \ldots, r$, the variety Y_j is regularly embedded in P_{j-1} and P_{j-1} is regularly embedded in $\mathbb{P}(U/D_{j-1})$. Furthermore, in $\Omega_*(P_{j-1})$, we have*

$$\iota_{j*}(1_{Y_j}) = \tilde{c}_{\lambda_j+n-k-j}\big((D_j/D_{j-1})^\vee \otimes (D_{\gamma_j}^\perp/F^{\chi_j})\big)(1_{P_{j-1}}).$$

Based on this lemma together with Theorem 2.10, we have the next pushforward formula for ϖ_j. For simplicity, let us introduce the following notation: for each $m \in \mathbb{Z}$ and $-n \le \ell \le n$, let

$$\widetilde{\mathscr{C}}_m^{(\ell)} := \widetilde{\mathscr{S}}_m\big(U^\vee - (E/F^\ell)^\vee\big).$$

In $\Omega^*(SG^k(E))$, we set $\mathscr{C}_m^{(\ell)} := \widetilde{\mathscr{C}}_m^{(\ell)}(1_{SG^k(E)})$.

Lemma 3.5. *In $\Omega_*(Y_{j-1})$, we have*

$$\varpi_{j*} \circ \tilde{\tau}_j^s(1_{Y_j}) = \sum_{p=0}^{j-1}\sum_{q=0}^{\infty}(-1)^p \tilde{c}_p(D_{j-1}^\vee - D_{\gamma_j}) \circ w_{-q}(D_{j-1}^\vee - D_{\gamma_j}) \circ \widetilde{\mathscr{C}}^{(\chi_j)}_{\lambda_j+s-p+q}(1_{Y_{j-1}}).$$

Proof. By Lemma 3.4, we have

$$\varpi_{j*} \circ \tilde{\tau}_j^s(1_{Y_j}) = \pi'_{j*} \circ \iota_{j*} \circ \tilde{\tau}_j^s(1_{Y_j}) = \pi'_{j*} \circ \tilde{\tau}_j^s \circ \iota_{j*}(1_{Y_j}) = \pi'_{j*} \circ \tilde{\tau}_j^s \circ \tilde{\alpha}_j(1_{P_{j-1}}),$$

where $\tilde{\alpha}_j := \tilde{c}_{\lambda_j+n-k-j}((D_j/D_{j-1})^\vee \otimes (D_{\gamma_j}^\perp/F^{\chi_j}))$. By Theorem 2.10, we have

$$\begin{aligned}\pi'_{j*} \circ \tilde{\tau}_j^s \circ \tilde{\alpha}_j(1_{P_{j-1}}) &= \widetilde{\mathscr{S}}_{s+\lambda_j}\big((U/D_{j-1})^\vee - (D_{\gamma_j}^\perp/F^{\chi_j})^\vee\big)(1_{Y_{j-1}}) \\ &= \widetilde{\mathscr{S}}_{s+\lambda_j}\big(U^\vee - (E/F^{\chi_j})^\vee - (D_{j-1} - D_{\gamma_j}^\vee)^\vee\big)(1_{Y_{j-1}}),\end{aligned}$$

where we have used $D_{\gamma_j}^\perp = E - D_{\gamma_j}^\vee$. Now the claim follows from Remark 2.9. □

For the umbral calculus mentioned above, we need to establish some notation. Let $R = \Omega^*(Gr_d(E))$, viewed as a graded algebra over $\mathbb{L}$, and let $t_1, \dots, t_r$ be indeterminates of degree 1. We use the multi-index notation $t^{\mathbf{s}} := t_1^{s_1} \cdots t_r^{s_r}$ for $\mathbf{s} = (s_1, \dots, s_r) \in \mathbb{Z}^r$. A formal Laurent series $f(t_1, \dots, t_r) = \sum_{\mathbf{s}\in\mathbb{Z}^r} a_{\mathbf{s}} t^{\mathbf{s}}$ is *homogeneous of degree $m \in \mathbb{Z}$* if $a_{\mathbf{s}}$ is zero unless $a_{\mathbf{s}} \in R_{m-|\mathbf{s}|}$ with $|\mathbf{s}| = \sum_{i=1}^r s_i$. Let $\operatorname{supp} f = \{\mathbf{s} \in \mathbb{Z}^r \mid a_{\mathbf{s}} \neq 0\}$. For each $m \in \mathbb{Z}$, define $\mathcal{L}_m^R$ to be the space of all formal Laurent series of homogeneous degree m such that there exists $\mathsf{n} \in \mathbb{Z}^r$ for which $\mathsf{n} + \operatorname{supp} f$ is contained in the cone in $\mathbb{Z}^r$ defined by $s_1 \ge 0,\ s_1 + s_2 \ge 0,\ \cdots,\ s_1 + \cdots + s_r \ge 0$. Then $\mathcal{L}^R := \bigoplus_{m\in\mathbb{Z}} \mathcal{L}_m^R$ is a graded ring over R with the obvious product. For each $i = 1, \dots, r$, let $\mathcal{L}^{R,i}$ be the R-subring of $\mathcal{L}^R$ consisting of series that do not contain any negative powers of $t_1, \dots, t_{i-1}$. In particular, $\mathcal{L}^{R,1} = \mathcal{L}^R$. A series $f(t_1, \dots, t_r)$ is a *power series* if it doesn't contain any negative powers of $t_1, \dots, t_r$. Let $R[\![t_1, \dots, t_r]\!]_m$ denote the set of all power series in $t_1, \dots, t_r$ of degree $m \in \mathbb{Z}$. We set $R[\![t_1, \dots, t_r]\!]_{\mathrm{gr}} := \bigoplus_{m\in\mathbb{Z}} R[\![t_1, \dots, t_r]\!]_m$.

Definition 3.6. Define a graded R-module homomorphism $\phi_1 : \mathcal{L}^R \to \Omega_*(SG^k(E))$ as

$$\phi_1^C(t_1^{s_1} \cdots t_r^{s_r}) = \widetilde{\mathscr{C}}_{s_1}^{(\chi_1)} \circ \cdots \circ \widetilde{\mathscr{C}}_{s_r}^{(\chi_r)}(1_{SG^k(E)}).$$

Similarly, for each $j = 2, \dots, d$, define a graded R-module homomorphism $\phi_j^C : \mathcal{L}^{R,j} \to \Omega_*(Y_{j-1})$ by setting

$$\phi_j^C(t_1^{s_1} \cdots t_r^{s_r}) = \tilde{\tau}_1^{s_1} \circ \cdots \circ \tilde{\tau}_{j-1}^{s_{j-1}} \circ \widetilde{\mathscr{C}}_{s_j}^{(\chi_j)} \circ \cdots \circ \widetilde{\mathscr{C}}_{s_r}^{(\chi_r)}(1_{Y_{j-1}}).$$

Remark 3.7. By regarding $\Omega^*(SG^k(E)) = \Omega_{\dim SG^k(E)-*}(SG^k(E))$, we have

$$\phi_1^C(t_1^{s_1} \cdots t_r^{s_r}) = \mathscr{C}_{s_1}^{(\chi_1)} \cdots \mathscr{C}_{s_r}^{(\chi_r)}.$$

Using ϕ_j^C, we can restate Lemma 3.5 as follows.

Lemma 3.8. *One has*

$$\varpi_{j*} \circ \tilde{\tau}_j^s(1_{Y_j}) = \phi_j^C \left(t_j^{\lambda_j+s} \frac{\prod_{i=1}^{j-1}(1 - t_i/t_j)P(t_j, t_i)}{\prod_{i=1}^{\gamma_j}(1 - \bar{t}_i/t_j)P(t_j, \bar{t}_i)} \right).$$

Proof. Consider the functions of $t_1, \dots, t_{j-1}$ defined by the following generating functions:

$$\sum_{p=0}^{\infty} H_p^\lambda(t_1, \dots, t_{j-1})u^p := \frac{e(t_1, \dots, t_{j-1}; u)}{e(\bar{t}_1, \dots, \bar{t}_{\gamma_j}; u)} = \frac{\prod_{i=1}^{j-1}(1 + t_i u)}{\prod_{i=1}^{\gamma_j}(1 + \bar{t}_i u)},$$

$$\sum_{q=0}^{\infty} W_{-q}^\lambda(t_1, \dots, t_{j-1})u^{-q} := \frac{w(t_1, \dots, t_{j-1}; u)}{w(\bar{t}_1, \dots, \bar{t}_{\gamma_j}; u)} = \frac{\prod_{i=1}^{j-1} P(u^{-1}, t_i)}{\prod_{i=1}^{\gamma_j} P(u^{-1}, \bar{t}_i)}.$$

Then we have

$$H_p^\lambda(\tilde{\tau}_1, \dots, \tilde{\tau}_{j-1}) = \tilde{c}_p(D_{j-1}^\vee - D_{\gamma_j}), \quad W_{-q}^\lambda(\tilde{\tau}_1, \dots, \tilde{\tau}_{j-1}) = \tilde{w}_{-q}(D_{j-1}^\vee - D_{\gamma_j}).$$

Thus, by Lemma 3.5 and the definition of ϕ_j^C, we have

$$\begin{aligned}
&\varpi_{j*} \circ \tilde{\tau}_j^s(1_{Y_j}) \\
&\quad = \phi_j^C \left(\sum_{p=0}^{j-1} \sum_{q=0}^{\infty} (-1)^p H_p^\lambda(t_1, \dots, t_{j-1}) W_{-q}^\lambda(t_1, \dots, t_{j-1}) t_j^{\lambda_j+s-p+q} \right) \\
&\quad = \phi_j^C \left(t_j^{\lambda_j+s} \left(\sum_{p=0}^{j-1} (-1)^p H_p^\lambda(t_1, \dots, t_{j-1}) t_j^{-p} \right) \left(\sum_{q=0}^{\infty} W_{-q}^\lambda(t_1, \dots, t_{j-1}) t_j^q \right) \right).
\end{aligned}$$

The claim follows from the definitions of H_p^λ and W_{-q}^λ in terms of the generating functions. □

Finally, we are able to prove the main theorem in the case of symplectic Grassmann bundles.

Theorem 3.9. *For a strict partition $\lambda \in \mathcal{SP}^k(n)$, the associated class κ_λ^C is given by*

$$\kappa_\lambda^C = \sum_{\mathbf{s}=(s_1,\ldots,s_r)\in\mathbb{Z}^r} f_{\mathbf{s}}^\lambda \mathscr{C}_{s_1+\lambda_1}^{(\chi_1)} \cdots \mathscr{C}_{s_r+\lambda_r}^{(\chi_r)},$$

where $f_{\mathbf{s}}^\lambda \in \mathbb{L}$ are the coefficients of the Laurent series

$$\frac{\prod_{1\le i<j\le r}(1-t_i/t_j)P(t_j,t_i)}{\prod_{(i,j)\in C(\lambda)}(1-\bar{t}_i/t_j)P(t_j,\bar{t}_i)} = \sum_{\mathbf{s}=(s_1,\ldots,s_r)\in\mathbb{Z}^r} f_{\mathbf{s}}^\lambda \cdot t_1^{s_1}\cdots t_r^{s_r} \tag{3-3}$$

as an element of $\mathcal{L}^{\mathbb{L}}$.

Proof. By Definition 3.2, it follows from successive applications of Lemma 3.8 (compare [Hudson et al. 2017]) that

$$\kappa_\lambda^C = \phi_1^C\left(t_1^{\lambda_1}\cdots t_r^{\lambda_r}\frac{\prod_{1\le i<j\le r}(1-t_i/t_j)P(t_j,t_i)}{\prod_{(i,j)\in C(\lambda)}(1-\bar{t}_i/t_j)P(t_j,\bar{t}_i)}\right).$$

Then, in view of the definition of the coefficients $f_{\mathbf{s}}$, it suffices to apply ϕ_1^C to obtain the claim. □

4. Odd orthogonal degeneracy loci

For this section we fix a nonnegative integer k.

Orthogonal degeneracy loci. Consider the vector bundle E of rank $2n+1$ over X with a symmetric non-degenerate bilinear form $\langle\ ,\ \rangle : E\otimes E \to O_X$ where O_X is the trivial line bundle over X. Let $\xi : OG^k(E) \to X$ be the Grassmann bundle consisting of pairs (x, U_x) where $x \in X$ and U_x is an $n-k$ dimensional isotropic subspace of E_x. Note that the bilinear form $\langle\ ,\ \rangle$ on E induces an isomorphism $F^\perp/F \otimes F^\perp/F \cong O_X$ for any maximal isotropic subbundle F of E where $F^\perp$ is the orthogonal complement of F with respect to $\langle\ ,\ \rangle$. This implies that $c_1(F^\perp/F)=0$ in $\Omega^*(X)\otimes_{\mathbb{Z}}\mathbb{Z}[1/2]$.

Fix a reference flag

$$0 = F^n \subset F^{n-1} \subset \cdots \subset F^1 \subset F^0 \subset (F^0)^\perp \subset F^{-1} \subset \cdots \subset F^{-n} = E,$$

such that $\operatorname{rk} F^i = n-i$ for $i \ge 0$ and $(F^i)^\perp = F^{-i}$ for all $i \ge 1$. For each $\lambda \in \mathcal{SP}^k(n)$ of length r, we define the associated degeneracy loci X_λ^B in $OG^k(E)$ is defined by

$$X_\lambda^B = \left\{(x, U_x) \in OG^k(E) \mid \dim(U_x \cap F^{\chi_i}) \ge i,\ i = 1,\ldots,r\right\},$$

where $\boldsymbol{\chi}$ is the characteristic index associated to λ.

Quadric bundle. The bundle $OG^{n-1}(E)$ is also known as the quadric bundle and we denote it by $Q(E)$. In this section, we do not assume that X is smooth as long as it is regularly embedded in a quasi-projective smooth variety. Let S be the tautological line bundle of $Q(E)$. In this particular case the Schubert varieties of $Q(E)$ are indexed by a single integer λ_1 and can be explicitly described as follows:

$$X^B_{\lambda_1} = \begin{cases} Q(E) \cap \mathbb{P}(F^{\lambda_1 - n}) & (0 \leq \lambda_1 < n), \\ \mathbb{P}(F^{\lambda_1 - n}) & (n \leq \lambda_1 < 2n). \end{cases} \tag{4-1}$$

It is worth noting that λ_1 represents the codimension of $X^B_{\lambda_1}$ in $Q(E)$.

Lemma 4.1. *The fundamental class of the subvariety $X^B_{\lambda_1}$ in $\Omega_*(Q(E))$ for $\lambda_1 < n$ is given by*

$$[X^B_{\lambda_1} \to Q(E)] = \tilde{c}_{\lambda_1}(S^\vee \otimes E/F^{\lambda_1 - n})(1_{Q(E)}). \tag{4-2}$$

Moreover the fundamental class of $X^B_{\lambda_1}$ in $\Omega_*(Q(E))$ for $\lambda_1 \geq n$ satisfies the identity

$$\begin{aligned} F^{(2)}\big(\tilde{c}_1(S^\vee \otimes (F^0)^\perp/F^0)\big)\big([X^B_{\lambda_1} \to Q(E)]\big) \\ = \tilde{c}_{\lambda_1}\big(S^\vee \otimes (E/(F^0)^\perp \oplus F^0/F^{\lambda_1 - n})\big)\big(1_{Q(E)}\big), \end{aligned} \tag{4-3}$$

where $F^{(2)}$ is a special case of the power series defined in (2-1).

Proof. The formula (4-2) follows from Lemma 2.2. For (4-3), first we show the case for $\lambda_1 = n$, by computing the class $[X^B_n \to Q(E)]$ in $\Omega^*(Q(E))$ in two different ways. The variety X^B_n is a divisor in $\mathbb{P}((F^0)^\perp)$, corresponding to the line bundle $S^\vee \otimes (F^0)^\perp/F^0$. Moreover, the scheme theoretic intersection $Q(E) \cap \mathbb{P}((F^0)^\perp)$ defines the Weil divisor $2X^B_n$ on $\mathbb{P}((F^0)^\perp)$ and in view of Lemma 2.3 we have

$$1_{Q(E)\cap\mathbb{P}((F^0)^\perp)} = \iota_*\big(F^{(2)}(\tilde{c}_1(S^\vee \otimes (F^0)^\perp/F^0))(1_{X^B_n})\big),$$

where $\iota : X^B_n \to Q(E) \cap \mathbb{P}((F^0)^\perp)$ is the obvious inclusion. Thus, by pushing forward this identity to $Q(E)$, we obtain the following identity in $\Omega_*(Q(E))$:

$$[Q(E) \cap \mathbb{P}((F^0)^\perp) \to Q(E)] = F^{(2)}\big(\tilde{c}_1(S^\vee \otimes (F^0)^\perp/F^0)\big)\big([X^B_n \to Q(E)]\big).$$

On the other hand, Lemma 2.2 implies that

$$[Q(E) \cap \mathbb{P}((F^0)^\perp) \to Q(E)] = \tilde{c}_n\big(S^\vee \otimes E/(F^0)^\perp\big)(1_{Q(E)}).$$

This proves (4-3) for $\lambda_1 = n$.

If $\lambda_1 > n$, again by Lemma 2.2 we have $[X^B_{\lambda_1} \to X^B_n] = \tilde{c}_i(S^\vee \otimes F^0/F^i)(1_{X^B_n})$ in $\Omega_*(X^B_n)$. Thus we have

$$\begin{aligned} F^{(2)}\big(\tilde{c}_1(S^\vee \otimes (F^0)^\perp/F^0)\big) &= \big([X^B_{\lambda_1} \to X^B_n]\big) \\ &= F^{(2)}\big(\tilde{c}_1(S^\vee \otimes (F^0)^\perp/F^0)\big) \circ \tilde{c}_i(S^\vee \otimes F^0/F^i)(1_{X^B_n}). \end{aligned}$$

By pushing it forward to $\Omega^*(Q(E))$ and applying (4-3) for $\lambda_1 = n$, we obtain

$$\begin{aligned}F^{(2)}\big(\tilde{c}_1(S^\vee \otimes (F^0)^\perp/F^0)\big)\big([X^B_{\lambda_1} \to Q(E)]\big) \\ &= \tilde{c}_n\big(S^\vee \otimes E/(F^0)^\perp\big) \circ \tilde{c}_{\lambda_1 - n}(S^\vee \otimes F^0/F^{\lambda_1 - n})(1_{Q(E)}) \\ &= \tilde{c}_{\lambda_1}\big(S^\vee \otimes (E/(F^0)^\perp \oplus F^0/F^i)\big)(1_{Q(E)}).\end{aligned}$$

This proves (4-3) for $\lambda_1 > n$. □

As mentioned above, we have $\tilde{c}_1((F^0)^\perp/F^0) = 0$ in $\Omega_*(Q(E)) \otimes_{\mathbb{Z}} \mathbb{Z}[1/2]$ so that $\tilde{c}_1(S^\vee \otimes (F^0)^\perp/F^0) = \tilde{c}_1(S^\vee)$. Therefore we have

$$F^{(2)}\big(\tilde{c}_1(S^\vee \otimes (F^0)^\perp/F^0)\big) = F^{(2)}\big(\tilde{c}_1(S^\vee)\big).$$

Notice that, since it is homogeneous of degree 0 with constant term 2, the series $F^{(2)}(u)$ is invertible in $\mathbb{L} \otimes_{\mathbb{Z}} \mathbb{Z}[1/2]$. Thus we have the following corollary.

Corollary 4.2. *In* $\Omega_*(Q(E)) \otimes_{\mathbb{Z}} \mathbb{Z}[1/2]$, *we have*

$$[X^B_{\lambda_1} \to Q(E)] = \begin{cases} \tilde{c}_{\lambda_1}(S^\vee \otimes E/F^{\lambda_1 - n})(1_{Q(E)}) & (0 \le \lambda_1 < n), \\ \dfrac{1}{F^{(2)}\big(\tilde{c}_1(S^\vee)\big)} \circ \tilde{c}_{\lambda_1}(S^\vee \otimes E/F^{\lambda_1 - n})(1_{Q(E)}) & (n \le \lambda_1 < 2n). \end{cases}$$

Remark 4.3. As mentioned in Remark 2.11, we have

$$[X^B_{\lambda_1} \to Q(E)] = \begin{cases} \widetilde{\mathscr{S}}_{\lambda_1}\big(S^\vee - (E/F^{\lambda_1 - n})^\vee\big)(1_{Q(E)}) & (0 \le \lambda_1 < n), \\ \dfrac{1}{F^{(2)}\big(c_1(S^\vee)\big)} \widetilde{\mathscr{S}}_{\lambda_1}\big(S^\vee - (E/F^{\lambda_1 - n})^\vee\big)(1_{Q(E)}) & (n \le \lambda_1 \le 2n). \end{cases}$$

Resolution of singularities and the class κ^B_λ. Consider the r-step flag bundle $\pi : Fl_r(U) \to OG^k(E)$ as before. We let $D_1 \subset \cdots \subset D_r$ be the tautological flag. Recall that $Fl_r(U)$ can be constructed as the tower of projective bundles

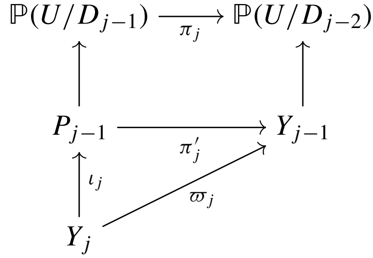
$$\pi : Fl_r(U) = \mathbb{P}(U/D_{r-1}) \xrightarrow{\pi_r} \cdots \xrightarrow{\pi_3} \mathbb{P}(U/D_1) \xrightarrow{\pi_2} \mathbb{P}(U) \xrightarrow{\pi_1} OG^k(E) \tag{4-4}$$

We regard D_j/D_{j-1} as the tautological line bundle of $\mathbb{P}(U/D_{j-1})$ where we let $D_0 = 0$. For each $j = 1, \ldots, r$, let $\tilde{\tau}_j := \tilde{c}_1((D_j/D_{j-1})^\vee)$ be the first Chern class operator of $(D_j/D_{j-1})^\vee$ on $\Omega_*(\mathbb{P}(U/D_{j-1}))$.

Definition 4.4. Let $\lambda \in \mathcal{SP}^k(n)$ be of length r. For each $j = 1, \ldots, r$, we define a subvariety Y_j of $\mathbb{P}(U/D_{j-1})$ by setting

$$Y_j := \big\{(x, U_x, (D_1)_x, \ldots, (D_j)_x) \in \mathbb{P}(U/D_{j-1}) \mid (D_i)_x \subset F^{\chi_i}_x,\ i = 1, \ldots, j\big\}.$$

We set $Y_0 := SG^k_r(U)$ and $Y^B_\lambda := Y_r$. Let $P_{j-1} := \pi_j^{-1}(Y_{j-1})$, $\pi'_j : P_{j-1} \to Y_{j-1}$ the projection and $\iota_j : Y_j \to P_{j-1}$ the obvious inclusion. Let $\varpi_j := \pi'_j \circ \iota_j$. We have the commutative diagram



As in the symplectic case we set $\varpi := \varpi_1 \circ \cdots \circ \varpi_r : Y_\lambda^B \to OG^k(E)$ and define

$$\kappa_\lambda^B := \varpi_*(1_{Y_\lambda^B}).$$

Computing κ_λ^B. The following lemma is known from [Hudson et al. 2017], where the computation of the fundamental class of Y_j in P_{j-1} is done in connective K-theory CK_*. However, the proof is valid in an arbitrary oriented BM homology and in particular in Ω_*.

Lemma 4.5. *For each $j = 1, \ldots, r$, the variety Y_j is regularly embedded in P_{j-1} and P_{j-1} is regularly embedded in $\mathbb{P}(U/D_{j-1})$, in particular they both belong to* $\mathbf{Lci_k}$. *Moreover we have*

$$\iota_{j*}(1_{Y_j}) = \tilde{\alpha}_j(1_{P_{j-1}})$$

in $\Omega_*(P_{j-1})$, where

$$\tilde{\alpha}_j = \begin{cases} \tilde{c}_{\lambda_j+n-k-j}\big((D_j/D_{j-1})^\vee \otimes (D_{\gamma_j}^\perp/F^{\chi_j})\big) & (-n \le \chi_j < 0), \\ \dfrac{1}{F^{(2)}(c_1((D_j/D_{j-1})^\vee))}\,\tilde{c}_{\lambda_j+n-k-j}\big((D_j/D_{j-1})^\vee \otimes (D_{\gamma_j}^\perp/F^{\chi_j})\big) & (0 \le \chi_j < n). \end{cases}$$

Definition 4.6. Let $-n \le \ell < n$. For each $m \in \mathbb{Z}$, we define the operators $\widetilde{\mathscr{B}}_m^{(\ell)}$ for $\Omega_*(OG^k(E)) \otimes_{\mathbb{Z}} \mathbb{Z}[1/2]$ by means of the following generating function

$$\sum_{m\in\mathbb{Z}} \widetilde{\mathscr{B}}_m^{(\ell)} u^m = \begin{cases} \widetilde{\mathscr{S}}\big(U^\vee - (E/F^\ell)^\vee; u\big) & (-n \le \ell < 0), \\ \dfrac{1}{F^{(2)}(u^{-1})}\widetilde{\mathscr{S}}\big(U^\vee - (E/F^\ell)^\vee; u\big) & (0 \le \ell < n). \end{cases}$$

If $\frac{1}{F^{(2)}(u^{-1})} = \sum_{s\ge 0} f_s u^{-s}$ with $f_s \in \mathbb{L} \otimes_{\mathbb{Z}} \mathbb{Z}[1/2]$, then we have

$$\widetilde{\mathscr{B}}_m^{(\ell)} = \sum_{s\ge 0} f_s \widetilde{\mathscr{S}}_{m+s}\big(U^\vee - (E/F^\ell)^\vee\big) \qquad (0 \le \ell < n).$$

Remark 4.7. If $\lambda = (\lambda_1) \in \mathcal{SP}^k(n)$, we have $\kappa_\lambda^B = \mathscr{B}_{\lambda_1}^{(\chi_1)}$.

Lemma 4.8. *For each $s \geq 0$, we have*

$$\varpi_{j*} \circ \tilde{\tau}_j^s(1_{Y_j}) = \sum_{p=0}^{\infty}\sum_{q=0}^{\infty}(-1)^p \tilde{c}_p(D_{j-1}^\vee - D_{\gamma_j}) \circ \widetilde{w}_{-q}(D_{j-1}^\vee - D_{\gamma_j}) \circ \widetilde{\mathscr{B}}^{(\chi_j)}_{\lambda_j+s-p+q}(1_{Y_{j-1}}).$$

Proof. By Lemma 4.5, we have

(4-5) $$\varpi_{j*}\circ\tilde{\tau}_j^s(1_{Y_j}) = \pi'_{j*}\circ\iota_{j*}\circ\tilde{\tau}_j^s(1_{Y_j}) = \pi'_{j*}\circ\tilde{\tau}_j^s\circ\iota_{j*}(1_{Y_j}) = \pi'_{j*}\circ\tilde{\tau}_j^s\circ\tilde{\alpha}_j(1_{P_{j-1}}).$$

Suppose that $\chi_j < 0$. By Theorem 2.10, the right-hand side of (4-5) equals

$$\widetilde{\mathscr{S}}_{\lambda_j+s}\big((U/D_{j-1}-D_{\gamma_j}^\perp/F^{\chi_j})^\vee\big)(1_{Y_{j-1}}) = \widetilde{\mathscr{S}}_{\lambda_j+s}\big((U-E/F^{\chi_j}-D_{j-1}+D_{\gamma_j}^\vee)^\vee\big)(1_{Y_{j-1}}),$$

where $D_{\gamma_j}^\perp = E - D_{\gamma_j}^\vee$. Then the claim follows from Remark 2.9. Similarly, if $0 \leq \chi_j$, Theorem 2.10 implies that the right-hand side of (4-5) equals

$$\sum_{s'=0}^{\infty} f_{s'}\widetilde{\mathscr{S}}_{\lambda_j+s+s'}\big((U/D_{j-1})^\vee - (D_{\gamma_j}^\perp/F^{\chi_j})^\vee\big)(1_{Y_{j-1}}),$$

where we set $F^{(2)}(u^{-1})^{-1} = \sum_{s'\geq 0} f_{s'}u^{-s'}$ with $f_{s'} \in \mathbb{L}\otimes_{\mathbb{Z}}[1/2]$ as above. Again, we use the identity $D_{\gamma_j}^\perp = E - D_{\gamma_j}^\vee$ and then the claim follows from Remark 2.9. □

Set $R := \Omega^*(OG^k(E))\otimes_{\mathbb{Z}}\mathbb{Z}[1/2]$ and let $\mathcal{L}^R$ be the ring of formal Laurent series with indeterminates $t_1, \ldots, t_r$ defined in the previous section.

Definition 4.9. Define a graded R-module homomorphism

$$\phi_1^B : \mathcal{L}^R \otimes_{\mathbb{Z}} \mathbb{Z}[1/2] \to \Omega_*(OG^k(E)) \otimes_{\mathbb{Z}} \mathbb{Z}[1/2]$$

by

$$\phi_1^B(t_1^{s_1}\cdots t_r^{s_r}) = \widetilde{\mathscr{S}}_{s_1}\big(U^\vee - (E/F^{\chi_1})^\vee\big)\circ\cdots\circ\widetilde{\mathscr{S}}_{s_r}\big(U^\vee - (E/F^{\chi_r})^\vee\big)(1_{OG^k(E)}).$$

Similarly, for each $j = 2, \ldots, r$, define a graded R-module homomorphism

$$\phi_j^B : \mathcal{L}^{R,j} \otimes_{\mathbb{Z}} \mathbb{Z}[1/2] \to \Omega_*(Y_{j-1}) \otimes_{\mathbb{Z}} \mathbb{Z}[1/2]$$

by

$$\phi_j^B(t_1^{s_1}\cdots t_r^{s_r}) = \tilde{\tau}_1^{s_1}\circ\cdots\circ\tilde{\tau}_{j-1}^{s_j}\circ\widetilde{\mathscr{S}}_{s_j}\big(U^\vee - (E/F^{\chi_j})^\vee\big)\circ\cdots\circ\widetilde{\mathscr{S}}_{s_r}\big(U^\vee - (E/F^{\chi_r})^\vee\big)(1_{Y_{j-1}}).$$

Remark 4.10. Note that ϕ_j^B replaces $\frac{t_i^m}{F^{(2)}(t_i)}$ by $\widetilde{\mathscr{B}}_m^{(\chi_i)}(1_{Y_{j-1}})$ for each i such that $j \leq i \leq r$ and $\chi_i \geq 0$, and $m \in \mathbb{Z}$.

As with Lemma 3.8, by making use of Lemma 4.8 we can prove the following lemma.

Lemma 4.11. *We have*

$$\varpi_{j*} \circ \tilde{\tau}_j^s(1_{Y_j}) = \begin{cases} \phi_j^B\left(t_j^{\lambda_j+s} \dfrac{\prod_{i=1}^{j-1}(1-t_i/t_j)P(t_j,t_i)}{\prod_{i=1}^{\gamma_j}(1-\bar{t}_i/t_j)P(t_j,\bar{t}_i)} \right) & (\chi_j < 0), \\ \phi_j^B\left(\dfrac{t_j^{\lambda_j+s}}{F^{(2)}(t_j)} \dfrac{\prod_{i=1}^{j-1}(1-t_i/t_j)P(t_j,t_i)}{\prod_{i=1}^{\gamma_j}(1-\bar{t}_i/t_j)P(t_j,\bar{t}_i)} \right) & (0 \le \chi_j), \end{cases}$$

for all $s \geq 0$.

A repeated application of Lemma 4.11 to the definition of κ_λ^B, together with Remark 4.10, allows us to obtain the main theorem for odd orthogonal Grassmannians.

Theorem 4.12. *We have*

$$\kappa_\lambda^B = \sum_{\mathbf{s}=(s_1,\dots,s_r)\in\mathbb{Z}^r} f_{\mathbf{s}}^\lambda \mathscr{B}_{\lambda_1+s_1}^{(\chi_1)} \cdots \mathscr{B}_{\lambda_r+s_r}^{(\chi_r)},$$

where the $f_{\mathbf{s}}^\lambda \in \mathbb{L}$ *are the coefficients of the Laurent series*

$$\frac{\prod_{1\le i<j\le r}(1-t_i/t_j)P(t_j,t_i)}{\prod_{(i,j)\in C(\lambda)}(1-\bar{t}_i/t_j)P(t_j,\bar{t}_i)} = \sum_{\mathbf{s}=(s_1,\dots,s_r)\in\mathbb{Z}^r} f_{\mathbf{s}}^\lambda \cdot t_1^{s_1}\cdots t_r^{s_r} \tag{4-6}$$

viewed as an element of $\mathcal{L}^{\mathbb{L}}$.

Acknowledgements

Both authors would like to thank the anonymous referee, whose helpful comments improved the overall presentation of this work. The early stages of this research were conducted while the first author was affiliated to KAIST, where he was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP)(ASARC, NRF-2007-0056093). This research was then completed within the framework of the research training group *GRK 2240: Algebro-Geometric Methods in Algebra, Arithmetic and Topology*, funded by the DFG. The second author is supported by Grant-in-Aid for Young Scientists (B) 16K17584.

References

[Anderson 2019] D. Anderson, "K-theoretic Chern class formulas for vexillary degeneracy loci", *Adv. Math.* **350** (2019), 440–485.

[Anderson and Fulton 2012] D. Anderson and W. Fulton, "Degeneracy Loci, Pfaffians, and vexillary signed permutations in types B, C, and D", preprint, 2012. arXiv

[Anderson and Fulton 2018] D. Anderson and W. Fulton, "Chern class formulas for classical-type degeneracy loci", *Compos. Math.* **154**:8 (2018), 1746–1774. MR Zbl

[Buch et al. 2017] A. S. Buch, A. Kresch, and H. Tamvakis, "A Giambelli formula for isotropic Grassmannians", *Selecta Math.* (*N.S.*) **23**:2 (2017), 869–914. MR Zbl

[Fulton 1998] W. Fulton, *Intersection theory*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 3 **2**, Springer, 1998. MR Zbl

[Hornbostel and Kiritchenko 2011] J. Hornbostel and V. Kiritchenko, "Schubert calculus for algebraic cobordism", *J. Reine Angew. Math.* **656** (2011), 59–85. MR Zbl

[Hornbostel and Perrin 2018] J. Hornbostel and N. Perrin, "Smooth Schubert varieties and generalized Schubert polynomials in algebraic cobordism of Grassmannians", *Pacific J. Math.* **294**:2 (2018), 401–422. MR Zbl

[Hudson and Matsumura 2018] T. Hudson and T. Matsumura, "Kempf–Laksov Schubert classes for even infinitesimal cohomology theories", pp. 127–151 in *Schubert varieties, equivariant cohomology and characteristic classes—IMPANGA 15*, edited by J. Buczyński et al., Eur. Math. Soc., Zürich, 2018. MR Zbl

[Hudson and Matsumura 2019] T. Hudson and T. Matsumura, "Segre classes and Damon–Kempf–Laksov formula in algebraic cobordism", *Math. Ann.* **374**:3 (2019), 1439–1457.

[Hudson et al. 2017] T. Hudson, T. Ikeda, T. Matsumura, and H. Naruse, "Degeneracy loci classes in K-theory—determinantal and Pfaffian formula", *Adv. Math.* **320** (2017), 115–156. MR Zbl

[Ikeda 2007] T. Ikeda, "Schubert classes in the equivariant cohomology of the Lagrangian Grassmannian", *Adv. Math.* **215**:1 (2007), 1–23. MR Zbl

[Ikeda and Matsumura 2015] T. Ikeda and T. Matsumura, "Pfaffian sum formula for the symplectic Grassmannians", *Math. Z.* **280**:1-2 (2015), 269–306. MR Zbl

[Kazarian 2000] M. Kazarian, "On Lagrange and symmetric degeneracy loci", preprint, Steklov Mathematical Institute, 2000, Available at http://www.pdmi.ras.ru/~arnsem/papers/shub.tex.gz.

[Kiritchenko and Krishna 2013] V. Kiritchenko and A. Krishna, "Equivariant cobordism of flag varieties and of symmetric varieties", *Transform. Groups* **18**:2 (2013), 391–413. MR Zbl

[Levine and Morel 2007] M. Levine and F. Morel, *Algebraic cobordism*, Springer, 2007. MR Zbl

[Mclaughlin 2010] E. V. Mclaughlin, *Equivariant Giambelli formulae for Grassmannians*, Ph.D. thesis, University of Maryland, 2010, Available at https://search.proquest.com/docview/855616668. MR

[Nakagawa and Naruse 2016] M. Nakagawa and H. Naruse, "Generalized (co)homology of the loop spaces of classical groups and the universal factorial Schur P- and Q-functions", pp. 337–417 in *Schubert calculus* (Osaka, 2012), edited by H. Naruse et al., Adv. Stud. Pure Math. **71**, Math. Soc. Japan, Tokyo, 2016. MR Zbl

[Nakagawa and Naruse 2018] M. Nakagawa and H. Naruse, "Universal Gysin formulas for the universal Hall–Littlewood functions", pp. 201–244 in *An alpine bouquet of algebraic topology*, edited by C. Ausoni et al., Contemp. Math. **708**, Amer. Math. Soc., Providence, RI, 2018. MR Zbl

[Pragacz 1991] P. Pragacz, "Algebro-geometric applications of Schur S- and Q-polynomials", pp. 130–191 in *Topics in invariant theory* (Paris, 1989/1990), edited by M.-P. Malliavin, Lecture Notes in Math. **1478**, Springer, 1991. MR Zbl

[Tamvakis and Wilson 2016] H. Tamvakis and E. Wilson, "Double theta polynomials and equivariant Giambelli formulas", *Math. Proc. Cambridge Philos. Soc.* **160**:2 (2016), 353–377. MR Zbl

[Vishik 2007] A. Vishik, "Symmetric operations in algebraic cobordism", *Adv. Math.* **213**:2 (2007), 489–552. MR Zbl

Received February 16, 2018. Revised January 25, 2019.

THOMAS HUDSON
FACHGRUPPE MATHEMATIK UND INFORMATIK
BERGISCHE UNIVERSITÄT WUPPERTAL
GAUSSSTRASSE 20
42119 WUPPERTAL
GERMANY

hudson@math.uni-wuppertal.de

TOMOO MATSUMURA
DEPARTMENT OF APPLIED MATHEMATICS
OKAYAMA UNIVERSITY OF SCIENCE
OKAYAMA 700-0005
JAPAN

matsumur@xmath.ous.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.119

A COMPACTNESS THEOREM ON BRANSON'S Q-CURVATURE EQUATION

GANG LI

Let (M, g) be a closed Riemannian manifold of dimension $5 \le n \le 7$. Assume that (M, g) is not conformally equivalent to the round sphere. If the scalar curvature R_g is greater than or equal to 0 and the Q-curvature Q_g is greater than or equal to 0 on M with $Q_g(p) > 0$ for some point $p \in M$, we prove that the set of metrics in the conformal class of g with prescribed constant positive Q-curvature is compact in $C^{4,\alpha}$ for any $0 < \alpha < 1$.

1. Introduction

On a manifold (M^n, g) of dimension $n \geq 5$, the Q-curvature of [1985] is defined by

$$Q_g = -\frac{2}{(n-2)^2}|\mathrm{Ric}_g|^2 + \frac{n^3-4n^2+16n-16}{8(n-1)^2(n-2)^2}R_g^2 - \frac{1}{2(n-1)}\Delta_g R_g,$$

where Ric_g is the Ricci curvature of g, R_g is the scalar curvature of g and Δ_g is the Laplacian operator with negative eigenvalues. The Paneitz operator [1983], which is the linear operator in the conformal transformation formula of the Q-curvature, is defined as

$$P_g = \Delta_g^2 - \mathrm{div}_g(a_n R_g g - b_n \mathrm{Ric}_g)\nabla_g + \frac{n-4}{2}Q_g, \tag{1-1}$$

with

$$a_n = \frac{(n-2)^2+4}{2(n-1)(n-2)} \quad \text{and} \quad b_n = \frac{4}{n-2}.$$

In fact, under the conformal change $\tilde{g} = u^{4/(n-4)}g$, the transformation formula of the Q-curvature is given by

$$P_g u = \frac{n-4}{2}Q_{\tilde{g}}u^{\frac{n+4}{n-4}}.$$

Research partially supported by China Postdoctoral Science Foundation Grant 2014M550540 and the National Natural Science Foundation of China No. 11701326.
MSC2010: primary 53C21; secondary 35B50, 35J61, 53A30.
Keywords: compactness, constant Q-curvature metrics, blowup argument, positive mass theorem, maximum principle, fourth order elliptic equations.

In comparison, for $n \geq 3$ the change of scalar curvature under the conformal change $\tilde{g} = u^{4/(n-2)} g$ satisfies

$$L_g u \equiv -\frac{4(n-1)}{(n-2)})\Delta_g u + R_g u = R_{\tilde{g}}\, u^{\frac{n+2}{n-2}}.$$

Let (M^n, g) be a closed Riemannian manifold of dimension $n \geq 5$. For existence of solutions u to the prescribed constant positive Q-curvature equation

$$P_g u = \frac{n-4}{2}\bar{Q} u^{\frac{n+4}{n-4}}, \tag{1-2}$$

with $\bar{Q} = \frac{1}{8} n(n^2-4)$, one may refer to [Esposito and Robert 2002; Qing and Raske 2006b; Hebey and Robert 2004; Gursky and Malchiodi 2015; Hang and Yang 2016a; 2016b; Gursky et al. 2016]. Recently, based on a version of maximum principle, Gursky and Malchiodi proved the following:

Theorem 1.1 [Gursky and Malchiodi 2015]. *For a closed Riemannian manifold (M^n, g) of dimension $n \geq 5$, if $R_g \geq 0$ and $Q_g \geq 0$ on M with Q_g not identically zero, then there is a conformal metric $h = u^{4/(n-4)} g$ with positive scalar curvature and constant Q-curvature $Q_h = \bar{Q}$.*

Moreover, they showed positivity of the Green's function of the Paneitz operator. Also, for $n = 5, 6, 7$, they proved a version of the positive mass theorem (see Theorem 2.1), which is important in proving compactness of the set of positive solutions to the prescribed constant Q-curvature problem in $C^{4,\alpha}(M)$ with $0 < \alpha < 1$. Note that when the pointwise condition in Theorem 1.1 is replaced by the requirement that the Yamabe constant $Y(M, [g])$ be greater than 0 and $Q_g \geq 0$, existence of solutions to (1-2) is proved in [Hang and Yang 2016b].

For compactness results of solutions to the prescribed constant Q-curvature equation under different conditions; see [Djadli et al. 2000; Hebey and Robert 2004; Humbert and Raulot 2009; Qing and Raske 2006a]. Djadli, Hebey and Ledoux [2000] studied the optimal Sobolev constant in the embedding $W^{2,2} \hookrightarrow L^{2n/(n-4)}$ when P_g has constant coefficients when g is an Einstein metric and also when P_g is replaced by a more general Paneitz-type operator. With some additional assumptions, they studied compactness of solutions to the related equations with $W^{2,2}$ bound and obtained existence of positive solutions for the corresponding equations. Under the assumption that the Paneitz operator is of strong positive type, Hebey and Robert [2004] considered compactness of positive solutions to (1-2) with $W^{2,2}$ bound in locally conformally flat manifolds with positive scalar curvature. They showed that under these conditions, when the Green's function of P_g satisfies a positive mass theorem, the compactness of solutions to (1-2) holds. Later, Humbert and Raulot [2009] showed that the positive mass theorem holds automatically under the assumption in [Hebey and Robert 2004]. Qing and Raske [2006a], with the use of the developing map and moving plane method, they showed an L^∞ bound of solutions to (1-2), for locally conformally flat manifolds with positive scalar curvature and an upper bound of the so-called Poincaré exponent (see [Chang et al. 2004]).

In this article we want to study compactness of solutions to (1-2) under the hypotheses in Theorem 1.1, following Schoen's outline of the proof of compactness of solutions to the prescribed scalar curvature problem. It is known that nonuniqueness of solutions to the prescribed scalar curvature problem (the Yamabe problem) could happen when the Yamabe constant of (M, g) is positive ([Schoen 1989; Pollack 1993]). In the conformal class of the round sphere metric, the solutions to the Yamabe problem are not uniformly bounded. Compactness of solutions to the Yamabe problem with positive Yamabe constant are well studied when g is not conformally equivalent to the round sphere metric. Following Schoen's original outline, one has the compactness of the solutions when (M^n, g) is locally conformally flat, or when $n \leq 24$ and the positive mass theorem holds on (M, g); see [Schoen 1991; Schoen and Zhang 1996; Li and Zhu 1999; Druet 2004; Chen and Lin 1998; Li and Zhang 2005; 2007; Marques 2005; Khuri et al. 2009]. It is interesting that when $n \geq 25$, there are conformal classes (which are not the round sphere metrics) with infinitely many solutions to the Yamabe problem which are not uniformly bounded; see [Brendle 2008; Brendle and Marques 2009]. In comparison, Wei and Zhao [2013] showed noncompactness of solutions to the positive constant Q-curvature equations for $n \geq 25$ in some conformal class different from that of the round sphere. For the compactness argument for the Nirenberg problem for a more general type conformal equation on the round sphere, see [Jin et al. 2017]. More precisely, we follow the approach in [Li and Zhu 1999] and [Marques 2005] for compactness of the set of solutions to the prescribed constant Q-curvature problem in dimension $5 \leq n \leq 7$ under the hypotheses of Theorem 1.1.

Our main theorem is the following:

Theorem 1.2. *Let (M^n, g) be a closed Riemannian manifold of dimension $5 \leq n \leq 7$ with $R_g \geq 0$, and also $Q_g \geq 0$ with $Q_g(p) > 0$ for some point $p \in M$. Assume that (M, g) is not conformally equivalent to the round sphere. Then there exists $C > 0$ depending on M and g such that for any positive solution u to (1-2), we have that*

$$C^{-1} \leq u \leq C,$$

and for any $0 < \alpha < 1$, there exists $C' > 0$ depending on M, g, and α such that

$$\|u\|_{C^{4,\alpha}} \leq C'.$$

We use a contradiction argument based on local information derived from a Pohozaev type identity for constant Q-curvature metrics and global information derived from the positive mass theorem of Gursky and Malchiodi [2015] (see Theorem 2.1). In comparison, for compactness of the Yamabe problem, the application of the positive mass theorem by Schoen and Yau [1979] (see also [Eichmair 2013; Eichmair et al. 2016; Witten 1981]) is crucial.

We extend the maximum principle in [Gursky and Malchiodi 2015] to manifolds with boundary under a Dirichlet-type condition and a scalar curvature condition restricted on the boundary; see Lemma 3.2. It turns out to be very useful and performs a role of a comparison theorem in the proof of the lower bound of the solutions away from the isolated blowup points (see Theorem 3.3) and in estimating upper bounds of solutions near blowup points (see Lemma 5.4). The Green's function is used as a comparison function in the uniform lower bound estimate Theorem 3.3. Note that Theorem 3.3 is important in the proof of the remark on page 138, Proposition 5.3 and Proposition 6.1. Since the main term of order $O(d_g^{-n})$ vanishes in $P_g d_g^{4-n}$, there is no comparison function to give the upper bound estimate in Proposition 5.3 directly. For that, the upper bound estimates of a sequence of blowup solutions near isolated simple blowup points are decomposed to a series of lemmas, following the approach in [Li and Zhu 1999] and in [Marques 2005]; see Section 5. We are able to prove a Harnack type inequality near the isolated blowup points for $5 \le n \le 9$; see Lemma 5.1. Besides the prescribed Q-curvature equation, nonnegativity of the scalar curvature is also important in the analysis of the blowing-up argument. With the aid of the Pohozaev type identity, we get a nice expansion of the limit of the blowing-up sequence near the blowup point, see Proposition 5.9, and using this we then show that in dimension $5 \le n \le 7$, each isolated blowup point is in fact an isolated simple blowup point. For the proof of Proposition 5.9, as in [Marques 2005], we need to estimate the speed of convergence of the rescaled functions to the limit, and for that, in Lemma 5.7 we need to classify bounded solutions to a linear fourth order elliptic equation on the Euclidean space which vanish uniformly at infinity, for $5 \le n \le 7$. The main difficulty for the classification problem in the Euclidean space is that the fourth order linear equation lacks the maximum principle, which is overcome by a combination of a comparison theorem for an initial value problem of ODEs, Kelvin transformation and an energy estimate; see Appendix B. After that, the proof of Theorem 1.2 is more or less standard, except that for the fourth order equation, more is involved for the blowing-up limit in ruling out the bubble accumulations; see Proposition 7.3. The Pohozaev type identity and the positive mass theorem in [Gursky and Malchiodi 2015] finally derive a contradiction on the sign of the constant term of the expansion of the singular limit function at the singular point in the proof of the main theorem. In Appendix A, we analyze the singular solutions to a linear fourth order elliptic equation near an isolated singular point, which is needed in Lemma 5.5 when finding the upper bound estimates of the solutions near the isolated simple blowup points. It is interesting to point out that in comparison with the proof of compactness of solutions to the Yamabe problem, here for compactness of positive constant Q-curvature metrics, no argument on vanishing of the Weyl tensor is needed for dimension $5 \le n \le 7$.

For $n \ge 8$, the Weyl tensor and its covariant derivatives are involved in the expansion of the Green's function and a vanishing argument of the Weyl tensor at the blowup points is needed (for instance, in Corollary 5.8 and Proposition 5.9), and yet a positive mass theorem for the Paneitz operator for cases which are not locally conformally flat in these dimensions is lacking. In this paper, for technical reasons, the Harnack inequality in Lemma 5.1 is only proved for $n \le 9$, the decay at infinity of the limit function $w(x)$ in Lemma 5.7 is only proved for $n \le 8$ due to the estimate (5-46), and the classification theorem (Corollary B.5) of solutions to the linear problem in Appendix B is given for $n \le 8$. But we believe that Lemma 5.1 and Corollary B.5 can be extended to high dimensions with some more discussion.

Remark. Let $Y(M,[g])$ be the Yamabe constant of (M,g) so that

$$Y(M,[g]) = \inf_{u\in C^\infty(M),\, u>0} \frac{\int_M \frac{4(n-1)}{n-2}|\nabla u|^2 + R_g u^2\, dV_g}{\left(\int_M u^{2n/(n-2)}\, dV_g\right)^{(n-2)/n}}.$$

Also, for $\alpha = \frac{4}{n-4}$ define

$$Y_4^*(M,[g]) = \inf_{u\in C^\infty(M),\, u>0,\, R_{u^\alpha g}>0} \frac{\int_M u\, P_g u\, dV_g}{\|u\|^2_{L^{2n/(n-4)}(M,g)}}.$$

From [Gursky et al. 2016], the following three statements are equivalent for dimension $n \ge 6$:

(1) $Y(M^n,[g]) > 0,\ \ P_g > 0$.

(2) $Y(M,[g]) > 0,\ \ Y_4^*(M,[g]) > 0$.

(3) There exists a metric $g_1 \in [g]$ such that $R_{g_1} > 0$ and $Q_{g_1} > 0$ on M.

As a corollary of Theorem 1.2, compactness of solutions to (1-2) holds for these conformal classes different from that of the round sphere for dimension $n = 6,\ 7$.

Remark. Recently, Li and Xiong [2019] proved compactness of prescribed constant Q metrics in a more general setting independently, by using the integration method developed from [Jin et al. 2017]. We follow the classical approach of [Li and Zhu 1999] and [Marques 2005].

To end the introduction, we introduce the definition of isolated blowup points and isolated simple blowup points.

Definition 1.3. Let g_j be a sequence of Riemannian metrics on a domain $\Omega \subseteq M$ with a uniform lower bound of injectivity radius $\bar{\delta} > 0$. Let $\{u_j\}_j$ be a sequence of positive solutions to (1-2) under the background metrics g_j in Ω. We call a point $\bar{x} \in \Omega$ an *isolated blowup point* of $\{u_j\}$ if there exist $\bar{C} > 0$, $0 < \delta < \min\{\frac{\bar{\delta}}{2}, \operatorname{dist}_{g_j}(\bar{x}, \partial\Omega)\}$ and $x_j \to \bar{x}$ as a local maximum of u_j with $u_j(x_j) \to \infty$ satisfying

(1-3) $B_\delta^{g_j}(\bar{x}),\ B_\delta^{g_j}(x_j) \subseteq \Omega$;

(1-4) $(B_\delta^{g_j}(x_j), x_j, g_j) \to (B_\delta^{g}(\bar{x}), \bar{x}, g)$ in $C^{k,\alpha}$ in the pointed Cheeger–Gromov sense, for $k > 0$ large and $0 < \alpha < 1$ and a smooth Riemannian metric g;

(1-5) $u_j(x) \le \bar{C} d_{g_j}(x, x_j)^{(4-n)/2}$ for $d_{g_j}(x, x_j) \le \delta$,

where $B_\delta^{g_j}$ is the δ-geodesic ball with respect to the metric g_j, and $d_{g_j}(x, x_j)$ is the geodesic distance between x and x_j with respect to the metric g_j.

In this paper, the sequence of metrics $\{g_j\}_j$ in the definition of the isolated blowup points are either a fixed metric on M, or the rescaled metrics $\{T_j g\}_j$ of g with a sequence of numbers $T_j \to \infty$, which converge to the flat metric as $j \to \infty$. Both these two cases satisfy the condition (1-4). For an isolated blowup point $x_j \to \bar{x}$ of u_j, we define

$$\bar{u}_j(r) = \frac{1}{|\partial B_r^{g_j}(x_j)|} \int_{\partial B_r^{g_j}(x_j)} u_j \, ds_{g_j}, \quad 0 < r < \delta,$$

and

$$\hat{u}_j(r) = r^{\frac{n-4}{2}} \bar{u}_j(r), \quad 0 < r < \delta, \tag{1-6}$$

with $B_r^{g_j}(x_j)$ that r-geodesic ball centered at x_j, ds_{g_j} the area element and $|\partial B_r^{g_j}(x_j)|$ the volume of $\partial B_r^{g_j}(x_j)$.

Definition 1.4. We call $\bar{x}$ an isolated simple blowup point if it is an isolated blowup point and there exists $0 < \delta_1 < \delta$ independent of j such that $\hat{u}_j$ has precisely one critical point in $(0, \delta_1)$, for j large.

2. The Green's representation

In this section, we assume that (M^n, g) is a closed Riemannian manifold of dimension $n \ge 5$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p) > 0$ for some point $p \in M$.

Theorem 2.1 [Gursky and Malchiodi 2015]. *For a closed Riemannian manifold (M^n, g) of dimension $n \ge 5$, if $R_g \ge 0$, $Q_g \ge 0$ on M and also $Q_g(p) > 0$ for some point $p \in M$, then*:

- *The scalar curvature R_g is greater than 0 in M.*
- *The Paneitz operator P_g is in fact positive and the Green's function G of P_g is positive where $G : M \times M - \{(q, q), q \in M\} \to \mathbb{R}$. Also, if $u \in C^4(M)$ and $P_g u \ge 0$ on M, then either $u \equiv 0$ or $u > 0$ on M.*

- *For any metric* g_1 *in the conformal class of* g, *if* $Q_{g_1} \geq 0$, *then* $R_{g_1} > 0$.
- *For any distinct points* $q_1, q_2 \in M$,

$$G(q_1, q_2) = G(q_2, q_1) = c_n d_g(q_1, q_2)^{4-n}(1 + f(q_1, q_2)), \tag{2-1}$$

with $c_n = \frac{1}{(n-2)(n-4)\omega_{n-1}}$, $\omega_{n-1} = |S^{n-1}|$, *and* $d_g(q_1, q_2)$ *the distance between* q_1 *and* q_2. *Here* f *is bounded and* $f \to 0$ *as* $d_g(q_1, q_2) \to 0$ *and*

$$|\nabla^j f| \leq C_j d_g(q_1, q_2)^{1-j} \tag{2-2}$$

for $1 \leq j \leq 4$.

- (*positive mass theorem*) *For* $5 \leq n \leq 7$, *or when* (M, g) *is locally conformally flat with dimension* $n \geq 5$, *for any point* $q_1 \in M$, *let* $x = (x^1, \ldots, x^n)$ *be the conformal normal coordinates constructed in* [Lee and Parker 1987] *centered at* q_1 *and* h *be the corresponding conformal metric. For* q_2 *close to* q_1, *the Green's function* $G_h(q_2, q_1)$ *of the Paneitz operator* P_h *has the expansion*

$$G_h(q_2, q_1) = c_n d_h(q_2, q_1)^{4-n} + \alpha + f(q_2)$$

with a constant $\alpha \geq 0$ *and* f *satisfying* (2-2) *and* $f(q_2) \to 0$ *as* $q_2 \to q_1$; *moreover,* $\alpha = 0$ *if and only if* (M^n, g) *is conformally equivalent to the round sphere.*

Let $u \in C^{4,\alpha}(M)$ be a solution to the equation

$$P_g u = f \geq 0.$$

Then we have the Green's representation

$$u(x) = \int_M G(x, y) f(y)\, dV_g(y)$$

for $x \in M$.

Now let $u > 0$ be a solution to the constant Q-curvature equation (1-2). Using the Green's representation

$$u(x) = \frac{n-4}{2} \bar{Q} \int_M G(x, y)\, u^{\frac{n+4}{n-4}}(y)\, dV_g(y),$$

we first show some basic estimates on the solution u.

Lemma 2.2. *Let* (M^n, g) *be a closed Riemannian manifold of dimension* $n \geq 5$ *with* $R_g > 0$, $Q_g \geq 0$ *on* M *and* $Q_g(p) > 0$ *for some point* $p \in M$. *Then there exist* C_1, $C_2 > 0$ *depending on* (M, g), *so that for any solution* u *to* (1-2), *we have*

$$\inf_M u \leq C_1, \quad \sup_M u \geq C_2.$$

Proof. Let $u(q) = \inf_M u$. Then by the Green's representation,

$$u(q) = \frac{n-4}{2}\bar{Q}\int_M G(q,y)\,u(y)^{\frac{n+4}{n-4}}\,dV_g(y)$$
$$\geq u(q)^{\frac{n+4}{n-4}} \times \frac{n-4}{2}\bar{Q}\int_M G(q,y)\,dV_g(y) \geq C_1^{-\frac{8}{n-4}} u(q)^{\frac{n+4}{n-4}}$$

with C_1 independent of the solution u and the point q, and the last inequality follows from (2-1). Therefore, the upper bound of $\inf_M u$ is established. A similar argument leads to the lower bound of $\sup_M u$. □

Next we give an integral type inequality, which shows that if u is bounded from above, then we get the lower bound of u.

Lemma 2.3. *Let (M^n, g) be a closed Riemannian manifold with dimension $n \geq 5$, $R_g > 0$, and also $Q_g \geq 0$ with $Q_g(p) > 0$ for some point $p \in M$. Then we have the inequality*

$$\inf_M u \geq C\left(\int_M G(z,y)^p\,u(y)^{\frac{8}{n-4}\alpha p}\,dV_g(y)\right)^{-\frac{q}{p}}$$

where $p = \frac{n+4}{n-4} - a$, $\frac{1}{p} + \frac{1}{q} = 1$, and $\alpha = \frac{(n-4)a}{8p}$, for any fixed number $\frac{4}{n-4} < a < \frac{8}{n-4}$, and z is the maximum point of u and $C = C(a, g) > 0$ is a constant. In particular, a uniform upper bound of u implies a uniform lower bound of u.

Proof. Let $u(x) = \inf_M u$ and $u(z) = \sup_M u$.

By the expansion formula (2-1), there exist two constants $C_3, C_4 > 0$ such that

$$0 < C_3 < \frac{1}{C_4} d_g(z_1, z_2)^{4-n} \leq G(z_1, z_2) \leq C_4 d_g(z_1, z_2)^{4-n} \tag{2-3}$$

for any two distinct points $z_1, z_2 \in M$.

By the Green's representation at the maximum point z we choose, we have

$$u(z) = \frac{n-4}{2}\bar{Q}\int_M G(z,y)\,u(y)^{\frac{n+4}{n-4}}\,dV_g(y)$$
$$\leq \frac{n-4}{2}\bar{Q}u(z)\int_M G(z,y)\,u(y)^{\frac{8}{n-4}}\,dV_g(y)$$

so that

$$1 \leq \frac{(n-4)}{2}\bar{Q}\int_M G(z,y)\,u(y)^{\frac{8}{n-4}(\alpha+(1-\alpha))}\,dV_g(y)$$
$$\leq \frac{(n-4)}{2}\bar{Q}\left(\int_M G(z,y)^p\,u(y)^{\frac{8}{n-4}\alpha p}\,dV_g(y)\right)^{\frac{1}{p}}\left(\int_M u(y)^{\frac{8}{(n-4)}(1-\alpha)q}\,dv_g(y)\right)^{\frac{1}{q}}$$
$$= \frac{(n-4)}{2}\bar{Q}\left(\int_M G(z,y)^p\,u(y)^{\frac{8}{n-4}\alpha p}\,dV_g(y)\right)^{\frac{1}{p}}\left(\int_M u(y)^{\frac{n+4}{n-4}}\,dv_g(y)\right)^{\frac{1}{q}},$$

with α, p, q chosen in the statement of the lemma. Here the second inequality is by Hölder's inequality. The range of a in the lemma keeps $0 < \alpha < 1$, $p > 1$ and $q > 1$, and also $p(4-n) > -n$ so that G^p is integrable.

Therefore, combining this with (2-3) we have

$$\begin{aligned}\inf_M u = u(x) &= \frac{n-4}{2}\bar{Q}\int_M G(x, y)u(y)^{\frac{n+4}{n-4}}\,dV_g(y)\\ &\geq C'\int_M u(y)^{\frac{n+4}{n-4}}\,dV_g(y) \geq C\left(\int_M G(z, y)^p\,u(y)^{\frac{8}{n-4}\alpha p}\,dV_g(y)\right)^{-\frac{q}{p}},\end{aligned}$$

where $C', C > 0$ are uniform constants independent of u, z and x. □

3. A maximum principle

In this section we prove a maximum principle for smooth domains with boundary in the manifold (M, g) defined in Lemma 2.2, which is a modification of the maximum principle given by Gursky and Malchiodi; see Lemma 3.2. As an application, we give a lower bound estimate of the blowing-up sequence.

Lemma 3.1. *Let $(\bar{\Omega}, g)$ be a compact Riemannian manifold of dimension $n \geq 5$ with boundary $\partial\Omega$. Let Ω be the interior of $\bar{\Omega}$. Assume the scalar curvature R_g is greater than or equal to 0 in $\bar{\Omega}$ and $R_g > 0$ at points on the boundary, and also $Q_g \geq 0$ in $\bar{\Omega}$. Then $R_g > 0$ in $\bar{\Omega}$.*

Proof. The proof is similar to that for closed manifolds. The Q-curvature is expressed as

$$Q_g = -\frac{1}{2(n-1)}\Delta_g R_g + c_1(n)R_g^2 - c_2(n)|\mathrm{Ric}|_g^2$$

with $c_1(n)$, $c_2(n)$ positive. By the nonnegativity of Q_g,

$$\frac{1}{2(n-1)}\Delta_g R_g \leq c_1(n)R_g^2.$$

By the strong maximum principle and the boundary condition, $R_g > 0$ in $\bar{\Omega}$. □

Lemma 3.2. *Let (M^n, g) be a closed Riemannian manifold of dimension $n \geq 5$ with $R_g \geq 0$, and $Q_g \geq 0$. Let $\Omega \subseteq M$ be an open domain with smooth boundary $\partial\Omega$ so that $\bar{\Omega} = \Omega \cup \partial\Omega$. Assume that $u \in C^4(\bar{\Omega})$ with $u > 0$ on $\partial\Omega$ satisfies*

$$P_g u \geq 0 \quad \text{in } \Omega. \tag{3-1}$$

Let $\tilde{g} = u^{4/(n-4)}g$ be the conformal metric in a neighborhood $\mathcal{U}$ of $\partial\Omega$ where $u > 0$. If the scalar curvature of $(\mathcal{U}, \tilde{g})$ satisfies $R_{\tilde{g}}(p) > 0$ for all points $p \in \partial\Omega$, then $u > 0$ in Ω.

Proof. Our conditions on the boundary guarantee that all the arguments are focused on the interior and then the argument is the same as in the proof of the maximum principle by Gursky and Malchiodi. For completeness, we present the proof.

We define the function

$$u_\lambda = (1-\lambda) + \lambda u$$

for $\lambda \in [0, 1]$, so that $u_0 = 1$ and $u_1 = u$. We assume

$$\min_{\bar{\Omega}} u \le 0.$$

Then there exists $\lambda_0 \in (0, 1]$ so that

$$\lambda_0 = \min\{\lambda \in (0, 1],\ \min_{\bar{\Omega}} u_\lambda = 0\}.$$

By definition, for $0 < \lambda < \lambda_0$, $u_\lambda > 0$. For the metric

$$g_\lambda = u_\lambda^{\frac{4}{n-4}} g,$$

the Q-curvature satisfies

$$Q_{g_\lambda} \ge 0 \quad \text{in}\, \Omega,$$

for $0 < \lambda < \lambda_0$. That follows from the conformal transformation formula

$$\begin{aligned} Q_{g_\lambda} &= \frac{2}{n-4} u_\lambda^{-\frac{n+4}{n-4}} P_g u_\lambda = \frac{2}{n-4} u_\lambda^{-\frac{n+4}{n-4}} ((1-\lambda) P_g(1) + \lambda P_g u) \\ &= \frac{2}{n-4} u_\lambda^{-\frac{n+4}{n-4}} ((1-\lambda) \frac{n-4}{2} Q_g + \lambda P_g u) \ge (1-\lambda) Q_g u_\lambda^{-\frac{n+4}{n-4}} \ge 0. \end{aligned}$$

Under the conformal transformation, the scalar curvature of g_λ satisfies

$$\begin{aligned} R_{g_\lambda} &= u_\lambda^{-\frac{n}{n-4}} \left(-\frac{4(n-1)}{n-4} \Delta_g u_\lambda - \frac{8(n-1)}{(n-4)^2} \frac{|\nabla_g u_\lambda|^2}{u_\lambda} + R_g u_\lambda \right) \\ &= u_\lambda^{-\frac{n}{n-4}} \left(-\frac{4(n-1)}{n-4} \lambda \Delta_g u - \frac{8(n-1)}{(n-4)^2} \frac{\lambda^2 |\nabla_g u|^2}{(1-\lambda) + \lambda u} + R_g u_\lambda \right) \\ &\ge u_\lambda^{-\frac{n}{n-4}} \left(-\frac{4(n-1)}{n-4} \lambda \Delta_g u - \frac{8(n-1)}{(n-4)^2} \frac{\lambda |\nabla_g u|^2}{u} + \lambda R_g u \right) \\ &= \lambda \left(\frac{u}{u_\lambda} \right)^{\frac{n}{n-4}} R_{\tilde{g}} > 0 \end{aligned}$$

on $\partial\Omega$ for $0 < \lambda < \lambda_0$. Then by Lemma 3.1,

$$R_{g_\lambda} > 0 \quad \text{in}\ \Omega,$$

for $0 < \lambda < \lambda_0$. Again by the conformal transformation formula of scalar curvature,

$$\Delta_g u_\lambda \le \frac{n-4}{4(n-1)} R_g u_\lambda \quad \text{in}\ \Omega.$$

By taking limit $\lambda \nearrow \lambda_0$, this also holds at $\lambda = \lambda_0$. But

$$u_\lambda = (1-\lambda) + \lambda u > 0$$

on $\partial\Omega$ for $0 \le \lambda \le 1$. By the strong maximum principle, $u_{\lambda_0} > 0$ in $\overline{\Omega}$, contradicting our choice of λ_0. Therefore, for all $0 \le \lambda \le 1$,

$$u_\lambda > 0 \quad \text{in } \Omega.$$

In particular, $u > 0$ in Ω. □

Theorem 3.3. *Let (M^n, g) be a closed Riemannian manifold of dimension $n \ge 5$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p_0) > 0$ for some point $p_0 \in M$. There exists $C > 0$ such that if there exists a sequence of positive solutions $\{u_j\}_{j=1}^\infty$ of (1-2) such that*

$$M_j = u_j(x_j) = \sup_M u_j \to \infty$$

as $j \to \infty$, then

$$u_j(p) \ge C M_j^{-1} d_g^{4-n}(p, x_j) \tag{3-2}$$

for any $p \in M$ such that $d_g(p, x_j) \ge M_j^{-2/(n-4)}$.

Proof. To prove the theorem, we only need to show that there exists $C > 0$ such that for any blowing-up sequence, there exists a subsequence such that (3-2) holds.

For each j, let $x = (x^1, \dots, x^n)$ be the corresponding normal coordinates in a small geodesic ball centered at x_j with radius $\delta > 0$ and x_j the origin. Let $y = M_j^{2/(n-4)} x$ and the metric h_j be given by $(h_j)_{pq}(y) = g_{pq}(M_j^{-2/(n-4)} y)$. Let

$$v_j(y) = M_j^{-1} u_j(\exp_{x_j}(M_j^{-\frac{2}{n-4}} y)) \quad \text{for } |y| \le \delta M_j^{\frac{2}{n-4}}.$$

Then,

$$0 < v_j(y) \le v_j(0) = 1,$$
$$P_{h_j} v_j(y) = \frac{n-4}{2} \bar{Q} v_j(y)^{\frac{n+4}{n-4}} \quad \text{for } |y| \le \delta M_j^{2/(n-4)}.$$

Here h_j converges to the Euclidean metric on $\mathbb{R}^n$ in C^k norm for any $k \ge 0$. By ellipticity, we have, after passing to a subsequence (still denoted as $\{v_j\}$), $v_j \to v$ in $C^4_{\text{loc}}(\mathbb{R}^n)$, and v satisfies

$$0 \le v(y) \le v(0) = 1 \quad \text{in } \mathbb{R}^n,$$
$$\Delta^2 v(y) = \frac{n-4}{2} \bar{Q} v(y)^{\frac{n+4}{n-4}} \quad \text{in } \mathbb{R}^n. \tag{3-3}$$

Also, since $R_{h_j} > 0$ and $R_{u_j^{4/(n-4)} g} > 0$ (by Theorem 2.1) on M, by the conformal transformation formula of scalar curvature,

$$\Delta_{h_j} v_j \le \frac{n-4}{4(n-1)} R_{h_j} v_j.$$

Passing to the limit we have

$$\Delta v(y) \le 0 \quad \text{in } \mathbb{R}^n.$$

By the strong maximum principle, since $v(0) = 1$, we have that $v(y) > 0$ in $\mathbb{R}^n$. Then by the classification theorem of C.S. Lin [1998], we have

$$v(y) = \left(\frac{1}{1+4^{-1}|y|^2}\right)^{\frac{n-4}{2}} \quad \text{in } \mathbb{R}^n.$$

We will abuse the notation with $v(|y|) = v(y)$. Thus, for fixed $R > 0$, for j large,

$$\tfrac{1}{2}\left(\frac{1}{1+4^{-1}R^2}\right)^{\frac{n-4}{2}} M_j \le u_j(\exp_{x_j}(x)) \le M_j \quad \text{for } |x| \le RM_j^{-\frac{2}{n-4}}.$$

For any $\epsilon > 0$, there exists $j_0 > 0$ such that, for $j > j_0$,

$$\|v_j - v\|_{C^4} \le \epsilon \quad \text{for } |y| \le 2.$$

We define $\phi_j : M - \{x_j\} \to \mathbb{R}$ as

$$\phi_j(p) = u_j(p) - \tau M_j^{-1} G_{x_j}(p),$$

with $G_{x_j}(p) = G(x_j, p)$ the Green's function of the Paneitz operator and $\tau > 0$ a small constant to be chosen. We will use the maximum principle to show that for $\epsilon, \tau > 0$ small,

$$\phi_j > 0 \ \text{ in } M - B_{M_j^{-2/(n-4)}}(x_j) \quad \text{for } j > j_0.$$

Here, we denote by $B_{M_j^{-2/(n-4)}}(x_j)$ the geodesic $M_j^{-2/(n-4)}$-ball centered at x_j in (M, g). If this holds, we will choose $\{u_j\}_{j>j_0}$ as the subsequence and the theorem is proved.

It is clear that

$$P_g\phi_j = P_g u_j = \frac{n-4}{2}\bar{Q} u_j^{\frac{n+4}{n-4}} > 0 \quad \text{in } M - B_{M_j^{-2/(n-4)}}(x_j).$$

To apply the maximum principle, we only need to verify the sign of ϕ_j and the related scalar curvature on $\partial B_{M_j^{-2/(n-4)}}(x_j)$.

First, for $|x| = M_j^{-\frac{2}{n-4}}$, we choose ϵ small so that for $j > j_0$,

$$u_j(\exp_{x_j}(x)) = M_j v_j(M_j^{\frac{2}{n-4}} x) \ge \tfrac{1}{2} v(1) M_j;$$

while by (2-3),

$$M_j^{-1} G_{x_j}(\exp_{x_j}(x)) \le C_4 M_j.$$

We take $\tau < v(1)/(4C_4)$. Then

$$\phi_j > 0 \quad \text{on } \partial B_{M_j^{-2/(n-4)}}(x_j) \quad \text{for } j > j_0.$$

Now let $\tilde{g}_j = \phi_j^{4/(n-4)} g_j$ in small neighborhood of $\partial B_{M_j^{-2/(n-4)}}(x_j)$ where $\phi_j > 0$. By conformal transformation,

$$R_{\tilde{g}_j} = \phi_j^{-\frac{n}{n-4}} \left(-\frac{4(n-1)}{n-4} \Delta_g \phi_j - \frac{8(n-1)}{(n-4)^2} \frac{|\nabla_g \phi_j|^2}{\phi_j} + R_g \phi_j \right).$$

Note that $R_g \phi_j > 0$ on $\partial B_{M_j^{-2/(n-4)}}(x_j)$. We only need to show that

$$-\frac{4(n-1)}{n-4} \left(\Delta_g \phi_j + \frac{2}{n-4} \frac{|\nabla_g \phi_j|^2}{\phi_j} \right) > 0 \quad \text{on } \partial B_{M_j^{-2/(n-4)}}(x_j) \quad \text{for } j > j_0. \tag{3-4}$$

Recall that

$$\left(\Delta_g u_j + \frac{2}{n-4} \frac{|\nabla_g u_j|^2}{u_j} \right) = M_j^{1+\frac{4}{n-4}} \left(\Delta_{h_j} v_j + \frac{2}{n-4} \frac{|\nabla_{h_j} v_j|^2}{v_j} \right).$$

Also,

$$\begin{aligned}
&\left(\Delta_{h_j} v_j + \frac{2}{n-4} \frac{|\nabla_{h_j} v_j|^2}{v_j} \right) \\
&\qquad \to \left(\Delta v + \frac{2}{n-4} \frac{|\nabla v|^2}{v} \right) \\
&\qquad = 2(4-n)(|y|^2+4)^{-\frac{n}{2}}(|y|^2+2n) \ + \frac{2}{n-4} \frac{(4-n)^2(|y|^2+4)^{2-n}|y|^2}{(|y|^2+4)^{(4-n)/2}} \\
&\qquad = 2(4-n)(|y|^2+4)^{-\frac{n}{2}}(|y|^2+2n) \ + 2(n-4)(|y|^2+4)^{-\frac{n}{2}}|y|^2 \\
&\qquad = 4n(4-n)(|y|^2+4)^{-\frac{n}{2}} < 0 \quad \text{at } |y| = 1.
\end{aligned}$$

Then we can choose $\epsilon < |v|_{C^4(B_1(0))}/100^n$. Combining this with the fact that

$$|D_g^k G_p(q)| \le C_k d_g^{4-n-k}(p,q) \quad \text{for } 0 \le k \le 4,$$

for any distinct points $p, q \in M$ with constants $C_k > 0$ independent of p and q, we have that there exists $\tau > 0$ only depending on C_k and ϵ so that

$$\begin{aligned}
\tau M_j^{-1} |\Delta_g G_{x_j}(\exp_{x_j}(M_j^{-\frac{2}{n-4}} y))| &< -M_j^{1+\frac{4}{n-4}} \frac{\Delta v}{4(2n+1)}, \quad \text{and} \\
\frac{|\nabla_g \phi_j|^2}{\phi_j} &\le \tfrac{5}{4} M_j^{1+\frac{4}{n-4}} \frac{|\nabla v|^2}{v} \quad \text{at } |y| = 1, \ \text{for } j > j_0.
\end{aligned}$$

Therefore, (3-4) holds for $j > j_0$, which implies

$$R_{\tilde{g}_j} > 0 \quad \text{on } \partial B_{M_j^{-2/(n-4)}}(x_j).$$

By Lemma 3.2, $\phi_j > 0$ in $M - B_{M_j^{-2/(n-4)}}(x_j)$. Recall that ϵ and τ are chosen independent of choice of the sequence. This completes the proof of the theorem. □

4. A Pohozaev type identity

In this section we introduce a Pohozaev type identity related to the constant Q-curvature equation. It will provide local information on the solutions in later use.

Let (M^n, g) be a closed Riemannian manifold of dimension $n \geq 5$ with $R_g \geq 0$, and $Q_g \geq 0$ with $Q_g(p_0) > 0$ for some point $p_0 \in M$. Let u be a positive solution to (1-2). For any geodesic ball $\Omega = B_\delta(q)$ in M with 2δ less than the injectivity radius of (M, g), we let

$$x = (x^1, \dots, x^n)$$

be the geodesic normal coordinates centered at q so that $g_{ij}(0) = \delta_{ij}$ and the Christoffel symbols $\Gamma^k_{ij}(0) = 0$. In this section, the gradient ∇, Laplacian Δ, divergence div, volume element dx, area element ds, σ-ball B_σ and

$$|x|^2 = (x^1)^2 + \cdots + (x^n)^2$$

are all with respect to the Euclidean metric. Define

$$\begin{aligned}
\mathcal{P}(u) &\equiv \int_\Omega \Big(x \cdot \nabla u + \frac{n-4}{2} u\Big) \Delta^2 u \, dx \\
&= \int_\Omega \Big[\frac{n-4}{2} \operatorname{div}(u \nabla(\Delta u) - \Delta u \nabla u) \\
&\qquad\qquad + \operatorname{div}((x \cdot \nabla u) \nabla(\Delta u) - \nabla(x \cdot \nabla u) \Delta u + \tfrac{1}{2} (\Delta u)^2 x)\Big] \, dx \\
&= \int_{\partial\Omega} \frac{n-4}{2} \Big(u \frac{\partial}{\partial \nu}(\Delta u) - \Delta u \frac{\partial}{\partial \nu} u\Big) \\
&\qquad\qquad + \Big((x \cdot \nabla u) \frac{\partial}{\partial \nu}(\Delta u) - \frac{\partial}{\partial \nu}(x \cdot \nabla u) \Delta u + \tfrac{1}{2} (\Delta u)^2 x \cdot \nu\Big) \, ds,
\end{aligned}$$

where ν is the outward-pointing normal vector of $\partial\Omega$ in the Euclidean metric. Then using (1-2), we have

$$\begin{aligned}
\mathcal{P}(u) &= \int_\Omega \Big(x \cdot \nabla u + \frac{n-4}{2} u\Big)(\Delta^2 - P_g) u + \Big(x \cdot \nabla u + \frac{n-4}{2} u\Big) P_g u \, dx \\
&= \int_\Omega \Big(x \cdot \nabla u + \frac{n-4}{2} u\Big)(\Delta^2 - P_g) u + \frac{n-4}{2} \bar{Q} \Big(x \cdot \nabla u + \frac{n-4}{2} u\Big) u^{\frac{n+4}{n-4}} \, dx \\
&= \int_\Omega \Big(x \cdot \nabla u + \frac{n-4}{2} u\Big)(\Delta^2 - P_g) u + \frac{(n-4)^2}{4n} \bar{Q} \operatorname{div}(u^{\frac{2n}{n-4}} x) \, dx \\
&= \int_\Omega \Big(x \cdot \nabla u + \frac{n-4}{2} u\Big)(\Delta^2 - P_g) u \, dx + \frac{(n-4)^2}{4n} \bar{Q} \int_{\partial\Omega} (x \cdot \nu) u^{\frac{2n}{n-4}} \, dx.
\end{aligned}$$

Using (1-1), we have

$$(\Delta^2 - P_g) u = (\Delta^2 - \Delta_g^2) u + \operatorname{div}_g(a_n R_g g - b_n \operatorname{Ric}_g) \nabla_g u - \frac{n-4}{2} Q_g u.$$

Since $\Gamma_{ij}^k(0)=0$ and $g_{ij}(0)=\delta_{ij}$,

$$\begin{aligned}(\Delta^2-\Delta_g^2)u\\ &=(\delta^{pq}\delta^{ij}\nabla_p\nabla_q\nabla_i\nabla_j-g^{pq}g^{ij}\nabla_p^g\nabla_q^g\nabla_i^g\nabla_j^g)u\\ &=(\delta^{pq}\delta^{ij}-g^{pq}g^{ij})\nabla_p\nabla_q\nabla_i\nabla_j u+O(|x|)|D^3u|+O(1)|D^2u|+O(1)|Du|\\ &=O(|x|^2)|D^4u|+O(|x|)|D^3u|+O(1)|D^2u|+O(1)|Du|.\end{aligned}$$

It follows that there exists $C>0$ which depends on $|Rm_g|_{L^\infty(\Omega)}$, $|Q_g|_{C(\Omega)}$ and $|\mathrm{Ric}_g|_{C^1(\Omega)}$ such that

$$|(\Delta^2-P_g)u|\le C(|x|^2|D^4u|+|x|\,|D^3u|+|D^2u|+|Du|+u). \tag{4-1}$$

5. Upper bound estimates near isolated simple blowup points

In this section we perform a parallel approach of [Li and Zhu 1999] to show the upper bound estimates of the solutions to (1-2) near an isolated simple blowup point; see Proposition 5.3. We start with a Harnack type inequality near an isolated blowup point.

Lemma 5.1. *Let (M^n,g) be a closed Riemannian manifold of dimension $5\le n\le 9$ with $R_g\ge 0$, and also $Q_g\ge 0$ with $Q_g(p_0)>0$ for some point $p_0\in M$. Let $\{u_j\}$ be a sequence of positive solutions to (1-2) and $x_j\to\bar{x}$ be an isolated blowup point. Then there exists a constant $C>0$ such that for any $0<r<\frac{\delta}{3}$ and $j>0$, we have*

$$\max_{q\in B_{2r}(x_j)-B_{r/2}(x_j)} u_j(q)\le C\min_{q\in B_{2r}(x_j)-B_{r/2}(x_j)} u_j(q). \tag{5-1}$$

Proof. Let $x=(x^1,\dots,x^n)$ be the geodesic normal coordinates centered at x_j. Here $\delta>0$ (see Definition 1.3) and 2δ is less than the injectivity radius. Let $y=r^{-1}x$. Define

$$v_j(y)=r^{\frac{n-4}{2}}u_j(\exp_{x_j}(ry))\quad\text{for } |y|<3.$$

Then by (1-5),

$$\begin{aligned}v_j(y)&\le\bar{C}|y|^{-\frac{n-4}{2}} &&\text{for } |y|<3,\\ v_j(y)&\le 3^{\frac{n-4}{2}}\bar{C} &&\text{for } \tfrac{1}{3}<|y|<3.\end{aligned}$$

We denote

$$\Omega_r=B_{3r}(x_j)-B_{\frac{r}{3}}(x_j).$$

By the Green's representation,

$$v_j(y)=r^{\frac{n-4}{2}}u_j(\exp_{x_j}(ry))=\frac{(n-4)\bar{Q}}{2}r^{\frac{n-4}{2}}\int_M G(\exp_{x_j}(ry),q)u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q)$$

$$= \frac{(n-4)\bar{Q}}{2} r^{\frac{n-4}{2}} \biggl(\int_{\Omega_r} G(\exp_{x_j}(ry), q) u_j(q)^{\frac{n+4}{n-4}} \, dV_g(q) + \int_{M-\Omega_r} G(\exp_{x_j}(ry), q) u_j(q)^{\frac{n+4}{n-4}} \, dV_g(q) \biggr).$$

We claim that for $\frac{5}{12} \le |y| \le \frac{12}{5}$, if

$$v_j(y) \ge 2 \times \frac{(n-4)\bar{Q}}{2} r^{\frac{n-4}{2}} \int_{\Omega_r} G(\exp_{x_j}(ry), q) u_j(q)^{\frac{n+4}{n-4}} \, dV_g(q), \tag{5-2}$$

then there exists $C > 0$ independent of j, x_j, r and y, such that for any $\frac{5}{12} \le |z| \le \frac{12}{5}$,

$$v_j(z) \ge C v_j(y). \tag{5-3}$$

In fact, by (2-3), there exists $C > 0$, such that

$$G(\exp_{x_j}(ry), q) \le C G(\exp_{x_j}(rz), q)$$

for $q \in M - \Omega_r$. Therefore,

$$\begin{aligned} \tfrac{1}{2} v_j(y) &\le \frac{(n-4)\bar{Q}}{2} r^{\frac{n-4}{2}} \int_{M-\Omega_r} G(\exp_{x_j}(ry), q) u_j(q)^{\frac{n+4}{n-4}} \, dV_g(q) \\ &\le C r^{\frac{n-4}{2}} \int_{M-\Omega_r} G(\exp_{x_j}(rz), q) u_j(q)^{\frac{n+4}{n-4}} \, dV_g(q) \\ &\le C v_j(z). \end{aligned}$$

This proves the claim.

We denote

$$\mathcal{C} = \Bigl\{ y \in \mathbb{R}^n, \ \tfrac{5}{12} \le |y| \le \tfrac{12}{5}, \ \text{so that (5-2) fails for } y \Bigr\}.$$

We choose $\frac{5}{12} \le |y| \le \frac{12}{5}$ with

$$v_j(y) \ge \tfrac{1}{2} \sup_{5/12 \le |z| \le 12/5} v_j(z).$$

If $y \notin \mathcal{C}$, then using the claim, we are done. If $y \in \mathcal{C}$, we will prove that the Harnack inequality (5-1) still holds.

By Hölder's inequality,

$$u_j(\exp_{x_j}(ry)) \le 2 \times \frac{(n-4)\bar{Q}}{2} \int_{\Omega_r} G(\exp_{x_j}(ry), q) u_j(q)^{\frac{n+4}{n-4}} \, dV_g(q)$$

$$
\begin{aligned}
&\le (n-4)\bar{Q}\biggl(\int_{\Omega_r} G(\exp_{x_j}(ry),q)^{\alpha}\,dV_g(q)\biggr)^{\frac{1}{\alpha}} \\
&\qquad\qquad \times\biggl(\int_{\Omega_r} u_j(q)^{\frac{n+4}{n-4}\beta}\,dV_g(q)\biggr)^{\frac{1}{\beta}} \\
&\le C(\alpha)r^{4-n+\frac{n}{\alpha}}\biggl(\int_{\Omega_r} u_j(q)^{\frac{n+4}{n-4}\beta}\,dV_g(q)\biggr)^{\frac{1}{\beta}} \\
&\le C(\alpha)r^{4-n+\frac{n}{\alpha}}\,(\bar{C}3^{\frac{n-4}{2}}r^{\frac{4-n}{2}})^{\frac{n+4}{n-4}\left(1-\frac{1}{\beta}\right)}\biggl(\int_{\Omega_r} u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q)\biggr)^{\frac{1}{\beta}} \\
&\le C(\alpha)r^{4-n+\frac{n}{\alpha}}\,(\bar{C}3^{\frac{n-4}{2}}r^{\frac{4-n}{2}})^{\frac{n+4}{n-4}\left(1-\frac{1}{\beta}\right)} \\
&\qquad\qquad \times\biggl(\int_{\Omega_r} C_4(4r)^{n-4}G(\exp_{x_j}(rz),q)u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q)\biggr)^{\frac{1}{\beta}} \\
&\le C(\alpha)r^{4-n+\frac{n}{\alpha}}(\bar{C}3^{\frac{n-4}{2}}r^{\frac{4-n}{2}})^{\frac{n+4}{n-4}\left(1-\frac{1}{\beta}\right)}r^{\frac{n-4}{\beta}}u_j(\exp_{x_j}(rz))^{\frac{1}{\beta}} \\
&= C(\alpha,\bar{C},n)r^{(2-\frac{n}{2})\left(1-\frac{1}{\beta}\right)}u_j(\exp_{x_j}(rz))^{\frac{1}{\beta}}
\end{aligned}
$$

for any $\frac{1}{3}\le|z|\le 3$, where $1<\alpha<\frac{n}{n-4}$, $\frac{1}{\alpha}+\frac{1}{\beta}=1$ such that $\beta>\frac{n}{4}$. Here we have used (1-5) and (2-3).

Since

$$\frac{n+4}{n-4}>\frac{n}{4}$$

for $5\le n\le 9$, we set $\beta=\frac{n+4}{n-4}$ and obtain

$$
\begin{aligned}
&u_j(\exp_{x_j}(rz))\ge C(\bar{C},n)\,r^4\,u_j(\exp_{x_j}(ry))^{\frac{n+4}{n-4}} && \text{(5-4)}\\
&\qquad\qquad\ge C(\bar{C},n)\,r^4\,(2^{-1}u_j(q))^{\frac{n+4}{n-4}}, && \text{(5-5)}
\end{aligned}
$$

for all $q\in B_{12r/5}(x_j)-B_{5r/12}(x_j)$ and $\frac{1}{2}\le|z|\le 2$, where $5\le n\le 9$.

For any $\frac{1}{2}\le|z|\le 2$,

$$
\begin{aligned}
\text{(5-6)}\quad &|\nabla_g u_j|(\exp_{x_j}(rz))\\
&\le\frac{n-4}{2}\bar{Q}\int_{B_{12r/5}(x_j)-B_{5r/12}(x_j)}|\nabla_g G(\exp_{x_j}(rz),q)|\,u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q)\\
&\quad+\frac{n-4}{2}\bar{Q}\int_{M-(B_{12r/5}(x_j)-B_{5r/12}(x_j))}|\nabla_g G(\exp_{x_j}(rz),q)|\,u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q).
\end{aligned}
$$

Note that for $\frac{1}{2} \le |z| \le 2$,

$$
\begin{aligned}
(5\text{-}7)\quad & u_j(\exp_{x_j}(rz)) \\
&\ge \frac{n-4}{2}\bar{Q}\int_{M-(B_{12r/5}(x_j)-B_{5r/12}(x_j))} G(\exp_{x_j}(rz),q)\,u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q) \\
&\ge Cr\int_{M-(B_{12r/5}(x_j)-B_{5r/12}(x_j))} |\nabla_g G(\exp_{x_j}(rz),q)|\,u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q),
\end{aligned}
$$

for a uniform constant C independent of j and the choice of points, where for the last inequality we have used (2-1).

Combining (5-4), (5-7) and (5-6), for $\frac{1}{2} \le |z| \le 2$ we have the gradient estimate

$$
\begin{aligned}
&|\nabla_g \log(u_j(\exp_{x_j}(rz)))| \\
&= \frac{|\nabla_g u_j(\exp_{x_j}(rz))|}{u_j(\exp_{x_j}(rz))} \\
&\le \frac{1}{u_j(\exp_{x_j}(rz))}\frac{n-4}{2}\bar{Q}\int_{B_{12r/5}(x_j)-B_{5r/12}(x_j)} |\nabla_g G(\exp_{x_j}(rz),q)|u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q) \\
&\quad + \frac{1}{u_j(\exp_{x_j}(rz))}\frac{n-4}{2}\bar{Q} \\
&\qquad \times \int_{M-(B_{12r/5}(x_j)-B_{5r/12}(x_j))} |\nabla_g G(\exp_{x_j}(rz),q)|u_j(q)^{\frac{n+4}{n-4}}\,dV_g(q) \\
&\le \frac{n-4}{2}\bar{Q}\int_{B_{12r/5}(x_j)-B_{5r/12}(x_j)} |\nabla_g G(\exp_{x_j}(rz),q)|C(\bar{C},n)^{-1}r^{-4}2^{-\frac{n+4}{n-4}}\,dV_g(q) \\
&\quad + C^{-1}r^{-1} \\
&\le C(\bar{C},n)(r^3r^{-4}+r^{-1}) \\
&= C(\bar{C},n)r^{-1},
\end{aligned}
$$

where $C(\bar{C},n)$ is some uniform constant depending on $\bar{C}$, the manifold and n. For any two points $p,\ q \in B_{2r}(x_j) - B_{r/2}(x_j)$, by the gradient estimate,

$$\frac{u_j(p)}{u_j(q)} \le e^{C(\bar{C},n)r^{-1}\,d_g(p,q)} \le e^{4nC(\bar{C},n)}.$$

This completes the proof of the Harnack inequality. □

Next we show that near an isolated blowup point, after rescaling the functions u_j converge to a standard solution to (3-3) in $\mathbb{R}^n$.

Lemma 5.2. *Let (M^n, g) be a closed Riemannian manifold of dimension $5 \le n \le 9$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p_0) > 0$ for some point $p_0 \in M$. Let $\{u_j\}$ be a sequence of positive solutions to (1-2) and $x_j \to \bar{x}$ be an isolated blowup point. Let $M_j = u_j(x_j)$. Assume $\{T_j\}_j$ and $\{\epsilon_j\}_j$ are any sequences of positive numbers* *such that $T_j \to +\infty$ and $\epsilon_j \to 0$ as $j \to \infty$. Then after possibly passing to a subsequence u_{k_j} and x_{k_j} (still denoted as u_j and x_j),*

$$\text{(5-8)} \quad \|M_j^{-1} u_j(\exp_{x_j}(M_j^{-\frac{2}{n-4}} y)) - (1+4^{-1}|y|^2)^{-\frac{n-4}{2}}\|_{C^4(B_{2T_j})} + \|M_j^{-1} u_j(\exp_{x_j}(M_j^{-\frac{2}{n-4}} y)) - (1+4^{-1}|y|^2)^{-\frac{n-4}{2}}\|_{H^4(B_{2T_j})} \le \epsilon_j,$$

and

$$\text{(5-9)} \qquad \frac{T_j}{\log(M_j)} \to 0 \quad \text{as } j \to \infty.$$

Proof. Let $x = (x^1, \dots, x^n)$ be geodesic normal coordinates centered at x_j, $y = r^{-1}x$ and the metric $h = r^{-2}g$ be the rescaled metric such that $(h_j)_{pq}(y) = (g_j)_{pq}(ry)$ in normal coordinates. Define

$$v_j(y) = M_j^{-1} u_j(\exp_{x_j}(M_j^{-\frac{2}{n-4}} y)) \quad \text{for } |y| < \delta\, M_j^{\frac{2}{n-4}}.$$

Then v_j satisfies

$$\text{(5-10)} \qquad P_{h_j} v_j(y) = \frac{n-4}{2} \bar{Q} v_j(y)^{\frac{n+4}{n-4}} \quad \text{for } |y| \le \delta M_j^{\frac{2}{n-4}},$$

$$\text{(5-11)} \qquad v_j(0) = 1, \quad \nabla_{h_j} v_j(0) = 0,$$

$$\text{(5-12)} \qquad 0 < v_j(y) \le \bar{C} |y|^{-\frac{n-4}{2}} \quad \text{for } |y| \le \delta M_j^{\frac{2}{n-4}}.$$

We next show that v_j is uniformly bounded. Since $R_{h_j} > 0$ and $R_{u_j^{4/(n-4)} g} > 0$ on M, by the conformal transformation formula of the scalar curvature,

$$\text{(5-13)} \qquad \Delta_{h_j} v_j \le \frac{n-4}{4(n-1)} R_{h_j} v_j,$$

where $R_{h_j} \to 0$ uniformly in $|y| \le 2$ as $j \to \infty$. Then the function $\eta_j(y) = (1+|y|^2)^{-1} v_j(y)$ satisfies

$$\Delta_{h_j} \eta_j + \sum_{k=1}^{n} b_k(y) \partial_k \eta_j(y) \le 0,$$

in $|y| \le 2$ with some function $b_k(y)$. By the maximum principle,

$$\text{(5-14)} \qquad \eta_j(0) \ge \inf_{|y|=r} \eta_j(y) \quad \text{for } 0 < r \le 1.$$

By the Harnack inequality (5-1) in Lemma 5.1,

$$\text{(5-15)} \qquad \max_{|y|=r} v_j(y) \le C \min_{|y|=r} v_j(y) \quad \text{for } 0 < r \le 1,$$

where C is independent of r and j. The inequalities (5-14) and (5-15) immediately lead to

$$\max_{|y|=r} v_j(y) \le C \min_{|y|=r} v_j(y) \le C v_j(0) = C \quad \text{for } 0 < r \le 1.$$

Combining this with (5-12), we have for $|y| \le \delta M_j^{2/(n-4)}$,

$$v_j(y) \le C,$$

with C independent of j, y and r.

Standard elliptic estimates of v_j imply that, after possibly passing to a subsequence, $v_j \to v$ in C^4_{loc} in $\mathbb{R}^n$ where, by (5-11) and (5-13), v satisfies

$$\Delta^2 v(y) = \frac{n-4}{2} \bar{Q} v^{\frac{n+4}{n-4}}, \quad \Delta v(y) \le 0, \quad v(y) \ge 0, \quad \text{for } y \in \mathbb{R}^n,$$
$$v(0) = 1, \ \nabla v(0) = 0.$$

By the strong maximum principle, $v(y) > 0$ in $\mathbb{R}^n$. Then the classification theorem in [Lin 1998] gives

$$v(y) = (1 + 4^{-1}|y|^2)^{-\frac{n-4}{2}}. \qquad \square$$

Remark. From Lemma 5.2, we can see that the proof of Theorem 3.3 still works at the isolated blowup point $x_j \to \bar{x}$. Therefore, there exists $C > 0$ independent of $j > 0$ such that for any isolated blowup point $x_j \to \bar{x}$,

$$u_j(q) \ge C u_j(x_j)^{-1} d_g^{4-n}(q, x_j)$$

for any $q \in M$ such that $d_g(q, x_j) \ge u_j(x_j)^{-2/(n-4)}$.

We now state the upper bound estimate of u_j near the isolated simple blowup points.

Proposition 5.3. *Let (M^n, g) be a closed Riemannian manifold of dimension $5 \le n \le 9$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p_0) > 0$ for some point $p_0 \in M$. Let $\{u_j\}$ be a sequence of positive solutions to (1-2) and $x_j \to \bar{x}$ be an isolated simple blowup point. Let δ_1 and $\bar{C}$ be the constants defined in Definition 1.4 and (1-5). Then there exists a constant C depending only on δ_1, $\bar{C}$, $\|R_g\|_{C^1(B_{\delta_1}(\bar{x}))}$ and $\|Q_g\|_{C^1(B_{\delta_1}(\bar{x}))}$, such that*

$$u_j(p) \le C u_j(x_j)^{-1} d_g(p, x_j)^{4-n} \quad \textit{for } d_g(p, x_j) \le \frac{\delta_1}{2}, \tag{5-16}$$

for $\delta_1 > 0$ small. Moreover, up to a subsequence,

$$u_j(x_j) u_j(p) \to a G(\bar{x}, p) + b(p) \quad \textit{in } C^4_{\text{loc}}(B_{\delta_1}(\bar{x}) - \{\bar{x}\}), \tag{5-17}$$

where G is the Green's function of the Paneitz operator P_g, $a > 0$ is a constant and $b(p) \in C^4(B_{\delta_1/2}(\bar{x}))$ satisfies $P_g b = 0$ in $B_{\delta_1/2}(\bar{x})$.

The proof of the proposition follows after a series of lemmas.

We first give a rough estimate on the upper bound of u_j near the isolated simple blowup points.

Lemma 5.4. *Under the condition in Proposition 5.3, assume $T_j \to \infty$ and $0 < \epsilon_j < e^{-T_j}$ satisfy (5-8) and (5-9). Denote $M_j = u_j(x_j)$. Then for any small number $0 < \sigma < \frac{1}{100}$, there exists $0 < \delta_2 < \delta_1$ and $C > 0$ independent of j such that*

$$M_j^{\lambda} u_j(p) \le C d_g(p, x_j)^{4-n+\sigma}, \tag{5-18}$$

$$M_j^{\lambda} |\nabla_g^k u_j(p)| \le C d_g(p, x_j)^{4-n-k+\sigma}, \tag{5-19}$$

for any p in $T_j M_j^{-2/(n-4)} \le d_g(p, x_j) \le \delta_2$ and $1 \le k \le 4$, where $\lambda = 1 - \frac{2}{n-4}\sigma$.

Proof. The outline of the proof is from [Li and Zhu 1999], while the use of our maximum principle here is more subtle. Let $x = (x^1, \ldots, x^n)$ be the geodesic normal coordinates centered at x_j for $d_g(p, x_j) \le \delta$. Let $r = |x|$. For any $\delta_2 \in (0, \delta_1)$ to be chosen, let

$$\Omega_j = \{p \in M,\ T_j M_j^{-\frac{2}{n-4}} \le d_g(p, x_j) \le \delta_2\}.$$

We want to use the maximum principle to get the upper bound of u_j. Before the construction of the barrier function on Ω_j, we first go through some properties of u_j.

From Lemma 5.2, we know that

$$u_j(p) \le C T_j^{4-n} M_j \quad \text{for } d_g(p, x_j) = T_j M_j^{-\frac{2}{n-4}}, \tag{5-20}$$

and there exists a critical point r_0 of $\hat{u}_j(r)$ defined in (1-6) in $0 < r < T_j M_j^{-2/(n-4)}$; moreover, for $r > r_0$, $\hat{u}_j(r)$ is decreasing. Using the assumption that $\bar{x}$ is an isolated simple blowup point, $\hat{u}_j$ is strictly decreasing for $T_j M_j^{-2/(n-4)} < r < \delta_1$. Therefore, combined with the Harnack inequality (5-1), for $p \in \Omega_j$ we have

$$\begin{aligned} d_g(p, x_j)^{\frac{n-4}{2}} u_j(p) &\le C \bar{u}_j(d_g(p, x_j)) \\ &\le C T_j^{\frac{n-4}{2}} M_j^{-1} \bar{u}_j(T_j M_j^{-\frac{2}{n-4}}) \\ &\le C T_j^{\frac{n-4}{2}} M_j^{-1} T_j^{4-n} M_j \\ &= C T_j^{-\frac{n-4}{2}}. \end{aligned}$$

This leads to

$$u_j(p)^{\frac{8}{n-4}} \le C T_j^{-4} d_g(p, x_j)^{-4} \quad \text{for } T_j M_j^{-\frac{2}{n-4}} < r < \delta_1. \tag{5-21}$$

We now define a linear elliptic operator on Ω_j,

$$L_j \phi = P_g \phi - \frac{n-4}{2} \bar{Q} u_j^{\frac{8}{n-4}} \phi \quad \text{for } \phi \in C^4(\Omega_j).$$

Therefore

$$L_j u_j = 0 \quad \text{in } \Omega_j.$$

Set

$$\varphi_j(p) = B\overline{M}_j \delta_2^\sigma d_g(p, x_j)^{-\sigma} + AM_j^{-1+\frac{2}{n-4}\sigma} d_g(p, x_j)^{-n+4+\sigma}, \ \ p \in \Omega_j,$$

where $A, B > 0$ are constants to be determined, $0 < \sigma < \frac{1}{100}$ and

$$\overline{M}_j = \sup_{d_g(p,x_j)=\delta_2} u_j \ \le \overline{C}\delta_2^{-\frac{n-4}{2}}.$$

There exists $C > 0$ such that for $m > 0,\ 1 \le k \le 4$, and any $p \in M$ fixed and $q \in M$ with $d_g(p, q) < \delta_2$ and δ_2 less than the injectivity radius, we have

$$|D_g^k d_g(p, q)^{-m}| \le Cm^k d_g(p, q)^{-m-k}. \tag{5-22}$$

It is easy to check that there exists $\delta_2 > 0$ independent of j so that in Ω_j,

$$|(P_g - \Delta_0^2)|x|^{-\sigma}| \le 100^{-1}|P_g(|x|^{-\sigma})|,$$
$$|(P_g - \Delta_0^2)|x|^{-n+4+\sigma}| \le 100^{-1}|P_g(|x|^{-n+4+\sigma})|,$$

where $|x| = d_g(p, x_j)$ and Δ_0 is the Euclidean Laplacian in the normal coordinates. It is easy to check that for $0 < m < n-4$ and $0 < r < \delta_2$,

$$-\Delta_0 r^{-m} = -m(m+2-n)r^{-m-2} > 0, \tag{5-23}$$

$$\Delta_0^2 r^{-m} = m(m+2-n)(m+2)(m+4-n)r^{-m-4} > 0. \tag{5-24}$$

But for $p \in \Omega_j$, by (5-21),

$$\frac{n-4}{2}\overline{Q}u_j(p)^{\frac{8}{n-4}}r^{-m} \le \frac{n-4}{2}\overline{Q}CT_j^{-4}r^{-m-4}.$$

Therefore,

$$L_j\varphi_j \ge 0 \quad \text{in } \Omega_j,$$

for j large. By (5-20), for $A > 1$,

$$u_j(p) < \varphi_j(p) \quad \text{for } d_g(p, x_j) = T_j M_j^{-\frac{2}{n-4}}. \tag{5-25}$$

Also, for $B > 1$,

$$u_j(p) < \varphi_j(p) \quad \text{for } d_g(p, x_j) = \delta_2. \tag{5-26}$$

We now want to check the sign of the scalar curvature $R_{(\varphi_j-u_j)^{4/(n-4)}g}$ near $\partial\Omega_j$. By the conformal transformation formula, it has the same sign as

$$-\frac{4(n-1)}{n-4}\Delta_g(\varphi_j - u_j) - \frac{8(n-1)}{(n-4)^2}\frac{|\nabla_g(\varphi_j - u_j)|^2}{(\varphi_j - u_j)} + R_g(\varphi_j - u_j).$$

Combining (1-5) and the standard interior estimate of (1-2), we have, for $k=1,\ 2$,

$$|D_g^k u_j(p)| \le C d_g(p, x_j)^{-\frac{n-4}{2}-k} \tag{5-27}$$

for some constant C independent of j and any $p \in \Omega_j$. It is easy to check that for $0 < m < n-4$,

$$\begin{aligned} \Delta_0 |x|^{-m} + \frac{2}{n-4} \frac{|\nabla_0 |x|^{-m}|^2}{|x|^{-m}} &= \left(m(m+2-n) + \frac{2m^2}{n-4} \right) |x|^{-m-2} \\ &= \frac{m(n-2)(m-(n-4))}{n-4} |x|^{-m-2} < 0. \end{aligned} \tag{5-28}$$

Also, note that for any positive functions $\phi_1,\ \phi_2 \in C^2$,

$$\begin{aligned} \Delta_0(\phi_1+\phi_2) &+ \frac{2}{n-4} \frac{|\nabla_0(\phi_1+\phi_2)|^2}{\phi_1+\phi_2} \\ &\le \left(\Delta_0 \phi_1 + \frac{2}{n-4} \frac{|\nabla_0(\phi_1)|^2}{\phi_1} \right) + \left(\Delta_0 \phi_2 + \frac{2}{n-4} \frac{|\nabla_0(\phi_2)|^2}{\phi_2} \right). \end{aligned} \tag{5-29}$$

Here we have used the fact that for any four positive numbers $a, b, c, d > 0$, we have

$$\frac{2cd}{a+b} \le \frac{bc^2}{a(a+b)} + \frac{ad^2}{b(a+b)}$$

so that

$$\frac{(c+d)^2}{a+b} = \frac{c^2+2c\,d+d^2}{a+b} \le \frac{c^2}{a} + \frac{d^2}{b}.$$

Using (5-25)–(5-29), we can choose $A, B > 100^n(1+C)$ independent of j and t with $C > 0$ in (5-27) so that

$$\begin{aligned} &-\frac{4(n-1)}{n-4} \Delta_g (t\varphi_j - u_j) \\ &\qquad - \frac{8(n-1)}{(n-4)^2} \frac{|\nabla_g(t\varphi_j - u_j)|^2}{(t\varphi_j - u_j)} + R_g(t\varphi_j - u_j) > 0 \quad \text{on } \partial\Omega_j, \end{aligned} \tag{5-30}$$

for all $t \ge 1$. Now for $t \ge 1$, we define

$$\phi_j^t(p) = t\varphi_j(p) - u_j(p), \quad p \in \Omega_j.$$

Then

$$0 \le L_j \phi_j^t = P_g \phi_j^t - \frac{n-4}{2} \bar{Q} \phi_j^t \quad \text{in } \Omega_j. \tag{5-31}$$

If

$$\phi_j^1 = \varphi_j - u_j \ge 0 \quad \text{in } \Omega_j, \tag{5-32}$$

then we are done. Otherwise, since Ω_j is compact, we pick the smallest number $t_j > 1$

so that $\phi_j^{t_j} \geq 0$. Therefore, by (5-31)

$$P_g \phi_j^{t_j} \geq \frac{n-4}{2} \bar{Q} \phi_j^{t_j} \geq 0. \tag{5-33}$$

Combining (5-25), (5-26), (5-30) and (5-33), the maximum principle in Lemma 3.2 implies

$$\phi_j^{t_j} > 0 \quad \text{in } \Omega_j,$$

contradicting the choice of t_j. Therefore, (5-32) holds. Now for $p \in \Omega_j$, we use Lemma 5.1, monotonicity of $\hat{u}_j$, and apply (5-32) at p to obtain

$$\begin{aligned} \delta_2^{\frac{n-4}{2}} \overline{M}_j &\leq C \hat{u}_j(\delta_2) \leq C \hat{u}_j(d_g(p, x_j)) \\ &\leq C d_g(p, x_j)^{\frac{n-4}{2}} (B \overline{M}_j \delta_2^\sigma d_g(p, x_j)^{-\sigma} + A M_j^{-\lambda} d_g(p, x_j)^{4-n+\delta}). \end{aligned}$$

Here $\frac{n-4}{2} > \sigma$. We choose p with $d_g(p, x_j)$ a small fixed number depending on n, σ, δ_2 to obtain

$$\overline{M}_j \leq C(n, \sigma, \delta_2) M_j^{-\lambda}.$$

The inequality (5-18) is then established from (5-32), and by the standard interior estimates for derivatives of u_j, the lemma is proved. □

Lemma 5.5. *Under the assumption in Proposition 5.3, for any* $0 < \rho \leq \delta_2/2$ *there exists a constant* $C(\rho) > 0$ *such that*

$$\limsup_{j \to \infty} \max_{p \in \partial B_\rho(x_j)} u_j(p) M_j \leq C(\rho),$$

where $M_j = u_j(x_j)$.

Proof. By Lemma 5.1, it suffices to show the inequality for some fixed small constant $\rho > 0$.

For any $p_\rho \in \partial B_\rho(x_j)$, we denote $\xi_j(p) = u_j(p_\rho)^{-1} u_j(p)$. Then ξ_j satisfies

$$P_g \xi_j(p) = \frac{n-4}{2} \bar{Q} u_j(p_\rho)^{\frac{8}{n-4}} \xi_j(p)^{\frac{n+4}{n-4}}.$$

For any compact subset $K \subseteq B_{\delta_2/2}(\bar{x}) - \{\bar{x}\}$, there exists $C(K) > 0$ such that for j large,

$$C(K)^{-1} \leq \xi_j \leq C(K) \quad \text{in } K.$$

Moreover, by Lemma 5.1, there exists $C > 0$ independent of $0 < r < \delta_2$ and j such that

$$\max_{B_r(x_j) - B_{r/2}(x_j)} u_j \leq C \inf_{B_r(x_j) - B_{r/2}(x_j)} u_j. \tag{5-34}$$

By the estimate (5-18), $u_j(p_\rho) \to 0$ as $j \to \infty$. Therefore, by the interior estimates of ξ_j, up to a subsequence,

$$\xi_j \to \xi \quad \text{in } C^4_{\mathrm{loc}}(B_{\delta_2/2}(\bar{x}) - \{\bar{x}\}),$$

with $\xi > 0$ such that

$$P_g \xi = 0 \quad \text{in } B_{\delta_2/2}(\bar{x}) - \{\bar{x}\},$$

and ξ satisfies (5-34) for $0 < r < \delta_2/2$. Moreover, for $0 < r < \rho$ and $\bar{\xi}(r) = |\partial B_r|^{-1} \int_{\partial B_r(\bar{x})} \xi \, ds_g$,

$$\lim_{j \to \infty} u_j(p_\rho)^{-1} r^{\frac{n-4}{2}} \bar{u}_j(r) = r^{\frac{n-4}{2}} \bar{\xi}(r).$$

Since $x_j \to \bar{x}$ is an isolated simple blowup point, $r^{(n-4)/2}\bar{\xi}(r)$ is nonincreasing in $0 < r < \rho$. Therefore, $\bar{x}$ is not a regular point of ξ.

Recall that

$$-\frac{4(n-1)}{n-2} \Delta_g u_j^{\frac{n-2}{n-4}} + R_g u_j^{\frac{n-2}{n-4}} = R_{u_j^{4/(n-4)} g} u_j^{\frac{n+2}{n-4}} \geq 0.$$

Passing to the limit, we have

$$-\frac{4(n-1)}{n-2} \Delta_g \xi^{\frac{n-2}{n-4}} + R_g \xi^{\frac{n-2}{n-4}} \geq 0, \tag{5-35}$$

in $B_{\delta_2/2}(\bar{x}) - \{\bar{x}\}$.

By Corollary A.5, for $\rho > 0$ small, there exists $m > 0$ independent of j such that for j large,

$$\begin{aligned} &\int_{B_\rho(x_j)} \Big(P_g \xi_j - \frac{n-4}{2} Q_g \xi_j \Big) \, dV_g \\ &\quad = \int_{\partial B_\rho(x_j)} \Big(\frac{\partial}{\partial \nu} \Delta_g \xi_j - \Big(a_n R_g \frac{\partial}{\partial \nu} \xi_j - b_n \mathrm{Ric}_g(\nabla_g \xi_j, \nu) \Big) \Big) \, ds_g \\ &\quad = \int_{\partial B_\rho(x_j)} \Big(\frac{\partial}{\partial \nu} \Delta_g \xi - \Big(a_n R_g \frac{\partial}{\partial \nu} \xi - b_n \mathrm{Ric}_g(\nabla_g \xi, \nu) \Big) \Big) \, ds_g + o(1) > m. \end{aligned} \tag{5-36}$$

On the other hand, nonnegativity of Q_g implies

$$\begin{aligned} &\int_{B_\rho(x_j)} \Big(P_g \xi_j - \frac{n-4}{2} Q_g \xi_j \Big) \, dV_g \\ &\quad = \int_{B_\rho(x_j)} \Big(\frac{n-4}{2} \bar{Q} u_j(p_\rho)^{-1} u_j(p)^{\frac{n+4}{n-4}} - \frac{n-4}{2} Q_g \xi_j \Big) \, dV_g \\ &\quad \leq \frac{n-4}{2} \bar{Q} \int_{B_\rho(x_j)} u_j(p_\rho)^{-1} u_j(p)^{\frac{n+4}{n-4}} \, dV_g. \end{aligned} \tag{5-37}$$

Using (5-8) and $\epsilon_j \le e^{-T_j}$, we have

$$\int_{B_{T_j M_j^{-2/(n-4)}}(x_j)} u_j^{\frac{n+4}{n-4}}\, dV_g \le C M_j^{-1},$$

while by (5-18) we have

$$\begin{aligned}\int_{B_\rho(x_j)-B_{T_j M_j^{-2/(n-4)}}(x_j)} u_j^{\frac{n+4}{n-4}}\, dV_g &\le C \int_{B_\rho(x_j)-B_{T_j M_j^{-2/(n-4)}}(x_j)} (M_j^{-\lambda} d_g(p,x_j)^{4-n+\sigma})^{\frac{n+4}{n-4}} \\ &\le C(T_j M_j^{-\frac{2}{n-4}})^{-4+\frac{n+4}{n-4}\sigma} M_j^{-\lambda\frac{n+4}{n-4}} \\ &= T_j^{-4+\frac{n+4}{n-4}\sigma} M_j^{-1} = o(1) M_j^{-1}.\end{aligned}$$

Therefore,

$$\int_{B_\rho(x_j)} u_j^{\frac{n+4}{n-4}}\, dV_g \le C M_j^{-1}. \tag{5-38}$$

Lemma 5.5 follows from (5-36)–(5-38). □

Proof of Proposition 5.3. Suppose (5-16) fails. Let $M_j = u_j(x_j)$. Then there exists a subsequence u_j and $\{p_j\}$ with $d_g(p_j, x_j) \le \delta_2/2$ with δ_2 in Lemma 5.4 such that

$$u_j(p_j) M_j d_g(p_j, x_j)^{n-4} \to \infty. \tag{5-39}$$

By Lemma 5.2 and $0 < \epsilon_j \le e^{-T_j}$,

$$T_j M_j^{-\frac{2}{n-4}} \le d_g(p_j, x_j) \le \frac{\delta_2}{2}.$$

For each j, let $x = (x^1, \ldots, x^n)$ be the geodesic normal coordinates centered at x_j. Denote $y = d_j^{-1} x$ where $d_j = d_g(p_j, x_j)$. We rescale:

$$v_j(y) = d_j^{\frac{n-4}{2}} u_j(\exp_{x_j}(d_j y)), \quad |y| \le 2.$$

Then v_j satisfies

$$P_{h_j} v_j(y) = \frac{n-4}{2} \bar{Q} v_j(y)^{\frac{n+4}{n-4}}, \quad |y| \le 2,$$

where $h_j = d_j^{-2} g$ so that $(h_j)_{pq}(y) = (g)_{pq}(d_j y)$. The metrics h_j depend on j. But since d_j has a uniform upper bound, the sequence of metrics stays in compact sets of $C^{k,\alpha}$ with $k > 4$ large and all the results in Lemma 5.5 hold uniformly for j. Also, the conclusion of Lemma 5.4 is scaling invariant. Note that the metrics h_j converge to a metric h in $C^{k,\alpha}$ with $k > 4$, and hence the Green's functions of Paneitz operators P_{h_j} converge to the Green's functions of Paneitz operators P_h uniformly away from the singularity. In particular, if $d_j \to 0$ then h_j converges to a flat metric on $B_2(0)$ so that in the proof of Proposition A.4, $G(p,\bar{x})$ will be replaced by $c_n|y|^{4-n}$ in Euclidean balls with c_n in (2-1). Therefore, Lemma 5.5 holds for v_j, and hence

$$\max_{|x|=1} v_j(0)v_j(x) \le C,$$

which shows that

$$M_j u_j(p_j) d_g(p_j, x_j)^{4-n} \le C,$$

contradicting (5-39). We have proved (5-16) in $B_{\delta_2/2}(\bar{x})$. By Lemma 5.1, the inequality (5-16) holds in $B_{\delta_1}(\bar{x})$.

The same properties for ξ_j in Lemma 5.5 now hold for $M_j u_j$ in $B_{\delta_2/2}(\bar{x})$. Up to a subsequence

$$M_j u_j \to v \quad \text{in } C^4_{\text{loc}}(B_{\delta_2/2}(\bar{x})),$$

and

$$P_g v = 0 \quad \text{in } B_{\delta_2/2}(\bar{x}).$$

By the remark on page 138, $v > 0$ in $B_{\delta_2/2}(\bar{x})$. Since $\bar{x}$ is an isolated simple blowup point, the same argument in Lemma 5.5 shows that $r^{(n-4)/2}\bar{v}(r)$ is nonincreasing for $0 < r < \delta_2/2$, where $\bar{v}(r) = |\partial B_r(\bar{x})|^{-1} \int_{\partial B_r(\bar{x})} v \, ds_g$. Combined with the Harnack inequality, it implies that v is not regular at $\bar{x}$. Also, v satisfies the condition in Proposition A.4. By Proposition A.4, we obtain (5-17). This completes the proof of Proposition 5.3. □

As an easy consequence of Proposition 5.3 and by the standard interior estimates of the elliptic equation (1-2), we have the following corollary:

Corollary 5.6. *Under the condition in Lemma 5.4, there exists $\delta_2 > 0$ independent of j such that for $T_j M_j^{-2/(n-4)} \le d_g(p, x_j) \le \delta_2$,*

$$|\nabla_g^k u_j(p)| \le C M_j^{-1} d_g(p, x_j)^{4-n-k} \quad \text{for } 0 \le k \le 4, \tag{5-40}$$

where $M_j = u_j(x_j)$, and C is a constant independent of j. For each j, let x be the geodesic normal coordinates of (Ω, g) centered at x_j. Then there exists $C > 0$ depending on $|g|_{C^3(\Omega)}$ such that for any fixed $r \le \delta_2$,

$$\left| \int_{d_g(p,x_j)\le r} \left(x \cdot \nabla u_j + \frac{n-4}{2} u_j \right) (\Delta^2 - P_g) u_j \, dx \right| \le C M_j^{-\frac{4}{n-4}+o(1)} \tag{5-41}$$

where $o(1) \to 0$ as $j \to \infty$.

Proof. Inequality (5-40) is a direct consequence of Proposition 5.3 and standard interior estimates of the elliptic equation (1-2). We will next establish (5-41). Note that $0 < \epsilon_j \le e^{-T_j}$. Using the estimates (5-40), (5-8) and (5-9), and recalling the error bound (4-1), we have

$$\begin{aligned}
&\int_{|x|\le T_j M_j^{-2/(n-4)}} \left| \left(x\cdot\nabla u_j + \frac{n-4}{2} u_j \right)(\Delta^2 - P_g)u_j \right| dx \\
&\quad \le \int_{|x|\le T_j M_j^{-2/(n-4)}} C(|x|\,|Du_j(x)| + u_j(x)) \\
&\qquad\quad \times \big(|x|^2 |D^4 u_j(x)| + |x|\,|D^3 u_j(x)| + |D^2 u_j(x)| + |Du_j(x)| + u_j(x) \big)\, dx \\
&\quad \le C \int_{|y|\le T_j} M_j (1+4^{-1}|y|^2)^{-\frac{n-4}{2}} M_j (1+4^{-1}|y|^2)^{-\frac{n-4}{2}-1} M_j^{\frac{4}{n-4}} M_j^{-\frac{2n}{n-4}}\, dy \\
&\quad = C M_j^{-\frac{4}{n-4}} \int_{|y|\le T_j} (1+4^{-1}|y|^2)^{3-n}\, dy = C M_j^{-\frac{4}{n-4}+o(1)}
\end{aligned}$$

and

$$\begin{aligned}
&\int_{T_j M_j^{-2/(n-4)} \le |x| \le r} \left| \left(x\cdot\nabla u_j + \frac{n-4}{2} u_j \right)(\Delta^2 - P_g)u_j \right| dx \\
&\quad \le \int_{T_j M_j^{-2/(n-4)} \le |x| \le r} C(|x|\,|Du_j(x)| + u_j(x)) \\
&\qquad\quad \times \big(|x|^2 |D^4 u_j(x)| + |x|\,|D^3 u_j(x)| + |D^2 u_j(x)| + |Du_j(x)| + u_j(x) \big)\, dx \\
&\quad \le C \int_{T_j M_j^{-2/(n-4)} \le |x| \le r} M_j^{-2} |x|^{6-2n}\, dx \\
&\quad \le C M_j^{-\frac{4}{n-4}+o(1)},
\end{aligned}$$

where $o(1) \to 0$ as $j \to \infty$ and $C > 0$ is a constant depending on $|g|_{C^3(\Omega)}$. Therefore,

$$\int_{|x|\le r} \left| \left(x\cdot\nabla u_j + \frac{n-4}{2} u_j \right)(\Delta^2 - P_g)u_j \right| dx \le C M_j^{-\frac{4}{n-4}+o(1)} \quad \text{for } T_j M_j^{-\frac{2}{n-4}} \le r,$$

where $C > 0$ is a constant independent of j and $o(1) \to 0$ as $j \to \infty$. □

For $n \ge 6$, a better estimate is needed in order to cancel the error terms in the Pohozaev identity. By (5-8),

$$u_j(\exp_{x_j}(x)) \le 2M_j (1 + 4^{-1} M_j^{\frac{4}{n-4}} |x|^2)^{-\frac{n-4}{2}} \quad \text{for } |x| \le T_j M_j^{-\frac{2}{n-4}}.$$

Combining this with Proposition 5.3, we have

$$\begin{aligned}
u_j(\exp_{x_j}(x)) &\le C \min\{ M_j (1 + 4^{-1} M_j^{\frac{4}{n-4}} |x|^2)^{-\frac{n-4}{2}},\ C M_j^{-1} |x|^{4-n} \} \\
&\le C\, M_j (1 + 4^{-1} M_j^{\frac{4}{n-4}} |x|^2)^{-\frac{n-4}{2}} \quad \text{for } |x| \le \delta_2.
\end{aligned}$$

For $n=6$ and $T_j M_j^{-2/(n-4)} \le r$,

$$\int_{|x|\le r}\left|\left(x\cdot\nabla u_j+\frac{n-4}{2}u_j\right)(\Delta^2-P_g)u_j\right|dx \le C\int_1^{M_j^{2/(n-4)}r} M_j^{-2}M_j^{\frac{2(n-6)}{n-4}}|y|^{5-n}\,d|y| \le CM_j^{-\frac{4}{n-4}}\ln(M_j^{\frac{2}{n-4}}r).$$

For $n\ge 7$ and $T_j M_j^{-2/(n-4)} \le r$,

$$\int_{|x|\le r}\left|\left(x\cdot\nabla u_j+\frac{n-4}{2}u_j\right)(\Delta^2-P_g)u_j\right|dx \le C\int_1^{M_j^{2/(n-4)}r} M_j^{-2}M_j^{\frac{2(n-6)}{n-4}}|y|^{5-n}\,d|y| \le CM_j^{-\frac{4}{n-4}}.$$

For the term $M_j^2\int_{|x|\le r}|Q_g|\,(u_j^2+|x|\,|Du_j|\,u_j)\,dx$ with $r>0$ fixed,

$$M_j^2\int_{|x|\le r}|Q_g|\,(u_j^2+|x|\,|Du_j|\,u_j)\,dx \le C\,M_j^2\int_0^{rM_j^{2/(n-4)}} M_j^2(1+|y|)^{8-2n}M_j^{-\frac{2n}{n-4}}|y|^{n-1}\,d|y| \le C\,M_j^{2-\frac{8}{n-4}}\int_0^{rM_j^{2/(n-4)}}(1+|y|)^{7-n}\,d|y|.$$

For $n=6$,

$$M_j^2\int_{|x|\le r}|Q_g|\,(u_j^2+|x|\,|Du_j|\,u_j)\,dx \le Cr^2.$$

For $n=7$,

$$M_j^2\int_{|x|\le r}|Q_g|\,(u_j^2+|x|\,|Du_j|\,u_j)\,dx \le Cr.$$

These are good terms. For later use, estimates on the error term

$$M_j^2\int_{|x|\le r}\left|\left(x\cdot\nabla u_j+\frac{n-4}{2}u_j\right)(\Delta^2-P_g)u_j\right|dx$$

are needed for $n\ge 6$.

For manifolds (M^n, g) of dimension $5\le n\le 7$, to estimate the error terms and to analyze the expansion of the limit function of $M_j u_j$ at the singular point, we have to work with the conformal normal coordinates. Let u_j be a sequence of positive solutions to (1-2) with isolated blowup points $x_j\to\bar{x}$. For each j, let $x=(x^1,\dots,x^n)$ be the conformal normal coordinates centered at x_j with the corresponding conformal metrics $g_j=\rho_j^{4/(n-4)}g$ constructed in [Lee and Parker 1987] such that

$$\det((g_j)_{pq}(x)) = 1+O(|x|^N),$$

with some large number N, say $N=10n$. We define $g_j=\rho_j^{4/(n-4)}g$ globally on M

by replacing the coefficient $\rho_j^{4/(n-4)}$ with $(\eta\rho_j+(1-\eta))^{4/(n-4)}$ which is still denoted as $\rho_j^{4/(n-4)}$ for simplicity, where η is a cut-off function supported in $B_{\delta_2}(x_j)$ under the metric g and $\eta = 1$ in $B_{\delta_2/(2)}(x_j)$. Recall that $\rho_j(x) = 1 + O(|x|^2)$ for $|x|$ small. Since $x_j \to \bar{x}$, by the construction of the conformal normal coordinates, $\rho_j(x) \to \rho(x)$ in $C^N(M)$ with $g_0 = \rho^{4/(n-4)}g$ the conformal metric corresponding to the conformal normal coordinates centered at $\bar{x}$. Let $\check{u}_j = \rho_j^{-1}u_j$. Then $\check{u}_j$ satisfies the equation

$$P_{g_j}\check{u}_j = \frac{n-4}{2}\bar{Q}\check{u}_j \quad \text{on } M.$$

Let

$$\hat{M}_j = \check{u}_j(x_j) = u_j(x_j)\rho_j(x_j)^{-1}.$$

We define the scaled coordinates $y = \hat{M}_j^{2/(n-4)}x$. Let $h_j = \hat{M}_j^{4/(n-4)}g_j$ and $v_j(y) = \hat{M}_j^{-1}\check{u}_j(\hat{M}_j^{-2/(n-4)}y)$. Denote

$$U_0(y) = (1+4^{-1}|y|^2)^{-\frac{n-4}{2}}, \quad y \in \mathbb{R}^n.$$

By the same argument as in Lemma 5.2, v_j converges to U_0 locally uniformly with the control as in (5-8) and (5-9). We will use this notation in Lemma 5.7.

Lemma 5.7. *Let (M^n, g) be a closed Riemannian manifold of dimension $5 \le n \le 7$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p) > 0$ for some point $p \in M$. Let $\{u_j\}$ be a sequence of positive solutions to (1-2) and $x_j \to \bar{x}$ be an isolated simple blowup point. For each j, let $x = (x^1, \dots, x^n)$ be the conformal normal coordinates at x_j with the corresponding conformal metric g_j. Denote $y = \hat{M}_j^{2/(n-4)}x$. Then there exist $\delta_2 > 0$ and $C > 0$ independent of j such that for $|y| \le \delta_2\hat{M}_j^{2/(n-4)}$,*

$$|v_j(y) - U_0(y)| \le C\hat{M}_j^{-2}, \tag{5-42}$$

where $\hat{M}_j = \check{u}_j(x_j)$.

Proof. The proof is a modification of Lemma 5.1 in [Marques 2005].

Let $s_j = \delta_2\hat{M}_j^{2/(n-4)}$ and

$$\Lambda_j = \max_{|y|\le s_j}|v_j - U_0| = |v_j(y_j) - U_0(y_j)|,$$

for some $|y_j| \le s_j$.

We claim that if there exists $c > 0$ such that $|y_j| \ge c\hat{M}_j^{2/(n-4)}$, there exists $C > 0$ such that (5-42) holds. To see this, observe that for $|y_j| \ge c\hat{M}_j^{2/(n-4)}$, by (5-16),

$$v_j(y_j) \le C|y_j|^{4-n} \le C\hat{M}_j^{-2},$$

and therefore

$$\Lambda_j \le C\hat{M}_j^{-2}.$$

This proves the claim.

Now assume $|y_j|\hat{M}_j^{-2/(n-4)} \to 0$ as $j \to \infty$. Then for $j > 0$ large, $|y_j| \le s_j/2$. Let

$$w_j(y) = \Lambda_j^{-1}(v_j(y) - U_0(y)).$$

Then $w_j(0) = 0$ and $Dw_j(0) = 0$.

We will argue by contradiction. If (5-42) fails, then, as $j \to \infty$,

$$\Lambda_j^{-1}\hat{M}_j^{-2} \to 0.$$

Let $h_j = \hat{M}_j^{4/(n-4)} g_j$. Then w_j satisfies the equation

$$P_{h_j} w_j - b_j w_j = H_j, \quad \text{for } |y| \le \delta_2 \hat{M}_j^{\frac{2}{n-4}},$$

where

$$b_j = \frac{(n-4)\overline{Q}(v_j^{(n+4)/(n-4)} - U_0^{(n+4)/(n-4)})}{2(v_j - U_0)} \ge 0,$$

and

$$\begin{aligned}
H_j(y) &= \Lambda_j^{-1}\Big(-P_{h_j}U_0 + \frac{n-4}{2}\overline{Q}U_0^{\frac{n+4}{n-4}}\Big) = \Lambda_j^{-1}(-P_{h_j} + \Delta_0^2)U_0(y) \\
&= \Lambda_j^{-1}\Big(\hat{M}_j^{-\frac{8}{n-4}} Q_{g_j}(\hat{M}_j^{-\frac{2}{n-4}} y)U_0(y) + \hat{M}_j^{-\frac{2}{n-4}N} O(|y|^N)(1+4^{-1}|y|^2)^{-\frac{n}{2}} \\
&\qquad + \hat{M}_j^{-\frac{2}{n-4}(1+N)} O(|y|^N)|y|(1+4^{-1}|y|^2)^{-\frac{n}{2}} \\
&\qquad + \hat{M}_j^{-\frac{2}{n-4}(2+N)} O(|y|^N)(1+4^{-1}|y|^2)^{1-\frac{n}{2}} \\
&\qquad + \hat{M}_j^{-\frac{2}{n-4}(3+N)} O(|y|^N)|y|(1+4^{-1}|y|^2)^{1-\frac{n}{2}}\Big) \\
&= \Lambda_j^{-1}(\hat{M}_j^{-\frac{8}{n-4}} Q_{g_j}(\hat{M}_j^{-\frac{2}{n-4}} y)U_0(y) + \hat{M}_j^{-\frac{2}{n-4}N} O(|y|^N)(1+4^{-1}|y|^2)^{-\frac{n}{2}}),
\end{aligned}$$

with $N = 10n$. By (5-16), for $|y| \le s_j$,

$$v_j(y) \le CU_0(y) \quad \text{and} \quad b_j(y) \le c\overline{Q}(1+4^{-1}|y|^2)^{-4} \quad \text{for some constant } c > 0.$$

By the interior estimates of the equation

$$P_{g_j} w_j = \hat{M}_j^{\frac{8}{n-4}} P_{h_j} w_j = \hat{M}_j^{\frac{8}{n-4}}(b_j w_j + H_j),$$

we have

$$\begin{aligned}
&|\nabla^k w_j(y)|_{h_j} \\
&\le C\hat{M}_j^{-\frac{2k}{n-4}}\big(\sup_{B_{\frac{1}{2}(\delta_2)2\hat{M}_j^{2/(n-4)}}(y)} |w_j| + \hat{M}_j^{\frac{8}{n-4}} \sup_{B_{\frac{1}{2}(\delta_2)2\hat{M}_j^{2/(n-4)}}(y)} |b_j w_j + H_j|\big) \\
&\le C(\hat{M}_j^{-\frac{2k}{n-4}} + \hat{M}_j^{\frac{8-2k}{n-4}}(1+|y|^2)^{-4}) \min\{1, \Lambda_j^{-1}(1+|y|^2)^{\frac{4-n}{2}}\} + C\hat{M}_j^{\frac{8-2k}{n-4}}\Lambda_j^{-1} \\
&\qquad \times \big(\hat{M}_j^{-\frac{8}{n-4}} Q_{g_j}(\hat{M}_j^{-\frac{2}{n-4}} y)U_0(y) + \hat{M}_j^{-\frac{2}{n-4}N} O(|y|^N)(1+4^{-1}|y|^2)^{-\frac{n}{2}}\big),
\end{aligned}$$

for $|\hat{M}_j^{-2/(n-4)} y| \le \delta_2$ and $1 \le k \le 3$.

For $\frac{1}{2}(\delta_2)\hat{M}_j^{2/(n-4)} \le |y| \le \delta_2 \hat{M}_j^{2/(n-4)}$, we have that $|w_j(y)| \le C\hat{M}_j^{-2}\Lambda_j^{-1}$, and then by a bootstrapping argument we get the estimate

$$|\nabla^k w_j(y)|_{h_j} \le C\hat{M}_j^{-\frac{2k}{n-4}}\hat{M}_j^{-2}\Lambda_j^{-1}, \tag{5-43}$$

for $1 \le k \le 5$.

Since $|w_j| \le 1$, by the interior estimates of the equation

$$P_{h_j} w_j = (b_j w_j + H_j),$$

we have that

$$|\nabla^k w_j(y)|_{h_j} \le C$$

where $|y| \le \delta_2 \hat{M}_j^{2/(n-4)}$ and $1 \le k \le 5$. Therefore, up to a subsequence, $w_j \to w$ in $C^4_{\mathrm{loc}}(\mathbb{R}^n)$. Moreover, $H_j(y) \to 0$ and w satisfies

$$\Delta^2 w(y) = \frac{n+4}{2}\bar{Q}U_0(y)^{\frac{8}{n-4}}w(y), \quad y \in \mathbb{R}^n. \tag{5-44}$$

For any fixed $y \in \mathbb{R}^n$, by the Green's representation, for j large,

$$\begin{aligned}
w_j(y) = &\int_\Omega G_{h_j}(y,z)P_{h_j}w_j(z)\,dV_{h_j}(z)\\
&-\int_{\partial\Omega} G_{h_j}(y,z)\Big[\frac{\partial}{\partial\nu}\Delta_{h_j}w_j - a_n\mathrm{Ric}_{h_j}(\nu,\nabla w_j) + b_n R_{h_j}\frac{\partial}{\partial\nu}w_j\Big]\,dS_{h_j}\\
&-\int_{\partial\Omega}\Big[-\frac{\partial}{\partial\nu}G_{h_j}(y,z)\Delta_{h_j}w_j\\
&\qquad + a_n\mathrm{Ric}_{h_j}(\nu,\nabla G_{h_j}(y,z))w_j - b_n R_{h_j}w_j\frac{\partial}{\partial\nu}G_{h_j}(y,z)\Big]\,dS_{h_j}\\
&-\int_{\partial\Omega}\Big[\Delta_{h_j}G_{h_j}(y,z)\frac{\partial}{\partial\nu}w_j - \frac{\partial}{\partial\nu}\Delta_{h_j}G_{h_j}(y,z)w_j\Big]\,dS_{h_j}\\
= &\int_\Omega G_{h_j}(y,z)P_{h_j}w_j(z)\,dV_{h_j}(z) + O(1)M_j^{-2}\Lambda_j^{-1},
\end{aligned}$$

as $j \to \infty$, where $\Omega = \{|z| \le \delta_2\hat{M}_j^{2/(n-4)}\}$ and the last equation is by (5-43). But for any $\delta > 0$, there exists $R(\delta) > |y| + 1 > 0$ independent of j such that

$$\begin{aligned}
&\int_{\Omega\cap\{|z|\ge R(\delta)\}} G_{h_j}(y,z)|P_{h_j}w_j(z)|\,dV_{h_j}(z)\\
&= \int_{\Omega\cap\{|z|\ge R(\delta)\}} G_{h_j}(y,z)|b_j w_j(z) + H_j(z)|\,dV_{h_j}(z)\\
&\le C(y)\int_R^{\delta_2\hat{M}_j^{2/(n-4)}} |z|^{4-n}\times\Big|\big(1+\tfrac{1}{4}|z|^2\big)^{-4}w_j + \Lambda_j^{-1}\hat{M}_j^{-\frac{8}{n-4}}|z|^{4-n}\\
&\qquad\qquad + \Lambda_j^{-1}\hat{M}_j^{-\frac{2N}{n-4}}|z|^N(1+|z|^2)^{-\frac{n}{2}}\Big|\times|z|^{n-1}\,d|z|
\end{aligned}$$

$$
\begin{aligned}
&\le C(y)\int_R^{\delta_2 \hat{M}_j^{2/(n-4)}} |z|^3(|z|^{-8}|w_j| + \Lambda_j^{-1}\hat{M}_j^{-2}\hat{M}_j^{\frac{-16+2n}{n-4}}|z|^{4-n} \\
&\qquad\qquad + \Lambda_j^{-1}\hat{M}_j^{-2}(\hat{M}_j^{-\frac{2}{n-4}}|z|)^{N-n+4}|z|^{-4})\,d|z| \\
&\le C(y)\int_R^{\delta_2 \hat{M}_j^{2/(n-4)}} (|z|^{-5} + \Lambda_j^{-1}\hat{M}_j^{-2}\hat{M}_j^{\frac{-16+2n}{n-4}}|z|^{7-n} \\
&\qquad\qquad + \Lambda_j^{-1}\hat{M}_j^{-2}(\hat{M}_j^{-\frac{2}{n-4}}|z|)^{N-n+4}|z|^{-1})\,d|z| \\
&\le C(y)(R^{-4} + \Lambda_j^{-1}\hat{M}_j^{-2}) \le \delta
\end{aligned}
$$

for j large and $5 \le n \le 7$.

Therefore,

$$
w(y) = c_n \int_{\mathbb{R}^n} |y-z|^{4-n}\Delta_0^2 w(z)\,dz = \frac{n+4}{2} c_n \int_{\mathbb{R}^n} |y-z|^{4-n} U_0(z)^{\frac{8}{n-4}} w(z)\,dz. \tag{5-45}
$$

Also, for $|y| \le \frac{1}{2}\delta_2 \hat{M}_j^{2/(n-4)}$, since $|w_j| \le 1$, we have

$$
\begin{aligned}
|w_j(y)| &= \left| \int_\Omega G_{h_j}(y,z) P_{h_j} w_j(z)\,dV_{h_j}(z) + O(1)\hat{M}_j^{-2}\Lambda_j^{-1} \right| \\
&= \left| \int_\Omega G_{h_j}(y,z)(b_j w_j + H_j)\,dV_{h_j}(z) + O(1)\hat{M}_j^{-2}\Lambda_j^{-1} \right| \\
&\le C\Big[(1+|y|)^{-4} + (1+|y|)^{4-n} + \Lambda_j^{-1}\hat{M}_j^{-2}(\hat{M}_j^{\frac{2n-16}{n-4}}(1+|y|)^{8-n} \\
&\qquad + (\hat{M}_j^{-\frac{2}{n-4}}|y|)^{N-n+4} + 1) + \Lambda_j^{-1}\hat{M}_j^{-2} \Big],
\end{aligned} \tag{5-46}
$$

with $N = 10n$. Therefore, for $5 \le n \le 7$, there exists $C > 0$ such that for $y \in \mathbb{R}^n$,

$$
|w(y)| \le C\,[\,(1+|y|)^{-4} + (1+|y|)^{4-n}].
$$

Since $v_j(0) = 1$ and $Dv_j(0) = 0$, we also have that $w(0) = 0$ and $Dw(0) = 0$.

Now by Corollary B.5, $w(y) = 0$ for $y \in \mathbb{R}^n$. Therefore, $y_j \to \infty$ as $j \to \infty$. But then by (5-46), $w_j(y_j) \to 0$ as $j \to \infty$, which is a contradiction with $w_j(y_j) = 1$ for $j \ge 1$. This completes the proof of Lemma 5.7. □

Remark. Using (5-42) and the equation satisfied by $(v_j - U_j)$ instead of that of w_j in the proof of Lemma 5.7, there exists a constant $C > 0$ independent of j such that

$$
|\nabla^k(v_j - U_j)| \le C\hat{M}_j^{-2}(1+|y|)^{-k},
$$

for $|y| \le \delta_2 \hat{M}_j^{2/(n-4)}$ and $1 \le k \le 4$.

Corollary 5.8. *Under the condition in Lemma 5.4, for each j let $x = (x^1, \dots, x^n)$ be the conformal normal coordinates of (Ω, g) centered at x_j constructed in [Lee* *and Parker 1987*], *and we denote* g_j *as the corresponding conformal metrics so that*

$$\det(g_j) = 1 + O(r^N),$$

where $N = 10n$. *Then there exists* $C > 0$ *such that for any fixed* $r \le \delta_2$,

$$\lim_{j\to\infty} \hat{M}_j^2 \left| \int_{d_{g_j}(p,x_j)\le r} \left(x \cdot \nabla \check{u}_j + \frac{n-4}{2} \check{u}_j \right) (\Delta^2 - P_{g_j}) \check{u}_j \, dx \right| \le Cr \tag{5-47}$$

for $5 \le n \le 7$, *where* $\check{u}_j = u_j \rho_j^{-1}$ *and* $\hat{M}_j = \check{u}_j(x_j)$ *are defined as in the paragraph preceding Lemma 5.7*, $N = 10n$ *and* $g_j = \rho_j^{4/(n-4)} g$.

Proof. Let

$$\tilde{u}_j(x) = \hat{M}_j^{-1} (|x|^2 + \hat{M}_j^{-\frac{4}{n-4}})^{\frac{4-n}{2}}.$$

We denote

$$\Lambda_j(r) = \hat{M}_j^2 \int_{d_{g_j}(p,x_j)\le r} \left(x \cdot \nabla \check{u}_j + \frac{n-4}{2} \check{u}_j \right) (\Delta^2 - P_{g_j}) \check{u}_j \, dx,$$

and

$$\tilde{\Lambda}_j(r) = \hat{M}_j^2 \int_{d_{g_j}(p,x_j)\le r} \left(x \cdot \nabla \tilde{u}_j + \frac{n-4}{2} \tilde{u}_j \right) \left(\Delta^2 - P_{g_j} + \frac{n-4}{2} Q_{g_j} \right) \tilde{u}_j \, dx$$

for $r < \delta_2$.

As in the discussion below Corollary 5.6, there exists a constant $C > 0$ independent of j such that

$$\hat{M}_j^2 \left| \int_{d_{g_j}(p,x_j)\le r} \left(x \cdot \nabla \check{u}_j + \frac{n-4}{2} \check{u}_j \right) Q_{g_j} \check{u}_j \, dx \right| \le C r^{8-n}$$

for $5 \le n \le 7$. Therefore,

$$\begin{aligned} &|\Lambda_j(r) - \tilde{\Lambda}_j(r)| \\ &\le \hat{M}_j^2 \left| \int_{|x|\le r} \left[\left(x \cdot \nabla \check{u}_j + \frac{n-4}{2} \check{u}_j \right) \left(\Delta^2 - \Delta_{g_j}^2 + \operatorname{div}_{g_j}(a_n R_{g_j} g_j - b_n \operatorname{Ric}_{g_j}) \nabla_{g_j} \right) \check{u}_j \right.\right. \\ &\qquad \left.\left. - \left(x \cdot \nabla \tilde{u}_j + \frac{n-4}{2} \tilde{u}_j \right) \left(\Delta^2 - \Delta_{g_j}^2 + \operatorname{div}_{g_j}(a_n R_{g_j} g_j - b_n \operatorname{Ric}_{g_j}) \nabla_{g_j} \right) \tilde{u}_j \right] dx \right| + C r^{8-n} \end{aligned}$$

for some constant $C > 0$ independent of j. The change of variables $y = \hat{M}_j^{2/(n-4)} x$ yields

$$\begin{aligned} \int_{|x|\le r} \Bigg\{ &\left(x \cdot \nabla \check{u}_j + \frac{n-4}{2} \check{u}_j \right) \left(\Delta^2 - \Delta_{g_j}^2 + \operatorname{div}_{g_j}(a_n R_{g_j} g_j - b_n \operatorname{Ric}_{g_j}) \nabla_{g_j} \right) \check{u}_j \\ &- \left(x \cdot \nabla \tilde{u}_j + \frac{n-4}{2} \tilde{u}_j \right) \left(\Delta^2 - \Delta_{g_j}^2 + \operatorname{div}_{g_j}(a_n R_{g_j} g_j - b_n \operatorname{Ric}_{g_j}) \nabla_{g_j} \right) \tilde{u}_j \Bigg\} \, dx \end{aligned}$$

$$
\begin{aligned}
&= \int_{|y|\le M_j^{2/(n-4)} r} \Big\{ \hat{M}_j \Big(y^k \partial_{y^k} v_j + \frac{n-4}{2} v_j \Big) \hat{M}_j^{\frac{8}{n-4}+1} \times \Big[\delta^{ab}\delta^{cd} \partial_{y^a}\partial_{y^b}\partial_{y^c}\partial_{y^d} v_j \\
&\quad - \big(g_j^{ab}(x)\partial_{y^a}\partial_{y^b} + \big(\partial_{y^a} g_j^{ap}(x) - \tfrac{1}{2} g_j^{ab} g_j^{ps} \partial_{y^s}(g_j)_{ab}\big)\partial_{y^p}\big) \\
&\qquad \times \big(g_j^{cd}\partial_{y^c}\partial_{y^d} + \big(\partial_{y^c} g_j^{cq} - \tfrac{1}{2} g_j^{cd} g_j^{qk} \partial_{y^k}(g_j)_{cd}\big)\partial_{y^q}\big) v_j \\
&\quad + (a_n - \tfrac{1}{2} b_n) \hat{M}_j^{-\frac{4}{n-4}} g_j^{pq}(x) \partial_{y^p} R_g(x) \partial_{y^q} v_j(y) \\
&\quad + a_n \hat{M}_j^{-\frac{4}{n-4}} R_{g_j}(x) \big(g_j^{pq} \partial_{y^p}\partial_{y^q} v_j + \big(\partial_{y^c} g_j^{cq} - \tfrac{1}{2} g_j^{cd} g_j^{qk} \partial_{y^k}(g_j)_{cd}\big)\partial_{y^q} v_j \big) \\
&\quad - b_n \hat{M}_j^{-\frac{4}{n-4}} \mathrm{Ric}_{g_j}^{pq}(x) \\
&\qquad \times \big(\partial_{y^p}\partial_{y^q} v_j - \tfrac{1}{2} g_j^{sk} (\partial_{y^p}(g_j)_{qk} + \partial_{y^q}(g_j)_{pk} - \partial_{y^k}(g_j)_{pq}) \partial_{y^s} v_j \big) \Big] \\
&\quad - \hat{M}_j \Big(y^k \partial_{y^k} U_0(y) + \frac{n-4}{2} U_0 \Big) \hat{M}_j^{\frac{8}{n-4}+1} \times \Big[\delta^{ab}\delta^{cd} \partial_{y^a}\partial_{y^b}\partial_{y^c}\partial_{y^d} U_0 \\
&\quad - \big(g_j^{ab}(x)\partial_{y^a}\partial_{y^b} + \big(\partial_{y^a} g_j^{ap}(x) - \tfrac{1}{2} g_j^{ab} g_j^{ps} \partial_{y^s}(g_j)_{ab}\big)\partial_{y^p}\big) \\
&\qquad \times \big(g_j^{cd}\partial_{y^c}\partial_{y^d} + \big(\partial_{y^c} g_j^{cq} - \tfrac{1}{2} g_j^{cd} g_j^{qk} \partial_{y^k}(g_j)_{cd}\big)\partial_{y^q}\big) U_0 \\
&\quad + \big(a_n - \tfrac{1}{2} b_n\big) \hat{M}_j^{-\frac{4}{n-4}} g_j^{pq}(x) \partial_{y^p} R_g(x) \partial_{y^q} U_0(y) \\
&\quad + a_n \hat{M}_j^{-\frac{4}{n-4}} R_{g_j}(x) \big(g_j^{pq} \partial_{y^p}\partial_{y^q} U_0(y) + \big(\partial_{y^c} g_j^{cq} - \tfrac{1}{2} g_j^{cd} g_j^{qk} \partial_{y^k}(g_j)_{cd}\big)\partial_{y^q} U_0 \big) \\
&\quad - b_n \hat{M}_j^{-\frac{4}{n-4}} \mathrm{Ric}_{g_j}^{pq}(x) \\
&\qquad \times \big(\partial_{y^p}\partial_{y^q} U_0 - \tfrac{1}{2} g_j^{sk} \big(\partial_{y^p}(g_j)_{qk} + \partial_{y^q}(g_j)_{pk} - \partial_{y^k}(g_j)_{pq}\big) \partial_{y^s} U_0 \big) \Big] \Big\} \hat{M}_j^{-\frac{2n}{n-4}} \, dy.
\end{aligned}
$$

Then by Lemma 5.7, one can check that

$$
\begin{aligned}
&|\Lambda_j(r) - \tilde{\Lambda}_j(r)| \\
&\quad \le c \hat{M}_j^2 \int_{|y|\le \hat{M}_j^{2/(n-4)} r} \Big[\big(|v_j(y) - U_0(y)| + |y|\,|D_y(v_j - U_0)| \big) \\
&\qquad\qquad \times \big(\hat{M}_j^{-\frac{2}{n-4}} (1+|y|)^{1-n} + \hat{M}_j^{-\frac{6}{n-4}} (1+|y|)^{3-n} \big) \\
&\qquad + |D_y(v_j - U_0)|\, \hat{M}_j^{-\frac{6}{n-4}} (1+|y|)^{4-n} + |D_y^2(v_j - U_0)| \hat{M}_j^{-\frac{4}{n-4}} (1+|y|)^{4-n} \\
&\qquad + |D_y^3(v_j - U_0)| \hat{M}_j^{-\frac{6}{n-4}} (1+|y|)^{4-n} \Big] \, dy + C r^{8-n} \\
&\quad \le cr + C r^{8-n} \le Cr.
\end{aligned}
$$

Also, by the construction of conformal normal coordinates,

$$
\begin{aligned}
&|\tilde{\Lambda}_j(r)| \\
&\quad = \hat{M}_j^2 \int_{|x|\le r} \Big| \Big(x \cdot \nabla \tilde{u}_j + \frac{n-4}{2} \tilde{u}_j \Big) \big(\Delta^2 - \Delta_{g_j}^2 + \mathrm{div}_{g_j} (a_n R_{g_j} g_j - b_n \mathrm{Ric}_{g_j}) \nabla_{g_j} \big) \tilde{u}_j \, dx \Big|
\end{aligned}
$$

$$\le c\hat{M}_j^2 \int_{|y|\le M_j^{2/(n-4)}r} \hat{M}_j(1+|y|)^{4-n}\hat{M}_j^{\frac{8}{n-4}+1}$$
$$\times\Big[\hat{M}_j^{-\frac{6}{n-4}}|x|^{N-3}(1+|y|)^{3-n} + \hat{M}_j^{-\frac{4}{n-4}}|x|^{N-2}(1+|y|)^{2-n}+\hat{M}_j^{-\frac{2}{n-4}}|x|^{N-1}(1+|y|)^{1-n}\Big]\hat{M}_j^{-\frac{2n}{n-4}}\,dy$$
$$\le C(r^{N+4-n}+\hat{M}_j^{2-\frac{2N}{n-4}}).$$

Therefore, (5-47) holds for $5\le n\le 7$. □

Proposition 5.9. *Let (M^n, g) be a closed Riemannian manifold of dimension $5\le n\le 7$ with $R_g\ge 0$, and also $Q_g\ge 0$ with $Q_g(p)>0$ for some point $p\in M$. Let $\{u_j\}$ be a sequence of positive solutions to (1-2) and $x_j\to\bar{x}$ be an isolated simple blowup point so that*

$$u_j(x_j)u_j(p)\to H(p)\quad \text{in } C^{4,\alpha}_{\text{loc}}(B_{\delta_2}(\bar{x})-\{\bar{x}\}),$$

for some $0<\alpha<1$. Assume that for some constants $a>0$ and A,

$$H(p)=\frac{a}{d_g(p,\bar{x})^{n-4}}+A+o(1)\quad \text{as } d_g(p,\bar{x})\to 0, \tag{5-48}$$

for $n=5$, or

$$\hat{H}(p)\equiv\rho^{-1}(\bar{x})\rho^{-1}(p)H(p)=\frac{a}{d_{g_0}(p,\bar{x})^{n-4}}+A+o(1)\quad \text{as } d_{g_0}(p,\bar{x})\to 0, \tag{5-49}$$

for $5\le n\le 7$, where $g_0=\rho^{4/(n-4)}g$ is the conformal metric corresponding to the conformal normal coordinates centered at $\bar{x}$. Then $A=0$.

Proof. Let us first consider $n=5$ under the condition (5-48).

Let $x=(x^1,\dots,x^n)$ be the geodesic normal coordinates at x_j for each j. Denote $\Omega_{\gamma,j}=B_\gamma(x_j)$ for $\gamma<\delta_2/(2)$. Then $\Omega_{\gamma,j}\to\Omega_\gamma=B_\gamma(\bar{x})$. By the Pohozaev identity,

$$\int_{\partial\Omega_{\gamma,j}}\frac{n-4}{2}\Big(u_j\frac{\partial}{\partial\nu}(\Delta u_j)-\Delta u_j\frac{\partial}{\partial\nu}u_j\Big) + \Big((x\cdot\nabla u_j)\frac{\partial}{\partial\nu}(\Delta u_j)-\frac{\partial}{\partial\nu}(x\cdot\nabla u_j)\Delta u_j+\tfrac12(\Delta u_j)^2x\cdot\nu\Big)\,ds$$
$$=\int_{\Omega_{\gamma,j}}\Big(x\cdot\nabla u_j+\frac{n-4}{2}u_j\Big)(\Delta^2-P_g)u_j\,dx+\frac{(n-4)^2}{4n}\bar{Q}\int_{\partial\Omega_{\gamma,j}}(x\cdot\nu)u_j^{\frac{2n}{n-4}}\,dx.$$

Multiplying $M_j^2=u_j(x_j)^2$ on both sides and taking $\lim_{\gamma\to0^+}\limsup_{j\to\infty}$ on both sides, we have that by Corollary 5.6,

$$\lim_{\gamma\to0}\limsup_{j\to\infty}M_j^2\int_{\Omega_{\gamma,j}}\Big(x\cdot\nabla u_j+\frac{n-4}{2}u_j\Big)(\Delta^2-P_g)u_j\,dx=0,$$

and

$$\lim_{\gamma\to 0}\bigg[\int_{\partial\Omega_\gamma} \frac{n-4}{2}\Big(H\frac{\partial}{\partial\nu}(\Delta H)-\Delta H\frac{\partial}{\partial\nu}H\Big) + \Big((x\cdot\nabla H)\frac{\partial}{\partial\nu}(\Delta H)-\frac{\partial}{\partial\nu}(x\cdot\nabla H)\Delta H+\tfrac{1}{2}(\Delta H)^2 x\cdot\nu\Big)\,ds\bigg]$$

$$= \lim_{\gamma\to 0}\limsup_{j\to\infty} M_j^2\int_{\partial\Omega_{\gamma,j}}\bigg[\frac{n-4}{2}\Big(u_j\frac{\partial}{\partial\nu}(\Delta u_j)-\Delta u_j\frac{\partial}{\partial\nu}u_j\Big) + \Big((x\cdot\nabla u_j)\frac{\partial}{\partial\nu}(\Delta u_j)-\frac{\partial}{\partial\nu}(x\cdot\nabla u_j)\Delta u_j+\tfrac{1}{2}(\Delta u_j)^2 x\cdot\nu\Big)\bigg]\,ds$$

$$= \lim_{\gamma\to 0}\limsup_{j\to\infty} M_j^{-\frac{8}{n-4}}\int_{\partial\Omega_{\gamma,j}}(x\cdot\nu)(M_j\,u_j)^{\frac{2n}{n-4}}\,dx = 0.$$

By assumption,

$$\lim_{\gamma\to 0}\bigg[\int_{\partial\Omega_\gamma} \frac{n-4}{2}\Big(H\frac{\partial}{\partial\nu}(\Delta H)-\Delta H\frac{\partial}{\partial\nu}H\Big) + \Big((x\cdot\nabla H)\frac{\partial}{\partial\nu}(\Delta H)-\frac{\partial}{\partial\nu}(x\cdot\nabla H)\Delta H+\tfrac{1}{2}(\Delta H)^2 x\cdot\nu\Big)\,ds\bigg]$$

$$= \lim_{\gamma\to 0}\int_{\partial\Omega_\gamma}(n-4)^2(n-2)aA|x|^{1-n}\,ds$$

$$= (n-4)^2(n-2)aA|\mathbb{S}^{n-1}|,$$

where $|\mathbb{S}^{n-1}|$ is the area of an $(n-1)$-dimensional round sphere. Therefore,

$$A = 0.$$

For $5\le n\le 7$ under the condition (5-49), for each j, let $x=(x^1,\dots,x^n)$ be the conformal normal coordinates of (Ω, g) centered at x_j and $g_j=\rho_j^{4/(n-4)}g$ the corresponding conformal metrics defined as in the paragraph preceding Lemma 5.7. Denote $\Omega_{\gamma,j}=B_\gamma(x_j)$ with respect to the metric g_j, for $\gamma<\delta_2/2$. Then

$$\Omega_{\gamma,j}\to\Omega_\gamma = B_\gamma(\bar{x}).$$

By the Pohozaev identity,

$$\int_{\partial\Omega_{\gamma,j}}\frac{n-4}{2}\Big(\check{u}_j\frac{\partial}{\partial\nu}(\Delta\check{u}_j)-\Delta\check{u}_j\frac{\partial}{\partial\nu}\check{u}_j\Big) + \Big((x\cdot\nabla\check{u}_j)\frac{\partial}{\partial\nu}(\Delta\check{u}_j)-\frac{\partial}{\partial\nu}(x\cdot\nabla\check{u}_j)\Delta\check{u}_j+\tfrac{1}{2}(\Delta\check{u}_j)^2 x\cdot\nu\Big)\,ds$$

$$=\int_{\Omega_{\gamma,j}}\Big(x\cdot\nabla\check{u}_j+\frac{n-4}{2}\check{u}_j\Big)(\Delta^2-P_{g_j})\check{u}_j\,dx+\frac{(n-4)^2}{4n}\,\overline{Q}\int_{\partial\Omega_{\gamma,j}}(x\cdot\nu)\check{u}_j^{\frac{2n}{n-4}}\,dx,$$

where $\check{u}_j = u_j \rho_j^{-1}$. Note that

$$\check{u}_j(p)\check{u}_j(x_j) \to H(p)\rho(\bar{x})^{-1}\rho(p)^{-1} = \hat{H}(p),$$

in

$$C^{4,\alpha}_{\mathrm{loc}}(B_{\delta_2/2}(\bar{x}) - \{\bar{x}\}).$$

Multiplying $\hat{M}_j^2 = \check{u}_j(x_j)^2$ on both sides of the identity and taking the limit $\lim_{\gamma\to 0^+} \limsup_{j\to\infty}$ on both sides, we have that by Corollary 5.8,

$$\lim_{\gamma\to 0} \limsup_{j\to\infty} \hat{M}_j^2 \int_{\Omega_{\gamma,j}} \Big(x\cdot\nabla\check{u}_j + \frac{n-4}{2}\check{u}_j \Big)(\Delta^2 - P_{g_j})\check{u}_j \, dx = 0,$$

and

$$\begin{aligned}
&\lim_{\gamma\to 0}\Big[\int_{\partial\Omega_\gamma} \frac{n-4}{2}\Big(\hat{H}\frac{\partial}{\partial\nu}(\Delta\hat{H}) - \Delta\hat{H}\frac{\partial}{\partial\nu}\hat{H}\Big) \\
&\qquad + \Big((x\cdot\nabla\hat{H})\frac{\partial}{\partial\nu}(\Delta\hat{H}) - \frac{\partial}{\partial\nu}(x\cdot\nabla\hat{H})\Delta\hat{H} + \tfrac{1}{2}(\Delta\hat{H})^2 x\cdot\nu\Big)\, ds\Big] \\
&\quad = \lim_{\gamma\to 0}\limsup_{j\to\infty} \hat{M}_j^2 \int_{\partial\Omega_{\gamma,j}} \Big[\frac{n-4}{2}\Big(\check{u}_j\frac{\partial}{\partial\nu}(\Delta\check{u}_j) - \Delta\check{u}_j\frac{\partial}{\partial\nu}\check{u}_j\Big) \\
&\qquad + \Big((x\cdot\nabla\check{u}_j)\frac{\partial}{\partial\nu}(\Delta\check{u}_j) - \frac{\partial}{\partial\nu}(x\cdot\nabla\check{u}_j)\Delta\check{u}_j + \tfrac{1}{2}(\Delta\check{u}_j)^2 x\cdot\nu\Big)\Big]\, ds \\
&\quad = \lim_{\gamma\to 0}\limsup_{j\to\infty} \hat{M}_j^{-\frac{8}{n-4}} \int_{\partial\Omega_{\gamma,j}} (x\cdot\nu)(\hat{M}_j\,\check{u}_j)^{\frac{2n}{n-4}}\, dx = 0.
\end{aligned}$$

By assumption,

$$\begin{aligned}
&\lim_{\gamma\to 0}\Big[\int_{\partial\Omega_\gamma} \frac{n-4}{2}\Big(\hat{H}\frac{\partial}{\partial\nu}(\Delta\hat{H}) - \Delta\hat{H}\frac{\partial}{\partial\nu}\hat{H}\Big) + \\
&\qquad \Big((x\cdot\nabla\hat{H})\frac{\partial}{\partial\nu}(\Delta\hat{H}) - \frac{\partial}{\partial\nu}(x\cdot\nabla\hat{H})\Delta\hat{H} + \tfrac{1}{2}(\Delta\hat{H})^2 x\cdot\nu\Big)\, ds\Big] \\
&\quad = \lim_{\gamma\to 0}\int_{\partial\Omega_\gamma} (n-4)^2(n-2)aA|x|^{1-n}\, ds \\
&\quad = (n-4)^2(n-2)aA|\mathbb{S}^{n-1}|,
\end{aligned}$$

where $|\mathbb{S}^{n-1}|$ is the area of an $(n-1)$-dimensional round sphere. Therefore,

$$A = 0. \qquad \square$$

Remark. It is easy to check that all conclusions in this section hold for an isolated (respectively, simple) blowup point $x_j \to \bar{x}$ of a sequence of solutions $\{v_j\}_j$ to (1-2), with the background metric g replaced by a sequence of rescaled metrics $g_j = T_j g$ corresponding to a sequence of positive numbers $T_j \to \infty$ as $j \to \infty$. In this situation, $\rho \equiv 1$ in (5-49) in Proposition 5.9.

6. From isolated blowup points to isolated simple blowup points

In this section we show that an isolated blowup point is an isolated simple blowup point.

Proposition 6.1. *Let (M^n, g) be a closed Riemannian manifold of dimension $5 \le n \le 7$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p_0) > 0$ for some point $p_0 \in M$. Let $\{u_j\}$ be a sequence of positive solutions to (1-2) and $x_j \to \bar{x}$ be an isolated blowup point. Let $M_j = u_j(x_j)$. Then $\bar{x}$ is an isolated simple blowup point.*

Proof. We prove the proposition by a contradiction argument. Assume that $\bar{x}$ is not an isolated simple blowup point. Then there exist two critical points of $r^{(n-4)/2}\bar{u}_j(r)$ in $(0, \bar{\mu}_j)$ with some $\bar{\mu}_j \to 0$ up to a subsequence as $j \to \infty$. By Lemma 5.2 with $0 < \epsilon_j < e^{-T_j}$, we have $r^{(n-4)/2}\bar{u}_j(r)$ has precisely one critical point in $(0, T_j M_j^{-2/(n-4)})$. We choose μ_j to be the second critical point of $r^{(n-4)/2}\bar{u}_j(r)$ so that $\mu_j \ge T_j M_j^{-2/(n-4)}$ and by assumption $\mu_j \to 0$.

For each j let $x = (x^1, \ldots, x^n)$ be the geodesic normal coordinates centered at x_j, and let $y = \mu_j^{-1} x$. For ease of notation, we assume $\delta_2 = 1$. We define the scaled metric $\tilde{g}_j = \mu_j^{-2} g$ so that $(\tilde{g}_j)_{pq}(\mu_j^{-1}x)dx^p dx^q = g_{pq}(x)dx^p dx^q$, and

$$\xi_j(y) = \mu_j^{\frac{n-4}{2}} u_j(\exp_{x_j}(\mu_j y)) \quad \text{for } |y| < \mu_j^{-1}.$$

We denote $\bar{\xi}_j$ as the spherical average of ξ_j. Then we have:

(6-1) $P_{\tilde{g}_j}\xi_j(y) = \frac{n-4}{2}\bar{Q}\xi_j(y)^{(n+4)/(n-4)}$, where $|y| < \mu_j^{-1}$,

(6-2) $|y|^{(n-4)/2}\xi_j(y) \le C$, where $|y| < \mu_j^{-1}$.

(6-3) $\lim_{j\to\infty} \xi_j(0) = \infty$.

(6-4) $-\frac{4(n-1)}{n-2}\Delta_{\tilde{g}_j}\xi_j^{(n-2)/(n-4)} + R_{\tilde{g}_j}\xi_j^{(n-2)/(n-4)} \ge 0$, where $|y| < \mu_j^{-1}$.

(6-5) $r^{(n-4)/2}\bar{\xi}_j(r)$ has precisely one critical point in $0 < r < 1$.

(6-6) $\frac{d}{dr}(r^{(n-4)/2}\bar{\xi}_j(r)) = 0$ at $r = 1$.

Therefore $\{0\}$ is an isolated simple blowup point of the sequence $\{\xi_j\}$. Note that the remark on page 138 holds for u_j so

$$\xi_j(0)\xi_j(y) \ge C|y|^{4-n} \quad \text{for } |y| \ge \mu_j^{-1} T_j M_j^{-\frac{2}{n-4}}, \tag{6-7}$$

where $\mu_j^{-1} T_j M_j^{-2/(n-4)} \le 1$. By Lemma 5.1, there exists $C > 0$ independent of j and k so that for any $k \in \mathbb{R}$,

$$\max_{2^k \le |y| \le 2^{k+1}} \xi_j(0)\xi_j(y) \;\le\; C \min_{2^k \le |y| \le 2^{k+1}} \xi_j(0)\xi_j(y), \;\text{ when } 2^{k+1} < \mu_j^{-1}\frac{\delta_2}{3}. \tag{6-8}$$

Note that $Q_{\tilde{g}_j} \ge 0$ and $R_{\tilde{g}_j} > 0$ in M. Also the rescaled metrics $\tilde{g}_j$ are all well controlled in $|y| \le 1$. In the proof of Lemma 5.4 the maximum principle holds for $\tilde{g}_j$ and the coefficients of the test function are still uniformly chosen for $\tilde{g}_j$ so that the estimate in Lemma 5.4 holds for each ξ_j in $|y| \le \tilde{\delta}_2$ for some $\tilde{\delta}_2 < 1$ independent of j. Hence Proposition 5.3 holds for ξ_j in $|y| \le \tilde{\delta}_2$. This combined with (6-7) and (6-8) implies

$$C(K)^{-1} \le \xi_j(0)\xi_j(y) \le C(K)$$

for $K \Subset \mathbb{R}^n - \{0\}$ when j is large; moreover, $\tilde{g}_j$ converges to the flat metric and there exists $a > 0$ such that $\xi_j(0)\xi_j(y)$ converges to

$$H(y) = a|y|^{4-n} + b(y) \quad \text{in } C^4_{\text{loc}}(\mathbb{R}^n - \{0\}),$$

where $b(y) \in C^4(\mathbb{R}^n)$ satisfies

$$\Delta^2 b = 0$$

in $\mathbb{R}^n$. Here $H > 0$ in $\mathbb{R}^n - \{0\}$. Also,

$$-\Delta H(y)^{\frac{n-2}{n-4}} \ge 0, \quad |y| > 0. \tag{6-9}$$

Moreover, for a fixed point y_0 in $|y| = 1$, by (6-8),

$$H(y) \le |y|^{2+\frac{\ln C}{\ln 2}} H(y_0)$$

for $|y| \ge 1$. Since $H > 0$ for $|y| > 0$, it follows that $b(y)$ is a polyharmonic function of polynomial growth on $\mathbb{R}^n$. Therefore, $b(y)$ must be a polynomial in $\mathbb{R}^n$; see [Armitage 1973]. Nonnegativity of H near infinity implies that $b(y)$ is of even order. Then either $b(y)$ is a nonnegative constant or $b(y)$ is a polynomial of even order with order at least two and $b(y)$ is nonnegative at infinity. The later case contradicts (6-9) for y near infinity. Therefore, $b(y)$ must be a nonnegative constant on $\mathbb{R}^n$ and

$$H(y) = a|y|^{4-n} + b$$

with a constant $a > 0$ and a constant b.

By (6-6),

$$\frac{d}{dr}(r^{\frac{n-4}{2}} H(r)) = 0 \quad \text{at } r = 1.$$

We then have $b = a > 0$, which contradicts Proposition 5.9. In fact, Proposition 5.9 applies to isolated simple blowup points with respect to the sequence of rescaled metrics $\{\tilde{g}_j\}$ with uniform curvature bound and uniform bound of injectivity radius with the property that $Q_{\tilde{g}_j} > 0$ and $R_{\tilde{g}_j} > 0$ (see the proof of Proposition 5.9). Here $\hat{H} = H$ in the condition (5-49). Indeed, for $n = 6, 7$, after rescaling, the conformal metric $g_j = \rho_j^{4/(n-4)} g$ corresponding to the conformal normal coordinates centered at x_j becomes $\hat{g}_j(y) = \mu_j^{-2} \rho_j(\mu_j y)^{4/(n-4)} g(\mu_j y)$ and the functions $\hat{\rho}_j(y) = \rho_j(\mu_j y) \to \rho(y) \equiv 1$ locally uniformly in C^N as $j \to +\infty$. This completes the proof of Proposition 6.1. $\square$

Remark. It is easy to check the proof of Proposition 6.1 shows that an isolated blowup point $x_j \to \bar{x}$ of a sequence of solutions $\{v_j\}_j$ to (1-2), with the background metric g replaced by a sequence of rescaled metrics $g_j = T_j g$ corresponding to a sequence of positive numbers $T_j \to \infty$ as $j \to \infty$, is in fact an isolated simple blowup point.

7. Compactness of solutions to the constant Q-curvature equations

Based on Propositions 5.3 and 6.1, the proof of compactness of the solutions is more or less standard; see, for example, [Li and Zhu 1999]. But again we need to deal with the limit of the blowup argument carefully, which satisfies a fourth order elliptic equation; see Lemma 7.1 and Proposition 7.3.

We first show that there are no bubble accumulations.

Lemma 7.1. *Let (M^n, g) be a closed Riemannian manifold of dimension $5 \le n \le 9$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p_0) > 0$ for some point $p_0 \in M$. For any given $\epsilon > 0$ and large constant $T > 1$, there exists some constant $C_1 > 0$ depending on M, g, ϵ, T, $\|Q_g\|_{C^1(M)}$ such that for any solution u to (1-2) and any compact subset $K \subset M$ satisfying*

$$\max_{p \in M - K} d(p, K)^{\frac{n-4}{2}} u(p) \ge C_1 \quad \text{if } K \neq \varnothing$$

and

$$\max_{p \in M} u(p) \ge C_1 \quad \text{if } K = \varnothing,$$

we have that there exists some local maximum point p' of u in $M - K$ with $B_{T\,u(p')^{-2/(n-4)}}(p') \subset M - K$ satisfying

$$\|u(p')^{-1} u(\exp_{p'}(u(p')^{-\frac{2}{n-4}} y)) - (1 + 4^{-1}|y|^2)^{-\frac{n-4}{2}}\|_{C^4(|y| \le 2T)} < \epsilon. \tag{7-1}$$

Proof. We argue by contradiction. That is to say, there exists a sequence of compact subsets K_j and a sequence of solutions u_j to (1-2) on M such that

$$\max_{p \in M - K_j} d(p, K_j)^{\frac{n-4}{2}} u(p) \ge j,$$

but no point satisfies (7-1) (here $d(p, K_j) = 1$ when $K_j = \varnothing$). We choose $x_j \in M - K_j$ satisfying

$$d_g(x_j, K_j)^{\frac{n-4}{2}} u_j(x_j) = \max_{p \in M - K_j} d_g(p, K_j)^{\frac{n-4}{2}} u_j(p).$$

Denote $T_j \equiv \frac{1}{4} u_j(x_j)^{2/(n-4)} d_g(x_j, K_j)$. We then define

$$v_j(y) = u_j(x_j)^{-1} u_j(\exp_{x_j}(u_j(x_j)^{-\frac{2}{n-4}} y)) \quad \text{for } |y| \le T_j.$$

Let $h_j = u_j(x_j)^{4/(n-4)} g$. The rescaled function v_j satisfies

$$P_{h_j} v_j = \frac{n-4}{2} \bar{Q} v_j^{\frac{n+4}{n-4}}, \tag{7-2}$$

and by Theorem 2.1,

$$\Delta_{h_j} v_j \le \frac{n-4}{4(n-1)} R_{h_j} v_j. \tag{7-3}$$

We will analyze the limit of the sequence $\{v_j\}$ as in Theorem 3.3 and conclude that (7-1) indeed holds. By assumption,

$$T_j \equiv \tfrac{1}{4} u_j(x_j)^{\frac{2}{n-4}} d_g(y_j, K_j) \ge \tfrac{1}{4} j^{\frac{2}{n-4}},$$

and

$$d_g(\exp_{x_j}(u_j(x_j)^{-\frac{2}{n-4}} y), K_j) \ge \tfrac{1}{2} d_g(x_j, K_j) \quad \text{for } |y| \le T_j.$$

It follows that

$$\begin{aligned} 0 < v_j(y) &= u_j(x_j)^{-1} u_j(\exp_{x_j}(u_j(x_j)^{-\frac{2}{n-4}} y)) \\ &\le u_j(x_j)^{-1} d_g(\exp_{x_j}(u_j(x_j)^{-\frac{2}{n-4}} y), K_j)^{-\frac{n-4}{2}} d_g(x_j, K_j)^{\frac{n-4}{2}} u_j(x_j) \\ &\le 2^{\frac{n-4}{2}} \quad \text{for } |y| \le T_j. \end{aligned}$$

Standard elliptic estimates imply that up to a subsequence,

$$v_j \to v \quad \text{in } C^4_{\text{loc}}(\mathbb{R}^n),$$

with v satisfying

$$\begin{aligned} \Delta^2 v &= \frac{n-4}{2} \bar{Q} v^{\frac{n+4}{n-4}} \quad \text{in } \mathbb{R}^n, \\ v(0) &= 1, \quad 0 \le v \le 2^{\frac{n-4}{2}} \quad \text{in } \mathbb{R}^n, \\ \Delta v &\le 0, \quad \text{in } \mathbb{R}^n. \end{aligned}$$

By the strong maximum principle, $v > 0$ in $\mathbb{R}^n$. Then by the classification theorem of C.S. Lin [1998]),

$$v(y) = \left(\frac{\lambda}{1 + 4^{-1} \lambda^2 |y - \bar{y}|^2} \right)^{\frac{n-4}{2}} \quad \text{in } \mathbb{R}^n,$$

with $v(0) = 1$ and $v(y) \le \lambda^{(n-4)/2} \le 2^{(n-4)/2}$. Therefore, $|\bar{y}| \le C(n)$ with $C(n) > 0$ only depending on n. We choose y_j to be the local maximum point of v_j converging to $\bar{y}$. Then $p_j = \exp_{x_j}(u_j(x_j)^{-2/(n-4)} y_j) \in M - K_j$ is a local maximum point of u_j. We now repeat the blowup argument with x_j replaced by p_j and $u_j(x_j)$ replaced by $u_j(p_j)$ and obtain the limit

$$v(y) = (1 + 4^{-1} |y|^2)^{-\frac{n-4}{2}} \quad \text{in } \mathbb{R}^n.$$

Therefore, for large j, there exists $p_j \in M - K_j$ such that (7-1) holds. This contradicts the assumption. Therefore, the proof of the lemma is completed. □

Lemma 7.2. *Let (M^n, g) be a closed Riemannian manifold of dimension $5 \le n \le 9$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p_0) > 0$ for some point $p_0 \in M$. For any given $\epsilon > 0$ and a large constant $T > 1$, there exist some constants $C_1 > 0$ and $C_2 > 0$ depending on $M, g, \epsilon, T, \|Q_g\|_{C^1(M)}$ such that for any solution u to (1-2) with*

$$\max_{p\in M} u(p) > C_1,$$

there exists some integer $N = N(u)$ depending on u and N local maximum points $\{p_1, \dots, p_N\}$ of u such that:

(i) *For $i \ne j$,*

$$\overline{B_{\gamma_i}(p_i)} \cap \overline{B_{\gamma_j}(p_j)} = \varnothing,$$

with $\gamma_j = Tu(p_j)^{-2/(n-4)}$ and $B_{\gamma_j}(p_j)$ the geodesic γ_j-ball centered at p_j, and

$$\|u(p_j)^{-1}u(\exp_{p_j}(u(p_j)^{-\frac{2}{n-4}}y)) - (1+4^{-1}|y|^2)^{-\frac{n-4}{2}}\|_{C^4(|y|\le 2R)} < \epsilon, \tag{7-4}$$

where $y = u(p_j)^{2/(n-4)}x$, with x geodesic normal coordinates centered at p_j, and $|y| = \sqrt{(y^1)^2 + \cdots + (y^n)^2}$.

(ii) *For $i < j$, we have $d_g(p_i, p_j)^{(n-4)/2}u(p_j) \ge C_1$, while for $p \in M$,*

$$d_g(p, \{p_1, \dots, p_n\})^{\frac{n-4}{2}} u(p) \le C_2.$$

Proof. We will use Lemma 7.1 and prove the lemma by induction. To start, we apply Lemma 7.1 with $K = \varnothing$. We choose p_1 to be a maximum point of u and thus (7-4) holds. Next we let $K = \overline{B_{\gamma_1}(p_1)}$.

Assume that for some $i_0 \ge 1$, (i) holds for $1 \le j \le i_0$ and $1 \le i < j$, and also $d_g(p_i, p_j)^{(n-4)/2}u(p_j) \ge C_1$ with p_j chosen as in Lemma 7.1 by induction (this holds for $i_0 = 1$). Then we let $K = \bigcup_{j=1}^{i_0} \overline{B_{\gamma_j}(p_j)}$. It follows that for $\epsilon > 0$ small, for any p such that $d_g(p, p_j) \le 2\gamma_j$ with $1 \le j \le i_0$, we have

$$\begin{aligned} d_g(p, \{p_1, \dots, p_{i_0}\})^{\frac{n-4}{2}} u(p) &\le d_g(p, p_j)^{\frac{n-4}{2}} u(p) \le 2 d_g(p, p_j)^{\frac{n-4}{2}} u(p_j) \\ &\le 2(2Tu(p_j)^{-\frac{2}{n-4}})^{\frac{n-4}{2}} u(p_j) = 2^{\frac{n-2}{2}} T^{\frac{n-4}{2}}, \end{aligned}$$

and therefore, for $p \in \bigcup_{j=1}^{i_0} \overline{B_{2\gamma_j}(p_j)}$,

$$d_g(p, \{p_1, \dots, p_{i_0}\})^{\frac{n-4}{2}} u(p) \le 2^{\frac{n-2}{2}} T^{\frac{n-4}{2}}. \tag{7-5}$$

If, for all $p \in M$, the inequality

$$d_g(p, \{p_1, \dots, p_{i_0}\})^{\frac{n-4}{2}} u(p) \le C_1,$$

holds then the induction stops. Otherwise, we apply Lemma 7.1, and we denote p_{i_0+1} as the local maximum point y_0 obtained in Lemma 7.1 so that

$$B_{T\,u(p_{i_0+1})^{-2/(n-4)}}(p_{i_0+1}) \subset M - K.$$

Thus, (i) holds for i_0+1. Also, by assumption, $d_g(p_j, p_{i_0+1})^{(n-4)/2}u(p_{i_0+1}) > C_1$. By the same argument, (7-5) holds for i_0 replaced by i_0+1. The induction must stop in a finite time $N = N(u)$, since $\int_M u^{2n/(n-4)}\,dV_g$ is bounded and

$$\int_{B_{\gamma_j}(p_j)} u^{\frac{2n}{n-4}}\,dV_g$$

is bounded below by a uniform positive constant. It is clear now that for $p \in M - \bigcup_{j=1}^N B_{\gamma_j}(p_j)$,

$$d(p, \{p_1, \dots, p_N\})^{\frac{n-4}{2}} u(p) \le 2^{\frac{n-4}{2}} d\biggl(p, \bigcup_{j=1}^N B_{\gamma_j}(p_j)\biggr)^{\frac{n-4}{2}} u(p) \le 2^{\frac{n-4}{2}} C_1.$$

By induction, (7-5) holds for i_0 replaced by N. We set

$$C_2 = 2^{\frac{n-2}{2}} T^{\frac{n-4}{2}} + 2^{\frac{n-4}{2}} C_1. \qquad \square$$

The next proposition rules out the bubble accumulations.

Proposition 7.3. *Let (M^n, g) be a closed Riemannian manifold of dimension $5 \le n \le 7$ with $R_g \ge 0$, and also $Q_g \ge 0$ with $Q_g(p_0) > 0$ for some point $p_0 \in M$. For $\epsilon > 0$ small enough and a constant $T > 1$ large enough, there exists $\gamma > 0$ depending on M, g, ϵ, T, $\|R_g\|_{C^1(M)}$ and $\|Q_g\|_{C^1(M)}$ such that for any solution u to (1-2) with $\max_{p\in M} u(p) > C_1$, we have*

$$d(p_i, p_j) \ge \gamma,$$

for $1 \le i,\ j \le N$ and $i \ne j$, where $N = N(u)$, $p_j = p_j(u)$, $p_i = p_i(u)$ and C_1 are defined in Lemma 7.2.

Proof. Suppose the proposition fails, which implies that there exist $\epsilon > 0$ small and $T > 0$ large and a sequence of solutions u_j to (1-2) such that $\max_{p\in M} u_j(p) > C_1$ and

$$\lim_{j\to\infty} \min_{i\ne k} d(p_i(u_j), p_k(u_j)) = 0.$$

We denote $p_{1,j}$ and $p_{2,j}$ to be the two points realizing the minimum distance in $\{p_1(u_j), \dots, p_N(u_j)\}$ of u_j constructed in Lemma 7.2. Let $\bar\gamma_j = d_g(p_{1,j}, p_{2,j})$. Since

$$B_{T u_j(p_{1,j})^{-2/(n-4)}}(p_{1,j}) \cap B_{T u_j(p_{2,j})^{-2/(n-4)}}(p_{2,j}) = \varnothing,$$

we have that $u_j(p_{1,j}) \to \infty$ and $u_j(p_{2,j}) \to \infty$.

For each j, let $x = (x^1, \dots, x^n)$ be the geodesic normal coordinates centered at $p_{1,j}$, $y = \bar{\gamma}_j^{-1} x$, and $\exp_{p_{1,j}}(x)$ be exponential map under the metric g. We define the scaled metric $h_j = \bar{\gamma}_j^{-2} g$, and the rescaled function

$$v_j(y) = \bar{\gamma}_j^{\frac{n-4}{2}} u_j(\exp_{p_{1,j}}(\bar{\gamma}_j y)).$$

It follows that v_j satisfies $v_j > 0$ in $|y| \le \bar{\gamma}_j^{-1} r_0$ and that

$$P_{h_j} v_j(y) = \frac{n-4}{2} \bar{Q} v_j(y)^{\frac{n+4}{n-4}} \quad \text{for } |y| \le \bar{\gamma}_j^{-1} r_0, \tag{7-6}$$

$$\Delta_{h_j} v_j \le \frac{(n-4)}{4(n-1)} R_{h_j} v_j \qquad \text{for } |y| \le \bar{\gamma}_j^{-1} r_0, \tag{7-7}$$

where r_0 is half of the injectivity radius of (M, g). We define $y_k = y_k(u_j) \in \mathbb{R}^n$ such that $\exp_{p_{1,j}}(\bar{\gamma}_j y_k) = p_k$ for the points $p_k(u_j)$. It follows that for $p_k \ne p_{1,j}$,

$$|y_k| \ge 1 + o(1)$$

with $o(1) \to 0$ as $j \to \infty$. Let $y_{2,j} \in \mathbb{R}^n$ be such that $p_{2,j} = \exp_{p_{1,j}}(\bar{\gamma}_j y_{2,j})$. Then

$$|y_{2,j}| \to 1 \quad \text{as } j \to \infty.$$

It follows that there exists $\bar{y} \in \mathbb{R}^n$ with $|\bar{y}| = 1$ such that up to a subsequence,

$$\bar{y} = \lim_{j \to \infty} y_{2,j}.$$

By Lemma 7.2,

$$\bar{\gamma}_j \ge C \max\{T u_j(p_{1,j})^{-\frac{2}{n-4}},\ T u_j(p_{2,j})^{-\frac{2}{n-4}}\}.$$

Thus, $v_j(0) \ge C_3$, $v_j(y_{2,j}) \ge C_3$ for some $C_3 > 0$ independent of j, y_k is a local maximum point of v_j for all $1 \le k \le N(u_j)$, and $\min_{1 \le k \le N(u_j)} |y - y_k|^{(n-4)/2} v_j(y) \le C_2$ for all $|y| \le \bar{\gamma}_j^{-1}$.

We claim that either

$$v_j(0) \to \infty \quad \text{and} \quad v_j(y_{2,j}) \to \infty, \tag{7-8}$$

or both of these two sequences are uniformly bounded. To see this, we first assume that one of them tends to infinity up to a subsequence, say $v_j(0) \to \infty$ for instance. It is clear that 0 is an isolated blowup point, and by Proposition 6.1 it is an isolated simple blowup point. Then $v_j(y_{2,j}) \to \infty$ in this subsequence since otherwise, by the control (7-4) at $p_{2,j}$ in Lemma 7.2 and the rescaling, the upper bound of v_j in the $\frac{1}{2}$-geodesic ball centered at $y_{2,j}$ under h_j is controlled by the lower bound of v_j in it up to a uniform multiplier, and thus by the Harnack inequality (5-1) in $B_{4/5}(0) - B_{1/5}(0)$ and Proposition 5.3, $v_j \to 0$ in $B_{1/2}(p_{2,j})$, contradicting $v_j(y_{2,j}) \ge C_3$. The claim is established. If v_j are uniformly bounded on any fixed compact subset of $\mathbb{R}^n$, then as discussed in Lemma 7.1, $v_j \to v$ in $C^4_{\mathrm{loc}}(\mathbb{R}^n)$ with $v > 0$ and

$$\Delta^2 v = \frac{n-4}{2}\bar{Q} v^{\frac{n+4}{n-4}}$$

in $\mathbb{R}^n$. Also, 0 and $\bar{y}$ are local maximum points of v. That contradicts the classification theorem in [Lin 1998]. Therefore, the set (denoted as K_0) of isolated blowup points of $\{v_j\}$ is nonempty. Hence v_j is uniformly bounded on any compact subset in $\mathbb{R}^n - K_0$. By a similar argument as the claim, there are at least two points in K_0 and for any two distinct points $y, z \in K$, $|y - z| \geq 1$. Also, by Proposition 6.1 (see also the remark on page 159), K_0 is a set of isolated simple blowup points.

Choose any two blowup points $y_{m,j} \to y_m$ and $y_{k,j} \to y_k \in K_0$. For j large, we pick a point p on the $\frac{1}{2}$-geodesic sphere of $y_{k,j}$. Now we apply Theorem 3.3 (see also the remark on page 138) about the blowup point y_m of v_j at p and Proposition 5.3 about the blowup point y_k of v_j at p; then we have that there exists a constant $C > 0$ independent of j such that

$$v_j(y_{m,j}) \geq C v_j(y_{k,j}).$$

Similarly, there exists a constant $C' > 0$ independent of j such that

$$v_j(y_{k,j}) \geq C' v_j(y_{m,j}).$$

For any point $y \in \mathbb{R}^n - K_0$, let y_k be one of the nearest points to y in K_0. Let Ω be the convex hull of $B_{1/2}(y) \cup B_{1/2}(y_k)$. The argument in Lemma 5.1 still holds with $B_{2r}(x_j)$ and $B_{2r}(x_j) - B_{r/2}(x_j)$ replaced by Ω and any compact subset of $\Omega - \{y_{k,j}\}$ containing y, and therefore the Harnack inequality holds uniformly for v_j on each compact subset of $\mathbb{R}^n - K_0$ when j is large. Therefore, by Proposition 5.3, for a given blowup point $y_{k,j} \to y_k \in K_0$, $v_j(y_{k,j})v_j$ is uniformly bounded in any fixed compact subset of $\mathbb{R}^n - K_0$. Multiplying $v_j(y_{k,j})$ on both sides of (7-6) and (7-7), we have that, up to a subsequence,

$$\lim_{j\to\infty} v_j(y_{k,j})v_j = F \geq 0 \quad \text{in } C^4_{\mathrm{loc}}(\mathbb{R}^n - K_0),$$

such that

$$\Delta^2 F = 0 \quad \text{in } \mathbb{R}^n - K_0, \tag{7-9}$$

$$\Delta F \leq 0 \quad \text{in } \mathbb{R}^n - K_0. \tag{7-10}$$

Pick a point $y_m \in K_0 - \{y_k\}$. Since all the blowup points in K_0 are isolated simple blowup points, by Proposition 5.3,

$$F(y) = a_1|y - y_k|^{4-n} + \Phi_1(y) = a_1|y - y_k|^{4-n} + a_2|y - y_m|^{4-n} + \Phi_2(y)$$

for $y \in \mathbb{R}^n - K_0$ with the constants $a_1, a_2 > 0$. Moreover,

$$\Phi_2 \in C^4(\mathbb{R}^n - (K_0 - \{y_k, y_m\}))$$

and Φ_2 satisfies (7-9) in $\mathbb{R}^n - (K_0 - \{y_k, y_m\})$. Define $\xi = \Delta\Phi_1$ in $\mathbb{R}^n - (K_0 - \{y_k\})$. By (7-10), $F > 0$ in $\mathbb{R}^n - K_0$. Therefore,

$$\liminf_{|y|\to\infty} \Phi_1(y) = \liminf_{|y|\to\infty} (F(y) - a_1|y - y_k|^{4-n}) \geq 0, \tag{7-11}$$

$$\limsup_{|y|\to\infty} \xi(y) = \limsup_{|y|\to\infty} \Delta(F(y) - a_1|y - y_k|^{4-n}) \leq 0, \tag{7-12}$$

where for (7-12) we have used (7-10). Moreover, $\xi < 0$ near any isolated singular point in $K_0 - \{y_k\}$ by Proposition 5.3. Applying the strong maximum principle to ξ and the equation

$$\Delta\xi = \Delta^2(F - a_1|y - y_k|^{4-n}) = 0$$

in $\mathbb{R}^n - (K_0 - \{y_k\})$,

$$\xi = \Delta\Phi_1 < 0$$

in $\mathbb{R}^n - (K_0 - \{y_k\})$. Since $\Phi_1 > 0$ near any isolated singular point in $K_0 - \{y_k\}$ by Proposition 5.3, and also (7-11) holds, applying the strong maximum principle to Φ_1 and $\Delta\Phi_1 < 0$ in $\mathbb{R}^n - (K_0 - \{y_k\})$, we have $\Phi_1 > 0$ in $\mathbb{R}^n - (K_0 - \{y_k\})$. It follows that

$$F(y) = a_1|y - y_k|^{4-n} + \Phi_1(0) + O(|y - y_k|) \text{ with } \Phi_1(y_k) > 0 \text{ near } y = y_k,$$

contradicting Proposition 5.9 (It is easy to check that Proposition 5.9 applies for the scaled metrics h_j instead of g.). Here in the statement of Proposition 5.9, $H = \hat{H} = F$. Indeed, for $5 \leq n \leq 7$, after rescaling, for each j the conformal metric $g_j = \rho_j^{4/(n-4)} g$ corresponding to the conformal normal coordinates centered at x_j becomes

$$\hat{g}_j(y) = \bar{\gamma}_j^{-2} \rho_j(\bar{\gamma}_j y)^{4/(n-4)} g(\bar{\gamma}_j y)$$

and the functions $\hat{\rho}_j(y) = \rho_j(\bar{\gamma}_j y) \to \rho(y) \equiv 1$ locally uniformly in C^N as $j \to +\infty$. Proposition 7.3 is then established. □

We are now ready to prove the compactness theorem of positive solutions to (1-2).

Proof of Theorem 1.2. By Lemma 2.3 and the ellipticity of (1-2), we only need to show that there is a constant $C > 0$ depending on M and g such that

$$u \leq C.$$

Suppose the contrary, then there exists a sequence of positive solutions u_j to (1-2) such that

$$\max_{p \in M} u_j \to \infty$$

as $j \to \infty$. By Proposition 7.3, after passing to a subsequence, there exist N distinct isolated simple blowup points $p_{1,j} \to p_1, \dots, p_{N,j} \to p_N$ with $N \geq 1$ independent of j. Applying Proposition 5.3, we have that up to a subsequence,

$$u_j(p_{1,j})u_j(p) \to F(p) = \sum_{k=1}^{N} a_k G_g(p_k, p) + b(p) \quad \text{in } C^4_{\text{loc}}(M - \{p_1, \dots, p_N\}),$$

where $a_1 > 0$, $a_2 \geq 0, \dots, a_N \geq 0$ are some constants, G_g is the Green's function of P_g under the metric g and $b(p) \in C^4(M)$ satisfying

$$P_g b = 0$$

on M. Since $Q_g \geq 0$ on M with $Q_g > 0$ at some point, by the strong maximum principle of P_g, $b \geq 0$ in M. We know that $G_g(p_k, p) > 0$ for $1 \leq k \leq N$ by Theorem 2.1. Let $x = (x^1, \dots, x^n)$ be the conformal normal coordinates centered at $p_{1,j}$ for each j (respectively, p_1) constructed in [Lee and Parker 1987] with respect to the conformal metric $h_j = \rho_j^{-4/(n-4)} g$ (respectively, $h = \rho^{-4/(n-4)} g$) such that

$$\det(h_{ij}) = 1 + O(|x|^{10n}).$$

Then there exists $C_1 > 0$ independent of j such that

$$C_1^{-1} \leq \rho_j \leq C_1,$$

and

$$\|\rho_j - \rho\|_{C^N(M)} \to 0 \quad \text{as } j \to \infty.$$

As shown in Theorem 2.1, under the conformal normal coordinates $x = (x^1, \dots, x^n)$ centered at p_1, the Green's function under metric h satisfies

$$G_h(p_1, p) = \rho^2(p) G_g(p_1, p) = d_h(p_1, p)^{4-n} + A + o(1)$$

near p_1 with the constant $A > 0$ and $o(1) \to 0$ as $p \to p_1$. Therefore,

$$\rho(p)^2 F(p) = a_1 d_h(p_1, p)^{4-n} + B + o(1)$$

$B = a_1 A + \sum_{k=2}^{N} a_k \rho(p_1)^2 G_g(p_k, p_1) + b(p_1) > 0$ and $o(1) \to 0$ as $p \to p_1$. That contradicts Proposition 5.9 with $\hat{H} = F$ in (5-49). Therefore, Theorem 1.2 is established. □

Appendix A: Positive solutions of certain linear fourth order elliptic equations in punctured balls

Assume $B_\delta(\bar{x})$ is a geodesic δ-ball on a complete Riemannian manifold (M^n, g) with 2δ less than the injectivity radius. For application, for $5 \leq n \leq 9$, (M, g) could either be the closed manifold in Proposition 5.3, or the Euclidean space.

Lemma A.1. *Let $u \in C^4(B_\delta(\bar{x}) - \{\bar{x}\})$ be a solution to*

$$P_g u = 0 \quad \text{in } B_\delta(\bar{x}) - \{\bar{x}\}. \tag{A-1}$$

If $u(p) = o(d_g(p, \bar{x})^{4-n})$ as $p \to \bar{x}$, then $u \in C^{4,\alpha}_{\text{loc}}(B_\delta(\bar{x}))$ for $0 < \alpha < 1$.

Proof. **Step 1.** We show that (A-1) holds in $B_\delta(\bar{x})$ in the distribution sense.

To see this, given any small $\epsilon > 0$, we define the cutoff function η_ϵ on $B_\delta(\bar{x})$ with $0 \le \eta_\epsilon \le 1$ so that

$$\begin{aligned} \eta_\epsilon(p) &= 1 && \text{for } d_g(p, \bar{x}) \le \epsilon, \\ \eta_\epsilon(p) &= 0 && \text{for } d_g(p, \bar{x}) \ge 2\epsilon, \\ |\nabla \eta_\epsilon(p)| &\le C\epsilon^{-1} && \text{for } \epsilon \le d_g(p, \bar{x}) \le 2\epsilon. \end{aligned}$$

For any given $\phi \in C_c^\infty(B_\delta(\bar{x}))$ we multiply by $\phi(1-\eta_\epsilon)$ on both sides of (A-1) and do integration by parts,

$$\int_{B_\delta(\bar{x})} P_g(\phi(1-\eta_\epsilon))u \, dV_g = 0.$$

Let $\epsilon \to 0$, then

$$\int_{B_\delta(\bar{x})} (1-\eta_\epsilon)u \, P_g \phi \, dV_g = O(1)\biggl(C\epsilon^{-4} \int_{B_{2\epsilon}(\bar{x}) - B_\epsilon(\bar{x})} |u|\biggr) + C \int_{B_\epsilon(\bar{x})} |u| \to 0,$$

where in the last step we have used $u(p) = o(d_g(p, \bar{x})^{4-n})$. Therefore, Step 1 is established.

Step 2. The assumption of u near $\bar{x}$ implies that $u \in L^p_{\text{loc}}(B_\delta(\bar{x}))$ for any $1 < p < \frac{n}{n-4}$. By $W^{4,p}$ estimates of the elliptic equation we obtain that $u \in W^{4,p}_{\text{loc}}(B_\delta(\bar{x}))$; see [Agmon 1959] for instance. The standard bootstrap argument gives $u \in C^{4,\alpha}_{\text{loc}}(B_\delta(\bar{x}))$. □

For later use, we now present Lemma 9.2 from [Li and Zhu 1999] without proof.

Lemma A.2. *There exists some constant $0 < \delta_0 \le \delta$ depending on n, $\|g_{ij}\|_{C^2(B_\delta(\bar{x}))}$ and $\|R_g\|_{L^\infty(B_\delta(\bar{x}))}$ such that the maximum principle for $-\frac{4(n-1)}{n-2}\Delta_g + R_g$ holds on $B_{\delta_0}(\bar{x})$, and there exists a unique $G_1(p) \in C^2(B_{\delta_0}(\bar{x}) - \{\bar{x}\})$ satisfying*

$$\begin{aligned} -\frac{4(n-1)}{n-2}\Delta_g G_1 + R_g G_1 &= 0 \quad \text{in } B_{\delta_0}(\bar{x}) - \{\bar{x}\}, \\ G_1 &= 0 \quad \text{on } \partial B_{\delta_0}(\bar{x}), \\ \lim_{p \to \bar{x}} d_g(p, \bar{x})^{n-2} G_1(p) &= 1. \end{aligned}$$

Furthermore, $G_1(p) = d_g(p, \bar{x})^{2-n} + \mathcal{R}(p)$ where, for all $0 < \epsilon < 1$, $\mathcal{R}(p)$ satisfies

$$d_g(p, \bar{x})^{n-4+\epsilon}|\mathcal{R}(p)| + d_g(p, \bar{x})^{n-3+\epsilon}|\nabla \mathcal{R}(p)| \le C(\epsilon), \quad p \in B_{\delta_0}(\bar{x}), \; n \ge 4,$$

where $C(\epsilon)$ depends on ϵ, n, $\|g_{ij}\|_{C^2(B_\delta(\bar{x}))}$ and $\|R_g\|_{L^\infty(B_\delta(\bar{x}))}$.

Lemma A.3. *Suppose a positive function $u \in C^4(B_\delta(\bar{x}) - \{\bar{x}\})$ satisfies (A-1) in $B_\delta(\bar{x}) - \{\bar{x}\}$, and assume that there exists a constant $C > 0$ such that for $0 < r < \delta$, the Harnack inequality holds*:

$$\max_{d_g(p,\bar{x})=r} u(p) \le C \min_{d_g(p,\bar{x})=r} u(p).$$

If moreover,

$$-\frac{4(n-1)}{n-2}\Delta_g u^{\frac{n-2}{n-4}} + R_g u^{\frac{n-2}{n-4}} \ge 0 \quad in\ B_\delta(\bar{x}) - \{\bar{x}\},$$

then

$$a = \limsup_{p\to\bar{x}} d_g(p,\bar{x})^{n-4} u(p) < +\infty.$$

Proof. If the lemma is not true, then for any $A > 0$, there exists $r_i \to 0^+$ satisfying

$$u(p) > \ A\, r_i^{4-n} \quad \text{for all } d_g(p,\bar{x}) = r_i.$$

Let $v_A = \frac{1}{2}A^{(n-2)/(n-4)}G_1$ with G_1 in Lemma A.2. For i large, by the maximum principle,

$$u(p)^{\frac{n-2}{n-4}} \ge v_A(p) \quad \text{for } r_i < d_g(p,\bar{x}) < \delta_0.$$

As $i \to \infty$,

$$u(p)^{\frac{n-2}{n-4}} \ge v_A(p) \quad \text{for } 0 < d_g(p,\bar{x}) < \delta_0.$$

Since A can be arbitrarily large, $u(p) = \infty$ in $0 < d_g(p,\bar{x}) < \delta_0$, which is a contradiction. □

Proposition A.4. *Let u be as in Lemma A.3. Then there exists a constant $b \ge 0$ such that*

$$u(p) = bG(p,\bar{x}) + E(p) \quad for\ p \in B_{\delta_0}(\bar{x}) - \{\bar{x}\}, \tag{A-2}$$

where G is the Green's function of P_g (for the existence of the Green's function in our application, G is the limit of the Green's function of the Paneitz operator of a sequence of metrics on M restricted to certain domains, and when g is the flat metric, let $G(x,y) = c_n|x-y|^{4-n}$), and δ_0 is defined in Lemma A.2. Here $E \in C^4(B_{\delta_0}(\bar{x}))$ satisfies $P_gE = 0$ in $B_{\delta_0}(\bar{x})$.

Proof. We rewrite (A-1) as

$$\Delta_g(\Delta_g u) = \operatorname{div}_g(a_nR_g g - b_n\mathrm{Ric}_g)\nabla_g u - \frac{n-4}{2}Q_g u.$$

By Lemma A.3, $0 < u(p) \le a_1G(p,\bar{x})$ with some constant $a_1 > a$ in $B_{\delta_0}(\bar{x}) - \{\bar{x}\}$ with $\delta_0 > 0$ in Lemma A.2. Combining this with the interior estimates, there exists a constant $C > 0$ such that

$$\left|\operatorname{div}_g(a_n R_g g - b_n \operatorname{Ric}_g)\nabla_g u - \frac{n-4}{2} Q_g u\right| \le C d_g^{2-n}(p, \bar{x}), \tag{A-3}$$

$$|\Delta_g u(p)| \le C\, d_g^{2-n}(p, \bar{x}), \tag{A-4}$$

for $p \in \bar{B}_{\delta_0}(\bar{x}) - \{0\}$. We define G_2 to be a Green's function of Δ_g on $\bar{B}_{\delta_0}(\bar{x})$ such that

$$0 < G_2(p, q) \le C d_g(p, q)^{2-n}, \tag{A-5}$$

for some constant $C > 0$ and any two distinct points p and q in $B_{\delta_0}(\bar{x})$. Then

$$\phi_1(p) = \int_{B_{\delta_0}(\bar{x})} G_2(p, q)\left(\operatorname{div}_g(a_n R_g g - b_n \operatorname{Ric}_g)\nabla_g u(q) - \frac{n-4}{2} Q_g u(q)\right) dV_g(q)$$

is a special solution to the equation

$$\Delta_g \phi = \operatorname{div}_g(a_n R_g g - b_n \operatorname{Ric}_g)\nabla_g u - \frac{n-4}{2} Q_g u \quad \text{in } B_{\delta_0}(\bar{x}) - \{\bar{x}\}.$$

Combining (A-3) and (A-5), we have that there exists a constant $C > 0$ such that

$$|\phi_1(p)| \le C d_g(p, \bar{x})^{4-n}$$

for $p \in B_{\delta_0}(\bar{x}) - \{\bar{x}\}$. Therefore,

$$\Delta_g(\Delta_g u - \phi_1) = 0 \quad \text{in } B_{\delta_0}(\bar{x}) - \{\bar{x}\}.$$

Since we also have (A-4), the proof of Proposition 9.1 in [Li and Zhu 1999] applies and there exists a constant $-C \le b_2 \le C$ such that

$$(\Delta_g u(p) - \phi_1(p)) = b_2 G_1(p) + \varphi_1(p) \quad \text{in } B_{\delta_0}(\bar{x}) - \{\bar{x}\},$$

with G_1 as in Lemma A.2 and φ_1 a harmonic function on $\bar{B}_{\delta_0}(\bar{x})$. Therefore,

$$\Delta_g u(p) = b_2 G_1(p) + \phi_1(p) + \varphi_1(p) \quad \text{in } B_{\delta_0}(\bar{x}) - \{\bar{x}\}.$$

By the same argument, there exists $b_3 \in \mathbb{R}$ such that

$$\begin{aligned} u(p) &= b_3 G_1(p) + \varphi_2(p) + \int_{B_{\delta_0}(\bar{x})} G_2(p, q)[b_2 G_1(q) + \phi_1(q) + \varphi_1(q)]\, dV_g(q) \\ &= b_3 G_1(p) + \varphi_2(p) + O(d_g(p, \bar{x})^{4-n}) \end{aligned}$$

in $B_{\delta_0}(\bar{x}) - \{\bar{x}\}$, with φ_2 a harmonic function on $B_{\delta_0}(\bar{x})$. But since $0 < u(p) \le a_1 G(p, \bar{x})$, we have $b_3 = 0$ and

$$u(p) = b_2 \int_{B_{\delta_0}(\bar{x})} G_2(p, q) G_1(q)\, dV_g(q) + o(d_g(p, \bar{x})^{4-n})$$

in $B_{\delta_0}(\bar{x})-\{\bar{x}\}$. Therefore, there exists a constant $b\ge 0$ such that

$$\begin{aligned} u(p) &= b d_g(p,\bar{x})^{4-n}+o(d_g(p,\bar{x})^{4-n}) \\ &= bG(p,\bar{x})+o(d_g(p,\bar{x})^{4-n}). \end{aligned}$$

Then by Lemma A.1, there exists a function $E\in C^4(B_{\delta_0}(\bar{x}))$ satisfying (A-1) and

$$u(p)=bG(p,\bar{x})+E(p)$$

for $p\in B_{\delta_0}(\bar{x})-\{\bar{x}\}$.

This completes the proof of the proposition. □

Using Proposition A.4, we immediately conclude the following corollary.

Corollary A.5. *For $n\ge 5$, assume that $u\in C^4(B_{\delta_0}(\bar{x})-\{\bar{x}\})$ is a positive solution of (A-1) with $\bar{x}$ a singular point, and also that the assumptions in Lemma A.3 hold for u. Then*

$$\begin{aligned} &\lim_{r\to 0}\int_{B_r(\bar{x})}\left(P_g u-\frac{n-4}{2}\bar{Q}u\right)dV_g \\ &\qquad = \lim_{r\to 0}\int_{\partial B_r(\bar{x})}\left(\frac{\partial}{\partial\nu}\Delta_g u-\left(a_n R_g\frac{\partial}{\partial\nu}u-b_n\operatorname{Ric}_g(\nabla_g u,\nu)\right)\right)ds_g \\ &\qquad = b\lim_{r\to 0}\int_{\partial B_r(\bar{x})}\frac{\partial}{\partial\nu}\Delta_g G(p,\bar{x})\,ds_g(p)=2(n-2)(n-4)|\mathbb{S}^{n-1}|\,b>0, \end{aligned}$$

where ν is the outer unit normal and $b>0$ is as in (A-2).

Appendix B: Classification of solutions with decay at infinity for a fourth order linear equation

Let $n\ge 5$. It is easy to check that $U_0=(1+4^{-1}|x|^2)^{-(n-4)/2}$ is a solution to the Q-curvature equation

$$\Delta^2 U_0=\frac{n-4}{2}\bar{Q}U_0^{\frac{n+4}{n-4}}$$

on $\mathbb{R}^n$ with $\bar{Q}=\frac{1}{8}n(n^2-4)$.

We now consider bounded solutions to the linearized equation

$$\Delta^2\phi(x)=\frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi(x),\quad x\in\mathbb{R}^n. \tag{B-1}$$

Chen and Lin [1998] classified bounded solutions to the linearized equation of the Yamabe equation in $\mathbb{R}^n$ with certain decay near infinity. Similarly, we want to show that if a solution ϕ to (B-1) has the decay $\phi\to 0$ uniformly as $|x|\to\infty$, then

$$\phi=c_0\left(x\cdot\nabla U_0+\frac{n-4}{2}U_0\right)+\sum_{j=1}^{n}c_j\partial_{x_j}U_0.$$

Let $\{\xi_{k,m}\}_m$ be the eigenfunctions of the Laplacian on $\mathbb{S}^{n-1}$, with respect to the eigenvalue $\lambda_k = k(n+k-2)$. Let $x = r\theta$ with $r = |x|$. Then we have the decomposition

$$\phi(r\theta) = \sum_{k=0}^{\infty} \sum_m \phi_{k,m}(r)\xi_{k,m}(\theta),$$

which converges locally uniformly, with $\phi_{k,m}(r) = \int_{S^{n-1}} \phi(r\theta)\xi_{k,m}(\theta)\,dS$. Let $u_{k,m}(r\theta) = \phi_{k,m}(r)\xi_{k,m}(\theta)$. Then $u_{k,m}$ satisfies the equation

$$\Delta^2 u_{k,m}(x) = \frac{n+4}{2}\bar{Q}U_0(x)^{\frac{8}{n-4}}u_{k,m}(x), \quad x \in \mathbb{R}^n, \tag{B-2}$$

and $\phi_{k,m}$ satisfies

$$\left(\partial_r^2 + \frac{n-1}{r}\partial_r - \frac{\lambda_k}{r^2}\right)\left(\partial_r^2 + \frac{n-1}{r}\partial_r - \frac{\lambda_k}{r^2}\right)\phi_{k,m} = \frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi_{k,m}, \quad r > 0, \tag{B-3}$$

with $\phi_{k,m}(0) = 0$ and $\phi'_{k,m}(0) = 0$. Equivalently, $\phi_{k,m}$ is a solution to the equation

$$\left(\Delta - \frac{\lambda_k}{r^2}\right)\left(\Delta - \frac{\lambda_k}{r^2}\right)\phi_{k,m} = \frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi_{k,m}. \tag{B-4}$$

Denote

$$v_{k,m}(r) = \left(\partial_r^2 + \frac{n-1}{r}\partial_r - \frac{\lambda_k}{r^2}\right)\phi_{k,m}.$$

Then

$$\left(\partial_r^2 + \frac{n-1}{r}\partial_r - \frac{\lambda_k}{r^2}\right)\phi_{k,m} = v_{k,m}, \tag{B-5}$$

$$\left(\partial_r^2 + \frac{n-1}{r}\partial_r - \frac{\lambda_k}{r^2}\right)v_{k,m} = \frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi_{k,m}, \tag{B-6}$$

where

$$\phi_{k,m}(0) = 0, \quad \phi'_{k,m}(0) = 0, \quad v_{k,m}(0) = 0 \quad \text{and} \quad v'_{k,m}(0) = 0. \tag{B-7}$$

By (B-2), we know that $u_{k,m}$ is analytic locally in $\mathbb{R}^n$. Then the solutions $\phi_{k,m}$ to (B-3) and (B-7) are generated linearly by the two solutions

$$\phi_{1,k,m}(r) = r^k + E_1 r^{k+4} + \sum_{j=2}^{\infty} E_j r^{k+2+2j},$$

$$\phi_{2,k,m}(r) = r^{k+2} + C_1 r^{k+6} + \sum_{j=2}^{\infty} C_j r^{k+4+2j},$$

with $E_1 > 0$ and $C_1 > 0$. The constants E_i and C_j can be calculated inductively using (B-3). It is easy to check that the radius of convergence of $\phi_{i,k,m}$ is positive for $i = 1, 2$ and $k \geq 1$. Therefore,

$$\phi_{k,m} = C\phi_{1,k,m}(r) + C'\phi_{2,k,m}(r),$$

with C and C' constants.

Now we employ a useful comparison theorem motivated by [Grunau et al. 2008]; see also [McKenna and Reichel 2003] and [Choi and Xu 2009].

Theorem B.1. *Let ϕ and v be a solution to (B-5) and (B-6) in $r > 0$. If it holds that for some $r_1 > 0$,*

$$\phi(r_1) \geq 0, \quad \phi'(r_1) \geq 0, \quad v(r_1) \geq 0 \quad \text{and} \quad v'(r_1) \geq 0,$$

with one of them nonzero, then

$$\text{(B-8)} \qquad \phi(r) > 0, \quad \phi'(r) > 0, \quad v(r) > 0 \quad \text{and} \quad v'(r) > 0$$

for $r > r_1$, and there exists a constant $C > 0$ such that $\phi(r) \geq C(r - r_1 - 1)^2$ for $r > r_1 + 1$. Moreover, there exists a positive constant $C' = C'(k)$ such that $\phi(r) \leq C'(r^{n+k+2} + 1)$. In particular, $\phi(r)$ is positive and exists for all $r > r_1$.

Proof. By the equations (B-5) and (B-6),

$$\partial_r(r^{n-1}\partial_r\phi) = r^{n-1}v + \frac{\lambda_k}{r^2}\phi r^{n-1},$$
$$\partial_r(r^{n-1}\partial_r v) = \frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi r^{n-1} + \frac{\lambda_k}{r^2}v r^{n-1}.$$

Using integration,

$$r^{n-1}\phi'(r) = r_1^{n-1}\phi'(r_1) + \int_{r_1}^{r} r^{n-1}v + \frac{\lambda_k}{r^2}\phi r^{n-1}\,dr,$$
$$r^{n-1}v'(r) = r_1^{n-1}v'(r_1) + \int_{r_1}^{r} \frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi r^{n-1} + \frac{\lambda_k}{r^2}v r^{n-1}\,dr.$$

Then it is easy to see that (B-8) holds for $r > r_1$. Also, for $r > r_1 + 1$,

$$\begin{aligned} r^{n-1}\phi'(r) &= (r_1+1)^{n-1}\phi'(r_1+1) + \int_{r_1+1}^{r} r^{n-1}v + \frac{\lambda_k}{r^2}\phi r^{n-1}\,dr \\ &\geq (r_1+1)^{n-1}\phi'(r_1+1) + \int_{r_1+1}^{r} r^{n-1}v(r_1+1)\,dr \\ &\geq v(r_1+1)\Big(\frac{1}{n}r^n - \frac{1}{n}(r_1+1)^n\Big), \end{aligned}$$

with $v(r_1+1) > 0$. Therefore, for $r > r_1 + 1$,

$$\phi'(r) \geq \frac{1}{n}v(r_1+1)r - \frac{1}{n}(r_1+1)v(r_1+1).$$

Therefore, ϕ grows at least quadratically.

Now let's see the upper bound of growth of ϕ. It is easy to check that

$$\left(\Delta - \frac{\lambda_k}{r^2}\right)\left(\Delta - \frac{\lambda_k}{r^2}\right) r^{n+k+2} \geq \frac{n+4}{2} \bar{Q} U_0^{\frac{8}{n-4}} r^{n+k+2}, \quad r > 0.$$

Also,

$$\frac{d}{dr} r^{n+k+2} > 0, \quad \left(\Delta - \frac{\lambda_k}{r^2}\right) r^{n+k+2} > 0, \text{ and } \frac{d}{dr}\left(\Delta - \frac{\lambda_k}{r^2}\right) r^{n+k+2} > 0 \text{ for } r > 0.$$

Therefore, for any $r_1 > 0$, there exists a constant $\delta = \delta(r_1) > 0$ such that the function $\varphi(r) = r^{n+k+2} - \delta\phi(r)$ satisfies (B-8) at $r = r_1$. Note that

$$\left(\Delta - \frac{\lambda_k}{r^2}\right)\left(\Delta - \frac{\lambda_k}{r^2}\right) \varphi(r) \geq \frac{n+4}{2} \bar{Q} U_0^{\frac{8}{n-4}} \varphi(r), \quad r > 0.$$

Denote

$$\tilde{v}(r) = \left(\Delta - \frac{\lambda_k}{r^2}\right) \varphi(r)$$

so that

$$\left(\Delta - \frac{\lambda_k}{r^2}\right) \tilde{v}(r) \geq \frac{n+4}{2} \bar{Q} U_0^{\frac{8}{n-4}} \varphi(r), \quad r > 0.$$

Using the same integration argument starting from $r = r_1$, we obtain that $\varphi(r) > 0$ for $r \geq r_1$. This completes the proof of Theorem B.1. □

Now we consider the behavior of $\phi_{1,k,m}$ and $\phi_{2,k,m}$.

Let $v_{1,k,m}$ and $v_{2,k,m}$ be defined as above with respect to $\phi_{1,k,m}$ and $\phi_{2,k,m}$:

$$v_{1,k,m}(r) = \left(\partial_r^2 + \frac{n-1}{r}\partial_r - \frac{\lambda_k}{r^2}\right)\phi_{1,k,m},$$

$$v_{2,k,m}(r) = \left(\partial_r^2 + \frac{n-1}{r}\partial_r - \frac{\lambda_k}{r^2}\right)\phi_{2,k,m}.$$

By the Taylor expansion, for $r > 0$ close to 0, $\phi_{1,k,m}(r) > 0$, $\phi'_{1,k,m}(r) > 0$, $v_{1,k,m}(r) > 0$ and $v'_{1,k,m}(r) > 0$. Then by Theorem B.1, $\phi_{1,k,m}(r)$ keeps increasing at least quadratically as r increases. Also, for any $\epsilon > 0$, there exists $C = C(\epsilon, k)$ such that $\phi_{1,k,m}(r)$ is bounded from above by Cr^{n+k+2} with some constant C for $r > \epsilon$. In particular, $\phi_{1,k,m}(r)$ is positive and exists for any $r > 0$. The same holds for $\phi_{2,k,m}$.

For any $r_1 > 0$, we know that $\phi_{i,k,m}$ satisfies (B-8) at $r = r_1$, for $i = 1, 2$ and $k \geq 1$. Then there exists $C > 0$ such that both $(\phi_{1,k,m} - C^{-1}\phi_{2,k,m})$ and $(\phi_{2,k,m} - C^{-1}\phi_{1,k,m})$ satisfy (B-8) at $r = r_1$. Then by Theorem B.1, for $r > r_1$,

$$\phi_{1,k,m}(r) - C^{-1}\phi_{2,k,m}(r) > 0 \quad \text{and} \quad \phi_{2,k,m}(r) - C^{-1}\phi_{1,k,m}(r) > 0.$$

That is to say, $\phi_{1,k,m}$ and $\phi_{2,k,m}$ are both positive on $(0, \infty)$ and they go to infinity as $r \to \infty$ in the same order. This leads to the following corollary.

Corollary B.2. *For any $k \geq 1$, there is at most one constant $C > 0$ such that $\phi_{1,k,m} - C\phi_{2,k,m}$ is bounded on $r \in (0, +\infty)$.*

Now we consider the asymptotic behavior of bounded solutions to (B-3) and (B-7) which vanish at infinity.

Lemma B.3. *Let $\phi_{k,m} = \phi_{1,k,m} - C\phi_{2,k,m}$ be a bounded solution to the initial value problem (B-3) and (B-7) such that $\phi_{k,m}(r) = o(1)$ as $r \to \infty$. Then $\phi_{k,m}(r) = O(r^{2-k-n})$ as $r \to +\infty$.*

Proof. We introduce

$$\phi^*_{k,m}(r) = r^{4-n}\phi_{k,m}(r^{-1}), \quad r > 0,$$

to be the Kelvin transformation of $\phi_{k,m}$ and

$$v^*_{k,m}(r) = \left(\Delta - \frac{\lambda_k}{r^2}\right)\phi^*_{k,m}(r), \quad r > 0.$$

Also, for $u_{k,m}(r\theta) = \phi_{k,m}(r)\xi_{k,m}(\theta)$, we denote

$$u^*_{k,m}(x) = |x|^{4-n} u_{k,m}\left(\frac{x}{|x|^2}\right), \quad x \in \mathbb{R}^n,$$

to be the Kelvin transformation of $u_{k,m}$. Then it is easy to check that $\phi^*_{k,m}$ is a solution to (B-3) and equivalently a solution to (B-4) in $(0, +\infty)$ and $u^*_{k,m}$ is a solution to (B-2) in $\mathbb{R}^n - \{0\}$. By our assumption on the decay of ϕ_k near infinity,

$$u^*_{k,m}(x) = o(|x|^{4-n})$$

as $x \to 0$. Then using the proof of Lemma A.1 in Appendix A we have that 0 is a removable singularity of $u^*_{k,m}$ and $u^*_{k,m}(x) = \phi^*_{k,m}(r)\xi_{k,m}(\theta)$ is a solution to (B-2) in $\mathbb{R}^n$. Therefore, $\phi^*_{k,m}$ and $v^*_{k,m}$ satisfy

$$\phi^*_{k,m}(0) = 0, \quad (\phi^*_{k,m})'(0) = 0, \quad v^*_{k,m}(0) = 0, \quad (v^*_{k,m})'(0) = 0.$$

Also, by the definition,

$$\phi^*_{k,m}(r) = r^{4-n}\phi_{k,m}(r^{-1}) = O(r^{4-k-n}) \quad \text{as } r \to +\infty.$$

Recall that $\phi_{1,k,m}$ and $\phi_{2,k,m}$ form a basis of the solution space to the problem (B-3) and (B-7). Since $\phi_{k,m}$ and $\phi^*_{k,m}$ are both bounded solutions to (B-3) and (B-7), by Corollary B.2 there exists a constant $a \in (-\infty, +\infty)$ such that $\phi^*_{k,m}(r) = a\phi_{k,m}(r)$ for $r > 0$. Note that $\phi^*_{k,m}(1) = \phi_{k,m}(1)$. If $\phi_{k,m}(1) \neq 0$, then $a = 1$. Otherwise, if also $\phi'_{k,m}(1) \neq 0$, then by L'Hospital's Rule, $a = -1$; else, if also $\phi'_{k,m}(1) = 0$ but $v_{k,m}(1) \neq 0$, then by L'Hospital's rule, $a = 1$; else, if also $\phi'_{k,m}(1) = 0$,

$v_{k,m}(1)=0$ but $v'_{k,m}(1)\neq 0$, then by L'Hospital's rule, $a=-1$ (In fact, by the comparison theorem Theorem B.1, since $\phi_{k,m}$ is bounded in $(0,+\infty)$, this could not happen). Since $\phi_{k,m}$ is assumed not to be identically zero, it is not possible that all the four data vanishes at $r=1$. Therefore, a is either 1 or -1. Therefore,

$$\begin{aligned}\phi_{k,m}(r)&=r^k+O(r^{k+2}) && \text{as } r\to 0,\\ \phi_{k,m}(r)&=\pm r^{4-k-n}+O(r^{2-n-k}) && \text{as } r\to +\infty.\end{aligned}$$

□

Let ϕ be a solution to (B-1) with the decay $\phi\to 0$ uniformly as $|x|\to\infty$. Let $u_{k,m}(r\theta)=\phi_{k,m}(r)\xi_{k,m}(\theta)=\int_{S^{n-1}}\phi(r)\xi_{k,m}(\theta)\,dS\,\xi_{k,m}(\theta),\ k\geq 1$. Then $\phi_{k,m}(r)=o(1)$ as $r\to\infty$. Using the energy method, in the following theorem we show that for $5\leq n\leq 8$, $\phi_{k,m}=0$ for $k\geq 2$.

Theorem B.4. *Let $\phi_{k,m}$ with $k\geq 2$ be a bounded solution to the initial value problem (B-3) and (B-7) for $5\leq n\leq 8$ such that $\phi_{k,m}(r)=o(1)$ as $r\to\infty$. Then $\phi_{k,m}=0$.*

Proof. By Lemma B.3, it is easy to check that $\phi_{k,m}\in H^2(\mathbb{R}^n)$, for $k\geq 2$.

By (B-4), for any $\epsilon>0$,

$$\int_{\mathbb{R}^n-B_\epsilon(0)}\phi_{k,m}\left(\Delta-\frac{\lambda_k}{r^2}\right)\left(\Delta-\frac{\lambda_k}{r^2}\right)\phi_{k,m}\,dx=\int_{\mathbb{R}^n-B_\epsilon(0)}\frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi_{k,m}^2\,dx.$$

Using integration by parts and letting $\epsilon\to 0$, we have that

$$\int_{\mathbb{R}^n}\left(\Delta-\frac{\lambda_k}{r^2}\right)\phi_{k,m}\left(\Delta-\frac{\lambda_k}{r^2}\right)\phi_{k,m}\,dx=\int_{\mathbb{R}^n}\frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi_{k,m}^2\,dx. \tag{B-9}$$

Note that

$$\begin{aligned}&\int_{\mathbb{R}^n}\left(\Delta-\frac{\lambda_k}{r^2}\right)\phi_{k,m}\left(\Delta-\frac{\lambda_k}{r^2}\right)\phi_{k,m}\,dx\\ &\qquad=\int_{\mathbb{R}^n}\left[(\Delta\phi_{k,m})^2-2\lambda_k r^{-2}\phi_{k,m}\Delta\phi_{k,m}+\lambda_k^2 r^{-4}\phi_{k,m}^2\right]dx,\end{aligned}$$

where by integration by parts,

$$\begin{aligned}\int_{\mathbb{R}^n}-2\lambda_k r^{-2}\phi_{k,m}\Delta\phi_{k,m}\,dx&=\int_{\mathbb{R}^n}2\lambda_k r^{-2}|\nabla\phi_{k,m}|^2\,dx+\int_{\mathbb{R}^n}2\lambda_k\phi_{k,m}\nabla\phi_{k,m}\cdot\nabla r^{-2}\,dx\\ &=\int_{\mathbb{R}^n}2\lambda_k r^{-2}|\nabla\phi_{k,m}|^2\,dx+\int_{\mathbb{R}^n}\lambda_k\nabla(\phi_{k,m}^2)\cdot\nabla r^{-2}\,dx\\ &=\int_{\mathbb{R}^n}2\lambda_k r^{-2}|\nabla\phi_{k,m}|^2\,dx-\int_{\mathbb{R}^n}\lambda_k\phi_{k,m}^2\Delta r^{-2}\,dx\\ &=\int_{\mathbb{R}^n}2\lambda_k r^{-2}|\nabla\phi_{k,m}|^2\,dx+(2n-8)\int_{\mathbb{R}^n}\lambda_k r^{-4}\phi_{k,m}^2\,dx\end{aligned}$$

for $n \geq 6$. Therefore,

$$\begin{aligned}
&\int_{\mathbb{R}^n} \left(\Delta - \frac{\lambda_k}{r^2}\right)\phi_{k,m}\left(\Delta - \frac{\lambda_k}{r^2}\right)\phi_{k,m}\,dx \\
&\quad = \int_{\mathbb{R}^n} (\Delta\phi_{k,m})^2\,dx + \int_{\mathbb{R}^n} 2\lambda_k r^{-2}|\nabla\phi_{k,m}|^2\,dx + (2n\lambda_k - 8\lambda_k + \lambda_k^2)\int_{\mathbb{R}^n} r^{-4}\phi_{k,m}^2\,dx \\
&\quad \geq (2n\lambda_k - 8\lambda_k + \lambda_k^2)\int_{\mathbb{R}^n} r^{-4}\phi_{k,m}^2\,dx.
\end{aligned}$$

Since $(1+4^{-1}r^2)^{-1} \leq r^{-1}$ for $r > 0$ and, for $k \geq 2$ and $5 \leq n \leq 8$,

$$2n\lambda_k - 8\lambda_k + \lambda_k^2 = (2n-8)k(n+k-2) + k^2(n+k-2)^2 > \frac{n+4}{2} \times \bar{Q} = \frac{n+4}{2} \times \frac{n(n^2-4)}{8},$$

we have that

$$\int_{\mathbb{R}^n} \left(\Delta - \frac{\lambda_k}{r^2}\right)\phi_{k,m}\left(\Delta - \frac{\lambda_k}{r^2}\right)\phi_{k,m}\,dx > \int_{\mathbb{R}^n} \frac{n+4}{2}\bar{Q}U_0^{\frac{8}{n-4}}\phi_{k,m}^2\,dx,$$

which contradicts (B-9) for $k \geq 2$ and $5 \leq n \leq 8$. Therefore, there exists no nontrivial bounded solution ϕ_k to (B-3) such that $\phi_k(r) = o(1)$ as $r \to +\infty$ for $k \geq 2$ and $5 \leq n \leq 8$. □

It is easy to check that

$$u_0 + \sum_m u_{1,m} = c_0\left(x \cdot \nabla U_0 + \frac{n-4}{2}U_0\right) + \sum_{j=1}^n c_j \partial_{x_j} U_0$$

with $c_0, \ldots, c_n$ some constants. As a direct corollary of Theorem B.4, we have:

Corollary B.5. *Let ϕ be a solution to (B-1) with the decay $\phi \to 0$ uniform as $|x| \to \infty$. Then for $5 \leq n \leq 8$, we have that*

$$\phi = c_0\left(x \cdot \nabla U_0 + \frac{n-4}{2}U_0\right) + \sum_{j=1}^n c_j \partial_{x_j} U_0$$

for some constants $c_0, c_1, \ldots, c_n$.

Acknowledgements

The author is partially supported by China Postdoctoral Science Foundation Grant 2014M550540. Most of the work is completed in BICMR. The author thanks Professor Fengbo Hang for discussion on an earlier version of the paper. The author would like to thank Professor Lei Zhang for helpful discussion on the paper [Li and Zhang 2005]. The author is grateful to Professor Chiun-Chuan Chen for helpful discussion for understanding [Chen and Lin 1998]. We wish to thank the referee for several useful comments which have improved the writing of the article.

References

[Agmon 1959] S. Agmon, "The L_p approach to the Dirichlet problem, I: Regularity theorems", *Ann. Scuola Norm. Sup. Pisa* (3) **13**:4 (1959), 405–448. MR Zbl

[Armitage 1973] D. H. Armitage, "A polyharmonic generalization of a theorem on harmonic functions", *J. London Math. Soc.* (2) **7**:2 (1973), 251–258. MR Zbl

[Branson 1985] T. P. Branson, "Differential operators canonically associated to a conformal structure", *Math. Scand.* **57**:2 (1985), 293–345. MR Zbl

[Brendle 2008] S. Brendle, "Blow-up phenomena for the Yamabe equation", *J. Amer. Math. Soc.* **21**:4 (2008), 951–979. MR Zbl

[Brendle and Marques 2009] S. Brendle and F. C. Marques, "Blow-up phenomena for the Yamabe equation, II", *J. Differential Geom.* **81**:2 (2009), 225–250. MR Zbl

[Chang et al. 2004] S.-Y. A. Chang, F. Hang, and P. C. Yang, "On a class of locally conformally flat manifolds", *Int. Math. Res. Not.* **2004**:4 (2004), 185–209. MR Zbl

[Chen and Lin 1998] C.-C. Chen and C.-S. Lin, "Estimate of the conformal scalar curvature equation via the method of moving planes, II", *J. Differential Geom.* **49**:1 (1998), 115–178. MR Zbl

[Choi and Xu 2009] Y. S. Choi and X. Xu, "Nonlinear biharmonic equations with negative exponents", *J. Differential Equations* **246**:1 (2009), 216–234. MR Zbl

[Djadli et al. 2000] Z. Djadli, E. Hebey, and M. Ledoux, "Paneitz-type operators and applications", *Duke Math. J.* **104**:1 (2000), 129–169. MR Zbl

[Druet 2004] O. Druet, "Compactness for Yamabe metrics in low dimensions", *Int. Math. Res. Not.* **2004**:23 (2004), 1143–1191. MR Zbl

[Eichmair 2013] M. Eichmair, "The Jang equation reduction of the spacetime positive energy theorem in dimensions less than eight", *Comm. Math. Phys.* **319**:3 (2013), 575–593. MR Zbl

[Eichmair et al. 2016] M. Eichmair, L.-H. Huang, D. A. Lee, and R. Schoen, "The spacetime positive mass theorem in dimensions less than eight", *J. Eur. Math. Soc.* **18**:1 (2016), 83–121. MR Zbl

[Esposito and Robert 2002] P. Esposito and F. Robert, "Mountain pass critical points for Paneitz–Branson operators", *Calc. Var. Partial Differential Equations* **15**:4 (2002), 493–517. MR Zbl

[Grunau et al. 2008] H.-C. Grunau, M. Ould Ahmedou, and W. Reichel, "The Paneitz equation in hyperbolic space", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **25**:5 (2008), 847–864. MR Zbl

[Gursky and Malchiodi 2015] M. J. Gursky and A. Malchiodi, "A strong maximum principle for the Paneitz operator and a non-local flow for the Q-curvature", *J. Eur. Math. Soc.* **17**:9 (2015), 2137–2173. MR Zbl

[Gursky et al. 2016] M. J. Gursky, F. Hang, and Y.-J. Lin, "Riemannian manifolds with positive Yamabe invariant and Paneitz operator", *Int. Math. Res. Not.* **2016**:5 (2016), 1348–1367. MR Zbl

[Hang and Yang 2016a] F. Hang and P. C. Yang, "Q curvature on a class of 3-manifolds", *Comm. Pure Appl. Math.* **69**:4 (2016), 734–744. MR Zbl

[Hang and Yang 2016b] F. Hang and P. C. Yang, "Q-curvature on a class of manifolds with dimension at least 5", *Comm. Pure Appl. Math.* **69**:8 (2016), 1452–1491. MR Zbl

[Hebey and Robert 2004] E. Hebey and F. Robert, "Compactness and global estimates for the geometric Paneitz equation in high dimensions", *Electron. Res. Announc. Amer. Math. Soc.* **10** (2004), 135–141. MR Zbl

[Humbert and Raulot 2009] E. Humbert and S. Raulot, "Positive mass theorem for the Paneitz–Branson operator", *Calc. Var. Partial Differential Equations* **36**:4 (2009), 525–531. MR Zbl

[Jin et al. 2017] T. Jin, Y. Li, and J. Xiong, "The Nirenberg problem and its generalizations: a unified approach", *Math. Ann.* **369**:1-2 (2017), 109–151. MR Zbl

[Khuri et al. 2009] M. A. Khuri, F. C. Marques, and R. M. Schoen, "A compactness theorem for the Yamabe problem", *J. Differential Geom.* **81**:1 (2009), 143–196. MR Zbl

[Lee and Parker 1987] J. M. Lee and T. H. Parker, "The Yamabe problem", *Bull. Amer. Math. Soc.* (*N.S.*) **17**:1 (1987), 37–91. MR Zbl

[Li and Xiong 2019] Y. Li and J. Xiong, "Compactness of conformal metrics with constant Q-curvature, I", *Adv. Math.* **345** (2019), 116–160. MR Zbl

[Li and Zhang 2005] Y. Y. Li and L. Zhang, "Compactness of solutions to the Yamabe problem, II", *Calc. Var. Partial Differential Equations* **24**:2 (2005), 185–237. MR Zbl

[Li and Zhang 2007] Y. Y. Li and L. Zhang, "Compactness of solutions to the Yamabe problem, III", *J. Funct. Anal.* **245**:2 (2007), 438–474. MR Zbl

[Li and Zhu 1999] Y. Li and M. Zhu, "Yamabe type equations on three-dimensional Riemannian manifolds", *Commun. Contemp. Math.* **1**:1 (1999), 1–50. 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

[Marques 2005] F. C. Marques, "A priori estimates for the Yamabe problem in the non-locally conformally flat case", *J. Differential Geom.* **71**:2 (2005), 315–346. MR Zbl

[McKenna and Reichel 2003] P. J. McKenna and W. Reichel, "Radial solutions of singular nonlinear biharmonic equations and applications to conformal geometry", *Electron. J. Differential Equations* **2003** (2003), art. id. 37. MR Zbl

[Paneitz 1983] S. M. Paneitz, "A quartic conformally covariant differential operator for arbitrary pseudo-Riemannian manifolds", preprint, 1983. Reprinted as art. id. 036 in *SIGMA* **4**, 2008.

[Pollack 1993] D. Pollack, "Nonuniqueness and high energy solutions for a conformally invariant scalar equation", *Comm. Anal. Geom.* **1**:3-4 (1993), 347–414. MR Zbl

[Qing and Raske 2006a] J. Qing and D. Raske, "Compactness for conformal metrics with constant Q curvature on locally conformally flat manifolds", *Calc. Var. Partial Differential Equations* **26**:3 (2006), 343–356. MR Zbl

[Qing and Raske 2006b] J. Qing and D. Raske, "On positive solutions to semilinear conformally invariant equations on locally conformally flat manifolds", *Int. Math. Res. Not.* **2006** (2006), art. id. 94172. MR Zbl

[Schoen 1989] R. M. Schoen, "Variational theory for the total scalar curvature functional for Riemannian metrics and related topics", pp. 120–154 in *Topics in calculus of variations*, vol. Montecatini Terme, Italy, 1987, edited by M. Giaquinta, Lecture Notes in Math. **1365**, Springer, 1989. MR Zbl

[Schoen 1991] R. M. Schoen, "On the number of constant scalar curvature metrics in a conformal class", pp. 311–320 in *Differential geometry*, edited by B. Lawson and K. Tenenblat, Pitman Monogr. Surveys Pure Appl. Math. **52**, Longman Sci. Tech., Harlow, 1991. MR Zbl

[Schoen and Yau 1979] R. Schoen and S. T. Yau, "On the proof of the positive mass conjecture in general relativity", *Comm. Math. Phys.* **65**:1 (1979), 45–76. MR Zbl

[Schoen and Zhang 1996] R. Schoen and D. Zhang, "Prescribed scalar curvature on the n-sphere", *Calc. Var. Partial Differential Equations* **4**:1 (1996), 1–25. MR Zbl

[Wei and Zhao 2013] J. Wei and C. Zhao, "Non-compactness of the prescribed Q-curvature problem in large dimensions", *Calc. Var. Partial Differential Equations* **46**:1-2 (2013), 123–164. MR Zbl

[Witten 1981] E. Witten, "A new proof of the positive energy theorem", *Comm. Math. Phys.* **80**:3 (1981), 381–402. MR Zbl

Received December 21, 2017. Revised January 4, 2019.

Gang Li
Beijing International Center for Mathematical Research
Peking University
Beijing
China
Current address:
Department of Mathematics
Shandong University
Ji'an
China
runxing3@sdu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.181

A CHARACTERIZATION OF FUCHSIAN ACTIONS BY TOPOLOGICAL RIGIDITY

KATHRYN MANN AND MAXIME WOLFF

We give a simple proof that any rigid representation of $\pi_1(\Sigma_g)$ in $\mathrm{Homeo}^+(S_1)$ with Euler number at least g is necessarily semiconjugate to a discrete, faithful representation into $\mathrm{PSL}(2, \mathbb{R})$. Combined with earlier work of Matsumoto, this precisely characterizes Fuchsian actions by a topological rigidity property. We have proved this result in greater generality, but with a much more involved proof, in arxiv:1710.04902.

1. Introduction

Let Σ_g be a surface of genus $g \geq 2$, and let $\Gamma_g = \pi_1(\Sigma_g)$. The *representation space* $\mathrm{Hom}(\Gamma_g, \mathrm{Homeo}^+(S^1))$ is the set of all actions of Γ_g on S^1 by orientation-preserving homeomorphisms, equipped with the compact-open topology. This is also the space of *flat topological circle bundles* over Σ_g, or equivalently, the space of circle bundles with a foliation transverse to the fibers. The *Euler class* of a representation $\rho \in \mathrm{Hom}(\Gamma_g, \mathrm{Homeo}^+(S^1))$ is defined to be the Euler class of the associated bundle, and the *Euler number* $\mathrm{eu}(\rho)$ is the integer obtained by pairing the Euler class with the fundamental class of the surface. The classical Milnor–Wood inequality [Milnor 1958; Wood 1971] is the statement that the absolute value of the Euler number of a flat bundle is bounded by the absolute value of the Euler characteristic of the surface.

While the Euler number determines the topological type of a flat S^1 bundle, it does not determine its flat structure — except in the special case where the Euler number is maximal, i.e., equal to $\pm(2g-2)$. In this case, a celebrated result of Matsumoto states that for any representation ρ with $\mathrm{eu}(\rho) = \pm(2g-2)$, there is a continuous, degree one, monotone map $h : S^1 \to S^1$ such that

$$h \circ \rho = \rho_F \circ h, \tag{1}$$

where ρ_F is *Fuchsian*, meaning a faithful representation of Γ_g onto a cocompact lattice in $\mathrm{PSL}(2, \mathbb{R})$. (We view $\mathrm{PSL}(2, \mathbb{R}) \subset \mathrm{Homeo}^+(S^1)$ via the action on $\mathbb{R}\mathrm{P}^1 \cong S^1$ by Möbius transformations.)

MSC2010: 20H10, 37E10, 37E45, 57S25, 58D29.
Keywords: rigidity, geometricity, Euler class, surface group actions on the circle.

An important consequence of Matsumoto's theorem is that representations with maximal Euler number are dynamically stable or rigid in the following sense.

Definition 1.1. Let Γ be a discrete group. A representation $\rho : \Gamma \to \mathrm{Homeo}^+(S^1)$ is called *path-rigid* if its path-component in $\mathrm{Hom}(\Gamma, \mathrm{Homeo}^+(S^1))$ consists of a single semiconjugacy class.

Semiconjugacy is the equivalence relation generated by the property shared by ρ and ρ_F in (1) above; we recall the precise definition in Section 2. As semiconjugacy classes are connected in $\mathrm{Hom}(\Gamma_g, \mathrm{Homeo}^+(S^1))$, path-rigid representations are precisely those whose path-component is as small as possible.

The purpose of this article is to prove the following converse to Matsumoto's result.

Theorem 1.2. *Let* $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$ *be a path-rigid representation, with* $|\mathrm{eu}(\rho)| \geq g$. *Then* $\mathrm{eu}(\rho)$ *is maximal, i.e.,* $|\mathrm{eu}(\rho)| = 2g - 2$, *and* ρ *is semiconjugate to a discrete, faithful representation into* $\mathrm{PSL}(2, \mathbb{R})$.

As shown in [Mann 2015], any 2-fold lift of a Fuchsian representation is path-rigid and has Euler class $g - 1$; hence the inequality $|\mathrm{eu}(\rho)| \geq g$ is optimal for this statement.

A stronger, but equally natural notion of rigidity comes from considering the *character space*, $X(\Gamma_g, \mathrm{Homeo}^+(S^1))$, defined as the largest Hausdorff quotient of the quotient $\mathrm{Hom}(\Gamma_g, \mathrm{Homeo}^+(S^1))/\mathrm{Homeo}^+(S^1)$. We say a representation is *rigid* if its image in $X(\Gamma_g, \mathrm{Homeo}^+(S^1))$ is an isolated point. In [Mann and Wolff 2017], we prove that all rigid representations are semiconjugate to the k-fold lift of a Fuchsian representation, for some divisor k of $2g - 2$; and that the weaker hypothesis of path-rigidity is sufficient provided the Euler class is nonzero. This is a more general statement than Theorem 1.2 here, but the proof in [Mann and Wolff 2017] is long and involved. This article gives a much easier, self-contained proof of this partial result. The assumption $|\mathrm{eu}(\rho)| \geq g$ greatly simplifies the situation, as it implies in particular that many elements of the group have north-south dynamics. In fact, our assumption here can be replaced with an a priori strictly weaker assumption on the dynamics of ρ, phrased in terms of rotation numbers of elements, as follows.

Theorem 1.3. *Suppose* $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$ *is path-rigid. If there exist based simple closed curves* $a, b \in \Gamma_g$ *with intersection number* 1 *and such that*

$$\widetilde{\mathrm{rot}}[\rho(a), \rho(b)] = \pm 1,$$

then $\mathrm{eu}(\rho) = \pm(2g - 2)$, *and* ρ *is semiconjugate to a Fuchsian representation.*

Commutators of elements of $\mathrm{Homeo}^+(S^1)$ have a well defined translation number, as we will recall in Section 2A. The hypothesis $\widetilde{\mathrm{rot}}[\rho(a), \rho(b)] = \pm 1$ is equivalent to the statement that the restriction of the representation to the torus defined by a

and b is semiconjugate to a standard Fuchsian one (see [Matsumoto 2016]). Thus, one can think of the statement above as a local-to-global result: the local condition that a torus is Fuchsian, together with path-rigidity, implies the global statement that the representation is Fuchsian.

Outline. In Section 2 we recall standard material on dynamics of groups acting on the circle, including rotation numbers and the Euler number for actions of surface groups. We then introduce important tools for the proof of Theorem 1.3, and give a quick proof that Theorem 1.3 implies Theorem 1.2.

Sections 3 through 5 are devoted to the proof of Theorem 1.3. Given a representation ρ satisfying the hypotheses of Theorem 1.3, we proceed as follows:

1. After modifying ρ by a semiconjugacy, we show there exists $a \in \Gamma_g$ represented by a nonseparating simple closed curve such that $\rho(a)$ is *hyperbolic*, meaning that it is conjugate to a hyperbolic element of $\mathrm{PSL}(2, \mathbb{R})$.

2. Using step 1, we show that (again after semiconjugacy of ρ), *any* $\gamma \in \Gamma_g$ represented by a nonseparating simple closed curve has the property that $\rho(\gamma)$ is hyperbolic. These two first steps are done in Section 3.

3. Next, in Section 4, we start to "reconstruct the surface", showing that the arrangement of attracting and repelling points of hyperbolic elements $\rho(\gamma)$, as γ ranges over simple closed curves, mimics that of a Fuchsian representation.

4. Finally, in Section 5 we show that the restriction of ρ to small subsurfaces is semiconjugate to a Fuchsian representation; this is then improved to a global result by additivity of the *relative Euler class*.

Throughout this paper, whenever we say "deformation", we mean deformation along a continuous path in $\mathrm{Hom}(\Gamma_g, \mathrm{Homeo}^+(S^1))$.

2. Preliminaries

This section gives a quick review of basic concepts used later in the text. The only material that is not standard is the *based intersection number* discussed in Section 2D.

2A. ***Rotation numbers and the Euler number.*** Most of the material in Sections 2A and 2B is covered in more detail in [Ghys 2001] and [Mann 2018].

Let $\mathrm{Homeo}^+_{\mathbb{Z}}(\mathbb{R})$ denote the group of homeomorphisms of $\mathbb{R}$ that commute with integer translations; this is a central extension of $\mathrm{Homeo}^+(S^1)$ by $\mathbb{Z}$. The primary dynamical invariant of such homeomorphisms is the translation or rotation number, whose use can be traced back to work of Poincaré [1885, Chapitre XV]. If $\tilde{g} \in \mathrm{Homeo}^+_{\mathbb{Z}}(\mathbb{R})$ and $x \in \mathbb{R}$, the *translation number* of $\tilde{g}$ is defined by $\widetilde{\mathrm{rot}}(\tilde{g}) := \lim_{n\to\infty}(\tilde{g}^n(x))/n$; this limit exists and does not depend on x. If $g \in \mathrm{Homeo}^+(S^1)$, its *rotation number* is defined by $\mathrm{rot}(g) := \widetilde{\mathrm{rot}}(\tilde{g}) \mod \mathbb{Z}$, where $\tilde{g}$ is any lift of g.

The translation number is invariant under conjugacy (and under semiconjugacy), and restricts to a morphism on every abelian subgroup of $\mathrm{Homeo}_{\mathbb{Z}}^{+}(\mathbb{R})$. On the whole group it is a *quasimorphism*, as it satisfies the following inequality.

Lemma 2.1 (see [Calegari and Walker 2011, Theorem 3.9]). *Let* $f, g \in \mathrm{Homeo}_{\mathbb{Z}}^{+}(\mathbb{R})$. *Then* $|\widetilde{\mathrm{rot}}(fg) - \widetilde{\mathrm{rot}}(f) - \widetilde{\mathrm{rot}}(g)| \le 1$, *and* $-1 \le \widetilde{\mathrm{rot}}([f, g]) \le 1$.

The second inequality is a direct consequence of the first. This in turn was implicit already in [Wood 1971]. An optimal inequality, which depends on the values of $\widetilde{\mathrm{rot}}(f)$ and $\widetilde{\mathrm{rot}}(g)$, is obtained in [Calegari and Walker 2011].

One way of defining the Euler number of a representation is in terms of translation numbers. This was perhaps first observed by Milnor and Wood [1958; 1971], who showed the following. For the purposes of this work, the reader may take this as the definition of the Euler number.

Proposition 2.2. *Consider a standard presentation*

$$\Gamma_g = \left\langle a_1, b_1, \ldots, a_g, b_g \mid \prod_i [a_i, b_i] \right\rangle.$$

Let $\rho \in \mathrm{Hom}(\Gamma_g, \mathrm{Homeo}^{+}(S^1))$, *and let* $\widetilde{\rho(a_i)}$ *and* $\widetilde{\rho(b_i)}$ *be any lifts of* $\rho(a_i)$ *and* $\rho(b_i)$ *to* $\mathrm{Homeo}_{\mathbb{Z}}^{+}(\mathbb{R})$. *Then the* ***Euler number*** $\mathrm{eu}(\rho)$ *is given by*

$$\mathrm{eu}(\rho) = \widetilde{\mathrm{rot}}([\widetilde{\rho(a_1)}, \widetilde{\rho(b_1)}] \cdots [\widetilde{\rho(a_g)}, \widetilde{\rho(b_g)}]).$$

Note that, for any f and g in $\mathrm{Homeo}^{+}(S^1)$, the value of the commutator $[\tilde{f}, \tilde{g}] \in \mathrm{Homeo}_{\mathbb{Z}}^{+}(\mathbb{R})$ is independent of the choice of lifts $\tilde{f}$ and $\tilde{g}$. Abusing notation slightly, we will often denote its translation number by $\widetilde{\mathrm{rot}}([f, g])$ (as in the statement of Theorem 1.3). Thus, in the statement above, the translation by an integer, $[\widetilde{\rho(a_1)}, \widetilde{\rho(b_1)}] \cdots [\widetilde{\rho(a_g)}, \widetilde{\rho(b_g)}]$, is independent of the choices of lifts. The Euler number $\mathrm{eu}(\rho)$ is then simply the magnitude of this translation.

As remarked in the introduction, the Milnor–Wood inequality is the statement that $|\mathrm{eu}(\rho)| \le 2g - 2$; it is a consequence of Lemma 2.1.

Though unimportant in the preceding remarks, in what follows we will need to fix a convention for commutators and group multiplication.

Convention 2.3. We read words in Γ_g from right to left, so that group multiplication coincides with function composition. We set the notation for a commutator as

$$[a, b] := b^{-1} a^{-1} b a.$$

2B. *Dynamics of groups acting on* S^1.

Definition 2.4 [Ghys 1987]. Let Γ be a group. Two representations ρ_1, ρ_2 in $\mathrm{Hom}(\Gamma, \mathrm{Homeo}^{+}(S^1))$ are *semiconjugate* if there is a monotone (possibly noncontinuous or noninjective) map $\tilde{h} : \mathbb{R} \to \mathbb{R}$ such that $\tilde{h}(x + 1) = \tilde{h}(x) + 1$ for all $x \in \mathbb{R}$, and such that, for all $\gamma \in \Gamma$, there are lifts $\widetilde{\rho_1(\gamma)}$ and $\widetilde{\rho_2(\gamma)}$ such that $\tilde{h} \circ \widetilde{\rho_1(\gamma)} = \widetilde{\rho_2(\gamma)} \circ \tilde{h}$.

Ghys gave an (incorrect, as he himself later noted [2001]) version of this definition in the introduction of [Ghys 1987]; but his text becomes correct and consistent upon replacing it by Definition 2.4. He proved that semiconjugacy is an equivalence relation on $\mathrm{Hom}(\Gamma, \mathrm{Homeo}^+(S^1))$, and it follows from his [1987, Propositions 2.2 and 2.3; 2001, Proposition 5.8] that this is the relation generated by the relationship shared by ρ and ρ_F in (1) of Section 1; this latter equivalence relation was used by other authors as a definition of semiconjugacy (see, e.g., [Calegari 2006]). Historical elements, and more discussion on the theme of semiconjugacy can be found in [Bucher et al. 2016].

The next proposition states a useful dynamical trichotomy for groups acting on the circle, which in particular can be used to explain when a semiconjugacy map can be taken to be continuous. As it is classical, we do not repeat the proof; the reader may refer to [Ghys 2001, Proposition 5.6].

Proposition 2.5. *Let $G \subset \mathrm{Homeo}^+(S^1)$. Then exactly one of the following holds*:

(i) *G has a finite orbit.*

(ii) *G is **minimal**, meaning that all orbits are dense.*

(iii) *There is a unique compact G-invariant subset of S^1 contained in the closure of any orbit, on which G acts minimally. This set is homeomorphic to a Cantor set and called the **exceptional minimal set** for G.*

In case (iii), defining h to be a map that collapses each interval in the complement of the exceptional minimal set to a point gives the following (we leave the proof as an exercise; see, e.g., [Ghys 2001, Proposition 5.8; 1987, Proposition 2.2] for more detail).

Proposition 2.6. *Let $\rho : G \to \mathrm{Homeo}^+(S^1)$ be a homomorphism such that $\rho(G)$ has an exceptional minimal set. Then ρ is semiconjugate to a homomorphism ν whose image is minimal. Moreover, provided that ν is minimal, any semiconjugacy h to any representation ρ' such that $h \circ \rho' = \nu \circ h$ is necessarily continuous.*

We will make frequent use of the following two consequences of Proposition 2.6.

Corollary 2.7. *Suppose that ρ and ρ' are semiconjugate representations. If both ρ and ρ' are minimal, then they are **conjugate**.*

Corollary 2.8. *Let $\rho \in \mathrm{Hom}(\Gamma_g, \mathrm{Homeo}^+(S^1))$ be a path-rigid representation. Then ρ is semiconjugate to a minimal representation.*

Proof. Corollary 2.7 follows immediately from Proposition 2.6. We now prove Corollary 2.8. Using Propositions 2.5 and 2.6, it suffices to show that a representation with a finite orbit is not path-rigid. If ρ has a finite orbit, then we may perform the Alexander trick, replacing the points of the periodic orbit with intervals and collapsing the complementary intervals, to continuously deform ρ into a representation with image in a conjugate K of $SO(2)$. As $\mathrm{Hom}(\Gamma_g, K) = K^{2g}$, the representation ρ can be deformed arbitrarily within this space, in particular to a representation which is not semiconjugate. □

Following Corollary 2.8, in the proof of Theorem 1.3 we will occasionally make the (justified) assumption that a path-rigid representation ρ is also minimal.

2C. *Deforming actions of surface groups.* Let $\gamma \in \Gamma_g$ be a based, simple loop. Cutting Σ_g along γ decomposes Γ_g into an amalgamated product $\Gamma_g = A *_{\langle\gamma\rangle} B$ if γ is separating, and an HNN-extension $A*_{\langle\gamma\rangle}$ if not. In both cases, A and B are free groups. As there is no obstruction to deforming a representation of a free group into any topological group, deforming a representation $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$ amounts to deforming the restriction(s) of ρ on A (and B, if γ separates), subject to the single constraint that these should agree on γ.

The following explicit deformations are analogous to special cases of *bending deformations* from the theory of quasi-Fuchsian and Kleinian groups.

Definition 2.9. (bending deformations) Let $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$.

(1) *Separating curves.* Let $\gamma = c \in \Gamma_g$ represent a separating simple closed curve with $\Gamma_g = A *_{\langle c\rangle} B$. Let c_t be a one-parameter group of homeomorphisms commuting with $\rho(c)$. Define ρ_t to agree with ρ on A, and to be equal to $c_t \rho c_t^{-1}$ on B.

(2) *Nonseparating curves.* Let $\gamma = a$ be a nonseparating curve, and let b be a nonseparating curve such that a and b are standard generators of a once-holed torus embedded in Σ_g (equivalently, the first two generators of a standard generating set of Γ_g). Let $c = [a, b]$, and let $A = \langle a, b\rangle \subset \Gamma_g$; we write again $\Gamma_g = A *_{\langle c\rangle} B$. Let a_t be a one-parameter group commuting with $\rho(a)$ and define ρ_t to agree with ρ on B and on $\langle a\rangle$, and define $\rho_t(b) = a_t \rho(b)$.

In both cases, we call this deformation of ρ a *bending along* γ.

In particular, if γ_t is a one-parameter group with $\gamma_1 = \rho(\gamma)$, then the deformation given above is the precomposition of ρ with $\tau_{\gamma *}$, where τ_γ is the Dehn twist along γ. Note that we have made a specific (though arbitrary) choice realizing the Dehn twist as an automorphism of Γ_g. This will allow us to do specific computations, for which having a twist defined only up to inner automorphism would not suffice. (See the discussion on based curves in the next subsection for more along these lines.)

Not every $f \in \mathrm{Homeo}^+(S^1)$ embeds in a one-parameter group. However, every element with at least one fixed point does. Indeed, $S^1 \setminus \mathrm{Fix}(f)$ is then a union of intervals on which the action of f is conjugated to the map $\mathbb{R} \to \mathbb{R}$, $t \mapsto t+1$ or its inverse, and it is easy to build a one-parameter group out of this observation; see, e.g., [Ghys 2001, Proposition 5.10] for more detail. This is the situation in which we will typically apply bending deformations in this article.

The next corollary is used frequently in the proof of Theorem 1.3.

Corollary 2.10. *Suppose that ρ is a path-rigid, minimal representation. Let ρ_t be a bending deformation along a, using a deformation a_t, with $a_1 = \rho(a)^N$ for some $N \in \mathbb{Z}$. Then ρ_1 is conjugate to ρ.*

Proof. By the discussion above, ρ_1 agrees with precomposition of ρ with an automorphism of Γ_g, so has the same image. Corollary 2.7 now implies that these are conjugate. □

2D. *Based curves, chains, and Fuchsian tori.* If a and b are simple closed curves on Σ_g, the familiar *geometric intersection number* is the minimum value of $|a' \cap b'|$, where a' and b' are any curves freely homotopic to a and b, respectively. It is well known that if a and b are nonseparating simple closed curves with geometric intersection number 1, then there is a subsurface $T \subset \Sigma$ homeomorphic to a torus with one boundary component with fundamental group (freely) generated by a and b. (See, e.g., [Farb and Margalit 2012, Section 1.2.3])

As mentioned earlier, the fact that we are working with specific representations, rather than conjugacy classes of elements, forces us to take the basepoint and orientation of curves into account. Although our notation $\Gamma_g = \pi_1(\Sigma_g)$ does not mention a basepoint, all elements of $\pi_1(\Sigma_g)$ will henceforth always be assumed based, and we will use the following variation on the standard definition of intersection number.

Definition 2.11 (based intersection number). Let $a, b \in \Gamma_g$. We write $i(a, b) = 0$ if we can represent a and b by differentiable maps $a, b : [0, 1] \to \Sigma_g$, based at the base point, whose restrictions to $[0, 1)$ are injective, and such that the cyclic order of their tangent vectors at the base point is either $(a'(0), -a'(1), b'(0), -b'(1))$ or $(a'(0), -a'(1), -b'(1), b'(0))$, or the reverse of one of these.
If, instead, the cyclic order of tangent vectors is $(a'(0), b'(0), -a'(1), -b'(1))$ or the reverse, we write $i(a, b) = 1$ and $i(a, b) = -1$, respectively.

This is a somewhat ad hoc definition. In particular, $i(a, b)$ is left undefined for many pairs (a, b).

Definition 2.12. A *directed k-chain* in Σ_g is a k-tuple $(\gamma_1, \dots, \gamma_k)$ of elements of Γ_g that can be represented by the images of the edges under an embedding (possibly orientation-reversing, but respecting the orientation of the edges) of the fat graph shown in Figure 1.

In particular, $i(\gamma_i, \gamma_j) = \pm 1$ if $|j - i| = 1$, and 0 otherwise. Note that we do not

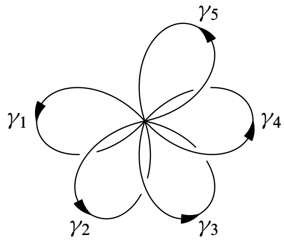


Figure 1. A directed chain of length 5.

require that the embedding be π_1-injective. For example, whenever $i(\gamma_1, \gamma_2) = 1$, then $(\gamma_1, \gamma_2, \gamma_1^{-1})$ is a (rather degenerate) directed 3-chain.

These k-chains will be useful especially to study bending deformations that realize sequences of Dehn twists. Whenever $(\gamma_1, \dots, \gamma_k)$ is a directed k-chain, the Dehn twist along the curve γ_i may be described by an automorphism of Γ_g leaving invariant the elements γ_j for $|j - i| \geq 2$ and $j = i$, and mapping γ_{i-1} to $\gamma_i^{-1}\gamma_{i-1}$, and γ_{i+1} to $\gamma_{i+1}\gamma_i$.

Notation 2.13. Let $i(a, b) = \pm 1$. Then their commutator $[a, b]$ bounds a genus 1 subsurface (well-defined up to homotopy) containing a and b. We denote this surface by $T(a, b)$.

Definition 2.14. We call any representation $\rho : \pi_1(T(a, b)) \to \mathrm{PSL}(2, \mathbb{R})$ arising from a complete hyperbolic structure of infinite volume on $T(a, b)$ a *standard Fuchsian representation of a once-punctured torus*. Similarly, we say that $\rho : \Gamma_g \to \mathrm{PSL}(2, \mathbb{R})$ is *standard Fuchsian* if it comes from a hyperbolic structure on Σ_g.

Convention 2.15. We assume Σ_g is oriented; hence standard Fuchsian representations of Γ_g have Euler number $-2g + 2$, and are all conjugate in $\mathrm{Homeo}^+(S^1)$. Similarly, $T(a, b)$ inherits an orientation, so all its standard Fuchsian representations are conjugate in $\mathrm{Homeo}^+(S^1)$.

Definition 2.16. We say that $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$ has a *Fuchsian torus* if there exist two simple closed curves $a, b \in \Gamma_g$, with $i(a, b) = \pm 1$ and such that $\widetilde{\mathrm{rot}}([\rho(a), \rho(b)]) = \pm 1$.

The terminology "Fuchsian torus" in Definition 2.16 comes from the following observation by Matsumoto.

Observation 2.17 [Matsumoto 1987]. Let $\alpha, \beta \in \mathrm{Homeo}^+(S^1)$ satisfy $\widetilde{\mathrm{rot}}([\alpha, \beta]) = \pm 1$. Then α and β generate a free group, and, up to reversing the orientation of S^1, this group is semiconjugate to a standard Fuchsian representation of a one-holed torus $T(a, b)$ with $\rho(a) = \alpha$ and $\rho(b) = \beta$.

The proof is not difficult; an easily readable sketch is given in [Matsumoto 2016, §3].

The next lemma shows the existence of such a torus is guaranteed, provided the absolute value of the Euler number of a representation is sufficiently high.

Lemma 2.18. *If* $|\mathrm{eu}(\rho)| \geq g$ *then* ρ *has a Fuchsian torus.*

Proof. If $\mathrm{eu}(\rho) \geq g$, then conjugating ρ by an orientation-reversing homeomorphism of S^1 gives a representation with Euler number at most $-g$. Thus, we may assume that $\mathrm{eu}(\rho) \leq -g$. Let $f \in \mathrm{Homeo}^+_{\mathbb{Z}}(\mathbb{R})$. It is an easy consequence of the definition of $\widetilde{\mathrm{rot}}$ that $\widetilde{\mathrm{rot}}(f) > 0$ if and only if $f(x) > x$ for all $x \in \mathbb{R}$. Hence if $f_1, \dots, f_g \in \mathrm{Homeo}^+_{\mathbb{Z}}(\mathbb{R})$ satisfy $\widetilde{\mathrm{rot}}(f_i) > 0$ for all i, then $\widetilde{\mathrm{rot}}(f_1 \cdots f_g) > 0$.

By composing such f_i by the translation by -1, which is central in $\mathrm{Homeo}^+_{\mathbb{Z}}(\mathbb{R})$, we deduce that if $\widetilde{\mathrm{rot}}(f_i) > -1$ for all i then $\widetilde{\mathrm{rot}}(f_1 \cdots f_g) > -g$. Now let ρ be a representation, and let $f_i = [\widetilde{\rho(a_i)}, \widetilde{\rho(b_i)}]$, where a_i, b_i are standard generators for Γ_g. Then the inequality $\mathrm{eu}(\rho) \leq -g$ implies $\widetilde{\mathrm{rot}}(f_i) \leq -1$ for some i. As the maximum absolute value of the rotation number of a commutator is 1 by Lemma 2.1, we in fact have $\widetilde{\mathrm{rot}}(f_i) = -1$ for some i. □

Lemma 2.18 immediately shows that Theorem 1.3 implies Theorem 1.2. The rest of this work is devoted to the proof of Theorem 1.3.

3. Steps 1 and 2: Existence and abundance of hyperbolic elements

Definition 3.1. We say a homeomorphism $f \in \mathrm{Homeo}^+(S^1)$ is *hyperbolic* if it is conjugate to a hyperbolic element of $\mathrm{PSL}(2, \mathbb{R})$, i.e., it has one *attracting fixed point* $f_+ \in S^1$ and one *repelling fixed point* $f_- \neq f_+$ such that $\lim_{n\to+\infty} f^n(x) = f_+$ for all $x \neq f_-$, and $\lim_{n\to+\infty} f^{-n}(x) = f_-$ for all $x \neq f_+$.

The first step of the proof of Theorem 1.3 is to show that a rigid, minimal representation has very many hyperbolic elements.

Lemma 3.2. *Let* $T(a, b)$ *be a one-holed torus subsurface, and let* $A = \pi_1 T(a, b)$. *Suppose* $\rho : A \to \mathrm{Homeo}^+(S^1)$ *is semiconjugate to a standard Fuchsian representation, as in Definition 2.14. Then there exists a continuous deformation* ρ_t *with* $\rho_0 = \rho$ *such that*

(i) $\rho_1(a)$ *is hyperbolic, and*

(ii) *there exists a continuous family of homeomorphisms* $f_t \in \mathrm{Homeo}^+(S^1)$ *such that* $\rho_t([a, b]) = f_t \rho([a, b]) f_t^{-1}$ *for all* t.

Proof. Let c denote the commutator $[a, b]$. Let $\bar{\rho}$ denote the minimal representation (unique up to conjugacy) that is semiconjugate to ρ. Since ρ is semiconjugate to a standard Fuchsian representation, we may suppose $\bar{\rho}$ is a representation corresponding to a *finite volume* complete hyperbolic structure on $T(a, b)$. By Proposition 2.6, there is a continuous map $h : S^1 \to S^1$, collapsing each component of the exceptional minimal set for ρ to a point, satisfying $h\rho = \bar{\rho}h$. Let x_+ and x_- be the endpoints of the axis of $\bar{\rho}(a)$, and X_+ and X_- the preimages under h of their orbits $\rho(A)x_+$ and $\rho(A)x_-$.

Note that X_+ and X_- are both $\rho(A)$-invariant sets and their images under h are the attractors (respectively, repellers) of closed curves in $T(a,b)$ conjugate to a. Moreover, for this reason, X_+ and X_- lie in a single connected component of $S^1 \smallsetminus \mathrm{Fix}(\rho(c))$, and the interiors of the intervals that make up X_+ and X_- are disjoint from the exceptional minimal set of ρ.

Define a continuous family of continuous maps $h_t : S^1 \to S^1$, with $h_0 = \mathrm{id}$, as follows: We define h_t to be the identity on the complement of the connected component of $S^1 \smallsetminus \mathrm{Fix}(\rho(c))$ containing X_+ and X_-, and for each interval I of X_+ or of X_-, have h_t be a homotopy contracting that interval so that $h_1(I)$ is a point. To make this precise, one needs to fix an identification of the target of h_t with the standard unit circle. Let J be the connected component of $S^1 \smallsetminus \mathrm{Fix}(\rho(c))$ that contains the exceptional minimal set of $\rho(A)$. Define h_t to rescale the length of each connected component of X_+ or X_- by a factor of $(1-t)$ and rescale the complement of $X_+ \cup X_-$ in J so that the total length of J remains unchanged. This gives us the desired map h_t which is the identity outside of J, and contracts intervals of X_+ and X_- to points.

Now define ρ_t by $h_t\rho(g)h_t^{-1} = \rho_t(g)$ for $t \in [0, 1)$. We claim that there is a unique $\rho_1(g)$ satisfying $h_1\rho(g) = \rho_1(g)h_1$. Indeed, $\rho(g)$ permutes the complementary intervals of the exceptional minimal set for ρ, so letting $h_1^{-1}(x)$ denote the preimage of x by h_1 (which is either a point or an open interval complementary to the exceptional minimal set), $h_1\rho(g)h_1^{-1}(x)$ is always a single point, and $h_1\rho(g)h_1^{-1}$ defines in this way a homeomorphism, which we denote by $\rho_1(g)$. It is easily verified that $\rho_t(g)$ approaches $\rho_1(g)$ as $t \to 1$. By construction, $\rho_1(a)$ is hyperbolic, and $\rho_t(c)$ is conjugate to a translation on the interval J defined above (and hence its restriction to J is conjugate to $\rho(c)|_J$), and $\rho_t(c)$ restricted to $S^1 \smallsetminus J$ agrees with $\rho(c)$. Let $f_t : S^1 \to S^1$ be a continuous family of homeomorphisms supported on J that conjugate the action of $\rho_t([a,b])$ to the action of $\rho(c)$ there. (For the benefit of the reader, justification of this step via a simple construction of such a family is given in Lemma 3.3 below.) Then $\rho_t(c) = f_t\rho(c)f_t^{-1}$, as claimed. □

Lemma 3.3. *Let g_t be a continuous family (though not necessarily a subgroup) of homeomorphisms of an open interval I, with $\mathrm{Fix}(g_t) \cap I = \varnothing$ for all $t \in [0, 1]$. There exists a continuous family of homeomorphisms f_t such that $f_t g_t f_t^{-1} = g_0$ for all t.*

Proof. Fix x in the interior of I, and let $D_t := [x, g_t(x)]$ be a fundamental domain for the action of g_t. Define the restriction of f_t to D_0 be the (unique) affine homeomorphism $D_0 \to D_t$, and extend f_t equivariantly to give a homeomorphism of I. □

Corollary 3.4. *Let $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$. Suppose that a and b are simple closed curves in Γ_g with $i(a,b) = \pm 1$ and $\widetilde{\mathrm{rot}}([\widetilde{\rho(a)}, \widetilde{\rho(b)}]) = \pm 1$. Then there exists a* *deformation ρ' of ρ such that $\rho'(a)$ is hyperbolic. If, additionally, ρ is assumed path-rigid and minimal, then $\rho(a)$ is hyperbolic.*

Proof. Let A denote the subgroup generated by a and b and let $c = [a, b]$, so $\Gamma_g = A *_{\langle c \rangle} B$. Let $\bar{\rho}$ denote the restriction of ρ to A. By Lemma 3.2, there exists a family of representations $\bar{\rho}_t : A \to \mathrm{Homeo}^+(S^1)$ such that $\bar{\rho}_t(c) = f_t \bar{\rho}(c) f_t^{-1}$ for some continuous family $f_t \in \mathrm{Homeo}^+(S^1)$, and such that $\bar{\rho}_1(a)$ is hyperbolic. As in the bending construction, define a deformation of ρ by

$$\rho_t(\gamma) = \begin{cases} \bar{\rho}_t(\gamma) & \text{for } \gamma \in A, \\ f_t \rho(\gamma) f_t^{-1} & \text{for } \gamma \in B. \end{cases}$$

By construction, ρ_t is a well-defined representation, and $\rho_1(a) = \bar{\rho}_1(a)$ is hyperbolic.

If ρ is assumed path-rigid, then this deformation ρ_1 is semiconjugate to ρ. If ρ is additionally known to be minimal, then there is a continuous map h satisfying $h \circ \rho_1 = \rho \circ h$. In particular, this implies that $\mathrm{Fix}(\rho(a)) = h\mathrm{Fix}(\rho_1(a))$, so $\rho(a)$ has at most two fixed points. In this case, if $\rho(a)$ does not have hyperbolic dynamics then it has a lift to $\mathrm{Homeo}^+_{\mathbb{Z}}(\mathbb{R})$ satisfying $|x - \widetilde{\rho(a)}(x)| \leq 1$ for all x. However, this easily implies that $|\widetilde{\mathrm{rot}}([\widetilde{\rho(a)}, \widetilde{\rho(b)}])| < 1$. (The reader may verify this as an exercise, or see the proof of Theorem 2.2 in [Matsumoto 1987] where this computation is carried out.) We conclude that $\rho(a)$ must be hyperbolic when ρ is path-rigid and minimal. □

Having found one hyperbolic element, our next goal is to produce many others. An important tool here, and in what follows, is the following basic observation on dynamics of circle homeomorphisms.

Observation 3.5. Let $f \in \mathrm{Homeo}^+(S^1)$ be hyperbolic, with attracting point f_+ and repelling point f_-, and let $g \in \mathrm{Homeo}^+(S^1)$. For any neighborhoods U_- and U_+ of f_- and f_+, respectively, and any neighborhoods V_- and V_+ of $g^{-1}(f_-)$ and $g(f_+)$, respectively, there exists $N \in \mathbb{N}$ such that

$$f^N g(S^1 \smallsetminus V_-) \subset U_+ \text{ and } g f^N (S^1 \smallsetminus U_-) \subset V_+.$$

The proof is a direct consequence of Definition 3.1. Note that, if f is hyperbolic, then f^{-1} is as well (with attracting point f_- and repelling point f_+), so an analogous statement holds with f^{-1} in place of f and the roles of f_+ and f_- reversed.

We now state two useful consequences of this observation. The proofs are elementary and left to the reader.

Corollary 3.6. *Let $f \in \mathrm{Homeo}^+(S^1)$ be hyperbolic, and suppose g does not exchange the fixed points of f. Then for N sufficiently large, either $f^N g$ or $f^{-N} g$ has a fixed point.*

Corollary 3.7. *Let $f \in \mathrm{Homeo}^+(S^1)$ be hyperbolic, and suppose $g^{-1}(f_-) \neq f_+$. Suppose also that $f^N g$ is known to be hyperbolic for large N. Then as $N \to \infty$,* *the attracting point of $f^N g$ approaches f_+ and the repelling point approaches $g^{-1}(f_-)$.*

With these tools in hand, we can use one hyperbolic element to find others.

Proposition 3.8. *Let ρ be path-rigid and minimal, and suppose that $i(a, b) = \pm 1$ and that $\rho(a)$ is hyperbolic. Then $\rho(b)$ is hyperbolic.*

Proof. We prove this under the assumption that $\rho(b)$ does not exchange the fixed points of $\rho(a)$. This assumption is justified by Lemma 3.9 below. Assuming $\rho(b)$ does not exchange the points of $\mathrm{Fix}(\rho(a))$, by Corollary 3.6, there exists some $N \in \mathbb{Z}$ such that $\rho(a^N b)$ has a fixed point. Since $\rho(a)$ is hyperbolic, $\rho(a)^N$ belongs to a one-parameter family of homeomorphisms, and a bending deformation using this family gives a deformation ρ_1 of ρ with $\rho_1(b) = \rho(a^N b)$. By Corollary 2.10, using the fact that ρ is minimal, ρ_1 and ρ are conjugate. Thus, $\rho(b)$ has a fixed point and belongs to a one-parameter group b_t.

Now we can build a bending deformation ρ'_t such that $\rho'_1(b) = \rho(b)$ and $\rho'_1(a) = \rho(ba)$. Thus, $\rho'_1(a^{-1}b) = \rho(a^{-1})$, which is hyperbolic. Since ρ'_1 and ρ are conjugate, this means that $\rho(a^{-1}b)$ is hyperbolic. Similarly, using the fact that a belongs to a one-parameter group, there exists a bending deformation ρ''_t with $\rho''_1(a^{-1}b) = \rho(b)$, and such that ρ''_1 is conjugate to ρ. This implies that $\rho(b)$ is hyperbolic. $\square$

Lemma 3.9. *Let $a, b \in \Gamma_g$ satisfy $i(a, b) = \pm 1$, and let $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$. Suppose that $\rho(a)$ is hyperbolic, and $\rho(b)$ exchanges the fixed points of $\rho(a)$. Then there is a deformation ρ' of ρ which is not semiconjugate to ρ.*

Proof. Note first that the property that $\rho(b)$ exchanges the fixed points of $\rho(a)$ implies that $\rho(b^{-1}a^{-1}b)$ is hyperbolic with the same attracting and repelling points as a. Hence $[\rho(a), \rho(b)]$ is hyperbolic with the same attracting and repelling points as well. We now produce a deformation ρ_1 of ρ such that $\rho_1(a)$ and $\rho_1(b)$ are in $\mathrm{PSL}(2, \mathbb{R})$, after this we will easily be able to make an explicit further deformation to a representation which is not semiconjugate.

First, conjugate ρ so that $\rho(a) \in \mathrm{PSL}(2, \mathbb{R})$ and so that the attracting and repelling fixed points of $\rho(a)$ are at 0 and $1/2$ respectively (thinking of S^1 as $\mathbb{R}/\mathbb{Z}$). Now choose a continuous path b_t from $b_0 = b$ to the order two rotation $b_1 : x \mapsto x + 1/2$, and such that $b_t(0) = 1/2$ and $b_t(1/2) = 0$ for all t. By the observation above, $[\rho(a), b_t]$ is hyperbolic with attracting fixed point 0 and repelling fixed point $1/2$ for all t, and so is conjugate to $\rho(a)$. By Lemma 3.3, applied separately to $(0, 1/2)$ and $(1/2, 1)$, there exists a continuous choice of conjugacies f_t such that $f_t[\rho(a), \rho(b)]f_t^{-1} = [\rho(a), b_t]$. Now to define ρ_t, we consider $\Gamma_g = A *_c B$ where $A = \langle a, b\rangle$ and $c = [a, b]$, and set

$$\rho_t(\gamma) = f_t \rho(\gamma) f_t^{-1} \text{ for } \gamma \in B, \qquad \rho_t(a) = \rho(a), \qquad \rho_t(b) = b_t.$$

This gives a continuous family of well-defined representations, with $\rho_1(b)$ the standard order 2 rotation, and $\rho_1(a) \in \mathrm{PSL}(2, \mathbb{R})$.

To finish the proof of the lemma, it suffices to note that, for a sufficiently small deformation b'_t of $\rho_1(b)$ in SO(2), the commutator $[\rho_1(a), b'_t]$ will remain hyperbolic, as the set of hyperbolic elements is open in $\mathrm{PSL}(2, \mathbb{R})$. Thus, there is a continuous path of conjugacies in $\mathrm{Homeo}^+(S^1)$ to $[\rho_1(a), b]$. This allows us to build a deformation ρ' of ρ with $\rho'(b) = b'_t \in \mathrm{SO}(2)$, using the strategy from Corollary 3.4. Since $\mathrm{rot}(b'_t) \neq \mathrm{rot}(b) = 1/2$, it follows that ρ' and ρ are not semiconjugate. □

The following corollary summarizes the results of this section.

Corollary 3.10. *Let $\sim_i$ denote the equivalence relation on nonseparating simple closed curves in Σ_g generated by $a \sim_i b$ if $i(a, b) = \pm 1$. Suppose $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$ is path-rigid, and suppose that there are simple closed curves a, b with $i(a, b) = \pm 1$ such that $\widetilde{\mathrm{rot}}[\rho(a), \rho(b)] = \pm 1$. Then ρ is semiconjugate to a (minimal) representation with $\rho(\gamma)$ hyperbolic for all $\gamma \sim_i a$.*

Remark 3.11. In fact, the relation $\sim_i$ has only a single equivalence class! This statement of connectedness of a certain complex of *based* curves can be proved using the connectedness of the arc complex of the once-punctured surface Σ_g^1; see [Mann and Wolff 2017, Section 2.1] for details. However, we will not need to use this fact here, so to keep the proof as self-contained and short as possible we will not refer to it further.

4. Step 3: Configuration of fixed points

The objective of this section is to organize the fixed points of the hyperbolic elements in a directed 5-chain; we will achieve this gradually by considering first 2-chains, then 3-chains, and finally 5-chains.

As in Definition 3.1, for a hyperbolic element $f \in \mathrm{Homeo}^+(S^1)$ we let f_+ denote the attracting fixed point of f, and f_- the repelling point. By "$\mathrm{Fix}(f)$ *separates* $\mathrm{Fix}(g)$" we mean that g_- and g_+ lie in different connected components of $S^1 \setminus \mathrm{Fix}(f)$. In particular, $\mathrm{Fix}(f)$ and $\mathrm{Fix}(g)$ are disjoint.

Lemma 4.1. *Let ρ be path-rigid and minimal, and let a, b be simple closed curves with $i(a, b) = \pm 1$ and $\rho(a)$ hyperbolic. Then $\rho(b)$ is hyperbolic, and $\mathrm{Fix}(\rho(a))$ separates $\mathrm{Fix}(\rho(b))$ in S^1.*

Proof. That $\rho(b)$ is hyperbolic follows from Proposition 3.8 above.

We prove the separation statement. As a first step, let us show that $\mathrm{Fix}(\rho(a))$ and $\mathrm{Fix}(\rho(b))$ are disjoint. Suppose for contradiction that they are not. Then, (after reversing orientations if needed) we have $\rho(a)_+ = \rho(b)_+$. Let I be a neighborhood of $\rho(a)_+$ with closure disjoint from $\{\rho(a)_-, \rho(b)_-\}$. Then, for $N > 0$ large enough, we have $\bar{I} \subset \rho(a^{-N}b)(I)$. Let ρ_t be a bending deformation with $\rho_0 = \rho$, $\rho_t(a) = \rho(a)$ and $\rho_1(b) = \rho(a^{-N}b)$. By Corollary 2.10, $\rho_1(b)$ is hyperbolic. Since $\bar{I} \subset \rho(a^{-N}b)(I)$, its attracting fixed point is outside I, and hence $\rho_1(b)_+ \neq \rho_1(a)_+$. But ρ and ρ_1 are conjugate by Corollary 2.10; this is a contradiction.

Now that we know that $\mathrm{Fix}(\rho(a)) \cap \mathrm{Fix}(\rho(b)) = \varnothing$, we will prove that they separate each other. Suppose for contradiction that $\mathrm{Fix}(\rho(a))$ does not separate $\mathrm{Fix}(\rho(b))$. Up to conjugating ρ by an orientation-reversing homeomorphism of S^1, and up to replacing b with b^{-1}, the fixed points of $\rho(a)$ and $\rho(b)$ have cyclic order $(a_+,\ a_-,\ b_+,\ b_-)$. (For simplicity, we have suppressed the notation ρ.)

Fix $N \in \mathbb{N}$ large, and let ρ' be a bending deformation of ρ so that $\rho'(b) = \rho(a^N)\rho(b)$, and $\rho'(a) = \rho(a)$. It follows from Corollaries 2.10 and 3.7 that, if N is large enough, the points $b'_+ = \rho'(b)_+$ and $b'_- = \rho'(b)_-$ can be taken arbitrarily close, respectively, to a_+ and $\rho(b)^{-1}(a_-)$. Since the cyclic order of fixed points is preserved under deformation, they are also in order (a_+, a_-, b'_+, b'_-). This is incompatible with the positions of a_+ and $\rho(b)^{-1}(a_-)$, unless perhaps if $\rho(b)^{-1}(a_-) = a_+$. But if $\rho(b)^{-1}(a_-) = a_+$, then $\rho'(b)$ has no fixed point in $(\rho(b)^{-1}(a_+), a_+)$ as this interval is mapped into (a_+, a_-) by $\rho(b')$. This again gives an incompatibility with the cyclic order. □

Lemma 4.2. *Let ρ be path-rigid and minimal, and let (a, b, c) be a directed* 3-*chain. Suppose that $\rho(a)$ is hyperbolic, and suppose that $\rho(a)$ and $\rho(c)$ do not have a common fixed point. Then $\rho(b)$ and $\rho(c)$ are hyperbolic, and, up to reversing the orientation of S^1, their fixed points are in the cyclic order*

$$(\rho(a)_-, \rho(b)_-, \rho(a)_+, \rho(c)_-, \rho(b)_+, \rho(c)_+).$$

Proof. It follows from Proposition 3.8 that $\rho(b)$ and $\rho(c)$ are hyperbolic, and from Lemma 4.1 that up to reversing orientation, the fixed points of $\rho(a)$ and $\rho(b)$ come in the cyclic order

$$(a_-, b_-, a_+, b_+).$$

(For simplicity we drop ρ from the notation for the fixed points.) As mentioned above, the effect of a bending deformation that realizes a power of a Dehn twist along a is to leave a and c invariant and to replace b with ba^N. Corollary 2.10 says that the resulting representation is conjugate to ρ. By doing this with $N > 0$ and $N < 0$ large, we get representations for which $b'_- = \rho(ba^N)_-$ can be taken arbitrarily close to a_+, as well as to a_-. This, and Lemma 4.1 applied to the curves (b, c), imply that the intervals (a_+, b_+) and (b_+, a_-) each contain one fixed point of c. To prove the lemma, it now suffices to prove the cyclic order of fixed points

$$(a_-, b_-, a_+, c_+, b_+, c_-)$$

cannot occur. Suppose for contradiction that this configuration holds, and apply a power of Dehn twist along b, replacing a with $b^{-N}a$ and c with cb^N (and leaving

b invariant), for $N > 0$ large. Denote by c'_+, c'_-, a'_- and a'_+ the resulting fixed points, i.e., the fixed points of $\rho(cb^N)$ and $\rho(b^{-N}a)$ for $N > 0$ large. If N is chosen large enough, then c'_+, c'_- and a'_- are arbitrarily close to $c(b_+)$, b_- and $a^{-1}(b_+)$, respectively. (See Corollary 3.7.) These three points are in the reverse cyclic order as c_+, c_- and a_-; hence, the representation ρ' obtained from this Dehn twist cannot be conjugate to ρ. This contradicts Corollary 2.10, and so eliminates the undesirable configuration. □

We are now ready to prove the main result of this section.

Proposition 4.3. *Let ρ be a path-rigid, minimal representation, and (a, b, c, d, e) be a directed 5-chain in Σ_g. Suppose $\rho(a)$ is hyperbolic. Then, $\rho(b), \dots, \rho(e)$ are hyperbolic as well, and up to reversing the orientation of the circle, their fixed points are in the following (total) cyclic order*:

$$(a_-, b_-, a_+, c_-, b_+, d_-, c_+, e_-, d_+, e_+).$$

In particular, these fixed points are all distinct. As before, for simplicity we have dropped ρ from the notation.

Proof. That $\rho(b), \dots, \rho(e)$ are all hyperbolic follows from Proposition 3.8. Next, using a bending deformation realizing a Dehn twist along d, we may change the action of c into $d^{-N}c$ without changing a, and without changing the conjugacy class of ρ. In particular, such a deformation moves the fixed points of c, so we can ensure that $\mathrm{Fix}(\rho(a))$ and $\mathrm{Fix}(\rho(c))$ are disjoint.

Similarly, for any two elements in the chain (a, b, c, d, e), there is a third one that intersects one but not the other. Thus, we may apply the same reasoning to show that all these five hyperbolic elements have pairwise disjoint fixed sets. It remains to order these fixed sets. For this, we will apply Lemma 4.2 repeatedly.

First, fix the orientation of S^1 so that, applying Lemma 4.2 to the directed 3-chain (a, b, c), we have the cyclic order of fixed points

$$(a_-, b_-, a_+, c_-, b_+, c_+).$$

Now, Lemma 4.2 applied to the directed 3-chain (b, c, d) implies that d_- lies in the interval (b_+, c_+) and d_+ in the interval (c_+, b_-). Applying the lemma to the directed 3-chain (a, cb, d) implies that d_+ in fact lies in the interval (c_+, a_-).

The same argument using Lemma 4.2 applied to the directed 3-chains (c, d, e) and (a, dcb, e) shows that e_- lies in the interval (c_+, d_+) and e_+ in the interval (d_+, a_-), as desired. □

5. Step 4: Maximality of the Euler number

In order to compute the Euler number of ρ, we will decompose Σ_g into subsurfaces and compute the contribution to $\mathrm{eu}(\rho)$ from each part. The proper framework for

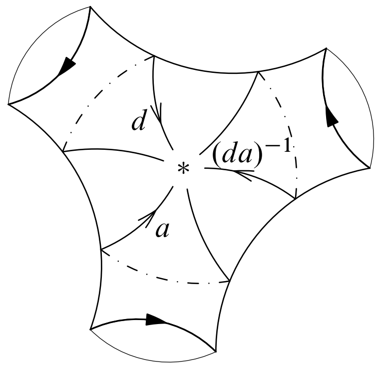


Figure 2. A pair of pants with standard generators of its fundamental group.

discussing this is the language of bounded cohomology: if Σ is a surface with boundary $\partial\Sigma$, and $\rho : \pi_1(\Sigma) \to \mathrm{Homeo}^+(S^1)$, one obtains a characteristic number by pulling back the *bounded Euler class* in $H_b^2(\mathrm{Homeo}^+(S^1); \mathbb{R})$ to $H_b^2(\Sigma, \partial\Sigma; \mathbb{R})$ and pairing it with the fundamental class $[\Sigma, \partial\Sigma]$. The contribution to the Euler number of $\rho : \Sigma_g \to \mathrm{Homeo}^+(S^1)$ from a subsurface Σ is simply this Euler number for the restriction of ρ to Σ.

However, in order to keep this work self-contained and elementary, we will avoid introducing the language of bounded cohomology, and give definitions in terms of rotation numbers alone. The reader may refer to [Burger et al. 2014, §4.3] for details on the cohomological framework.

Definition 5.1 (Euler number for pants). Let $\rho : \Gamma_g \to \mathrm{Homeo}^+(S^1)$, and let $P \subset \Sigma_g$ be a subsurface homeomorphic to a pair of pants, bounded by curves a, d and $(da)^{-1}$, with orientation induced from the boundary. Let $\widetilde{\rho(a)}$ and $\widetilde{\rho(d)}$ be any lifts of $\rho(a)$ and $\rho(d)$ to $\mathrm{Homeo}^+_{\mathbb{Z}}(\mathbb{R})$. The *Euler number of ρ on P* is the real number

$$\mathrm{eu}_P(\rho) = \widetilde{\mathrm{rot}}(\widetilde{\rho(a)}) + \widetilde{\mathrm{rot}}(\widetilde{\rho(d)}) - \widetilde{\mathrm{rot}}(\widetilde{\rho(d)}\widetilde{\rho(a)}).$$

An illustration in the case where P contains the basepoint is given in Figure 2.

Note that the number $\mathrm{eu}_P(\rho)$ is independent of the choice of lifts of $\rho(a)$ and $\rho(d)$. We also allow for the possibility that the image of P in Σ_g has two boundary curves identified, and so is a one-holed torus subsurface. In this case, one may choose free generators a, b for the fundamental group, with $i(a, b) = -1$ so the torus is $T(a, b)$ and the boundary of P is given by the curves b^{-1}, $a^{-1}ba$ and the commutator $[a, b]$. Then the definition above gives

$$\mathrm{eu}_P(\rho) = \widetilde{\mathrm{rot}}[\widetilde{\rho(a)}, \widetilde{\rho(b)}].$$

Now, the following is a restatement of Lemma 2.1 above.

Lemma 5.2. *Let P be any pants and ρ a representation. Then $|\mathrm{eu}_P(\rho)| \leq 1$.*

More generally, if $S \subset \Sigma_g$ is any subsurface, we define the Euler number $\mathrm{eu}_S(\rho)$ to be the sum of relative Euler numbers over all pants in a pants decomposition of S. From the perspective of bounded cohomology, it is immediate that this sum does not depend on the pants decomposition; however, since we are intentionally avoiding cohomological language, we give a short stand-alone proof.

Lemma 5.3. *For any subsurface $S \subseteq \Sigma_g$, the number $\mathrm{eu}_S(\rho)$ is well defined, i.e., independent of the decomposition of S into pants.*

Proof. Any two pants decompositions can be joined by a sequence of elementary moves; namely those of types (I) and (IV) as shown in [Hatcher and Thurston 1980]. A type (IV) move takes place within a pants-decomposed one-holed torus P and so does not change the value of eu_P, which is simply the rotation number of the boundary curve, as remarked above. A type (I) move occurs within a four-holed sphere S'; if the boundary of the sphere is given by oriented curves a, b, c, d with $dcba = 1$, then it consists of replacing the decomposition along da with a decomposition along ab. It is easy to verify by the definition that, in either case, the sum of the Euler numbers of the two pants on S' is given by

$$\widetilde{\mathrm{rot}}(\widetilde{\rho(a)}) + \widetilde{\mathrm{rot}}(\widetilde{\rho(b)}) + \widetilde{\mathrm{rot}}(\widetilde{\rho(c)}) + \widetilde{\mathrm{rot}}(\widetilde{\rho(d)}). \qquad \square$$

Corollary 5.4 (additivity of Euler number). *Let $\mathcal{P}$ be any decomposition of Σ into pants. Then*

$$\mathrm{eu}(\rho) = \sum_{P \in \mathcal{P}} \mathrm{eu}_P(\rho).$$

Proof. By Lemma 5.3, we may use any pants decomposition to compute the Euler class. By using a standard generating system $(a_1, \dots, b_g)$ and cutting Σ_g along geodesics freely homotopic to a_i, $c_i = [a_i, b_i]$, for $i = 1, \dots, g$ and $d_i = c_i \cdots c_1$ for $i = 2, \dots, g-1$, we recover the formula taken as a definition in Proposition 2.2. $\square$

We now return to our main goal: we prove that maximality of the Euler class holds first on small subsurfaces, then globally on Σ_g.

Proposition 5.5. *Let $S \subset \Sigma_g$ be a subsurface homeomorphic to a four-holed sphere. Suppose that none of the boundary components of S is separating in Σ_g, and let ρ be a path-rigid, minimal representation mapping one boundary component of S to a hyperbolic element of* $\mathrm{Homeo}^+(S^1)$. *Then, ρ maps all four boundary components of S to hyperbolic elements, and the relative Euler class* $\mathrm{eu}_S(\rho)$ *is equal to ± 2.*

In the statement above, we do not require that the boundary components are geodesics for some metric on Σ_g, in particular, two of them may well be freely homotopic.

Proof. Put the base point inside of S. The complement $\Sigma_g \setminus S$ may have one or two connected components, since none of the curves of ∂S are separating in Σ_g. In

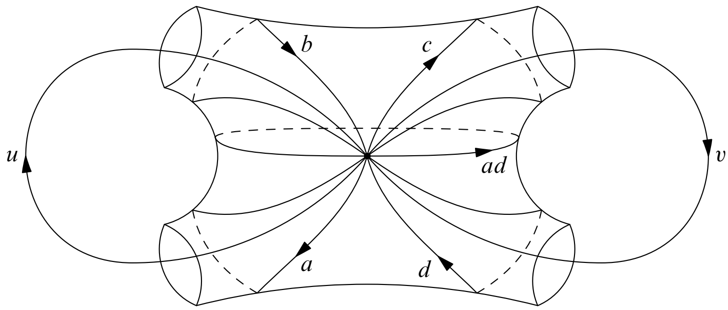


Figure 3. A four-holed sphere and two 5-chains.

either case, we may find two based, nonseparating, simple closed curves $u, v \in \Gamma_g$, with $i(u, v) = 0$, each having nonzero intersection number with exactly two of the boundary components of S, as shown in Figure 3. Additionally, we may fix orientations for u and v and choose four elements $a, b, c, d \in \pi_1 S$, each freely homotopic to a different boundary component of S, with $dcba = 1$, and such that $(a, u, d^{-1}a^{-1}, v, d)$ and (c, v, ad, u, b) are directed 5-chains in Σ_g. As we have assumed that the image under ρ of one of a, b, c or d is hyperbolic, Proposition 3.8 implies that all the curves appearing in these 5-chains are in fact hyperbolic.

Orient the circle so that $(u_-, (ad)_+, u_+, (ad)_-)$ are in cyclic order (as before, we drop the letter ρ from the notation, for better readability). Then, Proposition 4.3 applied to the two directed 5 chains above gives the cyclic orderings

$$(a_-, u_-, a_+, (ad)_+, u_+, v_-, (ad)_-, d_-, v_+, d_+)$$

and

$$(c_-, v_-, c_+, (ad)_-, v_+, u_-, (ad)_+, b_-, u_+, b_+).$$

These two orderings together yield the cyclic ordering

$$((ad)_-, d_-, d_+, a_-, a_+, (ad)_+, b_-, b_+, c_-, c_+).$$

We now use this ordering to prove maximality of the Euler class. Let α, β, γ and δ, respectively, denote the lifts of $\rho(a)$, $\rho(b)$, $\rho(c)$ and $\rho(d)$ to $\mathrm{Homeo}_{\mathbb{Z}}^+(\mathbb{R})$ with translation number zero. Let $x = (ad)_-$ be the repelling fixed point of ad.

Since x has a repelling fixed point of d immediately to the right, and an attracting fixed point of d to the left, we have $\delta(x) < x$. By the same reasoning, if y is any point in the interval between consecutive lifts of fixed points a_+ and a_- containing x, then $\alpha(y) < y$. Since $ad(x) = x$, it follows that $\delta(x)$ must lie to the left of the lift of a_+, and we have $\alpha\delta(x) = x - 1$.

Since $cbad = 1$, we also have that $cb(x) = x$. Considering the location of repelling points of b and c and imitating the argument above, we have again $\beta(x) < x$, and also $\gamma\beta(x) < x$. It follows that $\gamma\beta(x) = x - 1$, hence $\gamma\beta\alpha\delta(x) = x - 2$, and $\mathrm{eu}_S(\rho) = -2$. □

With this information on subsurfaces, we prove the Euler number of ρ is maximal.

Proposition 5.6. *Let ρ be path-rigid, and suppose that ρ admits a Fuchsian torus. Then ρ has Euler number $\pm(2g-2)$.*

Proof. After semiconjugacy, we may assume that ρ is minimal. Let $T(a, b)$ be a Fuchsian torus for ρ. By Corollary 3.4, we may suppose that $\rho(a)$ is hyperbolic. Ignoring the curve b, find a system of simple closed curves $a_1 = a, a_2, \dots, a_{g-1}$, with each a_i nonseparating, that decomposes Σ_g into a disjoint union of pairs of pants.

The dual graph of such a pants decomposition is connected (because Σ_g is connected), so we may choose a finite path that visits all the vertices. In other words, we may choose a sequence $P_1, \dots, P_N$ of pants from the decomposition (possibly with repetitions), that contains each of the pants of the decomposition, such that each two consecutive pants P_i and P_{i+1} are distinct, but share a boundary component. Let S_i denote the four-holed sphere obtained by taking the union of P_i and P_{i+1} along a shared boundary curve. (If P_i and P_{i+1} share more than one boundary component, choose only one). We may further assume that a is one of the boundary curves of S_1.

Starting with S_1 as the base case, and applying Proposition 5.5, we inductively conclude that all boundary components of all the S_i are hyperbolic, and that $\mathrm{eu}_{S_i}(\rho) = \pm 2$. Thus, the contributions of P_i and P_{i+1} are equal, and equal to ± 1, for all i. It follows that the contributions of all pairs of pants of the decomposition have equal contributions, equal to ± 1. By definition of the Euler class, we conclude that $\mathrm{eu}(\rho) = \pm(2g-2)$. □

The proof of Theorem 1.3 now concludes by citing Matsumoto's result [1987] that such a representation of maximal Euler number is semiconjugate to a Fuchsian representation. □

Acknowledgements

This work was started at MSRI during spring 2015 at a program supported by NSF grant 0932078. Both authors also acknowledge support of the U.S. National Science Foundation grants DMS 1107452, 1107263, 1107367 "RNMS: Geometric Structures and Representation Varieties" (the GEAR network). Mann was partially supported by NSF grant DMS-1606254, and thanks the Inst. Math. Jussieu and Foundation Sciences Mathématiques de Paris. This work was finished while Wolff was visiting the Universidad de la República, Montevideo, Uruguay.

References

[Bucher et al. 2016] M. Bucher, R. Frigerio, and T. Hartnick, "A note on semi-conjugacy for circle actions", *Enseign. Math.* **62**:3-4 (2016), 317–360. MR Zbl

[Burger et al. 2014] M. Burger, A. Iozzi, and A. Wienhard, "Higher Teichmüller spaces: from SL(2, ℝ) to other Lie groups", pp. 539–618 in *Handbook of Teichmüller theory, IV*, edited by A. Papadopoulos, IRMA Lect. Math. Theor. Phys. **19**, Eur. Math. Soc., Zürich, 2014. MR Zbl

[Calegari 2006] D. Calegari, "Dynamical forcing of circular groups", *Trans. Amer. Math. Soc.* **358**:8 (2006), 3473–3491. MR Zbl

[Calegari and Walker 2011] D. Calegari and A. Walker, "Ziggurats and rotation numbers", *J. Mod. Dyn.* **5**:4 (2011), 711–746. MR Zbl

[Farb and Margalit 2012] B. Farb and D. Margalit, *A primer on mapping class groups*, Princeton Math. Series **49**, Princeton Univ. Press, 2012. MR Zbl

[Ghys 1987] É. Ghys, "Groupes d'homéomorphismes du cercle et cohomologie bornée", pp. 81–106 in *The Lefschetz Centennial Conference, III* (Mexico City, 1984), edited by A. Verjovsky, Contemp. Math. **58**, Amer. Math. Soc., Providence, RI, 1987. MR Zbl

[Ghys 2001] É. Ghys, "Groups acting on the circle", *Enseign. Math.* (2) **47**:3-4 (2001), 329–407. MR Zbl

[Hatcher and Thurston 1980] A. Hatcher and W. Thurston, "A presentation for the mapping class group of a closed orientable surface", *Topology* **19**:3 (1980), 221–237. MR Zbl

[Mann 2015] K. Mann, "Spaces of surface group representations", *Invent. Math.* **201**:2 (2015), 669–710. MR Zbl

[Mann 2018] K. Mann, "Rigidity and flexibility of group actions on the circle", pp. 705–752 in *Handbook of group actions, IV*, edited by L. Ji et al., Adv. Lectures in Math. **41**, Int. Press, Somerville, MA, 2018. arXiv

[Mann and Wolff 2017] K. Mann and M. Wolff, "Rigidity and geometricity for surface group actions on the circle", preprint, 2017. arXiv

[Matsumoto 1987] S. Matsumoto, "Some remarks on foliated S^1 bundles", *Invent. Math.* **90**:2 (1987), 343–358. MR Zbl

[Matsumoto 2016] S. Matsumoto, "Basic partitions and combinations of group actions on the circle: a new approach to a theorem of Kathryn Mann", *Enseign. Math.* **62**:1-2 (2016), 15–47. MR Zbl

[Milnor 1958] J. Milnor, "On the existence of a connection with curvature zero", *Comment. Math. Helv.* **32** (1958), 215–223. MR Zbl

[Poincaré 1885] H. Poincaré, "Sur les courbes définies par les équations différentielles (III)", *Journal de mathématiques pures et appliquées* **4**:1 (1885), 167–244.

[Wood 1971] J. W. Wood, "Bundles with totally disconnected structure group", *Comment. Math. Helv.* **46** (1971), 257–273. MR Zbl

Received March 9, 2018. Revised January 16, 2019.

KATHRYN MANN
DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, NY
UNITED STATES
k.mann@cornell.edu

MAXIME WOLFF
SORBONNE UNIVERSITÉ
UNIVERSITÉ PARIS DIDEROT
CNRS, INSTITUT DE MATHÉMATIQUES DE JUSSIEU-PARIS RIVE GAUCHE, IMJ-PRG
PARIS
FRANCE
maxime.wolff@imj-prg.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.201

FUNDAMENTAL DOMAINS AND PRESENTATIONS FOR THE DELIGNE–MOSTOW LATTICES WITH 2-FOLD SYMMETRY

IRENE PASQUINELLI

In this work we will build a fundamental domain for Deligne–Mostow lattices in PU(2, 1) with 2-fold symmetry, which completes the list of Deligne–Mostow lattices in dimension 2. These lattices were introduced by Mostow, (1980; 1986) and Deligne and Mostow (1986) using monodromy of hypergeometric functions and have been reinterpreted by Thurston (1998) as automorphisms on a sphere with cone singularities. Following his approach, Parker (2006), Boadi and Parker (2015) and Pasquinelli (2016) built a fundamental domain for the class of lattices with 3-fold symmetry, i.e., when three of five cone singularities have the same cone angle. Here we extend this construction to the asymmetric case, where only two of the five cone points on the sphere have the same cone angle, hence building a fundamental domain for all remaining Deligne–Mostow lattices in PU(2, 1).

1. Introduction

Deligne–Mostow lattices first appeared in [Deligne and Mostow 1986; Mostow 1986; 1988]. They arise as monodromy of hypergeometric functions, a construction that dates back to Picard, Lauricella and others. More precisely, they start with a *ball N-tuple* $\mu = (\mu_1, \dots, \mu_N)$, i.e., a set of N real numbers between 0 and 1 such that $\sum \mu_i = 2$, from which they construct some lattices in $\mathrm{PU}(N-3, 1)$. Then Mostow deduced a sufficient condition on μ for the monodromy group to be

MSC2010: 22E40, 32M05, 51M10, 57M50.

Keywords: complex hyperbolic lattices, discrete subgroups of Lie groups, ball quotients, Deligne–Mostow lattices.

discrete, called condition ΣINT. This improved the sufficient condition called INT and introduced by Picard.

Thurston [1998] reinterpreted these lattices in terms of cone metrics on the sphere. First he considers a sphere with N cone singularities of cone angles θ_i between 0 and 2π satisfying the discrete Gauss–Bonnet formula (i.e., $\sum \alpha_i = 4\pi$, where $\alpha_i = 2\pi - \theta_i$ are the curvatures at the cone points). Then he proves that the moduli space of such cone metrics with prescribed cone angles and area 1 has a complex hyperbolic structure of dimension $N-3$. Considering the group of automorphisms of the sphere swapping cone points (and their squares), he gets some conditions on the cone angles for this group to be a lattice. Thurston's criterion corresponds to the ΣINT condition when taking $\mu_i = 2\pi\alpha_i$. Kojima [2001] proved that the two constructions are equivalent.

Combining the works of Deligne, Mostow and Thurston already mentioned with the work of Sauter [1990], one gets a finite and exhaustive list of ball N-tuples μ that give rise to a lattice using this construction (see also [Deligne and Mostow 1993]). This includes nine ball 5-tuples and one ball 6-tuple not satisfying the condition ΣINT, but commensurable to a monodromy group satisfying ΣINT. Any other value gives a nondiscrete quotient. In this work we will concentrate on the ball 5-tuples in Deligne and Mostow and Thurston's works and we will study the lattices in PU(2, 1) obtained. All of these lattices have an extra symmetry given by some of the μ_i's having the same value. In Thurston's approach, this means that some of the cone points on the sphere have the same cone angle. In particular, the lattices will have either 2-fold or 3-fold symmetry (i.e., they will have 2 or 3 cone points with same cone angle respectively). The latter case has been analysed in many different papers. We divide the 3-fold symmetry lattices into four classes (first, second, third and fourth type) according to the sign of certain parameters associated to them (see [Parker 2009] for details on the use of this terminology). First, Deraux, Falbel and Paupert [Deraux et al. 2005] built a fundamental domain for the lattices of second type. Using a different method, Parker [2006] built a fundamental domain for a class of these lattices, called Livné lattices (or lattices of third type). The lattices of first type were treated by Boadi and Parker [2015], using the same procedure as in [Parker 2006]. Later, in [Pasquinelli 2016], I explained how to use Parker's method to describe a single polyhedron that, appropriately modified, gives a fundamental domain for all lattices with 3-fold symmetry, including the cases already treated and the lattices of fourth type.

The goal of this paper is to show how to adapt Parker's construction in order to build a fundamental domain for the remaining Deligne–Mostow lattices, namely those with 2-fold symmetry. In the first part of the paper, we will forget about the symmetries and we will give a completely general construction that is valid whatever the initial cone points are. The construction consists in parametrising the cone metric and showing geometrically that Thurston's theorem holds, giving explicitly the Hermitian form that determines the complex hyperbolic structure. Then we will introduce the *moves*, which are maps on the sphere corresponding to swapping two cone points, i.e., applying a half Dehn twist along a curve containing two cone points or corresponding to a full Dehn twist. These are automorphisms of the sphere when the cone points which are swapped have the same cone angle. Here we will also consider maps that swap cone points with different cone angle. This means that we land on a new cone metric after applying the move. Moreover, we will show how one can build a polyhedron by studying what happens when pairs of cone points approach until they coalesce. We want to remark that this is completely general and can be built even if the cone angles we started from do not give a lattice. Then, in the 3-fold symmetry case the polyhedron is actually a fundamental domain for the lattices, when starting from the right set of cone singularities. In the 2-fold symmetry case this polyhedron is a building block for the fundamental domain, which will consist of the union of three copies of this polyhedron, each for a different ordering of the cone points. How to take these three copies is the topic of the second part of the paper, together with the proof that the new polyhedron built is the fundamental domain we wanted.

The paper is organised as follows. Section 1 is the present introduction.

Section 2 contains background information about the complex hyperbolic space, including the definition and the properties needed in the rest of the paper.

Section 3 considers a generic cone metric on the sphere (without any symmetry). To parametrise it, we cut along curves passing through the cone points and develop the metric on a plane in a polygonal form. One can recover the cone metrics by gluing the associated sides of the polygon back together. The parameters will be related to the sides of the polygon, which we use to give a set of projective coordinates. We will then describe the moves and the polyhedron. In the last part of the section, we will describe the two sets of coordinates that we will use to make the definition of the polyhedron clearer, we will analyse its cells and study its combinatorics.

In Section 4 we will specialise to the 2-fold symmetry case. First we will introduce the lattices we will be working with, listing the possible sets of cone angles we will be starting from. This is the original list from the works of Deligne and Mostow which can be found, for example, at the end of [Thurston 1998] or in [Mostow 1988]. Then we will build a new polyhedron as the union of three copies of the polyhedron described in Section 4B. It will be described using three sets of coordinates. At the end of the section we will describe its sides and use the moves of Section 3B to construct the side pairing maps that we need for Poincaré polyhedron theorem.

Section 5 is dedicated to our main theorem, which states that the polyhedron constructed (up to certain modifications) is indeed a fundamental domain for the lattices with 2-fold symmetry. This is proved using the Poincaré polyhedron theorem and in this section we will show that all conditions in the theorem are satisfied. Specifically we will prove that the polyhedron and its images under the side-pairing maps are disjoint and that they tessellate a neighbourhood of each ridge (2-dimensional facet) of the polyhedron. The theorem also gives us a presentation of the lattices, with the side pairings as generators. In the presentation the relations are given by cycle transformations which are complex reflections with certain parameters related to the lattice as their order. When an order is positive, we have a complex reflection with respect to a complex line, when it is negative we have a complex reflection in a point, while when it is ∞ we have a parabolic element and a fixed point on the boundary. The last two cases are related to the modifications of the polyhedron that we mentioned. In fact, when one of the parameters is negative or infinite, a ridge collapses to a single point, on the boundary when the parameter is infinite. In Section 5A we also explain in detail why this happens. In particular, when a particular one of the parameters is not positive and finite, we need to consider a different configuration (similar to the one in Section 3, but not quite the same), which is described in Section 5E. This explicit description of the polyhedron also allows us to calculate the orbifold Euler characteristic of the polyhedron, as the sum (with alternating signs with the dimension of the facets) of the order of stabiliser of one element for each orbit of facets. Then we calculate the volume of the quotient, which is a multiple of the orbifold Euler characteristic. Observe that the volume we calculated is consistent with the commensurability theorems we know for these lattices (see, for example, page 15 of [Parker 2009]) and the known volumes of the lattices.

Some of our proofs are very similar to the ones in [Parker 2006; Boadi and Parker 2015; Pasquinelli 2016]. When the exact same proof can be used, we will not rewrite it.

2. Complex hyperbolic space

Complex hyperbolic space is a natural generalisation to the complex world of real hyperbolic space. In fact, 1-dimensional complex hyperbolic space is just a 2-dimensional real hyperbolic space. The real hyperbolic plane is, in fact, an example of complex hyperbolic space of dimension 1. Generalising this construction to a complex vector space we get complex hyperbolic space. In this section we will introduce the background material about complex hyperbolic geometry that we will need in this work. More details about complex hyperbolic space can be found in [Goldman 1999].

2A. ***Definition and isometries.*** Consider $\mathbb{C}^{n,1}$, a complex vector space of dimension $n+1$ equipped with a Hermitian form H of signature $(n, 1)$.

Then one has a product law on $\mathbb{C}^{n,1}$ given by

$$\langle z, \boldsymbol{w}\rangle = \boldsymbol{w}^* H z,$$

where $\boldsymbol{w}^*$ denotes the complex conjugate of the transpose of the vector (and same for matrices). Since the norm $\langle z, z\rangle$ obtained using this product is always real, one can consider three subspaces of $\mathbb{C}^{n,1} \setminus \{0\}$, namely V_+, V_0, V_-, defined by the norm being positive, zero or negative respectively.

Consider the projection $\mathbb{P}$ of $\mathbb{C}^{n,1} \setminus \{0\}$ onto $\mathbb{CP}^n$.

Definition 2.1. The n-dimensional complex hyperbolic space for a Hermitian form H is $\boldsymbol{H}^n_{\mathbb{C}} = \mathbb{P}V_-$, i.e., the space of vectors of negative norm, up to scalar multiplication by complex numbers. Its boundary is $\partial \boldsymbol{H}^n_{\mathbb{C}} = \mathbb{P}V_0$.

On $\boldsymbol{H}^n_{\mathbb{C}}$ we consider the Bergman metric, given by the formula

$$ds^2 = \frac{-4}{\langle z, z\rangle^2} \det \begin{pmatrix} \langle z, z\rangle & \langle dz, z\rangle \\ \langle z, dz\rangle & \langle dz, dz\rangle \end{pmatrix}.$$

The -4 factor makes the curvature pinched between $-\frac{1}{4}$ and -1. For two points z and $\boldsymbol{w}$, their distance $\varrho(z, \boldsymbol{w})$ is given by

$$\cosh^2\left(\frac{\varrho(z, \boldsymbol{w})}{2}\right) = \frac{\langle z, \boldsymbol{w}\rangle\langle \boldsymbol{w}, z\rangle}{\langle z, z\rangle\langle \boldsymbol{w}, \boldsymbol{w}\rangle}. \tag{1}$$

Consider now $U(H)$ the group of $(n+1)\times(n+1)$ matrices A which are unitary with respect to H (i.e., $A^* H A = H$). The projectivisation of this group is

$$\mathrm{PU}(H) = U(H)/\{e^{i\theta} I : \theta \in [0, 2\pi)\}.$$

The full group of isometries of $\boldsymbol{H}^n_{\mathbb{C}}$ is generated by the complex conjugation and $\mathrm{PU}(H)$, which we denote as $\mathrm{PU}(n, 1)$, since it is independent on H.

In this work we will study some lattices in $\mathrm{PU}(2, 1)$, which are discrete subgroups Γ for which the quotient $\Gamma\backslash\boldsymbol{H}^n_{\mathbb{C}}$ has finite volume.

2B. *Bisectors.* Our goal is to build a fundamental domain for some lattices. In complex hyperbolic space this is very hard, since there are no totally geodesic real hypersurfaces. One possible substitute is bisectors, which have some interesting properties.

Given two points z and $\boldsymbol{w}$ in $\boldsymbol{H}^2_{\mathbb{C}}$, the bisector between (or equidistant from) z and $\boldsymbol{w}$ is the locus of points which are equidistant from the two points. The complex line L spanned by z and $\boldsymbol{w}$ is the *complex spine* of the bisector. The complex spine and the bisector intersect in the *spine* of the bisector, which is a geodesic $\gamma \in L$.

Bisectors are good to cut out a polyhedron because they are foliated by totally geodesic subspaces in two different ways. The first foliation is by slices and is a construction due to Giraud [1921] and Mostow [1980]. A *slice* is a complex line that is a fibre of the map Π_L, the orthogonal projection to the complex spine L. The bisector is the preimage by Π_L of γ and the preimage of each point of γ is a slice. The second foliation is by meridians and is due to Goldman [1999]. A *meridian* is a totally geodesic Lagrangian plane containing the spine γ. It can also be described as the set of fixed points of an antiholomorphic involution which swaps z and $\boldsymbol{w}$.

Bisectors also contain Giraud discs [1921]. For three points z, $\boldsymbol{w}$ and $\boldsymbol{x}$, not all contained in the same complex line, one can consider $B(z, \boldsymbol{w}, \boldsymbol{x})$, the set of points equidistant from these three points. Such set is a smooth disc, not totally geodesic, contained in exactly three bisectors ($B(z, \boldsymbol{w})$, $B(z, \boldsymbol{x})$ and $B(\boldsymbol{w}, \boldsymbol{x})$) and it's called a *Giraud disc*.

3. The general construction

This section follows [Parker 2006, Section 2; Boadi and Parker 2015, Section 2; Pasquinelli 2016, Section 4]. It generalises this procedure to when there is no symmetry given by cone points having same cone angle. The construction is based on Thurston's result:

Theorem 3.1 (Thurston). *Let $\alpha_1, \dots, \alpha_N$ be N real numbers in $(0, 2\pi)$ whose sum is 4π. The set of Euclidean cone metrics on the sphere with cone points of curvatures α_i and area 1 form a complex hyperbolic manifold of dimension $N-3$, whose metric completion is a complex hyperbolic cone manifold of finite volume.*

In this work we always have $N=5$. One can geometrically parametrise such cone metrics and that is the procedure that will be explained in this section. Thurston also showed that the metric completion is obtained by making pairs of cone point approach until they coalesce. Later in this section we will study this in order to build a polyhedron. Thurston also gave a sufficient condition on the curvatures for the cone manifold to be an orbifold, which corresponds to Mostow's ΣINT condition mentioned in the introduction. From Section 4 onwards we will restrict to initial cone angles whose curvatures satisfy the condition.

In the first part of this section we will show how to parametrise the cone metrics using suitable coordinates. This is a generic construction and does not depend on whether the cone angles we choose give a lattice or not, nor on whether the cone points have the same angle or not. The only restriction on the cone angles in this case is for the Hermitian form we obtain to have the required signature. We will also show how to explicitly see the complex hyperbolic structure in our coordinates. In the second part of the section we will show how to build a polyhedron in the moduli space starting from the cone metrics and using the coordinates we introduced. Finally we will describe some maps we will use, in the spirit of the moves in previous works. In the case of lattices with 3-fold symmetry treated in previous works, the polyhedron so constructed is a fundamental domain for the

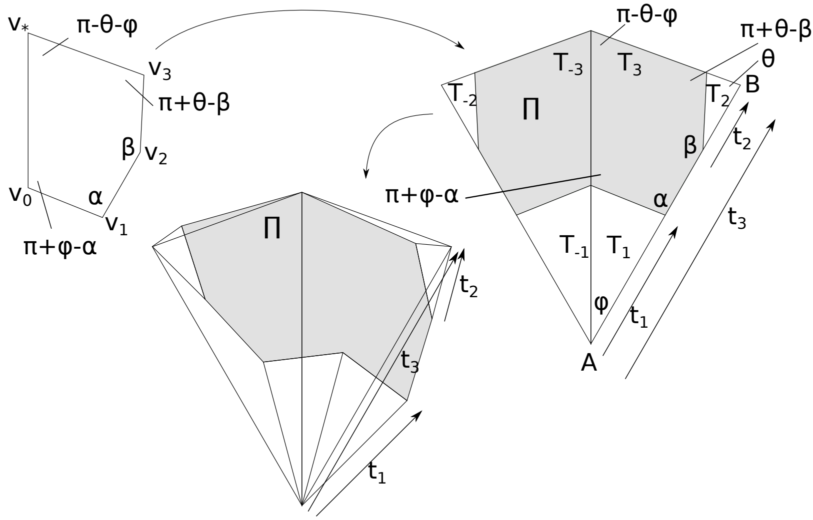


Figure 1. The configuration $(\alpha, \beta, \theta, \phi)$.

group generated by the moves, which are symmetries of the polyhedron. In the case of 2-fold symmetry treated in the Section 4, it will be a building block for the fundamental domain.

3A. ***Configuration space.*** Consider a cone metric on the sphere with cone points of angles $\theta_0, \theta_1, \theta_2, \theta_3$ and θ_4 with $0 < \theta_i < 2\pi$ and $\sum(2\pi - \theta_i) = 4\pi$ (discrete Gauss–Bonnet formula). Since we have 5 cone singularities, a priori the lattices are described by 5 parameters.

The discrete Gauss–Bonnet formula guarantees that the value of the fifth angle is determined by the previous four. To prescribe the cone angles we will use the parameters

$$\alpha = \frac{\theta_1}{2}, \quad \beta = \frac{\theta_2}{2}, \quad \theta = \frac{\theta_2}{2} + \frac{\theta_3}{2} - \pi, \quad \phi = \frac{\theta_0}{2} + \frac{\theta_1}{2} - \pi. \tag{2}$$

They have a geometric meaning which is made clear in Figure 1. Then we will denote a cone metric with these cone angles as $(\alpha, \beta, \theta, \phi)$. By definition of the parameters, we are considering a flat sphere with 5 cone singularities of angles

$$(2(\pi + \phi - \alpha)), \quad 2\alpha, \quad 2\beta, \quad 2(\pi + \theta - \beta), \quad 2(\pi - \theta - \phi)).$$

As one can see in the upper-left-hand side of Figure 1, the order of the angles is given by starting in the lower left corner and continuing anticlockwise. So the angle θ_i is the cone angle of the cone point v_i for $i = 0, 1, 2, 3$ and θ_4 is the cone angle of the cone point v_*.

We now fix the cone angles (so fix a configuration $(\alpha, \beta, \theta, \phi)$) and want to parametrise all possible positions of the cone points on the sphere. Let us first consider the easier case of when the five cone singularities are along the equator of the sphere. In other words, it is possible to draw a geodesic between the cone points which cuts the cone angles in half. Then one can cut along the geodesic passing through v_0, v_1, v_2, v_3 and v_* in order and open up in the complex plane the figure obtained, getting the upper-right-hand side of Figure 1. Remark that we are positioning v_* in the origin and v_0 along the negative imaginary axis. One can always recover the original cone metric by gluing back in pairs the sides of the octagonal shaped shaded polygon Π in Figure 1 by a vertical reflection.

Let us now consider the triangle T_3 which has a vertex in v_* and the three sides along the lines between v_*, v_3, between v_1, v_2, between v_*, v_0. Call A the vertex between the second and third lines mentioned and B the other vertex. Let T_2 be the triangle built on the segment between v_0, v_1 with the opposite vertex coinciding with A and T_1 the triangle built on the segment v_2, v_3 with opposite vertex B. Then T_{-i} is the reflection by the vertical axis of T_i, for $i = 1, 2, 3$. A way of describing Π is to consider the two triangles $T_{\pm 3}$ and removing from it copies of T_2 and T_{-2} and copies of T_1 and T_{-1}. Since we have fixed the positions of the points on the sphere to be on the equator, the possible variations in the cone metric are the possible distances between the cone points. Now since all the angles of the triangles are determined by the cone angles, it is enough to take as a parameter one side of each of the triangles in order to also have all the lengths determined and hence the distances between cone points. To describe the cone metric we will use the three aligned parameters t_1, t_2 and t_3 shown in the picture. These are the three sides of the triangles aligned along the line between A and B. Knowing the values of the parameters and the initial cone angles, one can recover the cone metric by building the triangles and gluing the octagon back into a sphere.

Let us now consider the general case, where the cone points are allowed to be anywhere on the sphere. By inspection on Figure 1 one can see that it is enough to allow the three real parameters t_1, t_2 and t_3 to be complex (i.e., not to be aligned). More precisely, we will fix the vertex of T_1 at A and the vertex of T_2 at B and allow the triangles to rotate around that vertex, moving v_i for $i = 0, 1, 2, 3$ along with it. The sphere can still be recovered by gluing the corresponding sides of Π. The length of the complex parameters encodes the information of the distances between cone points, while the angle of the parameters encodes the fact that two pieces of the geodesic might not divide the cone angle they share in two equal angles.

For more details on this description one can see [Parker 2006, Section 2.1; Boadi and Parker 2015, Section 2.1; Pasquinelli 2016, Section 4.1].

Since we are interested in the cone metrics up to rescaling, we will choose to assume that $t_3 = 1$. Remark that this is one of the possible normalisations, different from asking from the area to be 1 (like in [Thurston 1998]).

As proved by Thurston, the set of cone metrics of this type has a complex hyperbolic structure. The Hermitian form for the complex hyperbolic structure is given by the area of the cone metric. The area of the octagon Π is given by

$$\text{Area}\,\Pi = \frac{\sin\theta\sin\phi}{\sin(\theta+\phi)}|t_3|^2 - \frac{\sin\theta\sin\beta}{\sin(\beta-\theta)}|t_2|^2 - \frac{\sin\phi\sin\alpha}{\sin(\alpha-\phi)}|t_1|^2. \tag{3}$$

So one could write the Hermitian form in matrix form as

$$H = \begin{bmatrix} -\frac{\sin\phi\sin\alpha}{\sin(\alpha-\phi)} & 0 & 0 \\ 0 & -\frac{\sin\theta\sin\beta}{\sin(\beta-\theta)} & 0 \\ 0 & 0 & \frac{\sin\phi\sin\theta}{\sin(\theta+\phi)} \end{bmatrix},$$

and say that

$$\text{Area}\,\Pi = \boldsymbol{t}^* H \boldsymbol{t}.$$

Now, all this only makes sense if the area of Π (and so the area of the cone metric) is positive. Moreover, following Thurston [1998] we are considering the cone metrics of area 1, or equivalently, the coordinates up to projective equivalence. Finally, from the matrix form of H, it is clear that it has signature $(1,2)$ for suitable values of the parameters (see Section 4A for the values of the parameters that we will consider). Remembering the definition of $\boldsymbol{H}^2_{\mathbb{C}}$ presented in Section 2, we can write

$$\boldsymbol{H}^2_{\mathbb{C}} = \{z : \langle z, z\rangle = z^* H z > 0\}.$$

3B. ***Moves.*** In this section we will introduce some maps that will play a key role in the following sections, since their compositions will be the generators of the lattices with 2-fold symmetry. They generalise the maps used in [Parker 2006; Boadi and Parker 2015; Pasquinelli 2016].

The move R_1 exchanges the two cone points v_2 and v_3 with their cone angles, while R_2 exchanges v_1 and v_2. Since the moves change the values of our parameters, we will denote the move as $R_i(\alpha,\beta,\theta,\phi)$ to say that

$$R_i : (\alpha,\beta,\theta,\phi) \mapsto (\alpha',\beta',\theta',\phi'),$$

unless the angles of the configuration we start from is obvious. This means, for example, that when composing two maps $T(\alpha,\beta,\theta,\phi)$ and $S(\alpha,\beta,\theta,\phi)$, we need to consider that the second map is applied to the new angles, so we are doing the composition

$$S(\alpha',\beta',\theta',\phi') \circ T(\alpha,\beta,\theta,\phi) \tag{4}$$

because $(\alpha,\beta,\theta,\phi) \overset{T}{\mapsto} (\alpha',\beta',\theta',\phi') \overset{S}{\mapsto} (\alpha'',\beta'',\theta'',\phi'')$. Similarly, when calculating inverses we have

$$[T(\alpha,\beta,\theta,\phi)]^{-1} = T^{-1}(\alpha',\beta',\theta',\phi'), \tag{5}$$

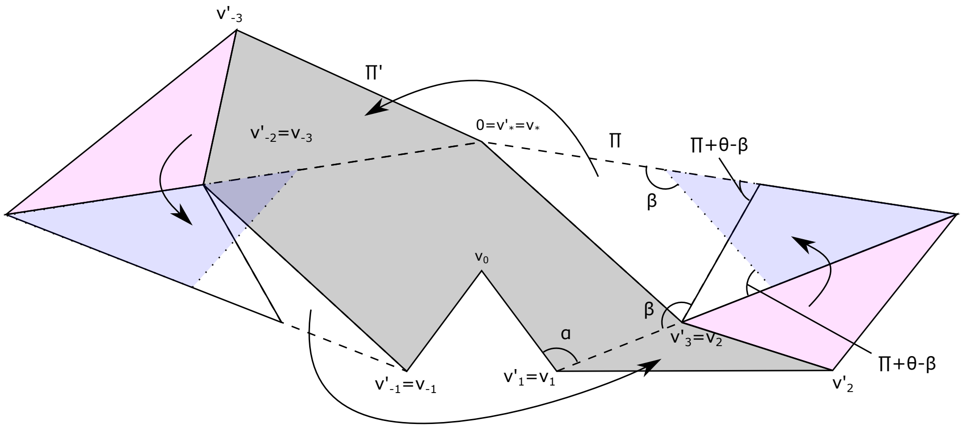


Figure 2. The move R_1.

since $T\colon (\alpha,\beta,\theta,\phi)\mapsto(\alpha',\beta',\theta',\phi')$ and $T^{-1}\colon(\alpha',\beta',\theta',\phi')\mapsto(\alpha,\beta,\theta,\phi)$.

The matrix of $R_1(\alpha,\beta,\theta,\phi)$ is now obtained from the equations $v_0'=v_0,\ v_*'=v_*,\ v_1'=v_1,\ v_3'=v_2$ and $v_{-2}'=v_3$, where the v_i's are the coordinates in the $(\alpha,\beta,\theta,\phi)$ configuration and the v_i''s in the $(\alpha',\beta',\theta',\phi')$ configuration (see Figure 2). The matrix of R_1 is then

(6)
$$R_1(\alpha,\beta,\theta,\phi)=\begin{bmatrix}1 & 0 & 0\\ 0 & e^{i\theta}\dfrac{\sin\beta}{\sin(\beta-\theta)} & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Similarly, one can find the matrix of R_2 by simultaneously solving the equations $v_0'=v_0,\ v_*'=v_*,\ v_2'=v_1,\ v_{-1}'=v_{-2}$ and $v_3'=v_3$ to get

(7)
$$\begin{aligned}&R_2(\alpha,\beta,\theta,\phi)\\ &=\frac{1}{\sin(\theta+\alpha-\beta)\sin(\phi+\beta-\alpha)}\\ &\quad\cdot\begin{bmatrix}\sin\alpha\sin\theta' e^{i(\alpha-\phi)} & \sin(\alpha-\phi)\sin\theta' e^{i\alpha} & -\sin(\alpha-\phi)\sin\theta' e^{i\alpha}\\ \sin(\beta-\theta)\sin\phi' e^{i\beta} & \sin\phi'\sin\beta e^{i(\beta-\theta)} & -\sin(\beta-\theta)\sin\phi' e^{i\beta}\\ \sin(\theta+\phi)\sin\alpha e^{i\beta} & \sin(\theta+\phi)\sin\beta e^{i\alpha} & A\end{bmatrix},\end{aligned}$$

with $\phi'=\phi+\beta-\alpha$ and $\theta'=\theta+\alpha-\beta$ and

(8)
$$\begin{aligned}A&=\sin\theta\sin\phi'-\sin(\theta+\phi)\sin\beta e^{i\alpha}\\ &=\sin\phi\sin\theta'-\sin(\theta+\phi)\sin\alpha e^{i\beta}\\ &=\sin\theta\sin\phi\cos(\alpha-\beta)-\sin\theta\cos\phi\sin\alpha e^{i\beta}-\cos\theta\sin\phi\sin\beta e^{i\alpha}.\end{aligned}$$

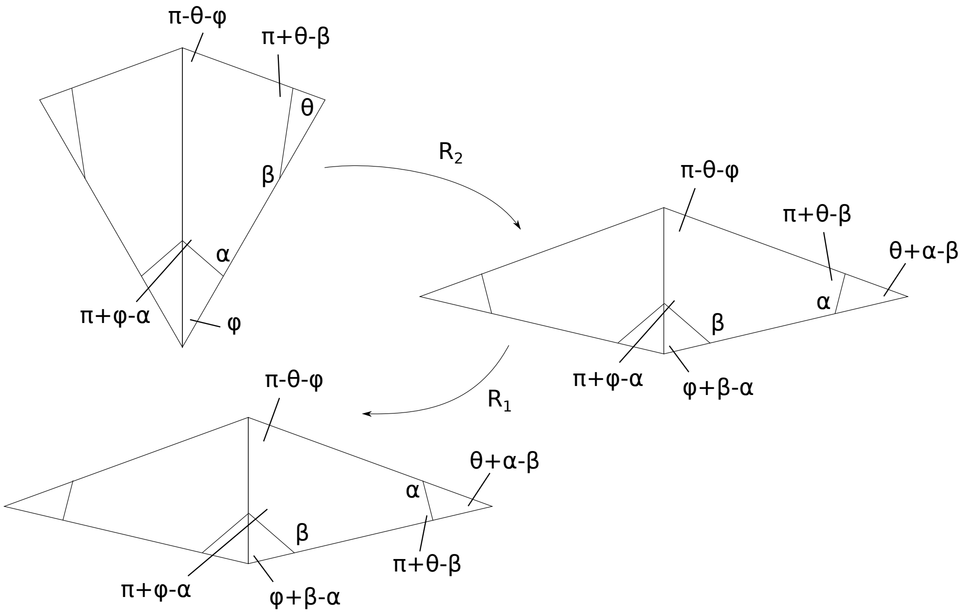


Figure 3. The action of P on the angles.

The third move A_1 is exactly as in previous papers because it starts and lands in the same configuration and its matrix is

$$A_1(\alpha,\beta,\theta,\phi)=\begin{bmatrix} e^{2i\phi} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}. \tag{9}$$

We now want to compose the two matrices to calculate $P=R_1R_2$ and $J=PA_1$. As we already mentioned, after applying the first transformation the angles change. Looking at Figure 3, one can deduce that

$$\begin{aligned} &P(\alpha,\beta,\theta,\phi)\\ &=R_1(\alpha',\beta',\theta',\phi')R_2(\alpha,\beta,\theta,\phi)\\ &=R_1(\beta,\alpha,\theta+\alpha-\beta,\phi+\beta-\alpha)R_2(\alpha,\beta,\theta,\phi)\\ &=\frac{1}{\sin(\theta+\alpha-\beta)\sin(\phi+\beta-\alpha)}\\ &\qquad\cdot\begin{bmatrix} \sin\alpha\sin\theta' e^{i(\alpha-\phi)} & \sin(\alpha-\phi)\sin\theta' e^{i\alpha} & -\sin(\alpha-\phi)\sin\theta' e^{i\alpha}\\ \sin\alpha\sin\phi' e^{i(\alpha+\theta)} & \dfrac{\sin\phi'\sin\beta\sin\alpha}{\sin(\beta-\theta)}e^{i\alpha} & -\sin\alpha\sin\phi' e^{i(\alpha+\theta)}\\ \sin(\theta+\phi)\sin\alpha e^{i\beta} & \sin(\theta+\phi)\sin\beta e^{i\alpha} & A\end{bmatrix}, \end{aligned}\tag{10}$$

where, as before, $\phi'=\phi+\beta-\alpha,\ \theta'=\theta+\alpha-\beta$ and A is as in (8).

On the other hand, $J = PA_1$ is easier to calculate, since A_1 does not change the type of the configuration. So

$$
\begin{aligned}
(11)\quad & J(\alpha,\beta,\theta,\phi) \\
&= P(\alpha,\beta,\theta,\phi)A_1(\alpha,\beta,\theta,\phi) \\
&= \frac{1}{\sin(\theta+\alpha-\beta)\sin(\phi+\beta-\alpha)} \\
&\quad\cdot \begin{bmatrix} \sin\alpha\sin\theta' e^{i(\alpha+\phi)} & \sin(\alpha-\phi)\sin\theta' e^{i\alpha} & -\sin(\alpha-\phi)\sin\theta' e^{i\alpha} \\ \sin\alpha\sin\phi' e^{i(\alpha+\theta+2\phi)} & \frac{\sin\phi'\sin\beta\sin\alpha}{\sin(\beta-\theta)} e^{i\alpha} & -\sin\alpha\sin\phi' e^{i(\alpha+\theta)} \\ \sin(\theta+\phi)\sin\alpha e^{i(\beta+2\phi)} & \sin(\theta+\phi)\sin\beta e^{i\alpha} & A \end{bmatrix},
\end{aligned}
$$

where again $\phi' = \phi+\beta-\alpha$, $\theta' = \theta+\alpha-\beta$ and A is as in (8).

We remark that if we define a second set of coordinates as $\boldsymbol{s} = P^{-1}\boldsymbol{t}$ (as we will do later), the action of R_2 is equivalent to applying R_1 on the $\boldsymbol{s}$-coordinates. In other words, $R_2 = PR_1P^{-1} = R_1R_2R_1R_2^{-1}R_1^{-1}$, which is equivalent to the braid relation

$$R_1R_2R_1 = R_2R_1R_2.$$

Again, to calculate this composition, we need to record how the configuration changes when applying the matrices so we need to prove that the following diagram commutes

$$
\begin{array}{ccc}
(\alpha,\beta,\theta,\phi) & \xrightarrow{R_1^{-1}} & (\alpha,\pi+\theta-\beta,\theta,\phi) \\
{\scriptstyle R_2}\downarrow & & \downarrow{\scriptstyle R_2^{-1}} \\
(\beta,\alpha,\alpha+\theta-\beta,\beta+\phi-\alpha) & & (\pi+\theta-\beta,\alpha,\alpha+\beta-\pi,\pi+\theta+\phi-\alpha-\beta) \\
{\scriptstyle R_1}\uparrow & & \downarrow{\scriptstyle R_1} \\
(\beta,\pi+\theta-\beta,\alpha+\theta-\beta,\beta+\phi-\alpha) & \xleftarrow[R_2]{} & (\pi+\theta-\beta,\beta,\alpha+\beta-\pi,\pi+\theta+\phi-\alpha-\beta)
\end{array}
$$

which is easy to verify by simple calculation.

3C. *The polyhedron.*

3C1. *Complex lines and vertices.* We want to study the metric completion of the moduli space. This means that we want to see what happens when two cone points get closer and closer until they coalesce. We define L_{ij} to be the complex line obtained when v_i and v_j coalesce, for $i, j \in \{0, 1, 2, 3, *\}$. Its normal vector will be denoted as $\boldsymbol{n}_{ij}$. They have equations described in Table 1.

The vertices of the polyhedron are obtained by intersecting pairs of these complex lines (i.e., by making two pairs of cone points coalesce) and they have coordinates given in Table 2.

L_{ij}	equations in terms of the $\boldsymbol{t}$-coordinates
L_{*0}	$t_1 = \dfrac{\sin(\alpha-\phi)\sin\theta}{\sin\alpha\,\sin(\theta+\phi)}$
L_{*1}	$t_1 = e^{-i\phi}\dfrac{\sin\theta}{\sin(\theta+\phi)}$
L_{*2}	$t_2 = e^{i\theta}\dfrac{\sin\phi}{\sin(\theta+\phi)}$
L_{*3}	$t_2 = \dfrac{\sin(\beta-\theta)\sin\phi}{\sin\beta\,\sin(\theta+\phi)}$
L_{01}	$t_1 = 0$
L_{02}	$\dfrac{\sin\alpha}{\sin(\alpha-\phi)}e^{i\phi}t_1 + t_2 = 1$
L_{03}	$\dfrac{\sin\alpha}{\sin(\alpha-\phi)}e^{i\phi}t_1 + e^{-i\theta}\dfrac{\sin\beta}{\sin(\beta-\theta)}t_2 = 1$
L_{12}	$t_1 + t_2 = 1$
L_{13}	$t_1 + e^{-i\theta}\dfrac{\sin\beta}{\sin(\beta-\theta)}t_2 = 1$
L_{23}	$t_2 = 0$

Table 1. The equations defining the complex lines of two cone points collapsing.

3C2. *Second set of coordinates.* It will be useful to define another set of coordinates in order to define the polyhedron explicitly. This is in the spirit of the $\boldsymbol{w}$-coordinates in the previous works and is given by

$$\boldsymbol{s} = \begin{bmatrix} s_1 \\ s_2 \\ 1 \end{bmatrix} = P^{-1}\begin{bmatrix} t_1 \\ t_2 \\ 1 \end{bmatrix}. \tag{12}$$

To calculate the $\boldsymbol{s}$-coordinates, the first thing to do is to calculate $P^{-1}(\alpha, \beta, \theta, \phi)$, with a similar argument as in Section 3B. We recall that this means that P^{-1} is applied to the configuration $(\alpha, \beta, \theta, \phi)$. As shown in Figure 4, P^{-1} acts as follows:

$$\begin{aligned}(\alpha, \beta, \theta, \phi) &\overset{R_1^{-1}}{\longmapsto} (\alpha', \beta', \theta', \phi') = (\alpha, \pi+\theta-\beta, \theta, \phi)\\ &\overset{R_2^{-1}}{\longmapsto} (\alpha'', \beta'', \theta'', \phi'') = (\pi+\theta-\beta, \alpha, \alpha+\beta-\pi, \pi+\theta+\phi-\alpha-\beta),\end{aligned} \tag{13}$$

so

$$P^{-1}(\alpha, \beta, \theta, \phi) = R_2^{-1}(\alpha, \pi+\theta-\beta, \theta, \phi) \circ R_1^{-1}(\alpha, \beta, \theta, \phi).$$

lines	t_k	t_1	t_2
$L_{01}\cap L_{23}$	t_1	0	0
$L_{03}\cap L_{12}$	t_2	$\frac{\sin(\alpha-\phi)(\sin(\beta-\theta)-e^{-i\theta}\sin\beta)}{e^{i\phi}\sin\alpha\sin(\beta-\theta)-e^{-i\theta}\sin\beta\sin(\alpha-\phi)}$	$\frac{e^{i\alpha}\sin(\beta-\theta)\sin\phi}{e^{i\phi}\sin\alpha\sin(\beta-\theta)-e^{-i\theta}\sin\beta\sin(\alpha-\phi)}$
$L_{*0}\cap L_{23}$	t_3	$\frac{\sin(\alpha-\phi)\sin\theta}{\sin\alpha\sin(\theta+\phi)}$	0
$L_{*0}\cap L_{12}$	t_4	$\frac{\sin(\alpha-\phi)\sin\theta}{\sin\alpha\sin(\theta+\phi)}$	$\frac{\sin(\alpha+\theta)\sin\phi}{\sin\alpha\sin(\theta+\phi)}$
$L_{*0}\cap L_{13}$	t_5	$\frac{\sin(\alpha-\phi)\sin\theta}{\sin\alpha\sin(\theta+\phi)}$	$e^{i\theta}\frac{\sin(\alpha+\theta)\sin(\beta-\theta)\sin\phi}{\sin\alpha\sin\beta\sin(\theta+\phi)}$
$L_{*1}\cap L_{23}$	t_6	$e^{-i\phi}\frac{\sin\theta}{\sin(\theta+\phi)}$	0
$L_{*1}\cap L_{02}$	t_7	$e^{-i\phi}\frac{\sin\theta}{\sin(\theta+\phi)}$	$\frac{\sin(\alpha-\theta-\phi)\sin\phi}{\sin(\alpha-\phi)\sin(\theta+\phi)}$
$L_{*1}\cap L_{03}$	t_8	$e^{-i\phi}\frac{\sin\theta}{\sin(\theta+\phi)}$	$e^{i\theta}\frac{\sin(\alpha-\theta-\phi)\sin(\beta-\theta)\sin\phi}{\sin(\alpha-\phi)\sin\beta\sin(\theta+\phi)}$
$L_{*3}\cap L_{01}$	t_9	0	$\frac{\sin(\beta-\theta)\sin\phi}{\sin\beta\sin(\theta+\phi)}$
$L_{*3}\cap L_{12}$	t_{10}	$\frac{\sin(\beta+\phi)\sin\theta}{\sin\beta\sin(\theta+\phi)}$	$\frac{\sin(\beta-\theta)\sin\phi}{\sin\beta\sin(\theta+\phi)}$
$L_{*3}\cap L_{02}$	t_{11}	$e^{-i\phi}\frac{\sin(\alpha-\phi)\sin(\beta+\phi)\sin\theta}{\sin\alpha\sin\beta\sin(\theta+\phi)}$	$\frac{\sin(\beta-\theta)\sin\phi}{\sin\beta\sin(\theta+\phi)}$
$L_{*2}\cap L_{01}$	t_{12}	0	$e^{i\theta}\frac{\sin\phi}{\sin(\theta+\phi)}$
$L_{*2}\cap L_{13}$	t_{13}	$\frac{\sin(\beta-\theta-\phi)\sin\theta}{\sin(\beta-\theta)\sin(\theta+\phi)}$	$e^{i\theta}\frac{\sin\phi}{\sin(\theta+\phi)}$
$L_{*2}\cap L_{03}$	t_{14}	$e^{-i\phi}\frac{\sin(\alpha-\phi)\sin(\beta-\theta-\phi)\sin\theta}{\sin\alpha\sin(\beta-\theta)\sin(\theta+\phi)}$	$e^{i\theta}\frac{\sin\phi}{\sin(\theta+\phi)}$

Table 2. The coordinates of the vertices.

Explicitly, we have

$$P^{-1}(\alpha,\beta,\theta,\phi) = \begin{bmatrix} -\sin\alpha\sin\theta' e^{-i(\alpha-\phi)} & -\frac{\sin(\alpha-\phi)\sin\theta'\sin\beta}{\sin(\beta-\theta)}e^{-i(\alpha+\theta)} & \sin(\alpha-\phi)\sin\theta' e^{-i\alpha} \\ \sin\beta\sin\phi' e^{i(\beta-\theta)} & \sin\beta\sin\phi' e^{i(\beta-\theta)} & -\sin\beta\sin\phi' e^{i(\beta-\theta)} \\ \sin(\theta+\phi)\sin\alpha e^{i(\beta-\theta)} & -\sin(\theta+\phi)\sin\beta e^{-i(\alpha+\theta)} & B \end{bmatrix}, \tag{14}$$

where $\phi' = \pi+\theta+\phi-\alpha-\beta$, $\theta' = \alpha+\beta-\pi$ and B is

$$\begin{aligned} B &= -\sin\theta'\sin\phi - \sin(\theta+\phi)\sin\alpha e^{i(\beta-\theta)} \\ &= -\sin\phi'\sin\theta + \sin(\theta+\phi)\sin(\beta-\theta)e^{-i\alpha}. \end{aligned} \tag{15}$$

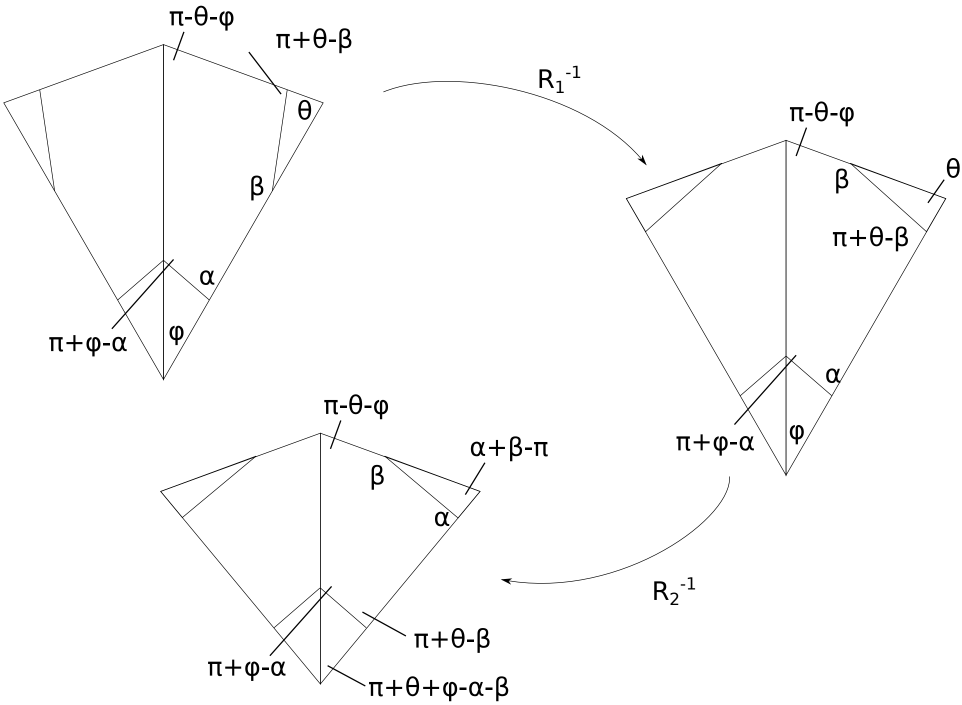


Figure 4. The action of P^{-1} on the angles.

One can easily verify that the matrices of $P(\alpha, \beta, \theta, \phi)$ and $P^{-1}(\alpha, \beta, \theta, \phi)$ in Equations (10) and (14) respectively satisfy Equation (5).

We now apply P^{-1} to the lines and vertices described in Tables 1 and 2 to obtain their $\boldsymbol{s}$-coordinates.

The complex lines have $\boldsymbol{s}$-coordinates described in Table 3, and the vertices have $\boldsymbol{s}$-coordinates shown in Table 4.

Remark 3.2. The equations of the lines are of the same form as the ones for the $\boldsymbol{t}$-coordinates, except for the sign of the exponential for the t_1-coordinate and up to substituting $(\alpha, \beta, \theta, \phi)$ with the new angles as in Figure 4, i.e., up to substituting $(\alpha, \beta, \theta, \phi)$ with $(\alpha', \beta', \theta', \phi') = (\pi+\theta-\beta, \alpha, \alpha+\beta-\pi, \pi+\theta+\phi-\alpha-\beta)$. The same is true for the coordinates of the vertices. In other words, up to remembering that

$$\alpha' = \pi+\theta-\beta, \quad \beta' = \alpha, \quad \theta' = \alpha+\beta-\pi \quad \text{and} \quad \phi' = \pi+\theta+\phi-\alpha-\beta$$

as in Figure 4, the $\boldsymbol{s}$-coordinates can be equivalently listed as in Tables 5 and 6.

3C3. *The polyhedron.* Many of the vertices are contained in bisectors (see Section 2B), and we use these bisectors to cut out a polyhedron, which will be called D. In

L_{ij}	equations in terms of the s-coordinates
L_{*0}	$s_1 = -\dfrac{\sin(\alpha-\phi)\sin(\alpha+\beta)}{\sin(\beta-\theta)\sin(\theta+\phi)}$
L_{*1}	$s_2 = -e^{i(\alpha+\beta)}\dfrac{\sin(\alpha+\beta-\theta-\phi)}{\sin(\theta+\phi)}$
L_{*2}	$s_2 = \dfrac{\sin(\alpha+\beta-\theta-\phi)\sin\beta}{\sin\alpha\sin(\theta+\phi)}$
L_{*3}	$s_1 = e^{-i(\alpha+\beta-\theta-\phi)}\dfrac{\sin(\alpha+\beta)}{\sin(\theta+\phi)}$
L_{01}	$-\dfrac{\sin(\beta-\theta)}{\sin(\alpha-\phi)}e^{i(\alpha+\beta-\theta-\phi)}s_1 + s_2 = 1$
L_{02}	$-\dfrac{\sin(\beta-\theta)}{\sin(\alpha-\phi)}e^{i(\alpha+\beta-\theta-\phi)}s_1 - e^{-i(\alpha+\beta)}\dfrac{\sin\alpha}{\sin\beta}s_2 = 1$
L_{03}	$s_1 = 0$
L_{12}	$s_2 = 0$
L_{13}	$s_1 + s_2 = 1$
L_{23}	$s_1 - e^{-i(\alpha+\beta)}\dfrac{\sin\alpha}{\sin\beta}s_2 = 1$

Table 3. The equations defining the complex lines of two cone points collapsing in terms of the $\boldsymbol{s}$-coordinates.

particular, and following [Parker 2006; Boadi and Parker 2015; Pasquinelli 2016], we will denote the bisectors as shown in Table 7.

The reason for the bisectors to be denoted as $B(T)$ is that we want the map T to send the side $B(T)$ to $B(T^{-1})$, for $T \in \{P^{\pm}, J^{\pm}, R_1^{\pm}, R_2^{\pm}\}$. The following lemma shows that this is the case.

Lemma 3.3. *In* $\boldsymbol{t}$*- and* $\boldsymbol{s}$*-coordinates and writing*

$$\theta' = \alpha+\beta-\pi \quad \textit{and} \quad \phi' = \pi+\theta+\phi-\alpha-\beta,$$

we have

$$\begin{aligned}
\operatorname{Im}(t_1) \le 0 \quad &\textit{if and only if} \quad |\langle \boldsymbol{t}, \boldsymbol{n}_{*1}\rangle| \le |\langle \boldsymbol{t}, P^{-1}(\boldsymbol{n}_{*3})\rangle|,\\
\operatorname{Im}(s_1) \ge 0 \quad &\textit{if and only if} \quad |\langle \boldsymbol{s}, \boldsymbol{n}_{*3}\rangle| \le |\langle \boldsymbol{s}, P(\boldsymbol{n}_{*1})\rangle|,\\
\operatorname{Im}(e^{i\phi}t_1) \ge 0 \quad &\textit{if and only if} \quad |\langle \boldsymbol{t}, \boldsymbol{n}_{*0}\rangle| \le |\langle \boldsymbol{t}, J^{-1}(\boldsymbol{n}_{*0})\rangle|,\\
\operatorname{Im}(e^{-i\phi'}s_1) \le 0 \quad &\textit{if and only if} \quad |\langle \boldsymbol{s}, \boldsymbol{n}_{*0}\rangle| \le |\langle \boldsymbol{s}, J(\boldsymbol{n}_{*0})\rangle|,\\
\operatorname{Im}(t_2) \ge 0 \quad &\textit{if and only if} \quad |\langle \boldsymbol{t}, \boldsymbol{n}_{*2}\rangle| \le |\langle \boldsymbol{t}, R_1^{-1}(\boldsymbol{n}_{*3})\rangle|,\\
\operatorname{Im}(e^{-i\theta}t_2) \le 0 \quad &\textit{if and only if} \quad |\langle \boldsymbol{t}, \boldsymbol{n}_{*3}\rangle| \le |\langle \boldsymbol{t}, R_1(\boldsymbol{n}_{*2})\rangle|,\\
\operatorname{Im}(s_2) \ge 0 \quad &\textit{if and only if} \quad |\langle \boldsymbol{s}, \boldsymbol{n}_{*1}\rangle| \le |\langle \boldsymbol{s}, R_2^{-1}(\boldsymbol{n}_{*2})\rangle|,\\
\operatorname{Im}(e^{-i\theta'}s_2) \le 0 \quad &\textit{if and only if} \quad |\langle \boldsymbol{s}, \boldsymbol{n}_{*2}\rangle| \le |\langle \boldsymbol{s}, R_2(\boldsymbol{n}_{*1})\rangle|.
\end{aligned}$$

$\boldsymbol{t}_k$	s_1	s_2
$\boldsymbol{t}_1$	$\dfrac{e^{-i\alpha}\sin(\alpha-\phi)\sin(\alpha+\beta)}{\sin(\alpha-\phi)\sin\beta-e^{-i(\theta+\phi)}\sin\alpha\sin(\beta-\theta)}$	$\dfrac{e^{i(\beta-\theta)}\sin\beta\sin(\alpha+\beta-\theta-\phi)}{\sin(\alpha-\phi)\sin\beta-e^{-i(\theta+\phi)}\sin\alpha\sin(\beta-\theta)}$
$\boldsymbol{t}_2$	0	0
$\boldsymbol{t}_3$	$-\dfrac{\sin(\alpha-\phi)\sin(\alpha+\beta)}{\sin(\beta-\theta)\sin(\theta+\phi)}$	$-e^{i(\alpha+\beta)}\dfrac{\sin(\alpha+\theta)\sin\beta\sin(\alpha+\beta-\theta-\phi)}{\sin(\beta-\theta)\sin\alpha\sin(\theta+\phi)}$
$\boldsymbol{t}_4$	$-\dfrac{\sin(\alpha-\phi)\sin(\alpha+\beta)}{\sin(\beta-\theta)\sin(\theta+\phi)}$	0
$\boldsymbol{t}_5$	$-\dfrac{\sin(\alpha-\phi)\sin(\alpha+\beta)}{\sin(\beta-\theta)\sin(\theta+\phi)}$	$\dfrac{\sin(\alpha+\theta)\sin(\alpha+\beta-\theta-\phi)}{\sin(\beta-\theta)\sin(\theta+\phi)}$
$\boldsymbol{t}_6$	$\dfrac{\sin(\alpha+\beta)\sin(\theta+\phi-\alpha)}{\sin\beta\sin(\theta+\phi)}$	$-e^{i(\alpha+\beta)}\dfrac{\sin(\alpha+\beta-\theta-\phi)}{\sin(\theta+\phi)}$
$\boldsymbol{t}_7$	$-e^{-i(\alpha+\beta-\theta-\phi)}\dfrac{\sin(\alpha-\phi)\sin(\alpha+\beta)\sin(\theta+\phi-\alpha)}{\sin(\beta-\theta)\sin\beta\sin(\theta+\phi)}$	$-e^{i(\alpha+\beta)}\dfrac{\sin(\alpha+\beta-\theta-\phi)}{\sin(\theta+\phi)}$
$\boldsymbol{t}_8$	0	$-e^{i(\alpha+\beta)}\dfrac{\sin(\alpha+\beta-\theta-\phi)}{\sin(\theta+\phi)}$
$\boldsymbol{t}_9$	$e^{-i(\alpha+\beta-\theta-\phi)}\dfrac{\sin(\alpha+\beta)}{\sin(\theta+\phi)}$	$\dfrac{\sin(\beta+\phi)\sin(\alpha+\beta-\theta-\phi)}{\sin(\alpha-\phi)\sin(\theta+\phi)}$
$\boldsymbol{t}_{10}$	$e^{-i(\alpha+\beta-\theta-\phi)}\dfrac{\sin(\alpha+\beta)}{\sin(\theta+\phi)}$	0
$\boldsymbol{t}_{11}$	$e^{-i(\alpha+\beta-\theta-\phi)}\dfrac{\sin(\alpha+\beta)}{\sin(\theta+\phi)}$	$-e^{i(\alpha+\beta)}\dfrac{\sin(\beta+\phi)\sin\beta\sin(\alpha+\beta-\theta-\phi)}{\sin(\alpha-\phi)\sin\alpha\sin(\theta+\phi)}$
$\boldsymbol{t}_{12}$	$-e^{-i(\alpha+\beta-\theta-\phi)}\dfrac{\sin(\alpha-\phi)\sin(\theta+\phi-\beta)\sin(\alpha+\beta)}{\sin(\beta-\theta)\sin\alpha\sin(\theta+\phi)}$	$\dfrac{\sin\beta\sin(\alpha+\beta-\theta-\phi)}{\sin\alpha\sin(\theta+\phi)}$
$\boldsymbol{t}_{13}$	$\dfrac{\sin(\theta+\phi-\beta)\sin(\alpha+\beta)}{\sin\alpha\sin(\theta+\phi)}$	$\dfrac{\sin\beta\sin(\alpha+\beta-\theta-\phi)}{\sin\alpha\sin(\theta+\phi)}$
$\boldsymbol{t}_{14}$	0	$\dfrac{\sin\beta\sin(\alpha+\beta-\theta-\phi)}{\sin\alpha\sin(\theta+\phi)}$

Table 4. The $\boldsymbol{s}$-coordinates of the vertices.

The proof follows the one of equivalent lemmas in the previous works. In particular, it can be found in [Parker 2006, Lemma 4.6; Boadi and Parker 2015, Lemma 4.2; Pasquinelli 2016, Lemma 7.2]. One just needs to remark that $\boldsymbol{n}_{*i}$ depends on the configuration we are using. So, for example, the first line of the lemma is

$$|\langle \boldsymbol{t}, \boldsymbol{n}_{*1}(\alpha,\beta,\theta,\phi)\rangle| \le \left|\left\langle \boldsymbol{t}, P^{-1}(\boldsymbol{n}_{*3}(\beta,\pi+\theta-\beta,\theta+\alpha-\beta,\phi+\beta-\alpha))\right\rangle\right|,$$

since $(\alpha,\beta,\theta,\phi)\overset{P}{\mapsto}(\beta,\pi+\theta-\beta,\theta+\alpha-\beta,\phi+\beta-\alpha)$. The rest is similar.

Now the polyhedron $D = D(\alpha,\beta,\theta,\phi)$ is defined as the intersection of all the half spaces in the lemma. More precisely, it will be

$$D(\alpha,\beta,\theta,\phi) = \left\{ \boldsymbol{t} = P(\boldsymbol{s}) : \begin{array}{ll} \arg(t_1)\in(-\phi,0), & \arg(t_2)\in(0,\theta), \\ \arg(s_1)\in(0,\phi'), & \arg(s_2)\in(0,\theta') \end{array} \right\}, \tag{16}$$

L_{ij}	equations in terms of the s-coordinates
L_{*0}	$s_1 = \dfrac{\sin(\alpha'-\phi')\sin\theta'}{\sin\alpha'\sin(\theta'+\phi')}$
L_{*1}	$s_2 = e^{i\theta'}\dfrac{\sin\phi'}{\sin(\theta'+\phi')}$
L_{*2}	$s_2 = \dfrac{\sin(\beta'-\theta')\sin\phi'}{\sin\beta'\sin(\theta'+\phi')}$
L_{*3}	$s_1 = e^{i\phi'}\dfrac{\sin\theta'}{\sin(\theta'+\phi')}$
L_{01}	$\dfrac{\sin\alpha'}{\sin(\alpha'-\phi')}e^{-i\phi'}s_1 + s_2 = 1$
L_{02}	$\dfrac{\sin\alpha'}{\sin(\alpha'-\phi')}e^{-i\phi'}s_1 + e^{-i\theta'}\dfrac{\sin\beta'}{\sin(\beta'-\theta')}s_2 = 1$
L_{03}	$s_1 = 0$
L_{12}	$s_2 = 0$
L_{13}	$s_1 + s_2 = 1$
L_{23}	$s_1 + e^{-i\theta'}\dfrac{\sin\beta'}{\sin(\beta'-\theta')}s_2 = 1$

Table 5. The equations defining the complex lines of two cone points collapsing in terms of the s-coordinates and of the angles in the target configuration.

where, as before, we have $\theta' = \alpha + \beta - \pi$ and $\phi' = \pi + \theta + \phi - \alpha - \beta$.

The sides (codimension 1 cells) of the polyhedron will be defined as $S(T) = D \cap B(T)$, for $T \in \{P^{\pm}, J^{\pm}, R_1^{\pm}, R_2^{\pm}\}$ again. Each of them is contained in one of the bisectors in the table.

3D. ***The combinatorial structure of the polyhedron.*** We now want to study the combinatorics of the polyhedron $D(\alpha, \beta, \theta, \phi)$.

The sides all have the same combinatorial structure. In particular, they will look like in Figure 5. This is the same structure as the one of the sides of the polyhedron in [Pasquinelli 2016] and first appeared as the combinatorial structure of 2 of the 10 sides in [Deraux et al. 2005]. Each side corresponds to fixing the argument of one of the coordinates. Then there will be one triangular ridge (e.g., the bottom one) where the coordinate is equal to zero and a second triangular ridge (e.g., the top one) where the coordinate has another fixed value. The complex lines interpolating between the two will be the slices of the foliation mentioned in Section 2B. The

t_k	s_1	s_2
t_1	$\dfrac{-e^{-i\beta'}\sin(\alpha'-\phi')\sin\theta'}{\sin(\alpha'-\phi')\sin(\beta'-\theta')-e^{-i(\theta'+\phi')}\sin\alpha'\sin\beta'}$	$\dfrac{-e^{-i\alpha'}\sin\phi'\sin(\beta'-\theta')}{\sin(\alpha'-\phi')\sin(\beta'-\theta')-e^{-i(\theta'+\phi')}\sin\alpha'\sin\beta'}$
t_2	0	0
t_3	$\dfrac{\sin(\alpha'-\phi')\sin\theta'}{\sin\alpha'\sin(\theta'+\phi')}$	$e^{i\theta'}\dfrac{\sin(\alpha'+\theta')\sin(\beta'-\theta')\sin\phi'}{\sin\alpha'\sin\beta'\sin(\theta'+\phi')}$
t_4	$\dfrac{\sin(\alpha'-\phi')\sin\theta'}{\sin\alpha'\sin(\theta'+\phi')}$	0
t_5	$\dfrac{\sin(\alpha'-\phi')\sin\theta'}{\sin\alpha'\sin(\theta'+\phi')}$	$\dfrac{\sin(\alpha'+\theta')\sin\phi'}{\sin\alpha'\sin(\theta'+\phi')}$
t_6	$\dfrac{\sin\theta'\sin(\beta'-\theta'-\phi')}{\sin(\beta'-\theta')\sin(\theta'+\phi')}$	$e^{i\theta'}\dfrac{\sin\phi'}{\sin(\theta'+\phi')}$
t_7	$e^{i\phi'}\dfrac{\sin(\alpha'-\phi')\sin\theta'\sin(\beta'-\theta'-\phi')}{\sin\alpha'\sin(\beta'-\theta')\sin(\theta'+\phi')}$	$e^{i\theta'}\dfrac{\sin\phi'}{\sin(\theta'+\phi')}$
t_8	0	$e^{i\theta'}\dfrac{\sin\phi'}{\sin(\theta'+\phi')}$
t_9	$e^{i\phi'}\dfrac{\sin\theta'}{\sin(\theta'+\phi')}$	$\dfrac{\sin(\alpha'-\theta'-\phi')\sin\phi'}{\sin(\alpha'-\phi')\sin(\theta'+\phi')}$
t_{10}	$e^{i\phi'}\dfrac{\sin\theta'}{\sin(\theta'+\phi')}$	0
t_{11}	$e^{i\phi'}\dfrac{\sin\theta'}{\sin(\theta'+\phi')}$	$e^{i\theta'}\dfrac{\sin(\alpha'-\theta'-\phi')\sin(\beta'-\theta')\sin\phi'}{\sin(\alpha'-\phi')\sin\beta'\sin(\theta'+\phi')}$
t_{12}	$e^{i\phi'}\dfrac{\sin(\alpha'-\phi')\sin(\beta'+\phi')\sin\theta'}{\sin\alpha'\sin\beta'\sin(\theta'+\phi')}$	$\dfrac{\sin(\beta'-\theta')\sin\phi'}{\sin\beta'\sin(\theta'+\phi')}$
t_{13}	$\dfrac{\sin(\beta'+\phi')\sin\theta'}{\sin\beta'\sin(\theta'+\phi')}$	$\dfrac{\sin(\beta'-\theta')\sin\phi'}{\sin\beta'\sin(\theta'+\phi')}$
t_{14}	0	$\dfrac{\sin(\beta'-\theta')\sin\phi'}{\sin\beta'\sin(\theta'+\phi')}$

Table 6. The $\boldsymbol{s}$-coordinates of the vertices in terms of the angles in the target configuration.

edge connecting the two triangles is contained in the spine of the bisector and always contains one of the vertices $\boldsymbol{t}_1$ or $\boldsymbol{t}_2$. The pentagonal side ridges containing the vertical edge are contained in totally geodesic Lagrangian planes and are the extremities of the foliation by meridians. We claim that in each side the modulus of the coordinate we are considering varies between the two values it assumes on the top and bottom triangular ridges. To check this, for example, in $S(J)$, we need to check that $|t_1|$ in $\boldsymbol{t}_{11}$ and $\boldsymbol{t}_{14}$ is smaller than $|t_1|$ in $\boldsymbol{t}_6$, $\boldsymbol{t}_7$ and $\boldsymbol{t}_8$ ($|t_1|$ has the same value in these three vertices, since they are contained in the complex line L_{*1}) and so on. It is easy to check that this is true for each side as long as

$$\begin{aligned} &\sin(\alpha+\beta-\pi)\geq 0, \quad \sin(\pi+\alpha+\beta-\theta-\phi)\geq 0,\\ &\sin(\alpha+\theta-\beta)\geq 0, \qquad\qquad \sin(\beta+\phi-\alpha)\geq 0. \end{aligned} \tag{17}$$

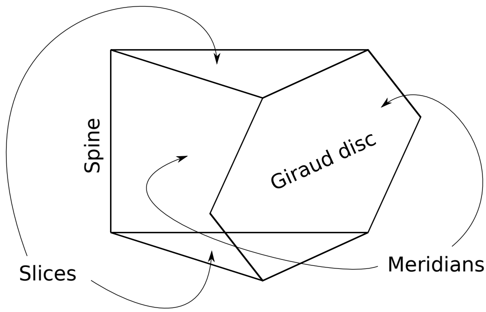


Figure 5. The combinatorial structure of a side.

Remembering the action of P and P^{-1} on the angles, this means that we are just asking for the configuration after applying these two maps to make sense in our coordinates and is always the case for the values we are considering.

This gives:

Lemma 3.4. *If the parameters satisfy (17), then*:

$$\begin{aligned}
&\textit{In } S(P), \textit{ we have} && |t_1| \le \frac{\sin(\alpha-\phi)\sin\theta}{\sin\alpha\,\sin(\theta+\phi)}.\\
&\textit{In } S(J), \textit{ we have} && |t_1| \le \frac{\sin\theta}{\sin(\theta+\phi)}.\\
&\textit{In } S(R_1), \textit{ we have} && |t_2| \le \frac{\sin(\beta-\theta)\sin\phi}{\sin\beta\,\sin(\theta+\phi)}.\\
&\textit{In } S(R_1^{-1}), \textit{ we have} && |t_2| \le \frac{\sin\phi}{\sin(\theta+\phi)}.\\
&\textit{In } S(P^{-1}), \textit{ we have} && |s_1| \le -\frac{\sin(\alpha-\phi)\sin(\alpha+\beta)}{\sin(\beta-\theta)\sin(\theta+\phi)}.\\
&\textit{In } S(J^{-1}), \textit{ we have} && |s_1| \le -\frac{\sin(\alpha+\beta)}{\sin(\theta+\phi)}.\\
&\textit{In } S(R_2), \textit{ we have} && |s_2| \le \frac{\sin(\alpha+\beta-\theta-\phi)\sin\beta}{\sin\alpha\,\sin(\theta+\phi)}.\\
&\textit{In } S(R_2^{-1}), \textit{ we have} && |s_2| \le \frac{\sin(\alpha+\beta-\theta-\phi)\sin\beta}{\sin\alpha\,\sin(\theta+\phi)}.
\end{aligned}$$

Now, following [Pasquinelli 2016], we will see that the combinatorics changes with the values of the angles. According to the values of the parameters, we will have occasions where the three vertices on L_{*i} collapse to a single vertex, for $i = 0, 1, 2, 3$.

Proposition 3.5. *We have*:

- *The vertices on* L_{*0} *collapse when* $\alpha-\phi \ge \pi-\theta-\phi$, *i.e., when* $\pi-\alpha-\theta \le 0$.
- *The vertices on* L_{*1} *collapse when* $\pi-\alpha \ge \pi-\theta-\phi$, *i.e.,* $\alpha-\theta-\phi \le 0$.

bisector	equation	points in the bisector
$B(P)$	$\mathrm{Im}(t_1)=0$	$\boldsymbol{t}_1, \boldsymbol{t}_3, \boldsymbol{t}_4, \boldsymbol{t}_5, \boldsymbol{t}_9, \boldsymbol{t}_{10}, \boldsymbol{t}_{12}, \boldsymbol{t}_{13}$
$B(P^{-1})$	$\mathrm{Im}(s_1)=0$	$\boldsymbol{t}_2, \boldsymbol{t}_3, \boldsymbol{t}_4, \boldsymbol{t}_5, \boldsymbol{t}_6, \boldsymbol{t}_8, \boldsymbol{t}_{13}, \boldsymbol{t}_{14}$
$B(J)$	$\mathrm{Im}(e^{i\phi}t_1)=0$	$\boldsymbol{t}_1, \boldsymbol{t}_6, \boldsymbol{t}_7, \boldsymbol{t}_8, \boldsymbol{t}_9, \boldsymbol{t}_{11}, \boldsymbol{t}_{12}, \boldsymbol{t}_{14}$
$B(J^{-1})$	$\mathrm{Im}(e^{-i\phi}s_1)=0$	$\boldsymbol{t}_2, \boldsymbol{t}_7, \boldsymbol{t}_8, \boldsymbol{t}_9, \boldsymbol{t}_{10}, \boldsymbol{t}_{11}, \boldsymbol{t}_{12}, \boldsymbol{t}_{14}$
$B(R_1)$	$\mathrm{Im}(t_2)=0$	$\boldsymbol{t}_1, \boldsymbol{t}_3, \boldsymbol{t}_4, \boldsymbol{t}_6, \boldsymbol{t}_7, \boldsymbol{t}_9, \boldsymbol{t}_{10}, \boldsymbol{t}_{11}$
$B(R_1^{-1})$	$\mathrm{Im}(e^{-i\theta}t_2)=0$	$\boldsymbol{t}_1, \boldsymbol{t}_3, \boldsymbol{t}_5, \boldsymbol{t}_6, \boldsymbol{t}_8, \boldsymbol{t}_{12}, \boldsymbol{t}_{13}, \boldsymbol{t}_{14}$
$B(R_2)$	$\mathrm{Im}(s_2)=0$	$\boldsymbol{t}_2, \boldsymbol{t}_4, \boldsymbol{t}_5, \boldsymbol{t}_9, \boldsymbol{t}_{10}, \boldsymbol{t}_{12}, \boldsymbol{t}_{13}, \boldsymbol{t}_{14}$
$B(R_2^{-1})$	$\mathrm{Im}(e^{-i\theta}s_2)=0$	$\boldsymbol{t}_2, \boldsymbol{t}_3, \boldsymbol{t}_4, \boldsymbol{t}_6, \boldsymbol{t}_7, \boldsymbol{t}_8, \boldsymbol{t}_{10}, \boldsymbol{t}_{11}$

Table 7. Bisector notation and description.

- *The vertices on* L_{*2} *collapse when* $\sin(\beta-\theta)/\sin\beta \le \sin\phi/\sin(\theta+\phi)$, *i.e.*, $\beta-\theta-\phi\le 0$.
- *The vertices on* L_{*3} *collapse when* $\pi-\beta\le\phi$, *i.e.*, $\pi-\beta-\phi\le 0$.

In fact, for example, the vertices $\boldsymbol{t}_3$, $\boldsymbol{t}_4$ and $\boldsymbol{t}_5$ on L_{*0} collapse if, when making T_1 as big as possible, before we can have $v_0\equiv v_*$, we have that v_1 hits the left-hand vertex of T_2 and so $v_1\equiv v_2\equiv v_3$. This implies that there is no other choice for z_2 but to be zero, instead of having the three choices that give the three possible vertices having $v_0\equiv v_*$. Translated on the parameters, this gives that $\alpha-\phi\ge\pi-\theta-\phi$. The others can be verified in a similar way.

Later we will. We now want to study all possible side (3-dimensional facets) intersections in order to be able to list all possible ridges (2-dimensional facets) and edges (1-dimensional facets) and hence describe the combinatorics of the polyhedron.

Proposition 3.6. *The following side intersections consist of the union of two edges*:

$$S(P)\cap S(J^{-1})=\gamma_{10,9}\cup\gamma_{9,12},\qquad S(R_1^{-1})\cap S(J^{-1})=\gamma_{8,14}\cup\gamma_{14,12},$$
$$S(P)\cap S(R_2^{-1})=\gamma_{3,4}\cup\gamma_{4,10},\qquad S(J)\cap S(R_2)=\gamma_{9,12}\cup\gamma_{12,14},$$
$$S(R_1)\cap S(R_2)=\gamma_{4,10}\cup\gamma_{10,9},\qquad S(J)\cap S(P^{-1})=\gamma_{6,8}\cup\gamma_{8,14},$$
$$S(R_1)\cap S(P^{-1})=\gamma_{4,3}\cup\gamma_{3,6},\qquad S(R_1^{-1})\cap S(R_2^{-1})=\gamma_{3,6}\cup\gamma_{6,8},$$

where $\gamma_{i,j}$ *is the geodesic segment between the vertices* $\boldsymbol{t}_i$ *and* $\boldsymbol{t}_j$.

The proof of this proposition follows exactly the ones in [Parker 2006, Appendix A; Pasquinelli 2016, Proposition 7.8].

Proposition 3.7. *The bisector intersections satisfy*:

- *A point* $\boldsymbol{t}$ *in the side intersection* $S(P)\cap S(P^{-1})$, *with*

$$t_1\ne\frac{\sin\theta\,\sin(\alpha-\phi)}{\sin\alpha\,\sin(\theta+\phi)}\quad\text{and}\quad s_1\ne-\frac{\sin(\alpha+\beta)\,\sin(\alpha-\phi)}{\sin(\beta-\theta)\,\sin(\theta+\phi)},$$

belongs to the edge $\gamma_{5,13}$.

- *A point* $\boldsymbol{t}$ *in the side intersection* $S(J)\cap S(R_2^{-1})$, *with*

$$t_1 \neq e^{-i\phi}\frac{\sin\theta}{\sin(\theta+\phi)} \quad \textit{and} \quad s_2 \neq -e^{i(\alpha+\beta)}\frac{\sin(\alpha+\beta-\theta-\phi)}{\sin(\theta+\phi)},$$

belongs to the edge $\gamma_{7,11}$.

- *Moreover*, *a point* $\boldsymbol{t}$ *in the side intersection* $S(R_2)\cap S(R_1^{-1})$, *with*

$$t_2 \neq e^{i\theta}\frac{\sin\phi}{\sin(\theta+\phi)} \quad \textit{and} \quad s_2 \neq \frac{\sin(\alpha+\beta-\theta-\phi)\sin\beta}{\sin\alpha\sin(\theta+\phi)},$$

belongs to the edge $\gamma_{5,13}$.

- *Finally*, *a point* $\boldsymbol{t}$ *in the side intersection* $S(R_1)\cap S(J^{-1})$, *with*

$$t_2 \neq \frac{\sin(\beta-\theta)\sin\phi}{\sin\beta\sin(\theta+\phi)} \quad \textit{and} \quad s_1 \neq e^{-i(\alpha+\beta-\theta-\phi)}\frac{\sin(\alpha+\beta)}{\sin(\theta+\phi)},$$

belongs to the edge $\gamma_{7,11}$.

We will prove the first point and the others are proved in the exact same way. The proof is very similar to the ones in [Parker 2006; Pasquinelli 2016].

Proof. Let us take $\boldsymbol{t}\in S(P)\cap S(P^{-1})$. Then

$$t_1 = x, \quad s_1 = u$$

and by hypothesis and using Lemma 3.4 we have

$$x \leq \frac{\sin\theta\sin(\alpha-\phi)}{\sin\alpha\sin(\theta+\phi)}, \quad u \leq -\frac{\sin(\alpha+\beta)\sin(\alpha-\phi)}{\sin(\beta-\theta)\sin(\theta+\phi)}. \tag{18}$$

Then using (12) one can express t_2 and s_2 in terms of x and u as follows:

$$\begin{aligned}
&\big(\sin(\theta+\phi)\sin\alpha x - \sin(\alpha-\phi)\sin\theta\big)s_2\\
&\quad = -\sin(\beta-\theta)\sin\theta e^{i(\alpha+\beta-\theta-\phi)}u + \sin(\theta+\phi)\sin(\beta-\theta)e^{i(\alpha+\beta-\theta)}ux\\
&\qquad\qquad + (\sin(\alpha+\beta)\sin\phi e^{i(\beta-\theta)} + \sin(\theta+\phi)\sin\alpha)x - \sin(\alpha-\phi)\sin\theta,
\end{aligned}$$

$$\begin{aligned}
&\big(-\sin(\theta+\phi)\sin(\beta-\theta)u - \sin(\alpha-\phi)\sin(\alpha+\beta)\big)t_2\\
&\quad = \frac{\sin(\beta-\theta)}{\sin\beta}e^{-i\theta}(\sin\theta\sin(\alpha+\beta-\theta-\phi)e^{i\alpha}u - \sin(\theta+\phi)\sin\alpha e^{i(\alpha+\beta-\theta)}ux\\
&\quad\quad + (\sin\alpha\sin(\alpha+\beta)e^{-i\phi} - \sin(\alpha-\phi)\sin(\alpha+\beta))x - \sin(\theta+\phi)\sin(\beta-\theta)).
\end{aligned}$$

Now, we know by Lemma 3.3 that inside D we have

$$\begin{aligned}
0 &\geq \operatorname{Im} e^{-i\theta}t_2\\
&= \frac{\sin(\beta-\theta)\sin\alpha}{\sin\beta}\\
&\qquad \cdot \frac{\sin(\alpha+\beta)\sin\phi x + \sin(\theta+\phi)\sin(\alpha+\beta-\theta)ux - \sin\theta\sin(\alpha+\beta-\theta-\phi)u}{\sin(\theta+\phi)\sin(\beta-\theta)u + \sin(\alpha-\phi)\sin(\alpha+\beta)},
\end{aligned}$$

but by (18) we know that the denominator is strictly negative and so the numerator must be positive.

Again by Lemma 3.3, $\boldsymbol{t}$ satisfies

$$\begin{aligned} 0 &\leq \operatorname{Im} s_2 \\ &= \sin(\beta-\theta) \\ &\quad \cdot \frac{\sin(\alpha+\beta)\sin\phi x + \sin(\theta+\phi)\sin(\alpha+\beta-\theta)ux - \sin\theta\sin(\alpha+\beta-\theta-\phi)u}{\sin(\theta+\phi)\sin\alpha x - \sin(\alpha-\phi)\sin\theta}, \end{aligned}$$

and since by (18) the denominator must be strictly negative, then the numerator must be negative.

But since the two numerators coincide, then they must both equal 0. This means that the point we are considering must be also in $S(R_1^{-1})$ and in $S(R_2)$, which means that we are on edge $\gamma_{5,13}$. □

Remark 3.8. The proof relies on Lemma 3.4. As we will see in Section 5E, there are cases in which (17) is not satisfied. In term of configurations, this means that one needs to consider a slightly different configuration of triangles (see Section 5E). Using the new configuration one can prove an equivalent statement using the exact same strategy of proof as in [Parker 2006; Pasquinelli 2016].

4. The polyhedron in the case of 2-fold symmetry

We will now consider the case where two of the cone points have the same cone angle. First we will describe which sets of cone angles give a lattice, then we will show how to use the polyhedron in Section 3 to build a fundamental domain for them.

4A. ***Lattices with 2-fold symmetry.*** As mentioned in the introduction (Section 1), the lattices we are considering were introduced by Deligne and Mostow starting from a ball 5-tuple. This is equivalent to considering a cone metric on a sphere of area 1 with prescribed cone singularities of angles $(\theta_0, \dots, \theta_4)$, with $0 < \theta_i < 2\pi$ and satisfying the discrete Gauss–Bonnet formula. A sphere with 5 cone points has the structure of a two dimensional complex hyperbolic space, as proved by Thurston [1998] and showed in Section 3.

Among these, we will consider the lattices with 2-fold symmetry, which means that two of the five cone points will have the same cone angle. We will assume that the 2-fold symmetry is given by $\theta_1 = \theta_2$. Occasionally the lattices will have an extra symmetry and we will also have $\theta_0 = \theta_3$. We will use the parameters in (2) to describe the lattices, except that now $\alpha = \beta$.

By similarity with the 3-fold symmetry case, to each lattice we will associate numbers p, p', k, k', l, l', d, which are the orders of some maps in the group and

lattice	θ_0	θ_1	θ_2	θ_3	θ_4
(6,6,3)	$2\pi/3$	$5\pi/3$	$5\pi/3$	$2\pi/3$	$4\pi/3$
(10,10,5)	$4\pi/5$	$7\pi/5$	$7\pi/5$	$4\pi/5$	$8\pi/5$
(12,12,6)	$5\pi/6$	$4\pi/3$	$4\pi/3$	$5\pi/6$	$5\pi/3$
(18,18,9)	$8\pi/9$	$11\pi/9$	$11\pi/9$	$8\pi/9$	$16\pi/9$
(4,4,3)	$5\pi/6$	$5\pi/3$	$5\pi/3$	$5\pi/6$	π
(4,4,5)	$11\pi/10$	$7\pi/5$	$7\pi/5$	$11\pi/10$	π
(4,4,6)	$7\pi/6$	$4\pi/3$	$4\pi/3$	$7\pi/6$	π
(3,3,4)	$7\pi/6$	$3\pi/2$	$3\pi/2$	$7\pi/6$	$2\pi/3$
(3,3,3)	π	$5\pi/3$	$5\pi/3$	π	$2\pi/3$
(2,6,6)	π	$4\pi/3$	$4\pi/3$	$5\pi/3$	$2\pi/3$
(2,4,3)	$5\pi/6$	$5\pi/3$	$5\pi/3$	$4\pi/3$	$\pi/2$
(2,3,3)	π	$5\pi/3$	$5\pi/3$	$4\pi/3$	$\pi/3$
(3,4,4)	π	$3\pi/2$	$3\pi/2$	$7\pi/6$	$5\pi/6$

Table 8. The lattices we are considering.

are defined as follows:

$$\begin{aligned} &\frac{\pi}{p}=\theta, && \frac{\pi}{k}=\phi, && \frac{\pi}{l}=\alpha-\theta-\phi, && \frac{\pi}{d}=\pi-\alpha-\theta \\ &\frac{\pi}{p'}=\alpha-\frac{\pi}{2}, && \frac{\pi}{k'}=\pi+\theta+\phi-2\alpha, && \frac{\pi}{l'}=\pi-\alpha-\phi. \end{aligned} \tag{19}$$

In particular, we will use (p, k, p') to denote the configuration (α, θ, ϕ) and give the other values in terms of these. Observe that in the double 2-fold symmetry case (i.e., when we also have $\theta_0 = \theta_3$), we have $\theta = \phi$ and so the lattice is of the form (p, p, p'). Notice also that in the 3-fold symmetry case one would have $k = k'$, $l = l'$ and $p = 2p'$. In fact k and k' will be the orders of $A_1(\alpha, \beta, \theta, \phi)$ for two of the different configurations we will consider (see (9)) which coincide in the 3-fold symmetry case. A similar thing happens for l and l'. The values p' and p here are the orders of $R_1(\pi+\theta-\alpha, \alpha, 2\alpha-\pi, \pi+\theta+\phi-2\alpha)$ and $R_1 \circ R_1(\alpha, \pi+\theta-\alpha, \theta, \phi)$ respectively (remember that the composition is done as in (4)), and notice that they are applied to different configurations. Since in the 3-fold symmetry case the three configurations we consider coincide, p will be the order of the square of R_1, which has order p' and hence $p = 2p'$.

In Table 8 we give the values of the cone angles for the lattice (p, k, p'). These are all the lattices with 2-fold symmetry in the original list by Deligne and Mostow and together with the 3-fold symmetry lattices form the whole list of Deligne–Mostow lattices in dimension 2.

The following table records the values of the parameters in (19) for each of the lattices we are considering.

lattice	p	k	p'	k'	l	l'	d
(6,6,3)	6	6	3	−3	2	∞	∞
(10,10,5)	10	10	5	−5	2	5	5
(12,12,6)	12	12	6	−6	2	4	4
(18,18,9)	18	18	9	−9	2	3	3
(4,4,3)	4	4	3	−6	3	−12	−12
(4,4,5)	4	4	5	10	5	20	20
(4,4,6)	4	4	6	6	6	12	12
(3,3,4)	3	3	4	6	12	−12	−12
(3,6,3)	3	6	3	−6	3	∞	−6
(2,6,6)	2	6	6	3	∞	6	−6
(2,4,3)	2	4	3	12	12	−12	−3
(2,3,3)	2	3	3	6	∞	−6	−3
(3,4,4)	3	4	4	12	6	∞	−12

Table 9. The values of the parameters for our lattices.

4B. *The fundamental polyhedron.*

4B1. *Definition.* In this section we will see how one can use the general polyhedron described in Section 3C to build a fundamental domain for Deligne–Mostow lattices with 2-fold symmetry. From now on we will consider a sphere with cone singularities $(\theta_0, \theta_1, \theta_2, \theta_3, \theta_4)$ in the list in Section 4A. This means that we have two equal angles at the vertices v_1 and v_2. In the configurations as described in Section 3, this means that $\alpha = \beta$. Since the case that we treated before is when the three angles at v_1, v_2 and v_3 were equal, by analogy we also want to consider the configurations where the two equal angles are at v_1 and v_2, at v_2 and v_3 or at v_1 and v_3. We will call these configurations of type ①, ② and ③ respectively. Remark that configuration of type ⓘ corresponds to having the cone angles satisfying $\theta_i = \theta_{i+1}$, for indices $i = 1, 2, 3$ taken mod 3.

We will build a polyhedron for each of these cases and use their union to build a fundamental domain for the lattices. On the parameters $(\alpha, \beta, \theta, \phi)$,

type ① corresponds to $(\alpha, \alpha, \theta, \phi)$,

type ② corresponds to $(\pi + \theta - \alpha, \alpha, 2\alpha - \pi, \pi + \theta + \phi - 2\alpha)$,

type ③ corresponds to $(\alpha, \pi + \theta - \alpha, \theta, \phi)$.

For each type, we will consider the $\boldsymbol{t}$-coordinates and $\boldsymbol{s}$-coordinates. We will have $\boldsymbol{x}$-, $\boldsymbol{y}$- and $\boldsymbol{z}$-coordinates as $\boldsymbol{t}$-coordinates of the configuration of type ①, ② and ③ respectively. We will also have $\boldsymbol{u}$-, $\boldsymbol{v}$- and $\boldsymbol{w}$-coordinates, representing copies of type ①, ② and ③ respectively and being the $\boldsymbol{s}$-coordinates of one of the previous ones. More precisely, the relation between $\boldsymbol{x}$-, $\boldsymbol{y}$-, $\boldsymbol{z}$- and $\boldsymbol{u}$-, $\boldsymbol{v}$-, $\boldsymbol{w}$-coordinates is as follows. Since P^{-1} acts on the copies as explained in Figure 4, then, for example, a configuration of type ① will be sent to one of type ②. This means that the coordinates defined as $P^{-1}(\boldsymbol{x})$ will be the $\boldsymbol{v}$-coordinates. With a similar argument, one gets

$$\boldsymbol{u} = P^{-1}(\boldsymbol{z}), \quad \boldsymbol{v} = P^{-1}(\boldsymbol{x}), \quad \boldsymbol{w} = P^{-1}(\boldsymbol{y}). \tag{20}$$

In other words, the $\boldsymbol{u}$-, $\boldsymbol{v}$- and $\boldsymbol{w}$-coordinates will be the coordinates for the configuration of type ①, ② and ③ respectively, obtained after applying P to the standard configuration of type ③, ① and ② respectively.

We will start from the configuration of type ③, with its $\boldsymbol{z}$-coordinates as the $\boldsymbol{t}$-coordinates of configuration $(\alpha, \pi+\theta-\alpha, \theta, \phi)$. The $\boldsymbol{x}$- and $\boldsymbol{y}$-coordinates will be determined by the action of the moves R_1 and R_2^{-1} respectively. See Figure 6 for more details. As mentioned, each configuration will give us a polyhedron of the same type as D in (16).

We will first explain what is the relation between the $\boldsymbol{x}$-, $\boldsymbol{y}$- and $\boldsymbol{z}$-coordinates. Since copies of type ① and ③ are swapped by R_1, it is natural to define

$$\boldsymbol{x} = R_1(\alpha, \alpha, \theta, \phi)\boldsymbol{z}. \tag{21}$$

Since the $\boldsymbol{w}$- and $\boldsymbol{u}$-coordinates are also of type ③ and ① respectively, one would also want

$$\boldsymbol{u} = R_1(\alpha, \pi+\theta-\alpha, \theta, \phi)\boldsymbol{w}. \tag{22}$$

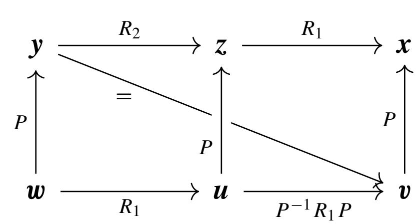
Using the definition of $\boldsymbol{u}$- and $\boldsymbol{w}$-coordinates, together with the previous formula, the $\boldsymbol{y}$-coordinates are defined as

$$\boldsymbol{z} = R_2(\pi+\theta-\alpha, \alpha, 2\alpha-\pi, \pi+\theta+\phi-2\alpha)\boldsymbol{y}. \tag{23}$$

Using Equations (20), (21) and (23), one can also see that

$$\boldsymbol{v} = P^{-1}\boldsymbol{x} = P^{-1}R_1\boldsymbol{z} = P^{-1}R_1R_2\boldsymbol{y} = \boldsymbol{y}. \tag{24}$$

The following digram summarises the relations on the coordinates.



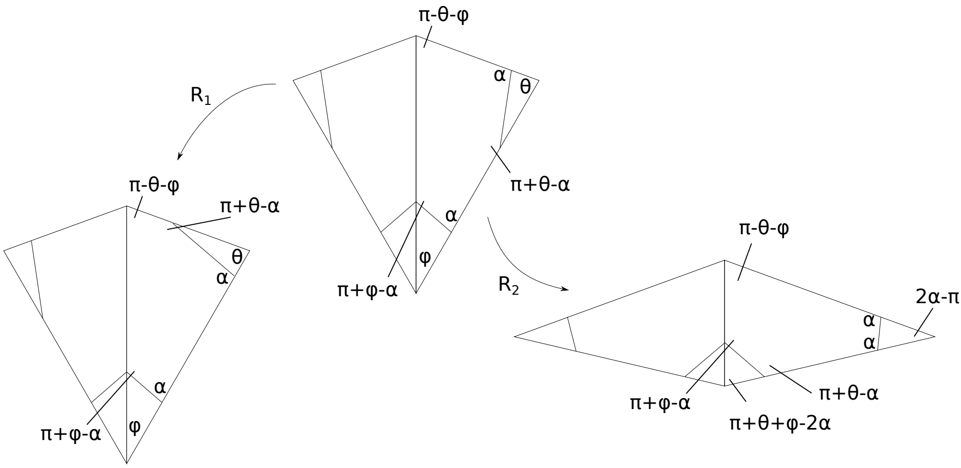


Figure 6. The representative for each configuration type.

For each coordinate type, we can define a polyhedron as in (16). This will give us three components of our fundamental polyhedron D and we will write

$$
(25)\quad D = D_1 \cup D_2 \cup D_3,\qquad \textit{with} \begin{cases} D_1 = D(\alpha, \alpha, \theta, \phi) = R_1^{-1}(D_3), \\ D_2 = D(\pi + \theta - \alpha, \alpha, 2\alpha - \pi, \pi + \theta + \phi - 2\alpha) = R_2(D_3), \\ D_3 = D(\alpha, \pi + \theta - \alpha, \theta, \phi). \end{cases}
$$

In coordinates, the polyhedron D_1 is defined as

$$
D_1 = \left\{ \boldsymbol{x} = P(\boldsymbol{v})\colon \begin{array}{ll} \arg(x_1) \in (-\phi, 0), & \arg(x_2) \in (0, \theta), \\ \arg(v_1) \in (0, \pi + \theta + \phi - 2\alpha), & \arg(v_2) \in (0, 2\alpha - \pi) \end{array} \right\},
$$

the polyhedron D_2 is

$$
D_2 = \left\{ \boldsymbol{y} = P(\boldsymbol{w})\colon \begin{array}{ll} \arg(y_1) \in (-(\pi + \theta + \phi - 2\alpha, 0), & \arg(y_2) \in (0, 2\alpha - \pi), \\ \arg(w_1) \in (0, \phi), & \arg(w_2) \in (0, \theta) \end{array} \right\}
$$

and the polyhedron D_3 is defined as

$$
D_3 = \left\{ \boldsymbol{z} = P(\boldsymbol{u})\colon \begin{array}{ll} \arg(z_1) \in (-\phi, 0), & \arg(z_2) \in (0, \theta), \\ \arg(u_1) \in (0, \phi), & \arg(u_2) \in (0, \theta) \end{array} \right\}.
$$

Due to the fact that the matrix for R_1 is extremely simple, we will keep track only of three sets of coordinates, namely $\boldsymbol{z}$-, $\boldsymbol{w}$- and $\boldsymbol{y}$-coordinates and use the relations in (21), (22) and (24) to give the other coordinates in term of these.

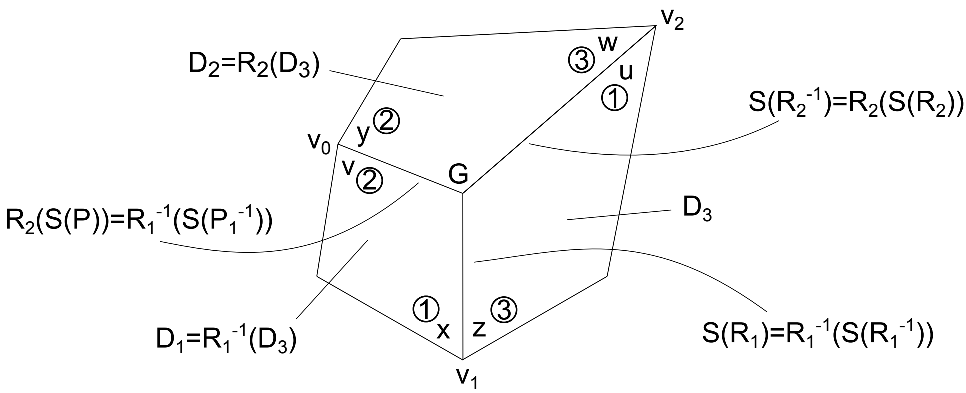


Figure 7. The interaction of the polyhedra and their coordinates.

Then we can describe the polyhedron as follows.

$$D = \left\{ z = R_2(\boldsymbol{y}) = R_2 P(\boldsymbol{w}) \colon \begin{array}{ll} \arg(z_1) \in (-\phi, 0), & \arg(z_2) \in (-\theta, \theta), \\ \arg(w_1) \in (0, \phi), & \arg(w_2) \in (-\theta, \theta), \\ \arg(y_1) \in (-\phi', \phi'), & \arg(y_2) \in (0, \theta') \end{array} \right\},$$

with $\phi' = \pi + \theta + \phi - 2\alpha$ and $\theta' = 2\alpha - \pi$.

In Figure 7 one can see how the polyhedra and the coordinates interact. The three polyhedra intersect pairwise in a side and all three have a common Giraud disc G. Passing from $\boldsymbol{t}$- to $\boldsymbol{s}$-coordinates changes the type of configuration from ⓙto ⓘ within the same polyhedron D_j, where $i = j - 2$, taken mod 3. The three special vertices $\boldsymbol{v}_0$, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are the origin of one of the coordinates.

4B2. *Vertices of D.* The vertices of D will be of three types. Some will come from D_1 and they will be called $\boldsymbol{x}_i$, for $i = 1, \dots, 14$, some will be the vertices of D_2 and we will denote them $\boldsymbol{y}_i$, for $i = 1, \dots, 14$ and finally there will be the vertices $\boldsymbol{z}_i$'s for $i = 1, \dots, 14$, coming from D_3. Since the three polyhedra intersect there will be some vertices that are repeated. Table 10 describes all the vertices. In the first column there will be the label we choose for the vertex, in the second, third and fourth column its name in D_3, D_1 and D_2 respectively (if there is one), and in the final columns we will record which coordinates have a "nice" form.

This reflects how the D_i's glue together. In particular, the polyhedra D_1 and D_3 glue along

$$\{\operatorname{Im} z_2 = 0\} \cap D_3 = \{\operatorname{Im} e^{-i\theta} x_2 = 0\} \cap D_1, \tag{26}$$

while D_2 and D_3 are glued along

$$\{\operatorname{Im} e^{-i\theta} u_2 = 0\} \cap D_3 = \{\operatorname{Im} w_2 = 0\} \cap D_2 \tag{27}$$

D	D_3	D_1	D_2	$\arg z_1$	$\arg z_2$	$\arg w_1$	$\arg w_2$	$\arg y_1$	$\arg y_2$
$\boldsymbol{v}_0$		$\boldsymbol{x}_2$	$\boldsymbol{y}_1$					$y_1=0$	$y_2=0$
$\boldsymbol{v}_1$	$\boldsymbol{z}_1$	$\boldsymbol{x}_1$		$z_1=0$	$z_2=0$				
$\boldsymbol{v}_2$	$\boldsymbol{z}_2$		$\boldsymbol{y}_2$			$w_1=0$	$w_2=0$		
$\boldsymbol{v}_3$	$\boldsymbol{z}_3$	$\boldsymbol{x}_3$	$\boldsymbol{y}_5$	0	$z_2=0$	0	0	0	θ'
$\boldsymbol{v}_4$	$\boldsymbol{z}_4$	$\boldsymbol{x}_5$	$\boldsymbol{y}_4$	0	0	0	$w_2=0$	0	0
$\boldsymbol{v}_5$	$\boldsymbol{z}_5$			0	θ	0	$-\theta$		
$\boldsymbol{v}_6$	$\boldsymbol{z}_6$	$\boldsymbol{x}_6$	$\boldsymbol{y}_{13}$	$-\phi$	$z_2=0$	0	0	0	θ'
$\boldsymbol{v}_7$	$\boldsymbol{z}_7$	$\boldsymbol{x}_8$	$\boldsymbol{y}_{12}$	$-\phi$	0	ϕ	0	$y_1=0$	θ'
$\boldsymbol{v}_8$	$\boldsymbol{z}_8$		$\boldsymbol{y}_{14}$	$-\phi$	θ	$w_1=0$	0	$-\phi'$	θ'
$\boldsymbol{v}_9$	$\boldsymbol{z}_9$	$\boldsymbol{x}_{12}$		$z_1=0$	0	ϕ	$-\theta$	ϕ'	0
$\boldsymbol{v}_{10}$	$\boldsymbol{z}_{10}$	$\boldsymbol{x}_{13}$	$\boldsymbol{y}_{10}$	0	0	ϕ	$w_2=0$	0	0
$\boldsymbol{v}_{11}$	$\boldsymbol{z}_{11}$	$\boldsymbol{x}_{14}$	$\boldsymbol{y}_9$	$-\phi$	0	ϕ	0	$y_1=0$	0
$\boldsymbol{v}_{12}$	$\boldsymbol{z}_{12}$			$z_1=0$	θ	ϕ	$-\theta$		
$\boldsymbol{v}_{13}$	$\boldsymbol{z}_{13}$			0	θ	0	$-\theta$		
$\boldsymbol{v}_{14}$	$\boldsymbol{z}_{14}$			$-\phi$	θ	0	$-\theta$		
$\boldsymbol{v}_{16}$		$\boldsymbol{x}_4$	$\boldsymbol{y}_3$	0	$-\theta$	0	θ	0	$y_2=0$
$\boldsymbol{v}_{17}$		$\boldsymbol{x}_7$		$-\phi$	$-\theta$			ϕ'	θ'
$\boldsymbol{v}_{18}$		$\boldsymbol{x}_9$		$z_1=0$	$-\theta$			ϕ'	0
$\boldsymbol{v}_{19}$		$\boldsymbol{x}_{10}$		0	$-\theta$			ϕ'	$y_2=0$
$\boldsymbol{v}_{20}$		$\boldsymbol{x}_{11}$		$-\phi$	$-\theta$			ϕ'	θ'
$\boldsymbol{v}_{21}$			$\boldsymbol{y}_6$			0	θ	$-\phi'$	$y_2=0$
$\boldsymbol{v}_{22}$			$\boldsymbol{y}_7$			ϕ	θ	$-\phi'$	0
$\boldsymbol{v}_{23}$			$\boldsymbol{y}_8$			$w_1=0$	θ	$-\phi'$	θ'
$\boldsymbol{v}_{24}$			$\boldsymbol{y}_{11}$			ϕ	θ	$-\phi'$	0

Table 10. The vertices of D.

and D_1 and D_2 intersect along

(28) $$\{\mathrm{Im}\, v_1 = 0\} \cap D_1 = \{\mathrm{Im}\, y_1 = 0\} \cap D_2.$$

Moreover, all three will intersect along the Giraud disc G containing the ridge bounded by vertices $\boldsymbol{v}_3$, $\boldsymbol{v}_4$, $\boldsymbol{v}_6$, $\boldsymbol{v}_7$, $\boldsymbol{v}_{10}$ and $\boldsymbol{v}_{11}$ (see Figure 7).

Remark 4.1. Using Table 1 one can obtain the equations of the complex lines for our three configurations and see that the following lines coincide:

(1) $$L_{*0}(\alpha, \pi+\theta-\alpha, \theta, \phi) = L_{*0}(\pi+\theta-\alpha, \alpha, 2\alpha-\pi, \pi+\theta+\phi-2\alpha) = L_{*0}(\alpha, \alpha, \theta, \phi),$$

$$\text{(2)}\quad L_{*3}(\alpha,\pi+\theta-\alpha,\theta,\phi)=L_{*3}(\pi+\theta-\alpha,\alpha,2\alpha-\pi,\pi+\theta+\phi-2\alpha)=L_{*2}(\alpha,\alpha,\theta,\phi),$$

$$\text{(3)}\quad L_{*1}(\alpha,\pi+\theta-\alpha,\theta,\phi)=L_{*2}(\pi+\theta-\alpha,\alpha,2\alpha-\pi,\pi+\theta+\phi-2\alpha)=L_{*1}(\alpha,\alpha,\theta,\phi).$$

4B3. *Sides and side pairing maps.* In view of applying Poincaré polyhedron theorem in Section 5, we need to analyse the sides of D and explain how we have some maps pairing them.

Clearly, the sides of D will be the union of all sides in D_i, with $i=1,2,3$, except for the three sides along which two of the copies glue. Some of the sides combine to create a single larger side. Remembering (25), the sides (illustrated in Figure 8 with their side pairings) will be as follows:

$$\begin{array}{llll}
S(J), & S(P), & S(R_1), & S(R_2),\\
S(J^{-1}), & S(P^{-1}), & S(R_1^{-1}), & S(R_2^{-1}),\\
R_1^{-1}S(J), & R_1^{-1}S(P), & R_1^{-1}S(R_1), & R_1^{-1}S(R_2),\\
R_1^{-1}S(J^{-1}), & R_1^{-1}S(P^{-1}), & R_1^{-1}S(R_1^{-1}), & R_1^{-1}S(R_2^{-1}),\\
R_2S(J), & R_2S(P), & R_2S(R_1), & R_2S(R_2),\\
R_2S(J^{-1}), & R_2S(P^{-1}), & R_2S(R_1^{-1}), & R_2S(R_2^{-1}).
\end{array}$$

Now the gluing of the three polyhedra (see Equations (26), (27) and (28)) tells us that

$$R_1^{-1}S(R_1^{-1})=S(R_1),\quad R_2S(R_2)=S(R_2^{-1}),\quad R_1^{-1}S(P^{-1})=R_2S(P),$$

so these sides are now internal (see Figure 7).

The side pairings will be obtained by adapting to the union of the three polyhedra the equivalent on each D_i of the side pairings in previous works (see [Parker 2006, Section 4.3; Boadi and Parker 2015, Section 5.3; Pasquinelli 2016, Section 8.3.1]). In other words, in each copy we need to consider R_1, R_2, P and J and adapt them to act on the sides of D. We will describe all of them treating the z-coordinates as the main coordinates. In other words, we will give the matrix as applied to the z-coordinates of the point.

First consider R_1 and R_2. Since applying $R_2(\alpha,\alpha,\theta,\phi)$ to a point in its $\boldsymbol{x}$-coordinates is equivalent to applying $R_1(\pi+\theta-\alpha,\alpha,2\alpha-\pi,\pi+\theta+\phi-2\alpha)$ to its $\boldsymbol{v}=\boldsymbol{y}$-coordinates, these combine to a single side pairing

$$R_1(\pi+\theta-\alpha,\alpha,2\alpha-\pi,\pi+\theta+\phi-2\alpha)=\begin{bmatrix}1&0&0\\0&-e^{2i\alpha}&0\\0&0&1\end{bmatrix}.$$

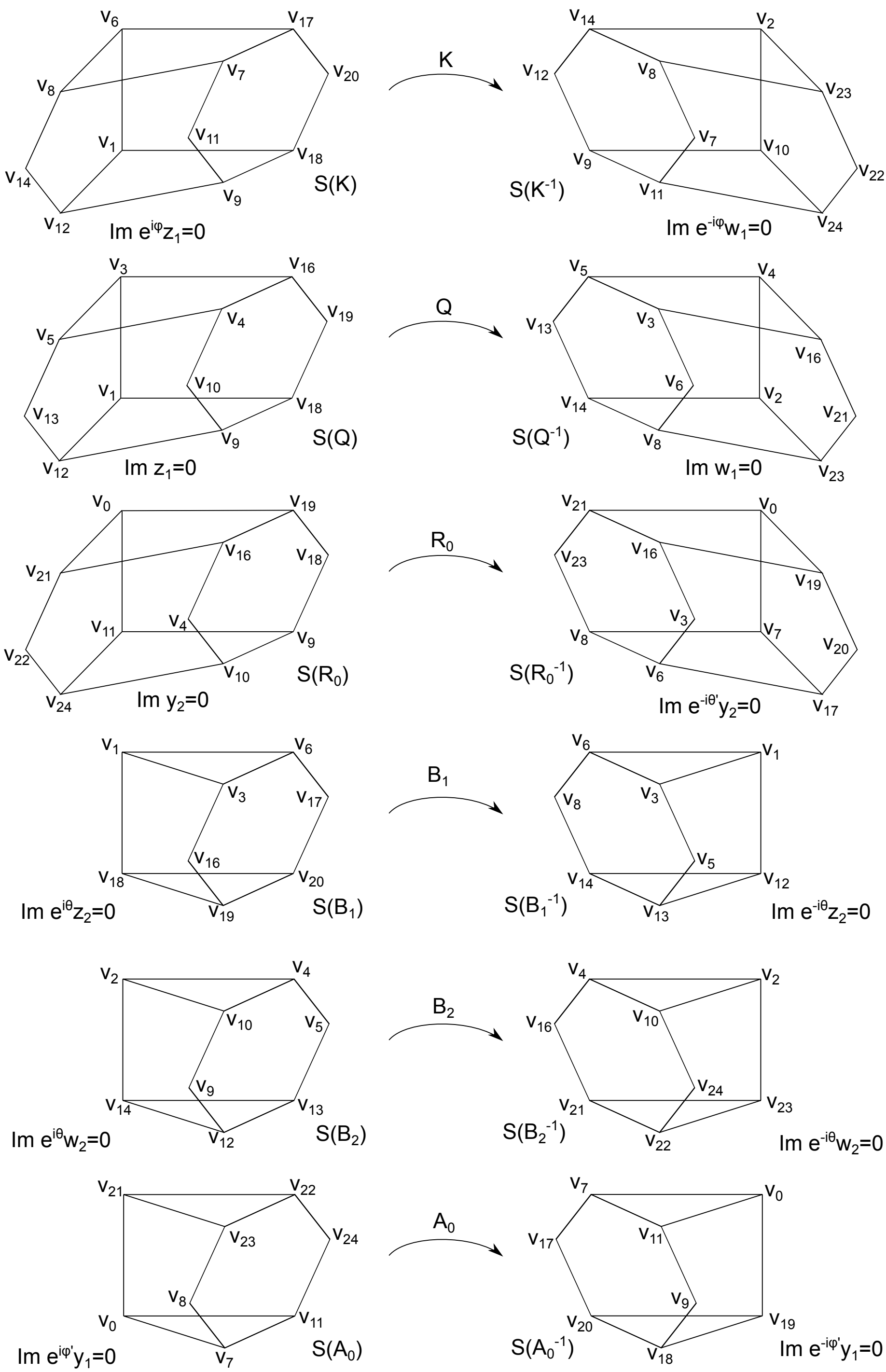


Figure 8. The sides of D.

This is the side pairing as applied on the $\boldsymbol{y}$-coordinates. We will hence consider

$$\begin{aligned} R_0 &= R_2 R_1 R_2^{-1} \\ &= R_2(\pi+\theta-\alpha, \alpha, 2\alpha-\pi, \pi+\theta+\phi-2\alpha) \\ &\qquad \circ R_1(\pi+\theta-\alpha, \alpha, 2\alpha-\pi, \pi+\theta+\phi-2\alpha) \circ R_2^{-1}(\alpha, \pi+\theta-\alpha, \theta, \phi), \end{aligned}$$

which includes the change of coordinates.

Now consider $R_1(\alpha, \alpha, \theta, \phi)$ and $R_1(\alpha, \pi+\theta-\alpha, \theta, \phi)$. The target side of the former coincides with the source side of the latter and is the (now internal) side $D_1 \cap D_3$. This means that we can compose the two maps and have a new side pairing

$$B_1 = R_1(\alpha, \pi+\theta-\alpha, \theta, \phi) R_1(\alpha, \alpha, \theta, \phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{2i\theta} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Remark that even though it looks like this is the matrix we use when applying the map to a point in its $\boldsymbol{x}$-coordinates, composing it with the change of coordinates from our coordinates (the $\boldsymbol{z}$-coordinates) one gets that in terms of matrices

$$\begin{aligned} B_1(\alpha, \pi+\theta-\alpha, \theta, \phi) &= R_1^{-1}(\alpha, \alpha, \theta, \phi) B_1(\alpha, \alpha, \theta, \phi) R_1(\alpha, \pi+\theta-\alpha, \theta, \phi) \\ &= B_1(\alpha, \alpha, \theta, \phi). \end{aligned}$$

Similarly, the target side of $R_2(\alpha, \pi+\theta-\alpha, \theta, \phi)$ and the starting side of $R_2(\pi+\theta-\alpha, \alpha, 2\alpha-\pi, \pi+\theta+\phi-2\alpha)$ is the common side of D_2 and D_3. We can then define

$$B_2 = R_2(\pi+\theta-\alpha, \alpha, 2\alpha-\pi, \pi+\theta+\phi-2\alpha) R_2(\alpha, \pi+\theta-\alpha, \theta, \phi).$$

The map B_2 is already defined to act on the $\boldsymbol{z}$-coordinates. As we said for R_2 and R_1, B_2 acts as B_1, but on the $\boldsymbol{u}$-coordinates.

The side pairings P and J related the $\boldsymbol{t}$- and $\boldsymbol{s}$-coordinates of the polyhedron, but the side pairing property relied on the fact that the source and target configurations were of the same type. Adapting this to our case means that we want to consider the maps relating $\boldsymbol{z}$- and $\boldsymbol{w}$-coordinates, $\boldsymbol{x}$- and $\boldsymbol{u}$-coordinates and $\boldsymbol{y}$- and $\boldsymbol{v}$-coordinates. The map relating $\boldsymbol{y}$- and $\boldsymbol{v}$-coordinates is the identity and it indeed maps the common side between D_1 and D_2 to itself. Since this side is in the interior of D, we can ignore it. Composing the map obtained with $A_1(\pi+\theta-\alpha, \alpha, 2\alpha-\pi, \pi+\theta+\phi-2\alpha)$ to compute the equivalent of J and applying the change of coordinates to our main coordinates, we obtain the side pairing

$$A_0 = R_2 A_1 R_2^{-1}.$$

Now, we have

$$\boldsymbol{w} = P^{-1}\boldsymbol{y} = P^{-1} R_2^{-1} \boldsymbol{z} = Q^{-1} \boldsymbol{z}$$

and

$$\boldsymbol{u} = R_2^{-1} R_1^{-1} R_1^{-1} x,$$

which translates to the z-coordinates as Q^{-1} again. Then $Q = R_1 R_2 R_1$ will be our new side pairing. Moreover, we will consider $K = Q A_1$.

Putting all this information together and remarking that $J^3 = \mathrm{Id}$, one gets that the side pairings are

$$\begin{aligned}
K = J R_1 = R_2 J &: R_1^{-1} S(J) \cup S(J) \mapsto S(J^{-1}) \cup R_2 S(J^{-1}),\\
Q = P R_1 = R_2 P &: R_1^{-1} S(P) \cup S(P) \mapsto S(P^{-1}) \cup R_2 S(P),\\
R_0 = R_1^{-1} R_2 R_1 = R_2 R_1 R_2^{-1} &: R_1^{-1} S(R_2) \cup R_2 S(R_1) \mapsto R_1^{-1} S(R_2^{-1}) \cup R_2 S(R_1^{-1}),\\
B_1 = R_1 R_1 &: R_1^{-1} S(R_1) \mapsto S(R_1^{-1}),\\
B_2 = R_2 R_2 &: S(R_2) \mapsto R_2 S(R_2^{-1}),\\
A_0 = R_1^{-1} J^{-1} J^{-1} R_2^{-1} &: R_2 S(J) \mapsto R_1^{-1} S(J^{-1}).
\end{aligned}$$

As mentioned for the general case, the sides are contained in bisectors. One can rewrite Lemma 3.3 for each copy and eliminate the inequalities related to the sides along which the polyhedra glue. Translating the inequalities on the right hand side into z-coordinates and giving all the $\boldsymbol{n}_{*i}$ in terms of the configuration $(\alpha, \pi + \theta - \alpha, \theta, \phi)$ (using Remark 4.1), we get the following lemma.

Lemma 4.2. *We have*:

$$\begin{aligned}
\mathrm{Im}(z_1) \le 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*1}\rangle| \le |\langle \boldsymbol{z}, P^{-1}(\boldsymbol{n}_{*3})\rangle|,\\
\mathrm{Im}(e^{i\phi} z_1) \ge 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*0}\rangle| \le |\langle \boldsymbol{z}, K^{-1}(\boldsymbol{n}_{*0})\rangle|,\\
\mathrm{Im}(e^{-i\theta} z_2) \le 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*3}\rangle| \le |\langle \boldsymbol{z}, B_1(\boldsymbol{n}_{*3})\rangle|,\\
\mathrm{Im}(e^{i\theta} z_2) \ge 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*3}\rangle| \le |\langle \boldsymbol{z}, B_1^{-1}(\boldsymbol{n}_{*3})\rangle|,\\
\mathrm{Im}(e^{i\phi'} y_1) \ge 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*0}\rangle| \le |\langle \boldsymbol{z}, K^{2}(\boldsymbol{n}_{*0})\rangle|,\\
\mathrm{Im}(y_2) \ge 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*1}\rangle| \le |\langle \boldsymbol{z}, Q^{-1} B_1(\boldsymbol{n}_{*3})\rangle|,\\
\mathrm{Im}(e^{-i\theta'} y_2) \le 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*3}\rangle| \le |\langle \boldsymbol{z}, B_1^{-1} Q(\boldsymbol{n}_{*1})\rangle|,\\
\mathrm{Im}(e^{-i\phi'} y_1) \le 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*0}\rangle| \le |\langle \boldsymbol{z}, K^{-2}(\boldsymbol{n}_{*0})\rangle|,\\
\mathrm{Im}(w_1) \ge 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*3}\rangle| \le |\langle \boldsymbol{z}, Q(\boldsymbol{n}_{*1})\rangle|,\\
\mathrm{Im}(e^{-i\phi} w_1) \le 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*0}\rangle| \le |\langle \boldsymbol{z}, K(\boldsymbol{n}_{*0})\rangle|,\\
\mathrm{Im}(e^{-i\theta} w_2) \le 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*1}\rangle| \le |\langle \boldsymbol{z}, B_2(\boldsymbol{n}_{*1})\rangle|,\\
\mathrm{Im}(e^{i\theta} w_2) \ge 0 &\quad \textit{if and only if} \quad |\langle \boldsymbol{z}, \boldsymbol{n}_{*1}\rangle| \le |\langle \boldsymbol{z}, B_2^{-1}(\boldsymbol{n}_{*1})\rangle|.
\end{aligned}$$

5. Main theorem

In this section we will state and prove that D (or a suitable modification of D) is a fundamental domain for Deligne–Mostow lattices with 2-fold symmetry as parametrised in Section 4A. To do this we will use the Poincaré polyhedron theorem, in the form given in [Parker 2006; Boadi and Parker 2015; Pasquinelli 2016]. It states that if one had a polyhedron D and a set of side pairing maps $\{T_i\}$ satisfying certain conditions, then D is a fundamental domain for the action of $\Gamma = \langle T_i \rangle$. The main condition to check in this case is that suitable images of the polyhedron under the side pairing maps tessellate around the ridges. The theorem also provides a presentation for the group, with the side pairings as generators and relations coming from the tessellation conditions.

5A. ***Main theorem.*** We can now state that D just defined or a suitable modification of it is a fundamental domain for the lattices we are considering.

Theorem 5.1. *Let Γ be one of the Deligne–Mostow lattices with 2-fold symmetry* (see Table 8). *Then a suitable modification of the polyhedron D defined in Section 4B is a fundamental domain for Γ. More precisely the fundamental domain is D up to collapsing some ridges to a point when some parameters are degenerate* (*negative of infinite*) *according to Table 11.*

Moreover, a presentation for Γ is given by

$$\Gamma = \left\langle \begin{matrix} K,\, Q,\, B_1, \\ B_2,\, R_0,\, A_0 \end{matrix} \;:\; \begin{matrix} B_1^p = B_2^p = R_0^{p'} = A_0^{k'} = (Q^{-1}K)^k = (R_0K)^l = I, \\ (A_0B_2B_1)^{l'} = Q^{2d} = I, \quad Q = B_1R_0 = R_0B_2 = B_2^{-1}QB_1, \\ R_0^{-1}A_0R_0 = A_0 = K^{-2}, \quad B_2K = KB_1 \end{matrix} \right\rangle,$$

where each of the relations involving k', l, l' and d hold as long as the corresponding parameter is finite and positive.

The reason for the ridges to collapse to a point (except for k', which is treated in Section 5E) relies on the combinatorial structure of the polyhedron as explained in Section 3D. More precisely:

- First consider the case when $d < 0$ or $d = \infty$. By definition (see (19)), this is equivalent to say that $\pi - \alpha - \theta \leq 0$. Remembering Proposition 3.5 and using the notation of Remark 4.1, one can see that the vertices on L_{*0} collapse when $\pi - \alpha - \theta \leq 0$. Since these three vertices form the ridge $F(Q, Q^{-1})$, this ridge collapses when $d < 0$ or $d = \infty$.
- Similarly, when $l < 0$ or $l = \infty$, by definition, we have $\alpha - \theta - \phi \leq 0$. Now the vertices on L_{*3} collapse when $\alpha - \theta - \phi \leq 0$ and so do the ones on L_{*1}. Since $F(K^{-1}, R_0)$ is formed of the vertices contained in L_{*3} and $F(K, R_0^{-1})$ of the ones contained in L_{*1}, they degenerate when $l < 0$ or $l = \infty$.

lattice	deg. par.	ridges collapsing
(4,4,6), (4,4,5)		
(3,4,4), (2,4,3), (3,3,4)	l', d	$F(A_0, B_2^{-1})$, $F(B_2, B_1^{-1})$, $F(A_0^{-1}, B_1)$, $F(Q, Q^{-1})$
(2,6,6)	l, d	$F(Q, Q^{-1})$, $F(K, R_0^{-1})$, $F(K^{-1}, R_0)$
(2,3,3)	l, l', d	$F(Q, Q^{-1})$, $F(K, R_0^{-1})$, $F(K^{-1}, R_0)$, $F(A_0, B_2^{-1})$, $F(B_2, B_1^{-1})$, $F(A_0^{-1}, B_1)$
(3,6,3), (4,4,3), (6,6,3)	k', l', d	$F(A_0, A_0^{-1})$, $F(Q, Q^{-1})$, $F(B_2, B_1^{-1})$, $F(A_0^{-1}, B_1)$
(2,3,3)	k'	$F(A_0, A_0^{-1})$

Table 11. Table for Theorem 5.1.

- Now assume $l' < 0$ or $l' = \infty$, i.e., $\pi - \alpha - \phi \leq 0$. By Proposition 3.5, the vertices on $L_{*1}(\pi + \theta - \alpha, 2\alpha - \pi, \pi + \theta + \phi - 2\alpha)$, $L_{*2}(\alpha, \pi + \theta - \alpha, \theta, \phi)$, and $L_{*3}(\alpha, \alpha, \theta, \phi)$ all degenerate when $\pi - \alpha - \phi \leq 0$. Then the claim of the theorem follows from the fact that $F(B_1, A_0^{-1})$, $F(B_1^{-1}, B_2)$ and $F(B_2^{-1}, A_0)$ are bounded by the vertices contained in $L_{*3}(\alpha, \alpha, \theta, \phi)$, $L_{*2}(\alpha, \pi+\theta-\alpha, \theta, \phi)$ and $L_{*1}(\pi + \theta - \alpha, 2\alpha - \pi, \pi + \theta + \phi - 2\alpha)$ respectively.
- Finally, the case of k' negative is treated in Section 5E.

An alternative presentation for the lattices can be obtained by remembering that $K = QA_1$ and substituting $Q = B_1 R_0$, $K = B_1 R_0 A_1$, $B_2 = R_0^{-1} B_1 R_0$ and $A_0 = (B_1 R_0 A_1)^{-2}$. Then

$$\Gamma = \left\langle B_1, R_0, A_1 \;:\; \begin{array}{l} B_1^p = R_0^{p'} = (B_1 R_0 A_1)^{2k'} = A_1^k = (R_0 B_1 R_0 A_1)^l = I, \\ (A_1 R_0)^{2l'} = (B_1 R_0)^{2d} = I, \quad \mathrm{br}_4(B_1, R_0), \\ \mathrm{br}_2((B_1 R_0 A_1)^{-2}, R_0), \quad \mathrm{br}_2(A_1, B_1) \end{array} \right\rangle,$$

where, following [Deraux et al. 2016], $\mathrm{br}_i(T, S)$ is the braid relation of length i on T and S, i.e., $(TS)^{n/2} = (ST)^{n/2}$, where the power $n/2$ with n odd means that the product has n factor (e.g., $(TS)^{3/2} = TST$).

5B. ***Volume.*** The volume of the quotient is a multiple of the orbifold Euler characteristic $\chi(\boldsymbol{H}^2_{\mathbb{C}}/\Gamma)$. This multiple is $8\pi^2/3$ when the holomorphic sectional curvature is normalised to -1 (see, for example, Section 8 of [McMullen 2017]). The orbifold Euler characteristic is calculated by taking the alternating sum of the reciprocal of the order or the stabilisers of each orbit of cell.

In Table 12 we list the orbits of facets by dimension, calculate the stabiliser of the first element in the orbit and give its order. Later, we will explain how the table changes when considering the degenerations of D.

orbit of the facet	stabiliser	order
$\boldsymbol{v}_1, \boldsymbol{v}_2$	$\langle A_1, B_1\rangle$	kp
$\boldsymbol{v}_3, \boldsymbol{v}_4$	$\langle Q^2, B_1\rangle$	pd
$\boldsymbol{v}_{16}, \boldsymbol{v}_5$	$\langle Q^2, R_0\rangle$	$p'd$
$\boldsymbol{v}_6, \boldsymbol{v}_{10}$	$\langle R_0K, B_1\rangle$	pl
$\boldsymbol{v}_7, \boldsymbol{v}_{11}$	$\langle R_0K, A_0\rangle$	$k'l$
$\boldsymbol{v}_8, \boldsymbol{v}_9, \boldsymbol{v}_{17}, \boldsymbol{v}_{24}$	$\langle QK^{-1}, R_0K\rangle$	kl
$\boldsymbol{v}_{18}, \boldsymbol{v}_{14}, \boldsymbol{v}_{20}, \boldsymbol{v}_{22}, \boldsymbol{v}_{23}, \boldsymbol{v}_{12}$	$\langle A_0B_2B_1, A_1\rangle$	$l'k$
$\boldsymbol{v}_{19}, \boldsymbol{v}_{13}, z_{21}$	$\langle A_0B_2B_1, R_0\rangle$	$p'l'$
$\boldsymbol{v}_0$	$\langle R_0, A_0\rangle$	$k'p'$
$\gamma_{1,3}, \gamma_{2,4}$	$\langle B_1\rangle$	p
$\gamma_{1,6}, \gamma_{2,10}$	$\langle B_1\rangle$	p
$\gamma_{1,12}, \gamma_{2,23}, \gamma_{2,14}, \gamma_{1,18}$	$\langle A_1\rangle$	k
$\gamma_{3,5}, \gamma_{4,16}, \gamma_{4,5}, \gamma_{3,16}$	$\langle Q^2\rangle$	d
$\gamma_{3,6}, \gamma_{4,10}$	$\langle B_1\rangle$	p
$\gamma_{5,13}, \gamma_{16,19}, \gamma_{16,21}$	$\langle R_0\rangle$	p'
$\gamma_{6,8}, \gamma_{10,24}, \gamma_{9,10}, \gamma_{6,17}$	$\langle R_0K\rangle$	l
$\gamma_{7,8}, \gamma_{11,24}, \gamma_{9,11}, \gamma_{7,17}$	$\langle R_0K\rangle$	l
$\gamma_{7,11}$	$\langle K\rangle$	$2k'$
$\gamma_{7,15}, \gamma_{11,15}$	$\langle A_0\rangle$	k'
$\gamma_{8,14}, \gamma_{22,24}, \gamma_{17,20}, \gamma_{9,18}, \gamma_{23,8}, \gamma_{9,12}$	$\langle A_1\rangle$	k
$\gamma_{12,13}, \gamma_{21,22}, \gamma_{18,19}, \gamma_{21,23}, \gamma_{19,20}, \gamma_{13,14}$	$\langle B_1A_0B_2\rangle$	l'
$\gamma_{12,14}, \gamma_{22,23}, \gamma_{18,20}$	$\langle B_2^{-1}K\rangle$	$2l'$
$\gamma_{15,19}, \gamma_{15,21}$	$\langle R_0\rangle$	p'
$F(K,Q),\ F(K^{-1},Q^{-1})$	$\langle A_1\rangle$	k
$F(K^{-1},R_0),\ F(K,R_0^{-1})$	$\langle KR_0\rangle$	l
$F(R_0,R_0^{-1})$	$\langle R_0\rangle$	p'
$F(Q,Q^{-1})$	$\langle Q\rangle$	$2d$
$F(B_1,A_0^{-1}),\ F(B_1^{-1},B_2),\ F(B_2^{-1},A_0)$	$\langle B_1A_0B_2\rangle$	l'
$F(B_1,B_1^{-1})$	$\langle B_1\rangle$	p
$F(B_2,B_2^{-1})$	$\langle B_2\rangle$	p
$F(A_0,A_0^{-1})$	$\langle A_1'\rangle$	k'

Table 12. The stabilisers when all values are positive and finite (continued on the next page).

$F(K, B_1)$, $F(K, B_1^{-1})$, $F(K^{-1}, B_2^{-1})$, $F(K^{-1}, B_2)$	1	1
$F(B_1, Q)$, $F(B_2, Q^{-1})$, $F(B_2^{-1}, Q^{-1})$, $F(B_1^{-1}, Q)$	1	1
$F(A_0, R_0)$, $F(A_0^{-1}, R_0)$, $F(A_0^{-1}, R_0^{-1})$, $F(A_0, R_0^{-1})$	1	1
$F(K, K^{-1})$, $F(K^{-1}, A_0)$, $F(K, A_0^{-1})$	1	1
$F(B_1, R_0^{-1})$, $F(B_1^{-1}, Q^{-1})$, $F(Q, R_0)$	1	1
$F(R_0^{-1}, Q^{-1})$, $F(Q, B_2)$, $F(B_2^{-1}, R_0)$	1	1
$S(K)$, $S(K^{-1})$	1	1
$S(Q)$, $S(Q^{-1})$	1	1
$S(B_2)$, $S(B_2^{-1})$	1	1
$S(B_1)$, $S(B_1^{-1})$	1	1
$S(R_0)$, $S(R_0^{-1})$	1	1
$S(A_0)$, $S(A_0^{-1})$	1	1
D	1	1

Table 12. (continued).

Then the orbifold Euler characteristic of D is given by

$$(29)\quad \chi(\boldsymbol{H}_{\mathbb{C}}^2/\Gamma) = \frac{1}{kp}+\frac{1}{pd}+\frac{1}{dp'}+\frac{1}{pl}+\frac{1}{k'l}+\frac{1}{kl}+\frac{1}{l'k}+\frac{1}{p'l'}+\frac{1}{k'p'} -\frac{1}{p}-\frac{1}{p}-\frac{1}{k}-\frac{1}{d}-\frac{1}{p}-\frac{1}{p'}-\frac{1}{l}-\frac{1}{l}-\frac{1}{2k'}-\frac{1}{k'} -\frac{1}{k}-\frac{1}{l'}-\frac{1}{2l'}-\frac{1}{p'}+\frac{1}{k}+\frac{1}{l}+\frac{1}{p'}+\frac{1}{2d}+\frac{1}{l'}+\frac{1}{p} +\frac{1}{p}+\frac{1}{k'}+1+1+1+1+1+1-1-1-1-1-1-1+1$$

and the volume is $(8\pi^2/3)\chi(\boldsymbol{H}_{\mathbb{C}}^2/\Gamma)$.

While it is easy to see that the stabiliser of each facet contains the group in the second column of the table, the converse inclusion required slightly more work and it follows from the cycles in the Poincaré polyhedron theorem. More specifically, to find the stabiliser, one needs to consider all the cycles of the cycle transformation and keep track of each facet. Then one considers all the transformations inside cycles that stabilise the facet and finds the map or maps that generate all of them. Since the cycles are composed of the side pairings, which are generators for the group, then this is the stabiliser required. Using this procedure, one can find the stabilisers in the second column of Table 12.

Now we need to explain how to modify the table when calculating the orbifold Euler characteristic for one of the degenerations of D.

• First consider the case when d is negative or infinite. Then the vertices $\boldsymbol{v}_3$, $\boldsymbol{v}_4$, $\boldsymbol{v}_5$ and $\boldsymbol{v}_{16}$ collapse to a single point. This means that the two orbits containing them will collapse to only one orbit. The new vertex is stabilised by $\langle B_1, R_0\rangle$ and we need to calculate its order. This is similar to the proof of Proposition 2.3 of [Deraux et al. 2016] (adapting the argument to complex reflections with different orders) and to proof of Propositions 4.4, 4.5 and 4.6 in [Parker 2009]. Now, R_0 has eigenvalues $e^{i\theta'}$, 1, 1, while B_1 has eigenvalues $e^{2i\theta}$, 1, 1. In other words, remembering $\theta' = 2\pi/p'$ and $\theta = \pi/p$, R_0 and B_1 have eigenvalues $e^{2i\pi/p'}$, 1, 1 and $e^{2i\pi/p}$, 1, 1 respectively. Now consider B_1R_0. It has eigenvalues 1, $e^{i(\alpha+\theta)}$, $-e^{i(\alpha+\theta)}$, which we can write as 1, $e^{i(\pi/p'+\pi/p+\pi/2)}$, $e^{i(\pi/p'+\pi/p-\pi/2)}$ because $\theta' = 2\alpha - \pi$. In this way the part acting on the sphere orthogonal to the common eigenspace is in SU(2). This means that $\langle R_0, B_1\rangle$ is a central extension of a $(2, p, p')$-triangle group. Remembering that a $(2, a, b)$-triangle group has order $4ab/(2a + 2b - ab)$ and the definition of the parameters (19), the order of the triangle group is $-2d$. Since $\pi - \alpha - \theta = \pi/d$, the eigenvalues of $(R_0A_1)^2$ are $e^{2\pi/d}$, $e^{2\pi/d}$, 1 and hence the order of the centre is $-d$. So the order of the stabiliser is $2d^2$. Moreover, the line of the table corresponding to the edges between these three points (so the line of the orbit of $\gamma_{3,5}$) needs to be eliminated and so does the line corresponding to the ridge $F(Q, Q^{-1})$. When d is ∞, the single point is on the boundary and the stabiliser has infinite order and the same lines of the table disappear.

• Now consider the case of l' negative or infinite. We have three triples of vertices collapsing to the three vertices $\boldsymbol{v}_{12,13,14}$, $\boldsymbol{v}_{18,19,20}$ and $\boldsymbol{v}_{21,22,23}$, where $\boldsymbol{v}_{i,j,k}$ is obtained collapsing vertices $\boldsymbol{v}_i$, $\boldsymbol{v}_j$ and $\boldsymbol{v}_k$. They are in a unique orbit and $\boldsymbol{v}_{18,19,20}$ is stabilised by $\langle R_0, Q^{-1}K\rangle = \langle R_0, A_1\rangle$. We need to calculate its order. Now, R_0 has eigenvalues $e^{i\theta'}$, 1, 1, while A_1 has eigenvalues $e^{2i\phi}$, 1, 1. In other words, remembering $\theta' = 2\pi/p'$ and $\phi = \pi/k$, R_0 and A_1 have eigenvalues $e^{2i\pi/p'}$, 1, 1 and $e^{2i\pi/k}$, 1, 1 respectively. Now consider R_0A_1. It has eigenvalues 1, $e^{i(\alpha+\phi)}$, $-e^{i(\alpha+\phi)}$, which we can write as 1, $e^{i(\pi/p'+\pi/k+\pi/2)}$, $e^{i(\pi/p'+\pi/k-\pi/2)}$. This means that $\langle R_0, A_1\rangle$ is a central extension of a $(2, p', k)$-triangle group, which has order $4p'k/(2p' + 2k - p'k) = -2l'$. Since $\alpha + \phi - \pi = \pi/l'$, the eigenvalues of $(R_0A_1)^2$ are $e^{2\pi/l'}$, $e^{2\pi/l'}$, 1 and hence the order of the centre is $-l'$. This means that the order of $\langle R_0, A_1\rangle$ is $2l'^2$. Moreover, the two lines of the table corresponding to edges between the three collapsing points need to be eliminated. In other words, the lines of the orbits of $\gamma_{12,13}$ and $\gamma_{12,14}$ disappear from the table, together with the orbit of the three ridges that collapse. Again, when l' is ∞, the vertices $\boldsymbol{v}_{12,13,14}$, $\boldsymbol{v}_{18,19,20}$ and $\boldsymbol{v}_{21,22,23}$ are on the boundary, the stabilisers have infinite order and the same lines of the table disappear.

• Now let us consider the parameter l. From Table 9 one can see that it is never negative. The only degeneration hence comes when it is infinite. This means that the two vertices obtained by triples collapsing are on the boundary and hence their stabiliser will have infinite order. So the orbit of these two vertices disappears in the calculation of the orbifold Euler characteristic. Similarly, the two orbits of edges between collapsing vertices disappear from the calculation (the orbits of $\gamma_{6,8}$ and $\gamma_{7,8}$) and so does the orbit containing the two ridges that collapse to the two new points on the boundary.

• When k' is negative, the vertices $\boldsymbol{v}_0$, $\boldsymbol{v}_7$ and $\boldsymbol{v}_{11}$ collapse to a point (see Section 5E). This means that the two orbits of these three points collapse to a single one. It is easy to see that the new point is stabilised by K, A_0 and R_0, so the stabiliser is $\langle R_0, K\rangle$. We now need to calculate the order of this group. Since $K^2 = A_0^{-1}$ and A_0 commutes with R_0, the centre is generated by K^2, which has order $-k'$. Now, we know that R_0K has order l, so $\langle R_0, K\rangle$ modulo the centre would is a $(2, p', l)$-triangle group, which has order $-2k'$. So the order of $\langle R_0, K\rangle$ is $2k'^2$. Moreover, the lines of the table corresponding to the two orbits of edges between these three points (i.e., the orbit of $\gamma_{7,11}$ and $\gamma_{7,0}$) disappear in the calculation and so does the line relative to $F(A_0, A_0^{-1})$. When k' is ∞, the vertices $\boldsymbol{v}_0$, $\boldsymbol{v}_7$ and $\boldsymbol{v}_{11}$ collapse to a point on the boundary, the stabiliser has infinite order and the same lines of the table disappear.

Remark 5.2. We remark that the calculation of the Euler orbifold characteristic is done for lattices with 2-fold symmetry but forgetting that some of them have 2-2-fold symmetry. These are the lattices in the first class of Table 8. In other words, we are calculating the volume of $\Gamma_{\mu\Sigma_1}$, with $\Sigma_1 = \langle(3, 4)\rangle \cong \mathbb{Z}_2$, rather than $\Gamma_{\mu\Sigma_2}$, with $\Sigma_2 = \langle(1, 2), (3, 4)\rangle \cong \mathbb{Z}_2 \times \mathbb{Z}_2$, which is the full symmetry group of the ball 5-tuple. When we have the extra symmetry, our polyhedron will contain two copies of a fundamental domain for the lattices. The Euler orbifold characteristic of the fundamental domain for $\Gamma_{\mu\Sigma}$, with Σ being the full symmetry group of the ball 5-tuple as usual, will hence be half the one found with the formulae.

The orbifold Euler characteristic χ calculated with the modification of Table 12 is consistent with the commensurability theorems we know between Deligne–Mostow lattices in PU(2, 1). Table 13 summarises the values found in relation with the ones previously known. The top left quadrant contains lattices that are commensurable according to Corollary 3.9 in [Parker 2009], which is Corollary 10.18 of [Deligne and Mostow 1993] and have index 3, because the theorem from which one deduces the corollary does not take into account the 3-fold symmetry. The top right quadrant correspond to commensurability stated in Theorem 3.10 in [Parker 2009] and they also have index 3. The bottom right quadrant of the table contains lattices which are commensurable according to Theorem 3.8 in [Parker 2009], which is Theorem 10.6 in [Deligne and Mostow 1993]. They have index 2 even though the theorem says they are isomorphic because here we only take into account the 2-fold symmetry

lattice	χ		χ	lattice
(6,6,3)	$\frac{1}{3\cdot 2^2}$	$\leftrightarrow$	$\frac{1}{3^2\cdot 2^3}$	(6,2)
(10,10,5)	$\frac{3}{2^3\cdot 5}$	$\leftrightarrow$	$\frac{1}{2^3\cdot 5}$	(10,2)
(12,12,6)	$\frac{7}{2^5\cdot 3}$	$\leftrightarrow$	$\frac{7}{2^5\cdot 3^2}$	(12,2)
(18,18,9)	$\frac{13}{2^3\cdot 3^3}$	$\leftrightarrow$	$\frac{13}{2^3\cdot 3^4}$	(18,2)
(3,4,4)	$\frac{17}{3\cdot 2^5}$	$\leftrightarrow$	$\frac{17}{2^5}$	$\mathcal{T}(4, E_2)$

lattice	χ		χ	lattice
(4,4,3)	$\frac{1}{3\cdot 2^3}$	$\leftrightarrow$	$\frac{1}{3^2\cdot 2^3}$	(4,3)
(4,4,5)	$\frac{11\cdot 3^2}{2^5\cdot 5^2}$	$\leftrightarrow$	$\frac{11\cdot 3}{2^5\cdot 5^2}$	(4,5)
(4,4,6)	$\frac{13}{2^5\cdot 3}$	$\leftrightarrow$	$\frac{13}{2^5\cdot 3^2}$	(4,6)
(2,6,6)	$\frac{1}{2^3}$	$\leftrightarrow$	$\frac{1}{3\cdot 2^2}$	(6,6)
(2,3,3)	$\frac{1}{3\cdot 2^3}$	$\leftrightarrow$	$\frac{1}{3\cdot 2^2}$	(3,6,3)
(3,3,4)	$\frac{7}{2^4\cdot 3}$	$\leftrightarrow$	$\frac{7}{2^5\cdot 3}$	(2,4,3)

Table 13. Values of χ found in relation with the ones previously known.

and ignore the 2-2-fold symmetry (see Remark 5.2). An exception is given by the first one in the list, where the extra term in the index is given by the fact that the theorem does not consider the 3-fold symmetry of the lattice. Finally, in the bottom left quadrant, Proposition 7.10 of [Deraux et al. 2016] shows that the Thompson group E_2 when $p=4$ is (conjugate to) a subgroup of index 3 in the Deligne–Mostow group where $\mu = (3, 3, 5, 6, 7)/12$, which is exactly our $(3, 4, 4)$.

5C. ***Cycles.*** The cycles given by the Poincaré polyhedron theorem are:

$$F(K, Q) \xrightarrow{K} F(K^{-1}, Q^{-1}) \xrightarrow{Q^{-1}} F(K, Q),$$
$$F(K^{-1}, R_0) \xrightarrow{R_0} F(K, R_0^{-1}) \xrightarrow{K} F(K^{-1}, R_0),$$
$$F(B_1, A_0^{-1}) \xrightarrow{B_1} F(B_1^{-1}, B_2) \xrightarrow{B_2} F(B_2^{-1}, A_0) \xrightarrow{A_0} F(B_1, A_0^{-1}),$$
$$F(R_0, R_0^{-1}) \xrightarrow{R_0} F(R_0, R_0^{-1}),$$
$$F(Q, Q^{-1}) \xrightarrow{Q} F(Q, Q^{-1}),$$
$$F(B_1, B_1^{-1}) \xrightarrow{B_1} F(B_1, B_1^{-1}),$$
$$F(B_2, B_2^{-1}) \xrightarrow{B_2} F(B_2, B_2^{-1}),$$
$$F(A_0, A_0^{-1}) \xrightarrow{A_0} F(A_0, A_0^{-1}),$$
$$F(K, B_1) \xrightarrow{B_1} F(K, B_1^{-1}) \xrightarrow{K} F(K^{-1}, B_2^{-1}) \xrightarrow{B_2^{-1}} F(K^{-1}, B_2) \xrightarrow{K^{-1}} F(K, B_1),$$
$$F(B_1, Q) \xrightarrow{Q} F(B_2, Q^{-1}) \xrightarrow{B_2} F(B_2^{-1}, Q^{-1}) \xrightarrow{Q^{-1}} F(B_1^{-1}, Q) \xrightarrow{B_1^{-1}} F(B_1, Q),$$

$$F(A_0, R_0) \xrightarrow{A_0} F(A_0^{-1}, R_0) \xrightarrow{R_0} F(A_0^{-1}, R_0^{-1}) \xrightarrow{A_0^{-1}} F(A_0, R_0^{-1}) \xrightarrow{R_0^{-1}} F(A_0, R_0),$$

$$F(K, K^{-1}) \xrightarrow{K} F(K^{-1}, A_0) \xrightarrow{A_0} F(K, A_0^{-1}) \xrightarrow{K} F(K, K^{-1}),$$

$$F(B_1, R_0^{-1}) \xrightarrow{B_1} F(B_1^{-1}, Q^{-1}) \xrightarrow{Q^{-1}} F(Q, R_0) \xrightarrow{R_0} F(B_1, R_0^{-1}),$$

$$F(R_0^{-1}, Q^{-1}) \xrightarrow{Q^{-1}} F(Q, B_2) \xrightarrow{B_2} F(B_2^{-1}, R_0) \xrightarrow{R_0} F(R_0^{-1}, Q^{-1}).$$

The cycles give the transformations in Table 14, where ℓ determines the power of T which fixes the ridge pointwise and ℓm is the order of T. Note that for all of the 2-fold symmetry values that we are considering, k, k', p, p', l, l' and d are all integers (positive or negative).

When the order of a cycle transformation is negative, the corresponding ridge collapses to a point and so the transformation is a complex reflection to a point. When the order is ∞, the cycle transformation is parabolic.

5D. ***Disjointness and tessellation.*** The proof of Theorem 5.1 consists in proving that D and our side pairings satisfy the hypothesis of the Poincaré polyhedron theorem. This is done in the same way as in [Parker 2006; Boadi and Parker 2015; Pasquinelli 2016]. The only conditions that are not obvious in our case are the disjointness of D and its images under the side pairings and the tessellation condition. We will include some proofs of the disjointness and tessellation conditions, since they are the hardest to prove. We will divide the ridges in three groups. Looking at the structure of sides in Figure 5, one can see that the ridges are contained in either a Giraud disc, a Lagrangian plane or a complex line. We will include the proofs for one ridge from each type.

Ridges contained in a Giraud disc. The ridges contained in a Giraud disc are $F(K, K^{-1})$, $F(K, A_0^{-1})$, $F(A_0, K^{-1})$, $F(B_1, R_0^{-1})$, $F(B_1^{-1}, Q^{-1})$, $F(Q, R_0)$, $F(R_0^{-1}, Q^{-1})$, $F(Q, B_2)$ and $F(B_2^{-1}, R_0)$. To prove the tessellation condition for them, we will use Lemma 4.2. The proof follows proofs of Propositions 4.5 and 4.7

cycle transformation T	ℓ	m	cycle transformation T	ℓ	m
$Q^{-1}K$	1	k	A_0	1	k'
R_0	1	p'	$B_1 A_0 B_2 = (B_2^{-1}K)^2$	1	l'
B_2	1	p	B_1	1	p
Q	2	d	$R_0 K$	1	l
$R_0 Q^{-1} B_1 = \mathrm{Id}$	1	1	$B_2 Q^{-1} R_0 = \mathrm{Id}$	1	1
$B_1 K^{-1} B_2^{-1} K = \mathrm{Id}$	1	1	$B_1^{-1} Q^{-1} B_2 Q = \mathrm{Id}$	1	1
$A_0 R_0^{-1} A_0^{-1} R_0 = \mathrm{Id}$	1	1	$K A_0 K = \mathrm{Id}$	1	1

Table 14. The cycle transformations and their orders.

of [Parker 2006], Proposition 5.3 (first part of the proof) in [Boadi and Parker 2015] and Proposition 8.7 of [Pasquinelli 2016].

Proposition 5.3. *The polyhedra* D, $K(D)$ *and* $KA_0(D) = K^{-1}(D)$ *are disjoint and tessellate around the ridge* $F(K, K^{-1})$.

Proof. Take $z \in D$. By the second point of Lemma 4.2, z is closer to $\boldsymbol{n}_{*0}$ than to $K^{-1}(\boldsymbol{n}_{*0})$. By the tenth point of the lemma, it is closer to $\boldsymbol{n}_{*0}$ than to $K(\boldsymbol{n}_{*0})$. Similarly, take a point $z \in K(D)$. This means that $K^{-1}(z) \in D$. By the second point of the lemma applied to $K^{-1}(z)$, z is closer to $K(\boldsymbol{n}_{*0})$ than to $\boldsymbol{n}_{*0}$. By the eighth point of the lemma, it is closer to $K(\boldsymbol{n}_{*0})$ than to $K^{-1}(\boldsymbol{n}_{*0})$. Finally, take a point $z \in K^{-1}(D)$. This means that $K(z) \in D$. By the fifth point of the lemma applied to $K(z)$, z is closer to $K^{-1}(\boldsymbol{n}_{*0})$ than to $K(\boldsymbol{n}_{*0})$. By the tenth point of the lemma, it is closer to $K^{-1}(\boldsymbol{n}_{*0})$ than to $\boldsymbol{n}_{*0}$.

This clearly implies that the three images are disjoint and since $F(K, K^{-1})$ is defined by $\mathrm{Im}(e^{i\phi} z_1) = 0$ and $\mathrm{Im}(e^{-i\phi} w_1) = 0$, a small enough neighbourhood of the ridge is covered by the three images. □

Ridges contained in a Lagrangian plane. The ones contained in a Lagrangian plane are ridges $F(K, B_1)$, $F(K^{-1}, B_2^{-1})$, $F(K^{-1}, B_2)$, $F(K, B_1^{-1})$, $F(B_1, Q)$, $F(B_2, Q^{-1})$, $F(B_2^{-1}, Q^{-1})$, $F(B_1^{-1}, Q)$, $F(A_0^{-1} R_0)$, $F(A_0^{-1}, R_0^{-1})$, $F(A_0, R_0^{-1})$ and $F(A_0, R_0)$. The proof is done by studying the sign of some of the coordinates and it follows proofs of Proposition 4.8 of [Parker 2006], Proposition 5.3 (end of the proof) of [Boadi and Parker 2015] and Proposition 8.8 of [Pasquinelli 2016]. We will prove the property for the first ridge mentioned. The others are done in a similar way.

Proposition 5.4. *The polyhedra* D, $B_1^{-1}(D)$, $K^{-1}(D)$ *and* $B_1^{-1}K^{-1}(D)$ *are disjoint and tessellate around the ridge* $F(K, B_1)$.

Proof. Let us consider points in D, $B_1^{-1}(D)$, $K^{-1}(D)$ and $B_1^{-1}K^{-1}(D)$ and record the sign of the values of $\mathrm{Im}(z_1)$, $\mathrm{Im}(e^{i\phi} z_1)$, $\mathrm{Im}(e^{i\theta} z_2)$ and $\mathrm{Im}(e^{-i\theta} z_2)$ for them. They are shown in the following table.

	$\mathrm{Im}(z_1)$	$\mathrm{Im}(e^{i\phi} z_1)$	$\mathrm{Im}(e^{i\theta} z_2)$	$\mathrm{Im}(e^{-i\theta} z_2)$
D	-	+	+	-
$B_1^{-1}(D)$	-	+	-	-
$K^{-1}(D)$	-	-	+	-
$B_1^{-1}K^{-1}(D)$	-	-	-	-

The first row can be deduced using the definition of D in terms of the arguments of the coordinates. The second row can be deduced by considering that the action of B_1 only consists in multiplying the coordinate z_2 by $e^{2i\theta}$. The third row can be deduced by the fact that applying K corresponds to first applying A_1, which multiplies the coordinate z_1 by $e^{2i\phi}$ and then applying Q, which relates the $\boldsymbol{z}$ coordinates to the $\boldsymbol{w}$ coordinates.

The ridge $F(K, B_1)$ is defined by $\mathrm{Im}(e^{i\phi}z_1) = 0$ and $\mathrm{Im}(e^{i\theta}z_2) = 0$ and in a neighbourhood of the ridge the images considered coincide with the sectors where the values are either positive or negative. Combining the information of the table one gets the tessellation as required. Moreover, since each pair of polyhedra has at least one value with opposite sign, they will always be separated by the subspace where that value is zero and hence they will be pairwise disjoint. □

Ridges contained in complex lines. The ridges contained in complex lines are $F(K, Q)$, $F(K^{-1}, Q^{-1})$, $F(K, R_0^{-1})$, $F(K^{-1}, R_0)$, $F(R_0, R_0)$, $F(Q, Q^{-1})$, $F(B_2, A_0^{-1})$, $F(B_1^{-1}, B_2)$, $F(B_1^{-1}, A_0)$, $F(B_1, B_1^{-1})$, $F(B_2, B_2^{-1})$ and $F(A_0, A_0^{-1})$. The strategy consists in showing that the polyhedron (and suitable images) cover a sector of amplitude ψ and that the cycle transformation acts on the orthogonal of the complex line as a rotation through angle ψ. Then each power of the cycle transformation covers a sector and since ψ is always $2\pi/a$ with a integer, we cover the whole space around the ridge. The proofs are similar to the ones of Proposition 4.11 of [Parker 2006], Proposition 5.3 of [Boadi and Parker 2015] (the middle part of the proof) and Proposition 8.10 of [Pasquinelli 2016].

The cases of $F(B_1, A_0^{-1})$, $F(Q, Q^{-1})$ and $F(K^{-1}, R_0)$ (and the ones in their cycles) are an exception because the procedure is the same but after applying a suitable change of coordinates.

The proofs for these cases are along the line of proof of Proposition 4.13 of [Parker 2006] and of Proposition 8.11 of [Pasquinelli 2016]. For completeness, we will include the proof of one of these ridges.

Proposition 5.5. *The polyhedra D, $A_0(D)$ and $A_0B_2(D)$ and their images under $A_0B_2B_1$ are disjoint and tessellate around the ridge $F(B_1, A_0^{-1})$.*

Proof. The ridge $F(B_1, A_0^{-1})$ is contained in $L_{*3}(\alpha, \alpha, \theta, \phi)$, and $e^{-2i(\theta-\alpha)}A_0B_2B_1$ fixes the ridge pointwise and rotates its normal vector $\boldsymbol{n}_{*3}(\alpha, \alpha, \theta, \phi)$ by $e^{2i(\alpha+\phi-\pi)}$.

The proof consists in changing the coordinates to have a similar situation as for the other ridges contained in a complex line. The new coordinates will be in terms of two vectors spanning the complex line and the vector normal to it, since the complex line is the mirror of the transformation $A_0B_2B_1$. More precisely, writing

$$\begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \frac{\sin\phi\sin(\alpha-\theta) - \sin\alpha\sin(\theta+\phi)x_2}{\sin\theta\sin(\alpha+\phi)} \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} + x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \frac{1-x_2}{\sin\theta\sin(\alpha+\phi)} \begin{pmatrix} 0 \\ \sin\phi\sin(\alpha-\theta) \\ \sin(\theta+\phi)\sin\alpha \end{pmatrix}, \tag{30}$$

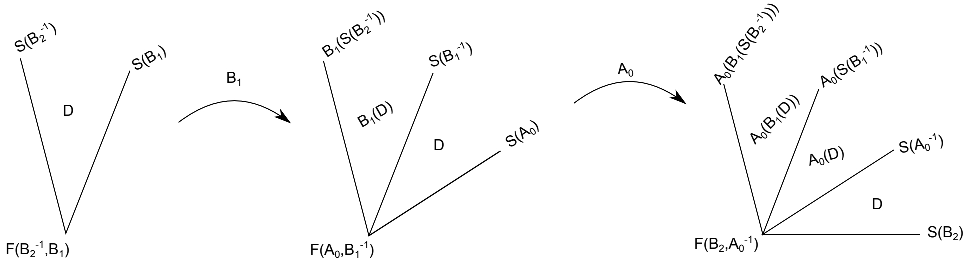


Figure 9. The polyhedra around $F(A_0^{-1}, B_2)$.

the new coordinates are

$$\begin{aligned} \xi_1 &= \frac{\sin\phi\sin(\alpha-\theta) - \sin\alpha\sin(\theta+\phi)x_2}{1-x_2}, \\ \xi_2 &= \frac{\sin\theta\sin(\alpha+\phi)x_1}{1-x_2}. \end{aligned} \tag{31}$$

This means that $A_0B_2B_1$ acts on the new coordinates by sending (ξ_1, ξ_2) to the point $(e^{2i(\alpha+\phi-\pi)}\xi_1, \xi_2)$. Since the configurations are as in Figure 9, if we prove that D, $A_0(D)$ and $A_0B_2(D)$ cover the sector defined by the argument of ξ_1 being between 0 and $2(\alpha+\phi-\pi)$, then the appropriate images under $A_0B_2B_1$ will cover a neighbourhood of $F(A_0^{-1}, B_2)$.

First notice that if we are in $S(B_1)$, then $x_2 \in \mathbb{R}$ and so $\arg\xi_1 = 0$. Moreover, if we take a point in $z \in S(B_1^{-})$, then $z_2 = e^{i\theta}u$ with $u \in \mathbb{R}$ and the coordinate ξ_1 of A_0B_2z is

$$\xi_1 = e^{2i(\alpha+\phi-\pi)}\frac{\sin(\theta+\phi)u + \sin\phi}{\sin(\alpha-\theta)u - \sin\alpha}$$

and so $\arg\xi_1 = 2(\alpha+\phi-\pi)$.

The last thing we need to prove is that such images are disjoint. Now the pairs D, A_0D and A_0D, B_2A_0D are disjoint because of tessellation property around $F(A_0, A_0^{-1})$ and $F(B_2, B_2^{-1})$. To prove that D and B_2A_0D are disjoint, it is enough to prove that the argument of the coordinate ξ_1 of points in D is smaller than $\alpha+\phi-\pi$, while the one of points in B_2A_0D is bigger than $\alpha+\phi-\pi$.

If one writes the coordinate ξ_1 in terms of the $\boldsymbol{v}$-coordinates, then a point in $S(A_0)$ has coordinate $v_1 = e^{i\phi'}u$, with $\mathbb{R} \ni u \le -\sin(2\alpha)/\sin(\theta+\phi)$ by Lemma 3.4 and

$$\xi_1 = e^{i(\alpha+\phi-\pi)}\frac{\sin\phi\sin(\alpha-\theta)(-\sin(2\alpha) - \sin(\theta+\phi)u)}{\sin(\alpha-\theta)u - \sin(\alpha+\phi)v_2 + \sin(\alpha-\phi)}.$$

Then

$$\operatorname{Im} e^{i(\alpha+\phi-\pi)}\xi_1 = \sin\phi\sin(\alpha-\theta)(-\sin(2\alpha) - \sin(\theta+\phi)u)\sin(\alpha+\phi)\operatorname{Im} v_2 \ge 0.$$

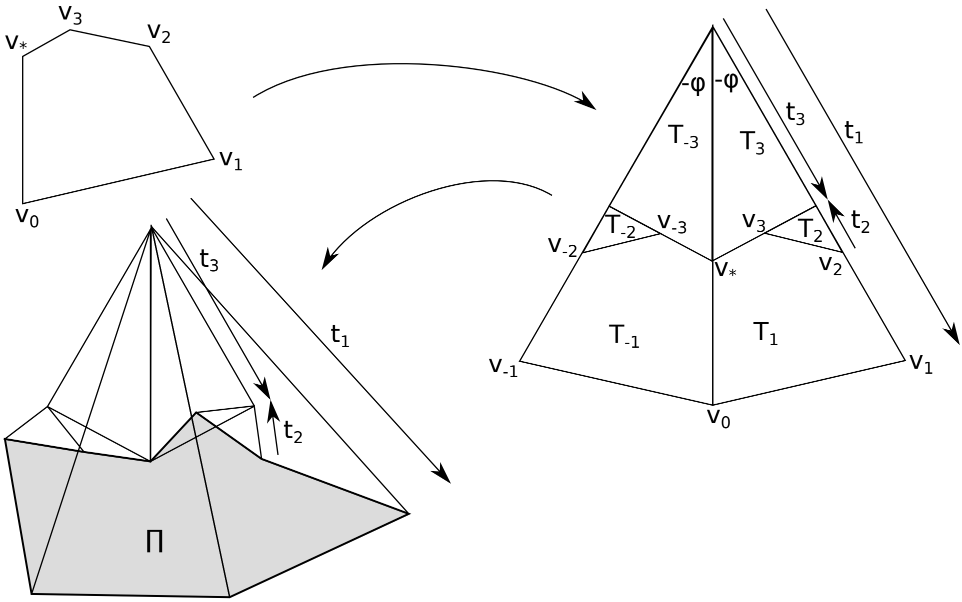


Figure 10. The configuration of triangles when k' is negative.

Similarly, for a point $z \in S(B_2^{-1})$ we have $w_2 = e^{-i\theta}v$ with $\mathbb{R} \ni v \le \sin\phi/\sin(\theta+\phi)$ and the coordinate ξ_1 of $A_0 z$ is

$$\xi_1 = e^{-i(\alpha+\phi-\pi)} \frac{\sin\phi}{\sin\alpha} \cdot \frac{\sin(\theta+\phi)u - \sin\phi}{\sin(\alpha+\phi)e^{-i\phi}w_1 + \sin(\alpha-\theta)u - \sin\alpha}$$

and

$$\operatorname{Im} e^{i(\alpha+\phi-\pi)}\xi_1 = \frac{\sin\phi}{\sin\alpha}(\sin\phi - \sin(\theta+\phi)u)\sin(\alpha+\phi)\operatorname{Im} e^{-i\phi}w_2 \le 0.$$

Observe that we are using the fact that $\sin(\alpha+\phi) > 0$, which is always the case when the ridge does not collapse (i.e., $l' > 0$ and finite). □

5E. ***The case k′ negative.*** When k' is negative, after applying P^{-1} to the configuration $(\alpha, \alpha, \theta, \phi)$ we obtain a configuration where the last angle is negative. This means that we cannot describe the configuration with the same coordinates and triangles as before. This doesn't stop us from doing everything in the same way, up to taking a slightly different configuration of triangles. By construction (see Figure 1), once we developed the cone metric on the plane, ϕ was the angle between the line passing through v_* and v_0 and the line passing through v_1 and v_2 on the side of v_0 and v_1. When this angle is negative, we will take $-\phi$ to be the angle between the same two lines, but on the side of v_2 and v_* (see Figure 10).

The area of the cone metric is the area of the shaded region Π. Using the coordinates as in the figure, this is

$$\text{Area} = \frac{-\sin\phi\sin\alpha}{\sin(\alpha-\phi)}|t_1|^2 - \frac{\sin\theta\sin\beta}{\sin(\beta-\theta)}|t_2|^2 - \frac{-\sin\phi\sin\theta}{\sin(\theta+\phi)}|t_3|^2.$$

Remembering that $-\sin\phi$ is positive, this is still a Hermitian form of the same signature, except that the roles of t_1 and t_3 are exchanged. This makes sense, since now the triangles T_2 and T_3 are "inside" the triangle T_1

When looking at the vertices, this tells us that the we cannot have the line L_{01}, since to make v_0 and v_1 collapse, one should take $x_1 = 0$ and the whole figure would collapse. We will hence have a new vertex $\boldsymbol{v}_{*23}$ obtained by taking $t_1 = t_3 = 0$ and so by making $v_* \equiv v_2 \equiv v_3$ instead of the three vertices $\boldsymbol{y}_1$, $\boldsymbol{y}_9$, $\boldsymbol{y}_{12}$. In terms of our polyhedron D, this means that v_0, v_7 and v_{11} collapse to this new point $\boldsymbol{v}_{*23}$, which is on the boundary (i.e., it makes the area be 0) if k' is infinite. All the other vertices remain the same and everything else in the study of the combinatorial structure of the polyhedron can be done in the exact same way. In particular, as in Proposition 3.5, we still have that the vertices on L_{*0} collapse to a single vertex if $\pi - \alpha' - \theta' \leq 0$ (i.e., if $d \leq 0$) and the vertices on L_{*1} collapse to a single vertex if $\alpha' - \theta' - \phi' \leq 0$ (i.e., if $l' \leq 0$). Note that the vertices on L_{*2} and L_{*3} never collapse, as $l > 0$ in all our cases. This analysis gives the remaining cases in Theorem 5.1.

Acknowledgements

This research was supported by a fellowship under the program "FY2015 JSPS Postoctoral Fellowship (short-term) for North American and European Researchers," while visiting Tokyo Institute of Technology and by a Doctoral EPSRC grant, awarded by Durham University. I would like to thank Professor Parker for his constant inspiration and encouragement. I would also like to thank Professor Kojima and his laboratory for hosting me during my fellowship, for supporting my work and for the useful discussions.

References

[Boadi and Parker 2015] R. K. Boadi and J. R. Parker, "Mostow's lattices and cone metrics on the sphere", *Adv. Geom.* **15**:1 (2015), 27–53. MR Zbl

[Deligne and Mostow 1986] P. Deligne and G. D. Mostow, "Monodromy of hypergeometric functions and nonlattice integral monodromy", *Inst. Hautes Études Sci. Publ. Math.* **63** (1986), 5–89. MR Zbl

[Deligne and Mostow 1993] P. Deligne and G. D. Mostow, *Commensurabilities among lattices in* PU$(1, n)$, Ann. of Math. Studies **132**, Princeton Univ. Press, 1993. MR Zbl

[Deraux et al. 2005] M. Deraux, E. Falbel, and J. Paupert, "New constructions of fundamental polyhedra in complex hyperbolic space", *Acta Math.* **194**:2 (2005), 155–201. MR Zbl

[Deraux et al. 2016] M. Deraux, J. R. Parker, and J. Paupert, "New non-arithmetic complex hyperbolic lattices, II", preprint, 2016. arXiv

[Giraud 1921] G. Giraud, "Sur certaines fonctions automorphes de deux variables", *Ann. Sci. École Norm. Sup.* (3) **38** (1921), 43–164. MR Zbl

[Goldman 1999] W. M. Goldman, *Complex hyperbolic geometry*, Oxford Univ. Press, 1999. MR Zbl

[Kojima 2001] S. Kojima, "Complex hyperbolic cone structures on the configuration spaces", *Rend. Istit. Mat. Univ. Trieste* **32**:Suppl. 1 (2001), 149–163. MR Zbl

[McMullen 2017] C. T. McMullen, "The Gauss–Bonnet theorem for cone manifolds and volumes of moduli spaces", *Amer. J. Math.* **139**:1 (2017), 261–291. MR Zbl

[Mostow 1980] G. D. Mostow, "On a remarkable class of polyhedra in complex hyperbolic space", *Pacific J. Math.* **86**:1 (1980), 171–276. MR Zbl

[Mostow 1986] G. D. Mostow, "Generalized Picard lattices arising from half-integral conditions", *Inst. Hautes Études Sci. Publ. Math.* **63** (1986), 91–106. MR Zbl

[Mostow 1988] G. D. Mostow, "On discontinuous action of monodromy groups on the complex n-ball", *J. Amer. Math. Soc.* **1**:3 (1988), 555–586. MR Zbl

[Parker 2006] J. R. Parker, "Cone metrics on the sphere and Livné's lattices", *Acta Math.* **196**:1 (2006), 1–64. MR Zbl

[Parker 2009] J. R. Parker, "Complex hyperbolic lattices", pp. 1–42 in *Discrete groups and geometric structures* (Kortrijk, Belgium, 2008), edited by K. Dekimpe et al., Contemp. Math. **501**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

[Pasquinelli 2016] I. Pasquinelli, "Deligne–Mostow lattices with three fold symmetry and cone metrics on the sphere", *Conform. Geom. Dyn.* **20** (2016), 235–281. MR Zbl

[Sauter 1990] J. K. Sauter, Jr., "Isomorphisms among monodromy groups and applications to lattices in $\mathrm{PU}(1,2)$", *Pacific J. Math.* **146**:2 (1990), 331–384. MR Zbl

[Thurston 1998] W. P. Thurston, "Shapes of polyhedra and triangulations of the sphere", pp. 511–549 in *The Epstein birthday schrift*, edited by I. Rivin et al., Geom. Topol. Monogr. **1**, Geom. Topol. Publ., Coventry, 1998. MR Zbl

Received September 4, 2017. Revised November 14, 2018.

IRENE PASQUINELLI
INSTITUT DE MATHÉMATIQUES DE JUSSIEU — PARIS RIVE GAUCHE
SORBONNE UNIVERSITÉ
PARIS
FRANCE
irene.pasquinelli@imj-prg.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.249

BINARY QUARTIC FORMS WITH BOUNDED INVARIANTS AND SMALL GALOIS GROUPS

CINDY (SIN YI) TSANG AND STANLEY YAO XIAO

We consider integral and irreducible binary quartic forms whose Galois group is isomorphic to a subgroup of the dihedral group of order eight. We first show that the set of all such forms is a union of families indexed by integral binary quadratic forms $f(x, y)$ of nonzero discriminant. Then, we shall enumerate the $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of all such forms associated to a fixed $f(x, y)$.

1. Introduction

The problem of enumerating $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of integral and irreducible binary forms of a fixed degree has a long history. The quadratic and cubic cases were solved in [Gauss 1801; Siegel 1944] and [Davenport 1951b; 1951c], respectively, where the forms are ordered by the natural height, namely the discriminant $\Delta(-)$. The quartic case turns out to be more challenging. This is because the ring of polynomial invariants of quartic forms have two independent generators, usually denoted $I(-)$ and $J(-)$. For

$$F(x, y) = a_4x^4 + a_3x^3y + a_2x^2y^2 + a_1xy^3 + a_0y^4, \tag{1-1}$$

they are given by the explicit formulae

$$I(F) = 12a_4a_0 - 3a_3a_1 + a_2^2,$$
$$J(F) = 72a_4a_2a_0 + 9a_3a_2a_1 - 27a_4a_1^2 - 27a_3^2a_0 - 2a_2^3,$$

MSC2010: 11E76, 11R45.

Keywords: binary quartic forms, coregular spaces, arithmetic statistics.

which are of degrees two and three, respectively. Bhargava and Shankar [2015], instead of using the discriminant, introduced the height function

$$H_{\mathrm{BS}}(F) = \max\{|I(F)|^3, J(F)^2/4\}. \tag{1-2}$$

For $X > 0$, let us define

$$N_{\mathbb{Z}}(X) = \#\big\{[F] : \text{integral and irreducible binary quartic forms } F \text{ such that } H_{\mathrm{BS}}(F) \le X\big\},$$

where $[-]$ denotes $\mathrm{GL}_2(\mathbb{Z})$-equivalence class. In [loc. cit.], they proved that

$$N_{\mathbb{Z}}(X) = \frac{44\zeta(2)}{135} X^{\frac{5}{6}} + O_\epsilon(X^{\frac{3}{4}+\epsilon}) \quad \text{for any } \epsilon > 0. \tag{1-3}$$

This is the first result ever obtained, and as far as we know, the only known result in the literature, for the quartic case.

1A. ***Set-up and notation.*** In this paper, we shall also be interested in the quartic case, but only the integral and irreducible binary quartic forms F with *small* Galois group $\mathrm{Gal}(F)$, which is defined to be the Galois group of the splitting field of $F(x, 1)$ over $\mathbb{Q}$. We know that $\mathrm{Gal}(F)$ is isomorphic to one of the following:

$$\begin{aligned} S_4 &= \text{the symmetric group on four letters,}\\ A_4 &= \text{the alternating group on four letters,}\\ D_4 &= \text{the dihedral group of order eight,}\\ C_4 &= \text{the cyclic group of order four,}\\ V_4 &= \text{the Klein-four group.} \end{aligned}$$

We shall say that $\mathrm{Gal}(F)$ is *small* if it is isomorphic to D_4, C_4, or V_4. Recall that the *cubic resolvent of F* is defined by

$$\mathcal{Q}_F(x) = x^3 - 3I(F)x + J(F).$$

Then, equivalently, we have the classical characterization that for irreducible F

$$\mathrm{Gal}(F) \text{ is small if and only if } \mathcal{Q}_F(x) \text{ is reducible.}$$

It turns out that whether $\mathrm{Gal}(F)$ is small or not may also be characterized in terms of binary quadratic forms and the following so-called *twisted action* of $\mathrm{GL}_2(\mathbb{R})$.

Given a complex binary form $\xi(x, y)$, let $\mathrm{GL}_2(\mathbb{R})$ act on it via

$$\xi_T(x, y) = \frac{1}{\det(T)^{\deg\xi/2}} \xi(t_1x + t_2y, t_3x + t_4y) \quad \text{for } T = \begin{pmatrix} t_1 & t_2 \\ t_3 & t_4 \end{pmatrix}.$$

Observe that this is only an action up to sign when $\deg\xi$ is odd, in the sense that for $T_1, T_2 \in \mathrm{GL}_2(\mathbb{R})$, we only have $\xi_{T_1T_2} = \pm(\xi_{T_1})_{T_2}$ in general. Now, given a real binary quadratic form $f(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ with $\Delta(f) \neq 0$, write

$$M_f = \begin{pmatrix} \beta & 2\gamma \\ -2\alpha & -\beta \end{pmatrix}$$

for its associated matrix in $\mathrm{GL}_2(\mathbb{R})$. Its action on binary quartic forms clearly remains unchanged if we scale $f(x, y)$ by a constant in $\mathbb{R}^\times$. The second author, Xiao, proved [2019] that for any real binary quartic form F with $\Delta(F) \neq 0$, elements of

$$\{T \in \mathrm{GL}_2(\mathbb{R}) : T \text{ is not a scalar multiple of } I_{2\times 2} \text{ and } F_T = F\}$$

all arise from binary quadratic forms in this way; see Proposition 2.1. Recall that an integral binary quadratic form is called *primitive* if its coefficients are coprime. Using this result from [Xiao 2019], in Section 2, we shall first show:

Theorem 1.1. *Let F be an integral binary quartic form with $\Delta(F) \neq 0$. Then, the following are equivalent.*

(1) *$\mathcal{Q}_F(x)$ is reducible.*

(2) *$F_T = F$ for some $T \in \mathrm{GL}_2(\mathbb{Q})$ which is not a scalar multiple of $I_{2\times 2}$.*

(3) *$F_{M_f} = F$ for an integral and primitive binary quadratic form f with $\Delta(f) \neq 0$.*

Moreover, in the case that $\mathcal{Q}_F(x)$ is reducible:

(a) *If $\Delta(F) \neq \square$, then there is a unique such f up to sign.*

(b) *If $\Delta(F) = \square$, then there are exactly three such f up to sign, among which one is definite and two are indefinite.*

Given a real binary quadratic form $f(x, y)$ with $\Delta(f) \neq 0$, let us further make the following definitions. First put

$$V_{\mathbb{R},f} = \{\text{real binary quartic forms } F \text{ such that } F_{M_f} = F\},$$

$$V_{\mathbb{Z},f} = \{\text{integral binary quartic forms } F \text{ such that } F_{M_f} = F\}.$$

Clearly $V_{\mathbb{R},f}$ is a vector space over $\mathbb{R}$ and $V_{\mathbb{Z},f}$ a lattice over $\mathbb{Z}$. A straightforward calculation shows that $\dim_{\mathbb{R}} V_{\mathbb{R},f}$ is three; see (3-1) and (3-2) below. Also, put

$$V^0_{\mathbb{R},f} = \{F \in V_{\mathbb{R},f} : \Delta(F) \neq 0\} \quad \text{and} \quad V^0_{\mathbb{Z},f} = \{F \in V_{\mathbb{Z},f} : \Delta(F) \neq 0\}.$$

For $F \in V^0_{\mathbb{R},f}$, we shall define two new invariants as follows. As we shall see in (2-3), there is a unique root $\omega_f(F)$ of $\mathcal{Q}_F(x)$ corresponding to f. Let $\omega'_f(F), \omega''_f(F)$ denote the other two roots of $\mathcal{Q}_F(x)$ and define

$$L_f(F) = \omega_f(F) \quad \text{and} \quad K_f(F) = -\omega'_f(F)\omega''_f(F). \tag{1-4}$$

By Proposition 3.2 below, they have degrees one and two, respectively, in the coefficients of F. Following (1-2), let us define the *height of F associated to f* by

$$H_f(F) = \max\{L_f(F)^2, |K_f(F)|\}.$$

This is comparable to the height (1-2) because by comparing coefficients in

$$x^3 - I(F)x + J(F) = (x - \omega_f(F))(x - \omega'_f(F))(x - \omega''_f(F)),$$

we easily deduce the relations

$$3I(F) = L_f(F)^2 + K_f(F) \quad \text{and} \quad J(F) = L_f(F)K_f(F), \tag{1-5}$$

which in turn imply that

$$(H_f(F)/10)^3 \le H_{\mathrm{BS}}(F) \le H_f(F)^3. \tag{1-6}$$

Let us note that

$$\begin{aligned}\Delta(F) &= \frac{4I(F)^3 - J(F)^2}{27} \\ &= \left(\frac{L_f(F)^2 + 4K_f(F)}{9}\right)\left(\frac{2L_f(F)^2 - K_f(F)}{9}\right)^2,\end{aligned} \tag{1-7}$$

where the first equality is well-known, and the second equality holds by (1-5). Also, our height $H_f(-)$ is an invariant in the sense that for any $T \in \mathrm{GL}_2(\mathbb{R})$, we have

$$H_{f_T}(F_T) = H_f(F),$$

as shown in Proposition 3.1 below. This implies that the map

$$V_{\mathbb{R},f} \to V_{\mathbb{R},f_T}, \quad F \mapsto F_T, \tag{1-8}$$

which is a well-defined bijection because $M_{f_T} = T^{-1}M_fT$, is height-preserving when restricted to the forms of nonzero discriminant.

Now, let us return to the integral and irreducible binary quartic forms with small Galois group. Write $V_{\mathbb{Z}}^{\mathrm{sm}}$ for the set of all such forms and set

$$V_{\mathbb{Z}}^{\mathrm{sm},\dagger} = \{F \in V_{\mathbb{Z}}^{\mathrm{sm}} : \mathrm{Gal}(F) \not\simeq V_4\}.$$

By Theorem 1.1, we know that

$$\begin{aligned} V_{\mathbb{Z}}^{\mathrm{sm}} &= \bigcup_{f \in \mathfrak{F}^*} \{F \in V_{\mathbb{Z},f}^0 : F \text{ is irreducible}\}, \\ V_{\mathbb{Z}}^{\mathrm{sm},\dagger} &= \bigsqcup_{f \in \mathfrak{F}^*} \{F \in V_{\mathbb{Z},f}^0 : F \text{ is irreducible and } \mathrm{Gal}(F) \not\simeq V_4\}, \end{aligned} \tag{1-9}$$

where $\mathfrak{F}^*$ denotes the set of all integral and primitive binary quadratic forms of nonzero discriminant, up to sign. In particular, given $F \in V_{\mathbb{Z}}^{\mathrm{sm},\dagger}$, there is a unique $f \in \mathfrak{F}^*$ such that $F \in V_{\mathbb{Z},f}^0$, and we may define the *height of* F by setting

$$H(F) = H_f(F).$$

For $X > 0$, let us define

$$\begin{aligned} N_{\mathbb{Z}}^{\dagger}(X) &= \#\{[F] : F \in V_{\mathbb{Z}}^{\mathrm{sm},\dagger} \text{ such that } H(F) \leq X\}, \\ N_{\mathbb{Z},f}^{\dagger}(X) &= \#\{[F] : F \in V_{\mathbb{Z}}^{\mathrm{sm},\dagger} \cap V_{\mathbb{Z},f}^0 \text{ such that } H(F) \leq X\}. \end{aligned}$$

Then, by (1-8) and (1-9), we have

$$N_{\mathbb{Z}}^{\dagger}(X) = \sum_{f \in \mathfrak{F}} N_{\mathbb{Z},f}^{\dagger}(X),$$

where $\mathfrak{F}$ denotes a set of representatives of the $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes on $\mathfrak{F}^*$. In Theorem 1.2, which is our main result, for $f \in \mathfrak{F}^*$, we shall determine the asymptotic formula for $N_{\mathbb{Z},f}^{\dagger}(X)$. In fact, we shall consider the finer counts

$$\begin{aligned} N_{\mathbb{Z},f}^{(D_4)}(X) &= \#\{[F] : F \in V_{\mathbb{Z}}^{\mathrm{sm}} \cap V_{\mathbb{Z},f}^0 \text{ such that } \mathrm{Gal}(F) \simeq D_4 \text{ and } H(F) \leq X\}, \\ N_{\mathbb{Z},f}^{(C_4)}(X) &= \#\{[F] : F \in V_{\mathbb{Z}}^{\mathrm{sm}} \cap V_{\mathbb{Z},f}^0 \text{ such that } \mathrm{Gal}(F) \simeq C_4 \text{ and } H(F) \leq X\}, \\ N_{\mathbb{Z},f}^{(V_4)}(X) &= \#\{[F] : F \in V_{\mathbb{Z}}^{\mathrm{sm}} \cap V_{\mathbb{Z},f}^0 \text{ such that } \mathrm{Gal}(F) \simeq V_4 \text{ and } H_f(F) \leq X\}, \end{aligned}$$

and show that the latter two are negligible compared to $N_{\mathbb{Z},f}^{(D_4)}(X)$. This means that most of the forms in $V_{\mathbb{Z}}^{\mathrm{sm}} \cap V_{\mathbb{Z},f}^0$ have Galois group isomorphic to D_4. However, all of our error estimates depend upon f. Currently, we do not know how to control them in a uniform way, and so we are unable to obtain an asymptotic formula for $N_{\mathbb{Z}}^{\dagger}(X)$ by summing over $f \in \mathfrak{F}$.

Finally, let us explain, for each $f \in \mathfrak{F}^*$, how counting forms in $V_{\mathbb{Z}}^{\mathrm{sm}} \cap V_{\mathbb{Z},f}^0$ may be reduced to counting lattice points. Write $f(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ with $\alpha, \beta, \gamma \in \mathbb{Z}$. By (3-1) and (3-2), the set $V_{\mathbb{R},f}$ is a vector space isomorphic to $\mathbb{R}^3$ via

$$\begin{aligned} \Theta_1 &: a_4x^4 + a_3x^3y + a_2x^2y^2 + a_1xy^3 + a_0y^4 \mapsto (a_4, a_3, a_2) \quad \text{if } \alpha \neq 0, \\ \Theta_2 &: a_4x^4 + a_3x^3y + a_2x^2y^2 + a_1xy^3 + a_0y^4 \mapsto (a_4, a_2, a_0) \quad \text{if } \beta, \beta^2 + 4\alpha\gamma \neq 0. \end{aligned}$$

Recall that the subset $V_{\mathbb{Z},f}$ has the structure of a rank-three $\mathbb{Z}$-lattice, which may be identified with the lattices

$$\Lambda_{f,1} = \Theta_1(V_{\mathbb{Z},f}) \quad \text{and} \quad \Lambda_{f,2} = \Theta_2(V_{\mathbb{Z},f}) \tag{1-10}$$

in $\mathbb{Z}^3$. Let us mention here that we shall use the isomorphism

$$\Theta_{w(f)}, \quad \text{where } w(f) = \begin{cases} 1 & \text{if } f \text{ is irreducible,} \\ 2 & \text{if } f \text{ is reducible.} \end{cases}$$

Thus, the problem is reduced to counting points in $\Lambda_{f,1}$ or $\Lambda_{f,2}$, and then sieving out those which come from reducible forms. In turn, counting lattice points amounts to computing certain volumes by a result of Davenport [1951a]; see Proposition 5.1.

1B. *Statement of the main theorem.* It is clear that we may choose the set $\mathfrak{F}$ of representatives to be such that for all $f \in \mathfrak{F}$, the x^2-coefficient is positive, and

$$f(x, y) = \alpha x^2 + \beta xy, \quad \text{where } \gcd(\alpha, \beta) = 1 \text{ and } 0 < \alpha \leq \beta \tag{1-11}$$

when f is reducible. Let $\sim$ denote $\mathrm{GL}_2(\mathbb{Z})$-equivalence. Then, our main result is:

Theorem 1.2. *Let $f(x, y)$ be an integral and primitive binary quadratic form of nonzero discriminant and with positive x^2-coefficient. Write $D_f = |\Delta(f)|$, and put*

$$s_f = \begin{cases} 8 & \text{if } D_f \text{ is odd,} \\ 1 & \text{if } D_f \text{ is even.} \end{cases}$$

(a) *Suppose that f is positive definite. Then, we have*

$$N_{\mathbb{Z},f}^{(D_4)}(X) = \frac{1}{s_f r_f} \frac{13\pi}{27 D_f^{3/2}} X^{3/2} + O_f(X^{1+\epsilon}) \quad \text{for any } \epsilon > 0,$$

where

$$r_f = \begin{cases} 6 & \text{if } f(x, y) \sim x^2 + xy + y^2, \\ 2 & \text{if } f(x, y) \sim ax^2 + cy^2 \text{ or } f(x, y) \sim ax^2 + bxy + ay^2 \text{ with } a \neq b, \\ 1 & \text{otherwise.} \end{cases}$$

(b) *Suppose that f is reducible and that f has the shape (1-11). Then, we have*

$$N_{\mathbb{Z},f}^{(D_4)}(X) = \frac{1}{s_f r_f} \frac{8}{9\beta^{3/2}} X^{3/2} \log X + O_f(X^{3/2}),$$

where

$$r_f = \begin{cases} 1 & \text{if } \beta \nmid \alpha^2 + 1 \text{ and } \beta \nmid \alpha^2 - 1, \\ 2 & \text{otherwise.} \end{cases}$$

(c) *Suppose that f is indefinite and irreducible. Define $t_{D_f} \in \mathbb{R}$ to be such that $e^{t_{D_f}}$ is the fundamental unit of the quadratic order $\mathbb{Z}[(D_f + \sqrt{D_f})/2]$, or equivalently*

$$t_{D_f} = \log((u_{D_f} + v_{D_f}\sqrt{D_f})/2),$$

where $(u_{D_f}, v_{D_f}) \in \mathbb{N}^2$ is the least solution to $x^2 - D_f y^2 = \pm 4$. Then, we have

$$N_{\mathbb{Z},f}^{(D_4)}(X) = \frac{1}{s_f r_f} \frac{32 t_{D_f}}{9 D_f^{3/2}} X^{3/2} + O_f(X^{1+\epsilon}) \quad \text{for any } \epsilon > 0,$$

where

$$r_f = \begin{cases} 2 & \text{if } f(x, y) \sim ax^2 + bxy - ay^2 \\ & \quad \text{or } f(x, y) \sim ax^2 + bxy + cy^2 \text{ with } a \mid b, \\ 1 & \text{otherwise.} \end{cases}$$

(d) *In all three cases, for any* $\epsilon > 0$, *we have*

$$N_{\mathbb{Z},f}^{(V_4)}(X) = O_{f,\epsilon}(X^{1+\epsilon}),$$

and also

$$N_{\mathbb{Z},f}^{(C_4)}(X) = \begin{cases} O_{f,\epsilon}(X^{1/2+\epsilon}) & \text{if } -\Delta(f) \neq \square, \\ O_f(X) & \text{if } -\Delta(f) = \square. \end{cases}$$

Notice that the error terms in Theorem 1.2 depend upon f. Hence, we are unable to obtain an asymptotic formula for $N_{\mathbb{Z}}^{\dagger}(X)$ by summing over $f \in \mathfrak{F}$. However, there are only three $f \in \mathfrak{F}$ that need to be considered if we restrict to the forms in

$$V_{\mathbb{Z}}^{\text{sm},*} = \{F \in V_{\mathbb{Z}}^{\text{sm}} : F_T = F \text{ for some } T \in \text{GL}_2(\mathbb{Z}) \setminus \{\pm I_{2\times 2}\}\}.$$

This is because by Proposition 2.1 below, such a matrix T must be of the shape M_f or $M_f/2$ up to sign, where $f \in \mathfrak{F}^*$. From (1-9), we then deduce that

$$V_{\mathbb{Z}}^{\text{sm},*} = \bigcup_{\substack{f \in \mathfrak{F}^* \\ \Delta(f) \in \{-4,1,4\}}} \{F \in V_{\mathbb{Z},f}^0 : F \text{ is irreducible}\},$$

$$V_{\mathbb{Z}}^{\text{sm},*,\dagger} = \bigsqcup_{\substack{f \in \mathfrak{F}^* \\ \Delta(f) \in \{-4,1,4\}}} \{F \in V_{\mathbb{Z},f}^0 : F \text{ is irreducible and } \text{Gal}(F) \not\simeq V_4\}.$$

For $X > 0$, let us put

$$N_{\mathbb{Z}}^{*,\dagger}(X) = \#\{[F] : F \in V_{\mathbb{Z}}^{\text{sm},*,\dagger} \text{ such that } H(F) \leq X\}.$$

Then, by (1-8) and the above discussion, we have

$$N_{\mathbb{Z}}^{*,\dagger}(X) = N_{\mathbb{Z},f^{(1)}}^{*,\dagger}(X) + N_{\mathbb{Z},f^{(2)}}^{*,\dagger}(X) + N_{\mathbb{Z},f^{(3)}}^{*,\dagger}(X),$$

where we may take

$$f^{(1)}(x,y) = x^2 + y^2, \quad f^{(2)}(x,y) = x^2 + xy, \quad f^{(3)}(x,y) = x^2 + 2xy,$$

whose discriminants are -4, 1, and 4, respectively. We have:

Corollary 1.3. *We have*

$$N_{\mathbb{Z}}^{*,\dagger}(X) = \tfrac{1}{9} X^{3/2} \log X + O(X^{3/2}).$$

Proof. Theorem 1.2 implies that

$$N_{\mathbb{Z},f^{(1)}}^{\dagger}(X) = O(X^{3/2}) \quad \text{and} \quad N_{\mathbb{Z},f^{(i)}}^{\dagger}(X) = \tfrac{1}{18} X^{3/2} \log X + O(X^{3/2}) \quad \text{for } i = 2, 3.$$

Summing these terms up then yields the claim. □

Finally, as a consequence of the proof of Theorem 1.2, we also have:

Theorem 1.4. *Let $D=\beta^2+4\alpha^2$, where $\alpha, \beta \in \mathbb{N}$ are coprime and D is not a square. Then, the negative Pell's equation $x^2-Dy^2=-4$ has integer solutions if and only if the integral binary quadratic form $\alpha x^2+\beta xy-\alpha y^2$ is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to a form of the shape $ax^2+bxy+cy^2$ with a dividing b.*

We now discuss some potential applications of our Theorem 1.2 and Corollary 1.3.

First, it is natural to ask whether the asymptotic formula (1-3), which was proven using Proposition 5.1, admits a secondary main term. From the arguments in [Bhargava and Shankar 2015], we see that the error term arising from volumes of the lower dimensional projections in Proposition 5.1 is only of order $O(X^{3/4})$. Thus, possibly $X^{3/4}$ is the order of a second main term, but it is dominated by another error term coming from

$$N^*_{\mathbb{Z},\mathrm{BS}}(X)=\#\{[F]: F\in V_{\mathbb{Z}}^{\mathrm{sm},*} \text{ such that } H_{\mathrm{BS}}(F)\le X\}.$$

In particular, it was shown in [Bhargava and Shankar 2015, Lemma 2.4] that

$$N^*_{\mathbb{Z},\mathrm{BS}}(X)=O_\epsilon(X^{3/4+\epsilon}) \quad \text{for any } \epsilon>0.$$

Our Corollary 1.3 removes this obstacle, because

$$N_{\mathbb{Z}}^{*,\dagger}(X^{1/3})\le N^*_{\mathbb{Z},\mathrm{BS}}(X)\le N_{\mathbb{Z}}^{*,\dagger}(10X^{1/3})+O_\epsilon(X^{1/3+\epsilon})$$

by (1-6) and Theorem 1.2(d), whence we have

$$N^*_{\mathbb{Z},\mathrm{BS}}(X)\asymp X^{1/2}\log X.$$

This improvement potentially allows one to prove a secondary main term for (1-3) by using similar methods from [Bhargava et al. 2013], where it was shown that the counting theorem in [Davenport and Heilbronn 1971] for cubic fields has a secondary main term of order $X^{5/6}$; this latter fact was proven independently in [Taniguchi and Thorne 2013] as well.

Next, integral binary quartic forms are closely related to quartic orders, and maximal irreducible quartic orders may be regarded as quartic fields. More generally, by the construction of Birch and Merriman [1972] or Nakagawa [1989], any integral binary form F gives rise to a $\mathbb{Z}$-order Q_F whose rank is the degree of F, where $\mathrm{GL}_2(\mathbb{Z})$-equivalence class of F corresponds to isomorphism class of Q_F. By [Delone and Faddeev 1964], it is well-known that all cubic orders come from integral binary cubic forms, which enabled the enumeration of cubic orders having a nontrivial automorphism as well as cubic fields by their discriminant; see [Bhargava and Shnidman 2014] and [Davenport and Heilbronn 1971], respectively. But this is not true for orders of higher rank. Parametrizations of quartic and quintic orders were given by Bhargava in his seminal work [2004; 2008]. Wood [2012] further showed that the quartic orders arising from integral binary quartic forms are exactly those having a monogenic *cubic resolvent*; see [Bhargava 2004] for the definition. This implies that the forms in

$$V_{\mathbb{Z}}^{\mathrm{sm},\star} = \{F \in V_{\mathbb{Z}}^{\mathrm{sm}} : Q_F \text{ is maximal}\}$$

correspond to quartic D_4-, C_4-, and V_4-fields whose ring of integers has a monogenic cubic resolvent. In our upcoming paper [Tsang and Xiao 2017], we shall enumerate $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of forms in $V_{\mathbb{Z}}^{\mathrm{sm},\star}$ with respect to a height corresponding to the conductor of fields, as motivated by [Altuğ et al. 2017]. In fact, we shall that show that

$$\text{for all } f \in \mathfrak{F}^* : \quad F \in V_{\mathbb{Z}}^{\mathrm{sm},\star} \cap V_{\mathbb{Z},f}^{0} \neq \varnothing \quad \text{if and only if} \quad \Delta(f) \in \{-4, 1, 4\}.$$

Thus, our counting theorem in [Tsang and Xiao 2017] may be regarded as a refinement and an extension of Corollary 1.3 above.

Last but not least, binary quartic forms are connected to elliptic curves as well. In particular, any integral binary quartic form F gives rise to an elliptic curve

$$E_F : y^2 = x^3 - \frac{I(F)}{3}x - \frac{J(F)}{27}$$

defined over $\mathbb{Q}$. Bhargava and Shankar [2015] applied (1-3) as well as a parametrization of 2-Selmer groups due to Birch and Swinnerton-Dyer to show that the average rank of elliptic curves over $\mathbb{Q}$, when ordered by a *naive* height analogous to (1-2), is at most $\frac{3}{2}$. This result is remarkable in that it is the first to show, unconditional on the BSD-conjecture and the Grand Riemann Hypothesis, boundedness of the average rank of large families of elliptic curves over $\mathbb{Q}$. Conditional bounds were obtained by Brumer [1992], Heath-Brown [2004], and Young [2006] previously. Now, the relations in (1-5) imply that for $F \in V_{\mathbb{Z}}^{\mathrm{sm}} \cap V_{\mathbb{Z},f}^{0}$ with $f \in \mathfrak{F}^*$, we have

$$E_F : y^2 = \left(x + \frac{L_f(F)}{3}\right)\left(x^2 - \frac{L_f(F)}{3}x - \frac{K_f(F)}{9}\right),$$

which has a rational 2-torsion point. Hence, our Theorem 1.2 potentially allows one to study arithmetic properties of elliptic curves with 2-torsion over $\mathbb{Q}$. Let us remark that unlike a *large* family of elliptic curves over $\mathbb{Q}$, in the sense of [Bhargava and Shankar 2015, Section 3], the family consisting of those curves with a rational 2-torsion exhibits a rather peculiar behavior. Indeed, Klagsbrun and Lemke Oliver [2014] proved that the average size of the 2-Selmer groups in this family is unbounded, and they conjectured an asymptotic growth rate. One might be able to obtain such an asymptotic growth rate using our Theorem 1.2 and a sieve that detects local solubility; this line of inquiry is pursued in an upcoming paper due to D. Kane and Z. Klagsbrun.

2. Characterization of forms with small Galois groups

2A. ***Cremona covariants.*** Let F be a real binary quartic form with $\Delta(F) \neq 0$. As Cremona defined [1999], we have three quadratic covariants $\mathfrak{C}_{F,\omega}(x, y)$, each of which is associated to a root ω of $\mathcal{Q}_F(x)$; see [Xiao 2019, Subsection 4.2] for the explicit definition. They satisfy the syzygy

$$\mathfrak{C}_{F,\omega}(x, y)^2 = \tfrac{1}{3}(F_4(x, y) + 4\omega F(x, y)), \tag{2-1}$$

where F_4 is the *Hessian covariant of* F and is given by

$$F_4(x, y) = 3(a_3^2 - 8a_4a_2)x^4 + 4(a_3a_2 - 6a_4a_1)x^3y + 2(2a_2^2 - 24a_4a_0 - 3a_3a_1)x^2y^2 \\ + 4(a_2a_1 - 6a_3a_0)xy^3 + (3a_1^2 - 8a_2a_0)y^4.$$

We shall label the roots $\omega_1(F), \omega_2(F), \omega_3(F)$ of $\mathcal{Q}_F(x)$ such that

$$\mathfrak{C}_{F,\omega_i(F)}(x, y) = \mathfrak{C}_{F,i}(x, y) \quad \text{for all } i = 1, 2, 3,$$

where $\mathfrak{C}_{F,i}(x, y)$ is defined as in [Xiao 2019, (4.6)]. Then, from (2-1) and the explicit expressions for $\mathfrak{C}_{F,\omega}(x, y)$ given in [Xiao 2019], we have the following observations:

(1) For $\omega = \omega_1(F)$, the binary quadratic form $\mathfrak{C}_{F,\omega}(x, y)$ has real coefficients.

(2) For $\omega = \omega_2(F), \omega_3(F)$, we have:
   - If $\Delta(F) > 0$, then $\lambda_\omega \cdot \mathfrak{C}_{F,\omega}(x, y)$ has real coefficients for some $\lambda_\omega \in \{1, \sqrt{-1}\}$.
   - If $\Delta(F) < 0$, then $\lambda \cdot \mathfrak{C}_{F,\omega}(x, y)$ does not have real coefficients for all $\lambda \in \mathbb{C}^\times$.

Also, it is easy to check that

$$\Delta(\mathfrak{C}_{F,\omega_1(F)}), \Delta(\mathfrak{C}_{F,\omega_3(F)}) > 0 \quad \text{and} \quad \Delta(\mathfrak{C}_{F,\omega_2(F)}) < 0. \tag{2-2}$$

We shall require the following result by Xiao [2019].

Proposition 2.1. *Let F be a real binary quartic form with $\Delta(F) \neq 0$. Then, a set of representatives for the quotient group*

$$\{T \in \mathrm{GL}_2(\mathbb{R}) : F_T = F\}/\{\lambda \cdot I_{2\times2} : \lambda \in \mathbb{R}^\times\}$$

is given by

$$\begin{cases} \{I_{2\times2}, M_f : f \in \{\mathfrak{C}_{F,\omega_1(F)},\ \lambda_{\omega_2(F)} \cdot \mathfrak{C}_{F,\omega_2(F)},\ \lambda_{\omega_3(F)} \cdot \mathfrak{C}_{F,\omega_3(F)}\}\} & \text{if } \Delta(F) > 0, \\ \{I_{2\times2}, M_f : f \in \{\mathfrak{C}_{F,\omega_1(F)}\}\} & \text{if } \Delta(F) < 0. \end{cases}$$

Furthermore, the quadratic forms $\mathfrak{C}_{F,\omega_1(F)}(x, y)$, $\mathfrak{C}_{F,\omega_2(F)}(x, y)$, and $\mathfrak{C}_{F,\omega_3(F)}(x, y)$, are pairwise nonproportional over $\mathbb{C}^\times$.

Proof. For the first statement, see [Xiao 2019, Proposition 4.6]. As for the second statement, since $\mathfrak{C}_{F,\omega_i(F)}(x, y)$ are covariants, replacing F by a $\mathrm{GL}_2(\mathbb{R})$-translate if necessary, we may assume that

$$F(x, y) = a_4x^4 + a_2x^2y^2 \pm a_4y^4.$$

In this special case, it is not hard to verify the claim using the explicit expressions for $\mathfrak{C}_{F,\omega_i(F)}(x, y)$ in [Xiao 2019, (4.6)]. □

Let F be a real binary quartic form with $\Delta(F) \neq 0$. Proposition 2.1 implies that for any real binary quadratic form f with $\Delta(f) \neq 0$, we have $F \in V_{\mathbb{R},f}$ if and only if

(2-3) $\qquad f(x, y)$ is proportional to $\mathfrak{C}_{F,\omega}(x, y)$ for a root ω of $\mathcal{Q}_F(x)$.

Moreover, this root ω is unique, and we shall denote it by $\omega_f(F)$. This was required in order to define the L_f- and K_f-invariants in (1-4).

2B. ***Proof of Theorem 1.1.*** The key is the following lemma.

Lemma 2.2. *Let F be an integral binary quartic form with $\Delta(F) \neq 0$ and let ω be a root of $\mathcal{Q}_F(x)$. Then, the quadratic form $\mathfrak{C}_{F,\omega}(x, y)$ is proportional over $\mathbb{C}^\times$ to a form with integer coefficients if and only if $\omega \in \mathbb{Z}$.*

Proof. If $\omega \in \mathbb{Z}$, then we easily see from (2-1) that $\lambda \cdot \mathfrak{C}_{F,\omega}(x, y)$ has integer coefficients for some $\lambda \in \mathbb{C}^\times$. Conversely, if $\lambda \cdot \mathfrak{C}_{F,\omega}(x, y)$ has integer coefficients for some $\lambda \in \mathbb{C}^\times$, then consider the action of an element $\sigma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, where $\bar{\mathbb{Q}}$ is an algebraic closure of $\mathbb{Q}$. It is clear from the definition of $\mathfrak{C}_{F,\omega}(x, y)$ that $\lambda \in \bar{\mathbb{Q}}$. From (2-1), we have

$$\tfrac{4}{3}(\omega - \sigma(\omega))F(x, y) = \mathfrak{C}_{F,\omega}(x, y)^2 - \sigma(\mathfrak{C}_{F,\omega}(x, y)^2) = \Big(1 - \frac{\lambda^2}{\sigma(\lambda)^2}\Big)\mathfrak{C}_{F,\omega}(x, y)^2,$$

and this last binary quartic form has zero discriminant. This shows that $\omega - \sigma(\omega) = 0$ for all $\sigma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. Thus, we have $\omega \in \mathbb{Q}$, and so $\omega \in \mathbb{Z}$ since $\mathcal{Q}_F(x)$ is monic. □

The first claim in Theorem 1.1 now follows from Proposition 2.1, Lemma 2.2, and (2-3). Note that

$$\Delta(F) = 27^2\Delta(\mathcal{Q}_F),$$

which means that $\mathcal{Q}_F(x)$ has three integer roots if and only if $\mathcal{Q}_F(x)$ is reducible and $\Delta(F) = \square$. The second claim then follows from this fact and (2-2).

3. Basic properties of forms in $V_{\mathbb{R},f}$ of nonzero discriminant

Throughout this section, let $f(x,y)=\alpha x^2+\beta xy+\gamma y^2$ be a real binary quadratic form with $\Delta(f)\neq 0$. It is not hard to check, by a direct calculation, that

$$V_{\mathbb{R},f}=\left\{Ax^4+Bx^3y+Cx^2y^2+\left(\frac{4\beta\gamma A-(\beta^2+2\alpha\gamma)B+2\alpha\beta C}{2\alpha^2}\right)xy^3 +\left(\frac{4\gamma(\beta^2+2\alpha\gamma)A-\beta(\beta^2+4\alpha\gamma)B+2\alpha\beta^2C}{8\alpha^3}\right)y^4 : A,B,C\in\mathbb{R}\right\} \tag{3-1}$$

if $\alpha\neq 0$, and similarly that

$$V_{\mathbb{R},f}=\left\{Ax^4+\left(\frac{\gamma(4\beta^2+8\alpha\gamma)A+2\alpha\beta^2B-8\alpha^3C}{\beta(\beta^2+4\alpha\gamma)}\right)x^3y+Bx^2y^2 -\left(\frac{8\gamma^3A-2\beta^2\gamma B-\alpha(4\beta^2+8\alpha\gamma)C}{\beta(\beta^2+4\alpha\gamma)}\right)xy^3+Cy^4 : A,B,C\in\mathbb{R}\right\} \tag{3-2}$$

if $\beta,\beta^2+4\alpha\gamma\neq 0$. Below, we shall give some basic properties of $V^0_{\mathbb{R},f}$ and $V^0_{\mathbb{Z},f}$.

3A. *The two new invariants.* Recall the definitions of the L_f- and K_f-invariants given in (1-4). First, we shall show that they are indeed invariants under the twisted action of $\mathrm{GL}_2(\mathbb{R})$ in the following sense.

Proposition 3.1. *For all $F\in V^0_{\mathbb{R},f}$ and $T\in\mathrm{GL}_2(\mathbb{R})$, we have*

$$L_{f_T}(F_T)=L_f(F) \quad \text{and} \quad K_{f_T}(F_T)=K_f(F).$$

Proof. Notice that $\mathcal{Q}_F(x)=\mathcal{Q}_{F_T}(x)$. For any root ω of $\mathcal{Q}_F(x)$, because $\mathfrak{C}_{F,\omega}(x,y)$ is a covariant up to sign by (2-1), if $\mathfrak{C}_{F,\omega}(x,y)$ is proportional to $f(x,y)$, then $\mathfrak{C}_{F_T,\omega}(x,y)$ is proportional to $f_T(x,y)$. It then follows from the definition that $L_{f_T}(F_T)=L_f(F)$. Since $I(F_T)=I(F)$, we also have $K_{f_T}(F_T)=K_f(F)$ by the first equality in (1-5). □

We shall give explicit formulae for $L_f(-)$ and $K_f(-)$ in two special cases.

Proposition 3.2. *The following holds.*

(a) *Assume that $\alpha\neq 0$. Then, for all $F\in V^0_{\mathbb{R},f}$ as in (3-1), we have*

$$L_f(F)=-(12\gamma A-3\beta B+2\alpha C)/(2\alpha),$$
$$K_f(F)=(72\beta^2\gamma A^2+9\alpha(\beta^2+4\alpha\gamma)B^2+8\alpha^3C^2 -18\beta(\beta^2+4\alpha\gamma)AB+12\alpha(3\beta^2-4\alpha\gamma)AC-24\alpha^2\beta BC)/(4\alpha^3).$$

Moreover, we have

$$\frac{4(L_f(F)^2+4K_f(F))}{9}=\frac{L_{f,1}(F)^2-\Delta(f)L_{f,2}(F)^2}{\alpha^4},$$

where

$$L_{f,1}(F)=4(\beta^2-\alpha\gamma)A-3\alpha\beta B+2\alpha^2C \quad \textit{and} \quad L_{f,2}(F)=2(2\beta A-\alpha B).$$

(b) *Assume that* $\gamma=0$. *Then, for all* $F\in V^0_{\mathbb{R},f}$ *as in* (3-2), *we have*

$$L_f(F)=(2\beta^2B-12\alpha^2C)/\beta^2,$$
$$K_f(F)=(-\beta^4B^2+144\alpha^4C^2+36\beta^4AC-24\alpha^2\beta^2BC)/\beta^4.$$

Moreover, we have

$$\frac{4(L_f(F)^2+4K_f(F))}{9}=\frac{8C}{\beta^2}\Big(8\beta^2A-8\alpha^2B+\frac{40\alpha^4}{\beta^2}C\Big).$$

Proof. This may be verified by explicit computation. □

We shall also need the following observation.

Proposition 3.3. *Assume that* f *is integral. Then, for all* $F\in V^0_{\mathbb{Z},f}$, *we have*

$$L_f(F),\ K_f(F),\ (L_f(F)^2+4K_f(F))/9,\ (2L_f(F)^2-K_f(F))/9\in\mathbb{Z}.$$

Moreover, when f *is primitive in addition, we have*

$$4(2L_f(F)^2-K_f(F))/(9\Delta(f))\in\mathbb{Z}.$$

Proof. We have $L_f(F)\in\mathbb{Z}$ by Lemma 2.2. Since $I(F)\in\mathbb{Z}$, we deduce from the first equality in (1-5) that $K_f(F)\in\mathbb{Z}$ holds as well. Observe that

$$I(F)+K_f(F)=(L_f(F)^2+4K_f(F))/3,$$
$$2I(F)-K_f(F)=(2L_f(F)^2-K_f(F))/3,$$

both of which are integers. Since $\Delta(F)\in\mathbb{Z}$, we deduce from (1-7) that at least one of the above expressions is divisible by 3. But again by (1-5), we have

$$3I(F)=(L_f(F)^2+4K_f(F))/3+(2L_f(F)^2-K_f(F))/3,$$

so in fact both expressions are divisible by 3. This proves the first claim.

Next, assume that f is primitive in addition. In view of Proposition 3.1, by applying a $\mathrm{GL}_2(\mathbb{Z})$-action on f if necessary, we may assume that $\alpha\neq0$ and that α is coprime to $\Delta(f)$. Using Proposition 3.2(a), we then compute that

$$\frac{4(2L_f(F)^2-K_f(F))}{9}=\Delta(f)\Big(\frac{\alpha(B^2-4AC)+2A(\beta B-4\gamma A)}{\alpha^3}\Big).$$

This expression is an integer by the first claim, and hence must be divisible by $\Delta(f)$, because α is taken to be coprime to $\Delta(f)$. This proves the second claim. □

3B. ***Determinants of the two lattices.*** In this subsection, assume that f is integral and primitive. Let $\Lambda_{f,1}$ and $\Lambda_{f,2}$ denote the lattices defined in (1-10). Below, we shall compute their determinants in terms of the number s_f as in Theorem 1.2.

Proposition 3.4. *We have* $\det(\Lambda_{f,1}) = s_f|\alpha|^3$ *and* $\det(\Lambda_{f,2}) = s_f|\beta(\beta^2+4\alpha\gamma)|/8$.

Proof. Observe that the linear transformation defined by the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ * & -\mathcal{B} & * \end{pmatrix}, \quad \text{where } \mathcal{B} = \frac{\beta(\beta^2+4\alpha\gamma)}{8\alpha^3},$$

has determinant $\mathcal{B}$, and it sends $\Lambda_{f,1}$ to $\Lambda_{f,2}$. Thus, it suffices to prove the first claim. Recall from (3-1) that $\Lambda_{f,1}$ is the set of tuples $(A, B, C) \in \mathbb{Z}^3$ satisfying

$$4\beta\gamma A - (\beta^2+2\alpha\gamma)B + 2\alpha\beta C \equiv 0 \pmod{2\alpha^2},$$
$$4\gamma(\beta^2+2\alpha\gamma)A - \beta(\beta^2+4\alpha\gamma)B + 2\alpha\beta^2 C \equiv 0 \pmod{8\alpha^3}.$$

If $\beta\gamma = 0$, then it is easy to check that $\det(\Lambda_{f,1}) = s_f|\alpha|^3$. If $\beta\gamma \neq 0$, then we shall use the fact that

$$\det(\Lambda_{f,1}) = \prod_p \det(\Lambda_{f,1}^{(p)}) = \prod_{p|2\alpha} \det(\Lambda_{f,1}^{(p)}), \quad \text{where } \Lambda_{f,1}^{(p)} = \mathbb{Z}_p \otimes_{\mathbb{Z}} \Lambda_{f,1},$$

and so $\det(\Lambda_{f,1}) = s_f|\alpha|^3$ indeed holds by Lemma 3.5 below. □

Lemma 3.5. *Let p be a prime dividing 2α and let $p^k \| \alpha$. Then, we have*

$$\det(\Lambda_{f,1}^{(p)}) = s_f^{\epsilon_p} p^{3k}, \quad \textit{where } \epsilon_p = \begin{cases} 1 & \textit{if } p = 2, \\ 0 & \textit{if } p \geq 3. \end{cases}$$

Proof. For brevity, write

$$\alpha = p^k a \quad \text{and} \quad \beta = p^\ell b, \quad \text{where } k, \ell, a, b \in \mathbb{Z} \text{ with } k, \ell \geq 0 \text{ and } p \nmid a, b.$$

Then, the claim may be restated as

$$\det(\Lambda_{f,1}^{(p)}) = \begin{cases} p^{3k+3\epsilon_p} & \text{if } \ell = 0, \\ p^{3k} & \text{if } \ell \geq 1. \end{cases}$$

By definition, the lattice $\Lambda_{f,1}^{(p)}$ is the set $(A, B, C) \in \mathbb{Z}_p^3$ of tuples satisfying

$$\mathcal{T}_1(A,B,C) \equiv 0 \pmod{p^{2k+\epsilon_p}} \quad \text{and} \quad \mathcal{T}_2(A,B,C) \equiv 0 \pmod{p^{3k+3\epsilon_p}},$$

where

$$\mathcal{T}_1(A,B,C) = p^\ell b(4\gamma A - p^\ell bB) - 2p^k a\gamma B + 2p^{k+\ell}abC,$$
$$\mathcal{T}_2(A,B,C) = (p^{2\ell}b^2 + 4p^k a\gamma)(4\gamma A - p^\ell bB) - 8p^k a\gamma^2 A + 2p^{k+2\ell}ab^2 C.$$

Observe that we have the relation

$$\mathcal{T}_2(A,B,C) - p^\ell b\mathcal{T}_1(A,B,C) = 2p^k a\gamma(4\gamma A - p^\ell bB). \tag{3-3}$$

For $\ell = 0$, we deduce from (3-3) that $\Lambda^{(p)}_{f,1}$ is defined solely by

$$\mathcal{T}_2(A,B,C) \equiv 0 \pmod{p^{3k+3\epsilon_p}}.$$

For $\ell \geq 1$ and $\ell \geq k + 2\epsilon_p$, it is easy to see that $\Lambda^{(p)}_{f,1}$ is in fact defined by

$$A \equiv 0 \pmod{p^{2k}} \quad \text{and} \quad B \equiv 0 \pmod{p^k}.$$

For $\ell \geq 1$ and $\ell \leq k + \epsilon_p$, we shall first show that $\Lambda^{(p)}_{f,1}$ is also defined by

$$\begin{cases} A \equiv 0 & \pmod{p^{2\ell-2\epsilon_p}}, \\ B \equiv 0 & \pmod{p^{\ell-\epsilon_p}}, \\ (4\gamma A - p^\ell bB)/p^{2\ell-\epsilon_p} \equiv 0 & \pmod{p^{k-\ell+\epsilon_p}}, \\ \mathcal{T}_2(A,B,C)/p^{k+2\ell+\epsilon_p} \equiv 0 & \pmod{p^{2k-2\ell+2\epsilon_p}}. \end{cases} \tag{3-4}$$

If (3-4) is satisfied, then from (3-3), it is easy to see that $(A,B,C) \in \Lambda^{(p)}_{f,1}$. Conversely, if $(A,B,C) \in \Lambda^{(p)}_{f,1}$, then the assumption $\ell \leq k + \epsilon_p$ implies that

$$\mathcal{T}_1(A,B,C) \equiv 0 \pmod{p^{k+\ell}} \quad \text{and} \quad \mathcal{T}_2(A,B,C) \equiv 0 \pmod{p^{k+2\ell+\epsilon_p}},$$

while reducing (3-3) mod $p^{2k+\ell+\epsilon_p}$ also yields

$$4\gamma A - p^\ell bB \equiv 0 \pmod{p^{k+\ell}}.$$

From these three congruence equations, it follows that (3-4) is indeed satisfied. In all cases, we then see that $\det(\Lambda^{(p)}_{f,1})$ is as claimed. □

3C. ***Forms with abelian Galois groups.*** In this subsection, assume that f is integral. Consider an irreducible form $F \in V^0_{\mathbb{Z},f}$. By Theorem 1.1, we have $\mathrm{Gal}(F) \simeq D_4$, C_4, or V_4. To distinguish among these three possibilities, note that the *cubic resolvent polynomial of* F, defined by

$$R_F(x) = a_4^3X^3 - a_4^2a_2X^2 + a_4(a_3a_1 - 4a_4a_0)X - (a_3^2a_0 + a_4a_1^2 - 4a_4a_2a_0)$$

when F has the shape (1-1), is reducible since $\mathrm{Gal}(F)$ is small. Also, it has a unique root $r_F \in \mathbb{Q}$ precisely when $\Delta(F) \neq \square$, in which case we define

$$\theta_1(F) = (a_3^2 - 4a_4(a_2 - r_Fa_4))\Delta(F) \quad \text{and} \quad \theta_2(F) = a_4(r_F^2a_4 - 4a_0)\Delta(F).$$

Then, we have the well-known criterion

$$\begin{aligned} &\mathrm{Gal}(F) \simeq V_4 \iff \Delta(F) = \square, \\ &\mathrm{Gal}(F) \simeq C_4 \iff \Delta(F) \neq \square \text{ and } \theta_1(F), \theta_2(F) = \square \text{ in } \mathbb{Q}. \end{aligned}$$

See [Conrad 2012], for example. We then deduce:

Proposition 3.6. *Let $F \in V^0_{\mathbb{Z},f}$ be an irreducible form. Then, we have*

$$\mathrm{Gal}(F) \simeq V_4 \iff L_f(F)^2 + 4K_f(F) = \square,$$

as well as

$$\mathrm{Gal}(F) \simeq C_4 \iff \begin{cases} L_f(F)^2 + 4K_f(F) \neq \square, \\ (L_f(F)^2 + 4K_f(F))(2L_f(F)^2 - K_f(F))/\Delta(f) = \square. \end{cases}$$

Proof. Observe that by (1-7), we have

$$\Delta(F) = \square \quad \text{if and only if} \quad L_f(F)^2 + 4K_f(F) = \square.$$

The first claim is then clear. Next, suppose that $\Delta(F) \neq \square$. By Proposition 3.1, we may assume that $\alpha \neq 0$. For F in the shape as in (3-1), a direct computation yields

$$r_F = (-4\gamma A + \beta B)/(2\alpha A).$$

Using Proposition 3.2 (a), we further compute that

$$\theta_1(F) = 4\alpha^2(2L_f(F)^2 - K_f(F))\Delta(F)/(9\Delta(f)),$$
$$\theta_2(F) = \beta^2(2L_f(F)^2 - K_f(F))\Delta(F)/(9\Delta(f)).$$

By (1-7) and the criterion above, it follows that $\theta_1(F), \theta_2(F)$ are squares if and only if $(L_f(F)^2 + 4K_f(F))(2L_f(F)^2 - K_f(F))/\Delta(f)$ is a square, as desired. □

3D. *Reducible forms.* In this subsection, assume that f is integral. We shall study the reducible forms in $V^0_{\mathbb{Z},f}$. Let us first make a definition and an observation.

Definition 3.7. Let $F \in V^0_{\mathbb{Z},f}$ be a reducible form.

(1) We say that F is of *type* 1 if $F = m \cdot pp_{M_f}$ for some $m \in \mathbb{Q}^\times$ and integral binary quadratic form p.

(2) We say that F is of *type* 2 if $F = pq$ for some integral binary quadratic forms p and q satisfying $p_{M_f} = -p$ and $q_{M_f} = -q$.

Lemma 3.8. *For all reducible forms $F \in V^0_{\mathbb{Z},f}$ of type* 1, *we have*

$$L_f(F)^2 + 4K_f(F) = \square.$$

Proof. This may be verified by a direct computation. □

Below, we shall show that the two reducibility types in Definition 3.7 are in fact the only possibilities. We shall require two further lemmas.

Lemma 3.9. *Let $\ell(x, y) = \ell_1 x + \ell_0 y$ be a nonzero complex binary linear form, and suppose that $\ell_{M_f} = \lambda \cdot \ell$ for some $\lambda \in \mathbb{C}^\times$. Then, we have $\lambda = \pm\sqrt{-1}$, with*

$$\lambda = \begin{cases} -\sqrt{-1} & \text{if and only if } \ell_0 = (\beta + \sqrt{\Delta(f)})\ell_1/(2\alpha), \\ \sqrt{-1} & \text{if and only if } \ell_0 = (\beta - \sqrt{\Delta(f)})\ell_1/(2\alpha), \end{cases}$$

in the case that $\alpha \neq 0$.

Proof. The hypothesis implies that

$$\frac{1}{\sqrt{-\Delta(f)}} \begin{pmatrix} \beta & -2\alpha \\ 2\gamma & -\beta \end{pmatrix} \begin{pmatrix} \ell_1 \\ \ell_0 \end{pmatrix} = \lambda \begin{pmatrix} \ell_1 \\ \ell_0 \end{pmatrix}.$$

Then, by computing the eigenvalues and eigenspaces of the 2×2 matrix above, we see that the claim holds. □

Lemma 3.10. *Let $p(x, y) = p_2 x^2 + p_1 xy + p_0 y^2$ be a nonzero complex binary quadratic form, and suppose that $p_{M_f} = \lambda \cdot p$ for some $\lambda \in \mathbb{C}^\times$. Then, we have $\lambda = \pm 1$, with*

$$\lambda = \begin{cases} -1 & \text{if and only if } p_0 = (\beta p_1 - 2\gamma p_2)/(2\alpha), \\ 1 & \text{if and only if } p = (p_2/\alpha) f, \end{cases}$$

in the case that $\alpha \neq 0$.

Proof. The hypothesis implies that

$$\frac{1}{-\Delta(f)} \begin{pmatrix} \beta^2 & -2\alpha & 4\alpha^2 \\ 4\beta\gamma & -(\beta^2 + 4\alpha\gamma) & 4\alpha\beta \\ 4\gamma^2 & -2\beta\gamma & \beta^2 \end{pmatrix} \begin{pmatrix} p_2 \\ p_1 \\ p_0 \end{pmatrix} = \lambda \begin{pmatrix} p_2 \\ p_1 \\ p_0 \end{pmatrix}.$$

Then, by computing the eigenvalues and eigenspaces of the 3×3 matrix above, it is not hard to check that the claim holds. □

Proposition 3.11. *Any reducible form $F \in V^0_{\mathbb{Z},f}$ is either of type 1 or of type 2.*

Proof. Write $F = g^{(1)} g^{(2)} g^{(3)} g^{(4)}$, where the $g^{(k)}$ are complex binary linear forms, and are pairwise nonproportional because $\Delta(F) \neq 0$. Since F is reducible, by renumbering if necessary, we may assume that

$$\begin{cases} g^{(1)}, g^{(2)} g^{(3)} g^{(4)} & \text{when } F \text{ has exactly one rational linear factor}, \\ g^{(1)}, g^{(2)}, g^{(3)} g^{(4)} & \text{when } F \text{ has exactly two rational linear factors}, \\ g^{(1)} g^{(2)}, g^{(3)} g^{(4)} & \text{when } F \text{ has no rational linear factor}, \\ g^{(1)}, g^{(2)}, g^{(3)}, g^{(4)} & \text{when } F \text{ has four rational linear factors}, \end{cases}$$

have integer coefficients and are irreducible. We have $M_f^2 = \Delta(f) \cdot I_{2\times 2}$ and $F_{M_f} = F$ by definition. Hence, up to scaling, the matrix M_f acts on the $g^{(k)}$ via a permutation σ on four letters of order dividing two. This has two consequences.

By (1-8), without loss of generality, we may assume that $\alpha \neq 0$. First, the form F cannot have exactly one rational linear factor, for otherwise

$$\sigma(1) = 1 \quad \text{and} \quad \sigma(k_0) = k_0 \text{ for at least one } k_0 \in \{2, 3, 4\}.$$

From Lemma 3.9, it would follow that $\Delta(f)$ is a square and that $g^{(k_0)}$ is proportional to a form with integer coefficients, which is a contradiction. Second, when F has four rational linear factors, by further renumbering if necessary, we may assume that

$$\sigma \in \{(1), (12), (12)(34)\}.$$

Now, in all three of the possible cases for the factorization of F, define

$$p = g^{(1)}g^{(2)} \quad \text{and} \quad q = g^{(3)}g^{(4)},$$

which are integral binary quadratic forms by definition. We then deduce that

$$(p_{M_f}, q_{M_f}) = (\lambda \cdot q, \lambda^{-1} \cdot p) \quad \text{or} \quad (p_{M_f}, q_{M_f}) = (\lambda \cdot p, \lambda^{-1} \cdot q)$$

for some $\lambda \in \mathbb{Q}^\times$. In the former case, it is clear that F is of type 1. In the latter case, we have $\lambda = -1$ by Lemma 3.10 and the fact that $\Delta(F) \neq 0$, so F is of type 2. □

4. Parametrizing forms in $V_{\mathbb{R},f}$ of nonzero discriminant

Throughout this section, let $f(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ be a real binary quadratic form with $\Delta(f) \neq 0$ and $\alpha > 0$. We shall give an alternative parametrization of $V^0_{\mathbb{R},f}$, different from (3-1) and (3-2), in terms of the regions

$$\begin{cases} \Omega^0 = \{(L, K) \in \mathbb{R}^2 : L^2 + 4K \neq 0 \text{ and } 2L^2 - K \neq 0\}, \\ \Omega^+ = \{(L, K) \in \mathbb{R}^2 : L^2 + 4K > 0 \text{ and } 2L^2 - K \neq 0\}, \\ \Omega^- = \{(L, K) \in \mathbb{R}^2 : L^2 + 4K < 0 \text{ and } 2L^2 - K > 0\}, \end{cases} \tag{4-1}$$

corresponding to the L_f- and K_f-invariants, as well as a parameter $t \in \mathbb{R}$ arising from the *orthogonal group of* f, defined by

$$O_f(\mathbb{R}) = \{T \in \mathrm{GL}_2(\mathbb{R}) : \det(T) = \pm 1 \text{ and } f_T = \pm f\}.$$

Note that by (1-7), for any $F \in V^0_{\mathbb{R},f}$, we have

$$(L_f(F), K_f(F)) \in \Omega^+ \iff \Delta(F) > 0,$$
$$(L_f(F), K_f(F)) \in \Omega^- \iff \Delta(F) < 0.$$

First, we shall show that it suffices to consider $x^2 + y^2$ and $x^2 - y^2$. It shall be helpful to recall (1-8) as well as the isomorphisms Θ_1 and Θ_2 defined in Section 1A.

Lemma 4.1. *Define a matrix*

$$T_f = \begin{pmatrix} \delta_f^{-1/4} & 0 \\ 0 & \delta_f^{1/4} \end{pmatrix} \cdot \frac{1}{2\sqrt{\alpha}} \begin{pmatrix} 2\alpha & \beta \\ 0 & 2 \end{pmatrix}, \quad \textit{where } \delta_f = \frac{|\Delta(f)|}{4}.$$

Then, we have a well-defined bijective linear map

$$\begin{cases} \Psi_f : V_{\mathbb{R},x^2+y^2} \to V_{\mathbb{R},f}, & \Psi_f(F) = F_{T_f} \quad \textit{if } f \textit{ is positive definite}, \\ \Psi_f : V_{\mathbb{R},x^2-y^2} \to V_{\mathbb{R},f}, & \Psi_f(F) = F_{T_f} \quad \textit{if } f \textit{ is indefinite}, \end{cases}$$

and we have $\det(\Psi_f) = 8\alpha^3|\Delta(f)|^{-3/2}$.

Proof. The first claim holds by (1-8) and the fact

$$\delta_f^{-1/2} \cdot f = \begin{cases} (x^2+y^2)_{T_f} & \text{if } f \text{ is positive definite}, \\ (x^2-y^2)_{T_f} & \text{if } f \text{ is indefinite}. \end{cases}$$

Identifying $V_{\mathbb{R},x^2\pm y^2}$ and $V_{\mathbb{R},f}$ with $\mathbb{R}^3$ via Θ_1, we see from (3-1) that

$$\text{(4-2)} \qquad \Psi_f : \begin{pmatrix} a_4 \\ a_3 \\ a_2 \end{pmatrix} \mapsto \begin{pmatrix} \alpha^2/\delta_f & 0 & 0 \\ 2\alpha\beta/\delta_f & \alpha/\sqrt{\delta_f} & 0 \\ 3\beta^2/2\delta_f & 3\beta/(2\sqrt{\delta_f}) & 1 \end{pmatrix} \begin{pmatrix} a_4 \\ a_3 \\ a_2 \end{pmatrix},$$

from which the second claim follows. □

In the subsequent subsections, we shall prove the following propositions.

Proposition 4.2. *There exists an explicit bijection*

$$\Phi : \Omega^+ \times [-\pi/4, \pi/4) \to V^0_{\mathbb{R},x^2+y^2},$$

defined as in (4-4), *such that*

(a) *we have* $L_{x^2+y^2}(\Phi(L,K,t)) = L$ *and* $K_{x^2+y^2}(\Phi(L,K,t)) = K$,

(b) *the Jacobian matrix of* $\Theta_1 \circ \Phi$ *has determinant* $-\frac{1}{18}$.

Proposition 4.3. *There exist explicit injections*

$$\Phi^{(1)}, \Phi^{(2)} : \Omega^+ \times \mathbb{R} \to V^0_{\mathbb{R},x^2-y^2} \quad \textit{and} \quad \Phi^{(3)}, \Phi^{(4)} : \Omega^- \times \mathbb{R} \to V^0_{\mathbb{R},x^2-y^2},$$

defined as in (4-6), *with*

$$V^0_{\mathbb{R},x^2-y^2} = \Phi^{(1)}(\Omega^+ \times \mathbb{R}) \sqcup \Phi^{(2)}(\Omega^+ \times \mathbb{R}) \sqcup \Phi^{(3)}(\Omega^- \times \mathbb{R}) \sqcup \Phi^{(4)}(\Omega^- \times \mathbb{R})$$

such that, for all $i = 1, 2, 3, 4$,

(a) *we have* $L_{x^2-y^2}(\Phi^{(i)}(L,K,t)) = L$ *and* $K_{x^2-y^2}(\Phi^{(i)}(L,K,t)) = K$,

(b) *the Jacobian matrix of* $\Theta_1 \circ \Phi^{(i)}$ *has determinant* $-\frac{1}{18}$.

In view of (1-11), we shall give another parametrization of $V_{\mathbb{R},f}$ when $\gamma = 0$, which does not require reducing to the form $x^2 - y^2$ via Lemma 4.1.

Proposition 4.4. *Suppose that* $\gamma = 0$. *Then, there exist explicit injections*

$$\Phi_f^{(1)}, \Phi_f^{(2)} : \Omega^0 \times \mathbb{R} \to V^0_{\mathbb{R},f},$$

defined as in (4-9), *with*

$$V^0_{\mathbb{R},f} = \Phi_f^{(1)}(\Omega^0 \times \mathbb{R}) \sqcup \Phi_f^{(2)}(\Omega^0 \times \mathbb{R})$$

such that, *for both* $i = 1, 2$,

(a) *we have* $L_f(\Phi^{(i)}(L, K, t)) = L$ *and* $K_f(\Phi^{(i)}(L, K, t)) = K$,

(b) *the Jacobian matrix of* $\Theta_2 \circ \Phi_f^{(i)}$ *has determinant* $-\frac{1}{18}$.

For $t \in \mathbb{R}$, we shall use the notation

$$T^+(t) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \quad \text{and} \quad T^-(t) = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}, \tag{4-3}$$

which is an element of $O_{x^2+y^2}(\mathbb{R})$ and $O_{x^2-y^2}(\mathbb{R})$, respectively.

4A. ***Positive definite case.*** Define

$$\Phi : \Omega^+ \times [-\pi/4, \pi/4) \to V^0_{\mathbb{R},x^2+y^2}, \quad \Phi(L, K, t) = (F_{(L,K)})_{T^+(t)}, \tag{4-4}$$

where

$$F_{(L,K)}(x, y) = \frac{-3L + \sqrt{L^2+4K}}{24}x^4 + \frac{-L - \sqrt{L^2+4K}}{4}x^2y^2 + \frac{-3L + \sqrt{L^2+4K}}{24}y^4.$$

The image of Φ lies in $V_{\mathbb{R},x^2+y^2}$ by (3-1) and (1-8). Using Propositions 3.1 and 3.2(a), it is easy to check that Proposition 4.2(a) holds.

Now, by (3-1), an arbitrary $F \in V^0_{\mathbb{R},x^2+y^2}$ has the shape

$$F(x, y) = a_4x^4 + a_3x^3y + a_2x^2y^2 - a_3xy^3 + a_4y^4.$$

Write $L = L_{x^2+y^2}(F)$ and $K = K_{x^2+y^2}(F)$. Note that $(L, K) \in \Omega^+$ because $\Delta(F) > 0$ by (1-7). For $t \in \mathbb{R}$, a direct computation yields

$$F_{T^+(t)}(x, y) = A(t)x^4 + B(t)x^3y + C(t)x^2y^2 - B(t)xy^3 + A(t)y^4,$$

where

$$\begin{cases} A(t) = \dfrac{6a_4+a_2}{8} + \dfrac{2a_4-a_2}{8}\cos(4t) - \dfrac{a_3}{4}\sin(4t), \\ B(t) = a_3\cos(4t) + \dfrac{2a_4-a_2}{2}\sin(4t), \\ C(t) = \dfrac{6a_4+a_2}{4} - \dfrac{3(2a_4-a_2)}{4}\cos(4t) + \dfrac{3a_3}{2}\sin(4t). \end{cases}$$

It is not hard to show that there exists a unique $t_0 \in (-\pi/4, \pi/4]$ such that $B(t_0) = 0$ and $2A(t_0) - C(t_0) > 0$. Put $(A, C) = (A(t_0), C(t_0))$. Then, we have

$$(L, K) = \big(L_{x^2+y^2}(F_{T^+(t_0)}), K_{x^2+y^2}(F_{T^+(t_0)})\big) = (-6A - C, -2C(6A - C))$$

by Propositions 3.1 and 3.2(a). We solve that $F_{T^+(t_0)} = F_{(L,K)}$, or equivalently

$$F = (F_{(L,K)})_{T^+(-t_0)} = \Phi(L, K, -t_0).$$

Since $-t_0 \in [-\pi/4, \pi/4)$ is uniquely determined by F, this shows that Φ is a bijection.

Finally, the above calculation also yields

$$(\Theta_1 \circ \Phi)(L, K, t) = \big(\Phi_1(L, K, t), \Phi_2(L, K, t), \Phi_3(L, K, t)\big),$$

where

$$\begin{cases} \Phi_1(L, K, t) = -\dfrac{L}{8} + \dfrac{\sqrt{L^2 + 4K}}{24} \cos(4t), \\ \Phi_2(L, K, t) = \dfrac{\sqrt{L^2 + 4K}}{6} \sin(4t), \\ \Phi_3(L, K, t) = -\dfrac{L}{4} - \dfrac{\sqrt{L^2 + 4K}}{4} \cos(4t). \end{cases} \tag{4-5}$$

By a direct computation, we then see that Proposition 4.2(b) holds.

4B. ***Indefinite case.*** Define

$$\begin{cases} \Phi^{(i)} : \Omega^+ \times \mathbb{R} \to V^0_{\mathbb{R},x^2-y^2}, & \Phi^{(i)}(L, K, t) = (F^{(i)}_{(L,K)})_{T^-(t)} \quad \text{for } i = 1, 2, \\ \Phi^{(i)} : \Omega^- \times \mathbb{R} \to V^0_{\mathbb{R},x^2-y^2}, & \Phi^{(i)}(L, K, t) = (F^{(i)}_{(L,K)})_{T^-(t)} \quad \text{for } i = 3, 4, \end{cases} \tag{4-6}$$

where

$$F^{(i)}_{(L,K)}(x, y) = \frac{3L + (-1)^i\sqrt{L^2 + 4K}}{24} x^4 + \frac{-L + (-1)^i\sqrt{L^2 + 4K}}{4} x^2y^2 + \frac{3L + (-1)^i\sqrt{L^2 + 4K}}{24} y^4$$

for $i = 1, 2$, and

$$F^{(i)}_{(L,K)}(x, y) = \frac{(-1)^i\sqrt{2L^2 - K}}{3} x^3y - Lx^2y^2 + \frac{(-1)^i\sqrt{2L^2 - K}}{3} xy^3$$

for $i = 3, 4$. The images of $\Phi^{(1)}, \Phi^{(2)}, \Phi^{(3)}, \Phi^{(4)}$ lie in $V_{\mathbb{R},x^2-y^2}$ by (3-1) and (1-8). Using Propositions 3.1 and 3.2(a), it is easy to check that Proposition 4.3(a) holds.

Now, by (3-1), an arbitrary $F \in V^0_{\mathbb{R},x^2-y^2}$ has the shape

$$F(x, y) = a_4x^4 + a_3x^3y + a_2x^2y^2 + a_3xy^3 + a_4y^4.$$

Write $L = L_{x^2-y^2}(F)$ and $K = K_{x^2-y^2}(F)$. For $t \in \mathbb{R}$, a direct computation yields

$$F_{T^-(t)}(x, y) = A(t)x^4 + B(t)x^3y + C(t)x^2y^2 + B(t)xy^3 + A(t)y^4,$$

where

$$\begin{cases} A(t) = \dfrac{6a_4 - a_2}{8} + \dfrac{2a_4 + a_2}{8}\cosh(4t) + \dfrac{a_3}{4}\sinh(4t), \\ B(t) = a_3\cosh(4t) + \dfrac{2a_4 + a_2}{2}\sinh(4t), \\ C(t) = -\dfrac{6a_4 - a_2}{4} + \dfrac{3(2a_4 + a_2)}{4}\cosh(4t) + \dfrac{3a_3}{2}\sinh(4t). \end{cases}$$

Note that $\frac{d}{dt}A(t) = \frac{1}{2}B(t)$. It is not hard to check that:

- If $\Delta(F) > 0$, then there is a unique $t_0 \in \mathbb{R}$ such that $B(t_0) = 0$.
- If $\Delta(F) < 0$, then $B(t) \neq 0$ for all $t \in \mathbb{R}$, and there is a unique $t_0 \in \mathbb{R}$ such that $A(t_0) = 0$.

Put $(A, B, C) = (A(t_0), B(t_0), C(t_0))$. Then, we have

$$\begin{aligned} (L, K) &= (L_{x^2-y^2}(F_{T^-(t_0)}), K_{x^2-y^2}(F_{T^-(t_0)})) \\ &= \begin{cases} (6A - C, 2C(6A + C)) & \text{if } \Delta(F) > 0, \\ (-C, -9B^2 + 2C^2) & \text{if } \Delta(F) < 0. \end{cases} \end{aligned}$$

by Propositions 3.1 and 3.2(a). We solve that $F_{T^-(t_0)} = F^{(i)}_{(L,K)}$, or equivalently

$$F = (F^{(i)}_{(L,K)})_{T^-(-t_0)} = \Phi^{(i)}(L, K, -t_0), \quad \text{for exactly one } i \in \{1, 2, 3, 4\}.$$

Since t_0 is uniquely determined by F, this shows that $\Phi^{(1)}, \Phi^{(2)}, \Phi^{(3)}, \Phi^{(4)}$ are all injections, and that the stated disjoint union holds.

Finally, the above calculation also yields

$$(\Theta_1 \circ \Phi^{(i)})(L, K, t) = \big(\Phi^{(i)}_1(L, K, t), \Phi^{(i)}_2(L, K, t), \Phi^{(i)}_3(L, K, t)\big),$$

where

$$\begin{cases} \Phi^{(i)}_1(L, K, t) = \dfrac{L}{8} + \dfrac{(-1)^i\sqrt{L^2+4K}}{24}\cosh(4t), \\ \Phi^{(i)}_2(L, K, t) = \dfrac{(-1)^i\sqrt{L^2+4K}}{6}\sinh(4t), \\ \Phi^{(i)}_3(L, K, t) = -\dfrac{L}{4} + \dfrac{(-1)^i\sqrt{L^2+4K}}{4}\cosh(4t), \end{cases} \tag{4-7}$$

for $i = 1, 2$, and

$$\text{(4-8)}\qquad \begin{cases} \Phi_1^{(i)}(L, K, t) = \dfrac{L}{8} - \dfrac{L}{8}\cosh(4t) + \dfrac{(-1)^i\sqrt{2L^2-K}}{12}\sinh(4t), \\ \Phi_2^{(i)}(L, K, t) = \dfrac{(-1)^i\sqrt{2L^2-K}}{3}\cosh(4t) - \dfrac{L}{2}\sinh(4t), \\ \Phi_3^{(i)}(L, K, t) = -\dfrac{L}{4} - \dfrac{3L}{4}\cosh(4t) + \dfrac{(-1)^i\sqrt{2L^2-K}}{2}\sinh(4t), \end{cases}$$

for $i = 3, 4$. By a direct computation, we then see that Proposition 4.3(b) holds.

4C. ***Reducible case.*** Suppose $\gamma = 0$. For $t \in \mathbb{R}$, put

$$T(t) = \begin{pmatrix} e^{-t} & 0 \\ \dfrac{2\alpha\sinh t}{\beta} & e^t \end{pmatrix},$$

which is an element of $O_f(\mathbb{R})$. Define

$$\text{(4-9)}\qquad \Phi_f^{(i)} : \Omega^0 \times \mathbb{R} \to V^0_{\mathbb{R},f}, \quad \Phi_f^{(i)}(L, K, t) = (F^{(i)}_{f,(L,K)})_{T(t)} \quad \text{for } i = 1, 2,$$

where

$$F^{(i)}_{f,(L,K)}(x, y) = \left(\frac{L^2 + (-1)^i 72\alpha^2 L + 4K + 144\alpha^4}{(-1)^i 144\beta^2}\right)x^4 + \left(\frac{\alpha L + (-1)^i 4\alpha^3}{\beta}\right)x^3 y + \left(\frac{L + (-1)^i 12\alpha^2}{2}\right)x^2y^2 + (-1)^i 4\alpha\beta x y^3 + (-1)^i\beta^2 y^4.$$

The images of $\Phi_f^{(1)}, \Phi_f^{(2)}$ lie in $V_{\mathbb{R},f}$ by (3-2) and (1-8). Using Propositions 3.1 and 3.2(b), it is easy to check that Proposition 4.4(a) holds.

Now, by (3-2), an arbitrary $F \in V^0_{\mathbb{R},f}$ has the shape

$$\text{(4-10)}\quad F(x, y) = a_4 x^4 + \left(\frac{2\alpha(\beta^2 a_2 - 4\alpha^2 a_0)}{\beta^3}\right)x^3 y + a_2 x^2 y^2 + \left(\frac{4\alpha a_0}{\beta}\right)x y^3 + a_0 y^4.$$

Write $L = L_f(F)$ and $K = K_f(F)$. For $t \in \mathbb{R}$, a direct computation yields

$$F_{T(t)}(x, y) = A(t)x^4 + (*)x^3 y + B(t)x^2y^2 + (*)xy^3 + C(t)y^4,$$

where

$$\begin{cases} A(t) = e^{-4t}a_4 + \dfrac{\alpha^2}{\beta^2}(e^{4t} - 1)e^{-4t}a_2 + \dfrac{\alpha^4}{\beta^4}(e^{4t} - 1)(e^{4t} - 5)e^{-4t}a_0, \\ B(t) = a_2 + \dfrac{6\alpha^2}{\beta^2}(e^{4t} - 1)a_0, \\ C(t) = e^{4t}a_0. \end{cases}$$

Since $\Delta(F) \neq 0$, we have $(-1)^i a_0 > 0$ for a unique $i \in \{1, 2\}$, and there is a unique $t_0 \in \mathbb{R}$ such that $C(t_0) = (-1)^i \beta^2$. Put $(A, B) = (A(t_0), B(t_0))$. Then, we have

$$\begin{aligned}(L, K) &= (L_f(F_{T(t_0)}), K_f(F_{T(t_0)}))\\ &= \big(2B - (-1)^i 12\alpha^2, -B^2 + (-1)^i 36\beta^2 A - (-1)^i 24\alpha^2 B + 144\alpha^4\big),\end{aligned}$$

by Propositions 3.1 and 3.2(b). We solve that $F_{T(t_0)} = F^{(i)}_{f,(L,K)}$, or equivalently

$$F = (F^{(i)}_{f,(L,K)})_{T(-t_0)} = \Phi^{(i)}_f(L, K, -t_0).$$

Since t_0 and i are uniquely determined by F, this shows that $\Phi^{(1)}_f$ and $\Phi^{(2)}_f$ are both injections, and that the stated disjoint union holds.

Finally, the above calculation also yields

$$(\Theta_2 \circ \Phi^{(i)}_f)(L, K, t) = \big(\Phi^{(i)}_{f,1}(L, K, t), \Phi^{(i)}_{f,2}(L, K, t), \Phi^{(i)}_{f,3}(L, K, t)\big),$$

where

$$\begin{cases} \Phi^{(i)}_{f,1}(L, K, t) = \dfrac{(-1)^i e^{-4t}}{144\beta^2}(L^2 + 4K) + \dfrac{\alpha^2}{2\beta^2} L + \dfrac{(-1)^i \alpha^4 e^{4t}}{\beta^2}, \\ \Phi^{(i)}_{f,2}(L, K, t) = \dfrac{L}{2} + (-1)^i 6\alpha^2 e^{4t}, \\ \Phi^{(i)}_{f,3}(L, K, t) = (-1)^i \beta^2 e^{4t}. \end{cases} \tag{4-11}$$

By a direct computation, we then see that Proposition 4.4(b) holds.

5. Definition of a bounded semialgebraic set

Throughout this section, let $f(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ be an integral and primitive binary quadratic form with $\Delta(f) \neq 0$ and $\alpha > 0$, in the shape (1-11) whenever f is reducible. As we have already explained in Section 1A, the proof of Theorem 1.2 is reduced to counting points in the lattices in (1-10), which in turn amounts to certain volume computations, by the result below.

Proposition 5.1 (Davenport's lemma). *Let $\mathcal{R}$ be a bounded semialgebraic multiset in $\mathbb{R}^n$ having maximum multiplicity m and which is defined by at most k polynomial inequalities, each having degree at most ℓ. Then, the number of integral lattice points (counted with multiplicity) contained in the region $\mathcal{R}$ is*

$$\mathrm{Vol}(\mathcal{R}) + O(\max\{\mathrm{Vol}(\bar{\mathcal{R}}), 1\}),$$

where $\mathrm{Vol}(\bar{\mathcal{R}})$ *denotes the greatest d-dimensional volume of any projection of $\mathcal{R}$ onto a coordinate subspace by equating $n-d$ coordinates to zero, with $1 \leq d \leq n-1$. The implied constant in the second summand depends only on n, m, k, ℓ.*

Proof. This is a result of Davenport [1951a], and the above formulation is due to Bhargava and Shankar [2015, Proposition 2.6]. □

For $X > 0$, define

$$V^0_{\mathbb{R},f}(X) = \{F \in V^0_{\mathbb{R},f} : H_f(F) \le X\} \quad \text{and} \quad V^0_{\mathbb{Z},f}(X) = \{F \in V^0_{\mathbb{Z},f} : H_f(F) \le X\}.$$

However, to prove Theorem 1.2, we cannot apply Proposition 5.1 directly to

$$\Theta_{w(f)}(V^0_{\mathbb{R},f}(X)), \quad \text{where } w(f) = \begin{cases} 1 & \text{if } f \text{ is irreducible,} \\ 2 & \text{if } f \text{ is reducible,} \end{cases}$$

as in Section 1A, to count the lattice points in $\Theta_{w(f)}(V^0_{\mathbb{Z},f}(X)) \subset \Lambda_{f,w(f)}$ because

(1) the set $\Theta_{w(f)}(V^0_{\mathbb{R},f}(X))$ is unbounded when f is indefinite,

(2) distinct forms in $V^0_{\mathbb{Z},f}(X)$ might be $\mathrm{GL}_2(\mathbb{Z})$-equivalent.

Recall (4-1) and define

$$\Omega^*(X) = \{(L, K) \in \Omega^* : \max\{L^2, |K|\} \le X\} \quad \text{for } * \in \{0, +, -\}.$$

In the notation of Lemma 4.1 as well as Propositions 4.2, 4.3, and 4.4, we have

$$(5\text{-}1) \quad V^0_{\mathbb{R},f}(X) = \begin{cases} (\Psi_f \circ \Phi)(\Omega^+(X) \times [-\pi/4, \pi/4)), \\ \displaystyle\bigsqcup_{i=1}^{2} (\Psi_f \circ \Phi^{(i)})(\Omega^+(X) \times \mathbb{R}) \sqcup \bigsqcup_{i=3}^{4} (\Psi_f \circ \Phi^{(i)})(\Omega^-(X) \times \mathbb{R}), \\ \displaystyle\bigsqcup_{i=1}^{2} \Phi_f^{(i)}(\Omega^0(X) \times \mathbb{R}), \end{cases}$$

respectively, if f is positive definite, indefinite, and reducible. We shall overcome the two issues above by restricting the values for $t \in \mathbb{R}$.

For brevity, in this section, write

$$D_f = |\Delta(f)| \quad \text{and} \quad \delta_f = D_f/4,$$

as in Theorem 1.2 and Lemma 4.1, respectively.

Definition 5.2. If f is positive definite, define

$$\mathcal{S}_f(X) = (\Psi_f \circ \Phi)(\Omega^+(X) \times [-\pi/4, \pi/4)).$$

If f is reducible, define

$$\mathcal{S}_f(X) = \bigsqcup_{i=1}^{2} \Phi_f^{(i)}(\Omega^0(X) \times [t_{f,1}, t_{f,2}]) \quad \text{for } t_{f,1} = -\frac{\log 8}{4} \text{ and } t_{f,2} = \frac{\log(5X/18)}{4}.$$

If f is indefinite and irreducible, define

$$\mathcal{S}_f(X) = \bigsqcup_{i=1}^{2} (\Psi_f \circ \Phi^{(i)})(\Omega^+(X) \times [0, t_{D_f})) \sqcup \bigsqcup_{i=3}^{4} (\Psi_f \circ \Phi^{(i)})(\Omega^-(X) \times [0, t_{D_f})),$$

where t_{D_f} is defined as in Theorem 1.2(c).

The goal of this section to prove the following preliminary results and estimates:

Proposition 5.3. *The set $\Theta_{w(f)}(\mathcal{S}_f(X))$ is bounded, semialgebraic, and definable by an absolutely bounded number of polynomial inequalities whose degrees are absolutely bounded.*

Proposition 5.4. *The following statements hold.*

(a) *A form in $V^0_{\mathbb{Z},f}(X)$ is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to at least one form in $\mathcal{S}_f(X)$.*

(b) *A form in $V^0_{\mathbb{Z},f}(X)$ for which $\Delta(F) \neq \Box$ is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to exactly r_f forms in $\mathcal{S}_f(X)$, where r_f is defined as in Theorem 1.2.*

5A. *Alternative description.* First, we shall give an alternative description of the set $\mathcal{S}_f(X)$ in terms of the coefficients of the forms in $V^0_{\mathbb{R},f}(X)$.

Lemma 5.5. *If f is positive definite, then $\mathcal{S}_f(X) = V^0_{\mathbb{R},f}(X)$.*

Proof. This is clear from (5-1). □

Lemma 5.6. *If f is reducible, then*

$$\mathcal{S}_f(X) = \{F \in V^0_{\mathbb{R},f}(X) : \beta^2/8 \leq |C_F| \leq 5\beta^2 X/18\},$$

where C_F denotes the y^4-coefficient of F.

Proof. For $i = 1, 2$ and for any $F = \Phi^{(i)}_f(L, K, t)$, we have $C_F = (-1)^i \beta^2 e^{4t}$ by (4-11), and the claim is then clear from (5-1). □

Lemma 5.7. *If f is an indefinite and irreducible, then*

$$\mathcal{S}_f(X) = \{F \in V^0_{\mathbb{R},f}(X) : 1 \leq E_{f,1}(F) Z_f(F)/E_{f,2}(F) < e^{8t_{D_f}}\},$$

where in the notation of Proposition 3.2(a), we define

$$E_{f,1}(F) = L_{f,1}(F) - \sqrt{D_f}\, L_{f,2}(F) \quad \text{and} \quad E_{f,2}(F) = L_{f,1}(F) + \sqrt{D_f}\, L_{f,2}(F),$$

and for F in the image of $\Psi_f \circ \Phi^{(i)}$, we define

$$Z_f(F) = \begin{cases} 1 & \text{for } i = 1, 2, \\ \dfrac{L_f(F)^2 + 4K_f(F)}{(4L_f(F) - (-1)^i 2\sqrt{2L_f(F)^2 - K_f(F)})^2} & \text{for } i = 3, 4. \end{cases}$$

Proof. For $i = 1, 2, 3, 4$, consider $F = (\Psi_f \circ \Phi^{(i)})(L, K, t)$. For $k = 1, 2$, we have

$$E_{f,k}(F) = \begin{cases} (-1)^i 2\alpha^2 \sqrt{L_f(F)^2 + 4K_f(F)}\, e^{(-1)^{k+1}4t}/3 & \text{if } i = 1, 2, \\ -2\alpha^2 (3L_f(F) + (-1)^{k+i} 2\sqrt{2L_f(F)^2 - K_f(F)}) e^{(-1)^{k+1}4t}/3 & \text{if } i = 3, 4, \end{cases}$$

by a direct computation using (4-2), (4-7), and (4-8). We then see that

$$E_{f,1}(F) Z_f(F)/E_{f,2}(F) = e^{8t},$$

from which the claim follows. □

5B. ***Proof of Proposition 5.3.*** From (4-5), (4-7), (4-8), and (4-11), it is clear that the set $\mathcal{S}_f(X)$ is bounded. Thus, it remains to show that $\mathcal{S}_f(X)$ is a semialgebraic set definable by an absolutely bounded number of polynomial inequalities whose degrees are absolutely bounded.

5B1. *The case when f is positive definite or reducible.* The claim follows immediately from Lemmas 5.5 and 5.6 as well as Proposition 3.2.

5B2. *The case when f is indefinite and irreducible.* The only problem is that, for F in the image of $\Psi_f \circ \Phi^{(i)}$ for $i = 3, 4$, the expression $Z_f(F)$ is not a polynomial in the x^4, x^3y, and x^2y^2-coefficients of F. We shall resolve this issue in Lemma 5.8 below. The claim then follows from Lemma 5.7 and Proposition 3.2.

Lemma 5.8. *For $i = 3, 4$, let $F \in (\Psi_f \circ \Phi^{(i)})(\Omega^- \times \mathbb{R})$. Then, the condition*

$$1 \le E_{f,1}(F) Z_f(F)/E_{f,2}(F) < e^{8t_{D_f}}$$

is equivalent to an absolutely bounded number of polynomial inequalities in the variables $L_f(F)$, $K_f(F)$, $E_{f,1}(F)$, $E_{f,2}(F)$ whose degrees are absolutely bounded.

Proof. For brevity, define

$$Y_{f,1}(F) = -E_{f,1}(F)(L_f(F)^2 + 4K(F)) + E_{f,2}(F)(17L_f(F)^2 - 4K_f(F)),$$
$$Y_{f,2}(F) = -E_{f,1}(F)(L_f(F)^2 + 4K_f(F)) + e^{8t_{D_f}} E_{f,2}(F)(17L_f(F)^2 - 4K_f(F)),$$

as well as write

$$(L, K, E_1, E_2, Z, Y_1, Y_2) = (L_f(F), K_f(F), E_{f,1}(F), E_{f,2}(F), Z_f(F), Y_{f,1}(F), Y_{f,2}(F)).$$

Note that $L^2 + 4K < 0$ by (1-7) because $\Delta(F) < 0$. This implies that $Z < 0$ and so the stated condition may be rewritten as

$$\begin{cases} E_2 \le E_1 Z < e^{8t_{D_f}} E_2 & \text{if } E_2 > 0, \text{ which is equivalent to } i = 3, \\ E_2 \ge E_1 Z > e^{8t_{D_f}} E_2 & \text{if } E_2 < 0, \text{ which is equivalent to } i = 4. \end{cases}$$

By rearranging, we may further rewrite the above as

$$12E_2 L\sqrt{2L^2 - K} \le (-1)^i Y_1 \quad \text{and} \quad 12e^{8t_{D_f}} E_2 L\sqrt{2L^2 - K} > (-1)^i Y_2.$$

From here, we shall consider the different possibilities for the signs of E_2, L, Y_1, Y_2. For example, when $E_2 > 0$ and $L \ge 0$, the above is equivalent to $Y_1 \le 0$ and

$$\begin{cases} (12E_2L)^2(2L^2 - K) \le Y_1^2 & \text{if } Y_2 > 0, \\ (12E_2L)^2(2L^2 - K) \le Y_1^2 \text{ and } (12e^{8t_{D_f}} E_2 L)^2(2L^2 - K) > Y_2^2 & \text{if } Y_2 \le 0. \end{cases}$$

The other cases are analogous. We then see that the claim holds. □

5C. *Integral orthogonal groups.* We shall require an explicit description of

$$O_f(\mathbb{Z}) = O_f(\mathbb{R}) \cap \mathrm{GL}_2(\mathbb{Z}).$$

In the notation of Lemma 4.1, observe that

$$O_f(\mathbb{R}) = \begin{cases} T_f^{-1}(O_{x^2+y^2}(\mathbb{R}))T_f & \text{if } f \text{ is positive definite,} \\ T_f^{-1}(O_{x^2-y^2}(\mathbb{R}))T_f & \text{if } f \text{ is indefinite.} \end{cases} \tag{5-2}$$

Moreover, it is well-known that

$$O_{x^2+y^2}(\mathbb{R}) = \{J_k T^+(t) : k \in \{1,4\} \text{ and } t \in \mathbb{R}\},$$
$$O_{x^2-y^2}(\mathbb{R}) = \{\pm J_k T^-(t) : k \in \{1,2,3,4\} \text{ and } t \in \mathbb{R}\},$$

where $T^+(t)$ and $T^-(t)$ are defined as in (4-3), and

$$J_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad J_4 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{5-3}$$

We shall need the following lemma.

Lemma 5.9. *Suppose that $T \in O_f(\mathbb{Z}) \setminus \{\pm I_{2\times 2}\}$ has finite order. Then, the form f is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to a form of the shape*

$$\begin{cases} x^2+y^2,\ x^2+xy+y^2,\ \text{or } ax^2+bxy-ay^2 & \text{if } \det(T)=1, \\ xy,\ x^2-y^2,\ ax^2+cy^2,\ \text{or } ax^2+bxy+ay^2 & \text{if } \det(T)=-1, \end{cases}$$

for some integers a, b, and c.

Proof. By [Newman 1972, Chapter IX], for example, a finite cyclic subgroup of $\mathrm{GL}_2(\mathbb{Z})$ not contained in $\{\pm I_{2\times 2}\}$ is conjugate to the subgroup generated by one of the following:

$$\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

We then deduce that there exists $P \in \mathrm{GL}_2(\mathbb{Z})$ such that $Q = P^{-1}TP$ is equal to one of the following matrices up to sign:

$$\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Since f is primitive with $\alpha > 0$ by assumption and $(f_P)_Q = \pm f_P$, we then check that f_P must have one of the stated shapes. □

Proposition 5.10. *Suppose that f is positive definite. Then, we have*

$$O_f(\mathbb{Z}) = \{\pm I_{2\times 2}\}$$

if f is not $\mathrm{GL}_2(\mathbb{Z})$*-equivalent to the forms below, and the group* $O_f(\mathbb{Z})$ *is equal to*

$$\begin{cases} \{\pm I_{2\times 2}, \pm\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\} & \text{if } f(x,y) = x^2+y^2, \\ \{\pm I_{2\times 2}, \pm\left(\begin{smallmatrix} 1 & 1 \\ -1 & 0 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 1 & 1 \\ 0 & -1 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} -1 & 0 \\ 1 & 1 \end{smallmatrix}\right)\} & \\ & \text{if } f(x,y) = x^2+xy+y^2, \\ \{\pm I_{2\times 2}, \pm\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\} & \text{if } f(x,y) = \alpha x^2+\gamma y^2 \text{ for } \alpha \neq \gamma, \\ \{\pm I_{2\times 2}, \pm\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\} & \text{if } f(x,y) = \alpha x^2+\beta xy+\alpha y^2 \text{ for } \beta \notin \{0,\alpha\}. \end{cases}$$

Proof. Elements in $O_f(\mathbb{Z})$ have finite order by (5-2) and so the first claim follows from Lemma 5.9. Using (5-2), we compute that elements in $O_f(\mathbb{R})$ are of the forms

$$\begin{pmatrix} \phi_t + \dfrac{\beta\psi_t}{2\sqrt{\delta_f}} & \dfrac{\gamma\psi_t}{\sqrt{\delta_f}} \\ -\dfrac{\alpha\psi_t}{\sqrt{\delta_f}} & \phi_t - \dfrac{\beta\psi_t}{2\sqrt{\delta_f}} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \phi_t - \dfrac{\beta\psi_t}{2\sqrt{\delta_f}} & \dfrac{\beta}{\alpha}\left(\phi_t - \dfrac{\beta\psi_t}{2\sqrt{\delta_f}}\right) + \dfrac{\gamma\psi_t}{\sqrt{\delta_f}} \\ \dfrac{\alpha\psi_t}{\sqrt{\delta_f}} & -\phi_t - \dfrac{\beta\psi_t}{2\sqrt{\delta_f}} \end{pmatrix},$$

where $t \in \mathbb{R}$ and $(\phi_t, \psi_t) = (\cos t, \sin t)$. With the help of the proof of Lemma 5.9, it is not hard to check that $O_f(\mathbb{Z})$ is as claimed. □

Proposition 5.11. *Suppose that f is reducible. Then, the group $O_f(\mathbb{Z})$ is equal to*

$$\begin{cases} \{\pm I_{2\times 2}\} & \text{if } \beta \nmid \alpha^2+1 \text{ and } \beta \nmid \alpha^2-1, \\ \{\pm I_{2\times 2}, \pm\left(\begin{smallmatrix} \alpha & \beta \\ -(\alpha^2+1)/\beta & -\alpha \end{smallmatrix}\right)\} & \text{if } \beta \mid \alpha^2+1 \text{ and } \beta \nmid \alpha^2-1, \\ \{\pm I_{2\times 2}, \pm\left(\begin{smallmatrix} \alpha & \beta \\ -(\alpha^2-1)/\beta & -\alpha \end{smallmatrix}\right)\} & \text{if } \beta \nmid \alpha^2+1 \text{ and } \beta \mid \alpha^2-1, \\ \{\pm I_{2\times 2}, \pm\left(\begin{smallmatrix} -1 & 0 \\ 2 & 1 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 1 & 1 \\ -2 & -1 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 1 & 1 \\ 0 & -1 \end{smallmatrix}\right)\} & \text{if } f(x,y) = x^2+xy, \\ \{\pm I_{2\times 2}, \pm\left(\begin{smallmatrix} -1 & 0 \\ 1 & 1 \end{smallmatrix}\right), \pm\left(\begin{smallmatrix} 1 & 2 \\ -1 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 2 \\ 0 & -1 \end{smallmatrix}\right)\} & \text{if } f(x,y) = x^2+2xy. \end{cases}$$

Proof. Using (5-2), we compute that elements in $O_f(\mathbb{R})$ are of the forms

$$\pm\begin{pmatrix} \phi_t - \psi_t & 0 \\ 2\alpha\psi_t/\beta & \phi_t + \psi_t \end{pmatrix} \quad \text{and} \quad \pm\begin{pmatrix} \phi_t + \psi_t & (\beta/\alpha)(\phi_t + \psi_t) \\ -2\alpha\psi_t/\beta & -\phi_t - \psi_t \end{pmatrix},$$

where $t \in \mathbb{R}$ and $(\phi_t, \psi_t) \in \{(\cosh t, \sinh t), (\sinh t, \cosh t)\}$. For the matrix on the left to have integer entries, necessarily

$$2\cosh t, 2\sinh t \in \mathbb{Z}, \quad \text{so} \quad (2\cosh t, 2\sinh t) = (2, 0).$$

Similarly, for the matrix on the right to have integer entries, necessarily

$$2\alpha\cosh t, 2\alpha\sinh t, (\cosh t + \sinh t)/\alpha \in \mathbb{Z}, \qquad \text{so} \quad (2\alpha\cosh t, 2\alpha\sinh t) = (\alpha^2+1, \alpha^2-1).$$

We then deduce that

$$O_f(\mathbb{Z}) = \left\{ \pm I_{2\times 2}, \pm\left(\begin{smallmatrix} -1 & 0 \\ 2\alpha/\beta & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} \alpha & \beta \\ -(\alpha^2\pm 1)/\beta & -\alpha \end{smallmatrix}\right) \right\} \cap \mathrm{GL}_2(\mathbb{Z}).$$

Since f has the shape (1-11) by assumption, we have

$$\beta \mid \alpha^2+1 \text{ and } \beta \mid \alpha^2-1 \Longleftrightarrow \alpha = 1 \text{ and } \beta \in \{1, 2\},$$

and we see that the claim indeed holds. □

Proposition 5.12. *Suppose that f is indefinite and irreducible. Define*

$$G_f(\mathbb{Z}) = \{\pm T_{D_f}^n : n \in \mathbb{Z}\}, \quad \textit{where } T_{D_f} = \begin{pmatrix} \frac{1}{2}(u_{D_f} - \beta v_{D_f}) & -\gamma v_{D_f} \\ \alpha v_{D_f} & \frac{1}{2}(u_{D_f} + \beta v_{D_f}) \end{pmatrix}$$

and $(u_{D_f}, v_{D_f}) \in \mathbb{N}^2$ is the least solution to $x^2 - D_f y^2 = \pm 4$. Then, we have

$$O_f(\mathbb{Z}) = G_f(\mathbb{Z})$$

if f is not $\mathrm{GL}_2(\mathbb{Z})$-equivalent to the forms below, and the group $O_f(\mathbb{Z})$ is equal to

$$\begin{cases} G_f(\mathbb{Z}) \sqcup G_f(\mathbb{Z})\left(\begin{smallmatrix} 1 & \beta/\alpha \\ 0 & -1 \end{smallmatrix}\right) & \textit{if } f(x,y) = \alpha x^2 + \beta xy + \gamma y^2 \textit{ with } \alpha \mid \beta, \\ G_f(\mathbb{Z}) \sqcup G_f(\mathbb{Z})\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) & \textit{if } f(x,y) = \alpha x^2 + \beta xy - \alpha y^2. \end{cases}$$

Proof. By (5-2), elements in $O_f(\mathbb{R})$ of infinite order are of the shape

$$\pm \begin{pmatrix} \phi_t - \beta\psi_t/(2\sqrt{\delta_f}) & -\gamma\psi_t/\sqrt{\delta_f} \\ \alpha\psi_t/\sqrt{\delta_f} & \phi_t + \beta\psi_t/(2\sqrt{\delta_f}) \end{pmatrix},$$

where $t \in \mathbb{R}$ and $(\phi_t, \psi_t) \in \{(\cosh t, \sinh t), (\sinh t, \cosh t)\}$. We then see that

$$G_f(\mathbb{Z}) = \{\pm I_{2\times 2}\} \sqcup \{T \in O_f(\mathbb{Z}) : T \text{ has infinite order}\}.$$

Hence, the first claim follows from Lemma 5.9 and the fact that $ax^2 + bxy + ay^2$ is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to the form

$$(2a-b)x^2 + (2a-b)xy + ay^2 \quad \text{via } \left(\begin{smallmatrix} -1 & -1 \\ 1 & 0 \end{smallmatrix}\right). \tag{5-4}$$

Now, again by (5-2), elements in $O_f(\mathbb{R})$ of finite order have the shape

$$\begin{pmatrix} \dfrac{-\beta}{\sqrt{D_f}} & -\dfrac{2\gamma}{\sqrt{D_f}} \\ \dfrac{2\alpha}{\sqrt{D_f}} & \dfrac{\beta}{\sqrt{D_f}} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \phi_t + \dfrac{\beta\psi_t}{2\sqrt{\delta_f}} & \dfrac{\beta}{\alpha}\left(\phi_t + \dfrac{\beta\psi_t}{2\sqrt{\delta_f}}\right) - \dfrac{\gamma\psi_t}{\sqrt{\delta_f}} \\ -\dfrac{\alpha\psi_t}{\sqrt{\delta_f}} & -\phi_t - \dfrac{\beta\psi_t}{2\sqrt{\delta_f}}, \end{pmatrix} \tag{5-5}$$

where $t \in \mathbb{R}$ and $(\phi_t, \psi_t) \in \{(\cosh t, \sinh t), (\sinh t, \cosh t)\}$. Notice that the matrix on the left cannot lie in $\mathrm{GL}_2(\mathbb{Z})$ because D_f is not square when f is irreducible. Using the description of $O_{x^2-y^2}(\mathbb{R})$, it is then not hard to check that $[O_f(\mathbb{Z}) : G_f(\mathbb{Z})] \leq 2$, from which the second claim follows. □

5D. *Proof of Theorem 1.4.* Suppose that $f(x, y) = \alpha x^2 + \beta xy - \alpha y^2$ and that D_f is not a square. In the notation of Proposition 5.12, we have

$$x^2 - D_f y^2 = -4 \quad \text{has integer solutions if and only if} \quad \det(T_{D_f}) = -1$$

by definition. But Proposition 5.12 also implies that $\det(T_{D_f}) = -1$ is equivalent to

$$O_f(\mathbb{Z}) \text{ has an element of finite order and negative determinant.}$$

The theorem now follows from Lemma 5.9 and (5-4).

5E. *Proof of Proposition 5.4.* We shall need the following lemma.

Lemma 5.13. *For all* $F \in V^0_{\mathbb{Z},f}$ *with* $\Delta(F) \neq \square$ *and* $T \in \mathrm{GL}_2(\mathbb{Z}) \setminus \{\pm I_{2\times 2}\}$, *we have*

(a) $\quad F_T \in V^0_{\mathbb{Z},f}$ *if and only if* $T \in O_f(\mathbb{Z})$,

(b) $\quad F_T = F$ *if and only if* $T = \pm D_f^{-1/2} M_f$.

Proof. Note that $F_T \in V^0_{\mathbb{Z},f_T}$ by (1-8). By Theorem 1.1(a), we then have $F_T \in V^0_{\mathbb{Z},f}$ if and only if $f_T = \pm f$, whence part (a) holds. By Theorem 1.1(a) and Proposition 2.1, we have $F_T = F$ if and only if T is proportional to M_f, from which part (b) follows since $\det(T) = \pm 1$. □

5E1. *The case when f is positive definite or reducible.* Let us first observe that:

Lemma 5.14. *We have* $V^0_{\mathbb{Z},f}(X) \subset \mathcal{S}_f(X)$.

Proof. Let $F \in V^0_{\mathbb{Z},f}(X)$ be given. If f is positive definite, then clearly $F \in \mathcal{S}_f(X)$ by Lemma 5.5. If f is reducible, then recall Lemma 5.6, and we have $F \in \mathcal{S}_f(X)$ since

$$\frac{8C_F}{\beta^2} \in \mathbb{Z} \quad \text{and} \quad \left|\frac{8C_F}{\beta^2}\right| \leq \left|\frac{4(L_f(F)^2 + 4K_f(F))}{9}\right| \leq \frac{20X}{9}$$

by (4-10) and Proposition 3.2(b), respectively. □

Lemma 5.14 implies that part (a) holds. Together with Lemma 5.13(a), it further implies that for $F \in V^0_{\mathbb{Z},f}(X)$ with $\Delta(F) \neq \square$, the number of forms in $\mathcal{S}_f(X)$ which are $\mathrm{GL}_2(\mathbb{Z})$-equivalent to F is equal to

$$[O_f(\mathbb{Z}) : \mathrm{Stab}_{O_f(\mathbb{Z})}(F)].$$

By Lemma 5.13(b), we in turn have

$$[O_f(\mathbb{Z}) : \mathrm{Stab}_{O_f(\mathbb{Z})}(F)] = [O_f(\mathbb{Z}) : O_f(\mathbb{Z}) \cap \{\pm I_{2\times 2}, \pm D_f^{-1/2} M_f\}],$$

which may be verified to be equal to r_f using Propositions 5.10 and 5.11.

5E2. *The case when f is indefinite and irreducible.* We shall use the notation from Lemma 4.1, Proposition 5.12, (4-3), and (5-3). Then, by definition, we have

$$T_{D_f} = T_f^{-1} J_{k(f)} T^-(t_{D_f}) T_f, \quad \text{where } k(f) = \begin{cases} 1 & \text{if } u_{D_f}^2 - D_f v_{D_f}^2 = -4, \\ 2 & \text{if } u_{D_f}^2 - D_f v_{D_f}^2 = 4. \end{cases}$$

Now, by (5-1) and (4-6), a form in $V^0_{\mathbb{Z},f}(X)$ is of the shape

$$F = (F^{(i)}_{(L,K)})_{T^-(t)T_f}, \quad \text{where } (L, K, t) \in \Omega^0(X) \times \mathbb{R} \text{ and } i \in \{1, 2, 3, 4\}.$$

Observe that J_1 and J_2 commute with $T^-(t)$ as well as fix the forms in $V_{\mathbb{R},x^2-y^2}$. For any $n \in \mathbb{Z}$, we then deduce that

$$F_{T^n_{D_f}} = (F^{(i)}_{(L,K)})_{T^-(t) J^n_{k(f)} T^-(nt_{D_f}) T_f} = (F^{(i)}_{(L,K)})_{T^-(t+nt_{D_f}) T_f}.$$

Let $n_1 \in \mathbb{Z}$ be the unique integer such that $0 \le t + n_1 t_{D_f} < t_{D_f}$. The existence of n_1 then implies part (a).

Next, suppose that $\Delta(F) \ne \square$, in which case

$$\text{for } T \in \mathrm{GL}_2(\mathbb{Z}): \qquad F_T \in V^0_{\mathbb{Z},f} \quad \text{if and only if} \quad T \in O_f(\mathbb{Z})$$

by Lemma 5.13(a). If $O_f(\mathbb{Z}) = G_f(\mathbb{Z})$, then part (b) holds by the uniqueness of n_1. If $O_f(\mathbb{Z}) \ne G_f(\mathbb{Z})$, then recall from Proposition 5.12 that

$$O_f(\mathbb{Z}) = G_f(\mathbb{Z}) \sqcup G_f(\mathbb{Z}) M, \quad \text{where } M \text{ has finite order.}$$

From (5-2), we see that

$$M = \pm T_f^{-1} J_{k_0} T^-(t_0) T_f, \quad \text{where } t_0 \in \mathbb{R} \text{ and } k_0 \in \{3, 4\}.$$

Then, for any $n \in \mathbb{Z}$, it is straightforward to verify that

$$\begin{aligned} F_{T^n_{D_f} M} &= (F^{(i)}_{(L,K)})_{T^-(t+nt_{D_f}) J_{k_0} T^-(t_0) T_f} \\ &= \begin{cases} (F^{(i)}_{(L,K)})_{T^-(-(t+nt_{D_f})+t_0) T_f} & \text{for } i \in \{1, 2\}, \\ (F^{(j)}_{(L,K)})_{T^-(-(t+nt_{D_f})+t_0) T_f} & \text{for } i \in \{3, 4\}, \text{ where } j \in \{3, 4\} \setminus \{i\}. \end{cases} \end{aligned}$$

There is a unique $n_2 \in \mathbb{Z}$ such that $0 \le -(t + n_2 t_{D_f}) + t_0 < t_{D_f}$. Observe that

$$F_{T^{n_1}_{D_f}} = F_{T^{n_2}_{D_f} M} \quad \text{would imply} \quad F_{T^{n_1}_{D_f}} = (F_{T^{n_1}_{D_f}})_{T^{n_2-n_1}_{D_f} M}.$$

But $T^{n_2-n_1}_{D_f} M$ has finite order, and so it cannot proportional to M_f by (5-5), which is a contradiction by Lemma 5.13(b). Then, we conclude from Proposition 5.12 that part (b) indeed holds.

6. Error estimates and the main theorem

Throughout this section, let $f(x, y) = \alpha x^2 + \beta xy + \gamma y^2$ be an integral and primitive binary quadratic form with $\Delta(f) \neq 0$ and $\alpha > 0$, in the shape (1-11) whenever f is reducible. Let D_f, r_f and s_f be as in Theorem 1.2.

In Subsections 6A and 6B, respectively, we shall first prove:

Proposition 6.1. *For any $\epsilon > 0$, we have*

$$\#\left\{F \in \mathcal{S}_f(X) \cap V^0_{\mathbb{Z},f} : L_f(F)^2 + 4K_f(F) = \square\right\} = O_{f,\epsilon}(X^{1+\epsilon}),$$

and

$$\#\Big\{F \in \mathcal{S}_f(X) \cap V^0_{\mathbb{Z},f} : (L_f(F)^2 + 4K_f(F))(2L_f(F) - K_f(F))/\Delta(f) = \square \text{ and } L_f(F) \neq 0\Big\} = O_f(X^{1/2+\epsilon}).$$

Further, the number

$$\#\left\{F \in \mathcal{S}_f(X) \cap V^0_{\mathbb{Z},f} : -4K_f(F)^2/\Delta(f) = \square \text{ and } L_f(F) = 0\right\}$$

is equal to zero if $-\Delta(f) \neq \square$, and is bounded by $O_f(X)$ otherwise.

Propositions 6.1, 3.6, and 5.4 then imply part (d) of Theorem 1.2.

Proposition 6.2. *We have*

$$\#\{F \in \mathcal{S}_f(X) \cap V^0_{\mathbb{Z},f} : F \text{ is reducible}\} = \begin{cases} O_f(X(\log X)^2) & \text{if } f \text{ is irreducible},\\ O_f(X(\log X)^3) & \text{if } f \text{ is reducible}.\end{cases}$$

Now, from Propositions 5.4, 6.1, and 6.2, we also easily see that

$$N^{(D_4)}_{\mathbb{Z},f}(X) = \frac{1}{r_f}\#(\mathcal{S}_f(X) \cap V^0_{\mathbb{Z},f}) + O_{f,\epsilon}(X^{1+\epsilon}) \quad \text{for any } \epsilon > 0. \tag{6-1}$$

Let $\mathcal{L}_{f,w(f)}$ be a linear transformation on $\mathbb{R}^3$ which takes $\Lambda_{f,w(f)}$ to $\mathbb{Z}^3$, and define

$$\mathcal{R}_f(X) = (\mathcal{L}_{f,w(f)} \circ \Theta_{w(f)})(\mathcal{S}_f(X)), \quad \text{where } w(f) = \begin{cases} 1 & \text{if } f \text{ is irreducible},\\ 2 & \text{if } f \text{ is reducible},\end{cases}$$

as before. Observe that then

$$\#(\mathcal{S}_f(X) \cap V^0_{\mathbb{Z},f}) = \#(\Theta_{w(f)}(\mathcal{S}_f(X)) \cap \Lambda_{f,w(f)}) = \#(\mathcal{R}_f(X) \cap \mathbb{Z}^3).$$

By Proposition 5.3, we may apply Proposition 5.1 to obtain

$$\begin{aligned}(6\text{-}2)\quad &\#(S_f(X)\cap V^0_{\mathbb{Z},f})\\ &= \mathrm{Vol}(\mathcal{R}_f(X)) + O(\max\{\mathrm{Vol}(\overline{\mathcal{R}_f(X)}), 1\})\\ &= \frac{1}{\det(\Lambda_{f,w(f)})}\mathrm{Vol}(\Theta_{w(f)}(\mathcal{S}_f(X))) + O_f(\max\{\mathrm{Vol}(\overline{\Theta_{w(f)}(\mathcal{S}_f(X))}), 1\}),\end{aligned}$$

where by Proposition 3.4, we know that

$$\det(\Lambda_{f,w(f)}) = \begin{cases} s_f\alpha^3 & \text{if } f \text{ is irreducible,}\\ s_f\beta^3/8 & \text{if } f \text{ is reducible.}\end{cases}$$

Hence, it remains to compute the above volumes, which we shall do in Section 6C.

6A. ***Proof of Proposition 6.1.*** Recall the notation from Proposition 3.2. By definition and Proposition 3.3, we then have a well-defined map

$$\iota : V^0_{\mathbb{Z},f} \to \mathbb{Z}^3,\quad \iota(F) = (L_f(F), L_{f,1}(F), L_{f,2}(F)).$$

Using Proposition 3.2, it is easy to verify that ι is in fact injective. We shall also need the following result.

Lemma 6.3 [Heath-Brown 2002, Corollary 2]. *Let $\xi(x_1, x_2, x_3)$ be a ternary quadratic form such that its corresponding matrix M_ξ has nonzero determinant. For $B_1, B_2, B_3 > 0$, let $N_\xi(B_1, B_2, B_3)$ denote the number of tuples $(x_1, x_2, x_3) \in \mathbb{Z}^3$ such that*

$$|x_1| \le B_1,\quad |x_2| \le B_2,\quad |x_3| \le B_3,\quad \gcd(x_1, x_2, x_3) = 1,\quad \xi(x_1, x_2, x_3) = 0.$$

Then, we have

$$N_\xi(B_1, B_2, B_3) \ll_\epsilon \left(1 + \left(B_1B_2B_3\cdot\frac{\det_0(M_\xi)^2}{|\det(M_\xi)|}\right)^{1/3+\epsilon}\right) d_3(|\det(M_\xi)|),$$

where $\det_0(M_\xi)$ denotes the greatest common divisor of the 2×2 minors of M_ξ, and $d_3(|\det(M_\xi)|)$ is the number of ways to write $|\det(M_\xi)|$ as a product of three positive integers.

In what follows, consider $F \in \mathcal{S}_f(X)\cap V^0_{\mathbb{Z},f}$, and for brevity, write

$$(L, K, L_1, L_2) = (L_f(F), K_f(F), L_{f,1}(F), L_{f,2}(F)).$$

Since ι is injective, it is enough to estimate the number of choices for (L, L_1, L_2). To that end, let us put $\mathcal{D}_f = \Delta(f)$. Recall from Propositions 3.2 and 3.3 that

$$L, K, L_1, L_2 \in \mathbb{Z},\qquad \text{as well as}\qquad L_1^2 - \mathcal{D}_f L_2^2 = 4\alpha^4(L^2+4K)/9,$$

which is nonzero by (1-7). By the definition of our height, we also have

$$\text{(6-3)} \qquad \begin{cases} L = O_f(X^{1/2}) \quad \text{and} \quad K = O_f(X) & \text{in all cases,} \\ L_1 = O_f(X^{1/2}) \quad \text{and} \quad L_2 = O_f(X^{1/2}) & \text{if } f \text{ is irreducible.} \end{cases}$$

The latter estimate holds by

$$\begin{cases} \text{(4-5), (4-2)} & \text{if } f \text{ is positive definite,} \\ \text{(4-7), (4-8), (4-2), and } 0 \le t < t_{D_f} & \text{if } f \text{ is indefinite and irreducible,} \end{cases}$$

as well as the fact that L_1 and L_2 are linear in the coefficients of F. Finally, we shall write $d(-)$ for the divisor function.

Proof of Proposition 6.1: first claim. Suppose that $L^2 + 4K = \square$. Then, we have

$$L_1^2 - \mathcal{D}_f L_2^2 = U^2, \quad \text{where } U \in \mathbb{N} \text{ is such that } U = O_f(X^{1/2}).$$

If f is reducible, then $\mathcal{D}_f = \square$ and so clearly there are

$$O_f\left(\sum_{U=1}^{X^{1/2}} d(U^2)\right) = O_{f,\epsilon}\left(\sum_{U=1}^{X^{1/2}} X^{\epsilon}\right) = O_{f,\epsilon}(X^{1/2+\epsilon})$$

choices for the pair (L_1, L_2). If f is irreducible, then note that

$$(L_1/n)^2 - \mathcal{D}_f(L_2/n)^2 = (U/n)^2, \quad \text{where } n = \gcd(L_1, L_2, U),$$

and applying Lemma 6.3 to the ternary quadratic form ξ with matrix

$$M_\xi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\mathcal{D}_f & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \text{with } \begin{cases} \det(M_\xi) = \mathcal{D}_f, \\ \det_0(M_\xi) = 1, \end{cases}$$

we deduce from (6-3) that there are

$$O_f\left(\sum_{n=1}^{X^{1/2}} N_\xi\left(\frac{X^{1/2}}{n}, \frac{X^{1/2}}{n}, \frac{X^{1/2}}{n}\right)\right) = O_{f,\epsilon}\left(\sum_{n=1}^{X^{1/2}}\left(1 + \frac{X^{1/2+\epsilon}}{n^{1+\epsilon}}\right)\right) = O_{f,\epsilon}(X^{1/2+\epsilon})$$

choices for the pair (L_1, L_2). In both cases, we see that there are

$$O_f(X^{1/2}) \cdot O_{f,\epsilon}(X^{1/2+\epsilon}) = O_{f,\epsilon}(X^{1+\epsilon})$$

choices for (L, L_1, L_2) in total, whence the claim. □

Proof of Proposition 6.1: second claim. Suppose that $(L^2+4K)(2L^2-K)/\mathcal{D}_f = \square$. By Proposition 3.3, we may write

$$\gcd(L^2 + 4K, 4(2L^2 - K)/\mathcal{D}_f) = 9ma^2, \quad \text{where } m, a \in \mathbb{N} \text{ and } m \text{ is square-free.}$$

From the hypothesis, we then easily see that

$$L^2+4K=9mU^2 \quad \text{and} \quad 4(2L^2-K)/\mathcal{D}_f=9mV^2, \qquad \text{where } U, V\in\mathbb{N},$$

as well as that m divides L. In particular, a simple calculation yields

$$L^2=m(U^2+\mathcal{D}_fV^2), \quad \text{whence } mW^2=U^2+\mathcal{D}_fV^2, \text{ where } W\in\mathbb{Z} \text{ with } L=mW.$$

Now, suppose also that $L\neq 0$, in which case $m=O_f(X^{1/2})$ by (6-3). Note also that

$$m(W/n)^2=(U/n)^2+\mathcal{D}_f(V/n)^2, \quad \text{where } n=\gcd(W,U,V).$$

Applying Lemma 6.3 to the ternary quadratic form ξ_m with matrix

$$M_{\xi_m}=\begin{pmatrix} m & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -\mathcal{D}_f\end{pmatrix}, \quad \text{with } \begin{cases}\det(M_{\xi_m})=m\mathcal{D}_f,\\ \det_0(M_{\xi_m})=\gcd(m,\mathcal{D}_f)\le|\mathcal{D}_f|,\end{cases}$$

we then see from (6-3) that there are

$$\begin{aligned} O_f\Bigl(\sum_{n=1}^{X^{1/2}/m} N_{\xi_m}\Bigl(\frac{X^{1/2}}{mn},\frac{X^{1/2}}{m^{1/2}n},\frac{X^{1/2}}{m^{1/2}n}\Bigr)\Bigr) &= O_{f,\epsilon}\Bigl(\sum_{n=1}^{X^{1/2}/m}\Bigl(1+\frac{X^{1/2+\epsilon}}{(mn)^{1+\epsilon}}\Bigr)m^{\epsilon}\Bigr)\\ &= O_{f,\epsilon}\Bigl(\frac{X^{1/2}}{m^{1-\epsilon}}+\frac{X^{1/2+\epsilon}}{m}\Bigr)\end{aligned}$$

choices for (x,u,v) when m is fixed. It follows that we have

$$O_{f,\epsilon}\Bigl(\sum_{m=1}^{X^{1/2}}\Bigl(\frac{X^{1/2}}{m^{1-\epsilon}}+\frac{X^{1/2+\epsilon}}{m}\Bigr)\Bigr)=O_{f,\epsilon}(X^{1/2+\epsilon})$$

choices for (m,x,u,v) and hence for (L,K).

Next, regard (L,K) as being fixed, and recall that

$$L_1^2-\mathcal{D}_fL_2^2=T, \quad \text{where } T=4\alpha^4(L^2+4K)/9.$$

We claim that there are $O_f(d(T))$ choices for (L_1,L_2). If f is positive definite or if f is reducible, then this is clear. If f is indefinite and irreducible, then by Definition 5.2 as well as Propositions 3.1 and 4.3, we have

$$F=(\Psi_f\circ\Phi^{(i)})(L,K,t), \quad \text{where } 0\le t<t_{D_f} \text{ and } i\in\{1,2,3,4\}.$$

Since $\mathcal{D}_f>0$, we must have $L^2+4K>0$ by the hypothesis, and so in fact $i\in\{1,2\}$. From the proof of Lemma 5.7, we know that

$$L_1-\sqrt{D_f}L_2=(-1)^i\sqrt{T}e^{4t} \quad \text{and} \quad L_1+\sqrt{D_f}L_2=(-1)^i\sqrt{T}e^{-4t},$$

which implies that

$$L_1 = (-1)^i \sqrt{T} \cosh(4t) \quad \text{and} \quad L_2 = (-1)^i \sqrt{T} \sinh(4t)/\sqrt{D_f}.$$

Since $t = O_f(1)$, we then deduce that indeed there are $O_f(d(T))$ choices for (L_1, L_2). Using the bound $d(T) = O_\epsilon(T^\epsilon) = O_{f,\epsilon}(X^\epsilon)$, we conclude that there are

$$O_{f,\epsilon}(X^{1/2+\epsilon}) \cdot O_{f,\epsilon}(X^\epsilon) = O_{f,\epsilon}(X^{1/2+\epsilon})$$

choices for (L, L_1, L_2) in total, whence the claim. □

Proof of Proposition 6.1: third claim. Suppose that $L = 0$ and that F is in the shape as in (3-1). Using Proposition 3.2, we then deduce that

$$C = (-12\gamma A + 3\beta B)/(2\alpha), \quad \text{and so} \quad K = -9\mathcal{D}_f(\alpha B^2 - 4\beta AB + 16\gamma A^2)/(4\alpha^3).$$

Clearly $-4K2/\Delta(f) = \square$ if and only if $-\Delta(f) = \square$.

We now suppose that $-\Delta(f) = \square$, so in particular f is positive definite. The form F is then determined by $(A, B) \in \mathbb{Z}^2$, and that $|K| \le X$ implies

$$\left| \left(B - \frac{2\beta}{\alpha} A \right)^2 - \frac{4\mathcal{D}_f}{\alpha^2} A^2 \right| \ll_f X.$$

Hence there are $O_f(X)$ choices for (A, B). It follows that the claim holds. □

6B. *Proof of Proposition 6.2.* By Lemma 3.8 and Proposition 6.1, we have

$$\#\{F \in \mathcal{S}_f(X) \cap V^0_{\mathbb{Z},f} : F \text{ is reducible of type } 1\} = O_{f,\epsilon}(X^{1+\epsilon}), \tag{6-4}$$

whence it is enough to consider the reducible forms in $\mathcal{S}_f(X) \cap V^0_{\mathbb{Z},f}$ of type 2; recall Definition 3.7. By definition, such a form has the shape

$$F(x, y) = p_2q_2x^4 + (p_2q_1 + p_1q_2)x^3y + (p_2q_0 + p_1q_1 + p_0q_2)x^2y^2 + (*)xy^3 + (*)y^4,$$

where $p_2, p_1, p_0, q_2, q_1, q_0 \in \mathbb{Z}$, and we have

$$p_0 = (\beta p_1 - 2\gamma p_2)/(2\alpha) \quad \text{and} \quad q_0 = (\beta q_1 - 2\gamma q_2)/(2\alpha)$$

by Lemma 3.10. We have the condition

$$|(\alpha p_1^2 - 2\beta p_1 p_2 + 4\gamma p_2^2)/\alpha|, |(\alpha q_1^2 - 2\beta q_1 q_2 + 4\gamma q_2^2)/\alpha|, |p_2|, \\ |\alpha p_1 - \beta p_2|, |q_2|, |\alpha q_1 - \beta q_2| \ge 1 \tag{6-5}$$

since the above numbers are all integers. Using Proposition 3.2(a), we compute that

$$\frac{L_f(F)^2 + 4K_f(F)}{9} = \frac{\alpha p_1^2 - 2\beta p_1 p_2 + 4\gamma p_2^2}{\alpha} \cdot \frac{\alpha q_1^2 - 2\beta q_1 q_2 + 4\gamma q_2^2}{\alpha}.$$

Now, by the definition of our height, we clearly have

$$|(\alpha p_1^2 - 2\beta p_1 p_2 + 4\gamma p_2^2)/\alpha|, \ |(\alpha q_1^2 - 2\beta q_1 q_2 + 4\gamma q_2^2)/\alpha| \le X. \tag{6-6}$$

Observe also that

(6-7) $\quad p_2q_2,\ p_2q_1+p_1q_2,\ p_1q_1 = O_f(X^{1/2}) \quad$ if f is indefinite and irreducible

by (4-7), (4-8), (4-2), and the bound $0 \le t < t_{D_f}$. We then deduce that

(6-8) $$\#\{F \in \mathcal{S}_f(X)\cap V^0_{\mathbb{Z},f} : F \text{ is reducible of type 2}\} \le \#(\mathcal{R}'_f(X)\cap \mathbb{Z}^4),$$

where we define

$$\mathcal{R}'_f(X) = \{(p_2,p_1,q_2,q_1)\in\mathbb{R}^4 : \text{(6-5), (6-6), and (6-7)}\}.$$

It is clear that this set is bounded and semialgebraic. Hence, we may apply Proposition 5.1 to estimate the number of integral points it contains.

6B1. *The case when f is irreducible.* Let us define

$$\mathcal{R}''_f(X) = \mathcal{L}_{D_f}(\mathbb{R}'_f(X)), \quad \text{where } \mathcal{L}_{D_f} = \begin{pmatrix} \sqrt{D_f} & 0 & 0 & 0 \\ -\beta & \alpha & 0 & 0 \\ 0 & 0 & \sqrt{D_f} & 0 \\ 0 & 0 & -\beta & \alpha \end{pmatrix}.$$

Applying Proposition 5.1, we then obtain

$$\begin{aligned} \#(\mathcal{R}'_f(X)\cap\mathbb{Z}^4) &= \mathrm{Vol}(\mathcal{R}'_f(X)) + O(\max\{\mathrm{Vol}(\overline{\mathcal{R}_f(X)}, 1\}) \\ &= \frac{1}{\det(\mathcal{L}_{D_f})}\mathrm{Vol}(\mathcal{R}''_f(X)) + O_f(\max\{\mathrm{Vol}(\overline{\mathcal{R}''_f(X)}), 1\}) \end{aligned}$$

For any $(u_2,u_1,v_2,v_1)\in\mathcal{R}''_f(X)$, from (6-5) and (6-6), we deduce that

$$|u_2|, |u_1|, |v_2|, |v_1| \ge 1$$

as well as that

(6-9) $$\begin{cases} 1 \le |u_1^2+u_2^2|, |v_1^2+v_2^2| \le \alpha^4 X & \text{if } f \text{ is positive definite,} \\ 1 \le |u_1^2-u_2^2|, |v_1^2-v_2^2| \le \alpha^4 X & \text{if } f \text{ is indefinite.} \end{cases}$$

This, together with (6-7), implies that in fact

$$1 \le |u_2|, |u_1|, |v_2|, |v_1|, |u_2v_2|, |u_1v_1| \ll_f X^{1/2}.$$

We then compute that

$$\begin{aligned} \mathrm{Vol}(\mathcal{R}''_f(X)) &= O_f\left(\prod_{i=1}^{2}\int_1^{X^{1/2}/v_i} du_i\, dv_i\right) = O_f(X(\log X)^2), \\ \mathrm{Vol}(\overline{\mathcal{R}''_f(X)}) &= O_f(X\log X). \end{aligned}$$

The claim now follows from (6-4) and (6-8).

6B2. *The case when f is reducible.* Let us define

$$\mathbb{R}''_f(X) = \mathcal{L}_{0,D_f}(\mathbb{R}'_f(X)), \quad \text{where } \mathcal{L}_{0,D_f} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{D_f} & 0 & 0 & 0 \\ -\beta & \alpha & 0 & 0 \\ 0 & 0 & \sqrt{D_f} & 0 \\ 0 & 0 & -\beta & \alpha \end{pmatrix}.$$

Since $D_f = \square$ in this case, we see that

$$\mathcal{L}_{0,D_f}(\mathcal{R}'_f(X) \cap \mathbb{Z}^4) \subset \mathcal{R}''_f(X) \cap \mathbb{Z}^4 \quad \text{and so} \quad \#(\mathcal{R}'_f(X) \cap \mathbb{Z}^4) \le \#(\mathcal{R}''_f(X) \cap \mathbb{Z}^4).$$

Now, applying Proposition 5.1, we have

$$\#(\mathcal{R}''_f(X) \cap \mathbb{Z}^4) = \operatorname{Vol}(\mathcal{R}''_f(X)) + O(\max\{\operatorname{Vol}(\overline{\mathcal{R}''_f(X)}), 1\}).$$

For any $(z_1, z_2, z_3, z_4) \in \mathcal{R}''_f(X)$, the conditions (6-5) and (6-6) imply that

$$|z_1|, |z_2|, |z_3|, |z_4| \ge 1 \quad \text{and} \quad |z_1 z_2 z_3 z_4| \le \alpha^4 X,$$

which is analogous to (6-9). We then compute that

$$\begin{aligned} \operatorname{Vol}(\mathcal{R}''_f(X)) &= O_f\left(\int_1^X \int_1^{X/z_4} \int_1^{X/(z_3 z_4)} \int_1^{X/(z_2 z_3 z_4)} dz_1\, dz_2\, dz_3\, dz_4\right) \\ &= O_f(X(\log X)^3), \\ \operatorname{Vol}(\overline{\mathcal{R}''_f(X)}) &= O_f(X(\log X)^2). \end{aligned}$$

The claim now follows from (6-4) and (6-8).

6C. ***Proof of Theorem 1.2.*** We have already proven part (d). To prove parts (a) through (c), it remains to compute the volumes in (6-2).

6C1. *The case when f is positive definite.* We have

$$\operatorname{Vol}(\Theta_1(\mathcal{S}_f(X))) = \frac{8\alpha^3}{D_f^{3/2}} \cdot \frac{1}{18} \cdot \operatorname{Vol}(\Omega^+(X) \times [-\pi/4, \pi/4))$$

by Lemma 4.1 and Proposition 4.2(b), as well as

$$\operatorname{Vol}\big(\Omega^+(X) \times [-\pi/4, \pi/4)\big) = \int_{-X^{1/2}}^{X^{1/2}} \int_{-L^2/4}^{X} \frac{\pi}{2}\, dK\, dL = \frac{13\pi}{12} X^{3/2}.$$

Observe also that

$$\operatorname{Vol}(\overline{\Theta_1(\mathcal{S}_f(X))}) = O_f(X)$$

because $\Theta_1(\mathcal{S}_f(X))$ lies in the cube centered at the origin of side length $O_f(X^{1/2})$ by (4-5) and (4-2). We then deduce part (a) from (6-1) and (6-2).

6C2. *The case when f is reducible.* We have

$$\mathrm{Vol}(\Theta_2(\mathcal{S}_f(X))) = \tfrac{1}{18}\cdot 2\cdot \mathrm{Vol}(\Omega^0(X)\times[t_{f,1},t_{f,2}])$$

by Proposition 4.4, as well as

$$\mathrm{Vol}(\Omega^0(X)\times[t_{f,1},t_{f,2}]) = \int_{-X^{1/2}}^{X^{1/2}}\int_{-X}^{X}\frac{1}{4}\log\Big(\frac{20X}{9}\Big)\,dK\,dL = X^{3/2}\log(20X/9).$$

We then deduce part (b) from Lemma 6.4 below as well as (6-1) and (6-2).

Lemma 6.4. *We have* $\mathrm{Vol}(\overline{\Theta_2(\mathcal{S}_f(X))}) = O_f(X^{3/2})$.

Proof. By Definition 5.2, an element in $\Theta_2(\mathcal{S}_f(X))$ takes the form

$$(A,B,C) = (\Theta_2\circ\Phi_f)(L,K,t),\quad \text{where } (L,K,t)\in\Omega^0(X)\times[t_{f,1},t_{f,2}].$$

Let us recall that

$$|L|\le X^{1/2},\quad |K|\le X,\quad 4t_{f,1} = -\log 8,\quad 4t_{f,2} = \log(5X/18). \tag{6-10}$$

Then, from (4-11), we see that 1-dimensional projections of $\Theta_2(\mathcal{S}_f(X))$ have lengths of order $O_f(X)$. As for the 2-dimensional projections, note that (5-1) and (6-10) yield

$$|C| = \beta^2 e^{4t}\quad\text{and}\quad 1\ll_f |C|\ll_f X,$$

as well as the estimates

$$\Big|B-\frac{6\alpha^2 C}{\beta^2}\Big|\le \tfrac{1}{2}X^{1/2}\quad\text{and}\quad \Big|A-\frac{\alpha^4 C}{\beta^4}\Big|\le\frac{5}{144|C|}X+\frac{\alpha^2}{2\beta^2}X^{1/2}.$$

Hence, the projections of $\Theta_2(\mathcal{S}_f(X))$ onto the BC-plane and AC-plane, respectively, have areas bounded by

$$O_f\Big(\int_1^X X^{1/2}\,dC\Big)\quad\text{and}\quad O_f\Big(\int_1^X\Big(\frac{1}{C}X+X^{1/2}\Big)\,dC\Big).$$

Similarly, from (5-1) and (6-10), we deduce that

$$|2B-L| = 12\alpha^2 e^{4t},\quad 1\ll_f|2B-L|\ll_f X,\quad |B|\ll_f X,$$

as well as the estimate

$$\Big|A-\frac{\alpha^2 B}{6\beta^2}\Big|\le\frac{5\alpha^2}{12\beta^2}\Big(\frac{1}{|2B-L|}X+X^{1/2}\Big).$$

Note that $|L| \leq X^{1/2}$ also implies that

$$|2B - L| \geq |2|B| - |L|| \geq 2|B| - X^{1/2} \quad \text{when } |B| \geq X^{1/2}/2.$$

Hence, the projection of $\Theta_2(\mathcal{S}_f(X))$ onto the AB-plane has area bounded by

$$O_f\biggl(\int_0^{1+X^{1/2}/2} (X + X^{1/2})\, dB + \int_{1+X^{1/2}/2}^{X} \Bigl(\frac{1}{2B - X^{1/2}} X + X^{1/2}\Bigr)\, dB\biggr).$$

It follows that all of the 2-dimensional projections of $\Theta_2(\mathcal{S}_f(X))$ have areas of order $O_f(X^{3/2})$, and this proves the lemma. □

6C3. *The case when f is indefinite and irreducible.* We have

$$\mathrm{Vol}(\Theta_1(\mathcal{S}_f(X))) = \frac{8\alpha^3}{D_f^{3/2}} \cdot \frac{1}{18} \cdot 2 \cdot \bigl(\mathrm{Vol}(\Omega^+(X) \times [0, t_{D_f})) + \mathrm{Vol}(\Omega^-(X) \times [0, t_{D_f}))\bigr)$$

by Lemma 4.1 and Proposition 4.3, as well as

$$\mathrm{Vol}\bigl(\Omega^+(X) \times [0, t_{D_f})\bigr) = \int_{-X^{1/2}}^{X^{1/2}} \int_{-L^2/4}^{X} t_{D_f}\, dK\, dL = \frac{13 t_{D_f}}{6} X^{3/2},$$

$$\mathrm{Vol}\bigl(\Omega^-(X) \times [0, t_{D_f})\bigr) = \int_{-X^{1/2}}^{X^{1/2}} \int_{-X}^{-L^2/4} t_{D_f}\, dK\, dL = \frac{11 t_{D_f}}{6} X^{3/2},$$

Observe also that

$$\mathrm{Vol}(\overline{\Theta_1(\mathcal{S}_f(X))}) = O_f(X)$$

because $\Theta_1(\mathcal{S}_f(X))$ lies in the cube centered at the origin of side length $O_f(X^{1/2})$ by (4-7), (4-8), (4-2), and the bound on t. We then deduce part (c) from (6-1) and (6-2).

Acknowledgments

Tsang was partially supported by the China Postdoctoral Science Foundation Special Financial Grant (grant number: 2017T100060). We would like to thank the referee for many useful suggestions which helped improve the exposition of the paper significantly.

References

[Altuğ et al. 2017] S. A. Altuğ, A. Shankar, I. Varma, and K. H. Wilson, "The number of quartic D_4-fields ordered by conductor", preprint, 2017. arXiv

[Bhargava 2004] M. Bhargava, "Higher composition laws, III: The parametrization of quartic rings", *Ann. of Math.* (2) **159**:3 (2004), 1329–1360. MR Zbl

[Bhargava 2008] M. Bhargava, "Higher composition laws, IV: The parametrization of quintic rings", *Ann. of Math.* (2) **167**:1 (2008), 53–94. MR Zbl

[Bhargava and Shankar 2015] M. Bhargava and A. Shankar, "Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves", *Ann. of Math.* (2) **181**:1 (2015), 191–242. MR Zbl

[Bhargava and Shnidman 2014] M. Bhargava and A. Shnidman, "On the number of cubic orders of bounded discriminant having automorphism group C_3, and related problems", *Algebra Number Theory* **8**:1 (2014), 53–88. MR Zbl

[Bhargava et al. 2013] M. Bhargava, A. Shankar, and J. Tsimerman, "On the Davenport–Heilbronn theorems and second order terms", *Invent. Math.* **193**:2 (2013), 439–499. MR Zbl

[Birch and Merriman 1972] B. J. Birch and J. R. Merriman, "Finiteness theorems for binary forms with given discriminant", *Proc. London Math. Soc.* (3) **24**:3 (1972), 385–394. MR Zbl

[Brumer 1992] A. Brumer, "The average rank of elliptic curves, I", *Invent. Math.* **109**:1 (1992), 445–472. MR Zbl

[Conrad 2012] K. Conrad, "Galois groups of cubics and quartics (not in characteristic 2)", expository notes, 2012, http://www.math.uconn.edu/~kconrad/blurbs/galoistheory/cubicquartic.pdf.

[Cremona 1999] J. E. Cremona, "Reduction of binary cubic and quartic forms", *LMS J. Comput. Math.* **2** (1999), 64–94. MR Zbl

[Davenport 1951a] H. Davenport, "On a principle of Lipschitz", *J. London Math. Soc.* **26**:3 (1951), 179–183. MR Zbl

[Davenport 1951b] H. Davenport, "On the class-number of binary cubic forms, I", *J. London Math. Soc.* **26**:3 (1951), 183–192. MR Zbl

[Davenport 1951c] H. Davenport, "On the class-number of binary cubic forms, II", *J. London Math. Soc.* **26**:3 (1951), 192–198. MR Zbl

[Davenport and Heilbronn 1971] H. Davenport and H. Heilbronn, "On the density of discriminants of cubic fields, II", *Proc. Roy. Soc. London Ser. A* **322**:1551 (1971), 405–420. MR Zbl

[Delone and Faddeev 1964] B. N. Delone and D. K. Faddeev, *The theory of irrationalities of the third degree*, Trans. Math. Monographs **10**, Amer. Math. Soc., Providence, RI, 1964. MR Zbl

[Gauss 1801] C. F. Gauss, *Disquisitiones arithmeticae*, Fleischer, Leipzig, 1801.

[Heath-Brown 2002] D. R. Heath-Brown, "The density of rational points on curves and surfaces", *Ann. of Math.* (2) **155**:2 (2002), 553–598. MR Zbl

[Heath-Brown 2004] D. R. Heath-Brown, "The average analytic rank of elliptic curves", *Duke Math. J.* **122**:3 (2004), 591–623. MR Zbl

[Klagsbrun and Lemke Oliver 2014] Z. Klagsbrun and R. J. Lemke Oliver, "The distribution of the Tamagawa ratio in the family of elliptic curves with a two-torsion point", *Res. Math. Sci.* **1**:1 (2014), art. id. 15. MR Zbl

[Nakagawa 1989] J. Nakagawa, "Binary forms and orders of algebraic number fields", *Invent. Math.* **97**:2 (1989), 219–235. Correction in **105**:2 (1991), 443. MR Zbl

[Newman 1972] M. Newman, *Integral matrices*, Pure Appl. Math. **45**, Academic Press, New York, 1972. MR Zbl

[Siegel 1944] C. L. Siegel, "The average measure of quadratic forms with given determinant and signature", *Ann. of Math.* (2) **45**:4 (1944), 667–685. MR Zbl

[Taniguchi and Thorne 2013] T. Taniguchi and F. Thorne, "Secondary terms in counting functions for cubic fields", *Duke Math. J.* **162**:13 (2013), 2451–2508. MR Zbl

[Tsang and Xiao 2017] C. Tsang and S. Y. Xiao, "The number of quartic D_4-fields with monogenic cubic resolvent ordered by conductor", preprint, 2017. arXiv

[Wood 2012] M. M. Wood, "Quartic rings associated to binary quartic forms", *Int. Math. Res. Not.* **2012**:6 (2012), 1300–1320. MR Zbl

[Xiao 2019] S. Y. Xiao, "On binary cubic and quartic forms", *J. Théor. Nombres Bordeaux* **31**:2 (2019), 323–341.

[Young 2006] M. P. Young, "Low-lying zeros of families of elliptic curves", *J. Amer. Math. Soc.* **19**:1 (2006), 205–250. MR Zbl

Received February 27, 2018. Revised October 21, 2018.

CINDY (SIN YI) TSANG
YAU MATHEMATICAL SCIENCES CENTER
TSINGHUA UNIVERSITY
BEIJING
CHINA
Current address:
SCHOOL OF MATHEMATICS (ZHUHAI)
SUN YAT-SEN UNIVERSITY
TANGJIAWAN, ZHUHAI
GUANGDONG
CHINA
zengshy26@mail.sysu.edu.cn

STANLEY YAO XIAO
MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD
UNITED KINGDOM
Current address:
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
CANADA
syxiao@math.toronto.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.293

OBSTRUCTIONS TO LIFTING ABELIAN SUBALGEBRAS OF CORONA ALGEBRAS

ANDREA VACCARO

Let A be a noncommutative, nonunital C*-algebra. Given a set of commuting positive elements in the corona algebra $Q(A)$, we study some obstructions to the existence of a commutative lifting of such a set to the multiplier algebra $M(A)$. Our focus is on the obstructions caused by the size of the collection we want to lift. It is known that no obstacles show up when lifting a countable family of commuting projections, or of pairwise orthogonal positive elements. However, this is not the case for larger collections. We prove in fact that for every primitive, nonunital, σ-unital C*-algebra A, there exists an uncountable set of pairwise orthogonal positive elements in $Q(A)$ such that no uncountable subset of it can be lifted to a set of commuting elements of $M(A)$. Moreover, the positive elements in $Q(A)$ can be chosen to be projections if A has real rank zero.

1. Introduction

Let A be a nonunital C*-algebra, denote its multiplier algebra by $M(A)$, its corona algebra (namely $M(A)/A$) by $Q(A)$, and the quotient map from $M(A)$ onto $Q(A)$ by π. A *lifting* in $M(A)$ of a set $B \subseteq Q(A)$ is a set $C \subseteq M(A)$ such that $\pi[C] = B$. The study of which properties of $B \subseteq Q(A)$ can be preserved in a lifting, as well as the analysis of the relations between B and its preimage $\pi^{-1}[B]$, has produced a rich theory with strong connections to the study of stable relations in C*-algebras. A general introduction to this subject can be found in [Loring 1997].

This note focuses on liftings of abelian subalgebras of corona algebras. This topic has been widely studied, for instance, as a means to producing interesting examples of $*$-algebras, and in the investigation of the masas (maximal abelian subalgebras) of the Calkin algebra $Q(H)$. In [Akemann and Doner 1979], for example, the authors produce, by means of a lifting, a nonseparable C*-algebra whose abelian subalgebras are all separable. Their proof assumes the continuum hypothesis, which was later shown to not be necessary; see [Popa 1983, Corollary 6.7; Bice and Koszmider 2017]. Another application of the continuum hypothesis to liftings of

MSC2010: primary 47C15; secondary 03E75.
Keywords: corona algebra, commuting self-adjoint elements, lifting.

abelian subalgebras of corona algebras can be found in [Anderson 1979]. The author builds a masa of $Q(H)$ which is generated by its projections and does not lift to a masa in $B(H)$. In this case it is not known whether the continuum hypothesis can be dropped; see [Shelah and Steprāns 2011]. More recently, the study of liftings led to the first example of an amenable nonseparable Banach algebra which is not isomorphic to a C*-algebra; see [Choi et al. 2014; Vignati 2015].

In this paper we focus on the following problem. Let A be a noncommutative, nonunital C*-algebra, and let B be a commutative family in $Q(A)$. What kind of obstructions could prevent the existence of a commutative lifting of B in $M(A)$? We consider collections with various properties, but our main concern and focus is the role played by the cardinality of the set that we want to lift. The following table summarizes all the cases that we are going to analyze. The symbols "✓" and "×" indicate whether it is possible or not to have a lifting for collections on the left column whose size is the cardinal in the top line.

$Q(A) \to M(A)$	$< \aleph_0$	$\aleph_0$	$\aleph_1$
Commuting self-adjoint → Commuting self-adjoint	×	×	×
Commuting projections → Commuting projections	✓ in $Q(H)$	✓ in $Q(H)$	×
Commuting projections → Commuting positive	✓	✓	×
Orthogonal positive → Orthogonal positive	✓	✓	×
Orthogonal positive → Commuting positive	✓	✓	×

It is clear from the table that starting with an uncountable collection is a fatal obstruction. We also remark that the two columns in the middle, representing the lifting problem for finite and countable collections, have the same values. One reason for this phenomenon is that the only obstructions in this scenario are of K-theoretic nature and involve only a finite number of elements, as we shall see in the next paragraph; see also [Davidson 1985]. This situation also relates to other compactness phenomena (at least at the countable level) that corona algebras of σ-unital algebras satisfy, due to their partial countable saturation; see [Farah and Hart 2013]. Most of the results in the table about finite and countable families are already known; see [Farah and Wofsey 2012, Lemma 5.34; Loring 1997, Lemma 10.1.12]. The main contribution of this paper concerns the right column, for which some theorems about projections in the Calkin algebra have already been proved; see [Farah and Wofsey 2012, Theorem 5.35; Bice and Koszmider 2017].

Let A be $K(H)$, the algebra of the compact operators on a separable Hilbert space H, so that $M(A) = B(H)$ and $Q(A) = Q(H)$. By a well-known K-theoretic obstruction, the unilateral shift is a normal element in $Q(H)$ which does not necessarily lift to a normal element in $B(H)$ (more on this in [Brown et al. 1977] and [Davidson 2010]). An element is normal if and only if its real and imaginary part commute. This proves that it is not always possible to lift a couple of commuting self-adjoint elements in a corona algebra to commuting self-adjoint elements in the multiplier algebra.

In order to bypass this obstruction, we add some conditions to the collection which we want to lift. In [Farah and Wofsey 2012, Lemma 5.34] it is proved that any countable family of commuting projections in the Calkin algebra can be lifted to a family of commuting projections in $B(H)$. Moreover, the authors provide a lifting of simultaneously diagonalizable projections. Proving a more general statement about liftings, in Section 2 we show that any countable collection of commuting projections in a corona algebra can be lifted to a commutative family of positive elements in the multiplier algebra.[1]

Two elements in a C*-algebra are *orthogonal* if their product is zero. Any countable family of orthogonal positive elements in a corona algebra admits a commutative lifting. This is a consequence of the more general result [Loring 1997, Lemma 10.1.12], which is relayed in this paper as Proposition 2.2.

In general, we cannot expect to be able to generalize verbatim the above result for uncountable families of orthogonal positive elements. This is the case since, by a cardinality obstruction, a multiplier algebra $M(A)$ which can be faithfully represented on a separable Hilbert space H, cannot contain an uncountable collection of orthogonal positive elements. The existence of such a collection in $M(A)$ (and thus in $B(H)$) would in fact imply the existence of an uncountable set of orthogonal vectors in H, contradicting the separability of H.

We could still ask whether it is possible to lift an uncountable family of orthogonal positive elements to a family of commuting positive elements. This leads to an obstruction of set-theoretic nature. In Theorem 5.35 of [Farah and Wofsey 2012], it is shown that there exists an $\aleph_1$-sized collection of orthogonal projections in the Calkin algebra whose uncountable subsets cannot be lifted to families of simultaneously diagonalizable projections in $B(H)$. This result is refined in Theorem 7 of [Bice and Koszmider 2017], where the authors provide an $\aleph_1$-sized set of orthogonal projections in $Q(H)$ which contains no uncountable subset that lifts to a collection of commuting operators in $B(H)$. The main result of this paper is a generalization of this theorem. A C*-algebra is *σ-unital* if it has a countable approximate unit, and it is *primitive* if it admits a faithful irreducible representation.

Theorem 1.1. *Assume A is a primitive, nonunital, σ-unital* C*-*algebra. Then there is a collection of $\aleph_1$ pairwise orthogonal positive elements of $Q(A)$ containing no uncountable subset that simultaneously lifts to commuting elements in $M(A)$.*

[1] We remark that it is not always possible to lift projections in a corona algebra to projections in the multiplier algebra. Such lifting is not possible for instance when $Q(A)$ has real rank zero but $M(A)$ has not, which is the case for $A = Q(H) \otimes K(H)$ (see [Zhang 1992, Example 2.7(iii)]) or $A = \mathcal{Z} \otimes K(H)$, where $\mathcal{Z}$ is the Jiang–Su algebra (see [Lin and Ng 2016]).

Corollary 1.2. *Assume A is a primitive, real rank zero, nonunital, σ-unital* C*-*algebra. Then there is a collection of $\aleph_1$ pairwise orthogonal projections of $Q(A)$ containing no uncountable subset that simultaneously lifts to commuting elements in $M(A)$.*

The proof of Theorem 1.1 is inspired by the combinatorics used in [Bice and Koszmider 2017] and [Farah and Wofsey 2012], which goes back to Luzin and Hausdorff, and to the study of uncountable almost disjoint families of subsets of $\mathbb{N}$ and Luzin's families; see [Luzin 1947]. We remark that no additional set theoretic assumption (such as the continuum hypothesis) is required in our proof.

The paper is structured as follows: in Section 2 we outline the results needed to settle the problem of liftings of countable families of commuting projections and of orthogonal positive elements. Section 3 is devoted to the proof of Theorem 1.1, while concluding remarks and questions can be found in Section 4.

2. Countable collections

Denote the set of self-adjoint and of positive elements of a C*-algebra A by A_{sa} and A_+, respectively. Given a compact Hausdorff space X, $C(X)$ is the C*-algebra of the continuous functions from X into $\mathbb{C}$.

Farah and Wofsey [2012, Lemma 5.34] proved that any countable set of commuting projections in the Calkin algebra can be lifted to a set of simultaneously diagonalizable projections in $B(H)$. The thesis of the following proposition is weaker, but it holds in a more general context.

Proposition 2.1. *Let $\varphi : A \to B$ be a surjective $*$-homomorphism between two* C*-*algebras and let $\{p_n\}_{n\in\mathbb{N}}$ be a collection of commuting projections of B. Then there exists a set $\{q_n\}_{n\in\mathbb{N}}$ of commuting positive elements of A such that $\varphi(q_n) = p_n$.*

Proof. We can assume that both A and B are unital, that $\varphi(1_A) = 1_B$ and that $1_B \in \{p_n\}_{n\in\mathbb{N}}$. Let $C \subseteq B$ be the abelian C*-algebra generated by the set $\{p_n\}_{n\in\mathbb{N}}$. Consider the element

$$b = \sum_{n\in\mathbb{N}} \frac{2p_n - 1}{3^n}.$$

Let X be the spectrum of b in A. The algebra C is generated by b (see [Rickart 1960, p. 293] for a proof), thus $C \cong C(X)$. Fix $a \in A$ such that $\varphi(a) = b$. The element $(a + a^*)/2$ is still in the preimage of b since b is self-adjoint, thus we can assume $a \in A_{sa}$. If Y is the spectrum of a, we have in general that $X \subseteq Y$. Fix $f_n \in C(X)_+$ such that $f_n(b) = p_n$. Since the range of f_n is contained in $[0, 1]$ and the spaces Y and X are compact and Hausdorff, by the Tietze extension theorem [Willard 1970, Theorem 15.8], for every $n \in \mathbb{N}$, there is a continuous $F_n : Y \to [0, 1]$ such that $F_n \restriction_X = f_n$. Set $q_n = F_n(a)$. The map φ acts on $C(Y)$ as the restriction on X

(here we identify $C^*(a)$ and $C^*(b)$ with $C(Y)$ and $C(X)$ respectively), therefore $\varphi(q_n) = p_n$ for every $n \in \mathbb{N}$. □

The q_n's can be chosen to be projections if there is a self-adjoint a in the preimage of b whose spectrum is X. By the Weyl–von Nuemann theorem, this is the case when φ is the quotient map from $B(H)$ onto the Calkin algebra; see [Davidson 1996, Theorem II.4.4].

We focus now on lifting sets of positive orthogonal elements, starting with a set of size two. Let therefore $\varphi : A \to B$ be a surjective $*$-homomorphism of C^*-algebras, and let $b_1, b_2 \in B_+$ be such that $b_1 b_2 = 0$. Consider the self-adjoint $b = b_1 - b_2$ and let $a \in A_{sa}$ be such that $\varphi(a) = b$. The positive and the negative part of a are two orthogonal positive elements of A such that

$$\varphi(a_+) = b_1, \quad \varphi(a_-) = b_2.$$

The situation is analogous when dealing with countable collections:

Proposition 2.2 [Loring 1997, Lemma 10.1.12]. *Assume $\varphi : A \to B$ is a surjective $*$-homomorphism between two C^*-algebras. Let $\{b_n\}_{n\in\mathbb{N}}$ be a collection of orthogonal positive elements in B. Then there exists a set $\{a_n\}_{n\in\mathbb{N}}$ of orthogonal positive elements in A such that $\varphi(a_n) = b_n$.*

3. Uncountable collections

Throughout this section, let A be a primitive, nonunital, σ-unital C^*-algebra. We can thus assume that A is a noncommutative strongly dense C^*-subalgebra of $B(H)$ for a certain Hilbert space H. A sequence of operators $\{x_n\}_{n\in\mathbb{N}}$ *strictly converges* to $x \in B(H)$ if and only if $x_n a \to xa$ and $a x_n \to ax$ in norm for all $a \in A$. In this scenario $M(A)$ can be identified with the idealizer

$$\{x \in B(H) : xA \subseteq A,\ Ax \subseteq A\}$$

or with the strict closure of A in $B(H)$. Given two elements a, b in a C^*-algebra A, we denote the commutator $ab - ba$ by $[a, b]$. From now on, let $(e_n)_{n\in\mathbb{N}}$ be an approximate unit of A such that:

(1) $e_0 = 0$;

(2) $\|e_i - e_j\| = 1$ for $i \neq j$;

(3) $e_i e_j = e_i$ for every $i < j$.

Such an approximate unit exists since A is σ-unital, as proved in Section 2 of [Pedersen 1990].

The proof of Theorem 1.1 follows closely the one given by Bice and Koszmider [2017, Theorem 7], and a lemma similar to their Lemma 6 is required.

Lemma 3.1. *Let A be a primitive, nonunital, σ-unital* C*-*algebra. There exists a family $(a_\beta)_{\beta\in\aleph_1} \subseteq M(A)_+ \setminus A$ such that*:

(1) $\|a_\beta\| = 1$ *for all* $\beta \in \aleph_1$;

(2) $a_\alpha a_\beta \in A$ *for all distinct* $\alpha, \beta \in \aleph_1$;

(3) *given* $d_1, d_2 \in M(A)$, *for all* $\beta \in \aleph_1$, *all* $n \in \mathbb{N}$, *and all but finitely many* $\alpha < \beta$:

$$\|[(a_\alpha + d_1 e_n), (a_\beta + d_2 e_n)]\| \geq \tfrac{1}{8}.$$

The rough idea to prove this lemma is to build, for every $\beta < \aleph_1$, a strictly increasing function $f_\beta : \mathbb{N} \to \mathbb{N}$ and a norm-bounded sequence $\{c_k^\beta\}_{k\in\mathbb{N}} \subseteq A_+$ to define

$$a_\beta = \sum_{k\in\mathbb{N}} (e_{f_\beta(2k+1)} - e_{f_\beta(2k)})^{\frac{1}{2}} c_k^\beta (e_{f_\beta(2k+1)} - e_{f_\beta(2k)})^{\frac{1}{2}}.$$

Note that this series belongs to $M(A)$ by Theorem 4.1 in [Pedersen 1990] (see also [Farah and Hart 2013, Item (10) p. 48]). In order to satisfy the thesis of the lemma, we will build each c_k^β so that, for some $\alpha < \beta$ and some $n \in \mathbb{N}$, the following holds

$$\|[(a_\alpha + e_n), (c_k^\beta + e_n)]\| \geq \tfrac{1}{8}.$$

The choice of f_β will guarantee orthogonality in $Q(A)$ exploiting, for $n_2 < n_1 < m_2 < m_1$, the following fact:

$$(e_{m_1} - e_{m_2})(e_{n_1} - e_{n_2}) = 0.$$

The main ingredient used to build c_k^β is Kadison's transitivity theorem, which we are allowed to use since A is primitive.

Proof of Lemma 3.1. Since the C*-algebra A is primitive, we can assume that there is a Hilbert space H such that $A \subseteq B(H)$ and A acts irreducibly on H. For each $n < m$, denote the space $\overline{(e_m - e_n)H}$ by $S_{n,m}$. We start by building a_0. Let $f : \mathbb{N} \to \mathbb{N}$ be defined as follows:

$$f(n) = \begin{cases} 2^{n+1} - 1 & \text{if } n \text{ is even,} \\ 2^n & \text{if } n \text{ is odd.} \end{cases}$$

For every $k \in \mathbb{N}$ there is a unit vector ξ in the range of $e_{f(2k+1)} - e_{f(2k)}$. By the definition of the approximate unit $(e_n)_{n\in\mathbb{N}}$, the vector ξ is a 1-eigenvector of $e_{f(2k+2)}$. This, along with the (algebraic) irreducibility of $A \subseteq B(H)$, entails that

$$AS_{f(2k+1),f(2k)} = H.$$

Denote the algebra $\overline{(e_{f(2k+1)} - e_{f(2k)})A(e_{f(2k+1)} - e_{f(2k)})}$ by A_k. We have that

$$A_k H \supseteq S_{f(2k),f(2k+1)}.$$

Let $\xi_k^0, \eta_k^0 \in S_{f(2k), f(2k+1)}$ be two orthogonal[2] norm one vectors. Since A acts irreducibly on H and A_k is a hereditary subalgebra of A, it follows that A_k acts irreducibly on $B(A_k H)$; see [Murphy 1990, Theorem 5.5.2]. Therefore, by Kadison's transitivity theorem, we can find a self-adjoint $c_k^0 \in A_k$ such that

$$c_k^0(\xi_k^0) = \xi_k^0,$$
$$c_k^0(\eta_k^0) = 0,$$

and $\|c_k^0\| = 1$. We can suppose that c_k^0 is positive by taking its square, doing so will not change its norm nor the image of ξ_k^0 and η_k^0. Consider the function

$$f_0(n) = \begin{cases} f(n) - 1 & \text{if } n \text{ is even,} \\ f(n) + 1 & \text{if } n \text{ is odd.} \end{cases}$$

We have that

$$e_{f_0(2k+1)} c_k^0 = c_k^0 e_{f_0(2k+1)} = c_k^0,$$
$$e_{f_0(2k)} c_k^0 = c_k^0 e_{f_0(2k)} = 0.$$

This entails

$$(e_{f_0(2k+1)} - e_{f_0(2k)}) c_k^0 = c_k^0 = c_k^0 (e_{f_0(2k+1)} - e_{f_0(2k)})$$

and therefore also

$$c_k^0 = (e_{f_0(2k+1)} - e_{f_0(2k)})^{1/2} c_k^0 (e_{f_0(2k+1)} - e_{f_0(2k)})^{1/2}.$$

The norm $\|c_k^0\|$ is bounded by 1 for every $k \in \mathbb{N}$, therefore the sum

$$a_0 = \sum_{k \in \mathbb{N}} c_k^0 = \sum_{k \in \mathbb{N}} (e_{f_0(2k+1)} - e_{f_0(2k)})^{1/2} c_k^0 (e_{f_0(2k+1)} - e_{f_0(2k)})^{1/2}$$

is strictly convergent (see [Pedersen 1990, Theorem 4.1] or [Farah and Hart 2013, Item (10) p. 48]), hence $a_0 \in M(A)_+$. Furthermore:

$$\begin{aligned} \|a_0\| &= \Big\| \sum_{k \in \mathbb{N}} (e_{f_0(2k+1)} - e_{f_0(2k)})^{1/2} c_k^0 (e_{f_0(2k+1)} - e_{f_0(2k)})^{1/2} \Big\| \\ &\le \Big\| \sum_{k \in \mathbb{N}} e_{f_0(2k+1)} - e_{f_0(2k)} \Big\| \le 1. \end{aligned}$$

In order to show that $a_0 \notin A$, first observe that

$$a_0(\xi_k^0) = \sum_{m<k} c_m^0(\xi_k^0) + c_k^0(\xi_k^0) + \sum_{m>k} c_m^0(\xi_k^0) = c_k^0(\xi_k^0) = \xi_k^0.$$

[2] We can always assume $S_{n,n+1}$ has at least 2 linearly independent vectors for each $n \in \mathbb{N}$ by taking, if necessary, a subsequence $(e_{k_j})_{j \in \mathbb{N}}$ from the original approximate unit.

The first sum is zero since $\xi_k^0 \in S_{f(2k),f(2k+1)}$ implies $\xi_k^0 = (e_{f_0(2k+1)} - e_{f_0(2k)})(\xi_k^0)$, and for $m < k$

$$c_m^0(e_{f_0(2k+1)} - e_{f_0(2k)})(\xi_k^0) = c_m^0 e_{f_0(2m+1)}(e_{f_0(2k+1)} - e_{f_0(2k)})(\xi_k^0) = 0,$$

which follows by $f_0(2m+1) < f_0(2k) < f_0(2k+1)$. The second series is also zero, indeed for $m > k$ we have

$$c_m^0 e_{f_0(2k+1)} = c_m^0 e_{f_0(2m)} e_{f_0(2k+1)} = 0$$

(the same equation also holds for $e_{f_0(2k)}$). Using the same argument, it can be proved that

$$a_0(\xi) = c_n^0(\xi)$$

for every $\xi \in S_{f_0(2n), f_0(2n+1)}$. Observe that $\|(a_0 - e_{f_0(2m+1)} a_0)(\xi_k^0)\| = 1$ for $k > m$, thus $a_0 \notin A$.

The construction proceeds by transfinite induction on $\aleph_1$, the first uncountable cardinal. At step $\beta < \aleph_1$ we assume we have a sequence of elements $(a_\alpha)_{\alpha<\beta}$ in $M(A)_+$ and functions $(f_\alpha)_{\alpha<\beta}$ such that:

(i) For all $\alpha < \beta$ the function $f_\alpha : \mathbb{N} \to \mathbb{N}$ is strictly increasing and, given any other $\gamma < \alpha$, for all $k \in \mathbb{N}$ there exists $N \in \mathbb{N}$ such that for all $j > N$ and all $i \in \mathbb{N}$ the following holds

$$|f_\alpha(j) - f_\gamma(i)| > 2^k.$$

Furthermore, we ask that for all $\alpha < \beta$ and all $k \in \mathbb{N}$:

$$f_\alpha(2(k+1)) - f_\alpha(2k+1) > 2^{2k+1}.$$

(ii) For each $\alpha < \beta$ there exists a sequence $(c_k^\alpha)_{k\in\mathbb{N}}$ of positive norm 1 elements in A such that

$$a_\alpha = \sum_{k\in\mathbb{N}} c_k^\alpha.$$

Moreover we require that

$$e_{f_\alpha(2k+1)} c_k^\alpha = c_k^\alpha e_{f_\alpha(2k+1)} = c_k^\alpha,$$
$$e_{f_\alpha(2k)} c_k^\alpha = c_k^\alpha e_{f_\alpha(2k)} = 0,$$

and that there exist $\xi_k^\alpha, \eta_k^\alpha \in S_{f_\alpha(2k), f_\alpha(2k+1)}$, two norm one orthogonal vectors, such that $c_k^\alpha(\xi_k^\alpha) = \xi_k^\alpha$ and $c_k^\alpha(\eta_k^\alpha) = 0$.

(iii) Given $\alpha < \beta$ and $d_1, d_2 \in M(A)$, for all $l \in \mathbb{N}$, and for all but possibly l many $\gamma < \alpha$ the following holds:

$$\|[(a_\alpha + d_1 e_l), (a_\gamma + d_2 e_l)]\| \geq \tfrac{1}{2}.$$

It can be shown, as we already did for a_0, that for all $\alpha < \beta$:

(a) $a_\alpha \in M(A)_+ \setminus A$;

(b) $\|a_\alpha\| = 1$;

(c) $a_\alpha(\xi) = c_k^\alpha(\xi) \in S_{f_\alpha(2k), f_\alpha(2k+1)}$ for every $\xi \in S_{f_\alpha(2k), f_\alpha(2k+1)}$.

Moreover, by items (i)–(ii), along with the fact that for $n_2 < n_1 < m_2 < m_1$

$$(e_{m_1} - e_{m_2})(e_{n_1} - e_{n_2}) = 0,$$

we have that $a_\alpha a_\gamma \in A$ for all $\alpha, \gamma < \beta$.

We want to find f_β and a_β such that the families $\{a_\alpha\}_{\alpha<\beta+1}$ and $\{f_\alpha\}_{\alpha<\beta+1}$ satisfy the three inductive hypotheses. This will be sufficient to continue the induction and to obtain the thesis of the lemma. Since β is a countable ordinal, the sequence $(a_\alpha)_{\alpha<\beta}$ is either finite or can be written as $(a_{\alpha_n})_{n<\mathbb{N}}$, where $n \mapsto \alpha_n$ is a bijection between $\mathbb{N}$ and β. We assume that β is infinite, since the finite case is easier. In order to ease the notation, we shall denote a_{α_n} by a_n (and similarly f_{α_n} by f_n, $c_{\alpha_n}^k$ by c_n^k, etc.).

The construction of a_β proceeds inductively on the set $\{(i, j) \in \mathbb{N} \times \mathbb{N} : i \le j\}$ ordered along with any well-ordering of type ω such that $(i, j) \le (i', j')$ implies $j \le j'$, like for example

$$(i, j) \le (i', j') \iff j \le j' \quad \text{or} \quad j = j', i \le i'.$$

Suppose we are at step M, which corresponds to a certain couple (i, j). At step M we provide a $c_M^\beta \in A_+$ such that, for every $d_1, d_2 \in M(A)$

$$\|[(a_j + d_1 e_i), (c_M^\beta + d_2 e_i)]\| \ge \tfrac{1}{2}$$

and we define two values of f_β. Assume that $f_\beta(n)$ has been defined for $n \le 2M-1$. Let $m \in \mathbb{N}$ be the smallest natural number such that

$$f_j(2m) > \max\{i + 2, f_\beta(2M-1) + 2^{2M-1} + 1\}$$

and such that, for $l \ge 2m$, the inequality $|f_j(l) - f_k(n)| > 2^M + 1$ holds for all $k \in \mathbb{N}$ such that $\alpha_k < \alpha_j$, and all $n \in \mathbb{N}$. By inductive hypothesis there are two norm one orthogonal vectors $\xi_m^j, \eta_m^j \in S_{f_j(2m), f_j(2m+1)}$ such that $c_m^j(\xi_m^j) = \xi_m^j$ and $c_m^j(\eta_m^j) = 0$. Set $\xi_M^\beta = \frac{1}{\sqrt{2}}(\xi_j^m + \eta_j^m)$ and $\eta_M^\beta = \frac{1}{\sqrt{2}}(\xi_j^m - \eta_j^m)$. Using Kadison's transitivity theorem, fix a positive, norm one element

$$c_M^\beta \in \overline{(e_{f_j(2m+1)} - e_{f_j(2m)})A(e_{f_j(2m+1)} - e_{f_j(2m)})}$$

such that

$$c_M^\beta(\xi_M^\beta) = \xi_M^\beta,$$
$$c_M^\beta(\eta_M^\beta) = 0.$$

Let $f_\beta(2M) = f_j(2m) - 1$ and $f_\beta(2M+1) = f_j(2m+1) + 1$. We have therefore that

$$e_{f_\beta(2M+1)} c_M^\beta e_{f_\beta(2M+1)} = c_M^\beta,$$
$$e_{f_\beta(2M)} c_M^\beta = c_M^\beta e_{f_\beta(2M)} = 0.$$

Moreover:

$$(*)\quad \|(a_j + d_1 e_i)(c_M^\beta + d_2 e_i)(\xi_M^\beta) - (c_M^\beta + d_2 e_i)(a_j + d_1 e_i)(\xi_M^\beta)\| \\ = \|a_j c_M^\beta(\xi_M^\beta) - c_M^\beta a_j(\xi_M^\beta)\| = \tfrac{1}{2\sqrt{2}}\|\xi_j^m - \eta_j^m\| = \tfrac{1}{2}.$$

This is the case since $e_i(\xi) = 0$ for every $\xi \in S_{f_j(2m), f_j(2m+1)}$ (we chose m so that $f_j(2m) > i+2$) and $c_M^\beta(\xi_M^\beta), a_j(\xi_M^\beta) = c_m^j(\xi_M^\beta) \in S_{f_j(2m), f_j(2m+1)}$. Define

$$a_\beta = \sum_{n\in\mathbb{N}} c_n^\beta = \sum_{n\in\mathbb{N}} (e_{f_\beta(2n+1)} - e_{f_\beta(2n)})^{\frac{1}{2}} c_n^\beta (e_{f_\beta(2n+1)} - e_{f_\beta(2n)})^{\frac{1}{2}}.$$

This series is strictly convergent because all c_n^β's have norm 1. The families $\{f_n\}_{n<\mathbb{N}} \cup \{f_\beta\}$ and $\{a_n\}_{n<\mathbb{N}} \cup \{a_\beta\}$ satisfy items (i)–(ii) of the inductive hypothesis.[3]

Finally we verify clause (iii). Notice that, by construction, for every $k \in \mathbb{N}$, given $\xi \in S_{f_\beta(2k), f_\beta(2k+1)}$ we have

$$a_\beta(\xi) = c_k^\beta(\xi).$$

Let $i \le j \in \mathbb{N}$, denote the step corresponding to the couple (i, j) by M, and let $m \in \mathbb{N}$ be such that $f_\beta(2M) = f_j(2m) - 1$ (by construction we can find such m). Remember that $\xi_M^\beta = \frac{1}{\sqrt{2}}(\xi_j^m + \eta_j^m) \in S_{f_\beta(2M), f_\beta(2M+1)}$. Given $d_1, d_2 \in M(A)$, we have that

$$\|(a_j + d_1 e_i)(a_\beta + d_2 e_i)(\xi_M^\beta) - (a_\beta + d_2 e_i)(a_j + d_1 e_i)(\xi_M^\beta)\| \\ = \|a_j a_\beta(\xi_M^\beta) - a_\beta a_j(\xi_M^\beta)\| = \tfrac{1}{2\sqrt{2}}\|\xi_j^m - \eta_j^m\| = \tfrac{1}{2}.$$

This equation can be shown using the same arguments used to prove $(*)$.

Notice that if β is finite, we only obtain a finite number of c_n^β, therefore their sum (which is finite) does not belong to $M(A) \setminus A$. In this case it is sufficient to add an infinite number of addends, as we did for a_0. Suppose that β is (the ordinal corresponding to) $N \in \mathbb{N}$, then the previous construction defines f_N only up until $2N+1$. Let $f_N(2(N+1))$ be the smallest integer such that

- $f_N(2(N+1)) - f_N(2N+1) > 2^{2N+1}$;
- $|f_N(2(N+1)) - f_j(n)| > 2^{2(N+1)}$ for all $j < N$, and for all $n \in \mathbb{N}$.

[3] The induction to define a_β and f_β is on the set $\{(i, j) \in \mathbb{N}\times\mathbb{N} : i \le j\}$ ordered with a well-ordering of type ω such that $(i, j) \le (i', j')$ implies $j \le j'$. This is used to show that f_β satisfies clause (i) of the inductive hypothesis.

Define $f_N(2(N+1)+1) = f_N(2(N+1))+3$ and continue inductively the definition of f_N. For each $n > N$ we can therefore, as we did for a_0 using Kadison's transitivity theorem, find a positive element

$$c_n^N \in \overline{(e_{f_N(2n+1)-1} - e_{f_N(2n)+1})A(e_{f_N(2n+1)-1} - e_{f_N(2n)+1})}$$

which moves a norm one vector $\xi_n^N \in S_{f_N(2n), f_N(2n+1)}$ into itself, and another orthogonal norm one vector η_n^N to zero. If we define a_N to be the sum of such c_n^N's, it is possible to show, using the same arguments exposed when β was assumed to be infinite, that the families $\{f_n\}_{n<\mathbb{N}} \cup \{f_\beta\}$ and $\{a_n\}_{n<N+1}$ satisfy items (i)–(iii) of the inductive hypothesis. □

The proof of Theorem 1.1 is analogous to the one given in Theorem 7 of [Bice and Koszmider 2017], but it uses our Lemma 3.1 instead of Lemma 6 of the same work.

Proof of Theorem 1.1. Let $(e_n)_{n\in\mathbb{N}} \subseteq A$ be the approximate unit defined at the beginning of this section, and let $(a_\beta)_{\beta\in\aleph_1}$ be the $\aleph_1$-sized collection obtained from Lemma 3.1. Suppose there is an uncountable $U \subseteq \aleph_1$ and $(d_\beta)_{\beta\in U} \subseteq A$ such that

$$[(a_\alpha + d_\alpha), (a_\beta + d_\beta)] = 0$$

for all $\alpha, \beta \in U$. By using the pigeonhole principle, we can suppose that $\|d_\beta\| \le M$ for some $M \in \mathbb{R}$, and that there is a unique $n \in \mathbb{N}$ such that $\|d_\beta - d_\beta e_n\| \le 1/(64(M+1))$ for all $\beta \in U$.

Therefore, for every $\beta \in U$ and all but finitely many $\alpha \in U$ such that $\alpha < \beta$, we have

$$0 = \|[(a_\alpha + d_\alpha), (a_\beta + d_\beta)]\| \ge \|[(a_\alpha + d_\alpha e_n), (a_\beta + d_\beta e_n)]\| - \tfrac{1}{16} \ge \tfrac{1}{16}.$$

This is a contradiction when $\{\alpha \in U : \alpha < \beta\}$ is infinite. □

Proof of Corollary 1.2. The proof follows verbatim the one given for Lemma 3.1 and Theorem 1.1. The only difference is that each time Kadison's transitivity theorem is invoked in Lemma 3.1, it is possible to use a stronger version of this theorem for C*-algebras with real rank zero (see for instance Theorem 6.5 of [Bice 2013]), which allows us to choose a projection at each step. This stronger version of Kadison's transitivity theorem can be used throughout the whole iteration since hereditary subalgebras of real rank zero C*-algebras have real rank zero. □

4. Concluding remarks and questions

If A is a commutative nonunital C*-algebra, then the problem of lifting commuting elements from $Q(A)$ to $M(A)$ is trivial, as both $M(A)$ and $Q(A)$ are abelian. In Section 3 we ruled out this possibility by asking for A to be primitive.

The other important feature we required to prove Theorem 1.1 is σ-unitality. We do not know whether this assumption could be weakened, but it certainly cannot be removed tout court. Indeed, there are extreme examples of primitive, non-σ-unital C*-algebras whose corona is finite-dimensional (see [Sakai 1971; Ghasemi and Koszmider 2018]), for which Theorem 1.1 is trivially false. Our conjecture is that there might be a condition on the order structure of the approximate unit of A which is weaker than σ-unitality, but still makes Theorem 1.1 true. For instance, it would be interesting to know whether the techniques used in Theorem 1.1 could be applied to the algebra of the compact operators on a nonseparable Hilbert space, or more generally to a C*-algebra A with a projection $p \in M(A)$ such that pAp is primitive, nonunital and σ-unital.

We remark that the proof of Theorem 1.1 we gave can be adapted to any primitive C*-algebra A which admits an increasing approximate unit $\{e_\alpha\}_{\alpha \in \kappa}$, for κ regular cardinal, to produce a κ^+-sized family of orthogonal positive elements in $Q(A)$ which cannot be lifted to a set of commuting elements in $M(A)$.

Another lifting problem that we want to discuss is the following:

Question 4.1. Assume $F \subseteq Q(A)_{sa}$ is a commutative family such that any smaller (in the sense of cardinality) subset can be lifted to a set of commuting elements in $M(A)_{sa}$. Can F be lifted to a collection of commuting elements in $M(A)_{sa}$?

Theorem 1.1 and Proposition 2.2 entail that this is not true in general for primitive, nonunital, σ-unital C*-algebras if $|F| = \aleph_1$, pointing out the set theoretic incompactness of $\aleph_1$ for this property.

If the family F is infinite and countable, then Question 4.1 has a positive answer in the Calkin algebra.

Proposition 4.2. *Suppose that A is a separable abelian* C*-*subalgebra of $Q(H)$ such that every finitely generated subalgebra of A has an abelian lift. Then A has an abelian lift.*

The proof of this proposition relies on Voiculescu's theorem [Higson and Roe 2000, Theorem 3.4.6], starting from the following lemma. Given a map $\varphi : A \to Q(H)$, we say that $\Phi : A \to B(H)$ *lifts* φ if $\varphi = \pi \circ \Phi$, where $\pi : B(H) \to Q(H)$ is the quotient map.

Lemma 4.3. *Let A be a separable unital abelian* C*-*subalgebra of $Q(H)$. If there exists a unital abelian* C*-*algebra $B \subseteq B(H)$ lifting A, then there is a unital $*$-homomorphism $\Phi : A \to B(H)$ lifting the identity map on A.*

Proof. Since B is abelian, there exists a masa (maximal abelian subalgebra) of $B(H)$ containing B. Masas in $B(H)$ are von Neumann algebras and, as such, they are generated by their projections. This entails that A is contained in a separable unital abelian subalgebra $C(Y)$ of $Q(H)$ which is generated by its projections. By [Brown et al. 1977, Theorem 1.15] there exists a unital $*$-homomorphism $\Psi : C(Y) \to B(H)$ lifting the identity on $C(Y)$. Let Φ be the restriction of Ψ to $C(X)$. □

Proof of Proposition 4.2. Suppose that $F = \{a_n\}_{n\in\mathbb{N}} \subseteq Q(H)_{sa}$ is an abelian family such that every finite subset of F has a commutative lift. Without loss of generality, we can assume that $a_0 = 1$. By Lemma 4.3 we can assume that, for every $k \in \mathbb{N}$, there is a unital $*$-homomorphism $\Phi_k : \mathrm{C}^*(\{a_n\}_{n\le k}) \to B(H)$ lifting the identity map on $\mathrm{C}^*(\{a_n\}_{n\le k})$. By Voiculescu's theorem [Higson and Roe 2000, Theorem 3.4.6] we can moreover assume that, for every $n \in \mathbb{N}$, the sequence $\{\Phi_k(a_n)\}_{k\ge n}$ converges to some self-adjoint operator A_n in $B(H)$ such that $A_n - \Phi_k(a_n)$ is compact for every $k \in \mathbb{N}$. The family $\{A_n\}_{n\in\mathbb{N}}$ is a commutative lifting of $\{a_n\}_{n\in\mathbb{N}}$. □

More general forms of Voiculescu's theorem are known to hold for extensions of various separable C*-algebras other than $K(H)$; see [Elliott and Kucerovsky 2001; Gabe 2016; Schafhauser 2018, Section 2.2]. Such generalizations could potentially be used to carry out the arguments exposed above for coronas of other separable nuclear stable C*-algebras. We note, however, the importance of being able to lift separable abelian subalgebras of $Q(H)$ to abelian algebras in $B(H)$ with the same spectrum, as guaranteed by Lemma 4.3. This is false in general in other coronas, as it happens for instance when $A = \mathcal{Z} \otimes K(H)$. In this case, projections in $Q(A)$ do not necessarily lift to projections in $M(A)$, since the former has real rank zero but the latter has not; see [Lin and Ng 2016].

The following example shows that Question 4.1 has a negative answer for finite families with an even number of elements.

Example 4.4. Let S^n be the n-dimensional sphere. The algebra $C(S^n)$ is generated by $n+1$ self-adjoint elements $\{h_i\}_{0\le i\le n}$ satisfying the relation

$$h_0^2 + \cdots + h_n^2 = 1.$$

Let $F = \{h_i\}_{0\le i\le n}$. The relation above implies that the joint spectrum of a subset of F of size $m \le n$ is the m-dimensional ball B^m. The space B^m is contractible, therefore the group $\mathrm{Ext}(B^m)$ is trivial; see [Higson and Roe 2000, Sections 2.6, 2.7] for the definition of the functor Ext and its basic properties. As a consequence, for any $[\tau] \in \mathrm{Ext}(S^n)$, any proper subset of $\tau[F]$ can be lifted to a set of commuting self-adjoint operators in $B(H)$. On the other hand $\mathrm{Ext}(S^{2k+1}) = \mathbb{Z}$ for every $k \in \mathbb{N}$. We conclude that any nontrivial extension τ of $C(S^{2k+1})$ produces, by Lemma 4.3, a family $\tau[F]$ of size $2k+2$ in the Calkin algebra for which Question 4.1 has a negative answer.

Since $\mathrm{Ext}(S^{2k}) = \{0\}$ for every $k \in \mathbb{N}$, this argument does not apply to families of odd cardinality. However, in [Davidson 1985] (see also [Voiculescu 1981; Loring 1988]), the author builds a set of three commuting self-adjoint elements in the corona algebra of $\bigoplus_{n\in\mathbb{N}} M_n(\mathbb{C})$ with no commutative lifting to the multiplier algebra, whose proper subsets of size two all admit a commutative lifting. The answer to Question 4.1 for larger finite families with an odd number of elements is, to the best of our knowledge, unknown.

Acknowledgements

My sincerest gratitude goes to Ilijas Farah, for his patience, and for the impressive amount of crucial suggestions he gave me while working on this problem and on the earlier drafts of this paper. I wish to thank Alessandro Vignati for the stimulating conversations we had, for the knowledge he shared with me, and for the suggestions on the earlier drafts of this paper. I wish to thank George Elliott for the incredible amount of interesting questions related to this topic he posed and for having suggested to me a more intuitive way to present this problem and a more readable form of Lemma 3.1. Finally I'd like to thank the anonymous referee for comments and remarks on the first draft of the paper.

References

[Akemann and Doner 1979] C. A. Akemann and J. E. Doner, "A nonseparable C^*-algebra with only separable abelian C^*-subalgebras", *Bull. London Math. Soc.* **11**:3 (1979), 279–284. MR Zbl

[Anderson 1979] J. Anderson, "Pathology in the Calkin algebra", *J. Operator Theory* **2**:2 (1979), 159–167. MR Zbl

[Bice 2013] T. Bice, "The projection calculus", *Münster J. Math.* **6**:2 (2013), 557–581. MR Zbl

[Bice and Koszmider 2017] T. Bice and P. Koszmider, "A note on the Akemann–Doner and Farah–Wofsey constructions", *Proc. Amer. Math. Soc.* **145**:2 (2017), 681–687. MR Zbl

[Brown et al. 1977] L. G. Brown, R. G. Douglas, and P. A. Fillmore, "Extensions of C^*-algebras and K-homology", *Ann. of Math.* (2) **105**:2 (1977), 265–324. MR Zbl

[Choi et al. 2014] Y. Choi, I. Farah, and N. Ozawa, "A nonseparable amenable operator algebra which is not isomorphic to a C^*-algebra", *Forum Math. Sigma* **2** (2014), art. id. e2. MR Zbl

[Davidson 1985] K. R. Davidson, "Almost commuting Hermitian matrices", *Math. Scand.* **56**:2 (1985), 222–240. MR Zbl

[Davidson 1996] K. R. Davidson, *C^*-algebras by example*, Fields Institute Monographs **6**, Amer. Math. Soc., Providence, RI, 1996. MR Zbl

[Davidson 2010] K. R. Davidson, "Essentially normal operators", pp. 209–222 in *A glimpse at Hilbert space operators*, edited by S. Axler et al., Oper. Theory Adv. Appl. **207**, Birkhäuser, Basel, 2010. MR Zbl

[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 Zbl

[Farah and Hart 2013] I. Farah and B. Hart, "Countable saturation of corona algebras", *C. R. Math. Acad. Sci. Soc. R. Can.* **35**:2 (2013), 35–56. MR Zbl

[Farah and Wofsey 2012] I. Farah and E. Wofsey, "Set theory and operator algebras", pp. 63–119 in *Appalachian set theory: 2006–2012*, edited by J. Cummings and E. Schimmerling, London Math. Soc. Lecture Note Ser. **406**, Cambridge Univ. Press, 2012. MR Zbl

[Gabe 2016] J. Gabe, "A note on nonunital absorbing extensions", *Pacific J. Math.* **284**:2 (2016), 383–393. MR Zbl

[Ghasemi and Koszmider 2018] S. Ghasemi and P. Koszmider, "An extension of compact operators by compact operators with no nontrivial multipliers", *J. Noncommut. Geom.* **12**:4 (2018), 1503–1529. MR Zbl

[Higson and Roe 2000] N. Higson and J. Roe, *Analytic K-homology*, Oxford Univ. Press, 2000. MR Zbl

[Lin and Ng 2016] H. Lin and P. W. Ng, "The corona algebra of the stabilized Jiang–Su algebra", *J. Funct. Anal.* **270**:3 (2016), 1220–1267. MR Zbl

[Loring 1988] T. A. Loring, "K-theory and asymptotically commuting matrices", *Canad. J. Math.* **40**:1 (1988), 197–216. MR Zbl

[Loring 1997] T. A. Loring, *Lifting solutions to perturbing problems in C^*-algebras*, Fields Institute Monographs **8**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[Luzin 1947] N. N. Luzin, "On subsets of the series of natural numbers", *Izv. Akad. Nauk SSSR. Ser. Mat.* **11**:5 (1947), 403–410. In Russian. MR

[Murphy 1990] G. J. Murphy, *C^*-algebras and operator theory*, Academic Press, Boston, 1990. MR Zbl

[Pedersen 1990] G. K. Pedersen, "The corona construction", pp. 49–92 in *Operator theory* (Indianapolis, 1988), edited by J. B. Conway and B. B. Morrel, Pitman Res. Notes Math. Ser. **225**, Longman Sci. Tech., Harlow, UK, 1990. MR Zbl

[Popa 1983] S. Popa, "Orthogonal pairs of $*$-subalgebras in finite von Neumann algebras", *J. Operator Theory* **9**:2 (1983), 253–268. MR Zbl

[Rickart 1960] C. E. Rickart, *General theory of Banach algebras*, van Nostrand, Princeton, 1960. MR Zbl

[Sakai 1971] S. Sakai, "Derivations of simple C^*-algebras, III", *Tôhoku Math. J.* (2) **23** (1971), 559–564. MR Zbl

[Schafhauser 2018] C. Schafhauser, "Subalgebras of simple AF-algebras", preprint, 2018. arXiv

[Shelah and Steprāns 2011] S. Shelah and J. Steprāns, "Masas in the Calkin algebra without the continuum hypothesis", *J. Appl. Anal.* **17**:1 (2011), 69–89. MR Zbl

[Vignati 2015] A. Vignati, "An algebra whose subalgebras are characterized by density", *J. Symb. Log.* **80**:3 (2015), 1066–1074. MR Zbl

[Voiculescu 1981] D. Voiculescu, "Remarks on the singular extension in the C^*-algebra of the Heisenberg group", *J. Operator Theory* **5**:2 (1981), 147–170. MR Zbl

[Willard 1970] S. Willard, *General topology*, Addison-Wesley, Reading, MA, 1970. MR Zbl

[Zhang 1992] S. Zhang, "Certain C^*-algebras with real rank zero and their corona and multiplier algebras, I", *Pacific J. Math.* **155**:1 (1992), 169–197. MR Zbl

Received February 5, 2017. Revised March 20, 2019.

ANDREA VACCARO
DEPARTMENT OF MATHEMATICS
BEN GURION UNIVERSITY OF THE NEGEV
BE'ER SHEVA
ISRAEL
vaccaro@post.bgu.ac.il

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.309

SCHWARZ LEMMA AT THE BOUNDARY ON THE CLASSICAL DOMAIN OF TYPE $\mathcal{IV}$

JIANFEI WANG, TAISHUN LIU AND XIAOMIN TANG

Let $\mathcal{R}_{\mathcal{IV}}(n)$ be the classical domain of type $\mathcal{IV}$ in $\mathbb{C}^n$ with $n \geq 2$. The purpose of this paper is twofold. The first is to investigate the boundary points of $\mathcal{R}_{\mathcal{IV}}(n)$. We give a sufficient and necessary condition such that the boundary points of $\mathcal{R}_{\mathcal{IV}}(n)$ are smooth. The second is to establish the boundary Schwarz lemma on the classical domain of type $\mathcal{IV}$. we obtain the optimal estimates of the eigenvalues of the Fréchet derivative for holomorphic self-mappings at the smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$.

1. Introduction

The Schwarz lemma is one of the most important results in the classical complex analysis, which has become a crucial theme in many branches of mathematics for over a hundred years. A great deal of work has been devoted to generalizations of Schwarz lemma to more general settings. These results are widely applied to many fields such as geometric function theory, harmonic analysis, complex dynamical systems and differential geometry. We refer to [Ahlfors 1938; Elin et al. 2014; Garnett 1981; Hua 1963; Kim and Lee 2011; Rodin 1987; Wu 1967; Xiao and Zhu 2011; Yau 1978] for a more complete insight on the Schwarz lemma.

As an elementary application of the Schwarz lemma, there is the boundary version of the Schwarz lemma as follows.

Lemma 1.1 [Garnett 1981]. *Assume a holomorphic function $f(z)$ maps the unit disk $\triangle$ into itself. If f is holomorphic at $z = 1$, with $f(0) = 0$ and $f(1) = 1$, then $f'(1) \geq 1$. Moreover, the inequality is sharp.*

If the condition $f(0) = 0$ is removed in Lemma 1.1, then one has the following estimate instead:

$$f'(1) \geq \frac{|1-\overline{f(0)}|^2}{1-|f(0)|^2} > 0. \tag{1-1}$$

Tang is the corresponding author.

MSC2010: primary 32H02; secondary 32H99, 30C80.

Keywords: holomorphic mapping, Schwarz lemma at the boundary, the classical domain of type $\mathcal{IV}$.

The Schwarz lemma at the boundary is also interesting and plays an important role in the complex analysis. For instance, since de Branges solved the Bieberbach conjecture, finding the exact value of the Bloch constant is the number one important problem in the geometric function theory of one complex variable. Though the precise value of the Bloch constant is still open, Bonk [1990] improved the Bloch constant by using Lemma 1.1. As for the geometric function theory of several complex variables, applying Lemma 1.1, Gong and Liu et al. [Gong and Liu 1999; Liu and Ren 1998b; Zhang and Liu 2002] have achieved breakthroughs on the growth, covering and distortion theorems for biholomorphic convex mappings or quasiconvex mappings on some domains.

A lot of attention [Cartan 1931; Look 1958; Burns and Krantz 1994; Huang 1993; 1994; 1995; Krantz 2011; Tang et al. 2017] has been paid to the multidimensional generalizations of the boundary Schwarz lemma. Recently, we first established the Schwarz lemma at the boundary on the open unit ball of $\mathbb{C}^n$ [Liu et al. 2015], and then extended the boundary Schwarz lemma to the strongly pseudoconvex domain and the unit polydisk in $\mathbb{C}^n$ [Liu and Tang 2016a; Tang et al. 2015]. These lemmas can be applied to study the distortion theorem of biholomorphic starlike mapping in [Liu and Tang 2016b; Liu et al. 2015].

As one generalization of the unit ball, it is natural to consider the boundary Schwarz lemma on the bounded symmetric domains of classical types which are Hermitian symmetric spaces of noncompact type with nonsmooth boundaries. Let us first recall the definition of the four classical domains in the sense of Hua. Details may be found in [Hua 1963]. Let $\mathbb{C}^{m\times n}$ be the set of all $m \times n$ complex matrices with $1 \le m \le n$.

- The classical domain of type $\mathcal{I}$, $\mathcal{R}_{\mathcal{I}}(m,n) \subset \mathbb{C}^{m\times n}$, consists of matrices Z such that $I_m - Z\bar{Z}' > 0$, where I_m is the unit square matrix of order m, and the inequality sign means that the left-hand side is positive definite.

- The classical domain of type $\mathcal{II}$, $\mathcal{R}_{\mathcal{II}}(n) \subset \mathbb{C}^{n\times n}$, consists of matrices Z such that $Z = Z'$ and $I_n - Z\bar{Z}' > 0$.

- The classical domain of type $\mathcal{III}$, $\mathcal{R}_{\mathcal{III}}(n) \subset \mathbb{C}^{n\times n}$, consists of matrices Z such that $Z = -Z'$ and $I_n - Z\bar{Z}' > 0$.

- The classical domain of type $\mathcal{IV}$, $\mathcal{R}_{\mathcal{IV}}(n) \subset \mathbb{C}^n$, is the set of $Z = (z_1, z_2, \dots, z_n)$ satisfying $1 - 2\|Z\|^2 + |ZZ'|^2 > 0$ and $|ZZ'| < 1$.

The classical domain of type $\mathcal{IV}$ is also called the Lie ball, which has attracted the attention of many mathematicians. Here we refer the reader to [Chu 2014; Morimoto 1999; Wang et al. 2009], for discussions related to such studies.

By [Liu 1989], it is easy to see that $\mathcal{R}_{\mathcal{IV}}(n)$ is a bounded convex circular domain in $\mathbb{C}^n$, the Minkowski functional

$$\rho(Z) = \sqrt{\|Z\|^2 + \sqrt{\|Z\|^4 - |ZZ'|^2}}$$

is a Banach norm of $\mathbb{C}^n$, and

$$\mathcal{R}_{\mathcal{IV}}(n) = \{Z \in \mathbb{C}^n : \rho(Z) < 1\}.$$

It is easy to verify that

$$\begin{aligned} &\mathcal{R}_{\mathcal{IV}}(1) = \triangle \\ &\mathcal{R}_{\mathcal{IV}}(2) \approx \triangle^2 \\ &\mathcal{R}_{\mathcal{IV}}(3) \approx \mathcal{R}_{\mathcal{II}}(2) \\ &\mathcal{R}_{\mathcal{IV}}(4) \approx \mathcal{R}_{\mathcal{I}}(2, 2) \\ &\mathcal{R}_{\mathcal{IV}}(6) \approx \mathcal{R}_{\mathcal{III}}(4). \end{aligned}$$

From now on, we assume that $n \geq 2$ for $\mathcal{R}_{\mathcal{IV}}(n)$.

Recently, we have established the Schwarz lemma at the boundary on the classical domain of type $\mathcal{I}$ [Liu and Tang 2017] and the classical domain of type $\mathcal{II}$ [Tang et al. 2018]. Because the classical domain of type $\mathcal{IV}$ is very different from the first three types, we will introduce a new analytic tool to deal with the Schwarz lemma at the boundary. Also, notice that $\mathcal{R}_{\mathcal{IV}}(n)$ is a convex domain, but $\mathcal{R}_{\mathcal{IV}}(n)$ is not a strongly pseudoconvex domain and $\partial\mathcal{R}_{\mathcal{IV}}(n)$ is not smooth. Hence, a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n) \subset \mathbb{C}^n$ is completely different from the unit ball and the strongly pseudoconvex domain in [Liu and Tang 2016a; Liu et al. 2015].

The aim of this paper is to prove the following theorem as a generalization of Lemma 1.1 which is concerned with the holomorphic mappings from $\mathcal{R}_{\mathcal{IV}}(n)$ to $\mathcal{R}_{\mathcal{IV}}(n)$.

Theorem 1.2. *Let $f : \mathcal{R}_{\mathcal{IV}}(n) \to \mathcal{R}_{\mathcal{IV}}(n)$ be a holomorphic mapping with $f(0) = a$, and let $\mathring{Z} = e^{i\theta}\big(\frac{1}{2}(r_1 + r_2), \frac{i}{2}(r_1 - r_2), 0, \ldots, 0\big)T$ be a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$, where $1 = r_1 > r_2 \geq 0$, $\theta \in \mathbb{R}$ and T is a real orthogonal square matrix of order n. If f is holomorphic at $\mathring{Z}$ and $f(\mathring{Z}) = \mathring{Z}$, then all the eigenvalues λ, $\mu_i (i = 1, \ldots, n-2)$ and ν of the linear transformation $Df(\mathring{Z})$ on $\mathbb{C}^n$ have the following properties.*

(1) *$\nabla\rho(\mathring{Z})$ is an eigenvector of $D^*f(\mathring{Z})$ and the corresponding eigenvalue is a real number λ that we just mentioned above. That is,*

$$D^*f(\mathring{Z})(\nabla\rho(\mathring{Z})) = \lambda\nabla\rho(\mathring{Z}).$$

(2)
$$\lambda \geq \frac{1 - \rho(a)}{1 + \rho(a)} > 0.$$

(3) $$T^{1,0}_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n)) = M \oplus N,$$

where

$$N = \{e^{i\theta}\alpha T : \alpha \in \mathbb{C}^n, \alpha_1 - i\alpha_2 = 0, (\alpha_3, \cdots, \alpha_n) = 0\}$$

is a one-dimensional invariant subspace of $Df(\mathring{Z})$, *and* M *is an* $(n-2)$-*dimensional invariant subspace of* $Df(\mathring{Z})$. *Moreover, the eigenvalues* μ_i *of* $Df(\mathring{Z})$, *which is a linear transformation on* M, *satisfy*

$$|\mu_i| \le \sqrt{\lambda}, \quad i = 1, \ldots, n-2;$$

and the eigenvalue ν *of* $Df(\mathring{Z})$, *which is a linear transformation on* N, *satisfies*

$$|\nu| \le 1.$$

(4) $$|\det Df(\mathring{Z})| \le \lambda^{n/2}, \quad |\operatorname{tr} Df(\mathring{Z})| \le \lambda + \sqrt{\lambda}(n-2) + 1.$$

Moreover, the inequalities in (2), (3) *and* (4) *are sharp.*

This paper is organized as follows. In Section 2, we investigate some properties of the smooth boundary points of $\mathcal{R}_{\mathcal{IV}}(n)$. In Section 3, we present four lemmas which are used to prove our main result. In Section 4, we give the complete proof of Theorem 1.2.

2. Smooth boundary points of $\mathcal{R}_{\mathcal{IV}}(n)$

In this section, we give some characterizations of the smooth boundary points of $\mathcal{R}_{\mathcal{IV}}(n)$, which will not only be used in the subsequent sections but also have its own interest.

Denote by $\mathbb{C}^n$ the n-dimensional complex Hilbert space with the inner product and the norm given by

$$\langle Z, W\rangle = \sum_{j=1}^{n} z_j \overline{w}_j, \quad \|Z\| = (\langle Z, Z\rangle)^{\frac{1}{2}},$$

where $Z, W \in \mathbb{C}^n$. As real vectors in $\mathbb{R}^{2n}$, Z and W are orthogonal if and only if $\Re\langle Z, W\rangle = 0$. Throughout this paper, we write a point $Z \in \mathbb{C}^n$ as a row vector in $1 \times n$ matrix form $Z = (z_1, z_2, \ldots, z_n)$, and the symbol $'$ stands for the transpose of vectors or matrices.

Suppose that $\theta \in \mathbb{R}$ and T is a real orthogonal square matrix of order n. Then

$$Z \in \mathcal{R}_{\mathcal{IV}}(n) \Longleftrightarrow e^{i\theta} ZT \in \mathcal{R}_{\mathcal{IV}}(n)$$

and

$$Z \in \partial\mathcal{R}_{\mathcal{IV}}(n) \Longleftrightarrow e^{i\theta} ZT \in \partial\mathcal{R}_{\mathcal{IV}}(n).$$

We also get

$$\rho(e^{i\theta} ZT) = \rho(Z)$$

for each $Z \in \mathbb{C}^n$. For $\mathring{Z} \in \mathbb{C}^n$, according to [Hua 1963], $\mathring{Z}$ has the following polar decompositions:

$$\mathring{Z} = e^{i\theta}\Big(\frac{r_1+r_2}{2}, i\frac{r_1-r_2}{2}, 0, \dots, 0\Big)T,$$

where $\theta \in \mathbb{R}$, $r_1 \geq r_2 \geq 0$ and T is a real orthogonal square matrix of order n. By an elementary calculation, we have

$$\rho(Z) = r_1.$$

Theorem 2.1. *Let $\mathring{Z} \in \mathbb{C}^n$ be the polar decomposition above. Then $\mathring{Z}$ is a smooth boundary point of $\partial\mathcal{R}_{\mathcal{IV}}(n)$ if and only if $1 = r_1 > r_2 \geq 0$. Furthermore, the Minkowski functional $\rho(Z)$ of $\mathcal{R}_{\mathcal{IV}}(n)$ has the following properties.*

(1) *$\rho(Z)$ is holomorphic about Z and $\bar{Z}$ near $\mathring{Z}$.*

(2) *The gradient of ρ at $\mathring{Z}$ is*

$$\nabla\rho(\mathring{Z}) = e^{i\theta}(1, i, 0, \dots, 0)T$$

and $\nabla\rho(\mathring{Z})$ is a nonzero outward normal vector to $\partial\mathcal{R}_{\mathcal{IV}}(n)$ at $\mathring{Z}$.

(3) $$\langle \mathring{Z}, \nabla\rho(\mathring{Z})\rangle = 1.$$

Proof. It is clear that $\mathring{Z} \in \partial\mathcal{R}_{\mathcal{IV}}(n)$ if and only if $r_1 = \rho(\mathring{Z}) = 1$. Suppose that $1 = r_1 > r_2 \geq 0$. Then

$$\|\mathring{Z}\|^2 = \frac{1+r_2^2}{2},$$
$$\mathring{Z}\mathring{Z}' = e^{i2\theta} r_2,$$
$$\sqrt{\|\mathring{Z}\|^4 - |\mathring{Z}\mathring{Z}'|^2} = \frac{1-r_2^2}{2}.$$

Notice that

$$\begin{aligned}
2\rho(\mathring{Z})\frac{\partial\rho}{\partial\bar{z}_j}(\mathring{Z}) &= \frac{\partial}{\partial\bar{z}_j}[(\rho(Z))^2]|_{Z=\mathring{Z}} \\
&= \frac{\partial}{\partial\bar{z}_j}\big(\|Z\|^2 + \sqrt{\|Z\|^4 - |ZZ'|^2}\big)\big|_{Z=\mathring{Z}} \\
&= \mathring{z}_j + \frac{1}{2}\frac{2\|\mathring{Z}\|^2\mathring{z}_j - 2(\mathring{Z}\mathring{Z}')\bar{\mathring{z}}_j}{\sqrt{\|\mathring{Z}\|^4 - |\mathring{Z}\mathring{Z}'|^2}}.
\end{aligned}$$

This gives

$$\begin{aligned}
\nabla\rho(\mathring{Z}) &= \mathring{Z} + 2\frac{\frac{1}{2}(1+r_2^2)\mathring{Z} - r_2 e^{i\theta}\big(\frac{1}{2}(1+r_2), -\frac{i}{2}(1-r_2), 0, \dots, 0\big)T}{1-r_2^2} \\
&= \mathring{Z} + \frac{(1+r_2^2)\mathring{Z} - r_2 e^{i\theta}\big(1+r_2, -i(1-r_2), 0, \dots, 0\big)T}{1-r_2^2} \\
&= \frac{2\mathring{Z} - r_2 e^{i\theta}\big(1+r_2, -i(1-r_2), 0, \dots, 0\big)T}{1-r_2^2} \\
&= e^{i\theta}\frac{\big(1+r_2, i(1-r_2), 0, \dots, 0\big) - \big(r_2(1+r_2), -ir_2(1-r_2), 0, \dots, 0\big)}{1-r_2^2}T \\
&= e^{i\theta}(1, i, 0, \dots, 0)T \neq 0.
\end{aligned}$$

Hence, $\mathring{Z}$ is a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$. It is obvious that $\rho(Z)$ is a holomorphic function about Z and $\bar{Z}$ near $\mathring{Z}$. Moreover,

$$\langle \mathring{Z}, \nabla\rho(\mathring{Z})\rangle = \mathring{Z}\overline{\nabla\rho(\mathring{Z})}' = \tfrac{1}{2}(1+r_2) + \tfrac{1}{2}(1-r_2) = 1.$$

Conversely, suppose that $\mathring{Z}$ is a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$. Assume

$$1 = r_1 = r_2.$$

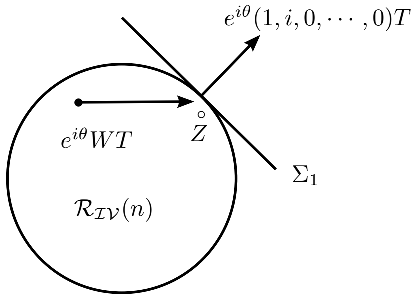
Then $\mathring{Z} = e^{i\theta}(1, 0, \dots, 0)T$, and any two nonzero outward normal vectors to $\partial\mathcal{R}_{\mathcal{IV}}(n)$ at $\mathring{Z}$ have the same direction. We consider the following two different $(2n-1)$-dimensional real affine spaces through $\mathring{Z}$ in $\mathbb{C}^n$:

$$\begin{aligned}
\Sigma_1 &= \{\mathring{Z} + e^{i\theta}\alpha T : \alpha \in \mathbb{C}^n,\ \Re(\alpha_1 - i\alpha_2) = 0\}, \\
\Sigma_2 &= \{\mathring{Z} + e^{i\theta}\alpha T : \alpha \in \mathbb{C}^n,\ \Re(\alpha_1 + i\alpha_2) = 0\}.
\end{aligned}$$

Given the nonzero vector $e^{i\theta}(1, i, 0, \dots, 0)T \in \mathbb{C}^n$, we have

$$\Re\langle e^{i\theta}\alpha T, e^{i\theta}(1, i, 0, \dots, 0)T\rangle = \Re(\alpha_1 - i\alpha_2) = 0$$

for any $\mathring{Z} + e^{i\theta}\alpha T \in \Sigma_1$. Thus $e^{i\theta}(1, i, 0, \dots, 0)T$ is a normal vector to Σ_1 at $\mathring{Z}$ (see figure). Similarly, $e^{i\theta}(1, -i, 0, \dots, 0)T \in \mathbb{C}^n$ is a normal vector to Σ_2 at $\mathring{Z}$.



Now, we claim that $|w_1 \pm iw_2| < 1$ for each $W \in \mathcal{R}_{\mathcal{IV}}(n)$. Write $\widetilde{W} = (w_3, \dots, w_n)$. Then

$$
\begin{aligned}
&|w_1 \pm iw_2|^2 \\
&= |w_1|^2 + |w_2|^2 \pm 2\Re \bar{w}_1 i w_2 \\
&= |w_1|^2 + |w_2|^2 \mp 2\Im \bar{w}_1 w_2 \\
&\le |w_1|^2 + |w_2|^2 + \sqrt{4(\Im \bar{w}_1 w_2)^2} \\
&= |w_1|^2 + |w_2|^2 + \sqrt{2\Re[\bar{w}_1 w_2 (w_1 \bar{w}_2 - \bar{w}_1 w_2)]} \\
&= |w_1|^2 + |w_2|^2 + \sqrt{|w_1|^4 + 2|w_1|^2|w_2|^2 + |w_2|^4 - |w_1|^4 - 2\Re \bar{w}_1^2 w_2^2 - |w_2|^4} \\
&= |w_1|^2 + |w_2|^2 + \sqrt{(|w_1|^2 + |w_2|^2)^2 - |w_1^2 + w_2^2|^2} \\
&\le |w_1|^2 + |w_2|^2 + \|\widetilde{W}\|^2 + \Big((|w_1|^2 + |w_2|^2)^2 + 2(|w_1|^2 + |w_2|^2)\|\widetilde{W}\|^2 + \|\widetilde{W}\|^4 \\
&\qquad\qquad - |w_1^2 + w_2^2|^2 - 2\Re(\overline{w_1^2 + w_2^2})\widetilde{W}\widetilde{W}' - |\widetilde{W}\widetilde{W}'|^2 \Big)^{\frac{1}{2}} \\
&= |w_1|^2 + |w_2|^2 + \|\widetilde{W}\|^2 + \sqrt{(|w_1|^2 + |w_2|^2 + \|\widetilde{W}\|^2)^2 - |w_1^2 + w_2^2 + \widetilde{W}\widetilde{W}'|^2} \\
&= \|W\|^2 + \sqrt{\|W\|^4 - |WW'|^2} \\
&= (\rho(W))^2 \\
&< 1.
\end{aligned}
$$

Thus, for any $e^{i\theta}WT \in \mathcal{R}_{\mathcal{IV}}(n)$, we get

$$\Re\langle \mathring{Z} - e^{i\theta}WT, e^{i\theta}(1, i, 0, \dots, 0)T \rangle = 1 - \Re(w_1 - iw_2) > 0.$$

This shows that $\mathcal{R}_{\mathcal{IV}}(n)$ is located on one side of Σ_1. That is, Σ_1 is an affine tangent space to $\partial\mathcal{R}_{\mathcal{IV}}(n)$ at $\mathring{Z}$. Similar to the proof above, we know that Σ_2 is also an affine tangent space to $\partial\mathcal{R}_{\mathcal{IV}}(n)$ at $\mathring{Z}$. Since $\mathring{Z}$ is a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$, this contradicts with $\Sigma_1 \neq \Sigma_2$. Hence, we have $1 = r_1 > r_2 \ge 0$. The proof is complete. □

3. Some lemmas

In this section, we exhibit some notations and collect several lemmas, which will be used in the subsequent section.

Let $f : \mathcal{R}_{\mathcal{IV}}(n) \to \mathbb{C}^n$ be a holomorphic mapping. The Fréchet derivative of f at $a \in \mathcal{R}_{\mathcal{IV}}(n)$ is defined by

$$(Df(a)(W))_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial z_j}(a) w_j, \quad W \in \mathbb{C}^n.$$

It is easy to see that $Df(a)$ is a linear transformation from $\mathbb{C}^n$ to $\mathbb{C}^n$ and $df(Z)\big|_{Z=a} = Df(a)(dZ)$. Let $D^*f(a)$ be the adjoint transformation of $Df(a)$ with respect to the inner product $\langle\cdot,\cdot\rangle$. That is,

$$\langle D^*f(a)(Z), W\rangle = \langle Z, Df(a)(W)\rangle, \quad Z, W \in \mathbb{C}^n.$$

$D^*f(a)$ is also a linear transformation from $\mathbb{C}^n$ to $\mathbb{C}^n$. Specifically,

$$(D^*f(a)(Z))_i = \sum_{j=1}^{n} \frac{\partial \bar{f}_j}{\partial \bar{z}_i}(a) z_j, \quad Z \in \mathbb{C}^n.$$

It is clear that λ is an eigenvalue of $Df(a)$ if and only if $\bar{\lambda}$ is an eigenvalue of $D^*f(a)$.

Lemma 3.1 [Liu 1989]. *Let $a = e^{i\phi}\big(\frac{1}{2}(l_1+l_2), \frac{i}{2}(l_1-l_2), 0, \dots, 0\big)A \in \mathcal{R}_{\mathcal{IV}}(n)$. Write*

$$Q = A'\left(\begin{array}{cc|c} 1+l_1l_2 & & 0 \\ & 1-l_1l_2 & \\ \hline \multicolumn{2}{c|}{0} & \sqrt{(1-l_1^2)(1-l_2^2)}I_{n-2} \end{array}\right)A,$$

where $1 > l_1 \geq l_2 \geq 0$, $\phi \in \mathbb{R}$ and A is a real orthogonal square matrix of order n. For any $Z \in \overline{\mathcal{R}_{\mathcal{IV}}(n)}$, define

$$\varphi_a(Z) = \frac{a + ZZ'\bar{a} - ZQ}{1 - 2Z\bar{a}' + ZZ'\overline{aa'}}.$$

Then the following statements hold:

1. *$\varphi_a(Z)$ is a holomorphic automorphism of $\mathcal{R}_{\mathcal{IV}}(n)$, and $\varphi_a(Z)$ is biholomorphic in a neighborhood of $\overline{\mathcal{R}_{\mathcal{IV}}(n)}$.*

(2) $$\varphi_a(0) = a, \quad \varphi_a(a) = 0, \quad \varphi_a^{-1} = \varphi_a.$$

(3) $$d\varphi_a(Z)|_{Z=a} = \frac{dZ(2a'\bar{a} - Q)}{(1-l_1^2)(1-l_2^2)}, \quad d\varphi_a(Z)|_{Z=0} = dZ(2\bar{a}'a - Q).$$

In what follows, we always denote by $F(Z,\xi)$ the infinitesimal form of the Carathéodory metric or the Kobayashi metric on $\mathcal{R}_{\mathcal{IV}}(n)$, where $Z \in \mathcal{R}_{\mathcal{IV}}(n)$ and $\xi \in \mathbb{C}^n$; see [Krantz 1982] for details.

Lemma 3.2. *Let $\rho(Z)$ be the Minkowski functional of $\mathcal{R}_{\mathcal{IV}}(n)$. Then with the notation of Lemma 3.1, for any $\xi \in \mathbb{C}^n$,*

$$F(a,\xi) = \frac{1}{(1-l_1^2)(1-l_2^2)} \times \rho\left[\xi A'\begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-l_1^2-l_2^2) & \frac{i}{2}(l_1^2-l_2^2) \\ -\frac{i}{2}(l_1^2-l_2^2) & \frac{1}{2}(2-l_1^2-l_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-l_1^2)(1-l_2^2)}I_{n-2} \end{pmatrix}\right].$$

Proof. Because $\varphi_a(Z)$ is a holomorphic automorphism of $\mathcal{R}_{\mathcal{IV}}(n)$, and $F(Z,\xi)$ is a biholomorphically invariant metric on $\mathcal{R}_{\mathcal{IV}}(n)$, we have

$$F(a,\xi)=F(0,D\varphi_a(a)(\xi)).$$

This, together with $D\varphi_a(a)(dZ)=d\varphi_a(Z)\big|_{Z=a}$ and Lemma 3.1, implies

$$F(a,\xi)=F(0,D\varphi_a(a)(\xi))=\frac{F(0,\xi(2a'\bar{a}-Q))}{(1-l_1^2)(1-l_2^2)}=\frac{F(0,\xi(Q-2a'\bar{a}))}{(1-l_1^2)(1-l_2^2)}.$$

Hence, by Lemma 3.2 in [Gong and Liu 1999], we obtain

$$F(a,\xi)=\frac{F(0,\xi(Q-2a'\bar{a}))}{(1-l_1^2)(1-l_2^2)}=\frac{\rho(\xi(Q-2a'\bar{a}))}{(1-l_1^2)(1-l_2^2)}.$$

Notice that

$$\begin{aligned}2a'\bar{a}&=2A'\begin{pmatrix}\frac{1}{2}(l_1+l_2)\\ \frac{i}{2}(l_1-l_2)\\ 0\\ \vdots\\ 0\end{pmatrix}\left(\tfrac{1}{2}(l_1+l_2),-\tfrac{i}{2}(l_1-l_2),0,\dots,0\right)A\\ &=A'\begin{pmatrix}\begin{pmatrix}\frac{1}{2}(l_1^2+l_2^2+2l_1l_2) & -\frac{i}{2}(l_1^2-l_2^2)\\ \frac{i}{2}(l_1^2-l_2^2) & \frac{1}{2}(l_1^2+l_2^2-2l_1l_2)\end{pmatrix} & 0\\ 0 & 0\end{pmatrix}A.\end{aligned}$$

Then

$$Q-2a'\bar{a}=A'\begin{pmatrix}\begin{pmatrix}\frac{1}{2}(2-l_1^2-l_2^2) & \frac{i}{2}(l_1^2-l_2^2)\\ -\frac{i}{2}(l_1^2-l_2^2) & \frac{1}{2}(2-l_1^2-l_2^2)\end{pmatrix} & 0\\ 0 & \sqrt{(1-l_1^2)(1-l_2^2)}I_{n-2}\end{pmatrix}A.$$

This yields

$$\begin{aligned}&F(a,\xi)\\ &=\frac{1}{(1-l_1^2)(1-l_2^2)}\\ &\quad\times\rho\left[\xi A'\begin{pmatrix}\begin{pmatrix}\frac{1}{2}(2-l_1^2-l_2^2) & \frac{i}{2}(l_1^2-l_2^2)\\ -\frac{i}{2}(l_1^2-l_2^2) & \frac{1}{2}(2-l_1^2-l_2^2)\end{pmatrix} & 0\\ 0 & \sqrt{(1-l_1^2)(1-l_2^2)}I_{n-2}\end{pmatrix}\right].\end{aligned}$$ □

Lemma 3.3. *Let $\mathring{Z}$ be a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$. Then for each $W \in \mathbb{C}^n$,*

$$|\langle W, \nabla\rho(\mathring{Z})\rangle| \le \rho(W).$$

Proof. Without loss of generality, we may assume $W \neq 0$. Then $\frac{W}{\rho(W)} \in \partial\mathcal{R}_{\mathcal{IV}}(n)$. Since $\mathcal{R}_{\mathcal{IV}}(n)$ is a bounded convex circular domain, we have

$$\Re\Big\langle \mathring{Z} - e^{i\theta}\frac{W}{\rho(W)}, \nabla\rho(\mathring{Z})\Big\rangle \ge 0$$

for any $\theta \in \mathbb{R}$. It follows from this and Theorem 2.1 that

$$\Re\frac{e^{i\theta}}{\rho(W)}\langle W, \nabla\rho(\mathring{Z})\rangle \le \Re\langle \mathring{Z}, \nabla\rho(\mathring{Z})\rangle = 1.$$

This gives $|\langle W, \nabla\rho(\mathring{Z})\rangle| \le \rho(W)$. □

Lemma 3.4 [Liu and Ren 1998a]. *Let $f : \mathcal{R}_{\mathcal{IV}}(n) \to \mathcal{R}_{\mathcal{IV}}(n)$ be a holomorphic mapping with $f(0) = 0$. Then for any $Z \in \mathcal{R}_{\mathcal{IV}}(n)$,*

$$\rho(f(Z)) \le \rho(Z).$$

4. Proof of Theorem 1.2

In this section, we will prove the Schwarz lemma at the smooth boundary point for holomorphic self-mappings of $\mathcal{R}_{\mathcal{IV}}(n)$. Firstly, we need the following notation.

Let $\mathring{Z} = e^{i\theta}\big(\frac{1}{2}(r_1 + r_2), \frac{i}{2}(r_1 - r_2), 0, \dots, 0\big)T$ be a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$, where $1 = r_1 > r_2 \ge 0$, $\theta \in \mathbb{R}$ and T is a real orthogonal square matrix of order n. Then by Theorem 2.1, we have

$$\langle e^{i\theta}\alpha T, \nabla\rho(\mathring{Z})\rangle = \alpha(1, -i, 0, \dots, 0)' = \alpha_1 - i\alpha_2$$

for any $\alpha \in \mathbb{C}^n$. Hence, the tangent space $T_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n))$ to $\partial\mathcal{R}_{\mathcal{IV}}(n)$ at $\mathring{Z}$ is

$$T_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n)) = \{\beta \in \mathbb{C}^n : \Re\langle \beta, \nabla\rho(\mathring{Z})\rangle = 0\} = \{e^{i\theta}\alpha T : \alpha \in \mathbb{C}^n,\ \Re(\alpha_1 - i\alpha_2) = 0\},$$

and the holomorphic tangent space $T^{1,0}_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n))$ to $\partial\mathcal{R}_{\mathcal{IV}}(n)$ at $\mathring{Z}$ is

$$T^{1,0}_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n)) = \{\beta \in \mathbb{C}^n : \langle \beta, \nabla\rho(\mathring{Z})\rangle = 0\} = \{e^{i\theta}\alpha T : \alpha \in \mathbb{C}^n, \alpha_1 - i\alpha_2 = 0\}.$$

Proof of Theorem 1.2. (1) For each $\beta \in T_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n))$, we have $Df(\mathring{Z})(\beta) \in T_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n))$. So

$$\Re\langle Df(\mathring{Z})(\beta), \nabla\rho(\mathring{Z})\rangle = \Re\langle \beta, D^*f(\mathring{Z})(\nabla\rho(\mathring{Z}))\rangle = 0.$$

Hence, there exists $\lambda \in \mathbb{R}$ such that

$$D^*f(\mathring{Z})(\nabla\rho(\mathring{Z})) = \lambda\nabla\rho(\mathring{Z}).$$

This means that λ is an eigenvalue of $D^*f(\mathring{Z})$ and $\nabla\rho(\mathring{Z})$ is an eigenvector of $D^*f(\mathring{Z})$ with respect to λ. Since $\lambda \in \mathbb{R}$, we know that λ is also an eigenvalue of $Df(\mathring{Z})$. The proof of (1) is complete.

(2) The proof of (2) is divided into two cases.

Case 1. $f(0) = a = 0$. For each $t \in (0, 1)$, by Lemma 3.4 we obtain

$$\rho(f(t\mathring{Z})) \le \rho(t\mathring{Z}) = t.$$

This, together with Lemma 3.3, yields

$$\Re\langle f(t\mathring{Z}), \nabla\rho(\mathring{Z})\rangle \le \rho(f(t\mathring{Z})) \le t. \tag{4-1}$$

By Theorem 2.1, $\langle \mathring{Z}, \nabla\rho(\mathring{Z})\rangle = 1$. Thus, combine

$$f(t\mathring{Z}) = \mathring{Z} - (1-t)Df(\mathring{Z})(\mathring{Z}) + O(|t-1|^2)$$

as $t \to 1^-$ and (4-1) to get

$$1 - (1-t)\Re\langle Df(\mathring{Z})(\mathring{Z}), \nabla\rho(\mathring{Z})\rangle + O(|t-1|^2) \le t.$$

This gives

$$\Re\langle \mathring{Z}, D^*f(\mathring{Z})(\nabla\rho(\mathring{Z}))\rangle + O(|t-1|) \ge 1.$$

It follows from $D^*f(\mathring{Z})(\nabla\rho(\mathring{Z})) = \lambda\nabla\rho(\mathring{Z})$ and $\langle \mathring{Z}, \nabla\rho(\mathring{Z})\rangle = 1$ that

$$\lambda + O(|t-1|) \ge 1.$$

Taking $t \to 1^-$, we have $\lambda \ge 1$.

Case 2. $f(0) = a \ne 0$. Suppose that $a = e^{i\phi}\big(\frac{1}{2}(l_1+l_2), \frac{i}{2}(l_1-l_2), 0, \dots, 0\big)A \in \mathcal{R}_{\mathcal{IV}}(n)$ and

$$Q = A'\left(\begin{array}{cc|c} 1+l_1l_2 & & 0 \\ & 1-l_1l_2 & \\ \hline \multicolumn{2}{c|}{0} & \sqrt{(1-l_1^2)(1-l_2^2)}\,I_{n-2} \end{array}\right)A,$$

where $1 > l_1 \ge l_2 \ge 0$, $\phi \in \mathbb{R}$ and A is a real orthogonal square matrix of order n. By Lemma 3.1, $g = \varphi_a \circ f : \mathcal{R}_{\mathcal{IV}}(n) \to \mathcal{R}_{\mathcal{IV}}(n)$ is a holomorphic mapping, $g(0) = 0$ and g is holomorphic at $\mathring{Z}$. Moreover,

$$\mathring{W} = g(\mathring{Z}) = \varphi_a(\mathring{Z}) = \frac{a + \mathring{Z}\mathring{Z}'\bar{a} - \mathring{Z}Q}{1 - 2\mathring{Z}\bar{a}' + \mathring{Z}\mathring{Z}'\overline{aa'}}$$

is also a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$. Because $D\varphi_a(\mathring{Z})(\beta) \in T_{\mathring{W}}(\partial\mathcal{R}_{\mathcal{IV}}(n))$ for each $\beta \in T_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n))$, we have

$$\Re\langle D\varphi_a(\mathring{Z})(\beta), \nabla\rho(\mathring{W})\rangle = 0, \quad \Re\langle \beta, D^*\varphi_a(\mathring{Z})(\nabla\rho(\mathring{W}))\rangle = 0.$$

It follows that there is $\mu \in \mathbb{R}$ such that

$$D^*\varphi_a(\mathring{Z})(\nabla\rho(\mathring{W})) = \mu\nabla\rho(\mathring{Z}). \tag{4-2}$$

Take

$$h_1(\zeta) = \langle \varphi_a(\zeta\mathring{Z}), \nabla\rho(\mathring{W})\rangle, \quad \zeta \in \triangle.$$

Then $h_1 : \triangle \to \triangle$ is a holomorphic function, and h_1 is holomorphic at $\zeta = 1$ with $h_1(1) = \langle \mathring{W}, \nabla\rho(\mathring{W})\rangle = 1$. This, together with (1-1) and (4-2), shows

$$\begin{aligned}\mu = \langle \mathring{Z}, \mu\nabla\rho(\mathring{Z})\rangle &= \langle \mathring{Z}, D^*\varphi_a(\mathring{Z})(\nabla\rho(\mathring{W}))\rangle \\ &= \langle D\varphi_a(\mathring{Z})(\mathring{Z}), \nabla\rho(\mathring{W})\rangle = h_1'(1) > 0.\end{aligned}$$

Set

$$h_2(\zeta) = \langle g(\zeta\mathring{Z}), \nabla\rho(\mathring{W})\rangle, \quad \zeta \in \triangle.$$

Then $h_2 : \triangle \to \triangle$ is a holomorphic function, and h_2 is holomorphic at $\zeta = 1$ with $h_2(0) = 0$ and $h_2(1) = 1$. It follows from Lemma 1.1, (4-2) and (1) that

$$\begin{aligned}1 \le h_2'(1) &= \langle Dg(\mathring{Z})(\mathring{Z}), \nabla\rho(\mathring{W})\rangle = \langle D\varphi_a(\mathring{Z})(Df(\mathring{Z})(\mathring{Z})), \nabla\rho(\mathring{W})\rangle \\ &= \langle Df(\mathring{Z})(\mathring{Z}), D^*\varphi_a(\mathring{Z})(\nabla\rho(\mathring{W}))\rangle = \mu\langle Df(\mathring{Z})(\mathring{Z}), \nabla\rho(\mathring{Z})\rangle \\ &= \mu\langle \mathring{Z}, D^*f(\mathring{Z})(\nabla\rho(\mathring{Z}))\rangle = \lambda\mu.\end{aligned}$$

This gives

$$\lambda \ge \frac{1}{\mu}.$$

By an elementary calculation (See Remark 4.1), we obtain

$$\mu = \langle D\varphi_a(\mathring{Z})(\mathring{Z}), \nabla\rho(\mathring{W})\rangle \le \rho(D\varphi_a(\mathring{Z})(\mathring{Z})) \le \frac{1+\rho(a)}{1-\rho(a)}.$$

It follows that

$$\lambda \ge \frac{1}{\mu} \ge \frac{1-\rho(a)}{1+\rho(a)}.$$

The proof of (2) is complete.

(3) It is clear that the $(n-1)$-dimensional space

$$T_{\mathring{Z}}^{1,0}(\partial\mathcal{R}_{\mathcal{IV}}(n)) = \left\{e^{i\theta}\alpha T : \alpha \in \mathbb{C}^n,\ \alpha_1 - i\alpha_2 = 0\right\}$$

is an invariant subspace of $Df(\mathring{Z})$. That is,

$$(e^{-i\theta}Df(\mathring{Z})(\beta)T')_1 - i(e^{-i\theta}Df(\mathring{Z})(\beta)T')_2 = 0$$

for any $\beta \in T_{\mathring{Z}}^{1,0}(\partial\mathcal{R}_{\mathcal{IV}}(n))$. Now, we claim that

$$N = \left\{e^{i\theta}\alpha T : \alpha \in \mathbb{C}^n,\ \alpha_1 - i\alpha_2 = 0,\ (\alpha_3, \dots, \alpha_n) = 0\right\}$$

is an invariant subspace of $Df(\mathring{Z})$. We only need to prove that for each

$$\beta = e^{i\theta}(\alpha_1, -i\alpha_1, 0, \dots, 0)T \in N,$$

if we set $\varepsilon = e^{-i\theta} Df(\mathring{Z})(\beta)T' \in \mathbb{C}^n$, then $\varepsilon_1 - i\varepsilon_2 = 0$ and $(\varepsilon_3, \dots, \varepsilon_n) = 0$.

Since $Df(\mathring{Z})(\beta) \in T^{1,0}_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n))$, we have $\varepsilon_1 - i\varepsilon_2 = 0$. For $t \in (0, 1)$, write the polar decompositions of $t\mathring{Z}$ and $f(t\mathring{Z})$ as

$$t\mathring{Z} = e^{i\theta}\big(\tfrac{1}{2}(tr_1 + tr_2), \tfrac{i}{2}(tr_1 - tr_2), 0, \dots, 0\big)T$$

and

$$f(t\mathring{Z}) = e^{i\theta(t)}\big(\tfrac{1}{2}(r_1(t) + r_2(t)), \tfrac{i}{2}(r_1(t) - r_2(t)), 0, \dots, 0\big)T(t)$$

respectively, where $1 > r_1(t) \ge r_2(t) \ge 0$, $\theta(t) \in \mathbb{R}$ and $T(t)$ is a real orthogonal square matrix of order n. By Lemma 3.1, corresponding to $a = t\mathring{Z}$ and $a = f(t\mathring{Z})$, take

$$Q = T' \left(\begin{array}{cc|c} 1+t^2r_2 & & \\ & 1-t^2r_2 & 0 \\ \hline 0 & & \sqrt{(1-t^2)(1-t^2r_2^2)}\,I_{n-2} \end{array}\right) T$$

and

$$Q(t) = T(t)' \left(\begin{array}{cc|c} 1+r_1(t)r_2(t) & & \\ & 1-r_1(t)r_2(t) & 0 \\ \hline 0 & & \sqrt{(1-r_1^2(t))(1-r_2^2(t))}\,I_{n-2} \end{array}\right) T(t).$$

Because $\lim_{t\to 1^-} f(t\mathring{Z}) = \mathring{Z}$, we obtain

$$\lim_{t\to 1^-} r_1(t) = 1, \qquad \lim_{t\to 1^-} r_2(t) = r_2.$$

Meanwhile, we get

$$\theta(t) = \theta + O(|t-1|), \qquad T(t) = T + O(|t-1|),$$
$$\text{and} \quad Df(t\mathring{Z})(\beta) = Df(\mathring{Z})(\beta) + O(|t-1|)$$

as $t \to 1^-$. Moreover, it follows from $f(t\mathring{Z}) = \mathring{Z} - (1-t)Df(\mathring{Z})(\mathring{Z}) + O(|t-1|^2)$ that

$$\begin{aligned} r_1(t) = \rho(f(t\mathring{Z})) &= 1 - (1-t)2\Re \textstyle\sum_{j=1}^n \frac{\partial\rho}{\partial z_j}(\mathring{Z}) Df_j(\mathring{Z})(\mathring{Z}) + O(|t-1|^2) \\ &= 1 - (1-t)\Re\langle Df(\mathring{Z})(\mathring{Z}), \nabla\rho(\mathring{Z})\rangle + O(|t-1|^2) \\ &= 1 - (1-t)\Re\langle \mathring{Z}, D^*f(\mathring{Z})(\nabla\rho(\mathring{Z}))\rangle + O(|t-1|^2) \\ &= 1 - \lambda(1-t) + O(|t-1|^2) \end{aligned}$$

as $t \to 1^-$. This implies

$$\begin{aligned} \sqrt{1 - r_1^2(t)} &= \sqrt{1 - [1 - \lambda(1-t) + O(|t-1|^2)]^2} \\ &= \sqrt{2\lambda(1-t) + O(|t-1|^2)} \end{aligned} \tag{4-3}$$

as $t \to 1^-$. By Lemma 3.2,

$$
\begin{aligned}
&F(t\mathring{Z}, \beta)\\
&= \frac{1}{(1-t^2)(1-t^2r_2^2)}\\
&\quad \times \rho\left[\beta T' \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-t^2-t^2r_2^2) & \frac{i}{2}(t^2-t^2r_2^2) \\ -\frac{i}{2}(t^2-t^2r_2^2) & \frac{1}{2}(2-t^2-t^2r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-t^2)(1-t^2r_2^2)}I_{n-2} \end{pmatrix}\right]\\
&= \frac{1}{(1-t^2)(1-t^2r_2^2)}\\
&\quad \times \rho\left[(\alpha_1, -i\alpha_1, 0, \dots, 0)\right.\\
&\qquad \left. \times \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-t^2-t^2r_2^2) & \frac{i}{2}(t^2-t^2r_2^2) \\ -\frac{i}{2}(t^2-t^2r_2^2) & \frac{1}{2}(2-t^2-t^2r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-t^2)(1-t^2r_2^2)}I_{n-2} \end{pmatrix}\right]\\
&= \frac{|\alpha_1|}{1-t^2r_2^2}\rho[(1, -i, 0, \dots, 0)].
\end{aligned}
$$

This gives

$$\lim_{t\to 1^-} \sqrt{1-t^2}F(t\mathring{Z}, \beta) = 0. \tag{4-4}$$

Similarly, we have

$$
\begin{aligned}
&F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta)]\\
&= \frac{1}{(1-r_1^2(t))(1-r_2^2(t))}\\
&\quad \times \rho\left[Df(t\mathring{Z})(\beta)T(t)'\right.\\
&\qquad \left. \times \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-r_1^2(t)-r_2^2(t)) & \frac{i}{2}(r_1^2(t)-r_2^2(t)) \\ -\frac{i}{2}(r_1^2(t)-r_2^2(t)) & \frac{1}{2}(2-r_1^2(t)-r_2^2(t)) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-r_1^2(t))(1-r_2^2(t))}I_{n-2} \end{pmatrix}\right].
\end{aligned}
$$

Notice that

$$Df(t\mathring{Z})(\beta)T(t)' = Df(\mathring{Z})(\beta)T' + O(|t-1|) = e^{i\theta}\varepsilon + O(|t-1|)$$

as $t \to 1^-$ and $\varepsilon_2 = -i\varepsilon_1$. This, together with (4-3), shows

$$
\begin{aligned}
&\lim_{t\to 1^-} \sqrt{1-r_1^2(t)}F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta)] \\
&= \lim_{t\to 1^-} \frac{1}{\sqrt{1-r_1^2(t)}(1-r_2^2(t))} \\
&\quad \times \rho\Bigg[\big(\varepsilon_1 + O(|t-1|), -i\varepsilon_1 + O(|t-1|), \varepsilon_3 + O(|t-1|), \dots, \varepsilon_n + O(|t-1|)\big) \\
&\quad\quad \times \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-r_1^2(t)-r_2^2(t)) & \frac{i}{2}(r_1^2(t)-r_2^2(t)) \\ -\frac{i}{2}(r_1^2(t)-r_2^2(t)) & \frac{1}{2}(2-r_1^2(t)-r_2^2(t)) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-r_1^2(t))(1-r_2^2(t))}I_{n-2} \end{pmatrix}\Bigg] \\
&= \lim_{t\to 1^-} \frac{1}{\sqrt{1-r_1^2(t)}(1-r_2^2(t))} \\
&\quad \times \rho\Big[\Big(\varepsilon_1(1-r_1^2(t)) + O(|t-1|), -i\varepsilon_1(1-r_1^2(t)) + O(|t-1|), \\
&\quad\quad \varepsilon_3\sqrt{(1-r_1^2(t))(1-r_2^2(t))} + O(|t-1|), \dots, \\
&\quad\quad \varepsilon_n\sqrt{(1-r_1^2(t))(1-r_2^2(t))} + O(|t-1|)\Big)\Big] \\
&= \frac{1}{\sqrt{1-r_2^2}}\rho[(0, 0, \varepsilon_3, \dots, \varepsilon_n)].
\end{aligned}
$$

By the contraction property of the Kobayashi metric, we get

$$
F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta)] \le F(t\mathring{Z}, \beta). \tag{4-5}
$$

Thus, by (4-3), (4-4) and (4-5), we obtain

$$
\begin{aligned}
\rho[(0, 0, \varepsilon_3, \dots, \varepsilon_n)] &= \sqrt{1-r_2^2} \lim_{t\to 1^-} \sqrt{1-r_1^2(t)}F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta)] \\
&\le \sqrt{1-r_2^2} \lim_{t\to 1^-} \frac{\sqrt{1-r_1^2(t)}}{\sqrt{1-t^2}}\sqrt{1-t^2}F(t\mathring{Z}, \beta) \\
&= \sqrt{1-r_2^2}\sqrt{\lambda} \lim_{t\to 1^-} \sqrt{1-t^2}F(t\mathring{Z}, \beta) \\
&= 0.
\end{aligned}
$$

That means $(\varepsilon_3, \dots, \varepsilon_n) = 0$. It follows that N is a one-dimensional invariant subspace of $Df(\mathring{Z})$. Hence, there is an $(n-2)$-dimensional invariant subspace M of $Df(\mathring{Z})$ such that

$$
T^{1,0}_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n)) = M \oplus N.
$$

Because $M \cap N = \{0\}$, we have $(\varepsilon_3, \dots, \varepsilon_n) \ne 0$ for any $\beta \in M \setminus \{0\}$.

For each eigenvalue μ_i of $Df(\mathring{Z})$ on M, suppose that $\beta^{(i)} = e^{i\theta}\alpha^{(i)}T \in M \setminus \{0\}$ is a nonzero eigenvector with respect to μ_i. Here,

$$(\alpha_3^{(i)}, \dots, \alpha_n^{(i)}) \neq 0 \quad \text{and} \quad e^{-i\theta} Df(\mathring{Z})(\beta^{(i)})T' = \mu_i \alpha^{(i)}, \quad i = 1, \dots, (n-2).$$

By Lemma 3.2, we get (omitting the dependence of r_1 and r_2 on t for brevity):

$$\begin{aligned}
&F(t\mathring{Z}, \beta^{(i)})\\
&= \frac{1}{(1-t^2)(1-t^2r_2^2)}\\
&\qquad \times \rho\left[\beta^{(i)}T' \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-t^2-t^2r_2^2) & \frac{i}{2}(t^2-t^2r_2^2) \\ -\frac{i}{2}(t^2-t^2r_2^2) & \frac{1}{2}(2-t^2-t^2r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-t^2)(1-t^2r_2^2)}\, I_{n-2} \end{pmatrix}\right]\\
&= \frac{1}{(1-t^2)(1-t^2r_2^2)}\rho\Bigg[\left(\alpha_1^{(i)}, -i\alpha_1^{(i)}, \alpha_3^{(i)}, \dots, \alpha_n^{(i)}\right)\\
&\qquad \times \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-t^2-t^2r_2^2) & \frac{i}{2}(t^2-t^2r_2^2) \\ -\frac{i}{2}(t^2-t^2r_2^2) & \frac{1}{2}(2-t^2-t^2r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-t^2)(1-t^2r_2^2)}\, I_{n-2} \end{pmatrix}\Bigg]\\
&= \frac{1}{1-t^2r_2^2}\rho\left[\left(\alpha_1^{(i)}, -i\alpha_1^{(i)}, \sqrt{\frac{1-t^2r_2^2}{1-t^2}}\alpha_3^{(i)}, \dots, \sqrt{\frac{1-t^2r_2^2}{1-t^2}}\alpha_n^{(i)}\right)\right].
\end{aligned}$$

This gives

$$\lim_{t\to 1^-} \sqrt{1-t^2}\, F(t\mathring{Z}, \beta^{(i)}) = \frac{1}{\sqrt{1-r_2^2}}\rho[(0, 0, \alpha_3^{(i)}, \dots, \alpha_n^{(i)})] \neq 0.$$

On the other hand,

$$\begin{aligned}
&F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta^{(i)})]\\
&= \frac{1}{(1-r_1^2)(1-r_2^2)}\\
&\qquad \times \rho\Bigg[Df(t\mathring{Z})(\beta^{(i)})T(t)'\\
&\qquad\qquad \times \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-r_1^2-r_2^2) & \frac{i}{2}(r_1^2-r_2^2) \\ -\frac{i}{2}(r_1^2-r_2^2) & \frac{1}{2}(2-r_1^2-r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-r_1^2)(1-r_2^2)}\, I_{n-2} \end{pmatrix}\Bigg].
\end{aligned}$$

Notice that

$$Df(t\mathring{Z})(\beta^{(i)})T(t)' = Df(\mathring{Z})(\beta^{(i)})T' + O(|t-1|) = \mu_i e^{i\theta}\alpha^{(i)} + O(|t-1|)$$

as $t \to 1^-$. Therefore we have

$$\begin{aligned}
&\lim_{t\to 1^-} \sqrt{1-r_1^2}\,F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta^{(i)})] \\
&= \lim_{t\to 1^-} \frac{1}{\sqrt{1-r_1^2}(1-r_2^2)} \\
&\quad \times \rho\Bigg[\big(\mu_i\alpha_1^{(i)} + O(|t-1|), -i\mu_i\alpha_1^{(i)} + O(|t-1|), \mu_i\alpha_3^{(i)} + O(|t-1|), \dots, \\
&\qquad\qquad \mu_i\alpha_n^{(i)} + O(|t-1|)\big) \\
&\qquad \times \left(\begin{matrix} \begin{pmatrix} \frac{1}{2}(2-r_1^2-r_2^2) & \frac{i}{2}(r_1^2-r_2^2) \\ -\frac{i}{2}(r_1^2-r_2^2) & \frac{1}{2}(2-r_1^2-r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-r_1^2)(1-r_2^2)}\,I_{n-2} \end{matrix} \right) \Bigg] \\
&= \lim_{t\to 1^-} \frac{1}{\sqrt{1-r_1^2}(1-r_2^2)} \\
&\quad \times \rho\Big[\Big(\mu_i\alpha_1^{(i)}(1-r_1^2) + O(|t-1|), -i\mu_i\alpha_1^{(i)}(1-r_1^2) + O(|t-1|), \\
&\qquad \mu_i\alpha_3^{(i)}\sqrt{(1-r_1^2)(1-r_2^2)} + O(|t-1|), \dots, \\
&\qquad\qquad \mu_i\alpha_n^{(i)}\sqrt{(1-r_1^2)(1-r_2^2)} + O(|t-1|)\Big)\Big] \\
&= \frac{|\mu_i|}{\sqrt{1-r_2^2}}\rho[(0,0,\alpha_3^{(i)}, \dots, \alpha_n^{(i)})] \\
&= |\mu_i| \lim_{t\to 1^-} \sqrt{1-t^2}\,F(t\mathring{Z}, \beta^{(i)}).
\end{aligned}$$

It follows from this and (4-3) that

$$\begin{aligned}
1 &\geq \lim_{t\to 1^-} \frac{F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta^{(i)})]}{F(t\mathring{Z}, \beta^{(i)})} \\
&= \lim_{t\to 1^-} \frac{\sqrt{1-t^2}}{\sqrt{1-r_1^2}} \frac{\sqrt{1-r_1^2}}{\sqrt{1-t^2}} \frac{F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta^{(i)})]}{F(t\mathring{Z}, \beta^{(i)})} = \frac{|\mu_i|}{\sqrt{\lambda}}.
\end{aligned}$$

This implies

$$|\mu_i| \leq \sqrt{\lambda}, \quad i = 1, \dots, n-2.$$

For the only eigenvalue ν of $Df(\mathring{Z})$ on N, suppose that $\beta = e^{i\theta}\alpha T \in N \setminus \{0\}$ is a nonzero eigenvector with respect to ν. Here, $\alpha = (\alpha_1, -i\alpha_1, 0, \dots, 0)$, $\alpha_1 \neq 0$ and $e^{-i\theta} Df(\mathring{Z})(\beta)T' = \nu\alpha$. Then by Lemma 3.2, we have

$$\begin{aligned}
&F(t\mathring{Z}, \beta)\\
&= \frac{1}{(1-t^2)(1-t^2r_2^2)}\\
&\qquad \times \rho\left[\beta T' \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-t^2-t^2r_2^2) & \frac{i}{2}(t^2-t^2r_2^2) \\ -\frac{i}{2}(t^2-t^2r_2^2) & \frac{1}{2}(2-t^2-t^2r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-t^2)(1-t^2r_2^2)}\, I_{n-2} \end{pmatrix}\right]\\
&= \frac{1}{(1-t^2)(1-t^2r_2^2)}\\
&\quad \times \rho\Bigg[(\alpha_1, -i\alpha_1, 0, \dots, 0)\\
&\qquad \times \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-t^2-t^2r_2^2) & \frac{i}{2}(t^2-t^2r_2^2) \\ -\frac{i}{2}(t^2-t^2r_2^2) & \frac{1}{2}(2-t^2-t^2r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-t^2)(1-t^2r_2^2)}\, I_{n-2} \end{pmatrix}\Bigg]\\
&= \frac{|\alpha_1|}{1-t^2r_2^2} \rho[(1, -i, 0, \dots, 0)].
\end{aligned}$$

Hence,

$$\lim_{t\to 1^-} F(t\mathring{Z}, \beta) = \frac{|\alpha_1|}{1-r_2^2} \rho[(1, -i, 0, \dots, 0)] = \frac{2|\alpha_1|}{1-r_2^2} \neq 0.$$

On the other hand,

$$\begin{aligned}
&F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta)]\\
&= \frac{1}{(1-r_1^2)(1-r_2^2)}\\
&\quad \times \rho\Bigg[Df(t\mathring{Z})(\beta)T(t)'\\
&\qquad \times \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-r_1^2-r_2^2) & \frac{i}{2}(r_1^2-r_2^2) \\ -\frac{i}{2}(r_1^2-r_2^2)2 & \frac{1}{2}(2-r_1^2-r_2^2) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-r_1^2)(1-r_2^2)}\, I_{n-2} \end{pmatrix}\Bigg].
\end{aligned}$$

By (4-3), we obtain

$$
\begin{aligned}
&Df(t\mathring{Z})(\beta)T(t)' \\
&\quad = Df(\mathring{Z})(\beta)T' + O(|t-1|) \\
&\quad = e^{i\theta}\nu\alpha + e^{i\theta}\big[b_1(1-r_1^2(t)) + O(|t-1|^2),\, b_2(1-r_1^2(t)) + O(|t-1|^2), \\
&\qquad\qquad O(|t-1|), \dots, O(|t-1|)\big]
\end{aligned}
$$

as $t \to 1^-$. This, together with (4-3), yields

$$
\begin{aligned}
&\lim_{t\to 1^-} F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta)] \\
&= \lim_{t\to 1^-} \frac{1}{(1-r_1^2(t))(1-r_2^2(t))} \\
&\quad \times \rho\Bigg[\Big(\nu\alpha_1 + b_1(1-r_1^2(t)) + O(|t-1|^2),\, -i\nu\alpha_1 + b_2(1-r_1^2(t)) + O(|t-1|^2), \\
&\qquad\qquad O(|t-1|), \dots, O(|t-1|)\Big) \\
&\qquad \times \begin{pmatrix} \begin{pmatrix} \frac{1}{2}(2-r_1^2(t)-r_2^2(t)) & \frac{i}{2}(r_1^2(t)-r_2^2(t)) \\ -\frac{i}{2}(r_1^2(t)-r_2^2(t)) & \frac{1}{2}(2-r_1^2(t)-r_2^2(t)) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-r_1^2(t))(1-r_2^2(t))}\, I_{n-2} \end{pmatrix}\Bigg] \\
&= \frac{1}{1-r_2^2}\rho\Big[\Big(\nu\alpha_1 + \frac{1-r_2^2}{2}(b_1-ib_2),\, -i\nu\alpha_1 + i\frac{1-r_2^2}{2}(b_1-ib_2), 0, \dots, 0\Big)\Big].
\end{aligned}
$$

Since

$$
\begin{aligned}
&\big|\nu\alpha_1 + \tfrac{i}{2}(1-r_2^2)(b_1-ib_2)\big|^2 + \big|-i\nu\alpha_1 + \tfrac{i}{2}(1-r_2^2)(b_1-ib_2)\big|^2 \\
&\qquad = 2|\nu|^2|\alpha_1|^2 + \tfrac{1}{2}(1-r_2^2)^2|b_1-ib_2|^2
\end{aligned}
$$

and

$$
\begin{aligned}
&\big[\nu\alpha_1 + \tfrac{i}{2}(1-r_2^2)(b_1-ib_2)\big]^2 + \big[-i\nu\alpha_1 + \tfrac{i}{2}(1-r_2^2)(b_1-ib_2)\big]^2 \\
&\qquad = 2\nu\alpha_1(1-r_2^2)(b_1-ib_2),
\end{aligned}
$$

we have

$$
\begin{aligned}
&\big(\rho\big[(\nu\alpha_1 + \tfrac{1}{2}(1-r_2^2)(b_1-ib_2), -i\nu\alpha_1 + \tfrac{i}{2}(1-r_2^2)(b_1-ib_2), 0, \dots, 0)\big]\big)^2 \\
&= 2|\nu|^2|\alpha_1|^2 + \tfrac{1}{2}(1-r_2^2)^2|b_1-ib_2|^2 \\
&\qquad + \sqrt{\big[2|\nu|^2|\alpha_1|^2 + \tfrac{1}{2}(1-r_2^2)^2|b_1-ib_2|^2\big]^2 - 4|\nu|^2|\alpha_1|^2(1-r_2^2)^2|b_1-ib_2|^2} \\
&= 2|\nu|^2|\alpha_1|^2 + \tfrac{1}{2}(1-r_2^2)^2|b_1-ib_2|^2 + \sqrt{\big[2|\nu|^2|\alpha_1|^2 - \tfrac{1}{2}(1-r_2^2)^2|b_1-ib_2|^2\big]^2} \\
&\geq 2|\nu|^2|\alpha_1|^2 + \tfrac{1}{2}(1-r_2^2)^2|b_1-ib_2|^2 + 2|\nu|^2|\alpha_1|^2 - \tfrac{1}{2}(1-r_2^2)^2|b_1-ib_2|^2 \\
&= 4|\nu|^2|\alpha_1|^2.
\end{aligned}
$$

It follows that

$$\begin{aligned}
&\lim_{t\to 1^-} F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta)] \\
&= \frac{1}{1-r_2^2}\rho\big[\big(\nu\alpha_1 + \tfrac{i}{2}(-r_2^2)(b_1 - ib_2), -i\nu\alpha_1 + \tfrac{i}{2}(1-r_2^2)(b_1 - ib_2), 0, \dots, 0\big)\big] \\
&\geq \frac{2|\nu||\alpha_1|}{1-r_2^2} \\
&= |\nu| \lim_{t\to 1^-} F(t\mathring{Z}, \beta),
\end{aligned}$$

which implies

$$1 \geq \lim_{t\to 1^-} \frac{F[f(t\mathring{Z}), Df(t\mathring{Z})(\beta)]}{F(t\mathring{Z}, \beta)} \geq |\nu|.$$

The proof of (3) is complete.

(4) Note that $T^{1,0}_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n)) = \{e^{i\theta}\alpha T : \alpha \in \mathbb{C}^n, \alpha_1 - i\alpha_2 = 0\} = M \oplus N$ is an $(n-1)$-dimensional invariant subspace of $Df(\mathring{Z})$. So, there is a one-dimensional invariant subspace L of $Df(\mathring{Z})$ such that

$$\mathbb{C}^n = L \oplus M \oplus N.$$

Since $L \cap T^{1,0}_{\mathring{Z}}(\partial\mathcal{R}_{\mathcal{IV}}(n)) = \{0\}$ we have $\alpha_1 - i\alpha_2 \neq 0$ for any $\beta = e^{i\theta}\alpha T \in L \setminus \{0\}$. Now, we prove that λ is just the eigenvalue of $Df(\mathring{Z})$ on L. Suppose that $\tilde{\lambda}$ is an eigenvalue of $Df(\mathring{Z})$ on L, and $\beta = e^{i\theta}\alpha T \in L \setminus \{0\}$ is a nonzero eigenvector of $Df(\mathring{Z})$ with respect to $\tilde{\lambda}$. Then Theorem 2.1 is utilized to derive

$$\begin{aligned}
\langle Df(\mathring{Z})(\beta), \nabla\rho(\mathring{Z})\rangle &= \tilde{\lambda}\langle \beta, \nabla\rho(\mathring{Z})\rangle \\
&= \tilde{\lambda}\langle e^{i\theta}\alpha T, e^{i\theta}(1, i, 0, \dots, 0)T\rangle = \tilde{\lambda}(\alpha_1 - i\alpha_2).
\end{aligned}$$

Meanwhile,

$$\langle Df(\mathring{Z})(\beta), \nabla\rho(\mathring{Z})\rangle = \langle \beta, D^*f(\mathring{Z})(\nabla\rho(\mathring{Z}))\rangle = \lambda\langle \beta, \nabla\rho(\mathring{Z})\rangle = \lambda(\alpha_1 - i\alpha_2).$$

This, together with $\alpha_1 - i\alpha_2 \neq 0$, gives $\tilde{\lambda} = \lambda$. Therefore λ, μ_i $(i = 1, \dots, n-2)$ and ν are all the eigenvalues of the linear transformation $Df(\mathring{Z})$ on $\mathbb{C}^n$. This implies

$$|\det Df(\mathring{Z})| \leq \lambda^{n/2}, \quad |\operatorname{tr} Df(\mathring{Z})| \leq \lambda + \sqrt{\lambda}(n-2) + 1.$$

The proof of (4) is complete. □

Finally, we give the following example to show that the inequalities in (2), (3) and (4) of Theorem 1.2 are sharp.

Example. Let $a = \big(\frac{\varepsilon}{2}, i\frac{\varepsilon}{2}, 0, \dots, 0\big) \in \mathcal{R}_{\mathcal{IV}}(n)$ and $0 < \varepsilon < 1$. According to Lemma 3.1, take $Q = \left(\begin{smallmatrix} I_2 & 0 \\ 0 & \sqrt{1-\varepsilon^2} I_{n-2} \end{smallmatrix}\right)$. Write $e_j = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{C}^n$, with the 1 in the j-th place. Let

$$\mathring{Z} = \left(\tfrac{1}{2}(1+r_2),\, \tfrac{i}{2}(1-r_2), 0, \dots, 0\right)$$

be a smooth boundary point of $\mathcal{R}_{\mathcal{IV}}(n)$, where $1 > r_2 \geq 0$. Define

$$f(Z) = -\varphi_{-a}(Z) = \frac{a + ZZ'\bar{a} + ZQ}{1 + 2Z\bar{a}'}, \quad Z \in \overline{\mathcal{R}_{\mathcal{IV}}(n)}.$$

Then $f : \mathcal{R}_{\mathcal{IV}}(n) \to \mathcal{R}_{\mathcal{IV}}(n)$ is a holomorphic mapping with $f(0) = a$, and f is holomorphic at $\mathring{Z}$. Moreover, f has the following properties.

(1) $$f(\mathring{Z}) = \mathring{Z}.$$

(2) For any $\beta \in \mathbb{C}^n$,

$$Df(\mathring{Z})(\beta) = \beta \begin{pmatrix} \begin{pmatrix} 1/(1+\varepsilon) & -i\varepsilon/(1+\varepsilon) \\ i\varepsilon/(1+\varepsilon) & 1/(1+\varepsilon) \end{pmatrix} & 0 \\ 0 & \sqrt{(1-\varepsilon)/(1+\varepsilon)}\, I_{n-2} \end{pmatrix}.$$

(3) $$Df(\mathring{Z})((1, i, 0, \dots, 0)) = \frac{1-\rho(a)}{1+\rho(a)}(1, i, 0, \dots, 0).$$

This shows that one of eigenvalues of $Df(\mathring{Z})$ is $(1-\rho(a))/(1+\rho(a))$.

(4) $$Df(\mathring{Z})(e_j) = \sqrt{\frac{1-\rho(a)}{1+\rho(a)}}\, e_j, \quad \text{where } j = 3, \dots, n.$$

This shows that the $n-2$ eigenvalues of $Df(\mathring{Z})$ are all $\sqrt{(1-\rho(a))/(1+\rho(a))}$.

(5) $$Df(\mathring{Z})((1, -i, 0, \dots, 0)) = (1, -i, 0, \dots, 0).$$

This shows that one of eigenvalues of $Df(\mathring{Z})$ is 1.

Proof. By Lemma 3.1, it is clear that $f : \mathcal{R}_{\mathcal{IV}}(n) \to \mathcal{R}_{\mathcal{IV}}(n)$ is a holomorphic mapping with $f(0) = a$, and f is holomorphic at $\mathring{Z}$.

(1) It follows from $\|a\|^2 = \varepsilon^2/2$ and $aa' = 0$ that $\rho(a) = \varepsilon$. Since $\mathring{Z}\mathring{Z}' = r_2$ and $\mathring{Z}\bar{a}' = \varepsilon/2$, we have

$$\begin{aligned} f(\mathring{Z}) &= \frac{a + \mathring{Z}\mathring{Z}'\bar{a} + \mathring{Z}Q}{1 + 2\mathring{Z}\bar{a}'} \\ &= \frac{(\varepsilon/2, i\varepsilon/2, 0, \dots, 0) + r_2(\varepsilon/2, -i\varepsilon/2, 0, \dots, 0) + \mathring{Z}}{1+\varepsilon} = \frac{\varepsilon\mathring{Z} + \mathring{Z}}{1+\varepsilon} = \mathring{Z}. \end{aligned}$$

(2) For any $\beta \in \mathbb{C}^n$, we get

$$\begin{aligned}
Df(\mathring{Z})(\beta) &= \frac{2\beta\mathring{Z}'\bar{a} + \beta Q}{1 + 2\mathring{Z}\bar{a}'} - \frac{2\beta\bar{a}'(a + \mathring{Z}\mathring{Z}'\bar{a} + \mathring{Z}Q)}{(1 + 2\mathring{Z}\bar{a}')^2} \\
&= \frac{\beta}{1 + 2\mathring{Z}\bar{a}'}(Q + 2\mathring{Z}'\bar{a} - 2\bar{a}' f(\mathring{Z})) \\
&= \frac{\beta}{1+\varepsilon}(Q + 2\mathring{Z}'\bar{a} - 2\bar{a}'\mathring{Z}) \\
&= \frac{\beta}{1+\varepsilon}\left[\begin{pmatrix} I_2 & 0 \\ 0 & \sqrt{1-\varepsilon^2} I_{n-2}\end{pmatrix} + \frac{\varepsilon}{2}\begin{pmatrix}\begin{pmatrix} 1+r_2 & -i(1+r_2) \\ i(1-r_2) & 1-r_2\end{pmatrix} & 0 \\ 0 & 0\end{pmatrix}\right. \\
&\qquad\qquad\left. - \frac{\varepsilon}{2}\begin{pmatrix}\begin{pmatrix} 1+r_2 & i(1-r_2) \\ -i(1+r_2) & 1-r_2\end{pmatrix} & 0 \\ 0 & 0\end{pmatrix}\right] \\
&= \frac{\beta}{1+\varepsilon}\begin{pmatrix}\begin{pmatrix} 1 & -i\varepsilon \\ i\varepsilon & 1\end{pmatrix} & 0 \\ 0 & \sqrt{1-\varepsilon^2} I_{n-2}\end{pmatrix} \\
&= \beta\begin{pmatrix}\begin{pmatrix} 1/(1+\varepsilon) & -i\varepsilon/(1+\varepsilon) \\ i\varepsilon/(1+\varepsilon) & 1/(1+\varepsilon)\end{pmatrix} & 0 \\ 0 & \sqrt{(1-\varepsilon)/(1+\varepsilon)} I_{n-2}\end{pmatrix}.
\end{aligned}$$

(3)–(5) Using (2) and making a straightforward calculation, we can easily obtain (3), (4) and (5). □

Remark 4.1. In the proof of (2) in Theorem 1.2, we need to estimate

$$\rho(D\varphi_a(\mathring{Z})(\mathring{Z})) \le \frac{1+\rho(a)}{1-\rho(a)}.$$

For simplicity we restrict attention to complex dimension 2. Let $\triangle(0, r)$ be the disk in $\mathbb{C}$ with the center 0 and the radius $r > 0$. Then A maps $\{Z \in \mathbb{C}^n : \rho(Z) < \delta\}$ to $\triangle^2(0, \delta)$. It is easy to check that the following relation holds:

$$\begin{array}{ccc}
\mathcal{R}_{\mathcal{IV}}(2) & \xrightarrow{\left(\begin{smallmatrix} 1 & 1 \\ -i & i\end{smallmatrix}\right) = A} & \triangle^2 \\
\varphi_a \big\downarrow & & \big\downarrow \varphi_b \\
\mathcal{R}_{\mathcal{IV}}(2) & \xrightarrow[\left(\begin{smallmatrix} 1 & 1 \\ -i & i\end{smallmatrix}\right) = A]{} & \triangle^2
\end{array}$$

Here, $\varphi_a(Z)$ is a holomorphic automorphism of $\mathcal{R}_{\mathcal{IV}}(2)$, and $\varphi_b(Z)$ is a holomorphic automorphism of $\triangle^2$. Thus, we have

$$\varphi_a(Z)A = \varphi_b(z_1 - iz_2, z_1 + iz_2), \quad D\varphi_a(\mathring{Z})(\mathring{Z}) = D\varphi_b(\mathring{\zeta})(\mathring{\zeta})A^{-1},$$

where $b = (a_1 - ia_2, a_1 + ia_2)$ and $\|b\|_\infty = \max(|a_1 - ia_2|, |a_1 + ia_2|) \le \rho(a)$. This implies

$$D\varphi_a(\mathring{Z})(\mathring{Z}) = \left(\frac{\varphi_{b_1}(\mathring{\zeta}_1) - (a_1 - ia_2)}{1 - \overline{(a_1 - ia_2)}\mathring{\zeta}_1}, \frac{\varphi_{b_2}(\mathring{\zeta}_2) - (a_1 + ia_2)}{1 - \overline{(a_1 + ia_2)}\mathring{\zeta}_2} \right) A^{-1}.$$

Because the pair on the right (multiplying A^{-1}) lies in

$$\triangle^2\left(0, \frac{1+\|b\|_\infty}{1-\|b\|_\infty}\right) \subset \triangle^2\left(0, \frac{1+\rho(a)}{1-\rho(a)}\right),$$

we obtain

$$\rho(D\varphi_a(\mathring{Z})(\mathring{Z})) \le \frac{1+\rho(a)}{1-\rho(a)}.$$

Remark 4.2. From the view of geometry, N is an invariant subspace of $Df(\mathring{Z})$ perhaps because the Levi form of ρ at $\mathring{Z}$ is positive semidefinite and not positive definite on N. We get the same conclusions from $|\mu_i| \le \sqrt{\lambda}$ $(i = 1, \dots, n-2)$ as Theorem 3.1 in [Liu and Tang 2016a] perhaps because the Levi form of ρ at $\mathring{Z}$ is positive definite on M.

Remark 4.3. From the proof of Theorem 1.2, it is clear that we need only to assume that the mapping f is C^1 up to the boundary of $\mathcal{R}_{\mathcal{IV}}(n)$ near $\mathring{Z}$.

Remark 4.4. When $\mathcal{R}_{\mathcal{IV}}(1) = \triangle$, Theorem 1.2 is just the classical Schwarz lemma at the boundary of the unit disk. When $\mathcal{R}_{\mathcal{IV}}(2) \approx \triangle^2$, Theorem 1.2 is just Theorem 3.1 in [Tang et al. 2015]. When $\mathcal{R}_{\mathcal{IV}}(3) \approx \mathcal{R}_{\mathcal{II}}(2)$, Theorem 1.2 is just Theorem 4.1 in [Tang et al. 2018]. When $\mathcal{R}_{\mathcal{IV}}(4) \approx \mathcal{R}_{\mathcal{I}}(2, 2)$, Theorem 1.2 is just Theorem 3.1 in [Liu and Tang 2017]. When $\mathcal{R}_{\mathcal{IV}}(6) \approx \mathcal{R}_{\mathcal{III}}(4)$, Theorem 1.2 is just the special case of the classical domain of type $\mathcal{III}$.

Acknowledgements

This research was supported by NNSF of China (Nos. 11571105, 11771139, 11501198, 11671362) and the Natural Science Foundation of Zhejiang Province (No. LY16A010004).

References

[Ahlfors 1938] L. V. Ahlfors, "An extension of Schwarz's lemma", *Trans. Amer. Math. Soc.* **43**:3 (1938), 359–364. MR Zbl

[Bonk 1990] M. Bonk, "On Bloch's constant", *Proc. Amer. Math. Soc.* **110**:4 (1990), 889–894. MR Zbl

[Burns and Krantz 1994] D. M. Burns and S. G. Krantz, "Rigidity of holomorphic mappings and a new Schwarz lemma at the boundary", *J. Amer. Math. Soc.* **7**:3 (1994), 661–676. MR Zbl

[Cartan 1931] H. Cartan, "Les fonctions de deux variables complexes et le problème de la represènta-tion analytique", *J. Math. Pure Appl.* **10** (1931), 1–116. Zbl

[Chu 2014] C.-H. Chu, "Iteration of holomorphic maps on Lie balls", *Adv. Math.* **264** (2014), 114–154. MR Zbl

[Elin et al. 2014] M. Elin, F. Jacobzon, M. Levenshtein, and D. Shoikhet, "The Schwarz lemma: rigidity and dynamics", pp. 135–230 in *Harmonic and complex analysis and its applications*, edited by A. Vasilev, Birkhäuser, Cham, 2014. MR Zbl

[Garnett 1981] J. B. Garnett, *Bounded analytic functions*, Pure Appl. Math. **96**, Academic Press, New York, 1981. MR Zbl

[Gong and Liu 1999] S. Gong and T. Liu, "Distortion theorems for biholomorphic convex mappings on bounded convex circular domains", *Chinese Ann. Math. Ser. B* **20**:3 (1999), 297–304. MR Zbl

[Hua 1963] L. K. Hua, *Harmonic analysis of functions of several complex variables in the classical domains*, Transl. Math. Monogr. **6**, Amer. Math. Soc., Providence, RI, 1963. MR Zbl

[Huang 1993] X. J. Huang, "Some applications of Bell's theorem to weakly pseudoconvex domains", *Pacific J. Math.* **158**:2 (1993), 305–315. MR Zbl

[Huang 1994] X. J. Huang, "A preservation principle of extremal mappings near a strongly pseudo-convex point and its applications", *Illinois J. Math.* **38**:2 (1994), 283–302. MR Zbl

[Huang 1995] X. J. Huang, "A boundary rigidity problem for holomorphic mappings on some weakly pseudoconvex domains", *Canad. J. Math.* **47**:2 (1995), 405–420. MR Zbl

[Kim and Lee 2011] K.-T. Kim and H. Lee, *Schwarz's lemma from a differential geometric viewpoint*, IISc Lecture Notes Series **2**, IISc Press, Bangalore, 2011. MR Zbl

[Krantz 1982] S. G. Krantz, *Function theory of several complex variables*, Wiley, New York, 1982. MR Zbl

[Krantz 2011] S. G. Krantz, "The Schwarz lemma at the boundary", *Complex Var. Elliptic Equ.* **56**:5 (2011), 455–468. MR Zbl

[Liu 1989] T. S. Liu, *The growth theorems, covering theorems and distortion theorems for biholomor-phic mappings on classical domains*, Ph.D. thesis, University of Science and Technology of China, 1989.

[Liu and Ren 1998a] T. Liu and G. Ren, "Decomposition theorem of normalized biholomorphic convex mappings", *J. Reine Angew. Math.* **496** (1998), 1–13. MR Zbl

[Liu and Ren 1998b] T. Liu and G. Ren, "Growth theorem of convex mappings on bounded convex circular domains", *Sci. China Ser. A* **41**:2 (1998), 123–130. MR Zbl

[Liu and Tang 2016a] T. Liu and X. Tang, "Schwarz lemma at the boundary of strongly pseudoconvex domain in $\mathbb{C}^n$", *Math. Ann.* **366**:1-2 (2016), 655–666. MR Zbl

[Liu and Tang 2016b] T. S. Liu and X. M. Tang, "Distortion theorem at extreme points for biholomor-phic starlike mappings on the unit ball", *Chinese Ann. Math. Ser. A* **37**:1 (2016), 47–54. In Chinese; translated in *Chinese J. Contemp. Math.* **37**:1 (2016), 45–52. MR

[Liu and Tang 2017] T. Liu and X. Tang, "A boundary Schwarz lemma on the classical domain of type $\mathcal{I}$", *Sci. China Math.* **60**:7 (2017), 1239–1258. MR Zbl

[Liu et al. 2015] T. Liu, J. Wang, and X. Tang, "Schwarz lemma at the boundary of the unit ball in $\mathbb{C}^n$ and its applications", *J. Geom. Anal.* **25**:3 (2015), 1890–1914. MR Zbl

[Look 1958] K. H. Look, "Schwarz lemma and analytic invariants", *Sci. Sinica* **7**:5 (1958), 453–504. MR Zbl

[Morimoto 1999] M. Morimoto, "A generalization of the Cauchy–Hua integral formula on the Lie ball", *Tokyo J. Math.* **22**:1 (1999), 177–192. MR Zbl

[Rodin 1987] B. Rodin, "Schwarz's lemma for circle packings", *Invent. Math.* **89**:2 (1987), 271–289. MR Zbl

[Tang et al. 2015] X. Tang, T. Liu, and J. Lu, "Schwarz lemma at the boundary of the unit polydisk in $\mathbb{C}^n$", *Sci. China Math.* **58**:8 (2015), 1639–1652. MR Zbl

[Tang et al. 2017] X. Tang, T. Liu, and W. Zhang, "Schwarz lemma at the boundary and rigidity property for holomorphic mappings on the unit ball of $\mathbb{C}^n$", *Proc. Amer. Math. Soc.* **145**:4 (2017), 1709–1716. MR Zbl

[Tang et al. 2018] X. Tang, T. Liu, and W. Zhang, "Schwarz lemma at the boundary on the classical domain of type $\mathcal{II}$", *J. Geom. Anal.* **28**:2 (2018), 1610–1634. MR Zbl

[Wang et al. 2009] J. Wang, T. Liu, and H. Xu, "Distortion theorems on the Lie ball $\mathcal{R}_{\mathrm{IV}}(n)$ in $\mathbb{C}^n$", *Sci. China Ser. A* **52**:12 (2009), 2743–2750. MR Zbl

[Wu 1967] H. Wu, "Normal families of holomorphic mappings", *Acta Math.* **119** (1967), 193–233. MR Zbl

[Xiao and Zhu 2011] J. Xiao and K. Zhu, "Volume integral means of holomorphic functions", *Proc. Amer. Math. Soc.* **139**:4 (2011), 1455–1465. MR Zbl

[Yau 1978] S. T. Yau, "A general Schwarz lemma for Kähler manifolds", *Amer. J. Math.* **100**:1 (1978), 197–203. MR Zbl

[Zhang and Liu 2002] W. Zhang and T. Liu, "On growth and covering theorems of quasi-convex mappings in the unit ball of a complex Banach space", *Sci. China Ser. A* **45**:12 (2002), 1538–1547. MR Zbl

Received May 2, 2018. Revised August 13, 2018.

JIANFEI WANG
SCHOOL OF MATHEMATICAL SCIENCES
HUAQIAO UNIVERSITY
QUANZHOU, FUJIAN
CHINA
wangjf@mail.ustc.edu.cn

TAISHUN LIU
DEPARTMENT OF MATHEMATICS
HUZHOU TEACHERS COLLEGE, HUZHOU UNIVERSITY
ZHEJIANG
CHINA
lts@ustc.edu.cn

XIAOMIN TANG
DEPARTMENT OF MATHEMATICS
HUZHOU TEACHERS COLLEGE, HUZHOU UNIVERSITY
ZHEJIANG
CHINA
txm@zjhu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.335

CYCLIC η-PARALLEL SHAPE AND RICCI OPERATORS ON REAL HYPERSURFACES IN TWO-DIMENSIONAL NONFLAT COMPLEX SPACE FORMS

YANING WANG

We consider three-dimensional real hypersurfaces in a nonflat complex space form of complex dimension two with cyclic η-parallel shape or Ricci operators and classify such hypersurfaces satisfying some other geometric restrictions. Some results extend those of Ahn et al. (1993), Lim et al. (2013), Kim et al. (2007) and Sohn (2007).

1. Introduction

A complex space form is a Kähler manifold of constant holomorphic sectional curvature c with complex dimension n and is denoted by $M^n(c)$, $n > 1$. If $M^n(c)$ is complete and simply connected, then it is complex analytically isometric to one of the following spaces:

- a complex projective space $\mathbb{C}P^n(c)$ when $c > 0$;
- a complex hyperbolic space $\mathbb{C}H^n(c)$ when $c < 0$;
- a complex Euclidean space $\mathbb{C}^n$ when $c = 0$.

Let M be a real hypersurface in a nonflat complex space form $M^n(c)$ whose Kähler metric and complex structure are denoted by $\bar{g}$ and J respectively. On M there exists an *almost contact metric structure* (ϕ, ξ, η, g) induced from $\bar{g}$ and J (for details see Section 2), where ξ is called a *structure vector field*. Let D be the distribution determined by tangent vectors orthogonal to ξ at each point of M. Let A be the shape operator of M in $M^n(c)$. If the structure vector field ξ is *principal*, that is, $A\xi = \alpha\xi$, where $\alpha = \eta(A\xi)$, then M is called a *Hopf hypersurface* and α is called *Hopf principal curvature*. Hopf real hypersurfaces in nonflat complex space forms with constant principal curvatures were classified in [Cecil and Ryan 1982; Kimura 1986; Takagi 1973; 1975a; 1975b] in the case of $\mathbb{C}P^n(c)$ and in [Berndt 1989] in the case of $\mathbb{C}H^n(c)$, respectively. For simplicity, a real hypersurface M in a nonflat complex space form is said to be of type (A) if it is locally congruent to a type (A_1) or (A_2) hypersurface in $\mathbb{C}P^n(c)$ (see [Takagi 1973; 1975a]) or type (A_0), (A_1)

MSC2010: 53B25, 53D15.
Keywords: real hypersurface, complex space forms, cyclic η-parallel, shape operator, Ricci operator.

or (A_2) hypersurface in $\mathbb{C}H^n(c)$ (see [Berndt 1989]). Similarly, M is said to be of type (B) if it is locally congruent to a type (B) hypersurface in $\mathbb{C}P^n(c)$ or $\mathbb{C}H^n(c)$.

Unlike the case of real space forms, the shape operator of real hypersurfaces in a nonflat complex space form can not be parallel (deduced directly from the Codazzi equation (2-8)). This motivates some authors to consider certain conditions weaker than parallel shape operators. One of the methods is to investigate cyclic parallel shape operators, i.e., $g((\nabla_X A)Y, Z) + g((\nabla_Y A)Z, X) + g((\nabla_Z A)X, Y) = 0$ for any vector fields X, Y and Z. Ki [1988] proved that the shape operator of a real hypersurface in a nonflat complex space form is cyclic parallel if and only if the hypersurface is of type (A); see also [Niebergall and Ryan 1997]. This implies that cyclic parallelism is still too strong and therefore cyclic η-parallelism for the shape operator, i.e., $g((\nabla_X A)Y, Z) + g((\nabla_Y A)Z, X) + g((\nabla_Z A)X, Y) = 0$ for any vector fields X, Y and Z orthogonal to the structure vector field, were studied in [Kim et al. 2007]. Another method is to study η-parallel shape operators, i.e., $g((\nabla_X A)Y, Z) = 0$ for any vector fields X, Y, Z orthogonal to the structure vector field. Because the η-parallelism condition is so weak, in history many authors investigated it together with some other conditions; see [Ahn et al. 1993; Kimura and Maeda 1989; Kon and Loo 2012; Niebergall and Ryan 1997; Suh 1990]. Finally, without any other redundant restrictions, Kon and Loo [2011] proved that a real hypersurface of dimension greater than three in nonflat complex space forms has η-parallel shape operator if and only if it is of type (A) or (B) or a ruled real hypersurface. However, the above result is still open for real hypersurfaces of dimension three. Among others, the third method to extend parallel shape operators was introduced by Cho [2015], who proved that there exist no real hypersurfaces with Killing type shape operator, i.e., $(\nabla_X A)Y + (\nabla_Y A)X = 0$ for any vector fields X, Y. However, there exist real hypersurfaces in a nonflat complex space form $M^n(c)$ with transversal Killing shape operators, i.e., $(\nabla_X A)Y + (\nabla_Y A)X = 0$ for any vector fields X, Y orthogonal to the structure vector field, which are of type (A) (see [Cho 2015] for the case of $n > 2$).

Just like the case of shape operators, the Ricci operator of real hypersurfaces in a nonflat complex space form can not be parallel (see [Niebergall and Ryan 1997] for dimension greater than three and [Kim 2004] for dimension three). However, there exist real hypersurfaces such that the Ricci operator is cyclic parallel (see [Kwon and Nakagawa 1988; Niebergall and Ryan 1997]), i.e.,

$$g((\nabla_X Q)Y, Z) + g((\nabla_Y Q)Z, X) + g((\nabla_Z Q)X, Y) = 0$$

for any vector fields X, Y, Z. Because Ricci cyclic parallelism is too strong, Kwon and Nakagawa [1989] considered cyclic η-parallel Ricci operator, i.e., $g((\nabla_X Q)Y, Z) + g((\nabla_Y Q)Z, X) + g((\nabla_Z Q)X, Y) = 0$ for any vector fields X, Y, Z orthogonal to the structure vector field, and proved that a Hopf real hypersurface of dimension greater than three in a nonflat complex space form has cyclic η-parallel Ricci operator if and only if it is of type (A) or (B). Real hypersurfaces with cyclic η-parallel and η-parallel Ricci operators were also studied in [Kim et al. 2007] and [Kim et al. 2006; Kon 2014; 2017; Maeda 2013; Pérez et al. 2001], respectively.

Many results mentioned above were obtained for hypersurfaces of dimension greater three. This motivates us to generalize those results for three-dimensional real hypersurfaces. Here we aim to classify real hypersurfaces in a nonflat complex space form of complex dimension two with cyclic η-parallel shape or Ricci operators satisfying some other geometric conditions. Some results in this paper are extensions of corresponding earlier results; see [Ahn et al. 1993; Lim et al. 2013a; 2013b; Kim et al. 2007; Sohn 2007].

2. Preliminaries

Let M be a real hypersurface immersed in a complex space form $M^n(c)$ and N be a unit normal vector field of M. We denote by $\overline{\nabla}$ the Levi-Civita connection of the metric $\bar{g}$ of $M^n(c)$ and J the complex structure respectively. Let g and ∇ be the induced metric from the ambient space and the Levi-Civita connection of g respectively. Then the Gauss and Weingarten formulas are given respectively as:

$$\overline{\nabla}_X Y = \nabla_X Y + g(AX, Y)N, \quad \overline{\nabla}_X N = -AX \tag{2-1}$$

for any vector fields X and Y tangent to M, where A denotes the shape operator of M in $M^n(c)$. For any vector field X tangent to M, we put

$$JX = \phi X + \eta(X)N, \quad JN = -\xi. \tag{2-2}$$

We can define on M an almost contact metric structure (ϕ, ξ, η, g) satisfying

$$\phi^2 = -\mathrm{id} + \eta \otimes \xi, \quad \eta(\xi) = 1, \quad \phi\xi = 0, \tag{2-3}$$

$$g(\phi X, \phi Y) = g(X, Y) - \eta(X)\eta(Y), \quad \eta(X) = g(X, \xi) \tag{2-4}$$

for any vector fields X and Y on M. Moreover, applying the parallelism of the complex structure (i.e., $\overline{\nabla} J = 0$) of $M^n(c)$ and using (2-1), (2-2), we have

$$(\nabla_X \phi)Y = \eta(Y)AX - g(AX, Y)\xi, \tag{2-5}$$

$$\nabla_X \xi = \phi AX \tag{2-6}$$

for any vector fields X and Y. We denote by R the Riemannian curvature tensor of M. Since $M^n(c)$ is assumed to be of constant holomorphic sectional curvature c, then the Gauss and Codazzi equations of M in $M^n(c)$ are given respectively as:

$$R(X,Y)Z=\frac{c}{4}\Big\{g(Y,Z)X-g(X,Z)Y+g(\phi Y,Z)\phi X-g(\phi X,Z)\phi Y \\ -2g(\phi X,Y)\phi Z\Big\}+g(AY,Z)AX-g(AX,Z)AY, \tag{2-7}$$

$$(\nabla_X A)Y-(\nabla_Y A)X=\frac{c}{4}\{\eta(X)\phi Y-\eta(Y)\phi X-2g(\phi X,Y)\xi\} \tag{2-8}$$

for any vector fields X, Y on M.

From (2-7) we see that the Ricci operator Q is given by

$$QX=\frac{c}{4}((2n+1)X-3\eta(X)\xi)+mAX-A^2X \tag{2-9}$$

for any vector field X tangent to the hypersurface, where $m:=\text{trace}A$ is the mean curvature.

In this paper, all manifolds are assumed to be connected and of class C^∞.

3. Cyclic η-parallel shape operator

Let M be a real hypersurface in a complex space form $M^n(c)$. We put

$$A\xi=\alpha\xi+\beta U, \tag{3-1}$$

where $\alpha=\eta(A\xi)$, U is a unit vector field orthogonal to ξ and β is a smooth function. Applying (2-1) and (2-2) we see that $\beta U=-\phi\nabla_\xi\xi$. We put

$$\Omega=\{p\in M \mid \beta(p)\neq 0\}.$$

Then Ω is an open subset of M.

Lemma 3.1 [Panagiotidou and Xenos 2012, Lemma 1]. *Suppose M is a three-dimensional real hypersurface in a nonflat complex plane $M^2(c)$. Then the following relations hold*:

$$\begin{gathered} AU=\gamma U+\delta\phi U+\beta\xi,\quad A\phi U=\delta U+\mu\phi U,\\ \nabla_U\xi=-\delta U+\gamma\phi U,\quad \nabla_{\phi U}\xi=-\mu U+\delta\phi U,\quad \nabla_\xi\xi=\beta\phi U,\\ \nabla_U U=\kappa_1\phi U+\delta\xi,\quad \nabla_{\phi U}U=\kappa_2\phi U+\mu\xi,\quad \nabla_\xi U=\kappa_3\phi U,\\ \nabla_U\phi U=-\kappa_1 U-\gamma\xi,\quad \nabla_{\phi U}\phi U=-\kappa_2 U-\delta\xi,\quad \nabla_\xi\phi U=-\kappa_3 U-\beta\xi, \end{gathered} \tag{3-2}$$

where γ, δ, μ, κ_i, $i=\{1,2,3\}$ are smooth functions on M and $\{\xi,U,\phi U\}$ is an orthonormal basis of the tangent space of M at a point of M.

Applying this lemma, from the Codazzi equation (2-8) for $X = U$ or $X = \phi U$ and $Y = \xi$ we have

(3-3) $$U(\beta) - \xi(\gamma) = \alpha\delta - 2\delta\kappa_3,$$

(3-4) $$\xi(\delta) = \alpha\gamma + \beta\kappa_1 + \delta^2 + \mu\kappa_3 + \tfrac{c}{4} - \gamma\mu - \gamma\kappa_3 - \beta^2,$$

(3-5) $$U(\alpha) - \xi(\beta) = -3\beta\delta,$$

(3-6) $$\xi(\mu) = \alpha\delta + \beta\kappa_2 - 2\delta\kappa_3,$$

(3-7) $$\phi U(\alpha) = \alpha\beta + \beta\kappa_3 - 3\beta\mu,$$

(3-8) $$\phi U(\beta) = \alpha\mu - 2\gamma\mu + 2\delta^2 + \tfrac{c}{2} + \alpha\gamma + \beta\kappa_1.$$

Similarly, from the Codazzi equation for $X = U$ and $Y = \phi U$ we have

(3-9) $$U(\delta) - \phi U(\gamma) = \mu\kappa_1 - \gamma\kappa_1 - \beta\gamma - 2\delta\kappa_2 - 2\beta\mu,$$

(3-10) $$U(\mu) - \phi U(\delta) = \gamma\kappa_2 + \beta\delta - \kappa_2\mu - 2\delta\kappa_1.$$

Moreover, applying Lemma 3.1, from the Gauss Equation (2-7) and the definition of the Riemannian curvature tensor $R(X, Y)Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$ we have

(3-11) $$U(\kappa_2) - \phi U(\kappa_1) = 2\delta^2 - 2\gamma\mu - \kappa_1^2 - \gamma\kappa_3 - \kappa_2^2 - \mu\kappa_3 - c.$$

(3-12) $$\phi U(\kappa_3) - \xi(\kappa_2) = 2\beta\mu - \mu\kappa_1 + \delta\kappa_2 + \kappa_3\kappa_1 + \beta\kappa_3.$$

The above relations can also be seen in [Panagiotidou and Xenos 2012; Wang 2018].

Proposition 3.2. *The shape operator of a real hypersurface in complex space forms is cyclic η-parallel if and only if it is η-parallel.*

Proof. From the Codazzi equation (2-8), we have $g((\nabla_X A)Y, Z) = g((\nabla_Y A)X, Z)$ for any vector fields X, Y, Z orthogonal to the structure vector field ξ. Since the shape operator is symmetric, we have $g((\nabla_X A)Y, Z) = g((\nabla_X A)Z, Y)$ for any vector fields X, Y, Z. Consequently, if the shape operator is cyclic η-parallel, i.e., $g((\nabla_X A)Y, Z) + g((\nabla_Y A)Z, X) + g((\nabla_Z A)X, Y) = 0$ for any vector fields X, Y, Z orthogonal to ξ, by the previous two properties we obtain $g((\nabla_X A)Y, Z) = 0$ for any vector fields X, Y, Z orthogonal to ξ. The converse is trivial. □

According to Proposition 3.2, we observe that [Kim et al. 2007, Theorem 1.1] and [Ahn et al. 1993, Theorem C] are the same. More precisely, as mentioned before, real hypersurfaces of dimension greater than three with η-parallel shape operators were completely classified by Kon and Loo [2011]. Thus, following their result and Proposition 3.2 we have:

Corollary 3.3. *A real hypersurface in a nonflat complex space form of complex dimension greater than two has cyclic η-parallel shape operator if and only if it is of type (A), (B) or is a ruled real hypersurface.*

However, the classification problem for real hypersurfaces of dimension three with η-parallel shape operators is still an open question; see also [Cho 2015; Niebergall and Ryan 1997]. Considering η-parallel shape operators together with some other conditions, we have:

Lemma 3.4 [Kimura and Maeda 1989; Niebergall and Ryan 1997; Suh 1990]. *A Hopf real hypersurface in a nonflat complex space form $M^n(c)$ with η-parallel shape operator is of type (A) or (B) for all $n \geq 2$.*

The next proposition shows that η-parallelism for shape operators is weak, so people studying this problem require some other geometric conditions (a recent paper on this topic is [Lim et al. 2013a]).

Proposition 3.5. *The shape operator of a real hypersurface in a nonflat complex space form of complex dimension two is η-parallel if and only if*

$$\begin{aligned} U(\gamma) - 2\kappa_1\delta - 2\beta\delta &= 0, \\ U(\delta) + \kappa_1\gamma + \beta\gamma - \kappa_1\mu &= 0, \\ \phi U(\delta) - \kappa_2\mu + \kappa_2\gamma + \beta\delta &= 0, \\ \phi U(\mu) + 2\kappa_2\delta &= 0, \\ U(\mu) + 2\kappa_1\delta &= 0, \\ \phi U(\gamma) - 2\beta\mu - 2\kappa_2\delta &= 0. \end{aligned} \tag{3-13}$$

Proof. According to Lemma 3.1 and relation (3-1), it follows directly that

$$\begin{aligned} (\nabla_U A)U &= (U(\beta) - \alpha\delta)\xi + (U(\gamma) - 2\kappa_1\delta - 2\beta\delta)U + (U(\delta) + \kappa_1\gamma + \beta\gamma - \kappa_1\mu)\phi U. \\ (\nabla_U A)\phi U &= (\delta^2 - \mu\gamma + \beta\kappa_1 + \alpha\gamma)\xi (U(\delta) - \kappa_1\mu + \kappa_1\gamma + \beta\gamma)U + (U(\mu) + 2\kappa_1\delta)\phi U. \\ (\nabla_{\phi U} A)U &= (\phi U(\beta) + \mu\gamma - \alpha\mu - \delta^2)\xi + (\phi U(\gamma) - 2\beta\mu - 2\kappa_2\delta)U \\ &\qquad + (\phi U(\delta) + \kappa_2\gamma + \beta\delta - \kappa_2\mu)\phi U. \\ (\nabla_{\phi U} A)\phi U &= (\alpha\delta + \beta\kappa_2)\xi + (\phi U(\delta) - \kappa_2\mu + \kappa_2\gamma + \beta\delta)U + (\phi U(\mu) + 2\kappa_2\delta)\phi U. \end{aligned}$$

The shape operator is said to be η-parallel if $g((\nabla_X A)Y, Z) = 0$ for any X, Y, Z orthogonal to ξ. Thus, the remainder of the proof follows immediately from the previous four equations. □

Next we present some solutions of the system of partial differential equations (3-13). But first, we need the following lemma:

Lemma 3.6 [Ahn et al. 1993; Kimura 1986]. *A real hypersurface in a nonflat space form $M^n(c)$, $n \geq 2$, is a ruled real hypersurface if and only if $g(AX, Y) = 0$ for any vector fields X, Y orthogonal to ξ.*

A ruled real hypersurface M in $\mathbb{C}P^n(c)$ or $\mathbb{C}H^n(c)$ can be characterized by using the shape operator A, namely,

$$A\xi = \alpha\xi + \beta U \ (\beta \neq 0), \quad AU = \beta\xi, \quad AZ = 0$$

for any $Z \perp \{\xi, U\}$, where U is a unit vector field orthogonal to ξ, both α and β are functions on M.

Theorem 3.7. *Let M be a real hypersurface in a nonflat complex space form of complex dimension two with η-parallel shape operator. Then, ξ is an eigenvector field of the Ricci operator if and only if M is of type (A), (B) or is a ruled real hypersurface.*

Proof. In view of Lemma 3.4, next we need only to consider the non-Hopf case. Assume that M is non-Hopf and hence

$$\Omega = \{p \in M \mid \beta(p) \neq 0 \text{ in a neighborhood of } p\}$$

is nonempty. On Ω, the applications of (3-1) and (3-2) in (2-9) give

$$Q\xi = \left(\tfrac{1}{2}c + \alpha(\gamma+\mu) - \beta^2\right)\xi + \beta\mu U - \beta\delta\phi U.$$

If ξ is an eigenvector field of the Ricci operator, it follows that $\mu = \delta = 0$ on Ω. In this context, the system of partial differential equations (3-13) becomes

$$U(\gamma) = 0, \quad (\kappa_1 + \beta)\gamma = 0, \quad \kappa_2\gamma = 0, \quad \phi U(\gamma) = 0. \tag{3-14}$$

From Lemma 3.1, we acquire $[U, \phi U] = -\gamma\xi - \kappa_1 U - \kappa_2\phi U$. By virtue of (3-14), taking differentiation of γ along $[U, \phi U]$ implies $\gamma\xi(\gamma) = 0$.

First of all, on Ω we suppose that $\gamma \neq 0$ holds on some open subset $\mathcal{Q}$ of Ω and in view of (3-14) we see that γ is a nonzero constant. Now, (3-14) becomes $\kappa_1 + \beta = 0$ and $\kappa_2 = 0$. With the aid of $\mu = \delta = 0$, the application of this in (3-11) implies $\phi U(\beta) = -\beta^2 - \gamma\kappa_3 - c$, which is compared with (3-8) giving

$$\tfrac{3}{2}c + \alpha\gamma + \gamma\kappa_3 = 0.$$

In view of $\delta = \mu = 0$ and $\kappa_1 = -\beta$, (3-4) becomes $\frac{1}{4}c + \alpha\gamma - 2\beta^2 - \gamma\kappa_3 = 0$, which is compared with the previous equation, giving

$$\tfrac{7}{8}c + \alpha\gamma - \beta^2 = 0. \tag{3-15}$$

Taking differentiation of (3-15) along ϕU yields that $\gamma\phi U(\alpha) = 2\beta\phi(\beta)$ which is analyzed with the aid of (3-7) and (3-8) giving

$$c + \alpha\gamma - \gamma\kappa_3 - 2\beta^2 = 0,$$

where we have applied the assumption $\beta \neq 0$ on Ω. In view of $\kappa_1 = -\beta$ and $\mu = \delta = 0$, comparing the previous equation with (3-4) yields $c = 0$, and we arrive at a contradiction and hence $\mathcal{Q}$ is empty.

Therefore, we conclude that $\gamma = 0$ on Ω. Moreover, in view of $\mu = \delta = 0$, from Lemma 3.1 we see that $g(AX, Y) = 0$ for any vector fields X, Y orthogonal to ξ on Ω. Next, we need only prove that $M - \Omega$ is empty. Actually, when $M - \Omega$ is nonempty, on this subset ξ is principal and hence α is a constant; see [Kon 1979; Maeda 1976; Niebergall and Ryan 1997]. Applying Lemma 3.4, we see that those principal curvatures on $M - \Omega$ have the same property as those of real hypersurfaces of type (A) or (B) in M. Consequently, it follows that all principal curvatures on $M - \Omega$ are constant. In view of continuity of principal curvatures and connectedness of the hypersurface, we conclude that $M - \Omega$ is empty, or equivalently, Ω coincides with the whole of M. Finally, according to Lemma 3.6 we see that the hypersurface is locally congruent to a ruled real hypersurface. The converse is easy to check. This completes the proof. □

We continue to solve the system of partial differential equations (3-13) under some other conditions. First, we consider

$$g((A\phi - \phi A)X, Y) = 0 \text{ for any } X, Y \perp \xi. \tag{3-16}$$

Note that (3-16) was investigated in [Ahn et al. 1993; Lim et al. 2013a; Kim et al. 2007] for real hypersurfaces of dimension greater than three. In what follows, we aim to generalize those results for real hypersurfaces of dimension three.

Theorem 3.8. *Let M be a real hypersurface in a nonflat complex space form of complex dimension two with η-parallel shape operator. Then, M satisfies* (3-16) *if and only if M is of type (A) or ruled real hypersurface.*

Proof. According to Lemma 3.1, we see that M satisfies (3-16) if and only if

$$\delta = 0 \quad \text{and} \quad \gamma = \mu. \tag{3-17}$$

In this case, by Proposition 3.5, the shape operator is η-parallel if and only if

$$U(\mu) = \phi U(\mu) = 0 \quad \text{and} \quad \beta\gamma = \beta\mu = 0. \tag{3-18}$$

If the hypersurface M is Hopf, using $\beta = 0$, with the aid of the first term of (3-17) and (3-6), we observe that $\xi(\mu) = 0$ which is combined with the first two terms of (3-18) giving that μ is a constant. This implies $A\phi = \phi A$ and this is a sufficient and necessary condition for a real hypersurface to be of type (A); for more details see [Montiel and Romero 1986; Okumura 1975].

If M is not a Hopf hypersurface, then

$$\Omega = \{p \in M \mid \beta(p) \neq 0 \text{ in a neighborhood of } p\}$$

is nonempty. As seen in proof of Theorem 3.7, we state that Ω coincides with the whole of M. Therefore, on M, using $\beta \neq 0$, from the second term of (3-18) we obtain $\gamma = \mu = 0$. In view of the first term of (3-17), by Lemmas 3.1 and 3.6, we see that M is locally congruent to a ruled real hypersurface. The converse is easy to check. □

Remark 3.9. The above theorem improves [Lim et al. 2013a, Theorem 1.8].

We consider the following condition studied in [Kim et al. 2006; Sohn 2007], i.e.,

$$g((Q\phi - \phi Q)X, Y) = 0 \text{ for any } X, Y \perp \xi. \tag{3-19}$$

By using (3-19), a solution for (3-13) is given and applying this we have:

Theorem 3.10. *Let M be a real hypersurface in a nonflat complex space form of complex dimension two with η-parallel shape operator. Then, M satisfies (3-19) if and only if M is of type (A) hypersurface.*

Proof. Because the proof is long, we divide the discussion into several steps.

Step 1. We assume that the hypersurface M is non-Hopf and hence

$$\Omega = \{p \in M \mid \beta(p) \neq 0 \text{ in a neighborhood of } p\}$$

is nonempty. On Ω, the application of (2-9) and Lemma 3.1 for the case $n = 2$ gives

$$\begin{aligned} Q\xi &= (\tfrac{1}{2}r - c - \gamma\mu + \delta^2)\xi + \beta\mu U - \beta\delta\phi U, \\ QU &= \beta\mu\xi + (\tfrac{1}{2}r - \tfrac{1}{4}c - \alpha\mu)U + \alpha\delta\phi U, \\ Q\phi U &= -\beta\delta\xi + \alpha\delta U + (\tfrac{1}{2}r - \tfrac{1}{4}c - \alpha\gamma + \beta^2)\phi U, \end{aligned} \tag{3-20}$$

where the scalar curvature is given by

$$r = 3c + 2\alpha\gamma + 2\alpha\mu + 2\gamma\mu - 2\delta^2 - 2\beta^2. \tag{3-21}$$

It is clear that M satisfies (3-19) if and only if

$$\alpha\delta = 0 \quad \text{and} \quad \alpha\mu - \alpha\gamma + \beta^2 = 0. \tag{3-22}$$

We shall prove that from the first term of (3-22) we have $\delta = 0$. In fact, let us assume that $\delta \neq 0$ on a certain open subset of Ω and we shall get a contradiction. On this subset, from (3-22) we have $\alpha = 0$ and hence $\beta = 0$, which is a contradiction. Therefore, (3-19) is true on Ω if and only if $\delta = 0$ and $\alpha\mu - \alpha\gamma + \beta^2 = 0$. Also, in this case (3-13) becomes

$$\begin{aligned} &U(\mu) = \phi U(\mu) = 0, \quad \kappa_1(\mu - \gamma) = \beta\gamma, \\ &\kappa_2(\mu - \gamma) = 0, \quad U(\gamma) = 0, \quad \phi U(\gamma) = 2\beta\mu. \end{aligned} \tag{3-23}$$

As seen before, from Lemma 3.1 we have $[U, \phi U] = -(\gamma + \mu)\xi - \kappa_1 U - \kappa_2 \phi U$. With the aid of (3-23) and (3-6), taking differentiation of μ along $[U, \phi U]$ we get

$$(\gamma + \mu)\beta\kappa_2 = 0. \tag{3-24}$$

In view of (3-24), next we consider the following subcases.

If $\kappa_2 \neq 0$ holds on some open subset of Ω, it follows from (3-24) that $\gamma + \mu = 0$ on this subset. Moreover, from (3-23) we have $\gamma = \mu$ which is compared with the previous relation giving $\gamma=\mu=0$. However, in this context the second term of (3-22) becomes $\beta=0$, contradicting the assumption. Thus, it follows from (3-24) that $\kappa_2=0$ on Ω.

By virtue of $\delta=\kappa_2=0$, (3-6) becomes $\xi(\mu)=0$ and hence by the first two terms of (3-23) we observe that μ is a constant. In this case, from Lemma 3.1 we have $[U,\phi U]=-(\gamma+\mu)\xi-\kappa_1 U$. Taking differentiation of γ along $[U,\phi U]$, with the aid of (3-23), we obtain $2\mu U(\beta)=-(\gamma+\mu)\xi(\gamma)$ which is simplified by (3-3) and $\delta=0$ giving

$$(\gamma+3\mu)\xi(\gamma)=0. \tag{3-25}$$

If $\gamma+3\mu=0$, from the second and last term of (3-23) we obtain $\mu=0$. However, this is impossible because in this case the second term of (3-22) becomes $\beta=0$, contradicting the assumption. Thus, from (3-25) we obtain $\gamma+3\mu\neq 0$ and hence $\xi(\gamma)=0$ on Ω. In this context, from Lemma 3.1 we have $[\xi,U]=(\kappa_3-\gamma)\phi U$. With the aid of $\xi(\gamma)=0$, (3-23) and the assumption $\beta\neq 0$ on Ω, the action of $[\xi,U]$ on γ gives

$$(\kappa_3-\gamma)\mu=0. \tag{3-26}$$

Because μ is a constant, with the aid of (3-7), (3-8), (3-23) and the assumption $\beta\neq 0$, taking differentiation of the second term of (3-22) along ϕU we get

$$\alpha\mu-\gamma\mu+c+\alpha\gamma+2\beta\kappa_1+\mu\kappa_3-3\mu^2-\gamma\kappa_3=0.$$

Comparing the above equation with (3-4), with the aid of $\delta=0$, we obtain

$$\beta\kappa_1+\tfrac{3}{4}c+\alpha\mu-3\mu^2+\beta^2=0. \tag{3-27}$$

We shall show that (3-26) implies only one case, i.e., $\mu=0$. Otherwise, when $\mu\neq 0$ holds on some open subset of Ω, we have $\kappa_3=\gamma$. Substituting this into (3-4), we have

$$\alpha\gamma+\beta\kappa_1+\tfrac{1}{4}c-\gamma^2-\beta^2=0. \tag{3-28}$$

Substituting the second term of (3-22) into this equation gives $\frac{1}{4}c+\alpha\mu+\beta\kappa_1-\gamma^2=0$, which is compared with (3-27) giving $\frac{1}{2}c+\gamma^2-3\mu^2+\beta^2=0$. In view of $\beta\neq 0$ on Ω and the fact that μ is a constant, taking differentiation of (3-28) along ϕU we get

$$\tfrac{1}{2}c+\alpha\mu+\alpha\gamma+\beta\kappa_1=0.$$

Comparing the above equation with (3-28) we obtain $\frac{1}{4}c+\alpha\mu+\gamma^2+\beta^2=0$, which is simplified by the second term of (3-22) giving

$$\tfrac{1}{4}c+\alpha\gamma+\gamma^2=0. \tag{3-29}$$

By virtue of $\delta = 0$ and (3-23), (3-9) becomes $\phi U(\gamma) = 2\beta\mu$. By applying this and (3-7) and taking differentiation of (3-29) along ϕU we obtain

$$2\alpha\mu + \alpha\gamma + \gamma^2 + \gamma\mu = 0, \tag{3-30}$$

which is combined with (3-29) giving

$$2\alpha\mu + \gamma\mu - \tfrac{1}{4}c = 0. \tag{3-31}$$

Because μ is a constant, taking differentiation of (3-31) along ϕU, with the aid of (3-7) and $\phi U(\gamma) = 2\beta\mu$, we obtain $\alpha + \gamma = 2\mu$. Consequently, substituting this into (3-30) we obtain

$$\alpha = -\tfrac{3}{2}\gamma \quad \text{and} \quad \mu = -\tfrac{1}{4}\gamma. \tag{3-32}$$

Finally, substituting (3-32) into (3-31) we get $\gamma^2 = \frac{1}{2}c$, a constant. Therefore, from the last term of (3-23) we have $\mu = 0$ because $\beta \neq 0$, contradicting our assumption.

Based on the above analyses, it follows from (3-26) that $\mu = 0$. From the second term of (3-13) we have $\gamma(\kappa_1 + \beta) = 0$ and hence $\gamma = 0$. In fact, if $\gamma \neq 0$ holds on some open subset of Ω, it follows that $\kappa_1 + \beta = 0$ and now the application of this and $\mu = 0$ in (3-27) implies $c = 0$, a contradiction. Taking into account $\gamma = \mu = 0$ and $\beta \neq 0$ on Ω, following Lemmas 3.4, 3.6 and the related statement in the proof of Theorem 3.7, we see that Ω coincides with the whole of M and hence the hypersurface is locally congruent to a ruled real hypersurface. However, with the aid of (3-20), one observes easily that ruled hypersurfaces do not satisfy (3-19). Thus, we conclude that the hypersurface M must be Hopf.

Step 2. Let the hypersurface M be Hopf. Using $\beta = 0$ and (3-20), we see that M satisfies (3-19) if and only if

$$\alpha\delta = 0 \quad \text{and} \quad \alpha(\mu - \gamma) = 0. \tag{3-33}$$

Moreover, using $\beta = 0$, (3-13) becomes

$$\begin{gathered} U(\mu) = -2\kappa_1\delta, \qquad \phi U(\mu) = -2\kappa_2\delta, \qquad U(\delta) = \kappa_1(\mu - \gamma), \\ \phi U(\delta) = \kappa_2(\mu - \gamma), \qquad U(\gamma) = 2\kappa_1\delta, \qquad \phi U(\gamma) = 2\kappa_2\delta. \end{gathered} \tag{3-34}$$

First, we show that on M there holds $\gamma = \mu$. If this is not true, then

$$\mathcal{W} = \{q \in M \mid (\gamma - \mu)(q) \neq 0 \text{ in a neighborhood of } q\}$$

is nonempty and an open subset of M. On $\mathcal{W}$, from the second term of (3-33) we have $\alpha = 0$. Using this and $\beta = 0$ on (3-8) we have

$$\delta^2 - \gamma\mu + \frac{c}{4} = 0. \tag{3-35}$$

Applying (3-35) together with $\alpha=\beta=0$ on (3-4) we obtain $\xi(\delta)=(\mu-\gamma)\kappa_3$. From Lemma 3.1 we obtain $[U,\phi U]=-(\gamma+\mu)\xi-\kappa_1 U-\kappa_2\phi U$. With the aid of (3-34), the action of this equation on δ reduces to $(\mu-\gamma)(U(\kappa_2)-\phi U(\kappa_1))=(\gamma-\mu)(\kappa_1^2+\kappa_2^2+(\gamma+\mu)\kappa_3)$. Because on $\mathcal{W}$ we have $\gamma\neq\mu$, it follows that

$$U(\kappa_2)-\phi U(\kappa_1)=-(\kappa_1^2+\kappa_2^2+(\gamma+\mu)\kappa_3),$$

which is simplified by (3-11) giving $\delta^2-\gamma\mu-\frac{c}{2}=0$. Comparing this with (3-35) we get $c=0$, which is a contradiction. This means that $\mathcal{W}$ is empty and we always have $\gamma=\mu$ on M.

Second, we show that on M we have $\delta=0$. If this is not true, then

$$\mathcal{N}=\{q\in M\mid \delta(q)\neq 0 \text{ in a neighborhood of } q\}$$

is nonempty and an open subset of M. It follows from the first term of (3-33) that $\alpha=0$ on $\mathcal{N}$. Notice that on $\mathcal{N}$ equation (3-35) is still true in this situation. The application of this and $\alpha=\beta=0$ on (3-4) gives $\xi(\delta)=0$. Moreover, the application of $\mu=\gamma$ on (3-34) gives $U(\delta)=\phi U(\delta)=0$, that is, δ is a constant.

Using $\alpha=\beta=0$ on $\mathcal{N}$, from (3-3) and (3-6) we have

$$\xi(\gamma)=2\kappa_3\delta \quad\text{and}\quad \xi(\mu)=-2\kappa_3\delta.$$

From the above two relations and (3-34), it is easily seen that $\mu+\gamma$ is a constant. Consequently, in view of constancy of δ, from (3-35) we see that both γ and μ are constant. The application of this on (3-34) and $\xi(\gamma)=2\kappa_3\delta$ yields

$$\kappa_1=\kappa_2=\kappa_3=0$$

on $\mathcal{N}$. However, using the above relations in (3-11) we have $\delta^2-\gamma\mu-\frac{c}{2}=0$, which is compared with (3-35), giving $c=0$, a contradiction. This means that $\mathcal{N}$ is empty and on the whole of M we always have $\delta=0$.

Finally, in view of $\delta=0$ and $\gamma=\mu$, by Lemma 3.1 we obtain $A\phi=\phi A$ on M. Following [Montiel and Romero 1986; Okumura 1975], we observe that the hypersurface M is locally congruent to a type (A) real hypersurface. The converse is easy to check. This completes the proof. □

The classification problem for η-parallel shape operators has existed for a long time, but it is hard to solve. Based on results shown in this section and the introduction, especially those of [Kon and Loo 2011] for dimension greater than three, we propose:

Conjecture 3.11. *A real hypersurface in a nonflat complex space form of complex dimension two has η-parallel shape operator if and only if it is of type (A), (B) or is a ruled real hypersurface.*

4. Cyclic η-parallel Ricci operator

The characterizations of real hypersurfaces in a nonflat complex space form by means of the Ricci operator were studied by many authors; see [Lim et al. 2013a; 2013b; Ki 1988; Kon 2014; 2017; Kwon and Nakagawa 1988; 1989; Niebergall and Ryan 1997]. Among others, Ricci η-parallelism was one of the most often discussed conditions. In this section, as applications of some results in Section 3, we aim to classify three-dimensional real hypersurfaces satisfying cyclic η-parallel Ricci operator. Unlike the case of shape operators, we show that the Ricci cyclic η-parallelism condition is much weaker than Ricci η-parallelism.

Proposition 4.1. *The Ricci operator of a real hypersurface in a nonflat complex space form of complex dimension two is cyclic η-parallel if and only if*

$$U(r-2\alpha\mu)-4\delta\beta\mu-4\alpha\delta\kappa_1=0, \tag{4-1}$$

$$\phi U(r-2\alpha\gamma+2\beta^2)-4\beta\delta^2+4\alpha\delta\kappa_2=0, \tag{4-2}$$

$$\begin{aligned}4(U(\alpha\delta)+\kappa_1(\alpha\gamma-\alpha\mu-\beta^2)+\beta\mu\gamma+\beta\delta^2)&\\ +\phi U(r-2\alpha\mu)-4\beta\mu^2-4\alpha\delta\kappa_2=0,&\end{aligned} \tag{4-3}$$

$$\begin{aligned}4(\phi U(\alpha\delta)+2\beta\mu\delta+\kappa_2(\alpha\gamma-\beta^2-\alpha\mu))&\\ +U(r-2\alpha\gamma+2\beta^2)-4\beta\delta\gamma+4\alpha\delta\kappa_1=0.&\end{aligned} \tag{4-4}$$

Proof. The application of Lemma 3.1 and (3-20) gives

$$\begin{aligned}(\nabla_U Q)U=&\\ \tfrac14(4U(\beta\mu)+\delta(3c-4\alpha\mu+4\gamma\mu-4\delta^2)+4\beta\delta\kappa_1-4\alpha\delta\gamma)\xi&\\ +\tfrac12(U(r-2\alpha\mu)-4\beta\delta\mu-4\alpha\delta\kappa_1)U&\\ +(U(\alpha\delta)+\kappa_1(\alpha\gamma-\alpha\mu-\beta^2)+\beta\gamma\mu+\beta\delta^2)\phi U.&\end{aligned} \tag{4-5}$$

$$\begin{aligned}(\nabla_U Q)\phi U=&\\ \tfrac14(4\alpha\delta^2-4U(\beta\delta)+\gamma(4\alpha\gamma-3c-4\beta^2-4\gamma\mu+4\delta^2)+4\beta\mu\kappa_1)\xi&\\ +(U(\alpha\delta)+\beta\delta^2+\kappa_1(\alpha\gamma-\alpha\mu-\beta^2)+\beta\mu\gamma)U&\\ +\tfrac12(U(r-2\alpha\gamma+2\beta^2)-4\beta\delta\gamma+4\alpha\delta\kappa_1)\phi U.&\end{aligned} \tag{4-6}$$

$$\begin{aligned}(\nabla_{\phi U} Q)U=&\\ \tfrac14(4\phi U(\beta\mu)+\mu(3c+4\gamma\mu-4\delta^2-4\alpha\mu)-4\alpha\delta^2+4\beta\delta\kappa_2)\xi&\\ +\tfrac12(\phi U(r-2\alpha\mu)-4\beta\mu^2-4\alpha\delta\kappa_2)U&\\ +(\phi U(\alpha\delta)+2\beta\delta\mu+\kappa_2(\alpha\gamma-\alpha\mu-\beta^2))\phi U.&\end{aligned} \tag{4-7}$$

(4-8) $(\nabla_{\phi U} Q)\phi U =$

$$\tfrac{1}{4}(4\alpha\mu\delta - 4\phi U(\beta\delta) + 4\beta\mu\kappa_2 + \delta(4\alpha\gamma - 3c - 4\beta^2 - 4\gamma\mu + 4\delta^2))\xi + (\phi U(\alpha\delta) + 2\beta\mu\delta + \kappa_2(\alpha\gamma - \beta^2 - \alpha\mu))U + \tfrac{1}{2}(\phi U(r - 2\alpha\gamma + 2\beta^2) - 4\beta\delta^2 + 4\alpha\delta\kappa_2)\phi U.$$

The Ricci tensor is cyclic η-parallel if and only if

$$g((\nabla_X Q)Y, Z) + g((\nabla_Y Q)Z, X) + g((\nabla_Z Q)X, Y) = 0$$

for any vector fields X, Y, Z orthogonal to ξ. Locally, this is also equivalent to

$$g((\nabla_U Q)U, U) = 0, \quad 2g((\nabla_{\phi U} Q)\phi U, U) + g((\nabla_U Q)\phi U, \phi U) = 0,$$
$$g((\nabla_{\phi U} Q)\phi U, \phi U) = 0, \qquad 2g((\nabla_U Q)U, \phi U) + g((\nabla_{\phi U} Q)U, U) = 0.$$

The proof follows directly from (4-5)–(4-8). □

Some solutions of the system of equations (4-1)–(4-4) are given as follows.

Theorem 4.2. *Let M be a real hypersurface in a nonflat complex space form of complex dimension two with cyclic η-parallel Ricci operator. Then, M satisfies (3-16) if and only if it is of type (A).*

Proof. We assume that $\Omega = \{p \in M \mid \beta(p) \neq 0 \text{ in a neighborhood of } p\}$ is nonempty. Suppose that M has a cyclic η-parallel Ricci operator and satisfies (3-16). As seen in the proof of Theorem 3.8, M satisfies (3-16) if and only if $\delta = 0$ and $\gamma = \mu$. With the aid of this, substituting (4-2) into (4-3) gives

$$\beta(\phi U(\beta) + \kappa_1\beta) = 0. \tag{4-9}$$

Next we show that it follows from (4-9) that Ω is empty. Otherwise, on Ω, from (4-9) we have $\phi U(\beta) = -\kappa_1\beta$, which is used in (3-8), giving

$$2\alpha\mu - 2\mu^2 + \frac{c}{2} + 2\beta\kappa_1 = 0, \tag{4-10}$$

where we have used $\gamma = \mu$ and $\delta = 0$. Applying this again and comparing (4-10) with (3-4) implies $\beta = 0$, which is a contradiction. We have proved that the hypersurface M must be Hopf.

In view of $\beta = 0$ and the above statement, we see that M satisfies (3-16) if and only if $\delta = 0$ and $\gamma = \mu$. In this context, from Lemma 3.1 we have $A\phi = \phi A$ and the proof follows directly from [Montiel and Romero 1986; Okumura 1975].

The converse is easy to check. This completes the proof. □

Remark 4.3. Kim, Kim and Sohn [Kim et al. 2007, Theorem 1.3] proved that real hypersurfaces of dimension > 3 satisfying (3-16) and cyclic η-parallel Ricci operators are of type (A). Our Theorem 4.2 extends their results for real hypersurfaces of dimension three.

Before giving another solution of partial differential equations (4-1)–(4-4), we need the following:

Lemma 4.4 [Maeda 2013; Suh 1990]. *Let M be a connected Hopf real hypersurface in $\mathbb{C}P^n(c)$ or $\mathbb{C}H^n(c)$, $c \neq 0$, $n \geq 2$. If M has η-parallel Ricci tensor, then it is locally congruent to a homogeneous type (A) or (B) hypersurface or a nonhomogeneous real hypersurface with vanishing Hopf principal curvature.*

With the aid of the above result, we have:

Theorem 4.5. *Let M be a real hypersurface in a nonflat complex space form of complex dimension two with cyclic η-parallel Ricci operator. Then, M satisfies* (3-19) *if and only if it is of homogeneous type (A) or (B) real hypersurface or a nonhomogeneous real hypersurface with vanishing Hopf principal curvature.*

Proof. We assume that M is non-Hopf, then

$$\Omega = \{p \in M \mid \beta(p) \neq 0 \text{ in a neighborhood of } p\}$$

is nonempty. Suppose that M has cyclic η-parallel Ricci operator and satisfies (3-19). As seen in the proof of Theorem 3.10, M satisfies (3-19) if and only if

$$\alpha\delta = 0 \quad \text{and} \quad \alpha\mu - \alpha\gamma + \beta^2 = 0. \tag{4-11}$$

If $\delta \neq 0$ holds on some open subset of Ω, it follows from (4-11) that $\alpha = \beta = 0$, contradicting the definition of Ω. Thus, on Ω we have $\delta = 0$.

Because on Ω we have $\beta \neq 0$, with the aid of the second term of (4-11) and $\delta = 0$, subtracting (4-3) from (4-2) we obtain

$$\mu(\gamma - \mu) = 0. \tag{4-12}$$

If $\mu = 0$, with the aid of $\delta = 0$, (4-1)–(4-4) and the second term of (4-11), from (4-5)–(4-8) one can check that the Ricci operator is η-parallel. However, from Lemma 4.4 we see that the hypersurface is Hopf, contradicting the assumption. Otherwise, if $\mu \neq 0$ holds on some open subset of Ω, from (4-12) we have $\gamma = \mu$. On such a subset, using $\delta = 0$ and $\gamma = \mu$, by Lemma 3.1 we obtain $A\phi = \phi A$ and hence M is of type (A) Hopf hypersurface (see [Montiel and Romero 1986; Okumura 1975]), a contradiction. Based on the above statement, we see that Ω is empty and the hypersurface M is Hopf.

In this context, M satisfies (3-19) if and only if

$$\alpha\delta = 0 \quad \text{and} \quad \alpha(\mu - \gamma) = 0. \tag{4-13}$$

Moreover, in this case, with the aid of (4-13), (4-1)–(4-4) become

$$\begin{aligned} U(r - 2\alpha\mu) &= U(r - 2\alpha\gamma) = 0, \\ \phi U(r - 2\alpha\mu) &= \phi U(r - 2\alpha\gamma) = 0. \end{aligned} \tag{4-14}$$

According to (4-13) and (4-14), from (4-5)–(4-8) it is easily seen that the Ricci operator is η-parallel. Applying Lemma 4.4, we see that M is locally congruent to a homogeneous type (A) or (B) real hypersurface or a nonhomogeneous real hypersurface with $A\xi = 0$.

The converse is easy to check. This completes the proof. □

Remark 4.6. Sohn [2007, Theorem 2] proved that real hypersurfaces of dimension greater than three satisfying (3-19) and η-parallel Ricci operators are of type (A) or (B). Later, Lim, Sohn and Song [Lim et al. 2013b] proved that three-dimensional real hypersurfaces satisfying (3-19) and η-parallel Ricci operators are of type (A) or satisfy $A\xi = 0$. Obviously, our Theorem 4.5 is an extension of these results.

Acknowledgements

This work was supported by the Key Scientific Research Program in Universities of Henan Province (No. 20A110023) and the Fostering Foundation of National Foundation in Henan Normal University (No. 2019PL22).

The author would like to thank the reviewer for a careful reading and comments which improved the original paper.

References

[Ahn et al. 1993] S.-S. Ahn, S.-B. Lee, and Y. J. Suh, "On ruled real hypersurfaces in a complex space form", *Tsukuba J. Math.* **17**:2 (1993), 311–322. MR Zbl

[Berndt 1989] J. Berndt, "Real hypersurfaces with constant principal curvatures in complex hyperbolic space", *J. Reine Angew. Math.* **395** (1989), 132–141. MR Zbl

[Cecil and Ryan 1982] T. E. Cecil and P. J. Ryan, "Focal sets and real hypersurfaces in complex projective space", *Trans. Amer. Math. Soc.* **269**:2 (1982), 481–499. MR Zbl

[Cho 2015] J. T. Cho, "Notes on real hypersurfaces in a complex space form", *Bull. Korean Math. Soc.* **52**:1 (2015), 335–344. MR Zbl

[Ki 1988] U.-H. Ki, "Cyclic-parallel real hypersurfaces of a complex space form", *Tsukuba J. Math.* **12**:1 (1988), 259–268. MR Zbl

[Kim 2004] U. K. Kim, "Nonexistence of Ricci-parallel real hypersurfaces in $P_2\mathbb{C}$ or $H_2\mathbb{C}$", *Bull. Korean Math. Soc.* **41**:4 (2004), 699–708. MR Zbl

[Kim et al. 2006] I.-B. Kim, H. J. Park, and W. H. Sohn, "On characterizations of real hypersurfaces with η-parallel Ricci operators in a complex space form", *Bull. Korean Math. Soc.* **43**:2 (2006), 235–244. MR Zbl

[Kim et al. 2007] I.-B. Kim, K. H. Kim, and W. H. Sohn, "Characterizations of real hypersurfaces in a complex space form", *Canad. Math. Bull.* **50**:1 (2007), 97–104. MR Zbl

[Kimura 1986] M. Kimura, "Real hypersurfaces and complex submanifolds in complex projective space", *Trans. Amer. Math. Soc.* **296**:1 (1986), 137–149. MR Zbl

[Kimura and Maeda 1989] M. Kimura and S. Maeda, "On real hypersurfaces of a complex projective space", *Math. Z.* **202**:3 (1989), 299–311. MR Zbl

[Kon 1979] M. Kon, "Pseudo-Einstein real hypersurfaces in complex space forms", *J. Differential Geom.* **14**:3 (1979), 339–354. MR Zbl

[Kon 2014] M. Kon, "3-dimensional real hypersurfaces and Ricci operator", *Differ. Geom. Dyn. Syst.* **16** (2014), 189–202. MR Zbl

[Kon 2017] M. Kon, "3-dimensional real hypersurfaces with η-harmonic curvature", pp. 155–164 in *Hermitian–Grassmannian submanifolds* (Daegu, South Korea, 2016), edited by Y. J. Suh et al., Springer Proc. Math. Stat. **203**, Springer, 2017. MR Zbl

[Kon and Loo 2011] S. H. Kon and T.-H. Loo, "Real hypersurfaces in a complex space form with η-parallel shape operator", *Math. Z.* **269**:1-2 (2011), 47–58. MR Zbl

[Kon and Loo 2012] S. H. Kon and T.-H. Loo, "On characterizations of real hypersurfaces in a complex space form with η-parallel shape operator", *Canad. Math. Bull.* **55**:1 (2012), 114–126. MR Zbl

[Kwon and Nakagawa 1988] J.-H. Kwon and H. Nakagawa, "Real hypersurfaces with cyclic-parallel Ricci tensor of a complex projective space", *Hokkaido Math. J.* **17**:3 (1988), 355–371. MR Zbl

[Kwon and Nakagawa 1989] J.-H. Kwon and H. Nakagawa, "Real hypersurfaces with cyclic η-parallel Ricci tensor of a complex space form", *Yokohama Math. J.* **37**:1 (1989), 45–55. MR Zbl

[Lim et al. 2013a] D. H. Lim, W. H. Sohn, and H. Song, "The Ricci operator and shape operator of real hypersurfaces in a non-flat 2-dimensional complex space form", *Adv. Pure Math.* **3**:2 (2013), 254–258.

[Lim et al. 2013b] D. H. Lim, W. H. Sohn, and H. Song, "A study of real hypersurfaces with Ricci operators in 2-dimensional complex space forms", *Pacific J. Math.* **266**:2 (2013), 305–311. MR Zbl

[Maeda 1976] Y. Maeda, "On real hypersurfaces of a complex projective space", *J. Math. Soc. Japan* **28**:3 (1976), 529–540. MR Zbl

[Maeda 2013] S. Maeda, "Hopf hypersurfaces with η-parallel Ricci tensors in a nonflat complex space form", *Sci. Math. Jpn.* **76**:3 (2013), 449–456. MR Zbl

[Montiel and Romero 1986] S. Montiel and A. Romero, "On some real hypersurfaces of a complex hyperbolic space", *Geom. Dedicata* **20**:2 (1986), 245–261. MR Zbl

[Niebergall and Ryan 1997] R. Niebergall and P. J. Ryan, "Real hypersurfaces in complex space forms", pp. 233–305 in *Tight and taut submanifolds* (Berkeley, CA, 1994), edited by T. E. Cecil and S.-s. Chern, Math. Sci. Res. Inst. Publ. **32**, Cambridge Univ. Press, 1997. MR Zbl

[Okumura 1975] M. Okumura, "On some real hypersurfaces of a complex projective space", *Trans. Amer. Math. Soc.* **212** (1975), 355–364. MR Zbl

[Panagiotidou and Xenos 2012] K. Panagiotidou and P. J. Xenos, "Real hypersurfaces in $\mathbb{C}P^2$ and $\mathbb{C}H^2$ whose structure Jacobi operator is Lie $\mathbb{D}$-parallel", *Note Mat.* **32**:2 (2012), 89–99. MR Zbl

[Pérez et al. 2001] J. d. D. Pérez, Y. J. Suh, and H. Y. Yang, "Characterizations of some pseudo-Einstein ruled real hypersurfaces in complex space forms in terms of Ricci tensor", *Nihonkai Math. J.* **12**:2 (2001), 197–217. MR Zbl

[Sohn 2007] W. H. Sohn, "Characterizations of real hypersurfaces of complex space forms in terms of Ricci operators", *Bull. Korean Math. Soc.* **44**:1 (2007), 195–202. MR Zbl

[Suh 1990] Y. J. Suh, "On real hypersurfaces of a complex space form with η-parallel Ricci tensor", *Tsukuba J. Math.* **14**:1 (1990), 27–37. MR Zbl

[Takagi 1973] R. Takagi, "On homogeneous real hypersurfaces in a complex projective space", *Osaka J. Math.* **10** (1973), 495–506. MR Zbl

[Takagi 1975a] R. Takagi, "Real hypersurfaces in a complex projective space with constant principal curvatures", *J. Math. Soc. Japan* **27** (1975), 43–53. MR Zbl

[Takagi 1975b] R. Takagi, "Real hypersurfaces in a complex projective space with constant principal curvatures, II", *J. Math. Soc. Japan* **27**:4 (1975), 507–516. MR Zbl

[Wang 2018] Y. Wang, "Conformally flat real hypersurfaces in nonflat complex planes", *C. R. Math. Acad. Sci. Paris* **356**:7 (2018), 823–829. MR Zbl

Received June 2, 2018. Revised September 8, 2018.

YANING WANG
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCES
HENAN NORMAL UNIVERSITY
XINXIANG
CHINA
wyn051@163.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.353

FINSLER SPHERES WITH CONSTANT FLAG CURVATURE AND FINITE ORBITS OF PRIME CLOSED GEODESICS

MING XU

In this paper, we consider a Finsler sphere $(M, F) = (S^n, F)$ with dimension $n > 1$ and flag curvature $K \equiv 1$. The action of the connected isometry group $G = I_o(M, F)$ on M, together with the action of $T = S^1$ shifting the parameter $t \in \mathbb{R}/\mathbb{Z}$ of the closed curve $c(t)$, define an action of $\hat{G} = G \times T$ on the free loop space ΛM of M. In particular, for each closed geodesic, we have a $\hat{G}$-orbit of closed geodesics. We assume the Finsler sphere (M, F) described above has only finite orbits of prime closed geodesics. Our main theorem claims that, if the subgroup H of all isometries preserving each close geodesic is of dimension m, then there exists m geometrically distinct orbits $\mathcal{B}_i$ of prime closed geodesics, such that for each i, the union B_i of geodesics in $\mathcal{B}_i$ is a totally geodesic submanifold in (M, F) with a nontrivial H_o-action. This theorem generalizes and slightly refines the one in a previous work, which only discussed the case of finite prime closed geodesics. At the end, we show that, assuming certain generic conditions, the Katok metrics, i.e., the Randers metrics on spheres with $K \equiv 1$, provide examples with the sharp estimate for our main theorem.

1. Introduction

In the recent work of R. L. Bryant, P. Foulon, S. Ivanov, V. S. Matveev and W. Ziller [Bryant et al. 2017], the authors classified Finsler spheres with constant flag curvature $K \equiv 1$ according to the behavior of geodesics. The Katok metric [1973] provides the most important key model for their classification. The celebrated Anosov conjecture [1975], claiming the minimal number of prime closed geodesics on a Finsler sphere (S^n, F) is $2[(n+1)/2]$, was based on the discovery of Katok metrics with only finite prime closed geodesics. There are many works using Morse theory and index theory to study the closed geodesics and Anosov conjecture in

Supported by National Natural Science Foundation of China (No. 11821101, No. 11771331), Beijing Natural Science Foundation (No. 00719210010001, No. 1182006), Capacity Building for Sci-Tech Innovation — Fundamental Scientific Research Funds (No. KM201910028021).
MSC2010: 22E46, 53C22, 53C60.
Keywords: Katok metric, Randers sphere, constant flag curvature, orbit of closed geodesics, totally geodesic submanifold, fixed point set.

Finsler geometry, assuming a pinch condition for the flag curvature, nondegenerating property for all closed geodesics, or using the specialty of low dimensions. See for example [Bangert and Long 2010; Duan 2016; Long and Duan 2009; Duan et al. 2016; Rademacher 1989; Wang 2012; 2015]. From the geometrical point of view, it was much later that people noticed that Katok metrics are Randers metrics on spheres with constant flag curvature [Rademacher 2004]. D. Bao, C. Robles and Z. Shen [Bao et al. 2004] provided a complete classification for all Randers metrics with constant flag curvature. The classification for the non-Randers case is still widely open. Bryant [1996; 1997; 2002]. provided many important examples of Finsler spheres with $K \equiv 1$.

However, one of the most important technique in [Bryant et al. 2017] is from Lie theory. The authors considered the antipodal map ψ for a Finsler sphere with $K \equiv 1$ (see [Bryant et al. 2017; Shen 1996] or Section 2 for its definition). It is a Clifford Wolf translation in the center of the isometry group $I(M, F)$. When ψ has an infinite order, after taking closure, it can be used to generate a closed abelian subgroup of isometries with a positive dimension.

For nonzero Killing vector fields on a Finsler sphere with $K \equiv 1$, we have the following totally geodesic technique. The common zero point set of Killing vector fields, or more generally the fixed point set of isometries, provide closed totally geodesic submanifolds. In particular, when the dimension of such a submanifold is one, it is a reversible geodesic, and when the dimension is even bigger, it is a Finsler sphere inheriting the curvature property and geodesic property from the ambient space. We can use this key observation to set up an inductive argument, when studying the geodesics on (S^n, F) with $n > 2$ and $K \equiv 1$, and generalizing some results in [Bryant et al. 2017] to high dimensions.

For example, we have proved the following lower bound estimate for the number of reversible prime closed geodesics in Finsler spheres with constant flag curvature.

Theorem 1.1 [Xu 2018b]. *Let $(M, F) = (S^n, F)$ with $n > 1$ be a Finsler sphere with $K \equiv 1$ and only finite prime closed geodesics. Then the number of geometrically distinct reversible closed geodesics is at least* $\dim I(M, F)$.

Recall that a geodesic $c(t)$ with constant speed is called *reversible* if $c(-t)$ also provides a geodesic with constant speed after a reparametrization by the new arc length. Two geodesics are *geometrically distinct* if and only if they are different subsets.

The assumption of only finite prime closed geodesics imposes a strong restriction on $I_o(M, F)$, which can only be a torus. A lot of important examples are excluded, for example, the standard unit spheres and the homogeneous non-Riemannian Randers spheres with $K \equiv 1$. So if we want more possibility for $I_o(M, F)$, the geodesic condition could be replaced by the assumption that there exist only finite orbits of prime closed geodesics, or Assumption (F) for simplicity. See Section 3 for its precise definition and detailed discussion.

The main purpose of this paper is to prove the following theorem.

Theorem 1.2. *Let $(M, F) = (S^n, F)$ be a Finsler sphere satisfying $n > 1$, $K \equiv 1$ and Assumption (F). Denote by H the subgroup of $G = I_o(M, F)$ preserving each closed geodesic, H_o its identity component and $m = \dim H$. Then there exist at least m geometrically distinct orbits $\mathcal{B}_i$'s of prime closed geodesics such that each union B_i of geodesics in $\mathcal{B}_i$ is a totally geodesic submanifold in M with a nontrivial H_o-action.*

When (M, F) has only finite prime closed geodesics, then Assumption (F) is satisfied, $H_o = G = I_o(M, F)$, and each orbit of closed geodesics consists of only one closed geodesic. So Theorem 1.2 generalizes Theorem 1.1. It even slightly refines Theorem 1.1 by claiming the totally geodesic B_i's found have nontrivial H_o-actions. So if the common zero point of H_o has a positive dimension, it provides one more totally geodesic B_i, which is either a reversible closed geodesic which length is a rational multiple of π, or isometric to a standard unit sphere.

By Theorem 1.1 and Theorem 1.2 in [Xu 2018a], each submanifold $(B_i, F|_{B_i})$ is in fact a non-Riemannian homogeneous Randers sphere with constant flag curvature. So Theorem 1.2 implies the existence of totally geodesic subspheres in which F has standard restrictions, though F itself may be strange.

This paper is organized as following. In Section 2, we recall some fundamental geometric properties of Finsler spheres with $K \equiv 1$, discussing their antipodal maps and totally geodesic submanifolds. In Section 3, we define Assumption (F), i.e., the assumption of only finite prime closed geodesics. In Section 4, we introduce the subgroup H of isometries which preserves each closed geodesic. In Section 5, we prove Theorem 1.2 by induction. In Section 6, we discuss the Katok metrics, and show that in some cases they provides examples for Theorem 1.2, with a sharp estimate.

2. Preliminaries: from antipodal map to Killing vector field

Let $(M, F) = (S^n, F)$ be a Finsler sphere satisfying the dimension $n > 1$ and the flag curvature $K \equiv 1$. Denote $G = I_o(M, F)$ the connected isometry group, i.e., the identity component of the isometry group $I(M, F)$ of (M, F).

We briefly recall the definition of the exponential map [Bao et al. 2000] and the antipodal map ψ [Bryant et al. 2017; Shen 1996] for (M, F).

For any $x \in M$ and nonzero $y \in T_xM$, the *exponential map* $\mathrm{Exp}_x : T_xM \to M$ is defined by $\mathrm{Exp}_x(y) = c(1)$ where $c(t)$ is the constant speed geodesic with $c(0) = x$ and $\dot{c}(0) = y$. When $y = 0 \in T_xM$, we define $\mathrm{Exp}_x(0) = x$. Notice that Exp_x is C^1 at $y = 0$ and C^∞ elsewhere.

The discussion for the Jacobi fields and conjugation points when $K \equiv 1$ indicates Exp_x maps the sphere

$$S_o^F(\pi) = \{y \in T_xM \mid F(y) = \pi\} \subset T_xM$$

to a single point $x^* \in M$. The map from x to x^* is an isometry of (M, F) in the center of $I(M, F)$ [Bryant et al. 2017]. Further more, it is easy to see that ψ is a Clifford Wolf translation for the (possibly nonreversible) distance $d_F(\cdot, \cdot)$ defined by the Finsler metric F. We will call it the *antipodal map* and always denote it as ψ. It is a generalization for the antipodal map for standard unit spheres but may not be an involution any more.

The above description immediately proves that any connected and simply connected Finsler manifold (M, F) with $\dim M > 1$ and $K \equiv 1$ is homeomorphic to a sphere. A more careful discussion with the local charts shows that the homeomorphism in this statement can be refined to be a diffeomorphism, and the argument is valid not only for M, but also any closed connected totally geodesic submanifold N with $\dim N > 1$, i.e., we have the following lemma.

Lemma 2.1 [Xu 2018b, Lemma 3.2]. *Let (M, F) be a closed connected and simply connected Finsler manifold with $K \equiv 1$ and N a closed connected totally geodesic submanifold with* $\dim N > 1$. *Then both M and N are diffeomorphic to standard spheres, and N is an imbedded submanifold in M.*

The fixed point set for a family of isometries in $I(M, F)$ is a closed, possibly disconnected, totally geodesic submanifold. We have the following lemma, indicating the connectedness of N, when its dimension is positive.

Lemma 2.2 [Xu 2018b, Lemma 3.5]. *Let $(M, F) = (S^n, F)$ be a Finsler sphere with $n > 1$ and $K \equiv 1$, and N the fixed point set of a family of isometries of (M, F). Then N must satisfy one of the following:*

(1) *N is a two-points ψ-orbit, i.e., $N = \{x', x''\}$ with $d_F(x', x'') = d_F(x'', x') = \pi$.*

(2) *N is a reversible closed geodesic.*

(3) *$(N, F|_N)$ is a Finsler sphere with* $\dim N > 1$ *and $K \equiv 1$.*

The space of Killing vector fields can be viewed as the Lie algebra of $I(M, F)$. So the common zero set of a family of Killing vector fields on (M, F) is a special case of fixed point sets for isometries.

In later discussions, we will need the following two lemmas for Killing vector fields.

Lemma 2.3. *Assume that X is a Killing vector field of the Finsler space (M, F), $f(\cdot) = F(X(\cdot))$ and $f(x) > 0$ at $x \in M$. Then the integration curve of X passing x is a geodesic if and only if x is a critical point of $f(\cdot)$.*

Lemma 2.4 (corollary of [Deng and Xu 2014, Lemma 3.1]). *Assume that $c = c(t)$ is a geodesic of positive constant speed on the Finsler space (M, F). Then restricted to $c(t)$, any Killing vector field X of (M, F) satisfies*

$$\langle X(c(t)), \dot{c}(t)\rangle^F_{\dot{c}(t)} \equiv \text{const}, \tag{2-1}$$

where $\langle u, v\rangle^F_y = g_{ij}(y)u^i v^j$ for $u, v, y \in T_x M$ and $y \neq 0$ is the inner product defined by the fundamental tensor.

Proof. Whenever the value of X is linearly independent of $\dot{c}(t)$, we can prove (2-1) by choosing a special local chart, such that $c = c(t)$ can be presented as $x^1 = t$ and $x^i = 0$ for $i > 1$, and $X = \partial_{x^2}$. Because X is Killing vector field, $F(x, y)$ is independent of x^2. The condition that $c = c(t)$ is a geodesic implies that for the coefficients $\boldsymbol{G}^i$ of the geodesic spray, we have

$$\begin{aligned}\boldsymbol{G}^i(c(t), \dot{c}(t)) &= \tfrac{1}{4}g^{il}([F^2]_{x^m y^l}y^m - [F^2]_{x^l})\\ &= \tfrac{1}{4}g^{il}([F^2]_{x^1 y^l} - [F^2]_{x^l}) = 0.\end{aligned}$$

In particular, on the geodesic $c = c(t)$, we have

$$\frac{d}{dt}\langle X(c(t)), \dot{c}(t)\rangle^F_{\dot{c}(t)} = \tfrac{1}{2}[F^2]_{x^1 y^2} = \tfrac{1}{2}[F^2]_{x^2} = 0,$$

which proves the lemma in this case.

When X is tangent to $c = c(t)$ for t in an interval I, we can easily get (2-1) for $t \in I$.

Summarizing this two cases and using the continuity, we have proved (2-1) along the whole geodesic $c = c(t)$. □

3. Orbit of closed geodesics and Assumption (F)

Now we define *Assumption (F)*, i.e., the condition that (M, F) has only finite orbits of prime closed geodesics. In later discussion, we will always assume it to be satisfied by (M, F) unless otherwise specified.

The free loop space ΛM of all piecewise smooth path $c = c(t)$ with $t \in \mathbb{R}/\mathbb{Z}$ (sometimes we will simply denote it as c or γ) admits the natural actions of $\hat{G} = G \times T$ such that

$$((g, t')\cdot c)(t) = g\cdot c(t + t'), \quad \text{for all } t.$$

So for each closed geodesic γ of constant speed, we have an $\hat{G}$-orbit $\hat{G}\cdot\gamma$ of closed geodesics with the same speed. The geodesic $c(t)$ (with $t \in \mathbb{R}/\mathbb{Z}$) is *prime*, i.e.,

$$\min\{t \mid t > 0 \text{ and } c(t) = c(0)\} = 1,$$

if and only if all the closed geodesics in $\hat{G}\cdot c$ are prime.

Definition 3.1. We say (M, F) has only finite orbits of prime closed geodesics if it satisfies

Assumption (F) all the prime closed geodesics of positive constant speed can be listed as a finite set of $\hat{G}$-orbits, $\mathcal{B}_i = \hat{G} \cdot \gamma_i,\ 1 \le i \le k$.

In Definition 3.1, we can equivalently list all the closed geodesics of constant speed $c(t)$ with $t \in \mathbb{R}/\mathbb{Z}$ as $\mathcal{B}_i^j = \hat{G} \cdot \gamma_i^j,\ 1 \le i \le k,\ j \in \mathbb{N}$. The orbit $\mathcal{B}_i$ in Definition 3.1 coincides with $\mathcal{B}_i^1$, for each i. The closed geodesics γ_i^j is the one which rotates j-times along the prime closed geodesic γ_i in Definition 3.1, i.e., if γ_i is presented as $c_i = c_i(t)$, then γ_i^j is $c_{i,j}(t) = c_i(jt)$.

We denote B_i the union of the geodesics in $\mathcal{B}_i$ or $\mathcal{B}_i^j$ for any $j \in \mathbb{N}$. Then we call $\mathcal{B}_i^j$ and $\mathcal{B}_{i'}^{j'}$ *geometrically distinct* (or *geometrically the same*), if B_i and $B_{i'}$ are different subsets (or the same subsets, respectively) of M.

The Assumption (F) for the ambient space can be inherited by some totally geodesic submanifolds, i.e., we have the following lemma.

Lemma 3.2. *Let (M, F) be any closed compact Finsler manifold satisfying Assumption (F), ϕ_α with $\alpha \in \mathcal{A}$ a family of isometries in the center of $I(M, F)$, and N the fixed point set for all ϕ_α's. Then each orbit of prime closed geodesic for $(N, F|_N)$ is also an orbit of prime closed geodesic for (M, F). In particular, $(N, F|_N)$ also satisfies Assumption (F).*

Proof. The fixed point set N for the isometries ϕ_α with $\alpha \in \mathcal{A}$ is a closed (possibly disconnected) totally geodesic submanifold of (M, F). Because each ϕ_α commutes with all isometries of (M, F), the fixed point set N for all ϕ_α's is preserved by the action of $G = I_o(M, F)$. The restriction of G-action to N defines isometries in $G' = I_o(N, F|_N)$. Denote $\hat{G}' = G' \times T$. Then for each prime closed geodesic γ in N, Assumption (F) implies that $\hat{G}' \cdot \gamma$ is a disjoint finite union of $\hat{G}$-orbits. Both $\hat{G}'$-orbits and $\hat{G}'$-orbits are compact and connected, so we get $\hat{G}' \cdot \gamma = \hat{G} \cdot \gamma$, which proves the first claim. The second claim follows immediately. □

The effect of Assumption (F) can be seen from the behavior of the antipodal map ψ. For example, when ψ has a finite order k, i.e., there exists a positive integer k, such that

$$\psi^k = \mathrm{id}, \quad \text{and} \quad \psi^i \neq \mathrm{id} \quad \text{when } 1 \le i < k,$$

we have the following lemma.

Lemma 3.3. *Let $(M, F) = (S^n, F)$ be a Finsler sphere satisfying $n > 1$, $K \equiv 1$ and Assumption (F). Assume that the antipodal map ψ has a finite order k. Then F must be the Riemannian metric for a standard unit sphere.*

Proof. Because ψ is a Clifford Wolf translation, and it has a finite order k, each geodesic of (M, F) is closed, and each prime closed geodesic admits a suitable multiple such that the length of the resulting closed geodesic is $k\pi$. By Assumption (F), the subset $B \subset \Lambda M$ of all closed geodesics with the length $k\pi$ can be listed as the disjoint union of $\mathcal{B}_i^{n_i} = \hat{G} \cdot \gamma_i^{n_i}$, $1 \leq i \leq k$, where each γ_i is a prime closed geodesic. Obviously B is connected and each $\mathcal{B}_i^{n_i}$ is compact, so we must have $k = 1$.

Then we prove (M, F) is G-homogeneous. Assume conversely that it is not, we consider a unit speed geodesic $c(t)$, and the G-orbit N passing $c(0)$, such that

$$\langle \dot{c}(0), T_{c(0)}N \rangle^F_{\dot{c}(0)} = 0. \tag{3-1}$$

Then by Lemma 2.4, for any Killing vector field $X \in \mathfrak{g}$, we have

$$\langle \dot{c}(t), X(c(t)) \rangle^F_{\dot{c}(t)} \equiv 0,$$

i.e., $c(t)$ meets each G-orbit orthogonally in the sense of (3-1). This property is preserved by $\hat{G}$-actions. So its $\hat{G}$-orbit can not exhaust all the geodesics, for example, those which does not satisfy (3-1). This is a contradiction to our previous observation that (M, F) can only have one orbit of prime closed geodesics, and it proves that (M, F) is homogeneous Finsler sphere.

Finally, we prove (M, F) is a standard unit sphere. Because (M, F) is a homogeneous Finsler space, it has at least one homogeneous geodesic $c(t) = \exp(tX) \cdot o$, in which $o \in M$ and $X \in \mathfrak{g} = \mathrm{Lie}(G)$ [Yan and Huang 2018]. Our previous observation that all geodesics belong to a single $\hat{G}$-orbit implies all geodesics are homogeneous. So for any $x \in M$ and any two F-unit tangent vectors y_1 and y_2 in T_xM, we have two unit speed geodesics $c_1(t)$ and $c_2(t)$ such that $c_1(0) = c_2(0) = x$ and $\dot{c}_i(0) = y_i$. Both geodesics belong to the same $\hat{G}$-orbit, so we can find $g_1 \in G$ such that $(g_1 \cdot c_1)(t) \equiv c_2(t + t_0)$ for some fixed t_0. Because the geodesic $c_2(t)$ is homogeneous, we can find another $g_2 \in G$ such that $(g_2 \cdot c_2)(t) = c_2(t - t_0)$. Then we have

$$(g_2 g_1 \cdot c_1)(t) = (g_2 \cdot c_2)(t + t_0) = c_2(t), \quad \text{for all } t.$$

So the isotropy action for (M, F) is transitive at each point. The only homogeneous spheres satisfying this property are Riemannian spheres of constant curvature. □

Using Lemmas 3.2 and 3.3, we can generalize Lemma 3.6 in [Xu 2018b] to the following.

Lemma 3.4. *Let $(M, F) = (S^n, F)$ be a Finsler sphere satisfying $n > 1$, $K \equiv 1$ and Assumption (F). Then the union N of all the finite orbits of ψ in M must be one of the following*:

(1) *A two-points ψ-orbit.*

(2) *A closed reversible geodesic which length is rational multiple of π.*

(3) *A Riemannian sphere of constant curvature isometrically imbedded in* (M, F) *as a totally geodesic submanifold. In this case we have* $k = 2$.

Proof. By the same argument as in the proof of Lemma 3.6 in [Xu 2018b], we can prove N is the fixed point set of ψ^k for some integer k, hence it is totally geodesic in (M, F). When $\dim N = 0$ or 1, we get the cases (1) and (2) respectively. The difference appears when $\dim N > 1$, which may happen with the finite orbit of prime closed geodesics condition. When $\dim N > 1$, by Lemma 2.2, $(N, F|_N)$ is a Finsler sphere satisfying $K \equiv 1$. By Lemma 3.2, $(N, F|_N)$ also satisfies Assumption (F). Then Lemma 3.3 provides the case (2) in the lemma. □

The cases (2) and (3) cover all the possibilities for the $\hat{G}$-orbit of a prime closed geodesic γ such that the length of γ is a rational multiple of π.

Next, we consider the $\hat{G}$-orbit of a prime closed geodesic γ such that the length of γ is an irrational multiple of π.

When the length of γ is an irrational multiple of π, any ψ-orbit in γ is dense. Following this observation, we can easily prove the following lemma.

Lemma 3.5. *Let* $(M, F) = (S^n, F)$ *be a Finsler sphere satisfying* $n > 1$, $K \equiv 1$ *and Assumption (F). Then two geometrically distinct closed geodesics can intersect if and only if they are intersecting geodesics in the totally geodesic submanifold in* (M, F) *which is isometric to a unit sphere, i.e., the case* (3) *in Lemma 3.4.*

Proof. Lemma 3.4 indicates that any two geometrically distinct closed geodesics γ_1 and γ_2 must satisfy one of the following. Either both lengths are 2π or one of them, for example γ_1, has a length which is an irrational multiple of π. In the first case, they are contained in a totally geodesic submanifold of (M, F) which is isometric to a unit sphere. In the second case, the intersection of the two geodesics contains a ψ-orbit, which is dense in γ_1. Both geodesics are closed, so does their intersection. So as subsets of M, we have $\gamma_1 \subset \gamma_2$ and furthermore the equality must happen because γ_2 is a closed connected curve. This is the contradiction ending the proof of the lemma. □

Using above lemmas, we can provide more explicit description for the orbits of prime closed geodesics by the following lemma.

Lemma 3.6. *Assume* $(M, F) = (S^n, F)$ *is a Finsler sphere satisfying* $n > 1$, $K \equiv 1$, *Assumption (F), and that it is not the standard unit sphere. Then we have the following*:

(1) *There exists closed geodesics whose lengths are irrational multiples of* π.

(2) *For the orbit of prime closed geodesics* $\mathcal{B}_i = \hat{G} \cdot \gamma_i$ *such that the length of* γ_i *is an irrational multiple of* π, *the corresponding* B_i *is an orbit for the action of* $G = I_o(M, F)$.

(3) *Two different orbits of prime closed geodesics*, $\mathcal{B}_i$ *and* $\mathcal{B}_j$, *are geometrically distinct if and only if* B_i *and* B_j *do not intersect.*

(4) *Two different orbits of prime closed geodesics* $\mathcal{B}_i$ *and* $\mathcal{B}_j$ *are geometrically the same if and only if we can find* $\gamma_i \in \mathcal{B}_i$ *and* $\gamma_j \in \mathcal{B}_j$ *such that* γ_i *and* γ_j *are the same curve with different directions.*

Proof. By Lemma 3.3 and the assumption that (M, F) is not the standard unit sphere, the antipodal map ψ generates an infinite subgroup in $I(M, F)$, which closure is a subgroup in the center of $I(M, F)$, corresponding to an abelian subalgebra $\mathfrak{c}' \subset \mathfrak{c}(\mathfrak{g})$ with $\dim \mathfrak{c}' > 0$. We can find a nonzero Killing vector field X from $\mathfrak{c}'$ which generates an S^1. Obviously, X is tangent to each closed geodesic. The restriction of X to each closed geodesic which length is a rational multiple of π is zero.

To prove (1), we only need to consider a maximum point x of $f(\cdot) = F(X(\cdot))$. By Lemma 2.3, the integration curve γ of X passing x is a geodesic, restricted to which X is nonzero. Because X generates an S^1, γ is closed. So it is a closed geodesic which length is an irrational multiple of π.

To prove (2), we consider a prime closed geodesic γ_i which length is an irrational multiple of π. Because the restriction of X to γ_i is a nonzero tangent vector field, γ_i is a homogeneous geodesic. In its $\hat{G} = G \times T$-orbit, The T-action on γ_i can be replaced by the actions of $\exp(tX) \in G$. So the union B_i for the geodesics in $\mathcal{B}_i$ is a G-orbit.

The statements (3) and (4) follows immediately from Lemma 3.5. □

Corollary 3.7. *Assume* $(M, F) = (S^n, F)$ *is a homogeneous Finsler sphere satisfying* $n > 1$, $K \equiv 1$ *and Assumption (F). Then all closed geodesics are reversible. Furthermore, one of the following two cases must happen*:

(1) (M, F) *is a standard unit sphere. It has exactly one orbit of prime closed geodesics and all geodesics are closed.*

(2) (M, F) *is a homogeneous non-Riemannian Randers sphere with an odd* n *and* $K \equiv 1$. *There exists exactly two orbits of prime closed geodesics* $\hat{G} \cdot \gamma_1$ *and* $\hat{G} \cdot \gamma_2$, *in which* γ_1 *and* γ_2 *are the same curve with different directions.*

Proof. If the antipodal map ψ has a finite order, then (M, F) is the standard unit sphere by Lemma 3.3. If ψ has an infinite order, then $G = I_o(M, F)$ has a one-dimensional center $\mathbb{R}X$, and $M = G/H$ must be

$$U(n')/U(n'-1) \quad \text{or} \quad Sp(n'')U(1)/Sp(n''-1)U(1).$$

By Theorems 1.1 and 1.2 in [Xu 2018a], when $K \equiv 1$, (M, F) is a geodesic orbit Finsler sphere and must be Randers. Integration curves of X and $-X$ provide prime closed geodesics whose lengths are different irrational multiples of π, belonging to two different orbits $\mathcal{B}_1$ and $\mathcal{B}_2$ with $B_1 = B_2 = M$. By Lemma 3.6, They are the only orbits of prime closed geodesics. □

4. Isometries preserving each closed geodesic

Assume $(M, F) = (S^n, F)$ is a Finsler sphere satisfying $n > 1$, $K \equiv 1$, and Assumption (F). Let ψ be its antipodal map. By Lemma 3.3, the case that ψ has a finite order is easy, so in the following discussion we assume that ψ has an infinite order.

Let H denote the subgroup of $G = I(M, F)$ which preserves each closed geodesic, H_o its identity component, and $\mathfrak{h}$ its Lie algebra. The group H is intersection of

$$G_\gamma = \{g \in G \mid (g \cdot \gamma)(t) \equiv \gamma(t + t_0) \text{ for some } t_0\}$$

for all closed geodesics γ. Each G_γ is a closed subgroup of G. So is H.

It should be remarked that the claim that G_γ is a closed subgroup of G is an easy fact in this case because γ is closed. In the recent work [Berestovskii and Nikonorov 2019], it has been proved that G_γ is still a Lie group when γ is not closed.

Obviously the antipodal map ψ belongs to H. Because ψ has an infinite order, then after taking closure, it generates an abelian subgroup of positive dimension, i.e., we have $\dim H > 0$. The following lemma claims that H_o commutes with all the G-actions.

Lemma 4.1. *The subgroup H_o is a closed subgroup in the center of $G = I_o(M, F)$.*

Proof. The previous observations have already proved that H_o is a closed subgroup of G. Because G is a compact Lie group, to prove this lemma we only need to prove $\mathfrak{h} = \mathrm{Lie}(G)$ is an abelian ideal of $\mathfrak{g}$.

The Lie algebra $\mathfrak{h} = \mathrm{Lie}(H)$ consists of all the Killing vector fields X which is tangent to each closed geodesic. Because the action of G permutes the closed geodesics in each orbit of prime closed geodesics, any Killing vector field of the form $\mathrm{Ad}(g)X$ for $g \in G$ and $X \in \mathfrak{h}$ is also tangent to each closed geodesic. So conjugations of G preserves $\mathfrak{h}$, i.e., $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.

Then we prove $\mathfrak{h}$ is abelian by contradiction. Assume conversely that $\mathfrak{h}$ is not abelian, then we can find a nonzero vector X from the compact semisimple Lie algebra $[\mathfrak{h}, \mathfrak{h}]$ which generates an S^1-subgroup. The Killing vector field on (M, F) induced by X has trivial restriction on each closed geodesic. By Lemma 2.3, the integration curve of X passing the maximum point of $f(\cdot) = F(X(\cdot))$ is a closed geodesic. This is a contradiction which ends the proof of this lemma. □

A direct consequence of Lemma 4.1 is the following lemma.

Lemma 4.2. *For any Killing vector field $X \in \mathfrak{h}$ and any orbit $\mathcal{B}_i$ of the prime closed geodesic $c = c(t)$, their exists a constant $\rho_{X,i} \in \mathbb{R}$ such that*

$$X|_{c(t)} \equiv \rho_{X,i}\dot{c}(t), \quad \textit{for all } c \in \mathcal{B}_i. \tag{4-1}$$

In particular, a Killing vector field $X \in \mathfrak{h}$ vanishes at some point $x \in \mathcal{B}_i$ if and only if $\rho_{X,i} = 0$, and if and only if X vanishes identically on B_i.

The last ingredient for the proof of Theorem 1.2 is the following lemma.

Lemma 4.3. *Let $(M, F) = (S^n, F)$ be a Finsler sphere satisfying $n > 1$, $K \equiv 1$ and Assumption (F). Then we have the following*:

1. *For any nonzero Killing vector field $X \in \mathfrak{h}$ which generates an S^1, there exists some orbit $\mathcal{B}_i$ of prime closed geodesics such that $\rho_{X,i} > 0$.*
2. *Any Killing vector field $X \in \mathfrak{h}$ vanishing on all closed geodesics must be a zero vector field.*
3. *The common zero set of all Killing vector fields in $\mathfrak{h}$ must be the fixed point set of ψ^k for some integer k. To be more precise, it is empty, a two-points ψ-orbit, some B_i which is a reversible closed geodesic which lengths for both directions are rational multiples of π, or some B_i which is a totally geodesic submanifold isometric to a standard unit sphere.*

Proof. (1) We consider the maximum point x for the function $f(\cdot) = F(X(\cdot))$. By Lemma 2.3, the integration curve of X passing x provide a prime closed geodesic γ, for which we have $X(c(t)) \equiv \rho_{X,\gamma}\dot{c}(t)$ with $\rho_{X,\gamma} > 0$.

(2) We assume conversely that there exists a nonzero Killing vector field on (M, F) such that it vanishes on all closed geodesics. Let $\mathfrak{k}$ be the space of all such Killing vector fields. It is a subalgebra of $\mathfrak{h}$ corresponding to a subtorus in H_o. We can find a nonzero Killing vector field X from $\mathfrak{k}$ which generates an S^1. The argument for (1) indicates X is not vanishing on some closed geodesic, which is the contradiction.

(3) Let N be the fixed point set of H_o, and assume N is not empty. By Lemma 2.2, N must be a two-points ψ-orbit, a reversible closed geodesic, or a Finsler sphere with $\dim N > 1$, $K \equiv 1$ isometrically imbedded in (M, F).

Obviously the action of ψ preserves N, i.e., N consists of ψ-orbits. Because H is compact, H/H_o is finite. We also have $\psi \in H$, and thus each ψ-orbit in N is finite. So when $\dim N = 1$, the lengths of N for both directions are rational multiples of π.

When $\dim N > 1$, we see $(N, F|_N)$ satisfies Assumption (F) by Lemma 3.2. Then Lemma 3.3 tells us that $(N, F|_N)$ is a standard unit sphere. ☐

5. Proof of Theorem 1.2

Now we are ready to prove Theorem 1.2, which applies a similar inductive argument as that for Theorem 1.2 in [Xu 2018b].

When ψ has a finite order, then by Lemma 3.3, (M, F) is the standard unit sphere. Obviously Theorem 1.2 is valid in this case. So in the following discussion, we assume ψ has an infinite order, and thus we have $m = \dim H > 0$.

We will prove Theorem 1.2 by an induction for $n = \dim M$.

When $n = 2$ and the antipodal map ψ has an infinite order, H_o coincides with $G = I_o(M, F) = S^1$. In [Bryant et al. 2017], it has been proved that geometrically there exists exactly one reversible closed geodesic γ with a nontrivial H_o-action. So Theorem 1.2 is valid in this case, and the estimate is sharp.

Now we assume Theorem 1.2 is valid when $n < l$ with $l > 3$ (the inductive assumption) and we will prove the theorem when $n = l$.

Firstly, we prove:

Claim 1. *When* $\dim H = 1$, *there exists at least one totally geodesic* B_i *with a nontrivial* H_o*-action.*

Let X be any nonzero Killing vector field from $\mathfrak{h} = \mathrm{Lie}(H)$. We list all the $\hat{G}$-orbits of prime closed geodesics as $\mathcal{B}_i$ with $1 \le i \le k$, such that when $1 \le i \le k'$ the coefficient $\rho_{X,i}$ in (4-1) is positive. Notice that by Lemma 4.3(1), we have $k' > 0$.

If the antipodal map ψ is not contained in H_o, we can find an isometry of (M, F) which is of the form $\phi = \psi \exp(t'X)$ such that its fixed point set contains B_1. By Lemma 2.2 (or see Lemma 3.5 in [Xu 2018b]), the fixed point set N of ϕ is a closed connected totally geodesic submanifold. It must have a positive codimension in M because $\phi \notin H_o$. When $\dim N = 1$, it is a reversible closed geodesic. When $\dim N > 1$, by Lemma 3.2 and the totally geodesic property, $(N, F|_N)$ is a Finsler sphere satisfying $K \equiv 1$ and Assumption (F). Using the inductive assumption, we can find some orbit of prime closed geodesic, $\mathcal{B}_i = \hat{G}' \cdot \gamma_i = \hat{G} \cdot \gamma_i$, where $\hat{G}' = G' \times T$ and $G' = I_o(N, F|_N)$, such that the corresponding $\mathcal{B}'_i$, is totally geodesic in $(N, F|_N)$ as well as in (M, F). The H_o-action on B_i is nontrivial because

$$\exp(t'X)|_{B_i} = \psi^{-1}\phi|_{B_i} = \psi^{-1}|_{B_i},$$

and ψ has no fixed point on any closed geodesic.

To summarize, this proves Claim 1 when $\psi \notin H_o$.

To continue the proof of Claim 1, we may assume $\psi \in H_o$. In this case, we can prove the zero set of X is empty as following. Assume conversely that the zero set of X is not empty, by Lemma 4.3, it is a two-points ψ-orbit, a reversible closed geodesic, or a connected totally geodesic standard unit sphere. For each possibility,

ψ can not be generated by X, which is a contradiction to the assumption $\psi \in H_o$. This fact implies that $f(\cdot) = F(X(\cdot))$ is a smooth function on M. By Lemma 2.3, the critical point set of $f(\cdot)$ consists of exactly all B_i's with $1 \leq i \leq k'$. Meanwhile, we see the H_o-action on each closed geodesic is nontrivial.

We take a prime closed geodesic $c_i(t)$ with $t \in \mathbb{R}/\mathbb{Z}$ from $\mathcal{B}_i$ for $1 \leq i \leq k'$, then $X|_{c_i} = \rho_{X,i}\dot{c}_i$ with $\rho_{X,i} > 0$. Because $H_o = S^1$, we can find some $t' > 0$ such that $\exp(t'X) = \mathrm{id}$, then we have

$$n_i = t'\rho_{X,i} \in \mathbb{N}, \quad \text{for all } 1 \leq i \leq k'.$$

We may reorder these c_i's such that

$$n_1 \leq n_2 \leq \cdots \leq n_{k'}.$$

There are two possibilities, all n_i's are not all the same, or all n_i's are all the same.

Assume all n_i's are not all the same, i.e., $n_1 < n_{k'}$. The fixed point set N of the isometry $\phi = \exp((t'/n_{k'})X) \in H_o$ contains $B_{k'}$ but not B_1. It is either a reversible closed geodesic, or a Finsler sphere satisfying $1 < \dim N < \dim M$, $K \equiv 1$ and Assumption (F). Applying the inductive assumption and Lemma 3.2, we can find a totally geodesic B_i for $(N, F|_N)$, as well as for (M, F).

Assume all n_i's are all the same, then all $\rho_{X,i}$'s are all the same as well. We may choose a suitable t' such that $n_i = 1$ for $1 \leq i \leq k'$. There exists $t'' \in (0, 1)$ such that $\psi(c_i(0)) = c_i(t'')$, i.e., $d_F(c_i(0), c_i(t'')) = \pi$, for $1 \leq i \leq k'$. Then we have

$$F(X|_{c_1}) = F(X|_{c_2}) = \cdots = F(X|_{c_{k'}}).$$

The function $f(\cdot) = F(X(\cdot))$ takes the same value on its critical point set, so it is a constant function. By Lemma 2.3, all integration curves of X are closed geodesics, which belongs to one $\hat{G}$-orbit. By Corollary 3.7, (M, F) is a non-Riemannian homogeneous Randers Finsler sphere with $K \equiv 1$ and exactly two $\hat{G}$-orbits of prime closed geodesics, $\mathcal{B}_1 = \hat{G} \cdot \gamma_1$ and $\mathcal{B}_2 = \hat{G} \cdot \gamma_2$ such that γ_1 and γ_2 are the same curve with different directions.

This ends the proof of Claim 1, i.e., Theorem 1.2 is valid when $m = \dim H = 1$.

Next we prove Theorem 1.2 assuming $m = \dim H > 1$.

Claim 2. *There exists at least $m - 1$ geometrically distinct orbits $\mathcal{B}_i$ such that each B_i is a totally geodesic submanifold with a nontrivial H_o-action.*

Let $\mathcal{B}_i$ with $1 \leq i \leq k'$ be all the geometrically distinct $\hat{G}$-orbits of prime closed geodesics such that the H_o-action on each B_i is not trivial. Let $\mathfrak{h}_i$ be the codimension one subalgebra of $\mathfrak{h}$ which restriction to B_i is zero. By Lemma 4.3, the intersection $\bigcap_{i=1}^{k'} \mathfrak{h}_i = 0$, from which we see that $m \leq k'$. We may reorder the orbits $\mathcal{B}_i$'s such that $\bigcap_{i=1}^{m} \mathfrak{h}_i = 0$. Take a nonzero Killing vector field $X \in \bigcap_{i=1}^{m-1} \mathfrak{h}_i$. Then the zero set N of X is a closed connected totally geodesic submanifold in M, containing B_i

for $1 \le i \le m-1$ but not B_m. Let H' be the subgroup of $I_o(N, F|_N)$ preserving all closed geodesics in N, and $\mathfrak{h}'$ its Lie algebra. The restriction from M to N defines a linear map from $\mathfrak{h}$ to $\mathfrak{h}'$ which kernel is spanned by X, so $\dim H' \ge m-1$.

If $\dim N = 1$, then $m = 2$, H_o has no fixed point, and N itself provides the totally geodesic B_i wanted by Claim 2.

If $\dim N > 1$, we can use the inductive assumption to find $m-1$ geometrically distinct orbits $\mathcal{B}_i$ of prime closed geodesics for $(N, F|_N)$, as well as for (M, F) by Lemma 3.2, such that the corresponding B_i's are totally geodesic submanifolds, with nontrivial H'_o-actions. Claim 2 is proved when each of these B_i's also has a nontrivial H_o-action.

But it is possible that there is some B_i in N on which the H'_o-action is nontrivial but the H_o-action is trivial. If it happens, this B_i is unique, and we must have $\dim H' > m-1$. So in this case, we can use the inductive assumption to find m geometrically distinct orbits of prime closed geodesics. At least $m-1$ geometrically distinct totally geodesic B_i's in N have nontrivial H_o-actions.

This proves Claim 2.

To finish the proof of Theorem 1.2 when $n = l$, we only need to find one more totally geodesic B_i with a nontrivial H_o-action.

We may reorder the orbits $\mathcal{B}_i$'s such that the first $m-1$ ones are those provided by Claim 2, and $\bigcap_{i=1}^{m} \mathfrak{h}_i = 0$. The nonzero Killing vector field X from $\bigcap_{i=1}^{m-1} \mathfrak{h}_i$ vanishes on B_i with $1 \le i \le m-1$, but not on B_m. We can find an isometry of the form $\phi = \psi \exp(t'X)$ such that it fixes each point of B_m. On the other hand, the fixed point set N of ϕ does not contain each $\mathcal{B}_i$ for $1 \le i \le m-1$.

The H_o-action on each closed geodesic in N is nontrivial. Assume conversely that there is a closed geodesic in N with a trivial H_o-action. Then the restriction of ψ to this geodesic coincides with that of ϕ, fixing each point of this geodesic. This is not true because ψ has no fixed points.

If $\dim N = 1$ it is a reversible closed geodesic, which is the extra B_i we want. If $\dim N > 1$ it is a Finsler sphere satisfying $K \equiv 1$ and Assumption (F), isometrically imbedded in (M, F) as a totally geodesic submanifold. In this situation we use the inductive assumption one more time, which provides one more totally geodesic B_i.

Summarizing above discussion, we have proved Theorem 1.2 when $n = l$.

This ends the proof of Theorem 1.2 by induction.

6. The example from Katok metrics

We conclude this paper by the examples from Katok metrics for which the estimate in Theorem 1.2 is sharp.

Let $(M, h) = (S^n, h)$ be a standard unit sphere with $n > 1$, W a Killing vector field on (M, h) such that $h(W, W) < 1$ everywhere.

Then the navigation process defines a Randers metric

$$F(y) = \frac{\sqrt{\lambda h(y, y) + h(W, y)^2}}{\lambda} - \frac{h(W, y)}{\lambda}$$

on M, in which $\lambda = 1 - h(W, W)$ is positive everywhere.

By the work of Bao, Robles and Shen [Bao et al. 2004], this construction provides all the Randers spheres with $K \equiv 1$. The behavior of the geodesics on (M, F) is determined by the choice of W.

We can find suitable coordinates $x = (\boldsymbol{x}_0, z_1, \ldots, z_k)$ for $x \in \mathbb{R}^{n+1}$, where

$$\boldsymbol{x}_0 = (x_{0,1}, \ldots, x_{0,n_0}) \in \mathbb{R}^{n_0} \quad \text{and} \quad z_i = (z_{i,1}, \ldots, z_{i,n_i}) \in \mathbb{C}^{n_i}$$

satisfy the following:

(A1) We permit $n_0 = 0$ and in this case $\boldsymbol{x}_0$ is always 0. All other n_i's are positive.

(A2) (M, h) is naturally identified as the unit sphere $S^n(1)$ defined by

$$|\boldsymbol{x}_0|^2 + |z_1|^2 + \cdots + |z|^2 = 1$$

in $\mathbb{R}^{n+1} = \mathbb{R}^{n_0} \oplus \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_k}$ with the standard product Euclidean metric.

(A3) W can be presented as

$$W(\boldsymbol{x}_0, z_0, \ldots, z_k) = (0, \sqrt{-1}\lambda_1 z_1, \ldots, \sqrt{-1}\lambda_k z_k), \tag{6-1}$$

such that $0 < \lambda_1 < \lambda_2 < \ldots < \lambda_k < 1$.

We further require one of the following is satisfied:

(A4) All λ_i's are irrational numbers. For any $1 \le i < j \le k$, 1, λ_i and λ_j are linearly independent over $\mathbb{Q}$.

(A5) All λ_i's are irrational numbers except one, $n_0 = 0$ and $n_i = 1$ if $\lambda_i \in \mathbb{Q}$. If λ_i and λ_j are irrational numbers, 1, λ_i and λ_j are linearly independent.

Then we have

Lemma 6.1. *For the Randers sphere (M, F) described above, satisfying (A1)–(A3) and one of (A4) and (A5), any closed geodesic on (M, F) must be contained in*

$$z_1 = \cdots = z_k = 0$$

or

$$\boldsymbol{x}_0 = 0 \quad \textit{and} \quad z_j = 0 \quad \textit{when } j \neq i,$$

for some i, $1 \le i \le k$.

Proof. Using (6-1), we can present the antipodal map as

$$\psi(\boldsymbol{x}_0, z_1, \ldots, z_k) = (\boldsymbol{x}_0, -e^{\sqrt{-1}\pi\lambda_1} z_1, \ldots, -e^{\sqrt{-1}\pi\lambda_k} z_k).$$

It is easy to check that finite ψ-orbits only appear in the situation that only $\boldsymbol{x}_0$ is nonzero or only z_i with $\lambda_i \in \mathbb{Q}$ is nonzero.

Let $x = (\boldsymbol{x}_0, z_1, \ldots, z_k)$ be a point on the closed geodesic γ. We only need to prove that only one of $\boldsymbol{x}_0$ and z_i's can be nonzero. Assume conversely this is not true. Then the length of γ can not be a rational multiple of π (i.e., consists of finite ψ-orbits), so the ψ-orbit of x is a dense subset in γ. There are three cases we need to consider.

In the first case, λ_i and λ_j are irrational numbers, $z_i \neq 0$, and $z_j \neq 0$. Then the condition that 1, λ_i and λ_j are linearly independent implies that the projection to the z_i- and z_j-factors maps the closed curve γ onto a two dimensional torus, which is a contradiction.

In the second case, λ_i is rational, λ_j is not, $z_i \neq 0$ and $z_j \neq 0$. Then the projection to the z_i-factor maps γ to a finite set with at least two points. This is impossible because γ is connected.

In the third case, $\boldsymbol{x}_0 \neq 0$ and $z_i \neq 0$. Then the projection to the $\boldsymbol{x}_0$-factor maps γ to two points. This is impossible for the same reason as the previous case.

To summarize, we have found contradiction for all the cases, and finished the proof of this lemma. □

Using Lemma 6.1, we can provides examples of Katok metrics such that the estimates in Theorem 1.2 are sharp.

Theorem 6.2. *Let F be the Randers metrics on S^n with $n > 1$ satisfying (A1)–(A3) and one of (A4) and (A5). Then it has only finite orbits of prime closed geodesics. Let H denote the subgroup of isometries preserving each closed geodesic, H_o its identity component, and $m = \dim H$. Then there exist exactly m geometrically distinct $\mathcal{B}_i$, such that the corresponding B_i's are totally geodesic with nontrivial H_o-actions.*

The proof is a case-by-case discussion. For each case, it is not hard to calculate $G = I_o(M, F)$, H_o and all the orbits of prime closed geodesics.

For example, when $n_0 > 2$ and all γ_i's are irrational numbers,

$$G = \mathrm{SO}(n_0) \times U(n_1) \times \cdots \times U(n_k), \quad \text{and}$$
$$H = C(U(n_1) \times \cdots \times U(n_k)) = U(1)^k,$$

so we have $\dim H = k$.

When $1 \leq i \leq k$,

$$B_i = \left\{x = (\boldsymbol{x}_0, z_1, \ldots, z_k) \in M \text{ with } \boldsymbol{x}_0 = 0 \text{ and } z_j = 0 \text{ when } j \neq i\right\}$$

is a homogeneous Randers sphere with exactly two orbits of prime closed geodesics. It is isometrically imbedded in (M, F) as a totally geodesic submanifold, because it is the fixed point set of the subgroup of G with the $U(n_i)$-factor removed. They provide all the different totally geodesic B_i's with nontrivial H_o-actions.

There exists one more totally geodesic B_{k+1} with a trivial H_o-action, i.e.,

$$B_{k+1} = \{x = (\boldsymbol{x}_0, z_1, \ldots, z_k) \in M \text{ with } z_1 = \cdots = z_k = 0\}.$$

It is a standard unit sphere with only one orbit of closed geodesics.

By Lemma 6.1, no other closed geodesics can be found.

Summarizing all these observations, we see that this Randers sphere (M, F) satisfies all the requirements in Theorem 1.2, and the estimate in Theorem 1.2 for the number of totally geodesic B_i's is sharp.

The discussion for other cases is similar, so we skip the details.

Acknowledgements

The author would like to thank sincerely the Chern Institute of Mathematics, Nankai University, and Shaoqiang Deng for the hospitality during the preparation for this paper. The author also thanks Yuri G. Nikonorov and Huagui Duan for helpful discussions.

References

[Anosov 1975] D. V. Anosov, "Geodesics in Finsler geometry", pp. 293–297 in *Proceedings of the International Congress of Mathematicians, II* (Vancouver, 1974), edited by R. D. James, Canad. Math. Congress, Montreal, 1975. In Russian; translated in *Amer. Math. Soc. Trans.* **109** (1977), 81–85. MR Zbl

[Bangert and Long 2010] V. Bangert and Y. Long, "The existence of two closed geodesics on every Finsler 2-sphere", *Math. Ann.* **346**:2 (2010), 335–366. MR Zbl

[Bao et al. 2000] D. Bao, S.-S. Chern, and Z. Shen, *An introduction to Riemann–Finsler geometry*, Graduate Texts in Math. **200**, Springer, 2000. MR Zbl

[Bao et al. 2004] D. Bao, C. Robles, and Z. Shen, "Zermelo navigation on Riemannian manifolds", *J. Differential Geom.* **66**:3 (2004), 377–435. MR Zbl

[Berestovskii and Nikonorov 2019] V. N. Berestovskii and Y. G. Nikonorov, "On homogeneous geodesics and weakly symmetric spaces", *Ann. Global Anal. Geom.* **55**:3 (2019), 575–589. MR

[Bryant 1996] R. L. Bryant, "Finsler structures on the 2-sphere satisfying $K = 1$", pp. 27–41 in *Finsler geometry* (Seattle, 1995), edited by D. Bao et al., Contemp. Math. **196**, Amer. Math. Soc., Providence, RI, 1996. MR Zbl

[Bryant 1997] R. L. Bryant, "Projectively flat Finsler 2-spheres of constant curvature", *Selecta Math. (N.S.)* **3**:2 (1997), 161–203. MR Zbl

[Bryant 2002] R. L. Bryant, "Some remarks on Finsler manifolds with constant flag curvature", *Houston J. Math.* **28**:2 (2002), 221–262. MR Zbl

[Bryant et al. 2017] R. L. Bryant, P. Foulon, S. Ivanov, V. S. Matveev, and W. Ziller, "Geodesic behavior for Finsler metrics of constant positive flag curvature on S^2", 2017. To appear in *J. Differential Geom.* arXiv

[Deng and Xu 2014] S. Deng and M. Xu, "Clifford–Wolf translations of Finsler spaces", *Forum Math.* **26**:5 (2014), 1413–1428. MR Zbl

[Duan 2016] H. Duan, "Two elliptic closed geodesics on positively curved Finsler spheres", *J. Differential Equations* **260**:12 (2016), 8388–8402. MR Zbl

[Duan et al. 2016] H. Duan, Y. Long, and W. Wang, "The enhanced common index jump theorem for symplectic paths and non-hyperbolic closed geodesics on Finsler manifolds", *Calc. Var. Partial Differential Equations* **55**:6 (2016), art. id. 145. MR Zbl

[Katok 1973] A. B. Katok, "Ergodic perturbations of degenerate integrable Hamiltonian systems", *Izv. Akad. Nauk SSSR Ser. Mat.* **37** (1973), 539–576. In Russian; translated in *Math. USSR-Izv.* **7** (1973) 535–572. MR Zbl

[Long and Duan 2009] Y. Long and H. Duan, "Multiple closed geodesics on 3-spheres", *Adv. Math.* **221**:6 (2009), 1757–1803. MR Zbl

[Rademacher 1989] H.-B. Rademacher, "On the average indices of closed geodesics", *J. Differential Geom.* **29**:1 (1989), 65–83. MR Zbl

[Rademacher 2004] H.-B. Rademacher, "A sphere theorem for non-reversible Finsler metrics", *Math. Ann.* **328**:3 (2004), 373–387. MR Zbl

[Shen 1996] Z. Shen, "Finsler manifolds of constant positive curvature", pp. 83–93 in *Finsler geometry* (Seattle, 1995), edited by D. Bao et al., Contemp. Math. **196**, Amer. Math. Soc., Providence, RI, 1996. MR Zbl

[Wang 2012] W. Wang, "On a conjecture of Anosov", *Adv. Math.* **230**:4-6 (2012), 1597–1617. MR Zbl

[Wang 2015] W. Wang, "Non-hyperbolic closed geodesics on Finsler spheres", *J. Differential Geom.* **99**:3 (2015), 473–496. MR Zbl

[Xu 2018a] M. Xu, "Geodesic orbit spheres and constant curvature in Finsler geometry", *Differential Geom. Appl.* **61** (2018), 197–206. MR Zbl

[Xu 2018b] M. Xu, "The number of geometrically distinct reversible closed geodesics on a Finsler sphere with $K \equiv 1$", preprint, 2018. arXiv

[Yan and Huang 2018] Z. Yan and L. Huang, "On the existence of homogeneous geodesic in homogeneous Finsler spaces", *J. Geom. Phys.* **124** (2018), 264–267. MR Zbl

Received April 29, 2018. Revised October 27, 2018.

MING XU
SCHOOL OF MATHEMATICAL SCIENCES
CAPITAL NORMAL UNIVERSITY
BEIJING
CHINA
mgmgmgxu@163.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 302, No. 1, 2019
dx.doi.org/10.2140/pjm.2019.302.371

DEGENERACY THEOREMS FOR TWO HOLOMORPHIC CURVES IN $\mathbb{P}^n(\mathbb{C})$ SHARING FEW HYPERSURFACES

KAI ZHOU AND LU JIN

In value distribution theory, many uniqueness and degeneracy theorems for holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ sharing hyperplanes or sharing sufficiently many hypersurfaces have been obtained in the last few decades. But there is no result concerning holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ sharing few hypersurfaces. We prove several degeneracy theorems for two algebraically nondegenerate holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ sharing $n+k$ hypersurfaces in general position.

1. Introduction

Since Fujimoto [1975] generalized Nevanlinna's uniqueness theorems of meromorphic functions sharing values to the case of meromorphic maps of $\mathbb{C}^m$ into $\mathbb{P}^n(\mathbb{C})$ sharing hyperplanes, plenty of uniqueness and degeneracy results for meromorphic maps sharing hyperplanes have been obtained; see for instance [Smiley 1983; Fujimoto 1998; Fujimoto 1999; Chen and Yan 2009; Si and Le 2015]. Some uniqueness theorems for holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ sharing sufficiently many hypersurfaces have also been proven; see [Dulock and Ru 2008; Phuong 2013; Quang and An 2017].

But as far as we know, there is no result concerning two holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ sharing $n+k$ hypersurfaces. This paper proves some degeneracy theorems for two holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ sharing $n+k$ hypersurfaces.

Now we introduce some notions. A holomorphic map $f:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})$ is said to be linearly (resp. algebraically) nondegenerate if its image is not contained in any proper linear subspace (resp. algebraic subset) of $\mathbb{P}^n(\mathbb{C})$. Hypersurfaces $D_1,\dots,D_q\,(q>n)$ in $\mathbb{P}^n(\mathbb{C})$ are said to be located in general position if $\bigcap_{k=1}^{n+1}\operatorname{Supp} D_{j_k}=\varnothing$ for any $n+1$ distinct indices $j_1,\dots,j_{n+1}\in\{1,\dots,q\}$. For a nonzero meromorphic function h on the complex plane $\mathbb{C}$, let ν_h^0 (resp. ν_h^∞) be the zero (resp. pole) divisor of h, and let $\nu_h=\nu_h^0-\nu_h^\infty$.

We may regard $\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ as a subvariety of $\mathbb{P}^{(n+1)^2-1}(\mathbb{C})$ via the Segre embedding $(a_0:\dots:a_n)\times(b_0:\dots:b_n)\mapsto(\dots:a_ib_j:\dots)$. And a holomorphic

This work was supported by National Natural Science Foundation of China (Grant No. 11331004).
MSC2010: primary 32H30; secondary 30D35.
Keywords: degeneracy theorem, hypersurface, holomorphic curve, Nevanlinna theory.

map $F:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ is said to be algebraically degenerate if its image is contained in a proper algebraic subset of $\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$.

We state our main theorems now. Let $f,g:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})$ be two algebraically nondegenerate holomorphic curves with reduced representations $\tilde{f}=(f_0,\dots,f_n)$ and $\tilde{g}=(g_0,\dots,g_n)$. Let $q>n$ and let D_j, $1\le j\le q$, be hypersurfaces of degrees d_j in $\mathbb{P}^n(\mathbb{C})$ located in general position. Let $Q_j\in\mathbb{C}[x_0,\dots,x_n]$, $1\le j\le q$, be the homogeneous polynomials of degrees d_j defining D_j. Let d be the least common multiple of the d_j's and set $\tilde{Q}_j=Q_j^{d/d_j}$ for $1\le j\le q$.

Theorem 1.1. *Assume that* $q=\max\{4,n+2\}$. *If*

(a) $f^{-1}(D_i)\cap f^{-1}(D_j)=\varnothing$ *for all* $i\in\{1,\dots,q\}$ *and* $j\in\{1,2,3,4\}\setminus\{i\}$,

(b) $\nu_{Q_j(\tilde{f})}=\nu_{Q_j(\tilde{g})}$ *for* $1\le j\le 4$ *and* $\min\{\nu_{Q_j(\tilde{f})},1\}=\min\{\nu_{Q_j(\tilde{g})},1\}$ *for* $4<j\le q$,

(c) $f=g$ *on* $\bigcup_{j=1}^{q}f^{-1}(D_j)$,

then there are three distinct indices $i,j,k\in\{1,2,3,4\}$ *such that*

$$\left(\frac{\tilde{Q}_i(\tilde{f})\cdot\tilde{Q}_k(\tilde{g})}{\tilde{Q}_i(\tilde{g})\cdot\tilde{Q}_k(\tilde{f})}\right)^s\cdot\left(\frac{\tilde{Q}_j(\tilde{f})\cdot\tilde{Q}_k(\tilde{g})}{\tilde{Q}_j(\tilde{g})\cdot\tilde{Q}_k(\tilde{f})}\right)^t\equiv 1$$

for some $(s,t)\in\mathbb{Z}\times\mathbb{Z}\setminus\{(0,0)\}$. *Consequently* $\{f_ug_v\}_{0\le u,v\le n}$ *satisfy a nontrivial homogeneous polynomial equation*; *thus* $f\times g:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ *is algebraically degenerate.*

Since the conditions "$f=g$ on $\bigcup_{j=1}^{q}f^{-1}(D_j)$" and "$f^{-1}(D_i)\cap f^{-1}(D_j)=\varnothing$ for $i\ne j$" are really rigid, it's natural to study the related problem without the two conditions. In this direction, Fujimoto [1999] proved a degeneracy theorem for sharing $2n+2$ hyperplanes with truncated multiplicities. In our case of sharing few hypersurfaces, we can only prove the following.

Theorem 1.2. *Assume that* $q=n+3$. *If*

(a) $\nu_{Q_j(\tilde{f})}=\nu_{Q_j(\tilde{g})}$ *for* $1\le j\le q$,

(b) $f=g$ *on* $\bigcup_{j=1}^{n+2}f^{-1}(D_j)$,

then there are three distinct indices $i,j,k\in\{1,\dots,q\}$ *such that*

$$\left(\frac{\tilde{Q}_i(\tilde{f})\cdot\tilde{Q}_k(\tilde{g})}{\tilde{Q}_i(\tilde{g})\cdot\tilde{Q}_k(\tilde{f})}\right)^s\cdot\left(\frac{\tilde{Q}_j(\tilde{f})\cdot\tilde{Q}_k(\tilde{g})}{\tilde{Q}_j(\tilde{g})\cdot\tilde{Q}_k(\tilde{f})}\right)^t\equiv 1$$

for some $(s,t)\in\mathbb{Z}\times\mathbb{Z}\setminus\{(0,0)\}$. *In particular* $f\times g:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ *is algebraically degenerate.*

In fact we prove a stronger theorem (see Theorem 4.1) with a weaker condition than "$f=g$ on $\bigcup_{j=1}^{n+2}f^{-1}(D_j)$" (see Remark 4.2).

If we require further that the order of f (see Definition 2.5) is less than 1, then we can get rid of both the two conditions; namely we have:

Theorem 1.3. *Assume that f is of order < 1. Let $q = n+2$. If $\nu_{Q_j(\tilde{f})} = \nu_{Q_j(\tilde{g})}$ for $j = 1, 2$ and $\min\{\nu_{Q_j(\tilde{f})}, 1\} = \min\{\nu_{Q_j(\tilde{g})}, 1\}$ for $2 < j \le q$, then there exists a nonzero constant C such that*

$$\frac{\tilde{Q}_1(\tilde{f}) \cdot \tilde{Q}_2(\tilde{g})}{\tilde{Q}_1(\tilde{g}) \cdot \tilde{Q}_2(\tilde{f})} \equiv C.$$

In particular, $f \times g : \mathbb{C} \to \mathbb{P}^n(\mathbb{C}) \times \mathbb{P}^n(\mathbb{C})$ is algebraically degenerate.

Remark 1.4. If all D_j's are hyperplanes in $\mathbb{P}^n(\mathbb{C})$, then the nondegeneracy assumption on f and g only needs to be linearly nondegenerate.

Our proof is based on the second main theorem for holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ intersecting hypersurfaces, which was first proved by Ru [2004], and a gcd bound for holomorphic units (see [Pasten and Wang 2017, Theorem 3.1]). The technique of using the gcd bound is due to Si [2013].

2. Preliminaries from Nevanlinna theory

For a divisor ν on $\mathbb{C}$, we define the counting function of ν by

$$N(r, \nu) = \int_0^r \frac{n(t, \nu) - n(0, \nu)}{t} dt + n(0, \nu) \log r,$$

where $n(t, \nu) := \sum_{|z| \le t} \nu(z)$.

Let $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be a holomorphic map and let $\tilde{f} = (f_0, \dots, f_n)$ be a reduced representation of f; namely, $f_0, \dots, f_n$ are entire functions on $\mathbb{C}$ without common zeros and $f(z) = [f_0(z) : \cdots : f_n(z)]$ for every $z \in \mathbb{C}$. The characteristic function of f is defined by

$$T_f(r) := \frac{1}{2\pi} \int_0^{2\pi} \log \|\tilde{f}(re^{i\theta})\| d\theta - \log \|\tilde{f}(0)\|,$$

where $\|\tilde{f}(z)\| = \sqrt{|f_0(z)|^2 + \cdots + |f_n(z)|^2}$. This definition is independent of the choice of the reduced representation. Let D be a hypersurface of degree d in $\mathbb{P}^n(\mathbb{C})$ with $f(\mathbb{C}) \not\subseteq D$. Let $Q \in \mathbb{C}[x_0, \dots, x_n]$ be the homogeneous polynomial of degree d defining D. Then the proximity function $m_f(r, D)$ is defined by

$$m_f(r, D) = \frac{1}{2\pi} \int_0^{2\pi} \log \frac{\|\tilde{f}(re^{i\theta})\|^d \|Q\|}{|Q(\tilde{f})(re^{i\theta})|} d\theta,$$

where $\|Q\|$ is the maximum of the absolute values of the coefficients of Q. And the counting function of f intersecting D with truncation level M, $M \in \mathbb{Z}^+ \cup \{+\infty\}$, is defined by

$$N_f^{[M]}(r, D) := N(r, \min\{\nu_{Q(\tilde{f})}, M\}).$$

We also write $N_f^{[1]}(r, D) = \overline{N}_f(r, D)$ and $N_f^{[+\infty]}(r, D) = N_f(r, D)$. If H is a hyperplane in $\mathbb{P}^n(\mathbb{C})$ defined by the linear form L, we also write $L(\tilde{f})$ as (f, H).

The Jensen formula (see [Ru 2001, Corollary A1.1.3]) implies the following first main theorem:

Theorem 2.1. *Let $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be a holomorphic map and let D be a hypersurface of degree d in $\mathbb{P}^n(\mathbb{C})$. If $f(\mathbb{C}) \not\subseteq D$, then there is a real constant C, such that for all $r > 0$,*

$$m_f(r, D) + N_f(r, D) = dT_f(r) + C.$$

The following is the well known second main theorem for holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ intersecting hyperplanes (see [Ru 2001, Theorem A3.2.2]) which was first proved by H. Cartan.

Theorem 2.2. *Let $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be a linearly nondegenerate holomorphic map and $\{H_j\}_{j=1}^q$ be q hyperplanes in $\mathbb{P}^n(\mathbb{C})$ located in general position. Then*

$$\Big\| \; (q - n - 1)T_f(r) \le \sum_{j=1}^{q} N_f^{[n]}(r, H_j) + o(T_f(r)),$$

where the notation "$\|$" means that the assertion holds for all $r > 0$ outside a set of finite Lebesgue measure.

Ru [2004] proved a second main theorem for holomorphic curves in $\mathbb{P}^n(\mathbb{C})$ intersecting hypersurfaces. The following version with truncation was proved in [Yan and Chen 2008; An and Phuong 2009].

Theorem 2.3. *Let $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be an algebraically nondegenerate holomorphic map. Let D_j, $1 \le j \le q$, be hypersurfaces of degrees d_j in $\mathbb{P}^n(\mathbb{C})$ located in general position. Then for any $\epsilon > 0$, there is a positive integer M_ϵ such that*

$$\Big\| \; (q - n - 1 - \epsilon)T_f(r) \le \sum_{j=1}^{q} d_j^{-1} N_f^{[M_\epsilon]}(r, D_j).$$

For a meromorphic function h on the complex plane $\mathbb{C}$, the Nevanlinna's characteristic function of h is defined by

$$T(r, h) := m(r, h) + N(r, h),$$

where $m(r, h) := \frac{1}{2\pi} \int_0^{2\pi} \log^+ |h(re^{i\theta})| d\theta$ with $\log^+ x = \max\{\log x, 0\}$ for $x \ge 0$, and $N(r, h) := N(r, \nu_h^\infty)$. It follows from the definition that for any meromorphic functions h_1, h_2 on $\mathbb{C}$, $T(r, h_1 + h_2) \le T(r, h_1) + T(r, h_2) + \ln 2$ and $T(r, h_1 h_2) \le T(r, h_1) + T(r, h_2)$ for $r \ge 1$. Furthermore we have the following first main theorem for meromorphic functions (see [Ru 2001, Theorem A1.1.5]).

Theorem 2.4. $T(r, h) = T\big(r, \frac{1}{h-a}\big) + O(1)$ *for any meromorphic function h on $\mathbb{C}$ and $a \in \mathbb{C}$ provided that $h \not\equiv a$.*

Definition 2.5. The order of a holomorphic map $f:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})$ is defined to be

$$\overline{\lim_{r\to+\infty}}\frac{\log^+ T_f(r)}{\log r}.$$

The order of a meromorphic function h on $\mathbb{C}$ can be similarly defined.

3. Proof of Theorem 1.1

We prove Theorem 1.1 in this section; in fact, we prove the following stronger theorem.

Theorem 3.1. *Let $f, g, \tilde f, \tilde g, d, D_j, Q_j, \tilde Q_j,\ 1\le j\le q$, be given as in Section 1. Assume there exist $I, J\subseteq\{1,\dots,q\}$ with $\#I\ge n+2$ and for any $i\in I$, $\#(J\setminus\{i\})\ge 3$, such that the following conditions are satisfied*:

(a) $f^{-1}(D_i)\cap f^{-1}(D_j)=\varnothing$ *for all* $i\in I$ *and* $j\in J\setminus\{i\}$,

(b) $\nu_{Q_j(\tilde f)}=\nu_{Q_j(\tilde g)}$ *for* $j\in J$ *and* $\min\{\nu_{Q_i(\tilde f)},1\}=\min\{\nu_{Q_i(\tilde g)},1\}$ *for* $i\in I$,

(c) $f=g$ *on* $\bigcup_{i\in I} f^{-1}(D_i)$.

Then there exist three distinct indices $i, j, k\in J$ such that

$$\left(\frac{\tilde Q_i(\tilde f)\cdot\tilde Q_k(\tilde g)}{\tilde Q_i(\tilde g)\cdot\tilde Q_k(\tilde f)}\right)^s\cdot\left(\frac{\tilde Q_j(\tilde f)\cdot\tilde Q_k(\tilde g)}{\tilde Q_j(\tilde g)\cdot\tilde Q_k(\tilde f)}\right)^t\equiv 1$$

for some $(s,t)\in\mathbb{Z}\times\mathbb{Z}\setminus\{(0,0)\}$. In particular, $f\times g:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ is algebraically degenerate.

Taking $q=\max\{4,n+2\}$, $I=\{1,\dots,q\}$, $J=\{1,2,3,4\}$, we get Theorem 1.1. Furthermore, we can deduce the following corollary by taking $q=n+5$, $I=\{4,\dots,q\}$, $J=\{1,2,3\}$.

Corollary 3.2. *Let $f, g, \tilde f, \tilde g, d, D_j, Q_j, \tilde Q_j,\ 1\le j\le q$, be given as in Section 1. Assume that $q=n+5$. If*

(a) $f^{-1}(D_i)\cap f^{-1}(D_j)=\varnothing$ *for* $i=1,2,3$ *and* $j=4,\dots,q$,

(b) $\nu_{Q_j(\tilde f)}=\nu_{Q_j(\tilde g)}$ *for* $j=1,2,3$, *and* $\min\{\nu_{Q_j(\tilde f)},1\}=\min\{\nu_{Q_j(\tilde g)},1\}$ *for* $j=4,\dots,q$,

(c) $f=g$ *on* $\bigcup_{j=4}^{q} f^{-1}(D_j)$,

then

$$\left(\frac{\tilde Q_1(\tilde f)\cdot\tilde Q_3(\tilde g)}{\tilde Q_1(\tilde g)\cdot\tilde Q_3(\tilde f)}\right)^s\cdot\left(\frac{\tilde Q_2(\tilde f)\cdot\tilde Q_3(\tilde g)}{\tilde Q_2(\tilde g)\cdot\tilde Q_3(\tilde f)}\right)^t\equiv 1$$

for some $(s,t)\in\mathbb{Z}\times\mathbb{Z}\setminus\{(0,0)\}$. In particular, $f\times g:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ is algebraically degenerate.

For our purpose, we need the following lemma on the gcd bound for holomorphic units; for the proof refer to [Pasten and Wang 2017, Theorem 3.1].

Lemma 3.3. *Let F, G be nowhere zero holomorphic functions on $\mathbb{C}$. If $F^s \cdot G^t$ is not constant for all $(s,t) \in \mathbb{Z} \times \mathbb{Z} \setminus \{(0,0)\}$, then for any $\epsilon > 0$,*

$$\| \; N(r, F-1, G-1) \le \epsilon \max\{T(r,F), T(r,G)\},$$

where $N(r, F-1, G-1)$ is the counting function of the common 1-points of F and G; namely, $N(r, F-1, G-1) := N(r, \min\{\nu^0_{F-1}, \nu^0_{G-1}\})$.

Remark 3.4. If $F^s \cdot G^t \equiv c \in \mathbb{C} \setminus \{1\}$ for some $(s,t) \in \mathbb{Z} \times \mathbb{Z} \setminus \{(0,0)\}$, then F and G have no common 1-points; namely, $N(r, F-1, G-1) \equiv 0$. So the conclusion of the above lemma actually holds when $F^s \cdot G^t \not\equiv 1$ for all $(s,t) \in \mathbb{Z} \times \mathbb{Z} \setminus \{(0,0)\}$.

Now we are going to prove Theorem 3.1. We give the following lemma first.

Lemma 3.5. *Let $f : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be a holomorphic map with reduced representation $\tilde{f} = (f_0, \dots, f_n)$. Let $Q_1, Q_2 \in \mathbb{C}[x_0, \dots, x_n]$ be two homogeneous polynomials of same degree $d > 0$ with $Q_2(\tilde{f}) \not\equiv 0$. Then there are constants $C_1, C_2 > 0$ such that for all $r > 0$ large enough,*

$$T\left(r, \frac{Q_1(\tilde{f})}{Q_2(\tilde{f})}\right) \le C_1 T_f(r) + C_2.$$

Proof. Take $k \in \{0, \dots, n\}$ such that $f_k \not\equiv 0$. Write $Q_1(\tilde{f}) = \sum a f_0^{i_0} \cdots f_n^{i_n}$ and $Q_2(\tilde{f}) = \sum b f_0^{j_0} \cdots f_n^{j_n}$, then

$$\frac{Q_1(\tilde{f})}{Q_2(\tilde{f})} = \frac{Q_1(\tilde{f})/f_k^d}{Q_2(\tilde{f})/f_k^d} = \frac{\sum a\left(\frac{f_0}{f_k}\right)^{i_0} \cdots \left(\frac{f_n}{f_k}\right)^{i_n}}{\sum b\left(\frac{f_0}{f_k}\right)^{j_0} \cdots \left(\frac{f_n}{f_k}\right)^{j_n}}.$$

Thus by the first main theorem and the properties of Nevanlinna's characteristic function, we conclude that

$$\begin{aligned} T\left(r, \frac{Q_1(\tilde{f})}{Q_2(\tilde{f})}\right) \le & T\left(r, \sum a\left(\frac{f_0}{f_k}\right)^{i_0} \cdots \left(\frac{f_n}{f_k}\right)^{i_n}\right) \\ & + T\left(r, \sum b\left(\frac{f_0}{f_k}\right)^{j_0} \cdots \left(\frac{f_n}{f_k}\right)^{j_n}\right) + O(1) \\ \le & \tilde{C}_1\left(T\left(r, \frac{f_0}{f_k}\right) + \cdots + T\left(r, \frac{f_n}{f_k}\right)\right) + \tilde{C}_2. \end{aligned}$$

By [Ru 2001, Theorem A3.1.2], we know that $T(r, f_t/f_k) \le T_f(r) + O(1)$ for $t = 0, \dots, n$, this together with the above inequality imply the desired conclusion. □

Proof of Theorem 3.1. Set $h_j = \tilde{Q}_j(\tilde{f})/\tilde{Q}_j(\tilde{g})$ for $j = 1, \dots, q$. Then by condition (b), h_j is a nowhere zero holomorphic function on $\mathbb{C}$ for every $j \in J$.

We argue by the method of contradiction. Assume that the conclusion doesn't hold, then for arbitrary three distinct indices $j_1, j_2, j_3 \in J$,

$$\left(\frac{h_{j_1}}{h_{j_3}}\right)^s \cdot \left(\frac{h_{j_2}}{h_{j_3}}\right)^t \not\equiv 1$$

for all $(s, t) \in \mathbb{Z} \times \mathbb{Z} \setminus \{(0, 0)\}$. So applying Lemma 3.3 (see Remark 3.4) to the functions h_{j_1}/h_{j_3} and h_{j_2}/h_{j_3}, we conclude that for any $\epsilon > 0$,

$$\Big\| \; N(r, h_{j_1}/h_{j_3} - 1, h_{j_2}/h_{j_3} - 1) \leq \epsilon \max\left\{T\left(r, \frac{h_{j_1}}{h_{j_3}}\right), T\left(r, \frac{h_{j_2}}{h_{j_3}}\right)\right\}. \tag{3.6}$$

Note that the $\tilde{Q}_j$'s are all of degree d, so by Lemma 3.5, we see that for any $i, j \in \{1, \ldots, q\}$,

$$T\left(r, \frac{h_i}{h_j}\right) \leq T\left(r, \frac{\tilde{Q}_i(\tilde{f})}{\tilde{Q}_j(\tilde{f})}\right) + T\left(r, \frac{\tilde{Q}_j(\tilde{g})}{\tilde{Q}_i(\tilde{g})}\right) \leq C_1(T_f(r) + T_g(r)) + C_2.$$

Combining this with inequality (3.6), we get that for arbitrary three distinct indices $j_1, j_2, j_3 \in J$, for any $\epsilon > 0$,

$$\| \; N(r, h_{j_1}/h_{j_3} - 1, h_{j_2}/h_{j_3} - 1) \leq \epsilon(T_f(r) + T_g(r)). \tag{3.7}$$

Take $i \in I$. By $\#(J \setminus \{i\}) \geq 3$, we can choose three distinct $j_1, j_2, j_3 \in J \setminus \{i\}$. By conditions (a), (b) and (c), if $z \in f^{-1}(D_i)$, then z is not the zero of $\tilde{Q}_{j_k}(\tilde{f})$ and $\tilde{Q}_{j_k}(\tilde{g})$, $k = 1, 2, 3$, and $\tilde{f}(z) = c\tilde{g}(z)$ for some nonzero constant c. So for $k = 1, 2, 3$,

$$h_{j_k}(z) = \frac{\tilde{Q}_{j_k}(\tilde{f})(z)}{\tilde{Q}_{j_k}(\tilde{g})(z)} = \frac{\tilde{Q}_{j_k}(\tilde{f}(z))}{\tilde{Q}_{j_k}(\tilde{g}(z))} = c^d,$$

thus

$$\frac{h_{j_1}}{h_{j_3}}(z) = \frac{h_{j_2}}{h_{j_3}}(z) = \frac{c^d}{c^d} = 1;$$

namely, z is a common 1-point of h_{j_1}/h_{j_3} and h_{j_2}/h_{j_3}. So combining this with inequality (3.7), we have for any $\epsilon > 0$,

$$\| \; \overline{N}_f(r, D_i) \leq N(r, \; h_{j_1}/h_{j_3} - 1, \; h_{j_2}/h_{j_3} - 1) \leq \epsilon(T_f(r) + T_g(r)).$$

Summing up the above inequality over $i \in I$ and noting that $\overline{N}_f(r, D_i) = \overline{N}_g(r, D_i)$, we get that for any $\epsilon > 0$,

$$\Big\| \; \sum_{i \in I}(\overline{N}_f(r, D_i) + \overline{N}_g(r, D_i)) \leq \epsilon(T_f(r) + T_g(r)). \tag{3.8}$$

On the other hand, by the second main theorem for holomorphic curves intersecting hypersurfaces (see Theorem 2.3) and the assumption $\#I \geq n+2$, and noting that $N_f^{[M]}(r, D) \le M\overline{N}_f(r, D)$, we deduce that there is a positive constant κ such that

$$\Big\| \sum_{i\in I}(\overline{N}_f(r, D_i)+\overline{N}_g(r, D_i)) \ge \kappa(T_f(r)+T_g(r)).$$

This contradicts (3.8).

Therefore we have proved that there exist three distinct indices $i, j, k \in J$ such that

$$\left(\frac{\tilde{Q}_i(\tilde{f})\cdot\tilde{Q}_k(\tilde{g})}{\tilde{Q}_i(\tilde{g})\cdot\tilde{Q}_k(\tilde{f})}\right)^s \cdot \left(\frac{\tilde{Q}_j(\tilde{f})\cdot\tilde{Q}_k(\tilde{g})}{\tilde{Q}_j(\tilde{g})\cdot\tilde{Q}_k(\tilde{f})}\right)^t \equiv 1$$

for some $(s, t) \in \mathbb{Z}\times\mathbb{Z}\setminus\{(0, 0)\}$. Now since all $\tilde{Q}_t$'s are of the same degree d, it is easy to see that the $(n+1)^2$ functions $\{f_u g_v\}_{0\le u,v\le n}$ satisfy a nontrivial homogeneous polynomial equation. This shows that the image of $f\times g:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ is contained in a proper algebraic subset of $\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$; in other words, $f\times g:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ is algebraically degenerate.

Furthermore from the above proof, we easily see that if all D_j's are hyperplanes, then the proof still works if f and g are only assumed to be linearly nondegenerate. This completes the proof. □

4. Proof of Theorem 1.2

We prove the following theorem which implies Theorem 1.2.

Theorem 4.1. *Let $f, g, \tilde{f}, \tilde{g}, d, D_j, Q_j, \tilde{Q}_j$, $1\le j\le q$, be given as in Section 1. Let $q=n+3$ and set $h_j=\tilde{Q}_j(\tilde{f})/\tilde{Q}_j(\tilde{g})$ for $j=1,\dots,q$. Assume that*

(a) $\nu_{Q_j(\tilde{f})}=\nu_{Q_j(\tilde{g})}$ *for $j=1,\dots,q$, and*

(b) *for every $i\in\{1,\dots,n+2\}$, the set*

$$A_i := \left\{\frac{h_j}{h_k}(z)\;\Big|\; z\in f^{-1}(D_i),\ 1\le j,k\le q \text{ with } z\notin f^{-1}(D_j\cup D_k)\cup g^{-1}(D_j\cup D_k)\right\}$$

is of finite cardinality.

Then there exist distinct indices $i, j, k\in\{1,\dots,q\}$ and constants

$$C_1, C_2 \in A := \{1\}\cup\bigcup_{i=1}^{n+2} A_i$$

such that

$$\left(\frac{\tilde{Q}_i(\tilde{f})\cdot\tilde{Q}_k(\tilde{g})}{C_1\tilde{Q}_i(\tilde{g})\cdot\tilde{Q}_k(\tilde{f})}\right)^s \cdot \left(\frac{\tilde{Q}_j(\tilde{f})\cdot\tilde{Q}_k(\tilde{g})}{C_2\tilde{Q}_j(\tilde{g})\cdot\tilde{Q}_k(\tilde{f})}\right)^t \equiv 1$$

for some $(s, t)\in\mathbb{Z}\times\mathbb{Z}\setminus\{(0,0)\}$. In particular $f\times g:\mathbb{C}\to\mathbb{P}^n(\mathbb{C})\times\mathbb{P}^n(\mathbb{C})$ is algebraically degenerate.

Remark 4.2. By the condition "$f = g$ on $\bigcup_{j=1}^{n+2} f^{-1}(D_j)$" of Theorem 1.2, one deduces as in the proof of Theorem 3.1 that for $i \in \{1, \dots, n+2\}$ and $j, k \in \{1, \dots, q\}$, $(h_j/h_k)(z) = 1$ for every point

$$z \in f^{-1}(D_i) \setminus (f^{-1}(D_j \cup D_k) \cup g^{-1}(D_j \cup D_k)).$$

So $A = \{1\}$. Thus the conclusion of Theorem 1.2 follows from Theorem 4.1.

Proof. By assumption, h_j is a nowhere zero holomorphic function on $\mathbb{C}$ for every $1 \le j \le q$ and A is a nonempty set consisting of finitely many nonzero complex numbers. So we may set $A = \{c_1, \dots, c_p\}$.

Assume that the conclusion doesn't hold, then for any distinct indices $i, j, k \in \{1, \dots, q\}$ and constants $c_u, c_v \in A$,

$$\left(\frac{h_i}{c_u h_k}\right)^s \cdot \left(\frac{h_j}{c_v h_k}\right)^t \not\equiv 1$$

for all $(s, t) \in \mathbb{Z} \times \mathbb{Z} \setminus \{(0, 0)\}$. Much as in the proof of Theorem 3.1, by making use of Lemmas 3.3 and 3.5, we conclude that for any $\epsilon > 0$,

$$\| \; N\left(r, \frac{h_i}{c_u h_k} - 1, \frac{h_j}{c_v h_k} - 1\right) \le \epsilon (T_f(r) + T_g(r)). \tag{4.3}$$

Let $\nu = \sum_{1 \le i < j < k \le q} \sum_{c_u, c_v \in A} \min\{\nu^0_{h_i/(c_u h_k)-1}, \nu^0_{h_j/(c_v h_k)-1}\}$, then

$$N(r, \nu) = \sum_{1 \le i < j < k \le q} \sum_{c_u, c_v \in A} N\left(r, \frac{h_i}{c_u h_k} - 1, \frac{h_j}{c_v h_k} - 1\right).$$

So (4.3) gives that for any $\epsilon > 0$,

$$\| \; N(r, \nu) \le \epsilon (T_f(r) + T_g(r)). \tag{4.4}$$

Now take $l \in \{1, \dots, n+2\}$. For a point $z \in f^{-1}(D_l)$, by the "in general position" assumption, we know that there are at most $n-1$ distinct $k \in \{1, \dots, q\} \setminus \{l\}$ such that $z \in f^{-1}(D_k)$. Since $q = n+3$, there are three distinct $i, j, k \in \{1, \dots, q\} \setminus \{l\}$ with $i < j < k$ such that $z \notin f^{-1}(D_i \cup D_j \cup D_k) \cup g^{-1}(D_i \cup D_j \cup D_k)$. Then

$$\frac{h_i}{h_k}(z), \frac{h_j}{h_k}(z) \in A_l \subseteq A.$$

Thus there are $c_u, c_v \in A$ such that z is a common 1-point of $h_i/(c_u h_k)$ and $h_j/(c_v h_k)$, so the point z is counted in $N(r, \nu)$. Consequently, for any $\epsilon > 0$,

$$\| \; \overline{N}_f(r, D_l) \le N(r, \nu) \le \epsilon (T_f(r) + T_g(r)).$$

From this we see that for any $\epsilon > 0$,

$$\Big\| \sum_{l=1}^{n+2}(\bar{N}_f(r, D_l) + \bar{N}_g(r, D_l)) \le \epsilon(T_f(r) + T_g(r)).$$

On the other hand, using the second main theorem (see Theorem 2.3), as in the proof of Theorem 3.1, we deduce that there exists a constant $\kappa > 0$ such that

$$\Big\| \kappa(T_f(r) + T_g(r)) \le \sum_{l=1}^{n+2}(\bar{N}_f(r, D_l) + \bar{N}_g(r, D_l)),$$

which contradicts the above inequality. This proves Theorem 4.1. □

Combining the proof of Theorem 4.1 with that of Theorem 3.1, one concludes easily the following theorem which is an improvement of Theorem 3.1.

Theorem 4.5. *Let $f, g, \tilde{f}, \tilde{g}, d, D_j, Q_j, \tilde{Q}_j$, $1 \le j \le q$, be given as in Section 1. Assume that there exist $I, J \subseteq \{1, \dots, q\}$ with $\#I \ge n+2$ and for any $i \in I$, $\#(J \setminus \{i\}) \ge 3$, such that the following conditions are satisfied*:

(a) $\nu_{Q_j(\tilde{f})} = \nu_{Q_j(\tilde{g})}$ *for $j \in J$ and* $\min\{\nu_{Q_i(\tilde{f})}, 1\} = \min\{\nu_{Q_i(\tilde{g})}, 1\}$ *for $i \in I$*;

(b) *for every $i \in I$, the set*

$$\left\{ \frac{h_j}{h_k}(z) \,\Big|\, z \in f^{-1}(D_i),\ j, k \in J \setminus \{i\} \right\} =: A_i$$

is of finite cardinality.

Then there exist three distinct indices $i, j, k \in J$ and constants

$$C_1, C_2 \in A := \{1\} \cup \bigcup_{u \in I} A_u$$

such that

$$\left(\frac{\tilde{Q}_i(\tilde{f}) \cdot \tilde{Q}_k(\tilde{g})}{C_1 \tilde{Q}_i(\tilde{g}) \cdot \tilde{Q}_k(\tilde{f})} \right)^s \cdot \left(\frac{\tilde{Q}_j(\tilde{f}) \cdot \tilde{Q}_k(\tilde{g})}{C_2 \tilde{Q}_j(\tilde{g}) \cdot \tilde{Q}_k(\tilde{f})} \right)^t \equiv 1$$

for some $(s, t) \in \mathbb{Z} \times \mathbb{Z} \setminus \{(0, 0)\}$.

5. Proof of Theorem 1.3

We prove Theorem 1.3 and then as a consequence we give a uniqueness theorem.

Proof of Theorem 1.3. Let $f, g, \tilde{f}, \tilde{g}, d, D_j, Q_j, \tilde{Q}_j$, $(1 \le j \le q)$ be given as in Section 1. We set $h_j = \tilde{Q}_j(\tilde{f})/\tilde{Q}_j(\tilde{g})$ for $j = 1, \dots, q$. Then the assumption shows that h_1 and h_2 are nowhere zero holomorphic functions on $\mathbb{C}$. We need to show that h_1/h_2 is constant.

Since $q = n+2$ and $\overline{N}_f(r, D_j) = \overline{N}_g(r, D_j)$ for $j = 1, \dots, q$, it follows from the first and the second main theorem that there are constants $C_1, C_2 > 0$ such that

$$\Big\| \; C_1 T_g(r) \leq \sum_{j=1}^{q} \overline{N}_g(r, D_j) = \sum_{j=1}^{q} \overline{N}_f(r, D_j) \leq q d T_f(r) + C_2;$$

therefore there is a constant $C > 0$ such that

$$\| \; T_g(r) \leq C T_f(r). \tag{5.1}$$

By Lemma 3.5, there is a constant $C_3 > 0$ such that for all large r,

$$T(r, h_1/h_2) \leq C_3(T_f(r) + T_g(r)).$$

Combining this with (5.1), we have

$$\| \; T(r, h_1/h_2) \leq C_4 T_f(r)$$

for some constant $C_4 > 0$. From this and the assumption that f is of order < 1, it follows that

$$\varliminf_{r\to+\infty} \frac{\log^+ T(r, h_1/h_2)}{\log r} \leq \varlimsup_{r\to+\infty} \frac{\log^+ T_f(r)}{\log r} < 1. \tag{5.2}$$

Since h_1/h_2 is nowhere zero holomorphic on $\mathbb{C}$, we may write $h_1/h_2 = e^H$ for some entire function H. If h_1/h_2 is nonconstant, then either H is a polynomial of degree ≥ 1 or H is a transcendental entire function; thus by [Yang and Yi 2003, Theorem 1.44] we have

$$\varliminf_{r\to+\infty} \frac{\log^+ T(r, h_1/h_2)}{\log r} \geq 1.$$

This contradicts (5.2). Thus h_1/h_2 is constant, which completes the proof. □

Remark 5.3. From the above proof, we easily see that if all D_j's are hyperplanes, then the conclusion still holds when f and g are only assumed to be linearly nondegenerate. So we have the following uniqueness theorem:

Corollary 5.4. *Let $f, g : \mathbb{C} \to \mathbb{P}^n(\mathbb{C})$ be two linearly nondegenerate holomorphic maps. Let $H_1, \dots, H_{n+2}$ be hyperplanes in $\mathbb{P}^n(\mathbb{C})$ located in general position. Suppose that f is of order < 1. If $\nu_{(f,H_j)} = \nu_{(g,H_j)}$ for every $j = 1, \dots, n+2$, then $f = g$.*

Proof. Take reduced representations $\tilde{f}, \tilde{g}$ for f and g respectively and let L_j, $1 \leq j \leq n+2$, be the linear forms that define H_j. Set $h_j = L_j(\tilde{f})/L_j(\tilde{g})$, $1 \leq j \leq n+2$. Then Theorem 1.3 shows that $h_i/h_1 = c_i$ is a constant for any $i \geq 2$, and $L_i(\tilde{f}) = h_1 c_i L_i(\tilde{g})$. Since the H_j's are in general position, we can write $L_1 = \sum_{i=2}^{n+2} b_i L_i$ for some nonzero constants b_i. Thus

$$h_1 \sum_{i=2}^{n+2} b_i L_i(\tilde{g}) = h_1 L_1(\tilde{g}) = L_1(\tilde{f}) = \sum_{i=2}^{n+2} b_i L_i(\tilde{f}) = h_1 \sum_{i=2}^{n+2} b_i c_i L_i(\tilde{g}),$$

which implies that

$$\left(\sum_{i=2}^{n+2} b_i(1-c_i)L_i\right)(\tilde{g}) = 0.$$

Now by the linearly nondegeneracy of g and the fact that $L_2, \dots, L_{n+2}$ are linearly independent, we conclude that

$$c_{n+2} = c_{n+1} = \cdots = c_2 = 1;$$

namely, $h_1 = h_2 = \cdots = h_{n+2}$. This implies that $f = g$. □

References

[An and Phuong 2009] T. T. H. An and H. T. Phuong, "An explicit estimate on multiplicity truncation in the second main theorem for holomorphic curves encountering hypersurfaces in general position in projective space", *Houston J. Math.* **35**:3 (2009), 775–786. MR Zbl

[Chen and Yan 2009] Z. Chen and Q. Yan, "Uniqueness theorem of meromorphic mappings into $\mathbb{P}^N(\mathbb{C})$ sharing $2N+3$ hyperplanes regardless of multiplicities", *Int. J. Math.* **20**:6 (2009), 717–726. MR Zbl

[Dulock and Ru 2008] M. Dulock and M. Ru, "A uniqueness theorem for holomorphic curves sharing hypersurfaces", *Complex Var. Elliptic Equ.* **53**:8 (2008), 797–802. MR Zbl

[Fujimoto 1975] H. Fujimoto, "The uniqueness problem of meromorphic maps into the complex projective space", *Nagoya Math. J.* **58** (1975), 1–23. MR Zbl

[Fujimoto 1998] H. Fujimoto, "Uniqueness problem with truncated multiplicities in value distribution theory", *Nagoya Math. J.* **152** (1998), 131–152. MR Zbl

[Fujimoto 1999] H. Fujimoto, "Uniqueness problem with truncated multiplicities in value distribution theory, II", *Nagoya Math. J.* **155** (1999), 161–188. MR Zbl

[Pasten and Wang 2017] H. Pasten and J. T.-Y. Wang, "GCD bounds for analytic functions", *Int. Math. Res. Not.* **2017**:1 (2017), 47–95. MR Zbl

[Phuong 2013] H. T. Phuong, "Uniqueness theorems for holomorphic curves sharing moving hypersurfaces", *Complex Var. Elliptic Equ.* **58**:11 (2013), 1481–1491. MR Zbl

[Quang and An 2017] S. D. Quang and D. P. An, "Second main theorem and unicity of meromorphic mappings for hypersurfaces in projective varieties", *Acta Math. Vietnam.* **42**:3 (2017), 455–470. MR Zbl

[Ru 2001] M. Ru, *Nevanlinna theory and its relation to Diophantine approximation*, World Sci., River Edge, NJ, 2001. MR Zbl

[Ru 2004] M. Ru, "A defect relation for holomorphic curves intersecting hypersurfaces", *Amer. J. Math.* **126**:1 (2004), 215–226. MR Zbl

[Si 2013] D. Q. Si, "Some extensions of the four values theorem of Nevanlinna–Gundersen", *Kodai Math. J.* **36**:3 (2013), 579–595. MR Zbl

[Si and Le 2015] D. Q. Si and N. Q. Le, "Algebraic dependences of meromorphic mappings sharing few hyperplanes counting truncated multiplicities", *Kodai Math. J.* **38**:1 (2015), 97–118. MR Zbl

[Smiley 1983] L. M. Smiley, "Geometric conditions for unicity of holomorphic curves", pp. 149–154 in *Value distribution theory and its applications* (New York, 1983), edited by C. C. Yang, Contemp. Math. **25**, Amer. Math. Soc., Providence, RI, 1983. MR Zbl

[Yan and Chen 2008] Q. M. Yan and Z. H. Chen, "Weak Cartan-type second main theorem for holomorphic curves", *Acta Math. Sin.* (*Engl. Ser.*) **24**:3 (2008), 455–462. MR Zbl

[Yang and Yi 2003] C.-C. Yang and H.-X. Yi, *Uniqueness theory of meromorphic functions*, Math. and its Appl. **557**, Kluwer, Dordrecht, 2003. MR Zbl

Received January 20, 2019.

KAI ZHOU
SCHOOL OF MATHEMATICAL SCIENCES
FUDAN UNIVERSITY
SHANGHAI
CHINA
17110180034@fudan.edu.cn

LU JIN
SCHOOL OF MATHEMATICAL SCIENCES
FUDAN UNIVERSITY
SHANGHAI
CHINA
jinlu@fudan.edu.cn

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 302 No. 1 September 2019