# *Pacific Journal of Mathematics*

Volume 280 No. 2 February 2016

# PACIFIC JOURNAL OF MATHEMATICS

msp.org/pjm

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





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

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

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.257

# TOPOLOGICAL MOLINO'S THEORY

JESÚS A. ÁLVAREZ LÓPEZ AND MANUEL F. MOREIRA GALICIA

**Molino's description of Riemannian foliations on compact manifolds is generalized to the setting of compact equicontinuous foliated spaces, in the case where the leaves are dense. In particular, a structural local group is associated to such a foliated space. As an application, we obtain a partial generalization of results by Carrière and Breuillard–Gelander, relating the structural local group to the growth of the leaves.**



## 1. Introduction

Riemannian foliations were introduced by Reinhart [1959] by requiring isometric transverse dynamics. It was pointed out by Ghys in [Molino 1988, Appendix E] (see also [Kellum 1993]) that equicontinuous foliated spaces should be considered as the "topological Riemannian foliations," and therefore many of the results about Riemannian foliations should have versions for equicontinuous foliated spaces. Some steps in this direction were given by Álvarez and Candel [2009; 2010], showing that, under reasonable conditions, their leaf closures are minimal foliated spaces, and their generic leaves are quasi-isometric to each other, like in the case of Riemannian foliations. In the same direction, Matsumoto [2010] proved that any minimal equicontinuous foliated space has a nontrivial transverse invariant measure, which is unique up to scaling if the space is compact — observe that this unicity

Álvarez is partially supported by the MICINN grant MTM2011-25656. Moreira was partially supported by the grant ERASMUS MUNDUS LOTE 20B, for the period 2010–2012.
*MSC2010:* 57R30.
*Keywords:* equicontinuous foliated space, equicontinuous pseudogroup, groupoid, germ, partial map, compact-open topology, local group, local action, growth.

implies ergodicity. The magnitude of the generalization from Riemannian foliations to equicontinuous foliated spaces was made precise by Álvarez and Candel [2010] (see also [Tarquini 2004]), giving a topological description of Riemannian foliations within the class of equicontinuous foliated spaces.

Most of the known properties of Riemannian foliations follow from a description due to Molino [1982; 1988]. However, so far, there was no version of Molino's description for equicontinuous foliated spaces — the indicated properties of equicontinuous foliated spaces were obtained by other means. The goal of our work is to develop such a version of Molino's theory, and use it to study the growth of their leaves, following the study of the growth of Riemannian foliations by Carrière [1988] and Breuillard and Gelander [2007]. To understand our results better, let us briefly recall Molino's theory.

**1A.** ***Molino's theory for Riemannian foliations.*** The necessary basic concepts from foliation theory can be seen in [Hector and Hirsch 1981; 1987; Candel and Conlon 2000].

Let $\mathcal{F}$ be a (smooth) foliation of codimension $q$ on a manifold $M$. Let $T\mathcal{F} \subset TM$ denote the vector subbundle of vectors tangent to the leaves, and $N\mathcal{F} = TM/T\mathcal{F}$ its normal bundle. Recall that there is a natural flat leafwise partial connection on $N\mathcal{F}$ such that any local normal vector field is leafwise parallel if and only if it is locally projectable by the distinguished submersions; terms like "leafwise flat," "leafwise parallel" and "leafwise horizontal" will refer to this partial connection. It is said that $\mathcal{F}$ is

- *Riemannian* if $N\mathcal{F}$ has a leafwise parallel Riemannian structure;
- *transitive* if the group of its foliated diffeomorphisms acts transitively on $M$;
- *transversely parallelizable* (*TP*) if there is a leafwise parallel global frame of $N\mathcal{F}$, called *transverse parallelism*; and a
- *Lie foliation* if moreover the transverse parallelism is a basis of a Lie algebra with the operation induced by the vector field bracket.

These conditions are successively stronger. Molino's theory describes Riemannian foliations on compact manifolds in terms of minimal Lie foliations, and using TP foliations as an intermediate step:

**1st step:** If $\mathcal{F}$ is Riemannian and $M$ compact, then there is an $\mathrm{O}(q)$-principal bundle $\hat{\pi} : \widehat{M} \to M$, with an $\mathrm{O}(q)$-invariant TP foliation $\widehat{\mathcal{F}}$, such that $\hat{\pi}$ is a foliated map whose restrictions to the leaves are the holonomy covers of the leaves of $\mathcal{F}$.

**2nd step:** If $\mathcal{F}$ is TP and $M$ compact, then there is a fiber bundle $\pi : M \to W$ whose fibers are the leaf closures of $\mathcal{F}$, and the restriction of $\mathcal{F}$ to each fiber is a Lie foliation.

Since the structure of Lie foliation is unique in the minimal case, we end up with a Lie algebra associated to $\mathcal{F}$, called the *structural Lie algebra*. The proofs of the above statements strongly use the differential structure of $\mathcal{F}$. In the first step, $\hat{\pi} : \widehat{M} \to M$ is the $\mathrm{O}(q)$-principal bundle of orthonormal frames for some leafwise parallel metric on $N\mathcal{F}$, and $\widehat{\mathcal{F}}$ is given by the corresponding flat leafwise horizontal distribution. Then $\widehat{\mathcal{F}}$ is TP by a standard argument. In the second step, foliated flows are used to produce the fiber bundle trivializations whose fibers are the leaf closures; this works because there are foliated flows in any transverse direction since $\mathcal{F}$ is TP.

When $\mathcal{F}$ is minimal (the leaves are dense), any leaf closure $\widehat{M}_0$ of $\widehat{\mathcal{F}}$ is a principal subbundle of $\hat{\pi} : \widehat{M} \to M$, obtaining the following:

**Minimal case:** If $\mathcal{F}$ is minimal and Riemannian and $M$ is compact, then, for some closed subgroup $H \subset \mathrm{O}(q)$, there is an $H$-principal bundle $\hat{\pi}_0 : \widehat{M}_0 \to M$ with an $H$-invariant minimal Lie foliation $\widehat{\mathcal{F}}_0$, such that $\hat{\pi}_0$ is a foliated map whose restrictions to the leaves are the holonomy covers of the leaves of $\mathcal{F}$.

A useful description of Lie foliations was also given by Fedida [1971; 1978], but it will not be considered here.

The differential structure cannot be used in our generalization; instead, we use the holonomy pseudogroup. Thus let us briefly indicate the holonomy properties of Riemannian foliations that will play an important role.

**1B.** ***Holonomy of Riemannian foliations.*** Recall that a *pseudogroup* is a maximal collection of local transformations of a space, which contains the identity map, and is closed under the operations of composition, inversion, restriction and combination. It can be considered as a generalized dynamical system, and all basic dynamical concepts have pseudogroup versions. They are relevant in foliation theory because the holonomy pseudogroup of a foliation $\mathcal{F}$ describes the transverse dynamics of $\mathcal{F}$. Such a pseudogroup is well determined up to certain *equivalence* of pseudogroups introduced by Haefliger [1985; 1988]. We may say that $\mathcal{F}$ is *transversely modeled* by a class of local transformations of some space if its holonomy pseudogroup can be generated by that type of local transformations. Riemannian, TP and Lie foliations can be respectively characterized by being transversely modeled by local isometries of some Riemannian manifold, by local parallelism preserving diffeomorphisms of some parallelizable manifold, and by local left translations of a Lie group. In this sense, Riemannian foliations are the transversely rigid ones, and TP foliations have a stronger type of transverse rigidity.

When the ambient manifold $M$ is compact, Haefliger [2002] has observed that the holonomy pseudogroup $\mathcal{H}$ of $\mathcal{F}$ satisfies a property called compact generation. If moreover $\mathcal{F}$ is Riemannian, then Haefliger [1988; 2002] has also strongly used the following properties of $\mathcal{H}$: completeness, quasianalyticity, and existence of a closure $\overline{\mathcal{H}}$, which is also complete and quasianalytic. Here, $\overline{\mathcal{H}}$ is defined by taking the closure of the set of 1-jets of maps in $\mathcal{H}$ in the space of 1-jets.

For a compactly generated pseudogroup $\mathcal{H}$ of local isometries of a Riemannian manifold $T$, Salem has given a version of Molino's theory [Salem 1988; Molino 1988, Appendix D] (see also [Álvarez and Masa 2008]). In particular, in the minimal case, it turns out that there is a Lie group $G$, a compact subgroup $K \subset G$ and a dense finitely generated subgroup $\Gamma \subset G$ such that $\mathcal{H}$ is equivalent to the pseudogroup generated by the action of $\Gamma$ on the homogeneous space $G/K$ (this was also observed by Haefliger [1988]).

**1C.** ***Growth of Riemannian foliations.*** Molino's theory has many consequences for a Riemannian foliation $\mathcal{F}$ on a compact manifold $M$: classification in particular cases, growth, cohomology, tautness, tenseness and global analysis. In all of them, Molino's theory is used to reduce the study to the case of Lie foliations with dense leaves, where it usually becomes a problem of Lie theory. We concentrate on the consequences about growth of the leaves and their holonomy covers. This study was begun by Carrière [1988], and recently continued by Breuillard and Gelander [2007], as a consequence of their study of a topological Tits alternative. Their results state the following, where $\mathfrak{g}$ is the structural Lie algebra of $\mathcal{F}$:

**Carrière's theorem.** *The holonomy covers of the leaves are Følner if and only if* $\mathfrak{g}$ *is solvable*, *and of polynomial growth if and only if* $\mathfrak{g}$ *is nilpotent. In the second case*, *the degree of polynomial growth is bounded by the nilpotence degree of* $\mathfrak{g}$.

**Breuillard and Gelander's theorem.** *The growth of the holonomy covers of the leaves is either polynomial or exponential.*

**1D.** ***Equicontinuous foliated spaces.*** A *foliated space* $X \equiv (X, \mathcal{F})$ is a topological space $X$ equipped with a partition $\mathcal{F}$ into connected manifolds (*leaves*), which can be locally described as the fibers of topological submersions. It will be assumed that $X$ is locally compact and Polish. A foliated space should be considered as a "topological foliation". In this sense, all topological notions of foliations have obvious versions for foliated spaces. In particular, the *holonomy pseudogroup* $\mathcal{H}$ of $X$ is defined on a locally compact Polish space $T$. Many basic results about foliations also have straightforward generalizations; for example, $\mathcal{H}$ is compactly generated if $X$ is compact. Even leafwise differential concepts are easy to extend. However this task may be difficult or impossible for transverse differential concepts. For instance, the normal bundle of a foliated space does not make any sense in general; it would be the tangent bundle of a topological space in the case of a space foliated by points. Thus the concept of Riemannian foliation cannot be extended by using the normal bundle. Instead, this can be done via the holonomy pseudogroup as follows.

The transverse rigidity of a Riemannian foliation can be translated to the foliated space $X$ by requiring equicontinuity of $\mathcal{H}$. In fact, the equicontinuity condition is not compatible with combinations of maps, and therefore it is indeed required for some generating subset $S \subset \mathcal{H}$ which is closed by the operations of composition and inversion. Such an $S$ is called a pseudo$*$group with the terminology of Matsumoto [2010]. This gives rise to the concept of equicontinuous foliated space.

In the topological setting, the quasianalyticity of $\mathcal{H}$ does not follow from the equicontinuity assumption. Thus it will be required as an additional assumption when needed. Indeed, it does not work well enough when $T$ is not locally connected, so we use a property called strong quasianalyticity, which is stronger than quasianalyticity only when $T$ is not locally connected.

Álvarez and Candel [2009] have proved that, if $\mathcal{H}$ is compactly generated, equicontinuous and strongly quasianalytic, then it is complete and has a closure $\overline{\mathcal{H}}$. Here, $\overline{\mathcal{H}}$ is the pseudogroup generated by the homeomorphisms on small enough open subsets $O$ of $T$ that are limits in the compact-open topology of maps in $\mathcal{H}$ defined on those sets $O$.

Transitive and Lie foliations have the following topological versions. It is said that the foliated space $X$ is

- *homogeneous* if its group of foliated transformations acts transitively on $X$;
- a *$G$-foliated space* if it is transversely modeled by local left translations in some locally compact Polish local group $G$ (if $X$ is minimal).

**1E. *Topological Molino's theory.*** Our first main result is the following topological version of the minimal case in Molino's theory.

**Theorem A.** *Let $X \equiv (X, \mathcal{F})$ be a compact Polish foliated space, and $\mathcal{H}$ its holonomy pseudogroup. Suppose that $X$ is minimal and equicontinuous, and $\overline{\mathcal{H}}$ is strongly quasianalytic. Then there is a compact Polish minimal foliated space $\widehat{X}_0 \equiv (\widehat{X}_0, \widehat{\mathcal{F}}_0)$, an open continuous foliated map $\hat{\pi}_0 : \widehat{X}_0 \to X$, and a locally compact Polish local group $G$ such that $\widehat{X}_0$ is a $G$-foliated space, the fibers of $\hat{\pi}_0$ are homeomorphic to each other, and the restrictions of $\hat{\pi}_0$ to the leaves of $\widehat{\mathcal{F}}_0$ are the holonomy covers of the leaves of $\mathcal{F}$.*

The proof of Theorem A is different from Molino's proof in the Riemannian foliation case because there may not be the normal bundle of $\mathcal{F}$. To define $\widehat{X}_0$, we first construct what should be its holonomy pseudogroup, $\widehat{\mathcal{H}}_0$ on a space $\widehat{T}_0$. To some extent, this was achieved by Álvarez and Candel [2010], proving that, with the assumptions of Theorem A, there is a locally compact Polish local group $G$, a compact subgroup $K \subset G$ and a dense finitely generated sub-local group $\Gamma \subset G$ such that $\mathcal{H}$ is equivalent to the pseudogroup generated by the local action of $\Gamma$ on $G/K$, like in the Riemannian foliation case. Hence $\widehat{\mathcal{H}}_0$ should be the pseudogroup generated by the local action of $\Gamma$ on $G$. This may look like a big step towards the proof, but the realization of compactly generated pseudogroups as holonomy pseudogroups of compact foliated spaces is impossible in general, as shown by Meigniez [Meigniez 2010]. This difficulty is overcome as follows.

Take a "good" cover of $X$ by distinguished open sets $\{U_i\}$, with corresponding distinguished submersions $p_i : U_i \to T_i$, and elementary holonomy transformations $h_{ij} : T_{ij} \to T_{ji}$, where $T_{ij} = p_i(U_i \cap U_j)$. Let $\mathcal{H}$ denote the corresponding representative of the holonomy pseudogroup on $T = \bigsqcup_i T_i$, generated by the maps $h_{ij}$. Then the construction of $\widehat{\mathcal{H}}_0$ must be associated to $\mathcal{H}$ in a natural way, so that it becomes induced by some "good" cover by distinguished open sets of a compact foliated space. In the Riemannian foliation case, the good choices of $\widehat{T}_0$ and $\widehat{\mathcal{H}}_0$ are the following:

- Let $P$ be the bundle of orthonormal frames for any $\mathcal{H}$-invariant metric on $T$. Fix $x_0 \in T$ and $\hat{x}_0 \in P_{x_0}$. Then, as a subspace of $P$,

$$\widehat{T}_0 = \overline{\{h_*(\hat{x}_0) \mid h \in \mathcal{H}, x_0 \in \operatorname{dom} h\}} = \{g_*(\hat{x}_0) \mid g \in \overline{\mathcal{H}}, x_0 \in \operatorname{dom} g\}. \tag{1}$$

- $\widehat{\mathcal{H}}_0$ is generated by the differentials of the maps in $\mathcal{H}$.

These differential concepts can be modified in the following way. In (1), each $g_*(\hat{x}_0)$ determines the germ $\boldsymbol{\gamma}(g, x_0)$ of $g$ at $x_0$, by the strong quasianalyticity of $\overline{\mathcal{H}}$. Therefore it also determines $\boldsymbol{\gamma}(f, x)$, where $f = g^{-1}$ and $x = g(x_0)$ — this little change, using $\boldsymbol{\gamma}(f, x)$ instead of $\boldsymbol{\gamma}(g, x_0)$, is not really necessary, but it helps to simplify the notation in some involved arguments. So

$$\widehat{T}_0 \equiv \{\boldsymbol{\gamma}(f, x) \mid f \in \overline{\mathcal{H}}, x \in \operatorname{dom} f, f(x) = x_0\}. \tag{2}$$

The projection $\hat{\pi}_0 : \widehat{T}_0 \to T$ corresponds via (2) to the source map $\boldsymbol{\gamma}(f, x) \mapsto x$. The differentials of maps $h \in \mathcal{H}$, acting on orthonormal references, correspond via (2) to the maps $\hat{h}$ defined by

$$\hat{h}(\boldsymbol{\gamma}(f, x)) = \boldsymbol{\gamma}(f h^{-1}, h(x)).$$

Let us describe the topology of $\widehat{T}_0$ using (2). Let $\overline{S}$ be a pseudo$*$group generating $\overline{\mathcal{H}}$ and satisfying the equicontinuity and strong quasianalyticity conditions. Endow $\overline{S}$ with the compact-open topology on partial maps with open domains, as defined by Abd-Allah and Brown [1980], and consider the subspace

$$\overline{S} * T = \{(f, x) \in \overline{S} \mid x \in \operatorname{dom} f\} \subset \overline{S} \times T.$$

Then the topology of $\widehat{T}_0$ corresponds via (2) to the quotient topology by the germ map $\boldsymbol{\gamma} : \overline{S} * T \to \boldsymbol{\gamma}(\overline{S} * T) \equiv \widehat{T}_0$, which is of course different from the sheaf topology on germs. This point of view, replacing orthonormal frames by germs, can be readily translated to the foliated space setting, obtaining good choices of $\widehat{T}_0$ and $\widehat{\mathcal{H}}_0$ under the conditions of Theorem A.

Now, consider triples $(x, \gamma, i)$ with $x \in U_i$, $\gamma \in \widehat{T}_{i,0} := \hat{\pi}_0^{-1}(T_i)$ and $p_i(x) = \hat{\pi}_0(\gamma)$. Declare $(x, \gamma, i)$ equivalent to $(y, \delta, j)$ if $x = y$ and $\widehat{h_{ij}}(\gamma) = \delta$. Then $\widehat{X}_0$ is defined as the corresponding quotient space. Let $[x, \gamma, i]$ denote the equivalence class of each triple $(x, \gamma, i)$. The foliated structure $\widehat{\mathcal{F}}_0$ on $\widehat{X}_0$ is determined by requiring that, for each fixed index $i$, the elements of the type $[x, \gamma, i]$ form a distinguished open set $\widehat{U}_{i,0}$, with distinguished submersion $\hat{p}_{i,0} : \widehat{U}_{i,0} \to \widehat{T}_{i,0}$ given by $\hat{p}_{i,0}([x, \gamma, i]) = \gamma$. The projection $\hat{\pi}_0 : \widehat{X}_0 \to X$ is defined by $\hat{\pi}_0([x, \gamma, i]) = x$. The properties stated in Theorem A are satisfied with these definitions.

It is also proved that, up to foliated homeomorphisms (respectively, local isomorphisms), $\widehat{X}_0$ (respectively, $G$) is independent of the choices involved. Hence $G$ can be called the *structural local group* of $\mathcal{F}$.

**1F.** ***Growth of equicontinuous foliated spaces.*** Our second main result is the following weak topological version of the above theorems of Carrière and Breuillard–Gelander.

**Theorem B.** *Let $X$ be a foliated space satisfying the conditions of Theorem A, and let $G$ be its structural local group. Then one of the following properties holds*:

- *$G$ can be approximated by nilpotent local Lie groups*; *or*
- *the holonomy covers of all leaves of $X$ have exponential growth.*

(The definition of *approximation* of a local group is given in Definition 2.25.) Like in the case of Riemannian foliations, Theorem A reduces the proof of Theorem B to the case of minimal $G$-foliated spaces, where it becomes a problem about local groups. Then, since any locally compact Polish local group can be approximated by local Lie groups in the above sense, the result follows by applying the same arguments as Breuillard and Gelander.

The paper concludes by indicating some examples where Theorems A and B may have interesting applications, and proposing some open problems.

## 2. Preliminaries on equicontinuous pseudogroups

**2A.** ***Compact-open topology on partial maps with open domains.*** (See [Abd-Allah and Brown 1980].) Given spaces $X$ and $Y$, let $C(X, Y)$ be the space of all continuous maps $X \to Y$; the notation $C_{\text{c-o}}(X, Y)$ may be used to indicate that $C(X, Y)$ is equipped with the compact-open topology. Let $Y^* = Y \cup \{\omega\}$, where $\omega \notin Y$, endowed with the topology in which $U \subset Y^*$ is open if and only if $U = Y^*$ or $U$ is open in $Y$. A *partial map* $X \rightarrowtail Y$ is a continuous map of a subset of $X$ to $Y$; the set of all partial maps $X \rightarrowtail Y$ is denoted by $\operatorname{Par}(X, Y)$. A partial map $X \rightarrowtail Y$ with open domain is called a *paro map*, and the set of all paro maps $X \rightarrowtail Y$ is denoted by $\mathrm{Paro}(X, Y)$. There is a bijection $\mu : \mathrm{Paro}(X, Y) \to C(X, Y^*)$ defined by

$$\mu(f)(x) = \begin{cases} f(x) & \text{if } x \in \operatorname{dom} f, \\ \omega & \text{if } x \notin \operatorname{dom} f. \end{cases}$$

The topology on $\mathrm{Paro}(X, Y)$ which makes $\mu : \mathrm{Paro}(X, Y) \to C_{\text{c-o}}(X, Y^*)$ a homeomorphism is called the *compact-open topology*, and the notation $\mathrm{Paro}_{\text{c-o}}(X, Y)$ may be used for the corresponding space. This topology has a subbasis of open sets of the form

$$\mathcal{N}(K, O) = \{h \in \mathrm{Paro}(X, Y) \mid K \subset \operatorname{dom} h, h(K) \subset O\},$$

where $K \subset X$ is compact and $O \subset Y$ is open.

**Proposition 2.1.** *If $X$ is second countable and locally compact*, *and $Y$ is second countable*, *then* $\mathrm{Paro}_{\text{c-o}}(X, Y)$ *is second countable.*

*Proof.* By hypothesis, there are countable bases of open sets, $\mathcal{V}$ of $X$ and $\mathcal{W}$ of $Y$, such that $\overline{V}$ is compact for all $V \in \mathcal{V}$. Then the sets $\mathcal{N}(\overline{V}, W)$ ($V \in \mathcal{V}$ and $W \in \mathcal{W}$) form a countable subbasis of open sets of $\mathrm{Paro}_{\text{c-o}}(X, Y)$. □

The following result is elementary.

**Proposition 2.2.** *For any open subset $U \subset X$, the restriction of the topology of* $\mathrm{Paro}_{\text{c-o}}(X, Y)$ *to the subset $C(U, Y)$ is its usual compact-open topology.*

Since paro maps are not globally defined, let us make precise the definition of their composition. Given spaces $X$, $Y$ and $Z$, the *composition* of two paro maps, $f \in \mathrm{Paro}(X, Y)$ and $g \in \mathrm{Paro}(Y, Z)$, is the paro map $gf \in \mathrm{Paro}(X, Z)$ defined as the usual composition of the maps

$$f^{-1}(\operatorname{dom} g) \xrightarrow{f} \operatorname{dom} g \xrightarrow{g} Z.$$

**Proposition 2.3** [Abd-Allah and Brown 1980, Proposition 3]. *The following properties hold*:

(i) *Let $h : T \rightarrowtail X$ and $g : Y \rightarrowtail Z$ be paro maps. Then the maps*

$$g_* : \mathrm{Paro}_{\text{c-o}}(X, Y) \to \mathrm{Paro}_{\text{c-o}}(X, Z), \quad f \mapsto gf,$$
$$h^* : \mathrm{Paro}_{\text{c-o}}(X, Y) \to \mathrm{Paro}_{\text{c-o}}(T, Y), \quad f \mapsto fh,$$

*are continuous.*

(ii) *Let $X' \subset X$ and $Y' \subset Y$ be subspaces such that $X'$ is open in $X$. Then the map*

$$\mathrm{Paro}_{\text{c-o}}(X', Y') \to \mathrm{Paro}_{\text{c-o}}(X, Y),$$

*mapping a paro map $X' \rightarrowtail Y'$ to the paro map $X \rightarrowtail Y$ with the same graph*, *is an embedding.*

**Proposition 2.4** [Abd-Allah and Brown 1980, Proposition 7]. *If $Y$ is locally compact, then the evaluation partial map*

$$\mathrm{ev} : \mathrm{Paro}_{\text{c-o}}(Y, Z) \times Y \rightarrowtail Z, \quad (f, y) \mapsto f(y),$$

*is a paro map; in particular, its domain is open.*

**Proposition 2.5** [Abd-Allah and Brown 1980, Proposition 9]. *If $X$ and $Y$ are locally compact, then the composition mapping*

$$\mathrm{Paro}_{\text{c-o}}(X, Y) \times \mathrm{Paro}_{\text{c-o}}(Y, Z) \to \mathrm{Paro}_{\text{c-o}}(X, Y), \quad (f, g) \mapsto gf,$$

*is continuous.*

Let $\mathrm{Loct}(T)$ be the family of all homeomorphisms between open subsets of a space $T$, which are called *local transformations*. For $h, h' \in \mathrm{Loct}(T)$, the composition $h'h \in \mathrm{Loct}(T)$ is the composition of maps

$$h^{-1}(\mathrm{im}\, h \cap \mathrm{dom}\, h') \xrightarrow{h} \mathrm{im}\, h \cap \mathrm{dom}\, h' \xrightarrow{h'} h'(\mathrm{im}\, h \cap \mathrm{dom}\, h').$$

Each $h \in \mathrm{Loct}(T)$ can be identified with the paro map $T \rightarrowtail T$ with the same graph. This gives rise to a canonical injection $\mathrm{Loct}(T) \to \mathrm{Paro}(T, T)$ compatible with composition. The corresponding restriction of the compact-open topology of $\mathrm{Paro}(T, T)$ to $\mathrm{Loct}(T)$ is also called *compact-open topology*, and the notation $\mathrm{Loct}_{\text{c-o}}(T)$ may be used for the corresponding space. The *bi-compact-open topology* is the smallest topology on $\mathrm{Loct}(X)$ such that the identity and inversion maps

$$\mathrm{Loct}(T) \to \mathrm{Loct}_{\text{c-o}}(T), \quad f \mapsto f^{\pm 1},$$

are continuous, and the notation $\mathrm{Loct}_{\text{b-c-o}}(T)$ will be used for the corresponding space. The following result is elementary.

**Proposition 2.6** [Abd-Allah and Brown 1980, Proposition 10]. *If $T$ is locally compact, then the composition and inversion maps,*

$$\mathrm{Loct}_{\text{b-c-o}}(T) \times \mathrm{Loct}_{\text{b-c-o}}(T) \to \mathrm{Loct}_{\text{b-c-o}}(T), \quad (g, f) \mapsto gf,$$
$$\mathrm{Loct}_{\text{b-c-o}}(T) \to \mathrm{Loct}_{\text{b-c-o}}(T), \quad f \mapsto f^{-1},$$

*are continuous.*

### 2B. *Pseudogroups.*

**Definition 2.7** [Sacksteder 1965; Haefliger 2002]. A *pseudogroup* on a space $T$ is a collection $\mathcal{H} \subset \mathrm{Loct}(T)$ such that

- the identity map of $T$ belongs to $\mathcal{H}$ ($\mathrm{id}_T \in \mathcal{H}$);
- if $h, h' \in \mathcal{H}$, then the composite $h'h$ is in $\mathcal{H}$ ($\mathcal{H}^2 \subset \mathcal{H}$);
- $h \in \mathcal{H}$ implies that $h^{-1} \in \mathcal{H}$ ($\mathcal{H}^{-1} \subset \mathcal{H}$);

- if $h \in \mathcal{H}$ and $U$ is open in $\operatorname{dom} h$, then the restriction $h : U \to h(U)$ is in $\mathcal{H}$; and
- if a combination (union) of maps in $\mathcal{H}$ is defined and is a homeomorphism, then it is in $\mathcal{H}$.

**Remark 1.** The following properties hold:

- $\mathrm{id}_U \in \mathcal{H}$ for every open subset $U \subset T$.
- A local transformation $h \in \mathrm{Loct}(T)$ belongs to $\mathcal{H}$ if and only if it locally belongs to $\mathcal{H}$ (any point $x \in \operatorname{dom} h$ has a neighborhood $V_x \subset \operatorname{dom} h$ such that $h|_{V_x} \in \mathcal{H}$).
- Any intersection of pseudogroups on $T$ is a pseudogroup on $T$.

**Example 2.8.** $\mathrm{Loct}(T)$ is the pseudogroup that contains every other pseudogroup on $T$.

**Definition 2.9.** A *subpseudogroup* of a pseudogroup $\mathcal{H}$ on $T$ is a pseudogroup on $T$ contained in $\mathcal{H}$. The *restriction* of $\mathcal{H}$ to an open subset $U \subset T$ is the pseudogroup

$$\mathcal{H}|_U = \{ h \in \mathcal{H} \mid \operatorname{dom} h \cup \operatorname{im} h \subset U \}.$$

The pseudogroup *generated* by a set $S \subset \mathrm{Loct}(T)$ is the intersection of all pseudogroups that contain $S$ (the smallest pseudogroup on $T$ containing $S$).

**Definition 2.10.** Let $\mathcal{H}$ be a pseudogroup on $T$. The *orbit* of each $x \in T$ is the set

$$\mathcal{H}(x) = \{ h(x) \mid h \in \mathcal{H}, x \in \operatorname{dom} h \}.$$

The orbits form a partition of $T$. The space of orbits, equipped with the quotient topology, is denoted by $T/\mathcal{H}$. It is said that $\mathcal{H}$ is

- (*topologically*) *transitive* if some orbit is dense; and
- *minimal* when all orbits are dense.

The following notion, less restrictive than the concept of pseudogroup, is useful to study some properties of pseudogroups.

**Definition 2.11** [Matsumoto 2010]. A *pseudo$*$group* on a space $T$ is a family $S \subset \mathrm{Loct}(T)$ that is closed by the operations of composition and inversion.

**Remark 2.** Any intersection of pseudo$*$groups on $T$ is a pseudo$*$group.

**Definition 2.12.** Any pseudo$*$group contained in another pseudo$*$group is called a *subpseudo$*$group*. The pseudo$*$group *generated* by a subset $S_0$ of $\mathrm{Loct}(T)$ is the intersection of all pseudo$*$groups containing $S_0$ (the smallest pseudo$*$group containing $S_0$).

**Remark 3.** Let $S$ be a pseudo$*$group on $T$, and let $S_1$ be the collection of restrictions of all maps in $S$ to all open subsets of their domains. Then $S_1$ is also a pseudo$*$group on $T$, and $S$ is a subpseudo$*$group of $S_1$.

**Definition 2.13.** In Remark 3, it will be said that $S_1$ is the *localization* of $S$. If $S = S_1$, then the pseudo$*$group $S$ is called *local*.

**Remark 4.** Let $S_0 \subset \mathrm{Loct}(T)$. The pseudo$*$group $S$ generated by $S_0$ consists of all compositions of maps in $S_0$ and their inverses. The pseudogroup $\mathcal{H}$ generated by $S_0$ consists of all $h \in \mathrm{Loct}(T)$ that locally belong to the localization of $S$.

**Remark 5.** If two local pseudo$*$groups, $S_1$ and $S_2$, generate the same pseudogroup $\mathcal{H}$, then $S_1 \cap S_2$ is also a local pseudo$*$group that generates $\mathcal{H}$.

Let $\mathcal{H}$ and $\mathcal{H}'$ be pseudogroups on respective spaces $T$ and $T'$.

**Definition 2.14** [Haefliger 1985; 1988]. A *morphism*[1] $\Phi\colon \mathcal{H} \to \mathcal{H}'$ is a maximal collection of homeomorphisms of open sets of $T$ to open sets of $T'$ such that

- if $\phi \in \Phi$, $h \in \mathcal{H}$ and $h' \in \mathcal{H}'$, then $h'\phi h \in \Phi$ ($\mathcal{H}'\Phi\mathcal{H} \subset \Phi$);
- the family of the domains of maps in $\Phi$ cover $T$; and
- if $\phi, \phi' \in \Phi$, then $\phi'\phi^{-1} \in \mathcal{H}'$ ($\Phi\Phi^{-1} \subset \mathcal{H}'$).

A morphism $\Phi$ is called an *equivalence* if the family $\Phi^{-1} = \{\phi^{-1} \mid \phi \in \Phi\}$ is also a morphism.

**Remark 6.** An equivalence $\Phi : \mathcal{H} \to \mathcal{H}'$ can be characterized as a maximal family of homeomorphisms of open sets of $T$ to open sets of $T'$ such that $\mathcal{H}'\Phi\mathcal{H} \subset \Phi$, and $\Phi^{-1}\Phi$ and $\Phi\Phi^{-1}$ generate $\mathcal{H}'$ and $\mathcal{H}'$, respectively.

**Remark 7.** Any morphism $\Phi : \mathcal{H} \to \mathcal{H}'$ induces a map between the corresponding orbit spaces, $T/\mathcal{H} \to T/\mathcal{H}'$. This map is a homeomorphism if $\Phi$ is an equivalence.

**Definition 2.15.** Let $\Phi_0$ be a family of homeomorphisms of open subsets of $T$ to open subsets of $T'$ such that

- the union of domains of maps in $\Phi_0$ meet all $\mathcal{H}$-orbits; and
- $\Phi_0\mathcal{H}\Phi_0^{-1} \subset \mathcal{H}'$.

Then there is a unique morphism $\Phi : \mathcal{H} \to \mathcal{H}'$ containing $\Phi_0$, which is said to be *generated* by $\Phi_0$. If, moreover,

- the union of images of maps in $\Phi_0$ meet all $\mathcal{H}'$-orbits; and
- $\Phi_0^{-1}\mathcal{H}\Phi_0 \subset \mathcal{H}$;

then $\Phi$ is an equivalence.

[1]This is usually called *étale morphism*. We simply call it morphism because no other type of morphism will be considered here.

**Definition 2.16** [Haefliger 2002]. A pseudogroup $\mathcal{H}$ on a locally compact space $T$ is said to be *compactly generated* if

- there is a relatively compact open subset $U \subset T$ meeting all $\mathcal{H}$-orbits;
- there is a finite set $S = \{h_1, \dots, h_n\} \subset \mathcal{H}|_U$ that generates $\mathcal{H}|_U$; and
- each $h_i$ is the restriction of some $\tilde{h}_i \in \mathcal{H}$ with $\overline{\operatorname{dom} h_i} \subset \operatorname{dom} \tilde{h}_i$.

**Remark 8.** Compact generation is very subtle (see [Ghys 1985; Meigniez 1995]). Haefliger asked when compact generation implies realizability as a holonomy pseudogroup of a compact foliated space. The answer is not always affirmative [Meigniez 2010].

**Definition 2.17** [Haefliger 1985]. A pseudogroup $\mathcal{H}$ is called *quasianalytic* if every $h \in \mathcal{H}$ is the identity around some $x \in \operatorname{dom} h$ whenever $h$ is the identity on some open set whose closure contains $x$.

If a pseudogroup $\mathcal{H}$ on a space $T$ is quasianalytic, then every $h \in \mathcal{H}$ with connected domain is the identity on $\operatorname{dom} h$ if it is the identity on some nonempty open set. Because of this, quasianalyticity is interesting when $T$ is locally connected, but local connectivity is too restrictive in our setting. Then, instead of requiring local connectivity, the following stronger version of quasianalyticity will be used.

**Definition 2.18** [Álvarez and Candel 2009]. A pseudogroup $\mathcal{H}$ on a space $T$ is said to be *strongly quasianalytic* if it is generated by some subpseudo$*$group $S \subset \mathcal{H}$ such that any transformation in $S$ is the identity on its domain if it is the identity on some nonempty open subset of its domain.

**Remark 9.** In [Álvarez and Candel 2009], the term used for the above property is “quasieffective”. However the term “strongly quasianalytic” seems to be more appropriate.

**Remark 10.** If the condition on $\mathcal{H}$ to be strongly quasianalytic is satisfied with a subpseudo$*$group $S$, it is also satisfied with the localization of $S$. It follows that this property is hereditary by taking subpseudogroups and restrictions to open subsets.

**Definition 2.19** [Haefliger 1985]. A pseudogroup $\mathcal{H}$ on a space $T$ is said to be *complete* if, for all $x, y \in T$, there are relatively compact open neighborhoods, $U_x$ of $x$ and $V_y$ of $y$, such that, for all $h \in \mathcal{H}$ and $z \in U_x \cap \operatorname{dom} h$ with $h(z) \in V_y$, there is some $g \in \mathcal{H}$ such that $\operatorname{dom} g = U_x$ and with the same germ as $h$ at $z$.

Since any pseudo$*$group $S$ on $T$ is a subpseudo$*$group of $\operatorname{Loct}(T)$, it can be endowed with the restriction of the (bi-)compact-open topology, also called the *(bi-)compact-open topology* of $S$, and the notation $S_{\text{(b-)c-o}}$ may be used for the corresponding space. In this way, according to Proposition 2.6, if $T$ is locally compact, then $S_{\text{b-c-o}}$ becomes a *topological pseudo$*$group* in the sense that the composition and inversion maps of $S$ are continuous. In particular, this applies to a pseudogroup $\mathscr{H}$ on $T$, obtaining $\mathscr{H}_{\text{(b-)c-o}}$; thus $\mathscr{H}_{\text{b-c-o}}$ is a *topological pseudogroup* in the above sense if $T$ is locally compact.

**Remark 11.** $S_{\text{(b-)c-o}} \hookrightarrow S'_{\text{(b-)c-o}}$ is continuous for pseudo$*$groups $S \subset S'$.

The pseudogroups considered from now on will be assumed to act on locally compact Polish[2] spaces; i.e., locally compact, Hausdorff and second countable spaces [Kechris 1991, Theorem 5.3].

## 2C. *The groupoid of germs of a pseudogroup.*

**Definition 2.20.** A *groupoid* $\mathfrak{G}$ is a small category where every morphism is an isomorphism. This means that $\mathfrak{G}$ is a set (of *morphisms*) equipped with the structure defined by an additional set $T$ (of *objects*), and the following *structural* maps:

- the *source* and *target* maps $s, t : \mathfrak{G} \to T$;
- the *unit* map $T \to \mathfrak{G}$, $x \mapsto 1_x$;
- the *operation* (or *multiplication*) map $\mathfrak{G} \times_T \mathfrak{G} \to \mathfrak{G}$, $(\delta, \gamma) \mapsto \delta\gamma$, where
$$\mathfrak{G} \times_T \mathfrak{G} = \{(\delta, \gamma) \in \mathfrak{G} \times \mathfrak{G} \mid t(\gamma) = s(\delta)\} \subset \mathfrak{G} \times \mathfrak{G};$$
- and the *inversion* map $\mathfrak{G} \to \mathfrak{G}$, $\gamma \mapsto \gamma^{-1}$;

such that the following conditions are satisfied:

- $s(\delta\gamma) = s(\gamma)$ and $t(\delta\gamma) = t(\delta)$ for all $(\delta, \gamma) \in \mathfrak{G} \times_T \mathfrak{G}$;
- for all $\gamma, \delta, \epsilon \in \mathfrak{G}$ with $t(\gamma) = s(\delta)$ and $t(\delta) = s(\epsilon)$, we have $\epsilon(\delta\gamma) = (\epsilon\delta)\gamma$ (associativity);
- $1_{t(\gamma)}\gamma = \gamma 1_{s(\gamma)} = \gamma$ (units or identity elements); and
- $s(\gamma) = t(\gamma^{-1})$, $t(\gamma) = s(\gamma^{-1})$, $\gamma^{-1}\gamma = 1_{s(\gamma)}$ and $\gamma\gamma^{-1} = 1_{t(\gamma)}$ for all $\gamma \in \mathfrak{G}$ (inverse elements).

If moreover $\mathfrak{G}$ and $T$ are equipped with topologies such that all of the above structural maps are continuous, then $\mathfrak{G}$ is called a *topological groupoid*.

**Remark 12.** For a groupoid $\mathfrak{G}$, observe that $s(1_x) = t(1_x) = x$ for all $x \in T$, and therefore the source and target maps $s, t : \mathfrak{G} \to T$ are surjective, and the unit map $T \to \mathfrak{G}$ is injective. If moreover $\mathfrak{G}$ is a topological groupoid, then the unit map $T \to \mathfrak{G}$ is a topological embedding, and therefore the topology of $T$ is determined by the topology of $\mathfrak{G}$; indeed, we can consider $T$ as a subspace of $\mathfrak{G}$ if desired.

**Definition 2.21.** A topological groupoid is called *étale* if the source and target maps are local homeomorphisms.

[2] Recall that a space is called *Polish* if it is separable and completely metrizable.

Let $\mathcal{H}$ be a pseudogroup on a space $T$. Note that the domain of the evaluation partial map $\mathrm{ev} : \mathcal{H} \times T \rightarrowtail T$ is

$$\mathcal{H} * T = \{ (h, x) \in \mathcal{H} \times T \mid x \in \operatorname{dom} h \} \subset \mathcal{H} \times T.$$

Define an equivalence relation on $\mathcal{H} * T$ by setting $(h, x) \sim (h', x')$ if $x = x'$ and $h = h'$ on some neighborhood of $x$ in $\operatorname{dom} h \cap \operatorname{dom} h'$. The equivalence class of each $(h, x) \in \mathcal{H} * T$ is called the *germ* of $h$ at $x$, which will be denoted by $\boldsymbol{\gamma}(h, x)$. The corresponding quotient set is denoted by $\mathfrak{G}$, and the quotient map, $\boldsymbol{\gamma} : \mathcal{H} * T \to \mathfrak{G}$, is called the *germ map*. It is well known that $\mathfrak{G}$ is a groupoid with set of units $T$, where the source and target maps $s, t : \mathfrak{G} \to T$ are given by $s(\boldsymbol{\gamma}(h, x)) = x$ and $t(\boldsymbol{\gamma}(h, x)) = h(x)$, the unit map $T \to \mathfrak{G}$ is defined by $1_x = \boldsymbol{\gamma}(\mathrm{id}_T, x)$, the operation map $\mathfrak{G} \times_T \mathfrak{G} \to \mathfrak{G}$ is given by $\boldsymbol{\gamma}(g, h(x))\, \boldsymbol{\gamma}(h, x) = \boldsymbol{\gamma}(gh, x)$, and the inversion map is defined by $\boldsymbol{\gamma}(h, x)^{-1} = \boldsymbol{\gamma}(h^{-1}, h(x))$.

For $x, y \in T$, let us use the notation $\mathfrak{G}_x = s^{-1}(x)$, $\mathfrak{G}^y = t^{-1}(y)$ and $\mathfrak{G}_x^y = \mathfrak{G}_x \cap \mathfrak{G}^y$; in particular, the group $\mathfrak{G}_x^x$ will be called the *germ group* of $\mathcal{H}$ at $x$. Points in the same $\mathcal{H}$-orbit have isomorphic germ groups (if $y \in \mathcal{H}(x)$, an isomorphism $\mathfrak{G}_y^y \to \mathfrak{G}_x^x$ is given by conjugation with any element in $\mathfrak{G}_x^y$); hence the germ groups of the orbits make sense up to isomorphism. Under pseudogroup equivalences, corresponding orbits have isomorphic germ groups. The set $\mathfrak{G}_x$ will be called the *germ cover* of the orbit $\mathcal{H}(x)$ with base point $x$. The target map restricts to a surjective map $\mathfrak{G}_x \to \mathcal{H}(x)$ whose fibers are bijective to $\mathfrak{G}_x^x$ (if $y \in \mathcal{H}(x)$, a bijection $\mathfrak{G}_x^x \to \mathfrak{G}_x^y$ is given by left product with any element in $\mathfrak{G}_x^y$); thus $\mathfrak{G}_x$ is finite if and only if both $\mathfrak{G}_x^x$ and $\mathcal{H}(x)$ are finite. Moreover germ covers based on points in the same orbit are also bijective (if $y \in \mathcal{H}(x)$, a bijection $\mathfrak{G}_y \to \mathfrak{G}_x$ is given by right product with any element in $\mathfrak{G}_x^y$); therefore the germ covers of the orbits make sense up to bijections.

**Definition 2.22.** It is said that $\mathcal{H}$ is

- *locally free* if all of its germ groups are trivial; and
- *strongly locally free* if $\mathcal{H}$ is generated by a subpseudo$*$group $S \subset \mathcal{H}$ such that, for all $h \in S$ and $x \in \operatorname{dom} h$, if $h(x) = x$ then $h = \mathrm{id}_{\operatorname{dom} h}$.

**Remark 13.** The condition of being (strongly) locally free is stronger than the condition of being (strongly) quasianalytic. If $\mathcal{H}$ is locally free and satisfies the condition of strong quasianalyticity with a subpseudo$*$group $S \subset \mathcal{H}$ generating $\mathcal{H}$, then $\mathcal{H}$ also satisfies the condition of being strongly locally free with $S$.

**Remark 14.** If $\mathcal{H}$ being strongly locally free is witnessed by a subpseudo$*$group $S$, then it is also witnessed by the localization of $S$. It follows that this property is hereditary by taking subpseudogroups and restrictions to open subsets.

The *sheaf topology* on $\mathfrak{G}$ has a basis consisting of the sets $\{\boldsymbol{\gamma}(h, x) \mid x \in \operatorname{dom} h\}$ for $h \in \mathcal{H}$. Equipped with the sheaf topology, $\mathfrak{G}$ is an étale groupoid.

Let us define another topology on $\mathfrak{G}$. Suppose that $\mathcal{H}$ is generated by some subpseudo$*$group $S \subset \mathcal{H}$. The set $S * T = (\mathcal{H} * T) \cap (S \times T)$ is open in $S_{\text{(b-)c-o}} \times T$ by Proposition 2.4. It will be denoted by $S_{\text{(b-)c-o}} * T$ when endowed with the restriction of the topology of $S_{\text{(b-)c-o}} \times T$. The induced quotient topology on $\mathfrak{G}$, via the germ map $\boldsymbol{\gamma} : S_{\text{(b-)c-o}} * T \to \mathfrak{G}$, will also be called the *(bi-)compact-open topology*. The corresponding space will be denoted by $\mathfrak{G}_{\text{(b-)c-o}}$, or by $\mathfrak{G}_{S,\text{(b-)c-o}}$ if reference to $S$ is needed. It follows from Proposition 2.6 that $\mathfrak{G}_{\text{b-c-o}}$ is a topological groupoid if $T$ is locally compact. We get a commutative diagram

$$\begin{array}{ccc} S_{\text{(b-)c-o}} * T & \xrightarrow{\text{inclusion}} & \mathcal{H}_{\text{(b-)c-o}} * T \\ {\scriptstyle\gamma}\downarrow & & \downarrow{\scriptstyle\gamma} \\ \mathfrak{G}_{S,\text{(b-)c-o}} & \xrightarrow{\text{identity}} & \mathfrak{G}_{\mathcal{H},\text{(b-)c-o}} \end{array}$$

where the top map is an embedding and the vertical maps are identifications. Hence the identity map $\mathfrak{G}_{S,\text{(b-)c-o}} \to \mathfrak{G}_{\mathcal{H},\text{(b-)c-o}}$ is continuous. Similarly, the identity map $\mathfrak{G}_{S,\text{b-c-o}} \to \mathfrak{G}_{S,\text{c-o}}$ is continuous.

**Question 2.23.** When are $\mathfrak{G}_{S,\text{(b-)c-o}} = \mathfrak{G}_{\mathcal{H},\text{(b-)c-o}}$ and $\mathfrak{G}_{S,\text{b-c-o}} = \mathfrak{G}_{S,\text{c-o}}$?

For the second equality, a partial answer will be given in Section 3B.

## 2D. *Local groups and local actions.* (See [Jacoby 1957].)

**Definition 2.24.** A *local group* is a quintuple $G \equiv (G, e, \cdot, ', \mathfrak{D})$ satisfying the following conditions:

(1) $(G, \mathfrak{D})$ is a topological space.

(2) $\cdot$ is a function from a subset of $G \times G$ to $G$.

(3) $'$ is a function from a subset of $G$ to $G$.

(4) There is a subset $O$ of $G$ such that
- $O$ is an open neighborhood of $e$ in $G$;
- $O \times O$ is a subset of the domain of $\cdot$;
- $O$ is a subset of the domain of $'$;
- for all $a, b, c \in O$, if $a \cdot b, b \cdot c \in O$, then $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;
- for all $a \in O$, we have $a' \in O$, $a \cdot e = e \cdot a = a$ and $a' \cdot a = a \cdot a' = e$;
- the map $\cdot : O \times O \to G$ is continuous; and
- the map $' : O \to G$ is continuous.

(5) The set $\{e\}$ is closed in $G$.

Asserting that a local group satisfies some topological property usually means that the property is satisfied on some open neighborhood of $e$.

A *local homomorphism* of a local group $G$ to a local group $H$ is a continuous partial map $\phi : G \rightarrowtail H$, whose domain is a neighborhood of $e$ in $G$, which is compatible in the usual sense with the identity elements, the operations and inversions. If moreover $\phi$ restricts to a homeomorphism between some neighborhoods of the identities in $G$ and $H$, then it is called a *local isomorphism*, and $G$ and $H$ are said to be *locally isomorphic*. A local group locally isomorphic to a Lie group is called a *local Lie group*.

The collection of all sets $O$ satisfying (4) is denoted by $\Psi G$. This is a neighborhood basis of $e$ in $G$; all of these neighborhoods are symmetric with respect to the inverse operation (3). Let $\Phi(G, n)$ denote the collection of subsets $A$ of $G$ such that the product of any collection of at most $n$ elements of $A$ is defined, and the set $A^n$ of such products is contained in some $O \in \Psi G$.

Let $H \subset G$. It is said that $H$ is a *subgroup* of $G$ if $H \in \Phi(G, 2)$, $e \in H$, $H' = H$ and $H^2 = H$; and $H$ is a *sub-local group* of $G$ if $H$ is itself a local group with respect to the induced operations and topology.

Let $\Upsilon G$ denote the set of all pairs $(H, V)$ of subsets of $G$ so that $e \in H$, $V \in \Psi G$, $a \cdot b \in H$ for all $a, b \in V \cap H$, and $c' \in H$ for all $c \in V \cap H$. Then a subset $H \subset G$ is a sub-local group if and only if there exists some $V$ such that $(H, V) \in \Upsilon G$ [Jacoby 1957, Theorem 26].

Let $\Pi G$ denote the family of pairs $(H, V)$ of subsets of $G$ such that

$$\begin{aligned}
&e \in H, && V \in \Psi G \cap \Phi(G, 6),\\
&a \cdot b \in H && \text{for all } a, b \in V^6 \cap H,\\
&c' \in H && \text{for all } c \in V^6 \cap H,\\
&V^2 \setminus H \text{ is open.}
\end{aligned}$$

Given $(H, V) \in \Pi G$, there is a (completely regular, Hausdorff) space $G/(V, H)$ and a continuous open surjection $T : V^2 \to G/(V, H)$ such that $T(a) = T(b)$ if and only if $a' \cdot b \in H$ (cf. [Jacoby 1957, Theorem 29]). For another pair in $\Pi G$ of the form $(H, W)$, the spaces $G/(H, V)$ and $G/(H, W)$ are locally homeomorphic at the identity class. Thus the concept of coset space of $H$ is well defined in this sense, as "a germ of a topological space". The notation $G/H$ may be used in this sense. It will be said that $G/H$ has a certain topological property when some $G/(H, V)$ has that property around $T(e)$.

Let $\Delta G$ be the set of pairs $(H, U)$ such that $(H, U) \in \Pi G$ and $b' \cdot (a \cdot b) \in H$ for all $a \in H \cap U^4$ and $b \in U^2$. A subset $H \subset G$ is called a normal sub-local group of $G$ if there exists $U$ such that $(H, U) \in \Delta G$. If $(H, U) \in \Delta G$ then the quotient space $G/(H, U)$ admits the structure of a local group (see [Jacoby 1957, Theorem 35] for details) and the natural projection $T : U^2 \to G/(H, U)$ is a local homomorphism. As before, another such pair $(H, V)$ produces a locally isomorphic quotient local group.

As usual, $a \cdot b$ and $a'$ will be denoted by $ab$ and $a^{-1}$.

Local groups were first studied by Jacoby [1957], giving local versions of important theorems for topological groups. For instance, Jacoby characterized local Lie groups as the locally compact local groups without small subgroups[3] [Jacoby 1957, Theorem 96]. Also, any finite dimensional metrizable locally compact local group is locally isomorphic to the direct product of a Lie group and a compact zero-dimensional topological group [Jacoby 1957, Theorem 107]. In particular, this property shows that any locally Euclidean local group is a local Lie group, which is an affirmative answer to a local version of Hilbert's 5th problem. However the proof of Jacoby is incorrect because he did not realize that, in local groups, associativity for three elements does not imply associativity for any finite sequence of elements [Plaut 1993; Olver 1996]. Fortunately, a completely new proof of the local Hilbert's 5th problem was given by Goldbring [2010]. Moreover van den Dries and Goldbring [2010; 2012] proved that any locally compact local group is locally isomorphic to a topological group, and therefore all other theorems for local groups of Jacoby hold as well because they are known for locally compact topological groups [Montgomery and Zippin 1955].

**Definition 2.25.** It is said that a local group $G$ can be *approximated* by a class $\mathscr{C}$ of local groups if, for all $W \in \Psi G \cap \Phi(G, 2)$, there is some $V \in \Psi G$ and a sequence of compact normal subgroups $F_n \subset V$ such that $V \subset W$, $F_{n+1} \subset F_n$, $\bigcap_n F_n = \{e\}$, $(F_n, V) \in \Delta G$ and $G/(F_n, V) \in \mathscr{C}$.

**Theorem 2.26** [Jacoby 1957, Theorems 97–103; van den Dries and Goldbring 2010; 2012]. *Any locally compact second countable local group $G$ can be approximated by local Lie groups.*

**Definition 2.27.** A *local action* of a local group $G$ on a space $X$ is a paro map $G \times X \rightarrowtail X$, $(g, x) \mapsto gx$, defined on some open neighborhood of $\{e\} \times X$, such that $ex = x$ for all $x \in X$, and $g_1(g_2x) = (g_1g_2)x$, provided both sides are defined.

**Remark 15.** The local transformations given by any local action of a local group on a space generate a pseudogroup.

A local action of a local group $G$ on a space $X$ is called *locally transitive* at some point $x \in X$ if there is a neighborhood $W$ of $e$ in $G$ such that the local action is defined on $W \times \{x\}$, and $Wx := \{gx \mid g \in W\}$ is a neighborhood of $x$ in $X$. Given another local action of $G$ on a space $Y$, a paro map $\phi : X \rightarrowtail Y$ is called *equivariant* if $\phi(gx) = g\phi(x)$ for all $x \in X$ and $g \in G$, provided both sides are defined.

[3] A local group is said to have no small subgroups when some neighborhood of the identity element contains no nontrivial subgroup.

**Example 2.28.** Let $H$ be a sub-local group of $G$. If $(H, V) \in \Pi G$, and if the map $T : V^2 \to G/(H, V)$ is the natural projection, then the map

$$V \times G/(H, V) \to G/(H, V), \quad (v, T(g)) \mapsto T(vg)$$

defines a local action of $G$ on $G/(H, V)$.

**Remark 16.** If $G$ is a local group locally acting on $X$ and the local action is locally transitive at $x \in X$, then there is a sub-local group $H$ of $G$ such that $(H, V) \in \Pi G$ for some $V$ and the orbit paro map $G \rightarrowtail X$, $g \mapsto gx$, induces an equivariant paro map $G/(H, V) \rightarrowtail X$, which restricts to a homeomorphism between neighborhoods of $T(e)$ and $x$.

**2E. *Equicontinuous pseudogroups.*** Álvarez and Candel [2009] introduced the following structure to define equicontinuity for pseudogroups. Let[4] $\{T_i, d_i\}$ be a family of metric spaces such that $\{T_i\}$ is a covering of a set $T$, each intersection $T_i \cap T_j$ is open in $(T_i, d_i)$ and $(T_j, d_j)$, and, for all $\epsilon > 0$, there is some $\delta(\epsilon) > 0$ such that the following property holds: for all $i$, $j$ and $z \in T_i \cap T_j$, there is some open neighborhood $U_{i,j,z}$ of $z$ in $T_i \cap T_j$ (with respect to the topology induced by $d_i$ and $d_j$) such that

$$d_i(x, y) < \delta(\epsilon) \Longrightarrow d_j(x, y) < \epsilon$$

for all $\epsilon > 0$ and all $x, y \in U_{i,j,z}$. Such a family is called a *cover of $T$ by quasilocally equal metric spaces*. Two such families are *quasilocally equal* when their union is also a cover of $T$ by quasilocally equal metric spaces. This is an equivalence relation whose equivalence classes are called *quasilocal metrics* on $T$. For each quasilocal metric $\mathfrak{Q}$ on $T$, the pair $(T, \mathfrak{Q})$ is called a *quasilocal metric space*. Such a $\mathfrak{Q}$ induces a topology[5] on $T$ so that, for each $\{T_i, d_i\}_{i \in I} \in \mathfrak{Q}$, the family of open balls of all metric spaces $(T_i, d_i)$ form a basis of open sets. Any topological concept or property of $(T, \mathfrak{Q})$ refers to this underlying topology. $(T, \mathfrak{Q})$ is locally compact and Polish if and only if it is Hausdorff, paracompact and separable [Álvarez and Candel 2009].

**Definition 2.29** [Álvarez and Candel 2009]. Let $\mathcal{H}$ be a pseudogroup on a quasilocal metric space $(T, \mathfrak{Q})$. Then $\mathcal{H}$ is said to be (*strongly*[6]) *equicontinuous* if there exists some $\{T_i, d_i\}_{i \in I} \in \mathfrak{Q}$ and some subpseudo$*$group $S \subset \mathcal{H}$ generating $\mathcal{H}$, such that, for every $\epsilon > 0$, there is some $\delta(\epsilon) > 0$ such that

$$d_i(x, y) < \delta(\epsilon) \Longrightarrow d_j(h(x), h(y)) < \epsilon$$

for all $h \in S$, $i, j \in I$ and $x, y \in T_i \cap h^{-1}(T_j \cap \operatorname{im} h)$.

[4]The notation will be simplified by using, for instance, $\{T_i, d_i\}$ instead of $\{(T_i, d_i)\}$.

[5]In fact, it induces a uniformity. We could even use any uniformity to define equicontinuity, but such generality will not be used here.

[6]This adverb, used in [Álvarez and Candel 2009], will be omitted for the sake of simplicity.

A pseudogroup $\mathcal{H}$ acting on a space $T$ will be called (*strongly*) *equicontinuous* when it is equicontinuous with respect to some quasilocal metric inducing the topology of $T$.

**Remark 17.** If the equicontinuity of $\mathcal{H}$ is witnessed by a subpseudo$*$group $S$, then it is also witnessed by the localization of $S$. It follows that equicontinuity is hereditary by taking subpseudogroups and restrictions to open subsets.

**Lemma 2.30** [Álvarez and Candel 2009, Lemma 8.8]. *Let $\mathcal{H}$ and $\mathcal{H}'$ be equivalent pseudogroups on locally compact Polish spaces. Then $\mathcal{H}$ is equicontinuous if and only if $\mathcal{H}'$ is equicontinuous.*

**Proposition 2.31** [Álvarez and Candel 2009, Proposition 8.9]. *Let $\mathcal{H}$ be a compactly generated and equicontinuous pseudogroup on a locally compact Polish quasilocal metric space $(T, \mathfrak{Q})$, and let $U$ be any relatively compact open subset of $(T, \mathfrak{Q})$ that meets every $\mathcal{H}$-orbit. Suppose that $\{T_i, d_i\}_{i\in I} \in \mathfrak{Q}$ satisfies the condition of equicontinuity. Let $E$ be any system of compact generation of $\mathcal{H}$ on $U$, and let $\bar{g}$ be an extension of each $g \in E$ with $\overline{\operatorname{dom} g} \subset \operatorname{dom} \bar{g}$. Also, let $\{T'_i\}_{i\in I}$ be any shrinking*[7] *of $\{T_i\}_{i\in I}$. Then there is a finite family $\mathcal{V}$ of open subsets of $(T, \mathfrak{Q})$ whose union contains $U$ and such that, for any $V \in \mathcal{V}$, $x \in U \cap V$ and $h \in \mathcal{H}$ with $x \in \operatorname{dom} h$ and $h(x) \in U$, the domain of $\tilde{h} = \bar{g}_n \cdots \bar{g}_1$ contains $V$ for any composite $h = g_n \cdots g_1$ defined around $x$ with $g_1, \dots, g_n \in E$. Moreover, $V \subset T'_{i_0}$ and $\tilde{h}(V) \subset T'_{i_1}$ for some $i_0, i_1 \in I$.*

**Remark 18.** The statement of Proposition 2.31 is stronger than the completeness of $\mathcal{H}|_U$. Since we can choose $U$ large enough to contain two arbitrarily given points of $T$, it follows $\mathcal{H}$ is complete.

**Proposition 2.32** [Álvarez and Candel 2009, Proposition 9.9]. *Let $\mathcal{H}$ be a compactly generated, equicontinuous and strongly quasianalytic pseudogroup on a locally compact Polish space $T$. Suppose that the conditions of equicontinuity and strong quasianalyticity are satisfied with a subpseudo$*$group $S \subset \mathcal{H}$ generating $\mathcal{H}$. Let $A$, $B$ be open subsets of $T$ such that $\overline{A}$ is compact and contained in $B$. If $x$ and $y$ are close enough points in $T$, then*

$$f(x) \in A \Rightarrow f(y) \in B$$

*for all $f \in S$ whose domain contains $x$ and $y$.*

**Theorem 2.33** [Álvarez and Candel 2009, Theorem 11.11]. *Let $\mathcal{H}$ be a compactly generated and equicontinuous pseudogroup on a locally compact Polish space $T$. If $\mathcal{H}$ is transitive, then $\mathcal{H}$ is minimal.*

---

[7] Recall that a *shrinking* of an open cover $\{U_i\}$ of a space $X$ is an open cover $\{U'_i\}$ of $X$, with the same index set, such that $\overline{U'_i} \subset U_i$ for all $i$. Similarly, if $\{U_i\}$ is a cover of a subset $A \subset X$ by open subsets of $X$, a *shrinking* of $\{U_i\}$, as a cover of $A$ by open subsets of $X$, is a cover $\{U'_i\}$ of $A$ by open subsets of $X$, with the same index set, such that $\overline{U'_i} \subset U_i$ for all $i$.

Theorem 2.33 can be restated by saying that the orbit closures form a partition of the space. The following result states that indeed the orbit closures are orbits of a pseudogroup if strong quasianalyticity is also assumed.

**Theorem 2.34** [Álvarez and Candel 2009, Theorem 12.1]. *Let $\mathcal{H}$ be a strongly quasianalytic, compactly generated and equicontinuous pseudogroup on a locally compact Polish space $T$. Let $S \subset \mathcal{H}$ be a subpseudo$*$group generating $\mathcal{H}$ such that $\mathcal{H}$ satisfies the conditions of equicontinuity and strong quasianalyticity with $S$. Let $\widetilde{\mathcal{H}}$ be the set of maps $h$ between open subsets of $T$ that satisfy the property that for every $x \in \operatorname{dom} h$, there exists a neighborhood $O_x$ of $x$ in $\operatorname{dom} h$ such that the restriction $h|_{O_x}$ is in the closure of $C(O_x, T) \cap S$ in $C_{\text{c-o}}(O_x, T)$. Then*

(i) *$\widetilde{\mathcal{H}}$ is closed by composition, combination and restriction to open sets*;

(ii) *any map in $\widetilde{\mathcal{H}}$ is a homeomorphism around every point of its domain*;

(iii) *$\overline{\mathcal{H}} = \widetilde{\mathcal{H}} \cap \operatorname{Loct}(T)$ is a pseudogroup that contains $\mathcal{H}$*;

(iv) *$\overline{\mathcal{H}}$ is equicontinuous*;

(v) *the orbits of $\overline{\mathcal{H}}$ are equal to the closures of the orbits of $\mathcal{H}$; and*

(vi) *$\widetilde{\mathcal{H}}$ and $\overline{\mathcal{H}}$ are independent of the choice of $S$.*

**Remark 19.** In Theorem 2.34, let $\overline{S}$ be the set of local transformations that are in the union of the closures of $C(O, T) \cap S$ in $C_{\text{c-o}}(O, T)$ with $O$ running on the open sets of $T$. According to the proof of [Álvarez and Candel 2009, Theorem 12.1], $\overline{S}$ is a pseudo$*$group that generates $\overline{\mathcal{H}}$. Moreover, if $\mathcal{H}$ satisfies the equicontinuity condition with $S$ and some representative $\{T_i, d_i\}$ of a quasilocal metric, then $\overline{\mathcal{H}}$ satisfies the equicontinuity condition with $\overline{S}$ and $\{T_i, d_i\}$.

**Remark 20.** From the proof of [Álvarez and Candel 2009, Theorem 12.1], it also follows that, with the notation of Remark 19, any $x \in \overline{U}$ has a neighborhood $O$ in $T$ such that the closure of

$$\{h \in C(O, T) \cap S \mid h(O) \cap \overline{U} \neq \varnothing\}$$

in $C_{\text{c-o}}(O, T)$ is contained in $\operatorname{Loct}(T)$, and therefore in $\overline{S}$.

**Example 2.35.** Let $G$ be a locally compact Polish local group with a left invariant metric, let $\Gamma \subset G$ be a dense sub-local group, and let $\mathcal{H}$ be the minimal pseudogroup generated by the local action of $\Gamma$ by local left translations on $G$. The local left and right translations in $G$ by each $g \in G$ will be denoted by $L_g$ and $R_g$. The restrictions of the local left translations $L_\gamma$ ($\gamma \in \Gamma$) to open subsets of their domains form a subpseudo$*$group $S \subset \mathcal{H}$ that generates $\mathcal{H}$. Obviously, $\mathcal{H}$ satisfies with $S$ the condition of being strongly locally free, and therefore strongly quasianalytic. Moreover $\mathcal{H}$ satisfies with $S$ the condition of being equicontinuous (indeed isometric) by considering any left invariant metric on $G$. Observe that any local right translation $R_g$ $(g \in G)$ generates an equivalence $\mathcal{H} \to \mathcal{H}$.

Now, suppose that $\mathcal{H}$ is compactly generated. Then $\overline{\mathcal{H}}$ is generated by the local action of $G$ on itself by local left translations. The subpseudo$*$group $\overline{S} \subset \overline{\mathcal{H}}$ consists of the restrictions of the local left translations $L_g$ $(g \in G)$ to open subsets of their domains. Observe that $\overline{\mathcal{H}}$ satisfies the condition of being strongly locally free, and therefore strongly quasianalytic, with $\overline{S}$.

**Lemma 2.36.** *Let $G$ and $G'$ be locally compact Polish local groups with left invariant metrics, let $\Gamma \subset G$ and $\Gamma' \subset G'$ be dense sub-local groups, and let $\mathcal{H}$ and $\mathcal{H}'$ be the pseudogroups generated by the local actions of $\Gamma$ and $\Gamma'$ by local left translations on $G$ and $G'$. Suppose that $\mathcal{H}$ and $\mathcal{H}'$ are compactly generated. Then $\mathcal{H}$ and $\mathcal{H}'$ are equivalent if and only if $G$ is locally isomorphic to $G'$.*

*Proof.* Consider the notation and observations of Example 2.35 for both $G$ and $G'$; in particular, $S \subset \mathcal{H}$ and $S' \subset \mathcal{H}'$ denote the subpseudo$*$groups of restrictions of local translations $L_\gamma$ and $L_{\gamma'}$ $(\gamma \in \Gamma$ and $\gamma' \in \Gamma')$ to open subsets of their domains. Let $e$ and $e'$ denote the identity elements of $G$ and $G'$. Let $\Phi : \mathcal{H} \to \mathcal{H}'$ be en equivalence. Since $\mathcal{H}'$ is minimal, after composing $\Phi$ with the equivalence generated by some local right translation in $G$ if necessary, we can assume that $\phi(e) = e'$ for some $\phi \in \Phi$ with $e \in \operatorname{dom} \phi$.

Let $U$ be a relatively compact open symmetric neighborhood of $e$ in $G$ with $\overline{U} \subset \operatorname{dom} \phi$. Let $\{f_1, \dots, f_n\}$ be a symmetric system of compact generation of $\mathcal{H}$ on $U$. Thus each $f_i$ has an extension $\tilde{f}_i \in \mathcal{H}$ such that $\overline{\operatorname{dom} f_i} \subset \operatorname{dom} \tilde{f}_i \subset \operatorname{dom} \phi$.

**Claim 1.** *We can assume that $\tilde{f}_i \in S$ and $\phi \tilde{f}_i \phi^{-1} \in S'$ for all $i$.*

Each point in $\operatorname{dom} \tilde{f}_i \cap \operatorname{dom} \phi$ has an open neighborhood $O$ such that $O \subset \operatorname{dom} \tilde{f}_i$, $\tilde{f}_i|_O \in S$ and $\phi \tilde{f}_i \phi^{-1}|_{\phi(O)} \in S'$. Take a finite covering $\{O_{ij}\}$ $(j \in \{1, \dots, k_i\})$ of the compact set $\overline{\operatorname{dom} f_i}$ by sets of this type. Let $\{P_{ij}\}$ be a shrinking of $\{O_{ij}\}$, as a cover of $\overline{\operatorname{dom} f_i}$ by open subsets of $\operatorname{dom} \tilde{f}_i$. Then the restrictions $g_{ij} = f_i|_{P_{ij} \cap U}$ $(i \in \{1, \dots, n\}$ and $j \in \{1, \dots, k_i\})$ generate $\mathcal{H}|_U$, each $\tilde{g}_{ij} = \tilde{f}_i|_{O_{ij}}$ is in $S$ and extends $g_{ij}$, $\overline{\operatorname{dom} g_{ij}} \subset \operatorname{dom} \tilde{g}_{ij}$, and $\phi \tilde{g}_{ij} \phi^{-1} \in S'$, showing Claim 1.

According to Claim 1, the maps $f'_i = \phi f_i \phi^{-1}$ form a symmetric system of compact generation of $\mathcal{H}'$ on $U' = \phi(U)$, which can be checked with the extensions $\tilde{f}'_i = \phi \tilde{f}_i \phi^{-1}$. Let $S_0 \subset S$ and $S'_0 \subset S'$ be the subpseudo$*$groups consisting of the restrictions of compositions of maps $f_i$ and $f'_i$ to open subsets of their domains, respectively. They generate $\mathcal{H}$ and $\mathcal{H}'$. It follows from Claim 1 that $\phi f \phi^{-1} \in S'$ for all $f \in S_0$. On the other hand, by Proposition 2.31, there is a smaller open neighborhood of the identity, $V \subset U$, such that, for all $h \in \mathcal{H}$ and all $x \in V \cap \operatorname{dom} h$ with $h(x) \in U$, there is some $f \in S_0$ such that $\operatorname{dom} f = V$ and $\boldsymbol{\gamma}(f, x) = \boldsymbol{\gamma}(h, x)$.

Let $W$ be another symmetric open neighborhood of the identity such that $W^2 \subset V$. Let us show that $\phi : W \to \phi(W)$ is a local isomorphism. Let $\gamma \in W \cap \Gamma$. The restriction $L_\gamma : W \to \gamma W$ is well defined and belongs to $S$. Hence there is some $f \in S_0$ such that $\operatorname{dom} f = V$ and $\boldsymbol{\gamma}(f, e) = \boldsymbol{\gamma}(L_\gamma, e)$. Since $f$ is also a restriction of a local left translation in $G$, it follows that $f = L_\gamma$ on $W$. So $\phi L_\gamma \phi^{-1}|_{\phi(W)} \in S'$; i.e., there is some $\gamma' \in \Gamma'$ such that $\phi L_\gamma \phi^{-1} = L_{\gamma'}$ on $\phi(W)$. In fact,

$$\phi(\gamma) = \phi L_\gamma(e) = \phi L_\gamma \phi^{-1}(e') = L_{\gamma'}(e') = \gamma'.$$

Hence, for all $\gamma, \delta \in \Gamma$,

$$\phi(\gamma\delta) = \phi L_\gamma(\delta) = L_{\phi(\gamma)}\phi(\delta) = \phi(\gamma)\phi(\delta),$$

$$\begin{aligned}\phi(\gamma)^{-1} &= L^{-1}_{\phi(\gamma)}(e') = (\phi L_\gamma \phi^{-1})^{-1}(e') \\ &= \phi L_{\gamma^{-1}}\phi^{-1}(e') = L_{\phi(\gamma^{-1})}(e') = \phi(\gamma^{-1}).\end{aligned}$$

Since $\phi$ and the product and inversion maps are continuous, it follows that, for all $g, h \in W$, we have $\phi(gh) = \phi(g)\phi(h)$ and $\phi(g^{-1}) = \phi(g)^{-1}$. □

**Example 2.37.** This generalizes Example 2.35. Let $G$ be a locally compact Polish local group with a left invariant metric, $K \subset G$ a compact subgroup, and $\Gamma \subset G$ a dense sub-local group. Take some $V$ such that $(H, V) \in \Pi(G)$. The left invariant metric on $G$ can be assumed to be also $K$-right invariant by the compactness of $K$, and therefore it defines a metric on $G/(K, V)$. Then the canonical local action of $\Gamma$ on some neighborhood of the identity class in $G/(K, V)$ induces a transitive equicontinuous pseudogroup $\mathcal{H}$ on a locally compact Polish space; in fact, this is a pseudogroup of local isometries.

Assume that $\mathcal{H}$ is compactly generated. Then $\overline{\mathcal{H}}$ is generated by the canonical local action of $G$ on some neighborhood of the identity class in $G/(K, V)$. Moreover the subpseudo$*$group $\overline{S} \subset \overline{\mathcal{H}}$ consists of the local translations of the local action of $G$ on $G/(K, V)$.

Examples 2.35 and 2.37 are particular cases of pseudogroups induced by local actions (Remark 15). The following result indicates their relevance.

**Theorem 2.38** [Álvarez and Candel 2010, Theorem 5.2]. *Let $\mathcal{H}$ be a transitive, compactly generated and equicontinuous pseudogroup on a locally compact Polish space, and suppose that $\overline{\mathcal{H}}$ is strongly quasianalytic. Then $\mathcal{H}$ is equivalent to a pseudogroup of the type described in Example 2.37.*

**Remark 21.** From the proof of [Álvarez and Candel 2010, Theorems 3.3 and 5.2], it also follows that, in Theorem 2.38, if moreover $\overline{\mathcal{H}}$ is strongly locally free, then $\mathcal{H}$ is equivalent to a pseudogroup of the type described in Example 2.35.

## 3. Molino's theory for equicontinuous pseudogroups

**3A.** ***Conditions on*** $\mathcal{H}$**.** Let $\mathcal{H}$ be a pseudogroup of local transformations of a locally compact Polish space $T$. Suppose that $\mathcal{H}$ is compactly generated, complete and equicontinuous, and that $\overline{\mathcal{H}}$ is also strongly quasianalytic.

Let $U$ be a relatively compact open set in $T$ that meets all the orbits of $\mathcal{H}$. The condition of compact generation is satisfied with $U$. Consider a representative $\{T_i, d_i\}$ of a quasilocal metric on $T$ satisfying the condition of equicontinuity of $\mathcal{H}$ with some subpseudo$*$group $S \subset \mathcal{H}$ that generates $\mathcal{H}$. We can also suppose that the condition of strong quasianalyticity of $\mathcal{H}$ is satisfied with $S$.

**Remark 22.** According to Theorem 2.34 and Remark 19, there is a mapping $\epsilon \mapsto \delta(\epsilon) > 0$ ($\epsilon > 0$) such that

$$d_i(x, y) < \delta(\epsilon) \Longrightarrow d_j(h(x), h(y)) < \epsilon$$

for all indices $i$ and $j$, every $h \in \overline{S}$, and $x, y \in T_i \cap h^{-1}(T_j \cap \operatorname{im} h)$.

**Remark 23.** By Remark 20 and refining $\{T_i\}$ if necessary, we can assume that $\overline{U}$ is covered by a finite collection $\{T_{i_1}, \dots, T_{i_r}\}$ of the sets $T_i$, such that the closure of

$$\{h \in C(T_{i_k}, T) \cap S \mid h(T_{i_k}) \cap \overline{U} \neq \varnothing\}$$

in $C_{\text{c-o}}(T_{i_k}, T)$ is contained in $\overline{S}$ for all $k \in \{1, \dots, r\}$.

**Remark 24.** By Proposition 2.31 and Remark 23, and refining $\{T_i\}$ if necessary, we can assume that, for all $h \in \overline{\mathcal{H}}$ and $x \in T_{i_k} \cap U \cap \operatorname{dom} h$ with $h(x) \in U$, there is some $\tilde{h} \in \overline{S}$ with $\operatorname{dom} \tilde{h} = T_{i_k}$ and $\boldsymbol{\gamma}(h, x) = \boldsymbol{\gamma}(\tilde{h}, x)$.

**Remark 25.** By Remarks 5, 10 and 17, and refining $\{T_i\}$ if necessary, we can assume that the strong quasianalyticity of $\overline{\mathcal{H}}$ is satisfied with $\overline{S}$.

**3B.** ***Coincidence of topologies.***

**Proposition 3.1.** $\overline{S}_{\text{b-c-o}} = \overline{S}_{\text{c-o}}$.

*Proof.* (This is inspired by [Arens 1946].) For each $g \in \overline{S}$, take any index $i$ and open sets $V, W \subset T$ such that $\overline{V} \subset W$ and $\overline{W} \subset \operatorname{im} g$. By Proposition 2.32, there is some $\epsilon(i, V, W) > 0$ such that, for all $x, y \in T_i$, if $d_i(x, y) < \epsilon(i, V, W)$, then

$$f(x) \in \overline{V} \Longrightarrow f(y) \in W$$

for all $f \in \overline{S}$ with $x, y \in \operatorname{dom} f$. Let $\mathcal{K}(g, i, V, W)$ be the family of compact subsets $K \subset T_i \cap \operatorname{dom} g$ such that

$$\mathring{K} \neq \varnothing, \quad \operatorname{diam}_{d_i}(K) < \epsilon(i, V, W), \quad g(K) \subset V,$$

where $\mathring{K}$ and $\operatorname{diam}_{d_i}(K)$ denote the interior and $d_i$-diameter of $K$. Moreover let $\mathcal{K}(g)$ denote the union of the families $\mathcal{K}(g, i, V, W)$ as above. Then a subbasis

$\mathcal{N}(g)$ of open neighborhoods of each $g$ in $\overline{S}_{\text{c-o}}$ is given by the sets $\mathcal{N}(K, O) \cap \overline{S}$, where $K \in \mathcal{K}(g)$ and $O$ is an open neighborhood of $g(K)$ in $T$.

We have to prove the continuity of the inversion map $\overline{S}_{\text{c-o}} \to \overline{S}_{\text{c-o}}$, $h \mapsto h^{-1}$. Let $h \in \overline{S}$ and let $\mathcal{N}(K, O) \in \mathcal{N}(h^{-1})$ with $K \in \mathcal{K}(h^{-1}, i, V, W)$, and fix any point $x \in \mathring{K}$. Then

$$\mathcal{V} = \mathcal{N}(\{h^{-1}(x)\}, \mathring{K}) \cap \mathcal{N}(\overline{W} \setminus O, T \setminus K)$$

is an open neighborhood of $h$ in $\mathcal{H}_{\text{c-o}}$. We have $d_i(fh^{-1}(x), y) < \epsilon(i, V, W)$ for all $f \in \mathcal{V} \cap \overline{S}$ and $y \in K$ since $fh^{-1}(x) \in \mathring{K}$ and $\operatorname{diam}_{d_i}(K) < \epsilon(i, V, W)$. So $f^{-1}(y) \in W$ by the definition of $\epsilon(i, V, W)$ since $f^{-1} \in \overline{S}$ and $h^{-1}(x) \in h^{-1}(K) \subset V$. Thus, if $f^{-1}(y) \notin O$, we get $f^{-1}(y) \in \overline{W} \setminus O$, obtaining $y \in T \setminus K$, which is a contradiction. Hence $f^{-1} \in \mathcal{N}(K, O)$ for all $f \in \mathcal{V} \cap \overline{S}$. □

Let $\overline{\mathfrak{G}}$ denote the groupoid of germs of $\overline{\mathcal{H}}$. The following direct consequence of Proposition 3.1 gives a partial answer to Question 2.23.

**Corollary 3.2.** $\overline{\mathfrak{G}}_{\overline{S},\text{b-c-o}} = \overline{\mathfrak{G}}_{\overline{S},\text{c-o}}$; *i.e.*, $\overline{\mathfrak{G}}_{\overline{S},\text{c-o}}$ *is a topological groupoid.*

**3C.** ***The space $\widehat{T}$.*** Recall that $s, t : \overline{\mathfrak{G}}_{\overline{S},\text{c-o}} \to T$ denote the source and target projections. Let $\widehat{T} = \mathfrak{G}_{\overline{S},\text{c-o}}$, where the following subsets are open:

$$\widehat{T}_U = s^{-1}(U) \cap t^{-1}(U), \quad \widehat{T}_{k,l} = s^{-1}(T_{i_k,i_l}) \cap t^{-1}(T_{i_k,i_l}), \quad \widehat{T}_{U,k,l} = \widehat{T}_U \cap \widehat{T}_{k,l}.$$

Observe that $\widehat{T}_U$ is an open subspace of $\widehat{T}$, and the family of sets $\widehat{T}_{U,k,l}$ form an open covering of $\widehat{T}_U$.

Let $\boldsymbol{\gamma}(h, x) \in \widehat{T}_{U,k,l}$. We can assume that $h \in \overline{S}$ and $\operatorname{dom} h = T_{i_k}$ according to Remark 24. Since $x \in T_{i_k} \cap U$ and $h(x) \in T_{i_l} \cap U$, there are relatively compact open neighborhoods, $V$ of $x$ and $W$ of $h(x)$, such that $\overline{V} \subset T_{i_k} \cap U$, $\overline{W} \subset T_{i_l} \cap U$ and $h(\overline{V}) \subset W$.

By Remark 24, for each $f \in \overline{S}$ with $x \in \operatorname{dom} f$, there is some $\tilde{f} \in \overline{S}$ with $\operatorname{dom} \tilde{f} = T_{i_k}$ and $\boldsymbol{\gamma}(\tilde{f}, x) = \boldsymbol{\gamma}(f, x)$.

**Lemma 3.3.** *We have $f = \tilde{f}$ on $V$.*

*Proof.* The composition $f|_V \tilde{f}^{-1}$ is defined on $\tilde{f}(V)$, belongs to $\overline{S}$, and is the identity on some neighborhood of $\tilde{f}(x) = f(x)$. So $f|_V \tilde{f}^{-1}$ is the identity on $\tilde{f}(V)$ because $\overline{\mathcal{H}}$ satisfies strong quasianalyticity with $\overline{S}$. Hence $f = \tilde{f}$ on $V$. □

Let

$$\overline{S}_0 = \{ f \in \overline{S} \mid \overline{V} \subset \operatorname{dom} f,\ f(\overline{V}) \subset W \}, \tag{3}$$

$$\overline{S}_1 = \{ f \in \overline{S} \mid \overline{V} \subset \operatorname{dom} f,\ f(\overline{V}) \subset \overline{W} \}, \tag{4}$$

equipped with the restriction of the compact-open topology. Notice that $\overline{S}_0$ is an open neighborhood of $h$ in $\overline{S}_{\text{c-o}}$. Consider the compact-open topology on $C(\overline{V}, \overline{W})$.

**Lemma 3.4.** *The restriction map* $\mathcal{R} : \overline{S}_1 \to C(\overline{V}, \overline{W})$, $\mathcal{R}(f) = f|_{\overline{V}}$, *defines an identification* $\mathcal{R} : \overline{S}_1 \to \mathcal{R}(\overline{S}_1)$.

*Proof.* The continuity of $\mathcal{R}$ is elementary.

Let $G \subset \mathcal{R}(\overline{S}_1)$ such that $\mathcal{R}^{-1}(G)$ is open in $\overline{S}_1$. For each $g_0 \in G$, there is some $g_0' \in \mathcal{R}^{-1}(G)$ such that $\mathcal{R}(g_0') = g_0$. Since $\mathcal{R}^{-1}(G)$ is open in $\overline{S}_1$, there are finite collections $\{K_1, \dots, K_p\}$ of compact subsets and $\{O_1, \dots, O_p\}$ of open subsets, such that

$$g_0' \in \left\{ f \in \overline{S}_1 \;\middle|\; \bigcup_{i=1}^p K_i \subset \operatorname{dom} f \text{ and } f(K_i) \subset O_i \text{ for each } i \right\} \subset \mathcal{R}^{-1}(G).$$

Then

$$g_0 \in \left\{ g \in \overline{S}_1 \;\middle|\; \bigcup_{i=1}^p K_i \cap \overline{V} \subset \operatorname{dom} g \text{ and } g(K_i \cap \overline{V}) \subset O_i \cap \overline{W} \text{ for each } i \right\} \subset G.$$

Since $K_1 \cap \overline{V}, \dots, K_p \cap \overline{V}$ are compact in $\overline{V}$ and $O_1 \cap \overline{W}, \dots, O_p \cap \overline{W}$ are open in $\overline{W}$, it follows that $g_0$ is in the interior of $G$ in $\mathcal{R}(\overline{S}_1)$. Hence $G$ is open in $\mathcal{R}(\overline{S}_1)$. □

**Lemma 3.5.** $\mathcal{R}(\overline{S}_1)$ *is closed in* $C(\overline{V}, \overline{W})$.

*Proof.* Observe that $C(\overline{V}, \overline{W})$ is second countable because $T$ is Polish. Take a sequence $g_n$ in $\mathcal{R}(\overline{S}_1)$ converging to $g$ in $C(\overline{V}, \overline{W})$. Then it easily follows that $g_n|_V$ converges to $g|_V$ in $C(V, T)$ with the compact-open topology. Thus $g|_V \in \overline{S}$ according to Remark 23. Let $f = \widetilde{g|_V}$. By Lemma 3.3, we have $g = f|_{\overline{V}}$. Therefore $f \in \overline{S}_1$ and $g = \mathcal{R}(f)$. □

**Corollary 3.6.** $\mathcal{R}(\overline{S}_1)$ *is compact in* $C(\overline{V}, \overline{W})$.

*Proof.* This follows by the Arzelà–Ascoli theorem and Lemma 3.5, because $\overline{V}$ and $\overline{W}$ are compact, and $\mathcal{R}(\overline{S}_1)$ is equicontinuous since $\overline{\mathcal{H}}$ satisfies the equicontinuity condition with $\overline{S}$ and $\{T_i, d_i\}$. □

Let $V_0$ be an open subset of $T$ such that $x \in V_0$ and $\overline{V_0} \subset V$. Since $\overline{V_0} \subset \operatorname{dom} f$ for all $f \in \overline{S}_1$, we can consider the restriction $\overline{S}_1 \times \overline{V_0} \to \widehat{T}$ of the germ map.

**Lemma 3.7.** *The image* $\boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0})$ *is compact in* $\widehat{T}$.

*Proof.* For each $g \in C(\overline{V}, \overline{W})$ and $y \in \overline{V}$, let $\bar{\boldsymbol{\gamma}}(g, y)$ denote the germ of $g$ at $y$, defining a germ map

$$\bar{\boldsymbol{\gamma}} : C(\overline{V}, \overline{W}) \times \overline{V} \to \bar{\boldsymbol{\gamma}}(C(\overline{V}, \overline{W}) \times \overline{V}).$$

Since $\overline{V_0} \subset V$, we get that $\boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0}) = \bar{\boldsymbol{\gamma}}(\mathcal{R}(\overline{S}_1) \times \overline{V_0})$ and the diagram

$$
\text{(5)} \qquad
\begin{array}{ccc}
\overline{S}_1 \times \overline{V_0} & \xrightarrow{\mathcal{R}\times\mathrm{id}} & \mathcal{R}(\overline{S}_1) \times \overline{V_0} \\
{\scriptstyle\boldsymbol{\gamma}}\downarrow & & \downarrow{\scriptstyle\bar{\boldsymbol{\gamma}}} \\
\boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0}) & = & \bar{\boldsymbol{\gamma}}(\mathcal{R}(\overline{S}_1) \times \overline{V_0})
\end{array}
$$

is commutative. Then

$$\bar{\boldsymbol{\gamma}} : \mathcal{R}(\overline{S}_1) \times \overline{V_0} \to \bar{\boldsymbol{\gamma}}(\mathcal{R}(\overline{S}_1) \times \overline{V_0})$$

is continuous because

$$\mathcal{R} \times \mathrm{id} : \overline{S}_1 \times \overline{V_0} \to \mathcal{R}(\overline{S}_1) \times \overline{V_0}$$

is an identification by Lemma 3.4, and

$$\boldsymbol{\gamma} : \overline{S}_1 \times \overline{V_0} \to \boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0})$$

is continuous. Hence $\boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0})$ is compact by Corollary 3.6. □

**Lemma 3.8.** *The image $\boldsymbol{\gamma}(\overline{S}_0 \times V_0)$ is open in $\widehat{T}$.*

*Proof.* This holds because $\overline{S}_0 \times V_0$ is open in $\overline{S}_{\text{c-o}} * T$ and saturated by the fibers of $\boldsymbol{\gamma} : \overline{S}_{\text{c-o}} * T \to \widehat{T}$. □

**Remark 26.** Observe that the proof of Lemma 3.8 does not require $\overline{V_0} \subset V$; it holds for any open $V_0 \subset V$.

**Corollary 3.9.** *$\widehat{T}_U$ is locally compact.*

*Proof.* We have that $\boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0})$ is compact by Lemma 3.7 and contains $\boldsymbol{\gamma}(\overline{S}_0 \times V_0)$, which is an open neighborhood of $\boldsymbol{\gamma}(h, x)$ by Lemma 3.8. Then the result follows because $\boldsymbol{\gamma}(h, x) \in \widehat{T}_U$ is arbitrary. □

**Lemma 3.10.** *The map $\bar{\boldsymbol{\gamma}} : \mathcal{R}(\overline{S}_1) \times \overline{V_0} \to \widehat{T}$ is injective.*

*Proof.* For $f_1, f_2 \in \overline{S}_1$ with $\bar{\boldsymbol{\gamma}}(\mathcal{R}(f_1), y_1) = \bar{\boldsymbol{\gamma}}(\mathcal{R}(f_2), y_2)$, suppose

$$(\mathcal{R}(f_1), y_1), (\mathcal{R}(f_2), y_2) \in \mathcal{R}(\overline{S}_1) \times \overline{V_0},$$

Thus, $y_1 = y_2 =: y$ and $\boldsymbol{\gamma}(f_1, y_1) = \boldsymbol{\gamma}(f_2, y_2)$; i.e., $f_1 = f_2$ on some neighborhood $O$ of $y$ in $\operatorname{dom} f_1 \cap \operatorname{dom} f_2$. Then $f_1(O) \subset \operatorname{dom}(f_2 f_1^{-1})$ and $f_2 f_1^{-1} = \mathrm{id}_T$ on $f_1(O)$. Since $f_2 f_1^{-1} \in \overline{S}$, we get $f_2 f_1^{-1} = \mathrm{id}_T$ on $\operatorname{dom}(f_2 f_1^{-1}) = f_1(\operatorname{dom} f_1 \cap \operatorname{dom} f_2)$ by the strong quasianalyticity of $\overline{S}$. Since $\overline{V} \subset \operatorname{dom} f_1 \cap \operatorname{dom} f_2$, it follows that $f_2 f_1^{-1} = \mathrm{id}_T$ on $f_1(\overline{V})$, and therefore $f_1 = f_2$ on $\overline{V}$; i.e., $\mathcal{R}(f_1) = \mathcal{R}(f_2)$. □

Let $\hat{\pi} := (s, t) : \widehat{T} \to T \times T$, which is continuous.

**Corollary 3.11.** *The restriction $\hat{\pi} : \widehat{T}_U \to U \times U$ is proper.*

*Proof.* Since $U \times U$ can be covered by sets of the form $V_0 \times W$, for $V_0$ and $W$ as above, it is enough to prove that $\hat{\pi}^{-1}(K_1 \times K_2)$ is compact for all compact sets $K_1 \subset V_0$ and $K_2 \subset W$. Then, with the above notation,

$$\hat{\pi}^{-1}(K_1 \times K_2) \subset \boldsymbol{\gamma}(\overline{S}_1 \times K_1) \subset \boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0}),$$

and the result follows from Lemma 3.7. □

**Corollary 3.12.** *The closure of $\widehat{T}_U$ in $\widehat{T}$ is compact.*

*Proof.* Take a relatively compact open subset $U' \subset T$ containing $\overline{U}$. By applying Corollary 3.11 to $U'$, it follows that $\hat{\pi} : \widehat{T}_{U'} \to U' \times U'$ is proper. Therefore $\hat{\pi}^{-1}(\overline{U} \times \overline{U})$ is compact and contains the closure of $\widehat{T}_U$ in $\widehat{T}$. □

**Lemma 3.13.** *$\widehat{T}_U$ is Hausdorff.*

*Proof.* Let $\boldsymbol{\gamma}(h_1, x_1) \neq \boldsymbol{\gamma}(h_2, x_2)$ in $\widehat{T}_U$.

Suppose first that $x_1 \neq x_2$. Since $T$ is Hausdorff, there are disjoint open subsets $V_1$ and $V_2$ such that $x_1 \in V_1$ and $x_2 \in V_2$. Then $\widehat{V}_1 = \widehat{T}_U \cap s^{-1}(V_1)$ and $\widehat{V}_2 = \widehat{T}_U \cap s^{-1}(V_2)$ are disjoint and open in $\widehat{T}_U$, and $\boldsymbol{\gamma}(h_1, x_1) \in \widehat{V}_1$ and $\boldsymbol{\gamma}(h_2, x_2) \in \widehat{V}_2$.

Now, assume that $x_1 = x_2 =: x$ but $h_1(x) \neq h_2(x)$. Take disjoint open subsets $W_1, W_2 \subset U$ such that $h_1(x) \in W_1$ and $h_2(x) \in W_2$. Then $\widehat{W}_1 = \widehat{T}_U \cap t^{-1}(W_1)$ and $\widehat{W}_2 = \widehat{T}_U \cap t^{-1}(W_2)$ are disjoint and open in $\widehat{T}_U$, and $\boldsymbol{\gamma}(h_1, x) \in \widehat{W}_1$ and $\boldsymbol{\gamma}(h_2, x) \in \widehat{W}_2$.

Finally, suppose that $x_1 = x_2 =: x$ and $h_1(x) = h_2(x) =: y$. Then $x \in T_{i_k} \cap U$ and $y \in T_{i_l} \cap U$ for some indices $k$ and $l$. Take open neighborhoods $V$ of $x$ and $W$ of $y$, such that $\overline{V} \subset T_{i_k} \cap U$, $\overline{W} \subset T_{i_l} \cap U$ and $h_1(\overline{V}) \cup h_2(\overline{V}) \subset W$. Define $\overline{S}_0$ and $\overline{S}_1$ by using $V$ and $W$ like in (3) and (4), and take an open subset $V_0 \subset T$ such that $x \in V_0$ and $\overline{V_0} \subset V$, as above. We can assume that $h_1, h_2 \in \overline{S}_1$. Then

$$\bar{\boldsymbol{\gamma}}(\mathcal{R}(h_1), x) = \boldsymbol{\gamma}(h_1, x_1) \neq \boldsymbol{\gamma}(h_2, x_2) = \bar{\boldsymbol{\gamma}}(\mathcal{R}(h_2), x),$$

and therefore $\mathcal{R}(h_1) \neq \mathcal{R}(h_2)$ in $\mathcal{R}(\overline{S}_1)$ by Lemma 3.10. Since $\mathcal{R}(\overline{S}_1)$ is Hausdorff (because it is a subspace of $C_{\text{c-o}}(\overline{V}, \overline{W})$), it follows that there are disjoint open subsets $\mathcal{N}_1, \mathcal{N}_2 \subset \mathcal{R}(\overline{S}_1)$ such that $\mathcal{R}(h_1) \in \mathcal{N}_1$ and $\mathcal{R}(h_2) \in \mathcal{N}_2$. So $\mathcal{R}^{-1}(\mathcal{N}_1)$ and $\mathcal{R}^{-1}(\mathcal{N}_2)$ are disjoint open subsets of $\overline{S}_1$ with $h_1 \in \mathcal{R}^{-1}(\mathcal{N}_1)$ and $h_2 \in \mathcal{R}^{-1}(\mathcal{N}_2)$. Hence $\mathcal{M}_1 = \mathcal{R}^{-1}(\mathcal{N}_1) \cap \overline{S}_0$ and $\mathcal{M}_2 = \mathcal{R}^{-1}(\mathcal{N}_2) \cap \overline{S}_0$ are disjoint and open in $\overline{S}_0$, and therefore they are open in $\overline{S}$. Moreover $\mathcal{M}_1 \times V_0$ and $\mathcal{M}_2 \times V_0$ are saturated by the fibers of $\boldsymbol{\gamma} : \overline{S}_0 \times V_0 \to \boldsymbol{\gamma}(\overline{S}_0 \times V_0)$; in fact, if $(f, z) \in \overline{S}_0 \times V_0$ satisfies $\boldsymbol{\gamma}(f, z) = \boldsymbol{\gamma}(f', z)$ for some $f' \in \mathcal{M}_a$ ($a \in \{1, 2\}$), then

$$\bar{\boldsymbol{\gamma}}(\mathcal{R}(f), z) = \boldsymbol{\gamma}(f, z) = \boldsymbol{\gamma}(f', z) = \bar{\boldsymbol{\gamma}}(\mathcal{R}(f'), z),$$

giving $\mathcal{R}(f) = \mathcal{R}(f') \in \mathcal{N}_a$ by Lemma 3.10. Therefore $f \in \mathcal{R}^{-1}(\mathcal{N}_a) \cap \overline{S}_0 = \mathcal{M}_a$. It follows that $\boldsymbol{\gamma}(\mathcal{M}_1 \times V_0)$ and $\boldsymbol{\gamma}(\mathcal{M}_2 \times V_0)$ are open in $\boldsymbol{\gamma}(\overline{S}_0 \times V_0)$, because the map $\boldsymbol{\gamma} : \overline{S}_0 \times V_0 \to \boldsymbol{\gamma}(\overline{S}_0 \times V_0)$ is an identification as $\overline{S}_0 \times V_0$ is open in $\overline{S}_{\text{c-o}} * T$ and saturated by the fibers of $\boldsymbol{\gamma} : \overline{S}_{\text{c-o}} * T \to \widehat{T}$. Furthermore, by the commutativity of the diagram (5),

$$\begin{aligned} \boldsymbol{\gamma}(\mathcal{M}_1 \times V_0) \cap \boldsymbol{\gamma}(\mathcal{M}_2 \times V_0) &= \bar{\boldsymbol{\gamma}}(\mathcal{N}_1 \times V_0) \cap \bar{\boldsymbol{\gamma}}(\mathcal{N}_2 \times V_0) \\ &= \bar{\boldsymbol{\gamma}}((\mathcal{N}_1 \cap \mathcal{N}_1) \times V_0) = \varnothing, \end{aligned}$$

and $\boldsymbol{\gamma}(h_1, x) \in \boldsymbol{\gamma}(\mathcal{M}_1 \times V_0)$ and $\boldsymbol{\gamma}(h_2, x) \in \boldsymbol{\gamma}(\mathcal{M}_2 \times V_0)$. □

**Corollary 3.14.** *The map* $\bar{\boldsymbol{\gamma}} : \mathcal{R}(\overline{S}_1) \times \overline{V_0} \to \bar{\boldsymbol{\gamma}}(\mathcal{R}(\overline{S}_1) \times \overline{V_0})$ *is a homeomorphism.*

**Lemma 3.15.** $\widehat{T}_U$ *is second countable.*

*Proof.* $\widehat{T}_U$ can be covered by a countable collection of open subsets of the type $\boldsymbol{\gamma}(\overline{S}_0 \times V_0)$ as above. But $\boldsymbol{\gamma}(\overline{S}_0 \times V_0)$ is second countable because it is a subspace of $\boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0}) = \bar{\boldsymbol{\gamma}}(\mathcal{R}(\overline{S}_1) \times \overline{V_0})$, which is homeomorphic to $\mathcal{R}(\overline{S}_1) \times \overline{V_0}$ by Corollary 3.14, and this space is second countable as a subspace of the second countable space $C(\overline{V_0}, \overline{W_0}) \times \overline{V_0}$. □

**Corollary 3.16.** $\widehat{T}_U$ *is Polish.*

*Proof.* This follows from Corollary 3.9, Lemmas 3.13 and 3.15, and [Kechris 1991, Theorem 5.3]. □

**Proposition 3.17.** $\widehat{T}$ *is Polish and locally compact.*

*Proof.* First, let us prove that $\widehat{T}$ is Hausdorff. Take different points $\boldsymbol{\gamma}(g, x)$ and $\boldsymbol{\gamma}(g', x')$ in $\widehat{T}$. Let $O$, $O'$, $P$ and $P'$ be relatively compact open neighborhoods of $x$, $x'$, $g(x)$ and $g(x')$, respectively. Then $U_1 = U \cup O \cup O' \cup P \cup P'$ is a relatively compact open subset of $T$ that meets all $\mathcal{H}$-orbits. By Lemma 3.13, $\widehat{T}_{U_1}$ is a Hausdorff open subset of $\widehat{T}$ that contains $\boldsymbol{\gamma}(g, x)$ and $\boldsymbol{\gamma}(g', x')$. Hence $\boldsymbol{\gamma}(g, x)$ and $\boldsymbol{\gamma}(g', x')$ can be separated in $\widehat{T}_{U_1}$ by disjoint open neighborhoods in $\widehat{T}_{U_1}$, and therefore also in $\widehat{T}$.

Second, let us show that $\widehat{T}$ is locally compact. For $\boldsymbol{\gamma}(g, x) \in \widehat{T}$, let $O$ and $P$ be relatively compact open neighborhoods of $x$ and $g(x)$, respectively. Then $U_1 = U \cup O \cup P$ is a relatively compact open set of $T$ that meets all $\mathcal{H}$-orbits. By Corollary 3.9, it follows that $\widehat{T}_{U_1}$ is a locally compact open neighborhood of $\boldsymbol{\gamma}(g, x)$ in $\widehat{T}$. Hence $\boldsymbol{\gamma}(g, x)$ has a compact neighborhood in $\widehat{T}_{U_1}$, and therefore also in $\widehat{T}$.

Finally, let us show that $\widehat{T}$ is second countable. Since $T$ is second countable (it is Polish) and locally compact, it can be covered by countably many relatively compact open subsets $O_n \subset T$. Then each $U_{n,m} = O_n \cup O_m \cup U$ is a relatively compact open set of $T$ that meets all $\mathcal{H}$-orbits. Hence, by Lemma 3.15, the sets $\widehat{T}_{U_{n,m}}$ are second countable and open in $\widehat{T}$. Moreover these sets form a countable cover of $\widehat{T}$ because, for any $\boldsymbol{\gamma}(g, x) \in \widehat{T}$, we have $x \in O_n$ and $g(x) \in O_m$ for some $n$ and $m$, obtaining $\boldsymbol{\gamma}(g, x) \in \widehat{T}_{U_{n,m}}$. So $\widehat{T}$ is second countable.

Now the result follows by [Kechris 1991, Theorem 5.3]. □

**Proposition 3.18.** *The map* $\hat{\pi} : \widehat{T} \to T \times T$ *is proper.*

*Proof.* Take any compact $K \subset T \times T$ and any relatively compact open $U' \subset T$ meeting all $\mathcal{H}$-orbits and such that $K \subset U' \times U'$. By applying Corollary 3.11 to $U'$, we get that $\hat{\pi}^{-1}(K)$ is compact in $\widehat{T}_{U'}$, and therefore in $\widehat{T}$. □

**3D. *The space $\widehat{T}_0$.*** From now on, assume that $\mathcal{H}$ is minimal, and therefore $\overline{\mathcal{H}}$ has only one orbit, the whole of $T$. Fix a point $x_0 \in U$, and let[8]

$$\widehat{T}_0 = t^{-1}(x_0) = \{\boldsymbol{\gamma}(g,x) \in \widehat{T} \mid g(x) = x_0\}, \quad \widehat{T}_{0,U} = \widehat{T}_0 \cap \widehat{T}_U.$$

Observe that $\widehat{T}_0$ is closed in $\widehat{T}$, whereas $\widehat{T}_{0,U}$ is open in $\widehat{T}_0$. Moreover, we have $\hat{\pi}(\widehat{T}_0) = T \times \{x_0\} \equiv T$ and $\hat{\pi}(\widehat{T}_{0,U}) = U \times \{x_0\} \equiv U$ because $T$ is the unique $\overline{\mathcal{H}}$-orbit; indeed, $\hat{\pi}(\boldsymbol{\gamma}(h,x)) = x$ for each $x \in T$ and any $h \in \overline{S}$ with $x \in \operatorname{dom} h$ and $h(x) = x_0$. Let $\hat{\pi}_0 := s : \widehat{T}_0 \to T$, which is continuous and surjective.

The following two corollaries are direct consequences of Proposition 3.17 (see [Kechris 1991, Theorem 3.11]) and Corollary 3.12.

**Corollary 3.19.** *$\widehat{T}_0$ is Polish and locally compact.*

**Corollary 3.20.** *The closure of $\widehat{T}_{0,U}$ in $\widehat{T}_0$ is compact.*

The following corollary is a direct consequence of Proposition 3.18 because $\hat{\pi}_0 : \widehat{T}_0 \to T$ can be identified with the restriction $\hat{\pi} : \widehat{T}_0 \to T \times \{x_0\} \equiv T$.

**Corollary 3.21.** *The map $\hat{\pi}_0 : \widehat{T}_0 \to T$ is proper.*

**Proposition 3.22.** *The fibers of $\hat{\pi}_0 : \widehat{T}_0 \to T$ are homeomorphic to each other.*

*Proof.* For each $x \in T$, there is some $f \in \overline{S}$ with $f(x) = x_0$. Then the mapping $\boldsymbol{\gamma}(g,x) \mapsto \boldsymbol{\gamma}(gf^{-1}, x_0)$ defines a homeomorphism $\hat{\pi}_0^{-1}(x) \to \hat{\pi}_0^{-1}(x_0)$ whose inverse is given by $\boldsymbol{\gamma}(g_0, x_0) \mapsto \boldsymbol{\gamma}(g_0 f, x)$. □

**Question 3.23.** When is $\hat{\pi}_0$ a fiber bundle?

**3E. *The pseudogroup $\widehat{\mathcal{H}}_0$.*** For $h \in S$, define

$$\hat{h} : \hat{\pi}_0^{-1}(\operatorname{dom} h) \to \hat{\pi}_0^{-1}(\operatorname{im} h), \quad \hat{h}(\boldsymbol{\gamma}(g,x)) = \boldsymbol{\gamma}(gh^{-1}, h(x)),$$

for $g \in S$ and $x \in \operatorname{dom} g \cap \operatorname{dom} h$ with $g(x) = x_0$. The following two results are elementary.

**Lemma 3.24.** *For any $h \in S$, we have $\hat{\pi}_0(\operatorname{dom} \hat{h}) = \operatorname{dom} h$ and $\hat{\pi}_0(\operatorname{im} \hat{h}) = \operatorname{im} h$, and the following diagram is commutative*:

$$\begin{tikzcd} \operatorname{dom} \hat{h} \arrow[r, "\hat{h}"] \arrow[d, "\hat{\pi}_0"'] & \operatorname{im} \hat{h} \arrow[d, "\hat{\pi}_0"] \\ \operatorname{dom} h \arrow[r, "h"] & \operatorname{im} h \end{tikzcd}$$

**Lemma 3.25.** *If $O \subset T$ is open with $\operatorname{id}_O \in S$, then $\widehat{\operatorname{id}_O} = \operatorname{id}_{\hat{\pi}_0^{-1}(O)}$.*

**Lemma 3.26.** *For $h, h' \in S$, we have $\widehat{h'h} = \widehat{h'}\hat{h}$.*

[8]The definition $\widehat{T}_0 = s^{-1}(x_0)$ would be valid too, of course, but it seems that the proofs in Sections 3D and 3E have a simpler notation with the choice $\widehat{T}_0 = t^{-1}(x_0)$.

*Proof.* By Lemma 3.24, we have

$$\begin{aligned}\operatorname{dom}(\widehat{h'}\hat{h}) &= \hat{h}^{-1}(\operatorname{dom}\widehat{h'} \cap \operatorname{im}\hat{h}) = \hat{h}^{-1}(\hat{\pi}_0^{-1}(\operatorname{dom}h' \cap \operatorname{im}h)) \\ &= \hat{\pi}_0^{-1}(h^{-1}(\operatorname{dom}h' \cap \operatorname{im}h)) = \hat{\pi}_0^{-1}(\operatorname{dom}(h'h)) = \operatorname{dom}\widehat{h'h}.\end{aligned}$$

Now let $\boldsymbol{\gamma}(g,x) \in \operatorname{dom}(\widehat{h'}\hat{h}) = \operatorname{dom}\widehat{h'h}$; therefore $g \in \overline{S}$, $x \in \operatorname{dom}g \cap \operatorname{dom}h$, $h(x) \in \operatorname{dom}h'$ and $g(x) = x_0$. Then

$$\begin{aligned}\widehat{h'h}(\boldsymbol{\gamma}(g,x)) &= \boldsymbol{\gamma}(g(h'h)^{-1}, h'h(x)) = \boldsymbol{\gamma}(gh^{-1}(h')^{-1}, h'h(x)) \\ &= \widehat{h'}\big(\boldsymbol{\gamma}(gh^{-1}, h(x))\big) = \widehat{h'}\hat{h}(\boldsymbol{\gamma}(g,x)).\end{aligned}$$ □

The following is a direct consequence of Lemmas 3.25 and 3.26.

**Corollary 3.27.** *For $h \in S$, the map $\hat{h}$ is bijective with $\hat{h}^{-1} = \widehat{h^{-1}}$.*

**Lemma 3.28.** *The map $\hat{h}$ is a homeomorphism for all $h \in S$.*

*Proof.* By Corollary 3.27, it is enough to prove that $\hat{h}$ is continuous, which holds because it can be expressed as the composition of continuous maps

$$\begin{aligned}\hat{\pi}_0^{-1}(\operatorname{dom}h) &\xrightarrow{(\mathrm{id},\mathrm{const},h\hat{\pi}_0)} \hat{\pi}_0^{-1}(\operatorname{dom}h) \times \{h^{-1}\} \times \operatorname{im}h \\ &\xrightarrow{\mathrm{id}\times\boldsymbol{\gamma}} \hat{\pi}_0^{-1}(\operatorname{dom}h) \times \boldsymbol{\gamma}(\{h^{-1}\} \times \operatorname{im}h) \\ &\xrightarrow{\text{product}} \hat{\pi}_0^{-1}(\operatorname{im}h).\end{aligned}$$

This can be checked on elements:

$$\begin{aligned}\boldsymbol{\gamma}(g,x) &\mapsto (\boldsymbol{\gamma}(g,x), h^{-1}, h(x)) \\ &\mapsto \big(\boldsymbol{\gamma}(g,x), \boldsymbol{\gamma}(h^{-1}, h(x))\big) \\ &\mapsto \boldsymbol{\gamma}(gh^{-1}, h(x)) = \hat{h}(\boldsymbol{\gamma}(g,x)).\end{aligned}$$ □

Set $\widehat{S}_0 = \{\hat{h} \mid h \in S\}$, and let $\widehat{\mathcal{H}}_0$ be the pseudogroup on $\widehat{T}_0$ generated by $\widehat{S}_0$. Lemmas 3.26 and 3.28 and Corollary 3.27 give the following.

**Corollary 3.29.** *$\widehat{S}_0$ is a pseudo∗group on $\widehat{T}_0$.*

**Lemma 3.30.** *$\widehat{T}_{0,U}$ meets all orbits of $\widehat{\mathcal{H}}_0$.*

*Proof.* Let $\boldsymbol{\gamma}(g,x) \in \widehat{T}_0$ with $g \in \overline{S}$; then $x \in \operatorname{dom}g$ and $g(x) = x_0$. Since $U$ meets all orbits of $\mathcal{H}$, there is some $h \in S$ such that $x \in \operatorname{dom}h$ and $h(x) \in U$. Then $\boldsymbol{\gamma}(g,x) \in \operatorname{dom}\hat{h}$ and $\hat{h}(\boldsymbol{\gamma}(g,x)) = \boldsymbol{\gamma}(gh^{-1}, h(x))$ satisfies

$$\hat{\pi}_0\big(\hat{h}(\boldsymbol{\gamma}(g,x))\big) = \hat{\pi}_0\big(\boldsymbol{\gamma}(gh^{-1}, h(x))\big) = h(x) \in U.$$

Hence $\hat{h}(\boldsymbol{\gamma}(g,x)) \in \widehat{T}_{0,U}$ as desired. □

**Lemma 3.31.** *The map $S_{\text{c-o}} \to \widehat{S}_{0,\text{c-o}}$, $h \mapsto \hat{h}$, is a homeomorphism.*

*Proof.* If $\widehat{h_1} = \widehat{h_2}$ for some $h_1, h_2 \in S$, then $h_1 = h_2$ by Lemma 3.24. So the stated map is injective, and therefore it is bijective by the definition of $\widehat{S}_0$.

Take a subbasic open set of $S_{\text{c-o}}$, which is of the form $S \cap \mathcal{N}(K, O)$ for some compact $K$ and open $O$ in $T$. The set $\hat{\pi}_0^{-1}(K)$ is compact by Corollary 3.21, and $\hat{\pi}_0^{-1}(O)$ is open. Then the map of the statement is open because

$$\{\hat{h} \mid h \in \mathcal{N}(K, O) \cap S\} = \widehat{\mathcal{N}}\big(\hat{\pi}_0^{-1}(K), \hat{\pi}_0^{-1}(O)\big) \cap \widehat{S}_0$$

by Lemma 3.24, which is open in $\widehat{S}_{0,\text{c-o}}$.

To prove its continuity, let us first show that its restriction to $S_U = S \cap \mathcal{H}|_U$ is continuous. Fix $h_0 \in S_U$, and take relatively compact open subsets

$$V, V_0, W, V', V_0', W' \subset U,$$

and indices $k$ and $k'$ such that

$$\overline{V_0} \subset V, \quad \overline{V} \subset T_{i_k} \cap \operatorname{dom} h_0, \tag{6}$$

$$\overline{V_0'} \subset V', \quad \overline{V'} \subset T_{i_{k'}} \cap \operatorname{im} h_0, \tag{7}$$

$$\overline{W} \subset W', \quad \overline{W'} \subset T_{i_{k_0}}, \tag{8}$$

$$h_0^{-1}(\overline{V'}) \subset V, \tag{9}$$

$$h_0(\overline{V_0}) \subset V'. \tag{10}$$

Let $\overline{S}_0$ and $\overline{S}_1$ (respectively, $\overline{S}_0'$ and $\overline{S}_1'$) be defined like in (3) and (4), by using $V$ and $W$ (respectively, $V'$ and $W'$). Then $\widehat{K} = \boldsymbol{\gamma}(\overline{S}_1 \times \overline{V_0})$ is compact in $\widehat{T}$ by Lemma 3.7, and $\widehat{O} = \boldsymbol{\gamma}(\overline{S}_0' \times V')$ is open in $\widehat{T}$ by Lemma 3.8 and Remark 26. Thus $\widehat{K}_0 = \widehat{K} \cap \widehat{T}_0$ is compact and $\widehat{O}_0 = \widehat{O} \cap \widehat{T}_0$ is open in $\widehat{T}_0$. So $\widehat{\mathcal{N}}(\widehat{K}_0, \widehat{O}_0) \cap \widehat{S}_0$ is a subbasic open set of $\widehat{S}_{0,\text{c-o}}$.

**Claim 1.** $\widehat{h_0} \in \widehat{\mathcal{N}}(\widehat{K}_0, \widehat{O}_0)$.

Let $\boldsymbol{\gamma}(g, x) \in \widehat{K}_0$; thus $g \in \overline{S}_1$, $x \in \overline{V}_0 \cap \operatorname{dom} g$ and $g(x) = x_0$. The condition $g \in \overline{S}_1$ means that $g \in \overline{S}$, $\overline{V} \subset \operatorname{dom} g$ and $g(\overline{V}) \subset \overline{W}$. By (7)–(9), it follows that $\overline{V'} \subset \operatorname{dom} gh_0^{-1}$ and

$$gh_0^{-1}(\overline{V'}) \subset g(\overline{V}) \subset \overline{W} \subset W'.$$

Hence $gh_0^{-1} \in \overline{S}_0'$, obtaining that

$$\widehat{h_0}(\boldsymbol{\gamma}(g, x)) = \boldsymbol{\gamma}(gh_0^{-1}, h_0(x)) \in \widehat{O},$$

which completes the proof of Claim 1.

**Claim 2.** *The sets $\widehat{\mathcal{N}}(\widehat{K}_0, \widehat{O}_0) \cap \widehat{S}_0$, constructed as above, form a local subbasis of $\widehat{S}_{0,\text{c-o}}$ at $\widehat{h_0}$.*

This assertion follows by Claim 1 and because the sets of the type $\widehat{O}_0$ form a basis of the topology of $\operatorname{im}\widehat{h_0}$, and any compact subset of $\operatorname{dom}\widehat{h_0}$ is contained in a finite union of sets of the type of $\widehat{K}_0$.

The sets
$$\mathcal{N} = \mathcal{N}(\overline{V_0}, V') \cap \mathcal{N}(\overline{V'}, V)^{-1} \cap S_U$$
are open neighborhoods of $h_0$ by (9), (10), and Propositions 2.6 and 3.1.

**Claim 3.** *We have $\hat{h} \in \widehat{\mathcal{N}}(\widehat{K}_0, \widehat{O}_0)$ for all $h \in \mathcal{N}$.*

Given $h \in \mathcal{N}$, we have $\overline{V'} \subset \operatorname{im} h$ and $h^{-1}(\overline{V'}) \subset V$. Let $\boldsymbol{\gamma}(g, x) \in \widehat{K}_0$; thus $x \in \overline{V_0} \cap \operatorname{dom} g$, $g(x) = x_0$, and we can assume that $g \in \overline{S}_1$, which means that $g \in \overline{S}$, $\overline{V} \subset \operatorname{dom} g$ and $g(\overline{V}) \subset \overline{W}$. Then $\overline{V'} \subset \operatorname{dom}(gh^{-1})$, $gh^{-1}(\overline{V'}) \subset \overline{W} \subset W'$ and $h(x) \in h(\overline{V_0}) \subset V'$. Therefore
$$\hat{h}(\boldsymbol{\gamma}(g, x)) = \boldsymbol{\gamma}(gh^{-1}, h(x)) \in \boldsymbol{\gamma}(\overline{S}'_0 \times V') \cap \widehat{T}_0 = \widehat{O}_0,$$
proving Claim 3.

Claims 2 and 3 show that the map $S_{U,\text{c-o}} \to \widehat{S}_{0,\text{c-o}}$, $h \mapsto \hat{h}$, is continuous at $h_0$. Now, let us prove that the whole map $S_{\text{c-o}} \to \widehat{S}_{0,\text{c-o}}$, $h \mapsto \hat{h}$, is continuous. Since the sets $\mathcal{N}(\widehat{K}, \widehat{O}) \cap \widehat{S}_0$, for small enough compact subsets $\widehat{K} \subset \widehat{T}_0$ and small enough open subsets $\widehat{O} \subset \widehat{T}_0$, form a subbasis of $\widehat{S}_{0,\text{c-o}}$, it is enough to prove that the inverse image of these subbasic sets are open in $S_{\text{c-o}}$. We can assume that $\widehat{K}, \widehat{O} \subset \hat{\pi}_0^{-1}(U')$ for some relatively compact open subset $U' \subset T$ that meets all $\mathcal{H}$-orbits. Consider the inclusion map $\iota : U' \hookrightarrow T$, and the paro map $\phi : T \rightarrowtail U'$ with $\operatorname{dom}\phi = U'$, where it is the identity map. According to Proposition 2.3, we get a continuous map $\phi_*\iota^* : \operatorname{Paro}_{\text{c-o}}(T, T) \to \operatorname{Paro}_{\text{c-o}}(U', U')$, which restricts to a continuous map $\phi_*\iota^* : S_{\text{c-o}} \to S_{U',\text{c-o}}$. Observe that $\phi_*\iota^*(h)$ is the restriction $h : U' \cap h^{-1}(U') \to h(U') \cap U'$ for each $h \in S$. Hence, since $\widehat{K}, \widehat{O} \subset \hat{\pi}_0^{-1}(U')$, it follows from Lemma 3.24 that $\mathcal{N}(\widehat{K}, \widehat{O}) \cap \widehat{S}_0$ has the same inverse image by the map $S_{\text{c-o}} \to \widehat{S}_{0,\text{c-o}}$, $h \mapsto \hat{h}$, and by the composition
$$S_{\text{c-o}} \xrightarrow{\phi_*\iota^*} S_{U',\text{c-o}} \longrightarrow \widehat{S}_{0,\text{c-o}},$$
where the second map is given by $h \mapsto \hat{h}$. This composition is continuous by the above case applied to $U'$, and therefore the inverse image of $\mathcal{N}(\widehat{K}, \widehat{O}) \cap \widehat{S}_0$ by $S_{\text{c-o}} \to \widehat{S}_{0,\text{c-o}}$, $h \mapsto \hat{h}$, is open in $S_{\text{c-o}}$. □

Since the compact generation of $\mathcal{H}$ is satisfied with the relatively compact open set $U$, there is a symmetric finite set $\{f_1, \dots, f_m\}$ generating $\mathcal{H}|_U$, which can be chosen in $S$, such that each $f_a$ has an extension $\tilde{f}_a$ with $\overline{\operatorname{dom} f_a} \subset \operatorname{dom} \tilde{f}_a$. We can also assume that $\tilde{f}_a \in S$. Let $\widehat{\mathcal{H}}_{0,U} = \widehat{\mathcal{H}}|_{\widehat{T}_{0,U}}$. Obviously, each $\widehat{\tilde{f}_a}$ is an extension of $\widehat{f_a}$. Moreover,
$$\overline{\operatorname{dom}\widehat{f_a}} = \overline{\hat{\pi}_0^{-1}(\operatorname{dom} f_a)} \subset \hat{\pi}_0^{-1}(\overline{\operatorname{dom} f_a}) \subset \hat{\pi}_0^{-1}(\operatorname{dom} \tilde{f}_a) = \operatorname{dom}\widehat{\tilde{f}_a}.$$

**Lemma 3.32.** *The maps* $\widehat{f_a}$ $(a \in \{1, \dots, m\})$ *generate* $\widehat{\mathscr{H}}_{0,U}$.

*Proof.* $\widehat{\mathscr{H}}_{0,U}$ is generated by the maps of the form $\hat{h}$ with $h \in S_U$, and any such $\hat{h}$ can be written as a composition of maps $\widehat{f_a}$ around any $\boldsymbol{\gamma}(g, x) \in \operatorname{dom} \hat{h} = \hat{\pi}_0^{-1}(\operatorname{dom} h)$ by Lemma 3.26. □

**Corollary 3.33.** $\widehat{\mathscr{H}}_0$ *is compactly generated.*

*Proof.* We saw that $\widehat{T}_{0,U}$ is relatively compact in $\widehat{T}_0$ (Corollary 3.20) and meets all $\widehat{\mathscr{H}}_0$-orbits (Lemma 3.30), the maps $\widehat{f_a}$ generate $\widehat{\mathscr{H}}_{0,U}$ (Lemma 3.32), and each $\widehat{\bar{f}_a}$ is an extension of each $\widehat{f_a}$ with $\overline{\operatorname{dom}\widehat{f_a}} \subset \operatorname{dom} \widehat{\bar{f}_a}$. □

Recall that the sets $T_{i_k}$ form a finite covering of $\overline{U}$ by open sets of $T$. Fix some index $k_0$ such that $x_0 \in T_{i_{k_0}}$. Let $\{W_k\}$ be a shrinking of $\{T_{i_k}\}$ as cover of $\overline{U}$ by open subsets of $T$; i.e., $\{W_k\}$ is a cover of $\overline{U}$ by open subsets of $T$ and $\overline{W_k} \subset T_{i_k}$ for all $k$. By applying Proposition 2.32 several times, we get finite covers, $\{V_a\}$ and $\{V'_u\}$, of $\overline{U}$ by open subsets of $T$, and shrinkings, $\{W_{0,k}\}$ of $\{W_k\}$ and $\{V_{0,a}\}$ of $\{V_a\}$, as covers of $\overline{U}$ by open subsets of $T$, such that the following properties hold:

- For all $h \in \mathscr{H}$ and $x \in \operatorname{dom} h \cap U \cap V_a \cap W_{0,k}$ with $h(x) \in U \cap W_{0,l}$, there is some $\tilde{h} \in S$ such that
$$\overline{V_a} \subset \operatorname{dom} \tilde{h} \cap W_k, \quad \boldsymbol{\gamma}(\tilde{h}, x) = \boldsymbol{\gamma}(h, x), \quad \tilde{h}(\overline{V_a}) \subset W_l.$$
- For all $h \in \mathscr{H}$ and $x \in \operatorname{dom} h \cap U \cap V'_u \cap V_{0,a}$ with $h(x) \in U \cap V_{0,b}$, there is some $\tilde{h} \in S$ such that
$$\overline{V'_u} \subset \operatorname{dom} \tilde{h} \cap V_a, \quad \boldsymbol{\gamma}(\tilde{h}, x) = \boldsymbol{\gamma}(h, x), \quad \tilde{h}(\overline{V'_u}) \subset V_b.$$

By the definition of $\overline{\mathscr{H}}$ and $\overline{S}$, it follows that these properties also hold for all $h \in \overline{\mathscr{H}}$ with $\tilde{h} \in \overline{S}$. Let $\{V'_{0,u}\}$ be a shrinking of $\{V'_u\}$ as a cover of $\overline{U}$ by open subsets of $T$. We have $x_0 \in W_{0,k_0} \cap V_{0,a_0} \cap V'_{0,u_0}$ for some indices $k_0$, $a_0$ and $u_0$. For each $a$, let $\overline{S}_{0,a}, \overline{S}_{1,a} \subset \overline{S}$ be defined like $\overline{S}_0$ and $\overline{S}_1$ in (3) and (4) by using $V_a$ and $W_{k_0}$ instead of $V$ and $W$. Take an index $u$ such that $\overline{V'_u} \subset V_a$. The sets $V_{0,a} \cap V'_{0,u}$, defined in this way, form a cover of $\overline{U}$, so that the sets $\widehat{T}_{a,u} = \boldsymbol{\gamma}(\overline{S}_{0,a} \times (V_{0,a} \cap V'_{0,u}))$ form a cover of $\overline{\widehat{T}_U}$ by open subsets of $\widehat{T}$ (Lemma 3.8), and thus the sets $\widehat{T}_{0,a,u} = \widehat{T}_{a,u} \cap \widehat{T}_0$ form a cover of $\overline{\widehat{T}_{0,U}}$ by open subsets of $\widehat{T}_0$. Let $\widehat{T}_{0,U,a,u} = \widehat{T}_{0,U} \cap \widehat{T}_{a,u}$. Like in Section 3C, let $\bar{\boldsymbol{\gamma}}$ denote the germ map defined on $C(\overline{V_a}, \overline{W_{k_0}}) \times \overline{V_a}$, and let $\mathscr{R}_a : \overline{S}_{1,a} \to C(\overline{V_a}, \overline{W_{k_0}})$ be the restriction map $f \mapsto f|_{\overline{V_a}}$. Then

$$\bar{\boldsymbol{\gamma}} : \mathscr{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}} \to \bar{\boldsymbol{\gamma}}\big(\mathscr{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}}\big) \tag{11}$$

is a homeomorphism by Corollary 3.14. Since $\overline{V_a}$ is compact, the compact-open topology on $\mathscr{R}_a(\overline{S}_{1,a})$ equals the topology induced by the supremum metric $d_a$

on $C(\overline{V_a}, \overline{W_{k_0}})$, defined with the metric $d_{i_{k_0}}$ on $T_{i_{k_0}}$. Take some index $k$ such that $\overline{V_a} \subset W_k$. Then the topology of

$$\mathcal{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}}$$

is induced by the metric $d_{a,u,k}$ given by

$$d_{a,u,k}((g, y), (g', y')) = d_{i_k}(y, y') + d_a(g, g')$$

(recall that $\overline{W_k} \subset T_{i_k}$). Let $\hat{d}_{a,u,k}$ be the metric on $\bar{\boldsymbol{\gamma}}(\mathcal{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}})$ that corresponds to $d_{a,u,k}$ by the homeomorphism (11); it induces the topology of $\bar{\boldsymbol{\gamma}}(\mathcal{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}})$. By the commutativity of (5),

$$\bar{\boldsymbol{\gamma}}\big(\mathcal{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}}\big) = \boldsymbol{\gamma}\big(\mathcal{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}}\big),$$

which is contained in $\widehat{T}$. Then the restriction $\hat{d}_{0,a,u,k}$ of $\hat{d}_{a,u,k}$ to

$$\bar{\boldsymbol{\gamma}}\big(\mathcal{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}}\big) \cap \widehat{T}_0$$

induces the topology of this space. Moreover, by the proof of Corollary 3.9, we get

$$\widehat{T}_{a,u} \subset \bar{\boldsymbol{\gamma}}\big(\mathcal{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}}\big),$$

and therefore

$$\widehat{T}_{0,a,u} \subset \bar{\boldsymbol{\gamma}}\big(\mathcal{R}_a(\overline{S}_{1,a}) \times \overline{V_{0,a} \cap V'_{0,u}}\big) \cap \widehat{T}_0.$$

For any index $v$, define $\overline{S}'_{0,v}$ and $\overline{S}'_{1,v}$ like $\overline{S}_0$ and $\overline{S}_1$ in (3) and (4) by using $V'_v$ and $W_{k_0}$ instead of $V$ and $W$. Let $\mathcal{R}'_v : \overline{S}'_{1,v} \to C(\overline{V'_v}, \overline{W_{k_0}})$ denote the restriction map. Again, the compact-open topology on $\mathcal{R}'_v(\overline{S}'_{1,v})$ equals the topology induced by the supremum metric $d'_v$ on $C(\overline{V'_v}, \overline{W_{k_0}})$, defined with the metric $d_{i_{k_0}}$ on $T_{i_{k_0}}$ (recall that $\overline{W_{k_0}} \subset T_{i_{k_0}}$). Take indices $b$ and $l$ such that $\overline{V'_v} \subset V_b$ and $\overline{V_b} \subset W_l$. Then we can consider the restriction map

$$\mathcal{R}^v_b : C(\overline{V_b}, \overline{W_{k_0}}) \to C(\overline{V'_v}, \overline{W_{k_0}}).$$

Its restriction $\mathcal{R}^v_b : \mathcal{R}_b(\overline{S}_{1,b}) \to \mathcal{R}'_v(\overline{S}'_{1,v})$ is injective by Remark 25, and surjective by Remark 24. So $\mathcal{R}^v_b : \mathcal{R}_b(\overline{S}_{1,b}) \to \mathcal{R}'_v(\overline{S}'_{1,v})$ is a continuous bijection between compact Hausdorff spaces, giving that it is a homeomorphism. Then, by compactness, it is a uniform homeomorphism with respect to the supremum metrics $d_b$ and $d'_v$. Since $b$ and $v$ run in finite families of indices, there is a mapping $\epsilon \mapsto \delta_1(\epsilon) > 0$ ($\epsilon > 0$) such that

$$d'_v(\mathcal{R}^v_b\mathcal{R}_b(f), \mathcal{R}^v_b\mathcal{R}_b(f')) < \delta_1(\epsilon) \Longrightarrow d_b(\mathcal{R}_b(f), \mathcal{R}_b(f')) < \epsilon \tag{12}$$

for all indices $v$ and $b$, and maps $f, f' \in \overline{S}_{1,b}$.

**Lemma 3.34.** *$\widehat{\mathcal{H}}_{0,U}$ satisfies the equicontinuity condition with $\widehat{S}_{0,U} = \widehat{S}_0 \cap \widehat{\mathcal{H}}_{0,U}$ and the quasilocal metric represented by the family $\{\widehat{T}_{0,U,a,u}, \hat{d}_{0,a,u,k}\}$.*

*Proof.* Let $h \in S$, and

$$\boldsymbol{\gamma}(g, y), \boldsymbol{\gamma}(g', y') \in \widehat{T}_{0,U,a,u} \cap \hat{h}^{-1}(\widehat{T}_{0,U,b,v}),$$

where $g, g' \in \overline{S}_{0,a}$ and $y, y' \in V_{0,a} \cap V'_{0,u}$ with $g(y) = g(y') = x_0$. Take some indices $k$ and $l$ such that $\overline{V_a} \subset W_k$ and $\overline{V_b} \subset W_l$ (recall that $\overline{W_k} \subset T_{i_k}$ and $\overline{W_l} \subset T_{i_l}$). By Remark 24, we can assume that $\operatorname{dom} h = T_{i_k}$. Then

$$\hat{h}(\boldsymbol{\gamma}(g, y)) = \boldsymbol{\gamma}(gh^{-1}, h(y)), \quad \hat{h}(\boldsymbol{\gamma}(g', y')) = \boldsymbol{\gamma}(g'h^{-1}, h(y'))$$

both belong to $\widehat{T}_{0,U,b,v}$, which means that $h(y), h(y') \in V_{0,b} \cap V'_{0,v}$ and there are $f, f' \in \overline{S}_{0,b}$ such that

$$\boldsymbol{\gamma}(f, h(y)) = \boldsymbol{\gamma}(gh^{-1}, h(y)), \quad \boldsymbol{\gamma}(f', h(y')) = \boldsymbol{\gamma}(g'h^{-1}, h(y')). \tag{13}$$

In particular, $\overline{V_b} \subset \operatorname{dom} f \cap \operatorname{dom} f'$. In fact, we can assume $\operatorname{dom} f = \operatorname{dom} f' = T_{i_l}$ by Remark 24. Observe that the image of $h$ may not be included in $T_{i_l}$, and the images of $f$, $f'$, $g$ and $g'$ may not be included in $T_{i_{k_0}}$.

**Claim 1.** $\overline{V'_v} \subset \operatorname{im} h$ *and* $h^{-1}(\overline{V'_v}) \subset V_a$.

By the assumptions on $\{V'_w\}$, since

$$h(y) \in U \cap V'_v \cap V_{0,b} \cap \operatorname{dom} h^{-1}, \quad h^{-1}h(y) = y \in U \cap V'_u \cap V_{0,a},$$

there is some $\widetilde{h^{-1}} \in S$ such that

$$\overline{V'_v} \subset \operatorname{dom} \widetilde{h^{-1}} \cap V_b, \quad \widetilde{h^{-1}}(\overline{V'_v}) \subset V_a, \quad \boldsymbol{\gamma}(\widetilde{h^{-1}}, h(y)) = \boldsymbol{\gamma}(h^{-1}, h(y));$$

indeed, we can suppose that $\operatorname{dom} \widetilde{h^{-1}} = T_{i_{k_0}}$ by Remark 24. Then

$$\widetilde{h^{-1}}(\overline{V'_v}) \subset V_a \subset T_{i_k} = \operatorname{dom} h,$$

obtaining $\overline{V'_v} \subset \operatorname{dom}(h\widetilde{h^{-1}})$. Moreover

$$\boldsymbol{\gamma}(h\widetilde{h^{-1}}, h(y)) = \boldsymbol{\gamma}(\mathrm{id}_T, h(y)).$$

Therefore $h\widetilde{h^{-1}} = \mathrm{id}_{\operatorname{dom}(h\widetilde{h^{-1}})}$ because $h\widetilde{h^{-1}} \in S$ since $h, \widetilde{h^{-1}} \in S$. So $h\widetilde{h^{-1}} = \mathrm{id}_T$ on some neighborhood of $\overline{V'_v}$, and therefore $\overline{V'_v} \subset \operatorname{im} h$ and $h^{-1} = \widetilde{h^{-1}}$ on $\overline{V'_v}$. Thus $h^{-1}(\overline{V'_v}) = \widetilde{h^{-1}}(\overline{V'_v}) \subset V_a$, which shows Claim 1.

By Claim 1 and since $\overline{V_a} \subset \operatorname{dom} g \cap \operatorname{dom} g'$ because $g, g' \in \overline{S}_{0,a}$, we get

$$\overline{V'_v} \subset \operatorname{dom}(gh^{-1}) \cap \operatorname{dom}(g'h^{-1}). \tag{14}$$

Since $f, f' \in \overline{S}_{0,b}$, we have $\overline{V_b} \subset \operatorname{dom} f \cap \operatorname{dom} f'$ and $f(\overline{V_b}) \cup f'(\overline{V_b}) \subset W_{k_0}$. On the other hand, it follows from (13) that $fh(y) = f'h(y') = x_0$ and

$$\boldsymbol{\gamma}(gh^{-1}f^{-1}, x_0) = \boldsymbol{\gamma}(g'h^{-1}f'^{-1}, x_0) = \boldsymbol{\gamma}(\mathrm{id}_T, x_0).$$

Moreover,

$$f(\overline{V'_v}) \subset \operatorname{dom}(gh^{-1}f^{-1}), \quad f'(\overline{V'_v}) \subset \operatorname{dom}(g'h^{-1}f'^{-1})$$

by (14). So, by Remark 25, $gh^{-1}f^{-1} = \mathrm{id}_T$ on some neighborhood of $f(\overline{V'_v})$, and $g'h^{-1}f'^{-1} = \mathrm{id}_T$ on some neighborhood of $f'(\overline{V'_v})$. Thus $gh^{-1} = f$ and $g'h^{-1} = f'$ on some neighborhood of $\overline{V'_v}$; in particular,

$$\mathcal{R}^v_b \mathcal{R}_b(f) = gh^{-1}|_{\overline{V'_v}}, \quad \mathcal{R}^v_b \mathcal{R}_b(f') = g'h^{-1}|_{\overline{V'_v}}.$$

Consider the mappings $\epsilon \mapsto \delta(\epsilon) > 0$ and $\epsilon \mapsto \delta_1(\epsilon) > 0$ satisfying Remark 22 and (12). Then, for each $\epsilon > 0$, define

$$\hat{\delta}(\epsilon) = \min\{\delta(\epsilon/2), \delta_1(\epsilon/2)\}.$$

Given any $\epsilon > 0$, suppose that

$$\hat{d}_{0,a,u,k}(\boldsymbol{\gamma}(g, y), \boldsymbol{\gamma}(g', y')) < \hat{\delta}(\epsilon).$$

This means that

$$d_{a,u,k}((\mathcal{R}_a(g), y), (\mathcal{R}_a(g'), y')) < \hat{\delta}(\epsilon),$$

or, equivalently,

$$d_{i_k}(y, y') + \sup_{x \in \overline{V_a}} d_{i_{k_0}}(g(x), g'(x)) < \hat{\delta}(\epsilon).$$

Therefore

$$d_{i_k}(y, y') < \delta(\epsilon/2), \tag{15}$$

$$\sup_{x \in \overline{V_a}} d_{i_{k_0}}(g(x), g'(x)) < \delta_1(\epsilon/2). \tag{16}$$

From (15) and Remark 22, it follows that

$$d_{i_l}(h(y), h(y')) < \epsilon/2 \tag{17}$$

since $h \in S \subset \overline{S}$ and $y, y' \in T_{i_k} \cap h^{-1}(T_{i_l} \cap \operatorname{im} h)$. On the other hand, by Claim 1 and (16), we get

$$\begin{aligned} d'_v\big(\mathcal{R}^v_b \mathcal{R}_b(f), \mathcal{R}^v_b \mathcal{R}_b(f')\big) & \\ &= \sup_{z \in \overline{V'_v}} d_{i_{k_0}}(gh^{-1}(z), g'h^{-1}(z)) = \sup_{x \in h^{-1}(\overline{V'_v})} d_{i_{k_0}}(g(x), g'(x)) \\ &\le \sup_{x \in \overline{V_a}} d_{i_{k_0}}(g(x), g'(x)) = d_a(\mathcal{R}_a(g), \mathcal{R}_a(g')) < \delta_1(\epsilon/2). \end{aligned}$$

So, by (12),

$$d_b(\mathcal{R}_b(f), \mathcal{R}_b(f')) < \epsilon/2. \tag{18}$$

From (17) and (18), we get

$$\begin{aligned}\hat{d}_{0,b,v,l}\big(\hat{h}(\boldsymbol{\gamma}(g,y)),\hat{h}(\boldsymbol{\gamma}(g',y'))\big) &= \hat{d}_{0,b,v,l}\big(\boldsymbol{\gamma}(f,h(y)),\boldsymbol{\gamma}(f',h(y'))\big)\\ &= d_{b,v,l}\big((\mathcal{R}_b(f),h(y)),(\mathcal{R}_b(f'),h(y'))\big)\\ &= d_{i_l}(h(y),h(y')) + d_b(\mathcal{R}_b(f),\mathcal{R}_b(f')) < \epsilon. \quad \square\end{aligned}$$

**Corollary 3.35.** *$\widehat{\mathcal{H}}_0$ is equicontinuous.*

*Proof.* $\widehat{\mathcal{H}}_0$ is equivalent to $\widehat{\mathcal{H}}_{0,U}$ by Lemma 3.30. Thus the result follows from Lemma 3.34 because equicontinuity is preserved by equivalences. □

**Lemma 3.36.** *$\widehat{\mathcal{H}}_0$ is minimal.*

*Proof.* By Lemma 3.30, it is enough to prove that $\widehat{\mathcal{H}}_{0,U}$ is minimal. Let the germs $\boldsymbol{\gamma}(g,y), \boldsymbol{\gamma}(g',y')$ be in $\widehat{T}_{0,U}$ with $g,g' \in \overline{S}$, $y \in \operatorname{dom} g \cap U$, $y' \in \operatorname{dom} g' \cap U$ and $g(y) = g'(y') = x_0$. Take indices $k$ and $k'$ such that $y \in T_{i_k}$ and $y' \in T_{i_{k'}}$. We can assume that $\operatorname{dom} g = T_{i_k}$ and $\operatorname{dom} g' = T_{i_{k'}}$ by Remark 24.

Let $f = g^{-1}g' \in \overline{S}$. We have $y' \in \operatorname{dom} f$ and $f(y') = y$. By Remark 24, there exists $\tilde{f} \in \overline{S}$ with $\operatorname{dom} \tilde{f} = T_{i_{k'}}$ and $\boldsymbol{\gamma}(\tilde{f},y') = \boldsymbol{\gamma}(f,y')$. By the definition of $\overline{S}$, there is a sequence $f_n$ in $S$ with $\operatorname{dom} f_n = T_{i_{k'}}$ and $f_n \to f$ in $C_{\text{c-o}}(T_{i_k},T)$ as $n \to \infty$; in particular, $f_n(y') \to f(y') = y$. So we can assume that $f_n(y') \in T_{i_k}$ for all $n$.

Take some relatively compact open neighborhood $V$ of $y'$ such that

$$\overline{V} \subset \operatorname{dom}(g\tilde{f}) \cap \operatorname{dom}(gf)$$

and $\tilde{f} = f$ in some neighborhood of $\overline{V}$. Since $f_n \to \tilde{f}$ in $\overline{S}_{\text{c-o}}$ as $n \to \infty$, we get $gf_n \to g\tilde{f}$ and $f_n^{-1} \to \tilde{f}^{-1}$ by Propositions 2.6 and 3.1. So $\overline{V} \subset \operatorname{dom}(gf_n)$ and $y \in \operatorname{dom} f_n^{-1} = \operatorname{im} f_n$ for $n$ large enough, and $f_n^{-1}(y) \to \tilde{f}^{-1}(y) = y'$. Moreover $gf_n|_V \to g\tilde{f}|_V = gf|_V = g'|_V$ in $C_{\text{c-o}}(V,T)$. So $\boldsymbol{\gamma}(gf_n, f_n^{-1}(y)) \to \boldsymbol{\gamma}(g',y')$ in $\widehat{T}_{0,U}$ by Proposition 2.2 and the definition of the topology of $\widehat{T}$. Thus, with $h_n = f_n^{-1} \in S$, we get

$$\widehat{h_n}(\boldsymbol{\gamma}(g,y)) = \boldsymbol{\gamma}(gh_n^{-1}, h_n(y)) = \boldsymbol{\gamma}(gf_n, f_n^{-1}(y)) \to \boldsymbol{\gamma}(g',y'),$$

and therefore $\boldsymbol{\gamma}(g',y')$ is in the closure of the $\widehat{\mathcal{H}}_{0,U}$-orbit of $\boldsymbol{\gamma}(g,y)$. □

**Remark 27.** By Lemma 3.24, the map $\hat{\pi}_0 : \widehat{T}_0 \to T$ generates a morphism of pseudogroups $\widehat{\mathcal{H}}_0 \to \mathcal{H}$ in the sense of [Álvarez and Masa 2008] — this morphism is not étale.

The following result is elementary.

**Proposition 3.37.** *In Example 2.37, if $\mathcal{H}$ is compactly generated and $\overline{\mathcal{H}}$ is strongly quasianalytic, then $\widehat{\mathcal{H}}_0$ is equivalent to the pseudogroup generated by the local action of $\Gamma$ on $G$ by local left translations, so that $\hat{\pi}_0 : \widehat{T}_0 \to T$ corresponds to the projection $T : V^2 \to G/(K,V)$.*

**Corollary 3.38.** *The map $\hat{\pi}_0 : \widehat{T}_0 \to T$ is open.*

*Proof.* This follows from Theorem 2.38 and Proposition 3.37 since, in Example 2.37, the projection $T : V^2 \to G/(K, V)$ is open. □

**3F.** ***The closure of $\widehat{\mathcal{H}}_0$.*** Let $\widehat{\overline{\mathcal{H}}}_0$ be the pseudogroup on $\widehat{T}_0$ defined like $\widehat{\mathcal{H}}$ by taking the maps $h$ in $\overline{S}$ instead of $S$; thus it is generated by $\widehat{\overline{S}}_0 = \{\hat{h} \mid h \in \overline{S}\}$. Observe that $\widehat{\overline{\mathcal{H}}}_0$, $\overline{S}$ and $\widehat{\overline{S}}_0$ satisfy the obvious versions of Lemmas 3.24–3.26, 3.28, 3.30 and 3.31, and Corollaries 3.27 and 3.29 (Section 3E). In particular, $\widehat{\overline{S}}_0$ is a pseudo$*$group, and $\widehat{T}_{0,U}$ meets all the orbits of $\widehat{\overline{\mathcal{H}}}_0$. The restriction of $\widehat{\overline{\mathcal{H}}}_0$ to $\widehat{T}_{0,U}$ will be denoted by $\widehat{\overline{\mathcal{H}}}_{0,U}$.

**Lemma 3.39.** $\overline{\widehat{\mathcal{H}}_0} = \widehat{\overline{\mathcal{H}}}_0$.

*Proof.* By the version of Lemma 3.31 for $\overline{S}$ and $\widehat{\overline{S}}_0$, the set $\widehat{S}_0$ is dense in $\widehat{\overline{S}}_{0,\text{c-o}}$. Then the result follows easily by Proposition 2.2 and the definition of $\overline{\widehat{\mathcal{H}}_0}$ (see Theorem 2.34 and Remark 19). □

**Lemma 3.40.** *$\widehat{\overline{\mathcal{H}}}_0$ is strongly locally free.*

*Proof.* Let $\hat{h} \in \widehat{\overline{S}}_0$ for $h \in \overline{S}$, and $\boldsymbol{\gamma}(g, x) \in \operatorname{dom} \hat{h}$ for $g \in \overline{S}$ and $x \in \operatorname{dom} g \cap \operatorname{dom} h$ with $g(x) = x_0$. Suppose that $\hat{h}(\boldsymbol{\gamma}(g, x)) = \boldsymbol{\gamma}(g, x)$. This means

$$\boldsymbol{\gamma}(gh^{-1}, h(x)) = \boldsymbol{\gamma}(g, x).$$

So $h(x) = x$ and $gh^{-1} = g$ on some neighborhood of $x$, and therefore $h = \mathrm{id}_T$ on some neighborhood of $x$. Then $h = \mathrm{id}_{\operatorname{dom} h}$ by the strong quasianalyticity condition of $\overline{\mathcal{H}}$ since $h \in \overline{S}$. Hence $\hat{h} = \mathrm{id}_{\operatorname{dom} \hat{h}}$ by Lemma 3.25. □

**Proposition 3.41.** *There is a locally compact Polish local group $G$ and some dense finitely generated sub-local group $\Gamma \subset G$ such that $\widehat{\mathcal{H}}_0$ is equivalent to the pseudogroup defined by the local action of $\Gamma$ on $G$ by local left translations.*

*Proof.* This follows from Remark 21 (see also Theorem 2.38) since $\widehat{\mathcal{H}}_0$ is compactly generated (Corollary 3.33) and equicontinuous (Corollary 3.35), and $\overline{\widehat{\mathcal{H}}_0}$ is strongly locally free (Lemma 3.40). □

**3G.** ***Independence of the choices involved.*** First, let us prove that $\widehat{T}_0$ and $\widehat{\mathcal{H}}_0$ are independent of the choice of the point $x_0$ up to an equivalence generated by a homeomorphism. Let $x_1$ be another point of $T$, and let $\widehat{T}_1$, $\hat{\pi}_1$, $\widehat{S}_1$ and $\widehat{\mathcal{H}}_1$ be constructed like $\widehat{T}_0$, $\hat{\pi}_0$, $\widehat{S}_0$ and $\widehat{\mathcal{H}}_0$ by using $x_1$ instead of $x_0$. Now, for each $h \in S$, let us use the notation $\hat{h}_0 := \hat{h} \in \widehat{S}_0$, and let $\hat{h}_1 : \hat{\pi}_1^{-1}(\operatorname{dom} h) \to \hat{\pi}_1^{-1}(\operatorname{im} h)$ be the map in $\widehat{S}_1$ defined like $\hat{h}$.

**Proposition 3.42.** *There is a homeomorphism $\theta : \widehat{T}_0 \to \widehat{T}_1$ that generates an equivalence $\Theta : \widehat{\mathcal{H}}_0 \to \widehat{\mathcal{H}}_1$ and such that $\hat{\pi}_0 = \hat{\pi}_1 \theta$.*

*Proof.* Since $\mathcal{H}$ is minimal, there is some $f_0 \in \overline{S}$ such that $x_0 \in \operatorname{dom} f_0$ and $f_0(x_0) = x_1$. Let $\theta : \widehat{T_0} \to \widehat{T_1}$ be defined by $\theta(\boldsymbol{\gamma}(f, x)) = \boldsymbol{\gamma}(f_0 f, x)$. This map is continuous because $\theta(\boldsymbol{\gamma}(f, x)) = \boldsymbol{\gamma}(f_0, x)\,\boldsymbol{\gamma}(f, x)$. So $\theta$ is a homeomorphism because $f_0^{-1}$ defines $\theta^{-1}$ in the same way. We also have $\hat{\pi}_0 = \hat{\pi}_1\theta$ since $\theta$ preserves the source of each germ. For each $h \in S$, we have $\operatorname{dom}\hat{h}_1 = \theta(\operatorname{dom}\hat{h}_0)$ because $\hat{\pi}_0 = \hat{\pi}_1\theta$, and $\hat{h}_1\theta = \theta$ since

$$\begin{aligned}\hat{h}_1\theta(\boldsymbol{\gamma}(f, x)) &= \hat{h}_1(\boldsymbol{\gamma}(f_0 f, x)) = \boldsymbol{\gamma}(f_0 f h^{-1}, h(x)) \\ &= \theta\big(\boldsymbol{\gamma}(f h^{-1}, h(x))\big) = \theta\big(\hat{h}_0(\boldsymbol{\gamma}(f, x))\big)\end{aligned}$$

for all $\boldsymbol{\gamma}(f, x) \in \operatorname{dom}\hat{h}_0$. It follows easily that $\theta$ generates an étale morphism $\Theta : \widehat{\mathcal{H}}_0 \to \widehat{\mathcal{H}}_1$, which is an equivalence since $\theta^{-1}$ generates $\Theta^{-1}$. □

Now, let us show that the topology of $\widehat{T}$ is independent of the choice of $S$. Therefore the topology of $\widehat{T_0}$ will be independent of the choice of $S$ as well. Let $S', S'' \subset \mathcal{H}$ be two subpseudo∗groups generating $\mathcal{H}$ and satisfying the conditions of Section 3A. With the notation of Section 3B, we have to prove the following.

**Proposition 3.43.** $\overline{\mathfrak{G}}_{\overline{S'},\text{c-o}} = \overline{\mathfrak{G}}_{\overline{S''},\text{c-o}}$.

*Proof.* First, up to solving the case where $S' \subset S''$, we can assume that $S'$ and $S''$ are local by Remarks 10 and 17. Second, if $S'$ and $S''$ are local, then the subpseudo∗group $S' \cap S''$ of $\mathcal{H}$ also generates $\mathcal{H}$. Moreover $S' \cap S''$ obviously satisfies all other properties required in Section 3A; note that a refinement of $\{T_i\}$ may be necessary to get the properties stated in Remarks 22–25 with $S' \cap S''$. Hence the result follows from the special case where $S' \subset S''$. With this assumption, the identity map $\overline{\mathfrak{G}}_{\overline{S'},\text{c-o}} \to \overline{\mathfrak{G}}_{\overline{S''},\text{c-o}}$ is continuous because the diagram

$$\begin{array}{ccc}\overline{S'}_{\text{c-o}} & \xrightarrow{\text{inclusion}} & \overline{S''}_{\text{c-o}} \\ {\scriptstyle\gamma}\big\downarrow & & \big\downarrow{\scriptstyle\gamma} \\ \overline{\mathfrak{G}}_{\overline{S'},\text{c-o}} & \xrightarrow{\text{identity}} & \overline{\mathfrak{G}}_{\overline{S''},\text{c-o}}\end{array}$$

is commutative, where the vertical maps are identifications and the top map is continuous.

For any compact subset $Q \subset T$, let $s^{-1}(Q)_{\overline{S'},\text{c-o}}$ and $s^{-1}(Q)_{\overline{S''},\text{c-o}}$ denote the spaces obtained by endowing $s^{-1}(Q)$ with the restriction of the topologies of $\overline{\mathfrak{G}}_{\overline{S'},\text{c-o}}$ and $\overline{\mathfrak{G}}_{\overline{S''},\text{c-o}}$, respectively. They are compact and Hausdorff by Propositions 3.17 and 3.18. It follows that $s^{-1}(Q)_{\overline{S'},\text{c-o}} = s^{-1}(Q)_{\overline{S''},\text{c-o}}$ because the identity map $s^{-1}(Q)_{\overline{S'},\text{c-o}} \to s^{-1}(Q)_{\overline{S''},\text{c-o}}$ is continuous. Hence, for any $\boldsymbol{\gamma}(f, x) \in \overline{\mathfrak{G}}$ and a compact neighborhood $Q$ of $x$ in $T$, the set $s^{-1}(Q)$ is a neighborhood of $\boldsymbol{\gamma}(f, x)$ in $\overline{\mathfrak{G}}_{\overline{S'},\text{c-o}}$ and $\overline{\mathfrak{G}}_{\overline{S''},\text{c-o}}$ with $s^{-1}(Q)_{\overline{S'},\text{c-o}} = s^{-1}(Q)_{\overline{S''},\text{c-o}}$. This shows that the identity map $\overline{\mathfrak{G}}_{\overline{S'},\text{c-o}} \to \overline{\mathfrak{G}}_{\overline{S''},\text{c-o}}$ is a local homeomorphism, and therefore a homeomorphism. □

Let $T'$ be an open subset of $T$ containing $x_0$, which meets all orbits because $\mathcal{H}$ is minimal. Then use $T'$, $\mathcal{H}' = \mathcal{H}|_{T'}$ and $S' = S \cap \mathcal{H}'$ to define $\widehat{T}_0'$, $\hat{\pi}_0'$, $\widehat{S}_0'$ and $\widehat{\mathcal{H}}_0'$ like $\widehat{T}_0$, $\hat{\pi}_0$, $\widehat{S}_0$ and $\widehat{\mathcal{H}}_0$. The proof of the following result is elementary.

**Proposition 3.44.** *There is a canonical identity of topological spaces, $\widehat{T}_0' \equiv \hat{\pi}_0^{-1}(T')$, such that $\hat{\pi}_0' \equiv \hat{\pi}_0|_{\widehat{T}_0'}$ and $\widehat{\mathcal{H}}_0' = \widehat{\mathcal{H}}_0|_{\widehat{T}_0'}$.*

**Corollary 3.45.** *Let $\mathcal{H}$ and $\mathcal{H}'$ be minimal equicontinuous compactly generated pseudogroups on locally compact Polish spaces such that $\overline{\mathcal{H}}$ and $\overline{\mathcal{H}'}$ are strongly quasianalytic. If $\mathcal{H}$ is equivalent to $\mathcal{H}'$, then $\widehat{\mathcal{H}}_0$ is equivalent to $\widehat{\mathcal{H}}_0'$.*

*Proof.* This is a direct consequence of Propositions 3.42–3.44. □

The following definition makes sense by Lemma 2.36, Propositions 3.42 and 3.43, and Corollary 3.45.

**Definition 3.46.** In Proposition 3.41, it is said that (the local isomorphism class of) $G$ is the *structural local group* of (the equivalence class of) $\mathcal{H}$.

## 4. Molino's theory for equicontinuous foliated spaces

**4A.** ***Preliminaries on equicontinuous foliated spaces.*** (See [Moore and Schochet 1988; Candel and Conlon 2000, Chapter 11; Ghys 1999].)

Let $X$ and $Z$ be locally compact Polish spaces. A *foliated chart* in $X$ of *leaf dimension $n$*, *transversely modeled* on $Z$, is a pair $(U, \phi)$, where $U \subseteq X$ is open and $\phi : U \to B \times T$ is a homeomorphism for some open $T \subset Z$ and some open ball $B$ in $\mathbb{R}^n$. It is said that $U$ is a *distinguished open set*. The sets $P_y = \phi^{-1}(B \times \{y\})$ for $y \in T$ are called *plaques* of this foliated chart. For every $x \in B$, the set $S_x = \phi^{-1}(\{x\} \times T)$ is called a *transversal* of the foliated chart. This local product structure defines a local projection $p : U \to T$, called *distinguished submersion*, given as composition of $\phi$ with the second factor projection $\mathrm{pr}_2 : B \times T \to T$.

Let $\mathcal{U} = \{U_i, \phi_i\}$ be a family of foliated charts in $X$ of leaf dimension $n$ modeled transversally on $Z$ and covering $X$. Assume further that the foliated charts are *coherently foliated* in the sense that, if $P$ and $Q$ are plaques in different charts of $\mathcal{U}$, then $P \cap Q$ is open both in $P$ and $Q$. Then $\mathcal{U}$ is called a *foliated atlas* on $X$ of *leaf dimension $n$* and *transversely modeled* on $Z$. A maximal foliated atlas $\mathcal{F}$ of leaf dimension $n$ and transversely modeled on $Z$ is called a *foliated structure* on $X$ of *leaf dimension $n$* and *transversely modeled* on $Z$. Any foliated atlas $\mathcal{U}$ of this type is contained in a unique foliated structure $\mathcal{F}$; then it is said that $\mathcal{U}$ *defines* (or is an atlas of) $\mathcal{F}$. If $Z = \mathbb{R}^m$, then $X$ is a manifold of dimension $n + m$, and $\mathcal{F}$ is traditionally called a *foliation* of *dimension $n$* and *codimension $m$*. The reference to $Z$ will be omitted.

For a foliated structure $\mathcal{F}$ on $X$ of dimension $n$, the plaques form a basis of a topology on $X$ called the *leaf topology*. With the leaf topology, $X$ becomes an

$n$-manifold whose connected components are called *leaves* of $\mathcal{F}$. $\mathcal{F}$ is determined by its leaves.

A foliated atlas $\mathcal{U} = \{U_i, \phi_i\}$ of $\mathcal{F}$ is called *regular* if

- each $\overline{U_i}$ is a compact subset of a foliated chart $(W_i, \psi_i)$ and $\phi_i = \psi_i|_{U_i}$;
- the cover $\{U_i\}$ is locally finite; and,
- if $(U_i, \phi_i)$ and $(U_j, \phi_j)$ are elements of $\mathcal{U}$, then each plaque $P$ of $(U_i, \phi_i)$ meets at most one plaque of $(U_j, \phi_j)$.

In this case, there are homeomorphisms $h_{ij} : T_{ij} \to T_{ji}$ such that $h_{ij} p_i = p_j$ on $U_i \cap U_j$, where $p_i : U_i \to T_i$ is the distinguished submersion defined by $(U_i, \phi_i)$ and $T_{ij} = p_i(U_i \cap U_j)$. Observe that the cocycle condition $h_{ik} = h_{jk} h_{ij}$ is satisfied on $T_{ijk} = p_i(U_i \cap U_j \cap U_k)$. For this reason, $\{U_i, p_i, h_{ij}\}$ is called a *defining cocycle* of $\mathcal{F}$ with values in $Z$ — we only consider defining cocycles induced by regular foliated atlases. The equivalence class of the pseudogroup $\mathcal{H}$ generated by the maps $h_{ij}$ on $T = \bigsqcup_{i \in I} T_i$ is called the *holonomy pseudogroup* of the foliated space $(X, \mathcal{F})$; $\mathcal{H}$ is the representative of the holonomy pseudogroup of $(X, \mathcal{F})$ induced by the defining cocycle $\{U_i, p_i, h_{ij}\}$. This $T$ can be identified with a *total* (or *complete*) *transversal* to the leaves in the sense that it meets all leaves and is locally given by the transversals defined by foliated charts. All compositions of maps $h_{ij}$ form a pseudo$*$group $S$ that generates $\mathcal{H}$, called the *holonomy pseudo$*$group* of $\mathcal{F}$ induced by $\{U_i, p_i, h_{ij}\}$. There is a canonical identity between the space of leaves and the space of $\mathcal{H}$-orbits, $X/\mathcal{F} \equiv T/\mathcal{H}$.

A foliated atlas (respectively, defining cocycle) contained in another one is called a *subfoliated atlas* (respectively, *subfoliated cocycle*).

The *holonomy group* of each leaf $L$ is defined as the germ group of the corresponding orbit. It can be considered as a quotient of $\pi_1(L)$ by taking "chains" of sets $U_i$ along loops in $L$; this representation of $\pi_1(L)$ is called the *holonomy representation*. The kernel of the holonomy representation is equal to $q_*\pi_1(\widetilde{L})$ for a regular covering space $q : \widetilde{L} \to L$, which is called the *holonomy cover* of $L$. If $\mathcal{F}$ admits a countable defining cocycle, then the leaves in some dense $G_\delta$ subset of $M$ have trivial holonomy groups [Hector and Hirsch 1981; 1987; Candel and Conlon 2000], and therefore they can be identified with their holonomy covers.

It is said that a foliated space is (*topologically*) *transitive* or *minimal* if any representative of its holonomy pseudogroup is such. Transitivity (respectively, minimality) of a foliated space means that some leaf is dense (respectively, all leaves are dense).

Haefliger [2002] has observed that, if $X$ is compact, then $\mathcal{H}$ is compactly generated, which can be seen as follows. There is some defining cocycle $\{U'_i, p'_i, h'_{ij}\}$, with $p'_i : U'_i \to T'_i$, such that $\overline{U_i} \subset U'_i$, $T_i \subset T'_i$ and $p'_i$ extends $p_i$. Therefore each $h'_{ij}$ is an extension of $h_{ij}$ so that $\overline{\operatorname{dom} h_{ij}} \subset \operatorname{dom} h'_{ij}$. Moreover $\mathcal{H}$ is the restriction to $T$ of the pseudogroup $\mathscr{H}'$ on $T' = \bigsqcup_i T'_i$ generated by the maps $h'_{ij}$, and $T$ is a relatively compact open subset of $T'$ that meets all $\mathscr{H}'$-orbits.

**Definition 4.1.** It is said that a foliated space is *equicontinuous* if any representative of its holonomy pseudogroup is equicontinuous.

**Remark 28.** The above definition makes sense by Lemma 2.30.

**Definition 4.2.** Let $G$ be a locally compact Polish local group. A minimal foliated space is called a *$G$-foliated space* if its holonomy pseudogroup can be represented by a pseudogroup given by Example 2.35 on a local group locally isomorphic to $G$.

**4B. *Molino's theory for equicontinuous foliated spaces.*** Let $(X, \mathscr{F})$ be a compact minimal foliated space that is equicontinuous and such that the closure of its holonomy pseudogroup is strongly quasianalytic. Let $\{U_i, p_i, h_{ij}\}$ be a defining cocycle of $\mathscr{F}$ induced by a regular foliated atlas, where $p_i : U_i \to T_i$. Let $\mathscr{H}$ denote the corresponding representative of the holonomy pseudogroup on $T = \bigsqcup_i T_i$, which satisfies the conditions of Section 3A. Let $S$ be the localization of the holonomy pseudo$*$group induced by $\{U_i, p_i, h_{ij}\}$. Fix an index $i_0$ and a point $x_0 \in U_{i_0}$. Let $\hat\pi_0 : \widehat{T}_0 \to T$ and $\widehat{\mathscr{H}}_0$ be defined like in Sections 3D and 3E, by using $T$, $\mathscr{H}$, the point $p_{i_0}(x_0) \in T_{i_0} \subset T$, and a local subpseudo$*$group $S \subset \mathscr{H}$.

With the notation $\widehat{T}_{i,0} = \hat\pi_0^{-1}(T_i) \subset \widehat{T}_0$, let

$$\check{X}_0 = \bigsqcup_i U_i \times \widehat{T}_{i,0} = \bigcup_i U_i \times \widehat{T}_{i,0} \times \{i\},$$

equipped with the corresponding topological sum of the product topologies, and consider its closed subspace

$$\widetilde{X}_0 = \{(x, \gamma, i) \in \check{X}_0 \mid p_i(x) = \hat\pi_0(\gamma)\} \subset \check{X}_0.$$

For $(x, \gamma, i), (y, \delta, j) \in \widetilde{X}_0$, write $(x, \gamma, i) \sim (y, \delta, j)$ if $x = y$ and $\gamma = \widehat{h_{ji}}(\delta)$. Since $h_{ij} p_i(x) = p_j(x)$, $h_{ji}^{-1} = h_{ij}$ and $h_{ik} = h_{jk} h_{ij}$, it follows that this defines an equivalence relation $\sim$ on $\widetilde{X}_0$. Let $\widehat{X}_0$ be the corresponding quotient space, $q : \widetilde{X}_0 \to \widehat{X}_0$ the quotient map, and $[x, \gamma, i]$ the equivalence class of each triple $(x, \gamma, i)$. For each $i$, let

$$\check{U}_{i,0} = U_i \times \widehat{T}_{i,0} \times \{i\}, \quad \widetilde{U}_{i,0} = \check{U}_{i,0} \cap \widetilde{X}_0, \quad \widehat{U}_{i,0} = q(\widetilde{U}_{i,0}).$$

**Lemma 4.3.** *$\widehat{U}_{i,0}$ is open in $\widehat{X}_0$.*

*Proof.* We have to check that $q^{-1}(\widehat{U}_{i,0}) \cap \widetilde{U}_{j,0}$ is open in $\widetilde{U}_{j,0}$ for all $j$, which is true because

$$q^{-1}(\widehat{U}_{i,0}) \cap \widetilde{U}_{j,0} = ((U_i \cap U_j) \times \widehat{T}_{j,0} \times \{j\}) \cap \widetilde{X}_0. \qquad \square$$

**Lemma 4.4.** *The quotient map $q : \widetilde{U}_{i,0} \to \widehat{U}_{i,0}$ is a homeomorphism.*

*Proof.* This map is surjective by the definition of $\widehat{U}_{i,0}$. On the other hand, two equivalent triples in $\widetilde{U}_{i,0}$ are of the form $(x,\gamma,i)$ and $(x,\delta,i)$ with $\gamma=\widehat{h_{ii}}(\delta)=\delta$. So $q:\widetilde{U}_{i,0}\to\widehat{U}_{i,0}$ is also injective. Since $q:\widetilde{U}_{i,0}\to\widehat{U}_{i,0}$ is continuous, it only remains to prove that this map is open. A basis of the topology of $\widetilde{U}_{i,0}$ consists of the sets of the form $(V\times W\times\{i\})\cap\widetilde{X}_0$, where $V$ and $W$ are open in $U_i$ and $\widehat{T}_{i,0}$, respectively. These basic sets satisfy

$$\widetilde{U}_{j,0}\cap q^{-1}q\big((V\times W\times\{i\})\cap\widetilde{X}_0\big)=\widetilde{U}_{j,0}\cap\big(V\times\widehat{h_{ij}}(W\cap\operatorname{dom}\widehat{h_{ij}})\times\{j\}\big)$$

for all $j$, which is open in $\widetilde{U}_{j,0}$. So $q^{-1}q((V\times W\times\{i\})\cap\widetilde{X}_0)$ is open in $\widetilde{X}_0$ and therefore $q((V\times W\times\{i\})\cap\widetilde{X}_0)$ is open in $\widehat{X}_0$. □

**Proposition 4.5.** *$\widehat{X}_0$ is compact and Polish.*

*Proof.* Let $\{U'_i,p'_i,h'_{ij}\}$ be a shrinking of $\{U_i,p_i,h_{ij}\}$; i.e., it is a defining cocycle of $\mathcal{F}$ such that $\overline{U'_i}\subset U_i$ and $p'_i:U'_i\to T'_i$ is the restriction of $p_i$ for all $i$. Therefore each $h'_{ij}$ is also a restriction of $h_{ij}$ and $T'_i$ is a relatively compact open subset of $T_i$. Then $\hat{\pi}_0^{-1}(\overline{T'_i})$ is a compact subset of $\widehat{T}_{i,0}$ by Corollary 3.21. Moreover $\widehat{X}_0$ is the union of the sets $q(\overline{U'_i}\times\hat{\pi}_0^{-1}(\overline{T'_i})\times\{i\})$. So $\widehat{X}_0$ is compact because it is a finite union of compact sets.

On the other hand, since $\widetilde{X}_0$ is closed in $\check{X}_0$, and $\check{U}_{i,0}$ is Polish and locally compact by Corollary 3.19, it follows that $\widetilde{U}_{i,0}$ is Polish and locally compact, and therefore $\widehat{U}_{i,0}$ is Polish and locally compact by Lemma 4.4. Then, by the compactness of $\widehat{X}_0$, Lemma 4.3 and [Kechris 1991, Theorem 5.3], it only remains to prove that $\widehat{X}_0$ is Hausdorff.

Let $[x,\gamma,i]\neq[y,\delta,j]$ in $\widehat{X}_0$. So $x\in U_i$ and $y\in U_j$. If $x=y$, then we have $[y,\delta,j]=[x,\widehat{h_{ji}}(\delta),i]\in\widehat{U}_{i,0}$. Thus, in this case, $[x,\gamma,i]$ and $[y,\delta,j]$ can be separated by open subsets of $\widehat{U}_{i,0}$ because $\widehat{U}_{i,0}$ is Hausdorff.

Now suppose that $x\neq y$. Then take disjoint open neighborhoods, $V$ of $x$ in $U_i$ and $W$ of $y$ in $U_j$. Let

$$\begin{aligned}
&\check{V}=V\times\widehat{T}_{i,0}\times\{i\}\subset\check{U}_{i,0}, &&\check{W}=V\times\widehat{T}_{j,0}\times\{j\}\subset\check{U}_{j,0},\\
&\widetilde{V}=\check{V}\cap\widetilde{X}_0\subset\widetilde{U}_{i,0}, &&\widetilde{W}=\check{W}\cap\widetilde{X}_0\subset\widetilde{U}_{j,0},\\
&\widehat{V}=q(\widetilde{V})\subset\widehat{U}_{i,0}, &&\widehat{W}=q(\widetilde{W})\subset\widehat{U}_{j,0}.
\end{aligned}$$

The sets $\widehat{V}$ and $\widehat{W}$ are open neighborhoods of $[x,\gamma,i]$ and $[y,\delta,j]$ in $\widehat{X}_0$. Suppose that $\widehat{V}\cap\widehat{W}\neq\varnothing$. Then there is a point $(x',\gamma',i)\in\widetilde{V}$ which is equivalent to some point $(y',\delta',j)\in\widetilde{W}$. This implies that $x'=y'\in V\cap W$, which is a contradiction because $V\cap W=\varnothing$. Therefore $\widehat{V}\cap\widehat{W}=\varnothing$. □

According to the above equivalence relation of triples, a map $\hat{\pi}_0:\widehat{X}_0\to X$ is defined by $\hat{\pi}_0([x,\gamma,i])=x$.

**Proposition 4.6.** *The map $\hat{\pi}_0 : \widehat{X}_0 \to X$ is continuous and surjective, and its fibers are homeomorphic to each other.*

*Proof.* Since each map $\hat{\pi}_0 : \widehat{T}_{i,0} \to T_i$ is surjective, we have $\hat{\pi}_0(\widehat{U}_{i,0}) = U_i$, obtaining that $\hat{\pi}_0 : \widehat{X}_0 \to X$ is surjective. Moreover the composition

$$\widetilde{U}_{i,0} \xrightarrow{\ q\ } \widehat{U}_{i,0} \xrightarrow{\ \hat{\pi}_0\ } U_i$$

is the restriction of the first factor projection $\check{U}_{i,0} \to U_i$, $(x, \gamma, i) \mapsto x$. Therefore, $\hat{\pi}_0 : \widehat{X}_0 \to X$ is continuous by Lemmas 4.3 and 4.4.

For $x \in U_i$, we have $\hat{\pi}_0^{-1}(x) \subset \widehat{U}_{i,0}$ and

$$\widetilde{U}_{i,0} \cap q^{-1}(\hat{\pi}_0^{-1}(x)) = \{x\} \times \hat{\pi}_0^{-1}(p_i(x)) \times \{i\} \equiv \hat{\pi}_0^{-1}(p_i(x)) \subset \widehat{T}_{i,0}.$$

So the last assertion follows from Lemma 4.4 and Proposition 3.22. □

Let $\tilde{p}_{i,0} : \widetilde{U}_{i,0} \to \widehat{T}_{i,0}$ denote the restriction of the second factor projection $\check{p}_{i,0} : \check{U}_{i,0} = U_i \times \widehat{T}_{i,0} \times \{i\} \to \widehat{T}_{i,0}$. By Lemma 4.4, $\tilde{p}_{i,0}$ induces a continuous map $\hat{p}_{i,0} : \widehat{U}_{i,0} \to \widehat{T}_{i,0}$.

**Proposition 4.7.** $\{\widehat{U}_{0.i}, \hat{p}_{i,0}, \widehat{h_{ij}}\}$ *is a defining cocycle of a foliated structure $\widehat{\mathcal{F}}_0$ on $\widehat{X}_0$.*

*Proof.* Let $\{U_i, \phi_i\}$ be a regular foliated atlas of $\mathcal{F}$ inducing the defining cocycle $\{U_i, p_i, h_{ij}\}$, where $\phi_i : U_i \to B_i \times T_i$ is a homeomorphism and $B_i$ is a ball in $\mathbb{R}^n$ ($n = \dim \mathcal{F}$). Then we get a homeomorphism

$$\check{\phi}_{i,0} = \phi_i \times \mathrm{id} \times \mathrm{id} : \check{U}_{i,0} = U_i \times \widehat{T}_{i,0} \times \{i\} \to B_i \times T_i \times \widehat{T}_{i,0} \times \{i\}.$$

Observe that $\check{\phi}_{i,0}(\widetilde{U}_{i,0})$ consists of the elements $(y, z, \gamma, i)$ with $\hat{\pi}_0(\gamma) = z$. So $\check{\phi}_{i,0}$ restricts to a homeomorphism

$$\tilde{\phi}_{i,0} : \widetilde{U}_{i,0} \to \check{\phi}_{i,0}(\widetilde{U}_{i,0}) \equiv B_i \times \widehat{T}_{i,0} \times \{i\} \equiv B_i \times \widehat{T}_{i,0}.$$

By Lemma 4.4, $\tilde{\phi}_{i,0}$ induces a homeomorphism $\hat{\phi}_{i,0} : \widehat{U}_{i,0} \to B_i \times \widehat{T}_{i,0}$. Moreover, $\check{p}_{i,0}$ corresponds to the third factor projection via $\check{\phi}_{i,0}$, obtaining that $\tilde{p}_{i,0}$ corresponds to the second factor projection via $\tilde{\phi}_{i,0}$, and therefore $\hat{p}_{i,0}$ also corresponds to the second factor projection via $\hat{\phi}_{i,0}$. Observe that $\hat{p}_{i,0} = \widehat{h_{ji}}\,\hat{p}_{j,0}$ on $\widehat{U}_{i,0} \cap \widehat{U}_{j,0}$ by the definition of $\sim$. The regularity of the foliated atlas $\{\widehat{U}_{0.i}, \hat{\phi}_{i,0}\}$ follows easily from the regularity of $\{U_i, \phi_i\}$. □

According to Proposition 4.7, the holonomy pseudogroup of $\widehat{\mathcal{F}}_0$ is represented by the pseudogroup on $\bigsqcup_i \widehat{T}_{i,0}$ generated by the maps $\widehat{h_{ij}}$, which is the pseudogroup $\widehat{\mathcal{H}}_0$ on $\widehat{T}_0$.

**Corollary 4.8.** *There is some locally compact Polish local group $G$ such that $(\widehat{X}_0, \widehat{\mathcal{F}}_0)$ is a minimal $G$-foliated space; in particular, it is equicontinuous.*

*Proof.* This follows from Propositions 4.7 and 3.41, and Lemma 3.36. □

**Proposition 4.9.** *The map* $\hat{\pi}_0 : (\widehat{X}_0, \widehat{\mathcal{F}}_0) \to (X, \mathcal{F})$ *is foliated.*

*Proof.* According to Proposition 4.7, this follows by checking the commutativity of each diagram

$$\begin{array}{ccc} \widehat{U}_{i,0} & \xrightarrow{\hat{p}_{i,0}} & \widehat{T}_{i,0} \\ {\scriptstyle \hat{\pi}_0}\downarrow & & \downarrow{\scriptstyle \hat{\pi}_0} \\ U_i & \xrightarrow{p_i} & T_i \end{array}$$

By Lemma 4.4, and the definition of $\hat{p}_{i,0}$ and $\hat{\pi}_{i,0}$, this commutativity follows from the commutativity of

$$\begin{array}{ccc} \widetilde{U}_{i,0} & \longrightarrow & \widehat{T}_{i,0} \\ \downarrow & & \downarrow{\scriptstyle \hat{\pi}_0} \\ U_i & \xrightarrow{p_i} & T_i \end{array}$$

where the left vertical and the top horizontal arrows denote the restrictions of the first and second factor projections of $\breve{U}_{i,0} = U_i \times \widehat{T}_{i,0} \times \{i\}$. But the commutativity of this diagram holds by the definition of $\widetilde{X}_0$ and $\widetilde{U}_{i,0}$. □

**Proposition 4.10.** *The restrictions of* $\hat{\pi}_0 : \widehat{X}_0 \to X$ *to the leaves are the holonomy covers of the leaves of* $\mathcal{F}$.

*Proof.* With the notation of the proof of Proposition 4.7, the diagram

$$\begin{array}{ccc} \widehat{U}_{i,0} & \xrightarrow{\hat{\phi}_{i,0}} & B_i \times \widehat{T}_{i,0} \\ {\scriptstyle \hat{\pi}_0}\downarrow & & \downarrow{\scriptstyle \mathrm{id}_{B_i} \times \hat{\pi}_0} \\ U_i & \xrightarrow{\phi_i} & B_i \times T_i \end{array} \tag{19}$$

is commutative, and $\widehat{U}_{i,0} = \hat{\pi}_0^{-1}(U_i)$. Hence, for corresponding plaques in $U_i$ and $\widehat{U}_{i,0}$, namely $P_z = \phi_0^{-1}(B_i \times \{\hat{z}\})$ and $\widehat{P}_{\hat{z}} = \hat{\phi}_0^{-1}(B_i \times \{z\})$ with $z \in T_i$ and $\hat{z} \in \hat{\pi}_0^{-1}(z) \subset \widehat{T}_{i,0}$, the restriction $\hat{\pi}_0 : \widehat{P}_{\hat{z}} \to P_z$ is a homeomorphism. It follows easily that $\hat{\pi}_0 : \widehat{X}_0 \to X$ restricts to covering maps of the leaves of $\widehat{\mathcal{F}}_0$ to the leaves of $\mathcal{F}$. In fact, these are the holonomy covers, which can be seen as follows.

According to the proof of Proposition 4.6 and the definition of the equivalence relation $\sim$ on $\widetilde{X}_0$, for each $x$ in $U_i \cap U_j$, we have homeomorphisms

$$\hat{\pi}_0^{-1}(p_i(x)) \xleftarrow{\hat{p}_{i,0}} \hat{\pi}_0^{-1}(x) \xrightarrow{\hat{p}_{j,0}} \hat{\pi}_0^{-1}(p_j(x))$$

satisfying $\hat{p}_{j,0}\hat{p}_{i,0}^{-1} = \widehat{h_{ij}}$. This easily implies the following. Given $x \in U_i$ and $\hat{x} \in \hat{\pi}_0^{-1}(X)$, denoting by $L$ and $\widehat{L}$ the leaves through $x$ and $\hat{x}$, respectively, and given a loop $c$ in $L$ based at $x$ inducing a local holonomy transformation $h \in S$

around $p_i(x)$ in $T_i$, the lift $\hat{c}$ of $c$ to $\widehat{L}$ with $\hat{c}(0)=\hat{x}$ satisfies $\hat{p}_{i,0}\hat{c}(1)=\hat{h}\hat{p}_{i,0}(\hat{x})$. Writing $\hat{p}_{i,0}(\hat{x})=\boldsymbol{\gamma}(f,p_i(x))$, we obtain

$$\hat{p}_{i,0}\hat{c}(1)=\hat{h}\big(\boldsymbol{\gamma}(f,p_i(x))\big)=\boldsymbol{\gamma}(fh,p_i(x)).$$

Thus $\hat{c}$ is a loop if and only if $\boldsymbol{\gamma}(fh,p_i(x))=\boldsymbol{\gamma}(f,p_i(x))$, which means that $\boldsymbol{\gamma}(h,p_i(x))=\boldsymbol{\gamma}(\mathrm{id}_T,p_i(x))$. So $\widehat{L}$ is the holonomy cover of $L$. $\square$

**Proposition 4.11.** *The map $\hat{\pi}_0:\widehat{X}_0\to X$ is open.*

*Proof.* This follows from Corollary 3.38 and the commutativity of (19). $\square$

Theorem A is the combination of the results of this section.

**4C. *Independence of the choices involved.*** Let $x_1$ be another point of $X$, and let $\widehat{X}_1$, $\widehat{\mathcal{F}}_1$ and $\hat{\pi}_1:\widehat{X}_1\to X$ be constructed like $\widehat{X}_0$, $\widehat{\mathcal{F}}_0$ and $\hat{\pi}_0:\widehat{X}_0\to X$ by using $x_1$ instead of $x_0$.

**Proposition 4.12.** *There is a foliated homeomorphism $\hat{\theta}:(\widehat{X}_0,\widehat{\mathcal{F}}_0)\to(\widehat{X}_1,\widehat{\mathcal{F}}_1)$ such that $\hat{\pi}_1F=\hat{\pi}_0$.*

*Proof.* Take an index $i_1$ such that $x_1\in U_{i_1}$. Let $\widehat{S}_1$, $\widehat{T}_1$, $\widehat{\mathcal{H}}_1$ and $\hat{\pi}_1:\widehat{T}_1\to T$ be constructed like $\widehat{S}_0$, $\widehat{T}_0$, $\widehat{\mathcal{H}}_0$ and $\hat{\pi}_0:\widehat{T}_0\to T$ by using $p_{i_1}(x_1)$ instead of $p_{i_0}(x_0)$, and let $\widehat{T}_{i,1}=\hat{\pi}_1^{-1}(T_i)$. Then the construction of $\widehat{X}_1$, $\widehat{\mathcal{F}}_1$ and $\hat{\pi}_1:\widehat{X}_1\to X$ involves the objects $\check{X}_1$, $\widetilde{X}_1$, $\check{U}_{i,1}$, $\widetilde{U}_{i,1}$, $\widehat{U}_{i,1}$, $\check{p}_{i,1}$, $\tilde{p}_{i,1}$, $\hat{p}_{i,1}$, $\check{\phi}_{i,1}$, $\tilde{\phi}_{i,1}$ and $\hat{\phi}_{i,1}$, defined like $\check{X}_0$, $\widetilde{X}_0$, $\check{U}_{i,0}$, $\widetilde{U}_{i,0}$, $\widehat{U}_{i,0}$, $\check{p}_{i,0}$, $\tilde{p}_{i,0}$, $\hat{p}_{i,0}$, $\check{\phi}_{i,0}$, $\tilde{\phi}_{i,0}$ and $\hat{\phi}_{i,0}$, by using $\widehat{T}_{i,1}$ and $\hat{\pi}_1:\widehat{T}_{i,1}\to T_i$ instead of $\widehat{T}_{i,0}$ and $\hat{\pi}_0:\widehat{T}_{i,0}\to T_i$. Let $\theta:\widehat{T}_0\to\widehat{T}_1$ be the homeomorphism given by Proposition 3.42, which obviously restricts to homeomorphisms $\theta_i:\widehat{T}_{i,0}\to\widehat{T}_{i,1}$. Since $\hat{\pi}_0=\hat{\pi}_1\theta$, it follows that each homeomorphism

$$\check{\theta}_i=\mathrm{id}_{U_i}\times\theta_i\times\mathrm{id}:\check{U}_{i,0}=U_i\times\widehat{T}_{i,0}\times\{i\}\to\check{U}_{i,1}=U_i\times\widehat{T}_{i,1}\times\{i\}$$

restricts to a homeomorphism $\tilde{\theta}_i=\widetilde{U}_{i,0}\to\widetilde{U}_{i,1}$. The combination of the homeomorphisms $\tilde{\theta}_i$ is a homeomorphism $\tilde{\theta}:\widetilde{X}_0\to\widetilde{X}_1$.

For each $h\in S$, use the notation $\hat{h}_0\in\widehat{S}_0$ and $\hat{h}_1\in\widehat{S}_1$ for the map $\hat{h}$ defined with $p_{i_0}(x_0)$ and $p_{i_1}(x_1)$, respectively. From the proof of Proposition 3.42, we get $\hat{h}_1\theta=\theta\hat{h}_0$ for all $h\in S$; in particular, this holds with $h=h_{ij}$. So $\tilde{\theta}:\widetilde{X}_0\to\widetilde{X}_1$ is compatible with the equivalence relations used to define $\widehat{X}_0$ and $\widehat{X}_1$, and therefore it induces a homeomorphism $\hat{\theta}:\widehat{X}_0\to\widehat{X}_1$. Note that $\hat{\theta}$ restricts to homeomorphisms $\hat{\theta}_i:\widehat{U}_{i,0}\to\widehat{U}_{i,1}$. Obviously, $\check{p}_{i,1}\check{\theta}_i=\theta_i\check{p}_{i,1}$, yielding $\tilde{p}_{i,1}\tilde{\theta}_i=\theta_i\tilde{p}_{i,1}$, and therefore $\hat{p}_{i,1}\hat{\theta}_i=\theta_i\hat{p}_{i,1}$. It follows that $\hat{\theta}$ is a foliated map. $\square$

Let $\{U'_a,p'_a,h'_{ab}\}$ be another defining cocycle of $\mathcal{F}$ induced by a regular foliated atlas. Then construct $\widehat{X}'_0$, $\widehat{\mathcal{F}}'_0$ and $\hat{\pi}'_0:\widehat{X}_0\to X$ like $\widehat{X}_0$, $\widehat{\mathcal{F}}_0$ and $\hat{\pi}_0:\widehat{X}_0\to X$ by using $\{U'_a,p'_a,h'_{ab}\}$ instead of $\{U_i,p_i,h_{ij}\}$.

**Proposition 4.13.** *There is a foliated homeomorphism* $F : (\widehat{X}_0, \widehat{\mathcal{F}}_0) \to (\widehat{X}'_0, \widehat{\mathcal{F}}'_0)$ *such that* $\hat{\pi}'_0 F = \hat{\pi}_0$.

*Proof.* By using a common refinement of the open coverings $\{U_i\}$ and $\{U'_a\}$, we can assume that $\{U'_a\}$ refines $\{U_i\}$. In this case, the union of the defining cocycles $\{U_i, p_i, h_{ij}\}$ and $\{U'_a, p'_a, h'_{ab}\}$ is contained in another defining cocycle induced by a regular foliated atlas. Thus the proof boils down to showing that a subdefining cocycle[9] $\{U_{i_k}, p_{i_k}, h_{i_k i_l}\}$ of $\{U_i, p_i, h_{ij}\}$ induces a foliated space homeomorphic to $(\widehat{X}_0, \widehat{\mathcal{F}}_0)$. But the pseudogroup $\mathcal{H}'$ induced by $\{U_{i_k}, p_{i_k}, h_{i_k i_l}\}$ is the restriction of $\mathcal{H}$ to an open subset $T' \subset T$, and the pseudo$*$group induced by $\{U_{i_k}, p_{i_k}, h_{i_k i_l}\}$ is $S' = S \cap \mathcal{H}'$. Then, by using the canonical identity given by Proposition 3.44, it easily follows that the foliated space $(\widehat{X}'_0, \widehat{\mathcal{F}}'_0)$ defined with $\{U_{i_k}, p_{i_k}, h_{i_k i_l}\}$ can be canonically identified with an open foliated subspace of $(\widehat{X}_0, \widehat{\mathcal{F}}_0)$, which indeed is the whole of $(\widehat{X}_0, \widehat{\mathcal{F}}_0)$ because $\{U_{i_k}\}$ covers $X$. □

The following definition makes sense by Propositions 4.12–4.13 and the results used to justify Definition 3.46.

**Definition 4.14.** In Corollary 4.8, (the local isomorphism class of) $G$ is called the *structural local group* of $(X, \mathcal{F})$.

## 5. Growth of equicontinuous pseudogroups and foliated spaces

**5A.** ***Coarse quasi-isometries and growth of metric spaces.*** A *net* in a metric space $M$, with metric $d$, is a subset $A \subset M$ that satisfies $d(x, A) \leq C$ for some $C > 0$ and all $x \in M$; the term $C$-*net* is also used. A *coarse quasi-isometry* between $M$ and another metric space $M'$ is a bi-Lipschitz bijection between nets of $M$ and $M'$; in this case, $M$ and $M'$ are said to be *coarsely quasi-isometric* (in the sense of Gromov) [Gromov 1993]. If such a bi-Lipschitz bijection, as well as its inverse, has dilation $\leq \lambda$, and it is defined between $C$-nets, then it will be said that the coarse quasi-isometry has *distortion* $(C, \lambda)$. A family of coarse quasi-isometries with a common distortion will be called a family of *equicoarse quasi-isometries*, and the corresponding metric spaces are called *equicoarsely quasi-isometric*.

The version of growth for metric spaces given here is taken from [Álvarez and Candel 2015; Álvarez and Wolak 2013].

Recall that, given nondecreasing functions[10] $u, v : [0, \infty) \to [0, \infty)$, it is said that $u$ is *dominated* by $v$, written $u \preccurlyeq v$, when there are $a, b \geq 1$ and $c \geq 0$ such that $u(r) \leq a v(br)$ for all $r \geq c$. If $u \preccurlyeq v \preccurlyeq u$, then it is said that $u$ and $v$ represent the same *growth type* or have *equivalent growth*; this is an equivalence relation, and $\preccurlyeq$ defines a partial order relation between growth types called *domination*. For

[9] A subdefining cocycle is a defining cocycle contained in another one.

[10] Usually, growth types are defined by using nondecreasing functions $\mathbb{Z}^+ \to [0, \infty)$, but nondecreasing functions $[0, \infty) \to [0, \infty)$ give rise to an equivalent concept.

a family of pairs of nondecreasing functions $[0,\infty)\to[0,\infty)$, *equidomination* means that those pairs satisfy the above condition of domination with the same constants $a, b, c$. A family of functions $[0,\infty)\to[0,\infty)$ will be said to have *equiequivalent growth* if they equidominate one another.

For a complete connected Riemannian manifold $L$, the growth type of each mapping $r\mapsto \operatorname{vol} B(x,r)$ is independent of $x$, and is called the *growth type* of $L$. For metric spaces whose bounded sets are finite, a similar definition of *growth type* can be given where the number of points is used in place of the volume.

Let $M$ be a metric space with metric $d$. A *quasilattice* $\Gamma$ of $M$ is a $C$-net of $M$ for some $C\ge 0$ such that, for every $r\ge 0$, there is some $K_r\ge 0$ such that $\operatorname{card}(\Gamma\cap B(x,r))\le K_r$ for every $x\in M$. It is said that $M$ is of *coarse bounded geometry* if it has a quasilattice. In this case, the *growth type* of $M$ can be defined as the growth type of any quasilattice $\Gamma$ of $M$; i.e., it is the growth type of the *growth function* $r\mapsto v_\Gamma(x,r)=\operatorname{card}(B(x,r)\cap\Gamma)$ for any $x\in\Gamma$. This definition is independent of $\Gamma$.

For a family of metric spaces, if they satisfy the above condition of coarse bounded geometry with the same constants $C$ and $K_r$, then they are said to have *equicoarse bounded geometry*. If moreover the lattices involved in this condition have growth functions with equiequivalent growth, then these metric spaces are said to have *equiequivalent growth*.

The condition of coarse bounded geometry is satisfied by complete connected Riemannian manifolds of bounded geometry, and also by discrete metric spaces with a uniform upper bound on the number of points in all balls of each given radius [Block and Weinberger 1997]. In those cases, the two given definitions of growth type are equal.

**Lemma 5.1** ([Álvarez and Candel 2009]; see also [Álvarez and Wolak 2013, Lemma 2.1]). *Two coarsely quasi-isometric metric spaces of coarse bounded geometry have the same growth type. Moreover*, *if a family of metric spaces are equicoarsely quasi-isometric to each other*, *then they have equiequivalent growth.*

**5B.** ***Quasi-isometry and growth types of orbits.*** Let $\mathscr{H}$ be a pseudogroup on a space $T$, and $E$ a symmetric set of generators of $\mathscr{H}$. Let $\mathfrak{G}$ be the groupoid of germs of maps in $\mathscr{H}$.

For each $h\in\mathscr{H}$ and $x\in\operatorname{dom} h$, let $|h|_{E,x}$ be the length of the shortest expression of $\boldsymbol{\gamma}(h,x)$ as a product of germs of maps in $E$ (being 0 if $\boldsymbol{\gamma}(h,x)=\boldsymbol{\gamma}(\mathrm{id}_T,x)$). For each $x\in T$, define metrics $d_E$ on $\mathscr{H}(x)$ and $\mathfrak{G}_x$ by

$$d_E(y,z)=\min\{\,|h|_{E,y}\mid h\in\mathscr{H},\ y\in\operatorname{dom}h,\ h(y)=z\,\},$$
$$d_E(\boldsymbol{\gamma}(f,x),\boldsymbol{\gamma}(g,x))=|fg^{-1}|_{E,g(x)}.$$

Notice that

$$d_E(f(x), g(x)) \le d_E(\boldsymbol{\gamma}(f, x), \boldsymbol{\gamma}(g, x)).$$

Moreover, on the germ covers, $d_E$ is right invariant in the sense that, if $y \in \mathcal{H}(x)$, the bijection $\mathfrak{G}_y \to \mathfrak{G}_x$, given by right multiplication with any element in $\mathfrak{G}_x^y$, is isometric; so the isometry types of the germ covers of the orbits make sense without any reference to base points. In fact, the definition of $d_E$ on $\mathfrak{G}_x$ is analogous to the definition of the right invariant metric $d_S$ on a group $\Gamma$ induced by a symmetric set of generators $S$: $d_S(\gamma, \delta) = |\gamma\delta^{-1}|$ for $\gamma, \delta \in \Gamma$, where $|\gamma|$ is the length of the shortest expression of $\gamma$ as a product of elements of $S$ (being 0 if $\gamma = e$).

Assume that $\mathcal{H}$ is compactly generated and $T$ locally compact. Let $U \subset T$ be a relatively compact open subset that meets all $\mathcal{H}$-orbits, let $\mathcal{G} = \mathcal{H}|_U$, and let $E$ be a symmetric system of compact generation of $\mathcal{H}$ on $U$. With these conditions, the quasi-isometry type of the $\mathcal{G}$-orbits with $d_E$ may depend on $E$ [Álvarez and Candel 2009, Section 6]. So the following additional condition on $E$ is considered.

**Definition 5.2** [Álvarez and Candel 2009, Definition 4.2]. With the above notation, it is said that $E$ is *recurrent* if, for any relatively compact open subset $V \subset U$ that meets all $\mathcal{G}$-orbits, there exists some $R > 0$ such that $\mathcal{G}(x) \cap V$ is an $R$-net in $\mathcal{G}(x)$ with $d_E$ for all $x \in U$.

The role played by $V$ in Definition 5.2 can be played by any relatively compact open subset meeting all orbits [Álvarez and Candel 2009, Lemma 4.3]. Furthermore there exists a recurrent system of compact generation on $U$ [Álvarez and Candel 2009, Corollary 4.5].

**Theorem 5.3** [Álvarez and Candel 2009, Theorem 4.6]. *Let $\mathcal{H}$ and $\mathcal{H}'$ be compactly generated pseudogroups on locally compact spaces $T$ and $T'$, let $U$ and $U'$ be relatively compact open subsets of $T$ and $T'$ that meet all orbits of $\mathcal{H}$ and $\mathcal{H}'$, let $\mathcal{G}$ and $\mathcal{G}'$ denote the restrictions of $\mathcal{H}$ and $\mathcal{H}'$ to $U$ and $U'$, and let $E$ and $E'$ be recurrent symmetric systems of compact generation of $\mathcal{H}$ and $\mathcal{H}'$ on $U$ and $U'$, respectively. Suppose that there exists an equivalence $\mathcal{H} \to \mathcal{H}'$, and consider the induced equivalence $\mathcal{G} \to \mathcal{G}'$ and homeomorphism $U/\mathcal{G} \to U'/\mathcal{G}'$. Then the $\mathcal{G}$-orbits with $d_E$ are equicoarsely quasi-isometric to the corresponding $\mathcal{G}'$-orbits with $d_{E'}$.*

An obvious modification of the arguments of the proof of [Álvarez and Candel 2009, Theorem 4.6] gives the following.

**Theorem 5.4.** *With the notation and conditions of Theorem 5.3, the germ covers of the $\mathcal{G}$-orbits with $d_E$ are equicoarsely quasi-isometric to the germ covers of the corresponding $\mathcal{G}'$-orbits with $d_{E'}$.*

**Corollary 5.5.** *With the notation and conditions of Theorem 5.3, the corresponding orbits of $\mathcal{G}$ and $\mathcal{G}'$, as well as their germ covers, have equiequivalent growth.*

*Proof.* This follows from Lemma 5.1 and Theorems 5.3 and 5.4. □

**Example 5.6.** Let $G$ be a locally compact Polish local group with a left-invariant metric, let $\Gamma \subset G$ be a dense finitely generated sub-local group, and let $\mathcal{H}$ denote the pseudogroup generated by the local action of $\Gamma$ on $G$ by local left translations. Suppose that $\mathcal{H}$ is compactly generated, and let $\mathcal{G} = \mathcal{H}|_U$ for some relatively compact open neighborhood $U$ of the identity element $e$ in $G$, which meets all $\mathcal{H}$-orbits because $\Gamma$ is dense. For every $\gamma \in \Gamma$ with $\gamma U \cap U \neq \varnothing$, let $h_\gamma$ denote the restriction $U \cap \gamma^{-1} U \to \gamma U \cap U$ of the local left translation by $\gamma$. There is a finite symmetric set $S = \{s_1, \dots, s_k\} \subset \Gamma$ such that $E = \{h_{s_1}, \dots, h_{s_k}\}$ is a recurrent system of compact generation of $\mathcal{H}$ on $U$; in fact, by reducing $\Gamma$ if necessary, we can assume that $S$ generates $\Gamma$. The recurrence of $E$ means that there is some $N \in \mathbb{N}$ such that

$$U = \bigcup_{h \in E^N} h^{-1}(V \cap \operatorname{im} h), \tag{20}$$

where $E^N$ is the family of compositions of at most $N$ elements of $E$.

For each $x \in U$, let

$$\Gamma_{U,x} = \{\gamma \in \Gamma \mid \gamma x \in U\}.$$

Let $\mathfrak{G}$ denote the topological groupoid of germs of $\mathcal{G}$. The map $\Gamma_{U,x} \to \mathfrak{G}_x$, $\gamma \mapsto \boldsymbol{\gamma}(h_\gamma, x)$ is bijective. For $\gamma \in \Gamma_{U,x}$, let $|\gamma|_{S,U,x} := |h_\gamma|_{E,x}$. Thus $|e|_{S,U,x} = 0$, and if $\gamma \neq e$, then $|\gamma|_{S,U,x}$ equals the minimum $n \in \mathbb{N}$ such that there are indices $i_1, \dots, i_n \in \{1, \dots, k\}$ with $\gamma = s_{i_n} \cdots s_{i_1}$ and $s_{i_m} \cdots s_{i_1} \cdot x \in U$ for all $1 \le m \le n$. Moreover $d_E$ on $\mathfrak{G}_x$ corresponds to the metric $d_{S,U,x}$ on $\Gamma_{U,x}$ given by

$$d_{S,U,x}(\gamma, \delta) = |\delta\gamma^{-1}|_{S,U,\gamma(x)}.$$

Observe that, for all $\gamma \in \Gamma_{U,x}$ and $\delta \in \Gamma_{U,\gamma\cdot x}$,

$$\delta\gamma \in \Gamma_{U,x}, \qquad |\delta\gamma|_{S,U,x} \le |\gamma|_{S,U,x} + |\delta|_{S,U,\gamma\cdot x}, \tag{21}$$

$$\gamma^{-1} \in \Gamma_{U,\gamma\cdot x}, \qquad |\gamma|_{S,U,x} = |\gamma^{-1}|_{S,U,\gamma\cdot x}. \tag{22}$$

In this example, we will be interested on the growth type of the orbits of $\mathcal{G}$ with $d_E$, or, equivalently, the growth type of the metric spaces $(\Gamma_{U,x}, d_{S,U,x})$. The following result was used by Breuillard and Gelander to study this growth type when $G$ is a Lie group.

**Proposition 5.7** [Breuillard and Gelander 2007, Proposition 10.5]. *Let $G$ be a nonnilpotent connected real Lie group and $\Gamma$ a finitely generated dense subgroup. For any finite set $S = \{s_1, \dots, s_k\}$ of generators of $\Gamma$, and any neighborhood $B$ of $e$ in $G$, there are elements $t_i \in \Gamma \cap s_i B$ $(i \in \{1, \dots, k\})$ which freely generate a free semigroup. If $G$ is not solvable, then we can choose the elements $t_i$ so that they generate a free group.*

**5C.** ***Growth of equicontinuous pseudogroups.*** Let $G$ be a locally compact Polish local group with a left-invariant metric, let $\Gamma \subset G$ be a dense finitely generated sub-local group, and let $\mathcal{H}$ denote the pseudogroup generated by the local action of $\Gamma$ on $G$ by local left translations. Suppose that $\mathcal{H}$ is compactly generated. Let $\mathcal{G} = \mathcal{H}|_U$ for some relatively compact open neighborhood $U$ of the identity element $e$ in $G$, which meets all $\mathcal{H}$-orbits because $\Gamma$ is dense. Let $E$ be a recurrent symmetric system of compact generation of $\mathcal{H}$ on $U$. Let $\mathfrak{G}$ be the groupoid of germs of maps in $\mathcal{G}$.

**Theorem 5.8.** *With the above notation and conditions*, *one of the following properties hold*:

- *$G$ can be approximated by nilpotent local Lie groups*; *or*
- *the germ covers of all $\mathcal{G}$-orbits have exponential growth with $d_E$.*

*Proof.* By Theorem 2.26, there is some $U_0 \in \Psi G$, contained in any given element of $\Psi G \cap \Phi(G, 2)$, and there exists a sequence of compact normal subgroups $F_n \subset U_0$ such that $F_{n+1} \subset F_n$, $\bigcap_n F_n = \{e\}$, $(F_n, U_0) \in \Delta G$, and $G/(F_n, U_0)$ is a local Lie group. Let $T_n : U_0^2 \to G/(F_n, U_0)$ denote the canonical projection. Take an open neighborhood $U_1$ of $e$ such that $\overline{U_1} \subset U_0$. Then $F_n\overline{U_1} \subset U_0$ for $n$ large enough by the properties of the sequence $F_n$. Let $U_2 = F_n U_1$ for such an $n$. Thus $U_2$ is saturated by the fibers of $T_n$, and $\overline{U_2} \subset U_0$. Then $U' := T_n(U_2)$ is a relatively compact open neighborhood of the identity in the local Lie group $G' := G/(F_n, U_0)$. Let $\Gamma' = T_n(\Gamma \cap U_0^2)$, which is a dense sub-local group of $G'$, and let $\mathcal{H}'$ denote the pseudogroup on $G'$ generated by the local action of $\Gamma'$ by local left translations.

For every $\gamma \in \Gamma \cap U_0$ for which $\gamma U_2 \cap U_2 \neq \varnothing$, let $h_\gamma$ denote the restriction $U_2 \cap \gamma^{-1} U_2 \to \gamma U_2 \cap U_2$ of the local left translation by $\gamma$. There is a finite symmetric set $S = \{s_1, \dots, s_k\} \subset \Gamma$ such that $E_2 = \{h_{s_1}, \dots, h_{s_k}\}$ is a recurrent system of compact generation of $\mathcal{H}$ on $U_2$. By reducing $\Gamma$ if necessary, we can suppose that $S$ generates $\Gamma$. For every $\delta \in \Gamma'$ with $\delta U' \cap U' \neq \varnothing$, let $h'_\delta$ denote the restriction $U' \cap \delta^{-1} U' \to \delta U' \cap U'$ of the local left translation by $\delta$. We can assume that $s_1, \dots, s_k$ are in $U_2$, and therefore we can consider their images $s'_1, \dots, s'_k$ by $T_n$. Moreover each $h_{s_i}$ induces via $T_n$ the map $h'_{s_i}$, and $E' = \{h'_{s_1}, \dots, h'_{s_k}\}$ is a system of compact generation of $\mathcal{H}'$ on $U'$. By increasing $E_2$ if necessary, we can assume that $E'$ is also recurrent. Fix any open set $V'$ in $G'$ with $\overline{V'} \subset U'$. Then $V = T_n^{-1}(V')$ satisfies $\overline{V} \subset U_2$.

**Claim 1.** *For each finite subset $F \subset \Gamma \cap U_2$, we have $U_2 \subset \bigcup_{\gamma \in \Gamma \setminus F} \gamma V$.*

Since $U_2$ and $V$ are saturated by the fibers of $T_n$, Claim 1 follows by showing that $U' \subset \bigcup_{\gamma \in \Gamma' \setminus F'} \gamma V'$, where $F' = T_n(F)$. Suppose that this inclusion is false. Then there is some finite symmetric subset $F \subset \Gamma \cap U_2$ and some $x \in U'$ such that $((\Gamma' \setminus F')x) \cap V' = \varnothing$. By the recurrence of $E'$, there is some $N \in \mathbb{N}$ satisfying (20) with $U'$ and $E'$. Since $\Gamma'_{U',x}$ is infinite because $\Gamma'$ is dense in $G'$, it follows that there is some $\gamma \in \Gamma'_{U',x} \setminus F'$ such that

$$|\gamma|_{S',U',x} > N + \max\{|\epsilon|_{S',U',x} \mid \epsilon \in F' \cap \Gamma'_{U',x}\}. \tag{23}$$

By (20), there is some $h \in E'^N$ such that $\gamma x \in h^{-1}(V' \cap \operatorname{im} h')$. We have $h = h'_\delta$ for some $\delta \in \Gamma'$. Note that $\delta \in \Gamma'_{U',\gamma' x}$ and $|\delta|_{S',U',\gamma x} \leq N$. Hence

$$|\gamma|_{S',U',x} \leq |\delta\gamma|_{S',U',x} + |\delta^{-1}|_{S',U',\delta\gamma x} = |\delta\gamma|_{S',U',x} + |\delta|_{S',U',\gamma' x} \leq |\delta\gamma|_{S',U',x} + N$$

by (21) and (22), obtaining that $\delta\gamma \notin F'$ by (23). However, $\delta\gamma x \in V'$, obtaining a contradiction, which completes the proof of Claim 1.

**Claim 2.** *For each finite subset $F \subset \Gamma \cap U_2$, we have $\overline{U_2} \subset \bigcup_{\gamma \in \Gamma \setminus F} \gamma V$.*

Take a relatively compact open subset $O_1 \subset G$ such that $\overline{U_1} \subset O_1$ and $F_n \overline{O_1} \subset U_0$. Let $O_2 = F_n O_1$ and $\mathcal{H} = \mathcal{H}|_{O_2}$. Then Claim 2 follows by applying Claim 1 to $O_2$.

According to Claim 2, by increasing $S$ if necessary, we can suppose that

$$\overline{U_2} \subset \bigcup_{i<j} (s_i \cdot V \cap s_j \cdot V) = \bigcup_{i<j} (s_i^{-1} \cdot V \cap s_j^{-1} \cdot V). \tag{24}$$

Suppose that $G$ cannot be approximated by nilpotent local Lie groups. Then we can assume that the local Lie group $G'$ is not nilpotent. Moreover we can suppose that $G'$ is a sub-local Lie group of a simply connected Lie group $L$. Let $\Delta$ be the dense subgroup of $L$ whose intersection with $G'$ is $\Gamma'$. Then, by Proposition 5.7, there are elements $t'_1, \dots, t'_k$ in $\Delta$, as close as desired to $s'_1, \dots, s'_k$, which are free generators of a free semigroup. If the elements $t'_i$ are close enough to $s'_i$, then they are in $U'$. So there are elements $t_i \in U_2$ such that $T_n(t_i) = t'_i$. By the compactness of $\overline{U_2}$, and because $U_2$ and $V$ are saturated by the fibers of $T_n$, if $t'_1, \dots, t'_k$ are close enough to $s'_1, \dots, s'_k$, then (24) gives

$$\overline{U_2} \subset \bigcup_{i<j} (t_i^{-1} V \cap t_j^{-1} V). \tag{25}$$

Now, we adapt the argument of the proof of [Breuillard and Gelander 2007, Lemma 10.6]. Let $\widehat{\Gamma} \subset \Gamma$ be the sub-local group generated by $t_1, \dots, t_k$; thus $\widehat{S} = \{t_1^{\pm 1}, \dots, t_k^{\pm 1}\}$ is a symmetric set of generators of $\widehat{\Gamma}$, and $S \cup \widehat{S}$ is a symmetric set of generators of $\Gamma$. With $\widehat{E} = \{h_{t_1}^{\pm 1}, \dots, h_{t_k}^{\pm 1}\}$, observe that $E_2 \cup \widehat{E}$ is a recurrent system of compact generation of $\mathcal{H}$ on $U_2$. Given $x \in U_2$, let $\mathsf{S}(n)$ be the sphere with center $e$ and radius $n \in \mathbb{N}$ in $\widehat{\Gamma}_{U_2,x}$ with $d_{\widehat{S},U,x}$. By (25), for each $\gamma \in \mathsf{S}(n)$, there are indices $i < j$ such that $\gamma x \in t_i^{-1} V \cap t_j^{-1} V$. So the points $t_i \gamma x$ and $t_j \gamma x$ are in $V$, obtaining that $t_i \gamma, t_j \gamma \in \mathsf{S}(n+1)$. Moreover all elements obtained in this way from elements of $\mathsf{S}(n)$ are pairwise distinct because $t'_1, \dots, t'_k$ freely generate a free semigroup. Hence $\operatorname{card}(\mathsf{S}(n+1)) \geq 2 \operatorname{card}(\mathsf{S}(n))$, giving $\operatorname{card}(\mathsf{S}(n)) \geq 2^n$. So $(\widehat{\Gamma}_{U_2,x}, d_{\widehat{S},U_2,x})$ has exponential growth. Since $\widehat{\Gamma}_{U_2,x} \subset \Gamma_{U_2,x}$

and $d_{S\cup\widehat{S},U_2,x} \le d_{\widehat{S},U_2,x}$ on $\widehat{\Gamma}_{U_2,x}$, it follows that $(\Gamma_{U_2,x}, d_{S\cup\widehat{S},U_2,x})$ also has exponential growth. So $(\mathfrak{G}_x, d_{E_2\cup\widehat{E}})$ has exponential growth, obtaining that $(\mathfrak{G}_x, d_E)$ has exponential growth by Corollary 5.5. $\square$

**5D. *Growth of equicontinuous foliated spaces.*** Let $X \equiv (X, \mathcal{F})$ be a compact Polish foliated space. Let $\{U_i, p_i, h_{ij}\}$ be a defining cocycle of $\mathcal{F}$, where $p_i : U_i \to T_i$ and $h_{ij} : T_{ij} \to T_{ji}$, and let $\mathcal{H}$ be the induced representative of the holonomy pseudogroup. As we saw in Section 4A, $\mathcal{H}$ can be considered as the restriction of some compactly generated pseudogroup $\mathcal{H}'$ to some relatively compact open subset, and $E = \{h_{ij}\}$ is a system of compact generation on $T$. Moreover, Álvarez and Candel [2009] observed that $E$ is recurrent. According to Theorems 5.3 and 5.4, it follows that the quasi-isometry type of the $\mathcal{H}$-orbits and their germ covers with $d_E$ are independent of the choice of $\{U_i, p_i, h_{ij}\}$ under the above conditions; thus they can be considered as quasi-isometry types of the corresponding leaves and their holonomy covers.

This has the following interpretation when $X$ is a smooth manifold. In this case, given any Riemannian metric $g$ on $X$, for each leaf $L$, the differentiable (and coarse) quasi-isometry type of $g|_L$ is independent of the choice of $g$; they depend only on $\mathcal{F}$ and $L$. In fact, it is coarsely quasi-isometric to the corresponding $\mathcal{H}$-orbit, and therefore they have the same growth type [Carrière 1988] (this is an easy consequence of the existence of a uniform bound of the diameter of the plaques). Similarly, the germ covers of the $\mathcal{H}$-orbits are also quasi-isometric to the holonomy covers of the corresponding leaves.

Theorem B follows from these observations and Theorem 5.8.

## 6. Examples and open problems

Theorems A and B may be relevant in the following examples; most of them are taken from [Candel and Conlon 2000, Chapter 11].

**Example 6.1.** Any locally free action of a connected Lie group on a locally compact Polish space, $\phi : H \times X \to X$, defines a foliated structure $\mathcal{F}$ on $X$ whose leaves are the orbits [Candel and Conlon 2000, Theorem 11.3.14; Palais 1961]. Moreover $\mathcal{F}$ is equicontinuous if $\phi$ is equicontinuous.

**Example 6.2.** A *matchbox manifold* is a foliated continuum[11] $X \equiv (X, \mathcal{F})$ transversely modeled on a totally disconnected space. The case of a single leaf is discarded, and it is assumed that $X$ is $C^1$ in the sense that the changes of foliated coordinates are $C^1$ along the leaves, with transversely continuous leafwise derivatives. An example of matchbox manifold is given by any inverse limit of smooth proper covering maps of compact $n$-manifolds, called an $n$-dimensional

[11] Recall that a *continuum* is a nonempty compact connected metrizable space.

*solenoid*; if moreover any composite of a finite number of bounding maps is a normal covering, then it is called a *McCord solenoid*. A matchbox manifold $X$ is equicontinuous if and only if it is a solenoid [Clark and Hurder 2013, Theorem 7.9]; and $X$ is homogeneous if and only if it is a McCord solenoid [Clark and Hurder 2013, Theorem 1.1]; this is the case where it is a $G$-foliated space. See [Alcalde Cuesta et al. 2011] for a generalization using inverse limits of compact branched manifolds.

**Example 6.3.** Let $C_b(\mathbb{R})$ be the space of continuous bounded functions $\mathbb{R} \to \mathbb{R}$, with the topology of uniform convergence. For a function $f \in C_b(\mathbb{R})$ and $t \in \mathbb{R}$, let $f_t \in C_b(\mathbb{R})$ be defined by $f_t(r) = f(r+t)$. It is said that $f$ is *almost periodic* if $\{f_t \mid t \in \mathbb{R}\}$ is equicontinuous [Besicovitch 1955; Gottschalk 1946], which means that $\mathfrak{M}(f) := \overline{\{f_t \mid t \in \mathbb{R}\}}$ is compact in $C_b(\mathbb{R})$. An equicontinuous flow

$$\Phi : \mathbb{R} \times \mathfrak{M}(f) \to \mathfrak{M}(f)$$

is defined by $\Phi_t(g) = g_t$. If $f$ is nonconstant, then $\Phi$ is nonsingular, defining an equicontinuous foliated structure $\mathcal{F}$ on $\mathfrak{M}(f)$. If $f$ is nonperiodic, then $\mathcal{F}$ does not reduce to a single leaf. With more generality, we can consider an almost-periodic nonperiodic continuous function $f$ on any connected Lie group with values in a Hilbert space.

**Example 6.4.** For each $n \in \mathbb{Z}_+$, let $\mathcal{M}_*(n)$ denote the set[12] of isometry classes $[M, x]$ of pointed complete connected Riemannian $n$-manifolds $(M, x)$. The $C^\infty$ convergence [Petersen 1998] defines a Polish topology on $\mathcal{M}_*(n)$ [Álvarez et al. 2016, Theorem 1.1]. The corresponding space is denoted by $\mathcal{M}_*^\infty(n)$, and its closure operator by $\mathrm{Cl}_\infty$. For any complete connected Riemannian manifold $M \equiv (M, g)$, a canonical continuous map $\iota : M \to \mathcal{M}_*^\infty(n)$ is defined by $\iota(x) = [M, x]$. A concept of *weak aperiodicity* of $M$ was introduced in [Álvarez et al. 2016]. On the other hand, $M$ is called *almost periodic* if, for all $m \in \mathbb{N}$, $\epsilon > 0$ and $x \in M$, there is a set $H$ of diffeomorphisms of $M$ such that $\sup |\nabla^k h^* g| < \epsilon$ for all $h \in H$ and $k \le m$, and $\{h(x) \mid h \in H\}$ is a net in $M$. If $M$ is weakly aperiodic and almost periodic, then $\mathrm{Cl}_\infty(\iota(M))$ canonically becomes a compact minimal equicontinuous foliated space of dimension $n$, as follows from [Álvarez et al. 2016, Theorem 1.2 and Lemma 12.5(ii); Lessa 2015, Lemma 7.2 and Theorem 4.1] (see also [Petersen 1998, Chapter 10, Theorem 3.3; Cheeger 1970]).

**Problem 6.5.** In the Examples 6.1–6.4, understand the specific application of Theorems A and B.

**Problem 6.6.** Use Theorem A to classify particular classes of equicontinuous foliated spaces.

[12]The logical problems of this definition can be avoided because any complete connected Riemannian manifold is equipotent to some subset of $\mathbb{R}$.

**Question 6.7.** Is it possible to improve Theorem B for special types of structural local groups?

**Question 6.8.** Is there any consequence of Theorems A and B in usual foliation theory, assuming that the minimal sets are equicontinuous?

The following questions refer to extensions of known properties of Riemannian foliations, where Theorem A could play an important role.

**Question 6.9.** For compact minimal equicontinuous foliated spaces, does the leafwise heat flow of leafwise differential forms preserve transverse continuity at infinite time?

**Question 6.10.** Is it possible to give useful extensions of tautness and tenseness to equicontinuous foliated spaces, and relate them to some kind of basic cohomology?

## Acknowledgement

We thank the referee for suggestions to improve the paper.

## References

[Abd-Allah and Brown 1980] A. M. Abd-Allah and R. Brown, "A compact-open topology on partial maps with open domain", *J. London Math. Soc.* (2) **21**:3 (1980), 480–486. MR 81i:54010 Zbl 0436.54012

[Alcalde Cuesta et al. 2011] F. Alcalde Cuesta, Á. Lozano Rojo, and M. Macho Stadler, "Transversely Cantor laminations as inverse limits", *Proc. Amer. Math. Soc.* **139**:7 (2011), 2615–2630. MR 2784831 Zbl 1229.57029

[Álvarez and Candel 2009] J. A. Álvarez López and A. Candel, "Equicontinuous foliated spaces", *Math. Z.* **263**:4 (2009), 725–774. MR 2010i:53040 Zbl 1177.53026

[Álvarez and Candel 2010] J. A. Álvarez López and A. Candel, "Topological description of Riemannian foliations with dense leaves", *Pacific J. Math.* **248**:2 (2010), 257–276. MR 2012b:53043 Zbl 1207.53033

[Álvarez and Candel 2015] J. A. Álvarez López and A. Candel, "Generic coarse geometry of leaves", preprint, 2015. arXiv 1406.1746

[Álvarez and Masa 2008] J. A. Álvarez López and X. M. Masa, "Morphisms between complete Riemannian pseudogroups", *Topology Appl.* **155**:6 (2008), 544–604. MR 2009c:58028 Zbl 1147.57025

[Álvarez and Wolak 2013] J. A. Álvarez López and R. Wolak, "Growth of some transversely homogeneous foliations", preprint, 2013. arXiv 1304.1939

[Álvarez et al. 2016] J. A. Álvarez López, R. Barral Lijó, and A. Candel, "A universal Riemannian foliated space", *Topology Appl.* **198** (2016), 47–85. MR 3433188 Zbl 06521163

[Arens 1946] R. Arens, "Topologies for homeomorphism groups", *Amer. J. Math.* **68** (1946), 593–610. MR 8,479i Zbl 0061.24306

[Besicovitch 1955] A. S. Besicovitch, *Almost periodic functions*, Dover, New York, 1955. MR 16,817a Zbl 0065.07102

[Block and Weinberger 1997] J. Block and S. Weinberger, "Large scale homology theories and geometry", pp. 522–569 in *Geometric topology* (Athens, GA, 1993), edited by W. H. Kazez, AMS/IP Stud. Adv. Math. **2**, American Mathematical Society, Providence, RI, 1997. MR 98h:57001 Zbl 0898.55006

[Breuillard and Gelander 2007] E. Breuillard and T. Gelander, "A topological Tits alternative", *Ann. of Math.* (2) **166**:2 (2007), 427–474. MR 2009a:20077 Zbl 1149.20039

[Candel and Conlon 2000] A. Candel and L. Conlon, *Foliations I*, Graduate Studies in Mathematics **23**, American Mathematical Society, Providence, RI, 2000. MR 2002f:57058 Zbl 0936.57001

[Carrière 1988] Y. Carrière, "Feuilletages Riemanniens à croissance polynômiale", *Comment. Math. Helv.* **63**:1 (1988), 1–20. MR 89a:57033 Zbl 0661.53022

[Cheeger 1970] J. Cheeger, "Finiteness theorems for Riemannian manifolds", *Amer. J. Math.* **92** (1970), 61–74. MR 41 #7697 Zbl 0194.52902

[Clark and Hurder 2013] A. Clark and S. Hurder, "Homogeneous matchbox manifolds", *Trans. Amer. Math. Soc.* **365**:6 (2013), 3151–3191. MR 3034462 Zbl 1285.57017

[van den Dries and Goldbring 2010] L. van den Dries and I. Goldbring, "Globalizing locally compact local groups", *J. Lie Theory* **20**:3 (2010), 519–524. MR 2011m:22009 Zbl 1203.22006

[van den Dries and Goldbring 2012] L. van den Dries and I. Goldbring, "Erratum to 'Globalizing locally compact local groups'", *J. Lie Theory* **22**:2 (2012), 489–490. MR 2976929 Zbl 1242.22005

[Fedida 1971] E. Fedida, "Sur les feuilletages de Lie", *C. R. Acad. Sci. Paris Sér. A* **272** (1971), 999–1001. MR 44 #2249 Zbl 0218.57014

[Fedida 1978] E. Fedida, "Sur la théorie des feuilletages associée au repère mobile: cas des feuilletages de Lie", pp. 183–195 in *Differential topology, foliations and Gelfand–Fuks cohomology* (Rio de Janeiro, 1976), edited by P. A. Schweitzer, Lecture Notes in Math. **652**, Springer, Berlin, 1978. MR 80a:57012 Zbl 0393.57006

[Ghys 1985] É. Ghys, "Groupes d'holonomie des feuilletages de Lie", *Indag. Math.* **88**:2 (1985), 173–182. MR 87d:57021 Zbl 0571.57020

[Ghys 1999] É. Ghys, "Laminations par surfaces de Riemann", pp. 49–95 in *Dynamique et géométrie complexes* (Lyon, 1997), edited by D. Cerveau et al., Panor. Synthèses **8**, Société Mathématique de France, Paris, 1999. MR 2001g:37068 Zbl 1018.37028

[Goldbring 2010] I. Goldbring, "Hilbert's fifth problem for local groups", *Ann. of Math.* (2) **172**:2 (2010), 1269–1314. MR 2011f:22010 Zbl 1219.22004

[Gottschalk 1946] W. H. Gottschalk, "Almost periodicity, equi-continuity and total boundedness", *Bull. Amer. Math. Soc.* **52** (1946), 633–636. MR 8,34b Zbl 0063.01711

[Gromov 1993] M. Gromov, *Geometric group theory, 2: Asymptotic invariants of infinite groups* (Brighton, 1991), edited by G. A. Niblo and M. A. Roller, London Math. Soc. Lecture Note Ser. **182**, Cambridge University Press, 1993. MR 95m:20041 Zbl 0841.20039

[Haefliger 1985] A. Haefliger, "Pseudogroups of local isometries", pp. 174–197 in *Differential geometry* (Santiago de Compostela, 1984), edited by L. A. Cordero, Res. Notes in Math. **131**, Pitman, Boston, 1985. MR 88i:58174 Zbl 0656.58042

[Haefliger 1988] A. Haefliger, "Leaf closures in Riemannian foliations", pp. 3–32 in *A fête of topology*, edited by Y. Matsumoto et al., Academic Press, Boston, 1988. MR 89c:57033 Zbl 0667.57012

[Haefliger 2002] A. Haefliger, "Foliations and compactly generated pseudogroups", pp. 275–295 in *Foliations: geometry and dynamics* (Warsaw, 2000), edited by P. Walczak et al., World Scientific, River Edge, NJ, 2002. MR 2003g:58029 Zbl 1002.57059

[Hector and Hirsch 1981] G. Hector and U. Hirsch, *Introduction to the geometry of foliations, A: Foliations on compact surfaces, fundamentals for arbitrary codimension, and holonomy*, Aspects of Mathematics **1**, Vieweg, Braunschweig, 1981. MR 83d:57019 Zbl 0486.57002

[Hector and Hirsch 1987] G. Hector and U. Hirsch, *Introduction to the geometry of foliations, B: Foliations of codimension one*, 2nd ed., Aspects of Mathematics **3**, Vieweg, Braunschweig, 1987. MR 92f:57037 Zbl 0704.57001

[Jacoby 1957] R. Jacoby, "Some theorems on the structure of locally compact local groups", *Ann. of Math.* (2) **66** (1957), 36–69. MR 19,751c Zbl 0084.03202

[Kechris 1991] A. S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics **156**, Springer, Berlin, 1991. MR 96e:03057 Zbl 0819.04002

[Kellum 1993] M. Kellum, "Uniformly quasi-isometric foliations", *Ergodic Theory Dynam. Systems* **13**:1 (1993), 101–122. MR 94c:57048 Zbl 0784.58060

[Lessa 2015] P. Lessa, "Reeb stability and the Gromov–Hausdorff limits of leaves in compact foliations", *Asian J. Math.* **19**:3 (2015), 433–463. MR 3361278 Zbl 06470466

[Matsumoto 2010] S. Matsumoto, "The unique ergodicity of equicontinuous laminations", *Hokkaido Math. J.* **39**:3 (2010), 389–403. MR 2011m:37038 Zbl 1213.37044

[Meigniez 1995] G. Meigniez, "Feuilletages de Lie résolubles", *Ann. Fac. Sci. Toulouse Math.* (6) **4**:4 (1995), 801–817. MR 99e:57042 Zbl 0867.57023

[Meigniez 2010] G. Meigniez, "A compactly generated pseudogroup which is not realizable", *J. Math. Soc. Japan* **62**:4 (2010), 1205–1218. MR 2011m:58028 Zbl 1207.57038

[Molino 1982] P. Molino, "Géométrie globale des feuilletages Riemanniens", *Indag. Math.* **85**:1 (1982), 45–76. MR 84j:53043 Zbl 0516.57016

[Molino 1988] P. Molino, *Riemannian foliations*, Progress in Mathematics **73**, Birkhäuser, Boston, 1988. MR 89b:53054 Zbl 0633.53001

[Montgomery and Zippin 1955] D. Montgomery and L. Zippin, *Topological transformation groups*, Interscience Publishers, New York, 1955. MR 17,383b Zbl 0068.01904

[Moore and Schochet 1988] C. C. Moore and C. Schochet, *Global analysis on foliated spaces*, Mathematical Sciences Research Institute Publications **9**, Springer, New York, 1988. MR 89h:58184 Zbl 0648.58034

[Olver 1996] P. J. Olver, "Non-associative local Lie groups", *J. Lie Theory* **6**:1 (1996), 23–51. MR 98d:22005 Zbl 0862.22005

[Palais 1961] R. S. Palais, "On the existence of slices for actions of non-compact Lie groups", *Ann. of Math.* (2) **73** (1961), 295–323. MR 23 #A3802 Zbl 0103.01802

[Petersen 1998] P. Petersen, *Riemannian geometry*, Graduate Texts in Mathematics **171**, Springer, New York, 1998. MR 98m:53001 Zbl 0914.53001

[Plaut 1993] C. Plaut, "Associativity and the local version of Hilbert's fifth problem", Notes, University of Tennessee, 1993.

[Reinhart 1959] B. L. Reinhart, "Foliated manifolds with bundle-like metrics", *Ann. of Math.* (2) **69** (1959), 119–132. MR 21 #6004 Zbl 0122.16604

[Sacksteder 1965] R. Sacksteder, "Foliations and pseudogroups", *Amer. J. Math.* **87** (1965), 79–102. MR 30 #4268 Zbl 0136.20903

[Salem 1988] É. Salem, "Une généralisation du théorème de Myers–Steenrod aux pseudogroupes d'isométries", *Ann. Inst. Fourier* (*Grenoble*) **38**:2 (1988), 185–200. MR 89i:58166 Zbl 0613.58041

[Tarquini 2004] C. Tarquini, "Feuilletages conformes", *Ann. Inst. Fourier* (*Grenoble*) **54**:2 (2004), 453–480. MR 2005f:53040 Zbl 1064.53014

Received May 22, 2014. Revised November 9, 2014.

Jesús A. Álvarez López
Departamento de Xeometría e Topoloxía
Facultade de Matemáticas
Universidade de Santiago de Compostela
15782 Santiago de Compostela
Spain

jesus.alvarez@usc.es

Manuel F. Moreira Galicia
Departamento de Xeometría e Topoloxía
Facultade de Matemáticas
Universidade de Santiago de Compostela
15782 Santiago de Compostela
Spain

manuel.moreira@usc.es

morgal2002@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.315

# EQUIVARIANT PRINCIPAL BUNDLES AND LOGARITHMIC CONNECTIONS ON TORIC VARIETIES

INDRANIL BISWAS, ARIJIT DEY AND MAINAK PODDAR

**Let $M$ be a smooth complex projective toric variety equipped with an action of a torus $T$, such that the complement $D$ of the open $T$-orbit in $M$ is a simple normal crossing divisor. Let $G$ be a complex reductive affine algebraic group. We prove that an algebraic principal $G$-bundle $E_G \to M$ admits a $T$-equivariant structure if and only if $E_G$ admits a logarithmic connection singular over $D$. If $E_H \to M$ is a $T$-equivariant algebraic principal $H$-bundle, where $H$ is any complex affine algebraic group, then $E_H$ in fact has a canonical integrable logarithmic connection singular over $D$.**

## 1. Introduction

Our aim is to give characterizations of the equivariant principal bundles on smooth complex projective toric varieties.

Let $M$ be a smooth complex projective toric variety equipped with an action

$$\rho : T \times M \to M$$

of a torus $T$. For any point $t \in T$, define the automorphism

$$\rho_t : M \to M, \quad x \mapsto \rho(t, x).$$

We assume that the complement $D$ of the open $T$-orbit in $M$ is a simple normal crossing divisor.

Let $G$ be a complex reductive affine algebraic group, and let $E_G$ be an algebraic principal $G$-bundle on $M$. In Proposition 4.1 we prove the following:

*The principal $G$-bundle $E_G$ admits a $T$-equivariant structure if and only if the pulled-back principal $G$-bundle $\rho_t^* E_G$ is isomorphic to $E_G$ for every $t \in T$.*

When $G = \mathrm{GL}(n, \mathbb{C})$, this result was proved by Klyachko [1989, p. 342, Proposition 1.2.1].

---

Biswas is supported by a J. C. Bose Fellowship. Poddar is supported by a FAPA grant from Universidad de los Andes.
*MSC2010:* primary 14L30, 14M27; secondary 14M17.
*Keywords:* smooth toric variety, logarithmic connection, equivariant principal bundle.

Using the above characterization of $T$-equivariant principal $G$-bundles on $M$, we prove the following (see Theorem 4.2):

*The principal $G$-bundle $E_G$ admits a logarithmic connection singular over $D$ if and only if $E_G$ admits a $T$-equivariant structure.*

The "if" part of Theorem 4.2 does not require $G$ to be reductive. More precisely, any $T$-equivariant principal $H$-bundle $E_H \to M$, where $H$ is any complex affine algebraic group, admits a canonical integrable logarithmic connection singular over $D$ (see Proposition 3.2).

## 2. Equivariant bundles

Let $\mathbb{G}_m = \mathbb{C} \setminus \{0\}$ be the multiplicative group. Take a complex algebraic group $T$ which is isomorphic to a product of copies of $\mathbb{G}_m$. Let $M$ be a smooth irreducible complex projective variety equipped with an algebraic action of $T$

$$\rho : T \times M \to M \tag{2-1}$$

such that

- there is a Zariski open dense subset $M^0 \subset M$ with $\rho(T, M^0) = M^0$,
- the action of $T$ on $M^0$ is free and transitive, and
- the complement $M \setminus M^0$ is a simple normal crossing divisor of $M$.

In particular, $M$ is a smooth projective toric variety. Note that $M^0$ is the unique $T$-orbit in $M$ with trivial isotropy.

Let $G$ be a connected complex affine algebraic group. A *$T$-equivariant* principal $G$-bundle on $M$ is a pair $(E_G, \tilde{\rho})$, where

$$p : E_G \to M$$

is an algebraic principal $G$-bundle, and

$$\tilde{\rho} : T \times E_G \to E_G$$

is an algebraic action of $T$ on the total space of $E_G$, such that

- $p \circ \tilde{\rho} = \rho \circ (\mathrm{Id}_T \times p)$, where $\rho$ is the action in (2-1), and
- the actions of $T$ and $G$ on $E_G$ commute.

Fix a point $x_0 \in M^0 \subset M$. Let

$$\iota : \rho(T, x_0) = M^0 \hookrightarrow M \tag{2-2}$$

be the inclusion map. Let $M^0 \times G$ be the trivial principal $G$-bundle on $M^0$. It has a tautological integrable algebraic connection given by its trivialization.

Let $(E_G, \tilde{\rho})$ be a $T$-equivariant principal $G$-bundle on $M$. Fix a point $z_0 \in (E_G)_{x_0}$. Using $z_0$, the action $\tilde{\rho}$ produces an isomorphism of principal $G$-bundles between $M^0 \times G$ and the restriction $E_G|_{M^0}$. This isomorphism of principal $G$-bundles is uniquely determined by the following two conditions:

- this isomorphism is $T$-equivariant (the action of $T$ on $M^0 \times G$ is given by the action of $T$ on $M^0$), and
- it takes the point $z_0 \in E_G$ to $(x_0, e) \in M^0 \times G$.

Using this trivialization of $E_G|_{M^0}$, the tautological integrable algebraic connection on $M^0 \times G$ produces an integrable algebraic connection $\mathcal{D}^0$ on $E_G|_{M^0}$. We note that the connection $\mathcal{D}^0$ is independent of the choice of the points $x_0$ and $z_0$. Indeed, the flat sections for $\mathcal{D}^0$ are precisely the orbits of $T$ in $E_G|_{M^0}$. Note that this description of $\mathcal{D}^0$ does not require choosing base points in $M^0$ and $E_G|_{M^0}$.

In Proposition 3.2 it will be shown that $\mathcal{D}^0$ extends to a logarithmic connection on $E_G$ over $M$ singular over the simple normal crossing divisor $M \setminus M^0$.

## 3. Logarithmic connections

***A canonical trivialization.*** The Lie algebra of $T$ will be denoted by $\mathfrak{t}$. Let

$$\mathcal{V} := M \times \mathfrak{t} \to M \tag{3-1}$$

be the trivial vector bundle with fiber $\mathfrak{t}$. The holomorphic tangent bundle of $M$ will be denoted by $TM$. Consider the action of $T$ on $M$ in (2-1). It produces a homomorphism of $\mathcal{O}_M$-coherent sheaves

$$\beta : \mathcal{V} \to TM. \tag{3-2}$$

Let

$$D := M \setminus M^0$$

be the simple normal crossing divisor of $M$. Let

$$TM(-\log D) \subset TM \tag{3-3}$$

be the corresponding logarithmic tangent bundle. Recall that $TM(-\log D)$ is characterized as the maximal coherent subsheaf of $TM$ that preserves $\mathcal{O}_M(-D) \subset \mathcal{O}_M$ for the derivation action of $TM$ on $\mathcal{O}_M$.

**Lemma 3.1.**

(1) *The image of $\beta$ in* (3-2) *is contained in the subsheaf $TM(-\log D) \subset TM$.*

(2) *The resulting homomorphism $\beta : \mathcal{V} \to TM(-\log D)$ is an isomorphism.*

*Proof.* The divisor $D$ is preserved by the action of $T$ on $M$. Therefore, the action of $T$ on $\mathcal{O}_M$, given by the action of $T$ on $M$, preserves the subsheaf $\mathcal{O}_M(-D)$. From this it follows immediately that the subsheaf $\mathcal{O}_M(-D) \subset \mathcal{O}_M$ is preserved by the derivation action of the subsheaf

$$\beta(\mathcal{V}) \subset TM.$$

Therefore, we conclude that $\beta(\mathcal{V}) \subset TM(-\log D)$.

It is known that the vector bundle $TM(-\log D)$ is holomorphically trivial. This follows from Proposition 2 in [Fulton 1993, p. 87], which says that $\Omega^1_M(\log D)$ is holomorphically trivial, together with the equality $\Omega^1_M(\log D)^* = TM(-\log D)$.

So, both $\mathcal{V}$ and $TM(-\log D)$ are trivial vector bundles, and $\beta$ is a homomorphism between them which is an isomorphism over the open subset $M^0$. From this it can be deduced that $\beta$ is an isomorphism over entire $M$. To see this, consider the homomorphism

$$\bigwedge^r \beta : \bigwedge^r \mathcal{V} \to \bigwedge^r TM(-\log D)$$

induced by $\beta$, where $r = \dim_{\mathbb{C}} T = \operatorname{rank}(\mathcal{V})$. So $\bigwedge^r \beta$ is a holomorphic section of the line bundle $\left(\bigwedge^r TM(-\log D)\right) \otimes \left(\bigwedge^r \mathcal{V}\right)^*$. This line bundle is holomorphically trivial because both $\mathcal{V}$ and $TM(-\log D)$ are holomorphically trivial. Fixing a trivialization of $\left(\bigwedge^r TM(-\log D)\right) \otimes \left(\bigwedge^r \mathcal{V}\right)^*$, we consider $\bigwedge^r \beta$ as a holomorphic function on $M$. This function is nowhere vanishing because it does not vanish on $M^0$ and holomorphic functions on $M$ are constants. Since $\bigwedge^r \beta$ is nowhere vanishing, the homomorphism $\beta$ is an isomorphism. □

***A canonical logarithmic connection on equivariant bundles.*** The Lie algebra of $G$ will be denoted by $\mathfrak{g}$.

Let $p : E_G \to M$ be an algebraic principal $G$-bundle. Consider the differential

$$dp : TE_G \to p^* TM, \tag{3-4}$$

where $TE_G$ is the algebraic tangent bundle of $E_G$. The kernel of $dp$ will be denoted by $T_{E_G/M}$. Using the action of $G$ on $E_G$, the subbundle $T_{E_G/M} \subset TE_G$ is identified with the trivial vector bundle over $E_G$ with fiber $\mathfrak{g}$.

The action of $G$ on $E_G$ produces an action of $G$ on $TE_G$. So we get an action of $G$ on the quasicoherent sheaf $p_* TE_G$ on $M$. The invariant part

$$\operatorname{At}(E_G) := (p_* TE_G)^G \subset p_* TE_G$$

is a locally free coherent sheaf; its coherence property follows from the fact that the action of $G$ on the fibers of $p$ is transitive, implying that a $G$-invariant section of $(TE_G)|_{p^{-1}(x)}$, $x \in M$, is uniquely determined by its evaluation at just one point of the fiber $p^{-1}(x)$. Also note that $\operatorname{At}(E_G) = (TE_G)/G$. This $\operatorname{At}(E_G)$ is known as the *Atiyah bundle* for $E_G$. Since $T_{E_G/M}$ is identified with $E_G \times \mathfrak{g}$, the invariant direct image $(p_* T_{E_G/M})^G$ is identified with the adjoint vector bundle

$$\mathrm{ad}(E_G) := E_G \times^G \mathfrak{g} \to M$$

associated to $E_G$ for the adjoint action of $G$ on $\mathfrak{g}$. We note that $\mathrm{ad}(E_G) = T_{E_G/M}/G$. Now the differential $dp$ in (3-4) produces a short exact sequence of holomorphic vector bundles on $M$

$$0 \to \mathrm{ad}(E_G) \to \mathrm{At}(E_G) \xrightarrow{\phi} TM \to 0, \tag{3-5}$$

which is known as the Atiyah exact sequence. A holomorphic connection on $E_G$ over $M$ is a holomorphic splitting

$$TM \to \mathrm{At}(E_G)$$

of (3-5) [Atiyah 1957].

As before, setting $D = M \setminus M^0$, define

$$\mathrm{At}(E_G)(-\log D) := \phi^{-1}(TM(-\log D)) \subset \mathrm{At}(E_G),$$

where $\phi$ is the projection in (3-5) and $TM(-\log D)$ is the subsheaf in (3-3). So (3-5) gives the following short exact sequence of holomorphic vector bundles on $M$:

$$0 \to \mathrm{ad}(E_G) \to \mathrm{At}(E_G)(-\log D) \xrightarrow{\phi} TM(-\log D) \to 0. \tag{3-6}$$

A *logarithmic connection* on $E_G$, with singular locus $D$, is a holomorphic homomorphism

$$\delta : TM(-\log D) \to \mathrm{At}(E_G)(-\log D)$$

such that $\phi \circ \delta$ is the identity automorphism of $TM(-\log D)$, where $\phi$ is the homomorphism in (3-6). Just like the curvature of a connection, the curvature of a logarithmic connection $\delta$ on $E_G$ is the obstruction for the homomorphism $\delta$ to preserve the Lie algebra structure of the sheaf of sections of $TM(-\log D)$ and $\mathrm{At}(E_G)(-\log D)$ given by the Lie bracket of vector fields. In particular, $\delta$ is called *integrable* (or *flat*) if it preserves the Lie algebra structure of the sheaf of sections of $TM(-\log D)$ and $\mathrm{At}(E_G)(-\log D)$ given by the Lie bracket of vector fields.

**Proposition 3.2.** *Let $(E_G, \tilde{\rho})$ be a T-equivariant principal G-bundle on $M$. Then $E_G$ admits an integrable logarithmic connection that restricts to the connection $\mathcal{D}^0$ on $M^0$ constructed in Section 2.*

*Proof.* Let

$$\tilde{\mathcal{V}} := E_G \times \mathfrak{t} \to E_G$$

be the trivial vector bundle over $E_G$ with fiber $\mathfrak{t}$. Note that $p^*\mathcal{V} = \tilde{\mathcal{V}}$, where $\mathcal{V}$ is the vector bundle in (3-1), and $p$, as before, is the projection of $E_G$ to $M$.

The action $\tilde{\rho}$ of $T$ on $E_G$ produces a homomorphism

$$\tilde{\beta} : \tilde{\mathcal{V}} \to TE_G. \tag{3-7}$$

Since $p^{-1}(D)$ is preserved by the action of $T$ on $E_G$, the induced action of $T$ on $\mathcal{O}_{E_G}$ preserves the subsheaf $\mathcal{O}_{E_G}(-p^{-1}(D))$. Hence the image of $\tilde{\beta}$ lies inside the subsheaf

$$TE_G(-\log p^{-1}(D)) \subset TE_G.$$

Note that $p^{-1}(D)$ is a simple normal crossing divisor on $E_G$ because $D$ is a simple normal crossing divisor on $M$.

In Lemma 3.1(2) we saw that $\beta$ is an isomorphism. Consider

$$p^*\beta^{-1} : p^*(TM(-\log D)) \to p^*\mathcal{V} = \tilde{\mathcal{V}}.$$

Precomposing this with $\tilde{\beta}$ in (3-7), we have

$$\tilde{\beta} \circ (p^*\beta^{-1}) : p^*(TM(-\log D)) \to TE_G(-\log p^{-1}(D)).$$

We observe that the homomorphism $\tilde{\beta} \circ (p^*\beta^{-1})$ is $G$-equivariant for the trivial action of $G$ on $p^*(TM(-\log D))$ and the action of $G$ on $TE_G(-\log p^{-1}(D))$ induced by the action of $G$ on $E_G$. Therefore, taking the $G$-invariant parts of the direct images by $p$, the above homomorphism $\tilde{\beta} \circ (p^*\beta^{-1})$ produces a homomorphism

$$\begin{aligned} \beta' : TM(-\log D) = \big(p_* p^*(TM(-\log D))\big)^G \\ \to \big(p_* TE_G(-\log p^{-1}(D))\big)^G = \mathrm{At}(E_G)(-\log D). \end{aligned}$$

It is now straightforward to check that the homomorphism $\beta'$ produces a holomorphic splitting of the exact sequence in (3-6). Therefore, $\beta'$ defines a logarithmic connection on $E_G$ singular on $D$. The restriction of this logarithmic connection to $M^0$ clearly coincides with the connection $\mathcal{D}^0$ constructed in Section 2. □

## 4. A criterion for equivariance

For each point $t \in T$, define the automorphism

$$\rho_t : M \to M, \quad x \mapsto \rho(t, x),$$

where $\rho$ is the action in (2-1). If $(E_G, \tilde{\rho})$ is a $T$-equivariant principal $G$-bundle on $M$, then clearly the map

$$E_G \to E_G, \quad z \mapsto \tilde{\rho}(t, z)$$

is an isomorphism of the principal $G$-bundle $\rho_t^* E_G$ with $E_G$. The aim in this section is to prove a converse of this statement.

Take an algebraic principal $G$-bundle

$$p : E_G \to M.$$

Let $\mathcal{G}$ be the set of all pairs of the form $(t, f)$, where $t \in T$ and where

$$f : E_G \to E_G$$

is an algebraic automorphism of the variety $E_G$ satisfying the following conditions:

- $p \circ f = \rho_t \circ p$, and
- $f$ intertwines the action of $G$ on $E_G$.

Note that the above two conditions imply that $f$ is an algebraic isomorphism of the principal $G$-bundle $\rho_t^* E_G$ with $E_G$.

We have the following composition on the set $\mathcal{G}$:

$$(t_1, f_1) \cdot (t_2, f_2) := (t_1 \circ t_2, f_1 \circ f_2).$$

The inverse of $(t, f)$ is $(t^{-1}, f^{-1})$. These operations make $\mathcal{G}$ a group. In fact, $\mathcal{G}$ has the structure of an affine algebraic group defined over $\mathbb{C}$. Let $\mathcal{A}$ denote the group of all algebraic automorphisms of the principal $G$-bundle $E_G$. So $\mathcal{A}$ is a subgroup of $\mathcal{G}$ with the inclusion map being $f \mapsto (e, f)$. We have a natural projection

$$h : \mathcal{G} \to T, \quad (t, f) \mapsto t$$

which fits in the following exact sequence of complex affine algebraic groups:

$$0 \to \mathcal{A} \to \mathcal{G} \xrightarrow{h} T. \tag{4-1}$$

We note that there is a tautological action of $\mathcal{G}$ on $E_G$; the action of any $(t, f) \in \mathcal{G}$ on $E_G$ is given by the map defined by $y \mapsto f(y)$.

Now assume that $E_G$ satisfies the condition that, for every $t \in T$, the pulled-back principal $G$-bundle $\rho_t^* E_G$ is isomorphic to $E_G$. This assumption is equivalent to the statement that the homomorphism $h$ in (4-1) is surjective.

In view of the above assumption, the sequence in (4-1) becomes the following short exact sequence of complex affine algebraic groups:

$$0 \to \mathcal{A} \to \mathcal{G} \xrightarrow{h} T \to 0. \tag{4-2}$$

Let $\mathcal{G}^0 \subset \mathcal{G}$ be the connected component containing the identity element. Since $T$ is connected and $h$ is surjective, the restriction of $h$ to $\mathcal{G}^0$ is also surjective. Therefore, from (4-2) we have the short exact sequence of affine complex algebraic groups

$$0 \to \mathcal{A}^0 \xrightarrow{\iota_{\mathcal{A}}} \mathcal{G}^0 \xrightarrow{h^0} T \to 0, \tag{4-3}$$

where $\mathcal{A}^0 := \mathcal{A} \cap \mathcal{G}^0$ and $h^0 := h|_{\mathcal{G}^0}$.

Take a maximal torus $T_{\mathcal{G}} \subset \mathcal{G}^0$. From (4-3) it follows that the restriction

$$h' := h^0|_{T_{\mathcal{G}}} : T_{\mathcal{G}} \to T$$

is surjective. Define $T_{\mathcal{A}} := \mathcal{A}^0 \cap T_{\mathcal{G}} \subset T_{\mathcal{G}}$ using the homomorphism $\iota_{\mathcal{A}}$ in (4-3). Therefore, from (4-3) we have the short exact sequence of algebraic groups

$$0 \to T_{\mathcal{A}} \xrightarrow{\iota_{\mathcal{A}}|_{T_{\mathcal{A}}}} T_{\mathcal{G}} \xrightarrow{h'} T \to 0. \tag{4-4}$$

Recall that $\mathcal{G}$ has a tautological action on $E_G$. Therefore, the subgroup $T_{\mathcal{G}}$ has a tautological action on $E_G$ which is the restriction of the tautological action of $\mathcal{G}$.

Now we assume that the group $G$ is reductive.

A parabolic subgroup of $G$ is a connected Zariski closed subgroup $P \subset G$ such that the variety $G/P$ is projective. For a parabolic subgroup $P$, its unipotent radical will be denoted by $R_u(P)$. A Levi subgroup of $P$ is a connected reductive subgroup $L(P) \subset P$ such that the composition

$$L(P) \hookrightarrow P \to P/R_u(P)$$

is an isomorphism. Levi subgroups exist, and any two Levi subgroups of $P$ differ by conjugation by an element of $R_u(P)$ [Humphreys 1975, p. 184–185, §30.2; Borel 1991, p. 158, 11.22 and 11.23].

Let $\mathrm{Ad}(E_G) := E_G \times^G G \to M$ be the adjoint bundle associated to $E_G$ for the adjoint action of $G$ on itself. The fibers of $\mathrm{Ad}(E_G)$ are groups identified with $G$ up to an inner automorphism; the corresponding Lie algebra bundle is $\mathrm{ad}(E_G)$. We note that $\mathcal{A}$ in (4-2) is the space of all algebraic sections of $\mathrm{Ad}(E_G)$.

Using the action of $T_{\mathcal{A}}$ on $E_G$, we have

- a Levi subgroup $L(P)$ of a parabolic subgroup $P$ of $G$, and
- an algebraic reduction of structure group $E_{L(P)} \subset E_G$ of $E_G$ to $L(P)$ which is preserved by the tautological action of $T_{\mathcal{G}}$ on $E_G$,

such that the image of $T_{\mathcal{A}}$ in $\mathrm{Ad}(E_G)$ (recall that the elements of $\mathcal{A}$ are sections of $\mathrm{Ad}(E_G)$) lies in the connected component, containing the identity element, of the center of each fiber of $\mathrm{Ad}(E_{L(P)}) \subset \mathrm{Ad}(E_G)$ (see [Balaji et al. 2005; Biswas and Parameswaran 2006] for the construction of $E_{L(P)}$). The construction of $E_{L(P)}$ requires fixing a point $z_0$ of $E_G$, where $E_{L(P)}$ contains $z_0$. Using $z_0$, the fiber $(E_{L(P)})_{p(z_0)}$ is identified with $L(P)$. Moreover, the evaluation, at $p(z_0)$, of the sections of $\mathrm{Ad}(E_G)$ corresponding to the elements of $T_{\mathcal{A}}$ makes $T_{\mathcal{A}}$ a subgroup of the connected component, containing the identity element, of the center of $E_{L(P)}$; in particular, this evaluation map on $T_{\mathcal{A}}$ is injective (see the second paragraph in [Balaji et al. 2005, p. 230, Section 3]). We briefly recall (from [Balaji et al. 2005; Biswas and Parameswaran 2006]) the argument that the evaluation map on semisimple elements of $\mathcal{A}$ is injective. Let $\xi$ be a semisimple element of $\mathcal{A} = \Gamma(M, \mathrm{Ad}(E_G))$.

Since $\xi$ is semisimple, for each point $x \in M$, the evaluation $\xi(x)$ is a semisimple element of $\mathrm{Ad}(E_G))_x$. The group $\mathrm{Ad}(E_G))_x$ is identified with $G$ up to an inner automorphism of $G$. All conjugacy classes of a semisimple element of $G$ are parametrized by $T_G/W_{T_G}$, where $T_G$ is a maximal torus in $G$, and $W_{T_G} = N(T_G)/T_G$ is the Weyl group with $N(T_G)$ being the normalizer of $T_G$ in $G$. We note that $T_G/W_{T_G}$ is an affine variety. Therefore, we get a morphism $\xi' : M \to T_G/W_{T_G}$ that sends any $x \in M$ to the conjugacy class of $\xi(x)$. Since $M$ is a projective variety and $T_G/W_{T_G}$ is an affine variety, we conclude that $\xi'$ is a constant map. So if $\xi(x) = e$ for some $x \in M$, then $\xi = e$ identically.

Let $Z^0_{L(P)} \subset L(P)$ be the connected component, containing the identity element, of the center. We note that $Z^0_{L(P)}$ is a product of copies of $\mathbb{G}_m$. Therefore, the above injective homomorphism $T_{\mathcal{A}} \to Z^0_{L(P)}$ extends to a homomorphism

$$\eta : T_{\mathcal{G}} \to Z^0_{L(P)}.$$

Define

$$\eta' := \tau \circ \eta, \tag{4-5}$$

where $\tau$ is the inversion homomorphism of $Z^0_{L(P)}$ defined by $g \mapsto g^{-1}$.

Consider the action of $T_{\mathcal{G}}$ on $E_{L(P)}$; recall that $E_{L(P)}$ is preserved by the tautological action of $T_{\mathcal{G}}$ on $E_G$. We can twist this action on $E_{L(P)}$ by $\eta'$ in (4-5), because the actions of $Z^0_{L(P)}$ and $L(P)$ on $E_{L(P)}$ commute. For this new action, the group $T_{\mathcal{A}}$ clearly acts trivially on $E_{L(P)}$.

Consider the above action of $T_{\mathcal{G}}$ on $E_{L(P)}$ constructed using $\eta'$. Since $T_{\mathcal{A}}$ acts trivially on $E_{L(P)}$, the action of $T_{\mathcal{G}}$ on $E_{L(P)}$ descends to an action of $T$ on $E_{L(P)}$ (see (4-4)). The principal $G$-bundle $E_G$ is the extension of the structure group of $E_{L(P)}$ using the inclusion of $L(P)$ in $G$. Therefore, the above action of $T$ on $E_{L(P)}$ produces an action of $T$ on $E_G$. More precisely, the total space of $E_G$ is the quotient of $E_{L(P)} \times G$ where two elements $(z_1, g_1)$ and $(z_2, g_2)$ of $E_{L(P)} \times G$ are identified if there is an element $g \in L(P)$ such that $z_2 = z_1 g$ and $g_2 = g^{-1} g_1$. Now the action of $T$ on $E_{L(P)} \times G$, given by the above action of $T$ on $E_{L(P)}$ and the trivial action of $T$ on $G$, descends to an action of $T$ on the quotient space $E_G$. Consequently, $E_G$ admits a $T$-equivariant structure.

Therefore, we have proved the following:

**Proposition 4.1.** *Let $G$ be reductive, and let $E_G \to M$ be a principal $G$-bundle such that, for every $t \in T$, the pulled-back principal $G$-bundle $\rho_t^* E_G$ is isomorphic to $E_G$. Then $E_G$ admits a $T$-equivariant structure.*

For vector bundles on $M$, Proposition 4.1 was proved by Klyachko [1989, p. 342, Proposition 1.2.1].

***Equivariance property from a logarithmic connection.***

**Theorem 4.2.** *Let $G$ be reductive, and let $p : E_G \to M$ be a principal $G$-bundle admitting a logarithmic connection whose singularity locus is contained in the divisor $D = M \setminus M^0$. Then $E_G$ admits a $T$-equivariant structure.*

*Proof.* Since $E_G$ admits a logarithmic connection, by definition, there is a homomorphism of coherent sheaves

$$\delta : TM(-\log D) \to \mathrm{At}(E_G)(-\log D)$$

such that $\phi \circ \delta$ is the identity automorphism of $TM(-\log D)$, where $\phi$ is the homomorphism in (3-6). Let

$$\hat{\delta} : H^0(M, TM(-\log D)) \to H^0(M, \mathrm{At}(E_G)(-\log D))$$

be the homomorphism of global sections given by $\delta$. From Lemma 3.1(2) we know that $H^0(M, TM(-\log D))$ is the Lie algebra $\mathfrak{t}$ of $T$.

We will now show that there is a natural injective homomorphism

$$\theta : H^0(M, \mathrm{At}(E_G)(-\log D)) \to \mathrm{Lie}(\mathcal{G}), \tag{4-6}$$

where $\mathrm{Lie}(\mathcal{G})$ is the Lie algebra of the group $\mathcal{G}$ in (4-1).

The elements of $\mathrm{Lie}(\mathcal{G})$ are all holomorphic sections $s \in H^0(M, \mathrm{At}(E_G))$ such that the vector field $\phi(s)$, where $\phi$ is the projection in (3-5), is of the form $\beta(s')$, where $s' \in \mathfrak{t}$ and where $\beta$ is the homomorphism in (3-2). Now, if

$$s \in H^0(M, \mathrm{At}(E_G)(-\log D)) \subset H^0(M, \mathrm{At}(E_G)),$$

then $\phi(s)$ is a holomorphic section of $TM(-\log D)$ (see (3-6)). From Lemma 3.1(2) it now follows that $\phi(s)$ is of the form $\beta(s')$, where $s' \in \mathfrak{t}$. This gives us the injective homomorphism in (4-6).

Finally, consider the composition

$$\theta \circ \hat{\delta} : \mathfrak{t} = H^0(M, TM(-\log D)) \to \mathrm{Lie}(\mathcal{G}).$$

From its construction it follows that

$$(dh) \circ \theta \circ \hat{\delta} = \mathrm{Id}_{\mathfrak{t}},$$

where $dh : \mathrm{Lie}(\mathcal{G}) \to \mathfrak{t}$ is the homomorphism of Lie algebras given by $h$ in (4-1). In particular, $dh$ is surjective. Since $T$ is connected, this immediately implies that the homomorphism $h$ is surjective. Now from Proposition 4.1 it follows that $E_G$ admits a $T$-equivariant structure. $\square$

## References

[Atiyah 1957] M. F. Atiyah, "Complex analytic connections in fibre bundles", *Trans. Amer. Math. Soc.* **85** (1957), 181–207. MR 19,172c Zbl 0078.16002

[Balaji et al. 2005] V. Balaji, I. Biswas, and D. S. Nagaraj, "Krull–Schmidt reduction for principal bundles", *J. Reine Angew. Math.* **578** (2005), 225–234. MR 2006e:14061 Zbl 1084.14041

[Biswas and Parameswaran 2006] I. Biswas and A. J. Parameswaran, "On the equivariant reduction of structure group of a principal bundle to a Levi subgroup", *J. Math. Pures Appl.* (9) **85**:1 (2006), 54–70. MR 2006k:14081 Zbl 1159.14307

[Borel 1991] A. Borel, *Linear algebraic groups*, 2nd ed., Graduate Texts in Mathematics **126**, Springer, New York, 1991. MR 92d:20001 Zbl 0726.20030

[Fulton 1993] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies **131**, Princeton Univ. Press, 1993. MR 94g:14028 Zbl 0813.14039

[Humphreys 1975] J. E. Humphreys, *Linear algebraic groups*, Graduate Texts in Mathematics **21**, Springer, New York, 1975. MR 53 #633 Zbl 0325.20039

[Klyachko 1989] A. A. Klyachko, "Equivariant bundles on toral varieties", *Izv. Akad. Nauk SSSR Ser. Mat.* **53**:5 (1989), 1001–1039. In Russian; translated in *Math. USSR Izv.* **35**:2 (1990), 337–375. MR 91c:14064 Zbl 0706.14010

Received March 11, 2015. Revised June 8, 2015.

INDRANIL BISWAS
SCHOOL OF MATHEMATICS
TATA INSTITUTE OF FUNDAMENTAL RESEARCH
HOMI BHABHA ROAD
MUMBAI 400005
INDIA
indranil@math.tifr.res.in

ARIJIT DEY
DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY MADRAS
CHENNAI 600036
INDIA
arijitdey@gmail.com

MAINAK PODDAR
DEPARTAMENTO DE MATEMÁTICAS
UNIVERSIDAD DE LOS ANDES
BOGOTA
COLOMBIA
mainakp@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.327

# ON A SPECTRAL THEOREM IN PARAORTHOGONALITY THEORY

KENIER CASTILLO, RUYMÁN CRUZ-BARROSO
AND FRANCISCO PERDOMO-PÍO

**Motivated by the works of Delsarte and Genin (1988, 1991), who studied paraorthogonal polynomials associated with positive definite Hermitian linear functionals and their corresponding recurrence relations, we provide paraorthogonality theory, in the context of quasidefinite Hermitian linear functionals, with a recurrence relation and the analogous result to the classical Favard's theorem or spectral theorem. As an application of our results, we prove that for any two monic polynomials whose zeros are simple and strictly interlacing on the unit circle, with the possible exception of one of them which could be common, there exists a sequence of paraorthogonal polynomials such that these polynomials belong to it. Furthermore, an application to the computation of Szegő quadrature formulas is also discussed.**

## 1. Introduction

The *paraorthogonal polynomials on the unit circle* (POPUC), in the context of quasidefinite (or regular) moment linear functionals, were introduced for the first time by Jones, Njåstad and Thron in their excellent survey paper [Jones et al. 1989]. The main objective of the authors was to construct quadrature formulas for the approximation of an integral with respect to a measure whose support is contained in the unit circle, analogous to the generalized Gaussian rules and, as a consequence, solve the trigonometric moment problem [Geronimus 1946]. In this respect, nodes on the unit circle, positive weights and maximal domain of validity are required. As a result, the so-called *Szegő quadrature* (SQ) formulas are introduced and characterized: their nodes are zeros of a special class of POPUC, known as invariant (or self-inverse). But moreover, [Jones et al. 1989] served to demand a deeper study of the properties of this new family of polynomials since, contrary to *orthogonal polynomials on the unit circle* (OPUC), the invariant POPUC with respect to a measure supported on the unit circle have simple zeros on the unit circle with many additional

*MSC2010:* 42C05, 30C15, 26C10.
*Keywords:* paraorthogonal polynomials, quasidefinite Hermitian linear functionals, spectral theorem, Geronimus–Wendroff theorem, Szegő quadrature formulas.

properties [Cantero et al. 2002; Golinskii 2002; Simon 2005a; 2005b; 2007; 2011; Wong 2007]. These polynomials play in the unit circle the same role as *orthogonal polynomials on the real line* (OPRL) from the perspective of quadrature formulas.

The earliest reference to invariant POPUC is due to Geronimus [1946, Theorem III]. However, in [Jones et al. 1989] they are defined in a more general setting from their orthogonality conditions and characterized in terms of the corresponding OPUC. The counterpart to the deficiency in the orthogonality conditions for POPUC, which are not orthogonal to the constants, is the fact that for a given measure and a fixed $n$, the POPUC of degree $n$ is not unique, and basically depends on one unimodular free parameter. Equivalently, in quadrature terminology, we have a one-parameter family of $n$-point SQ formulas, exact in a subspace of Laurent polynomials of dimension $2n-1$, instead of $2n$; see, e.g., [Cruz-Barroso et al. 2007; Peherstorfer 2011].

Beyond their essential role in the development of quadrature formulas, the theory of invariant POPUC is significantly enriched from both theoretical and practical points of view. The uses of their zeros instead of zeros of OPUC in frequency analysis problems [Daruis et al. 2003] and their appearance as the minimizer of the isometric Arnoldi minimization problem [Helsen et al. 2005] represent some of their best applications. On the other hand, the works of Cantero, Moral and Velázquez [Cantero et al. 2002], Golinskii [2002], Simon [2007], and Wong [2007] are essential to understand the behavior of the zeros of POPUC. Recently in [Simanek 2015], it was also proved that the zeros of invariant POPUC designate the location of a set of particles that are in electrostatic equilibrium with respect to a particular external field. Furthermore, after their formal introduction, POPUC were defined in the more general context of orthogonal rational functions [Bultheel et al. 1999, Chapter 5].

One of the main algebraic properties in the study of orthogonal polynomials has not been established yet for general POPUC in the context of quasidefinite linear functionals: a recurrence relation and its corresponding Favard's theorem or spectral theorem. And this is precisely the starting point of this work, even though for the positive definite case it is very well known that POPUC satisfy a three-term recurrence, which is the key for the *tridiagonal approach* developed by Delsarte and Genin [1988; 1991a; 1991b] to solve the standard linear prediction problem. Similar results can be also found in [Castillo et al. 2014] where the corresponding OPUC and the nontrivial probability measure supported on the unit circle are deduced. The reader whose interest concerns particularly the applications of POPUC to digital signal processing can find a survey in [Delsarte and Genin 1990]. We recall that in [Delsarte and Genin 1988; 1991a; 1991b], the authors considered POPUC associated with positive definite moment linear functionals [Delsarte and Genin 1988, (4.13)] and, particularly, in [loc. cit.] they say that presumably the quasidefinite case can be traced with the help of the theory of pseudo-Carathéodory functions. Motivated by this last observation, in the present work, we study some properties of POPUC for the quasidefinite case using only standard techniques from the theory of OPRL and OPUC. We focus our attention on the analogs of the spectral theorem and the Geronimus–Wendroff theorem for POPUC, expecting them to be as useful as these results in the theory of OPRL and OPUC. The importance of the previous results for OPRL and OPUC is summarized in the survey [Marcellán and Álvarez-Nodarse 2001].

In Section 3, we prove that three consecutive POPUC are connected by a simple relation which we can derive in a straightforward way. Moreover, the spectral theorem is also proved. In Section 4, we present an example of the applicability of the spectral theorem by proving the Geronimus–Wendroff theorem for POPUC. Furthermore, an application to the computation of SQ formulas is considered.

In the next section, we fix the notation used in this work and present some preliminaries, which will help to make our original results self-contained and accessible to the reader not familiar with the theory of OPUC and POPUC.

## 2. Orthogonality and paraorthogonality

We denote by $\Lambda := \mathbb{C}[z, z^{-1}]$ the complex vector space of Laurent polynomials in the variable $z$. Associated with every pair of integer numbers $(p, q)$, with $p \leq q$, we define the vector subspace $\Lambda_{p,q}$ of Laurent polynomials of the form

$$\sum_{n=p}^{q} \varsigma_n z^n, \qquad \varsigma_n \in \mathbb{C}.$$

The vector subspace of complex polynomials will be denoted by $\mathbb{P} := \mathbb{C}[z]$ and we write $\mathbb{P}_q \equiv \Lambda_{0,q}$ for the vector subspace of polynomials of degree (at most) $q$, while $\mathbb{P}_{-1} \equiv \{0\}$ is the trivial subspace.

Let us introduce the moment linear functional $\mu$ on $\Lambda$ such that

$$c_n := \mu(z^n) = \overline{\mu(z^{-n})} =: \bar{c}_{-n}, \qquad n \geq 0, \tag{2-1}$$

i.e., $\mu$ is an Hermitian linear functional. The complex numbers $\{c_n\}_{n=-\infty}^{\infty}$ are called the moments associated with $\mu$. In terms of $\mu$, we consider a sesquilinear functional $\langle \cdot, \cdot \rangle$ on $\Lambda \times \Lambda$ defined by

$$\langle f, g \rangle := \mu(f(z)\bar{g}(z^{-1})), \qquad f, g \in \Lambda.$$

The Gram matrix associated with the inner product $\langle \cdot, \cdot \rangle$ in terms of $1, z, z^2, \ldots$ is the Toeplitz matrix

$$\boldsymbol{T} = [\langle z^l, z^j \rangle]_{l,j \geq 0} = \begin{bmatrix} c_0 & c_1 & \cdots & c_n & \cdots \\ c_{-1} & c_0 & \cdots & c_{n-1} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ c_{-n} & c_{-n+1} & \cdots & c_0 & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{bmatrix}.$$

Denote by $\boldsymbol{T}_n$ the $(n+1)\times(n+1)$ principal leading submatrix of $\boldsymbol{T}$. If $\det(\boldsymbol{T}_n) \neq 0$ for every $n \geq 0$, then $\mu$ is said to be *quasidefinite* and the existence of a sequence of monic polynomials, orthogonal with respect to $\mu$, is guaranteed. On the other hand, by the Carathéodory–Toeplitz theorem [Simon 2005a, Section 1.3], if $\det(\boldsymbol{T}_n) > 0$ for every $n \geq 0$, then (2-1) are the moments of a nontrivial (i.e., with infinitely many points of increase) probability measure $d\sigma$ supported on the unit circle $\partial\mathbb{D}$, that is, the boundary of the open unit disk $\mathbb{D} := \{z \in \mathbb{C};\ |z| < 1\}$ parametrized by $z = e^{i\theta}$, $\theta \in [0, 2\pi)$, and the converse is also true. In mathematical terms, $\mu$ has the integral representation

$$\mu(f) = \int f\, d\sigma, \qquad f \in \Lambda.$$

In this case, $\mu$ is called *positive definite*.

The application of the Gram–Schmidt process to $1, z, z^2, \ldots$ (a linearly independent system in the Hilbert space $L^2(\partial\mathbb{D}, d\sigma)$ with the norm induced by our inner product) yields the sequence of monic polynomials, $\{\Phi_n\}_{n\geq 0}$, orthogonal with respect to $d\sigma$ (or equivalently with respect to $\mu$) called the sequence of OPUC (see [Simon 2005a; 2005b; 2011] for a recent account of the theory). In other words, there exists a unique sequence of monic polynomials such that

$$\langle \Phi_n, z^m\rangle = \int \Phi_n(z) z^{-m}\, d\sigma(z) = \kappa_n \delta_{n,m}, \quad \kappa_n > 0, \quad 0 \leq m \leq n, \tag{2-2}$$

with $\Phi_n \in \mathbb{P}_n \setminus \mathbb{P}_{n-1}$ and $\delta_{n,m}$ the Kronecker delta symbol. We recall that the solution of the trigonometric moment problem is always unique [Geronimus 1946; Jones et al. 1989]. The associated orthonormal polynomials are given by

$$\varphi_n = \kappa_n \Phi_n, \quad \kappa_n := \prod_{j=0}^{n-1} \rho_j^{-1}, \quad \rho_j := (1 - |\Phi_{j+1}(0)|^2)^{1/2}. \tag{2-3}$$

The monic OPUC satisfy the following recurrence relation (Szegő's recurrence):

$$\begin{bmatrix} \Phi_{n+1}(z) \\ \Phi_{n+1}^*(z) \end{bmatrix} = \begin{bmatrix} z & -\bar{\alpha}_n \\ -\alpha_n z & 1 \end{bmatrix} \begin{bmatrix} \Phi_n(z) \\ \Phi_n^*(z) \end{bmatrix}, \tag{2-4}$$

with initial condition $\Phi_0 :\equiv 1$. The numbers $\{\alpha_n\}_{n\geq 0} \in \mathbb{D}^\infty$ are known as Verblunsky coefficients and, as usual, if $f \in \mathbb{P}_n \setminus \mathbb{P}_{n-1}$, then $f^*$ denotes its reversed polynomial, defined by $f^*(z) := z^n \overline{f(1/\bar{z})}$. By Szegő's recurrence, we get $\alpha_n = -\overline{\Phi_{n+1}(0)}$ since $\Phi_n^*(0) = 1$; thus, we set $\alpha_{-1} :\equiv -1$. We recall that for the quasidefinite case $\{\alpha_n\}_{n\geq 0} \notin \partial\mathbb{D}^\infty$.

The orthogonality conditions (2-2) can be weakened adequately in order to overcome the apparent difference between OPRL and OPUC [Simon 2011, Theorem 1.2.6 and Theorem 2.14.2]. In this way, the corresponding polynomials will be the POPUC introduced by Jones, Njåstad and Thron [Jones et al. 1989].

**Definition 2.1.** A sequence of polynomials $\{\Phi_n(\,\cdot\,,\tau_n)\}_{n\geq 0}$ is said to be a sequence of POPUC if

$$\begin{gathered}\langle \Phi_n(\,\cdot\,,\tau_n), 1\rangle \neq 0,\\ \langle \Phi_n(z,\tau_n), z^m\rangle = 0, \quad 1 \leq m \leq n-1, \qquad \langle \Phi_n(z,\tau_n), z^n\rangle \neq 0.\end{gathered} \tag{2-5}$$

In general, we follow the notation from standard literature, as in [Simon 2005a; 2005b; 2011]. The presence of the parameter $\tau_n$ in Definition 2.1 will be fully clarified in (2-6); see further. It is worth pointing out that in the applications (see among others [Jones et al. 1989; Peherstorfer 2011]), it is useful to have a POPUC such that the distribution of its zeros behaves as in the case of OPRL. This allows us to introduce the concept of invariance.

**Definition 2.2.** A sequence of polynomials $\{f_n\}_{n\geq 0}$ is said to be invariant if there exists $\chi_n \in \partial\mathbb{D}$ such that $f_n^* = \chi_n f_n$.

By [Jones et al. 1989, Theorem 6.1(B)], if the polynomials $\{\Phi_n(\,\cdot\,,\tau_n)\}_{n\geq 1}$ are (monic) invariant POPUC, then

$$\Phi_n(z,\tau_n) = z\Phi_{n-1}(z) - \bar{\tau}_n\Phi_{n-1}^*(z), \tag{2-6}$$

with $\tau_n = -\chi_n \in \partial\mathbb{D}$. Clearly, the converse is also true. Based on the previous assertion, $\Phi_n(\,\cdot\,,\tau_n)$ is completely determined by the parameter $\tau_n$ and the first $n-1$ Verblunsky coefficients associated with the corresponding sequence of OPUC.

For numbers on $\partial\mathbb{D}$, we can define a cyclic order in terms of their arguments [Simon 2007]. An ordered set of points $(z_1,\ldots,z_n)\in\partial\mathbb{D}^n$ is called *cyclicly ordered* if $(z_j,z_{j+1})_{j=1}^n$ and $(z_n,z_1)$ contain no other $z_j$. Two cyclicly ordered sets of points on $\partial\mathbb{D}$, $(z_1,\ldots,z_n)$ and $(\zeta_1,\ldots,\zeta_m)$, are said to *strictly interlace* if after a cyclic permutation of the $\zeta_j$, we have $\zeta_j\in(z_j,z_{j+1})$, $j=1,\ldots,m$, and $\zeta_n\in(z_n,z_1)$ if $n=m$. This definition can be naturally extended to two cyclicly ordered sets of zeros with different numbers of elements. We recall here that the zeros of two consecutive invariant POPUC, $\Phi_{n+1}(\,\cdot\,,\tau_{n+1})$ and $\Phi_n(\,\cdot\,,\tau_n)$, associated with the measure $d\sigma$ (positive definite case) have at most one zero in common, namely $\zeta$. In other words, one of two following possibilities holds: the zeros of $\Phi_{n+1}(z,\tau_{n+1})/(z-\zeta)$ and $\Phi_n(\,\cdot\,,\tau_n)$, or the zeros of $\Phi_{n+1}(\,\cdot\,,\tau_{n+1})$ and $\Phi_n(\,\cdot\,,\tau_n)$, strictly interlace on $\partial\mathbb{D}$ (see [Simon 2007] and the references given there). We must once again urge the reader to consult the monographs [Simon 2005a; 2005b; 2011] where all the previous results can be found.

## 3. Recurrence relation and spectral theorem

It is very well known that the OPRL (and also the OPUC for nonzero Verblunsky coefficients) satisfy a linear recurrence relation [Chihara 1978; Szegő 1975] which plays a crucial role in the subsequent behavior of their theory. Such a recurrence relation does not hold for POPUC, but we can obtain a similar recurrence formula. To do this, we follow an analogous procedure to one pointed out by Atkinson [1964] and recovered by Simon [2005a, Theorem 1.5.2] to obtain Szegő's recurrence.

**Theorem 3.1.** *Given a quasidefinite moment functional $\mu$, there always exist three consecutive monic POPUC such that*

$$\Phi_{n+1}(z,\tau_{n+1}) = (z+\beta_n)\Phi_n(z,\tau_n) - \gamma_n z\Phi_{n-1}(z,\tau_{n-1}), \tag{3-7}$$

*where $\beta_n, \gamma_n \in \mathbb{C}\setminus\{0\}$ are given by*

$$\gamma_n = \frac{\langle \Phi_n(\,\cdot\,,\tau_n),1\rangle}{\langle \Phi_{n-1}(\,\cdot\,,\tau_{n-1}),1\rangle}, \quad \beta_n = \gamma_n \frac{\langle \Phi_{n-1}(z,\tau_{n-1}),z^{n-1}\rangle}{\langle \Phi_n(z,\tau_n),z^n\rangle}. \tag{3-8}$$

*Proof.* Our proof starts with the observation that

$$Q_{n+1}(z) := (z+\beta_n)\Phi_n(z,\tau_n) - \gamma_n z\Phi_{n-1}(z,\tau_{n-1}), \quad \beta_n,\gamma_n \in \mathbb{C},$$

is a monic polynomial of degree $n+1$ which is orthogonal to $\operatorname{span}\{z^2, z^3, \ldots, z^{n-1}\}$. The important point to notice here is that for constants $\beta_n$ and $\gamma_n$ given as in (3-8), $Q_{n+1}(z)$ is orthogonal to $\operatorname{span}\{z, z^2, \ldots, z^n\}$, which proves the theorem. □

In terms of the parameters $\{\tau_n\}_{n\ge1}$, the previous theorem says that given two numbers $\tau_{n-1}$ and $\tau_n$, a third number $\tau_{n+1}$ can be found such that the corresponding POPUC satisfies (3-7). We are now interested in the expression of the recurrence coefficients (3-8) in terms of the parameters $\{\tau_n\}_{n\ge1} \in \partial\mathbb{D}^\infty$ and the Verblunsky coefficients.

**Corollary 3.2.** *With reference to the recurrence formula* (3-7) *for the invariant case*, *the following holds*:

$$\beta_n = \frac{\tau_n}{\tau_{n+1}} \in \partial\mathbb{D}, \quad \gamma_n = \frac{\tau_n - \alpha_{n-1}}{\tau_{n+1} - \alpha_{n-2}}\rho_{n-2}^2 \in \mathbb{C}\setminus\{0\}.$$

*Proof.* From (2-6), we get

$$\Phi_n^*(\,\cdot\,,\tau_n) = -\tau_n\Phi_n(\,\cdot\,,\tau_n),$$

which gives

$$\tau_{n+1}\Phi_{n+1}(z,\tau_{n+1}) = (1+\bar\beta_n z)\tau_n\Phi_n(z,\tau_n) - \bar\gamma_n z\tau_{n-1}\Phi_{n-1}(z,\tau_{n-1}), \tag{3-9}$$

when substituted in the reversed (3-7). The expression for $\beta_n$ follows from the above equation comparing the leading coefficients.

By inverting the Szegő recurrence (2-4) [Simon 2005a, Theorem 1.5.4], it is easy to check that (2-6) in the monic case can be expressed by

$$\Phi_n(z,\tau_n)=\frac{1}{1-\omega_n\alpha_{n-1}}\big(\Phi_n(z)+\omega_n\Phi_n^*(z)\big),\qquad \omega_n:=\frac{\bar\alpha_{n-1}-\bar\tau_n}{1-\bar\tau_n\alpha_{n-1}}.$$

Thus, $\gamma_n$ follows from here as a consequence of Theorem 3.1, (2-3), the paraorthogonality conditions for $\Phi_n(\,\cdot\,,\tau_n)$ and the orthogonality conditions for $\Phi_n$ and $\Phi_n^*$. $\square$

The following result will be useful in determining the relation between $\tau_{n-1}$, $\tau_n$, and $\tau_{n+1}$ for the invariant case.

**Lemma 3.3.** *Let $\{\beta_n\}_{n\ge 0}$ be an arbitrary sequence on $\partial\mathbb{D}$ and let $\{\gamma_n\}_{n\ge 1}$ be an arbitrary sequence on $\mathbb{C}\setminus\{0\}$. Any sequence of polynomials $\{\Psi_n\}_{n\ge 0}$ defined by*

$$\Psi_{n+1}(z)=(z+\beta_n)\Psi_n(z)-\gamma_n z\Psi_{n-1}(z),\tag{3-10}$$

*with initial conditions $\Psi_0:\equiv 1$ and $\Psi_1(z):=z+\beta_0$, is a sequence of invariant polynomials if and only if the recurrence coefficients satisfy*

$$\frac{\gamma_n}{\bar\gamma_n}=\beta_{n-1}\beta_n.\tag{3-11}$$

*Proof.* As a direct consequence of Theorem 3.1, we get

$$\Psi_n(z)=\det(\boldsymbol{J}_n(z)),\qquad n\ge 1,$$

where the matrix $\boldsymbol{J}_n(z)$ is given by

$$\boldsymbol{J}_n(z)=\begin{bmatrix} z+\beta_0 & -\gamma_1 & 0 & \cdots & 0 & 0\\ -z & z+\beta_1 & -\gamma_2 & \cdots & 0 & 0\\ 0 & -z & z+\beta_2 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & z+\beta_{n-1} & -\gamma_{n-1}\\ 0 & 0 & 0 & \cdots & -z & z+\beta_{n-1}\end{bmatrix}.$$

For a polynomial $f$ to be invariant, it is necessary and sufficient that $|f|\equiv|f^*|$. Let us define the matrices $\boldsymbol{D}_n:=\operatorname{diag}[\beta_0^{-1},\beta_1^{-1},\dots,\beta_{n-1}^{-1}]$ and

$$\boldsymbol{J}_n^{(0)}(z):=\begin{bmatrix} z+\beta_0 & \beta_0 & 0 & \cdots & 0 & 0\\ -\bar\gamma_1\beta_1 z & z+\beta_1 & \beta_1 & \cdots & 0 & 0\\ 0 & -\bar\gamma_2\beta_2 z & z+\beta_2 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & z+\beta_{n-2} & \beta_{n-2}\\ 0 & 0 & 0 & \cdots & \bar\gamma_{n-1}\beta_{n-1}z & z+\beta_{n-1}\end{bmatrix}.$$

Notice that $z\boldsymbol{J}_n^H(1/\bar{z}) = \boldsymbol{D}_n\boldsymbol{J}_n^{(0)}(z)$, where the $^H$-operator denotes the conjugate transpose. Hence,

$$|\det(\boldsymbol{J}_n(z))| = |\det(\boldsymbol{D}_n)|\,|\det(\boldsymbol{J}_n^{(0)}(z))| = |\det(z\boldsymbol{J}_n^H(1/\bar{z}))|, \quad n \geq 1, \tag{3-12}$$

holds, if and only if (3-11) holds. □

Notice that under the hypothesis of Corollary 3.2, we get by Lemma 3.3 that

$$\tau_{n+1} = \frac{\bar{\gamma}_n}{\gamma_n}\tau_{n-1}.$$

As a consequence, we can deduce the following (forward) recurrence relation for the paraorthogonality parameters,

$$\tau_{n+1} = \frac{\tau_n - \alpha_{n-1}}{\bar{\tau}_n - \bar{\alpha}_{n-1}}\,\frac{\bar{\tau}_{n-1} - \bar{\alpha}_{n-2}}{\tau_{n-1} - \alpha_{n-2}}\tau_{n-1}, \tag{3-13}$$

which is equivalent to the (backward) recurrence

$$\tau_{n-1} = \frac{\tau_{n+1}\alpha_{n-2}(\bar{\alpha}_{n-1} - \bar{\tau}_n) + \alpha_{n-1} - \tau_n}{\tau_{n+1}(\bar{\alpha}_{n-1} - \bar{\tau}_n) + \bar{\alpha}_{n-2}(\alpha_{n-1} - \tau_n)}. \tag{3-14}$$

We next point out the important converse of Theorem 3.1. In agreement with classical literature, we refer to this result as Favard's theorem (see, among others, [Chihara 1978; Erdélyi et al. 1991; Favard 1935; Marcellán and Álvarez-Nodarse 2001; Szegő 1975]) or the spectral theorem [Ismail 2005], even though this result is previously contained in the works of Stieltjes [1895; 1894] and Stone [1932]. In the positive definite case, an analog of Favard's theorem for OPUC based on the construction of a sequence of absolutely continuous measures whose limit is the spectral measure is presented in [Erdélyi et al. 1991]. In this paper the authors follow a method used previously by Delsarte, Genin and Kamp [Delsarte et al. 1978] who consider the matrix-valued case. Some extensions to the quasidefinite case have been analyzed in [Marcellán and Álvarez-Nodarse 2001]. Our proof follows a standard scheme (constructive approach) which goes back at last to [Chihara 1978]; see also [Marcellán and Álvarez-Nodarse 2001].

**Theorem 3.4** (spectral theorem). *Let $\{\beta_n\}_{n\geq 0}$ be an arbitrary sequence on $\partial\mathbb{D}$ and let $\{\gamma_n\}_{n\geq 1}$ be an arbitrary sequence on $\mathbb{C}\setminus\{0\}$. Set $c_0 \in \mathbb{R}\setminus\{0\}$ and let $\{\Psi_n\}_{n\geq 0}$ be a sequence of invariant polynomials satisfying a recurrence relation as* (3-10), *with $\Psi_0 :\equiv 1$ and $\Psi_1(z) := z + \beta_0$. Then there exists a unique quasidefinite moment functional $\mu$ such that $\mu(1) = c_0$ and $\{\Psi_n\}_{n\geq 0}$ is the corresponding sequence of POPUC. Moreover, if $\{\Psi_n\}_{n\geq 0}$ is a sequence of polynomials with all its zeros on $\partial\mathbb{D}$, then there exists a unique measure $d\sigma$ such that $\int d\sigma = c_0 > 0$ and $\{\Psi_n\}_{n\geq 0}$ is the corresponding sequence of POPUC.*

*Proof.* We begin by constructing for $n \geq 2$ the moment linear functional $\mu^{(n)}$ on $\Lambda_{-(n-1),n-1}$ by

$$c_k := \mu^{(n)}(z^k) = \overline{\mu^{(n)}(z^{-k})} =: \bar{c}_{-k}, \quad 0 \leq k \leq n-1,$$

such that

$$\mu^{(n)}(\Psi_\ell(z)z^{-1}) = 0, \quad 2 \leq \ell \leq n. \tag{3-15}$$

So we proceed by induction on $n$. Notice that for $n = 2$,

$$\mu^{(2)}(\Psi_2(z)z^{-1}) = c_1 + \beta_0\beta_1 c_{-1} + (\beta_0 + \beta_1 - \gamma_1)c_0,$$

which allows us to define $c_1 = \bar{c}_{-1} \in \mathbb{C}$. More precisely, we let $c_1 := (1/2\gamma_1 - \beta_0)c_0$, which by Lemma 3.3 implies

$$\mu^{(2)}(\Psi_2(z)z^{-1}) = (1/2\gamma_1 + 1/2\bar{\gamma}_1\beta_0\beta_1 - \gamma_1)c_0 = 0.$$

In order to prove (3-15), write

$$\Psi_n(z) = z^n - \sum_{k=0}^{n-1} a_{n,k}\Psi_k(z),$$

where $\{a_{n,k}\}_{k=0}^{n-1}$ is uniquely determined. Let us now define $\mu^{(n)}$ on $\Lambda_{-(n-1),n-1}$ as an extension of $\mu^{(n-1)}$ such that $c_{n-1} := (a_{n,0} + a_{n,1}\beta_0)c_{-1} + a_{n,1}c_0$. In other words, we assume that for some $n \geq 3$, $c_{-(n-2)}, \dots, c_{n-2}$ have been determined such that $\mu^{(n-1)}$ defined on $\Lambda_{-(n-2),n-2}$ satisfies

$$c_m = \bar{c}_{-m}, \quad 0 \leq m \leq n-2,$$

and

$$\mu^{(n-1)}(\Psi_\ell(z)z^{-1}) = 0, \quad 2 \leq \ell \leq n-1.$$

Hence, by our assumption,

$$\mu^{(n)}(\Psi_n(z)z^{-1}) = 0.$$

This completes the induction. Therefore, it follows that $\mu$ defined on $\Lambda$ by (2-1) is an extension of $\mu^{(n)}$ defined on $\Lambda_{-n,n}$, and consequently

$$\mu(\Psi_n(z)z^{-1}) = 0, \quad n \geq 2. \tag{3-16}$$

According to the paraorthogonality conditions (2-5), it remains to check that

$$\langle \Psi_n, 1 \rangle \neq 0, \qquad \langle \Psi_n(z), z^q \rangle = 0, \quad 2 \leq q \leq n-1, \qquad \langle \Psi_n(z), z^n \rangle \neq 0.$$

From (3-7) and (3-16), we obtain

$$\begin{aligned}\langle \Psi_n, 1\rangle &= \gamma_n \langle \Psi_{n-1}, 1\rangle \\ &= c_0 \prod_{k=1}^{n} \gamma_k =: b_0 \neq 0.\end{aligned}$$

On the other hand, from (3-16) and the Hermitian character of the functional,

$$\langle \Psi_n(z), z\rangle = \langle \Psi_n^*(z), z^{n-1}\rangle = 0,$$

which by the invariant hypothesis implies

$$\langle \Psi_n(z), z^{n-1}\rangle = 0. \tag{3-17}$$

Now, we define an appropriate statement $(I_r)$ by

$$\langle \Psi_n(z), z^q\rangle = 0, \quad 1 \leq q \leq r, \quad n \geq r+1, \tag{$I_r$}$$

and prove by induction that the statement is valid for all $r$. Obviously, $(I_1)$ is (3-16). Assuming $(I_r)$ holds for some $r \geq 2$, we will prove $(I_{r+1})$, that is,

$$\langle \Psi_n(z), z^{r+1}\rangle = 0, \quad n \geq r+2.$$

Since $(I_r)$ holds, (3-7) yields

$$\langle \Psi_{n+1}(z), z^{r+1}\rangle = \beta_n \langle \Psi_n(z), z^{r+1}\rangle.$$

Taking into account that (3-17) holds,

$$\langle \Psi_{r+2}(z), z^{r+1}\rangle = 0,$$

which yields $\langle \Psi_{r+3}(z), z^{r+1}\rangle = 0$. Continuing in this manner, we conclude that $(I_{r+1})$ is valid. Furthermore, it follows easily that

$$\beta_n \langle \Psi_n(z), z^n\rangle = \gamma_n \langle \Psi_{n-1}(z), z^{n-1}\rangle.$$

Hence,

$$\langle \Psi_n(z), z^n\rangle = c_0 \prod_{k=1}^{n} \frac{\gamma_k}{\beta_k} =: b_n \neq 0.$$

If a sequence of POPUC associated with $\mu$ exists, it is uniquely determined by $b_0$ and $b_n$. Therefore,

$$\langle \Psi_n, 1\rangle = b_0, \qquad \langle \Psi_n, z^k\rangle = 0, \quad 1 \leq k \leq n-1, \qquad \langle \Psi_n, z^n\rangle = b_n,$$

is a consistent system of $n+1$ linear equations with $n+1$ unknowns. Notice that the coefficient matrix of this system is the moment matrix $\boldsymbol{T}_n$. Then it has a unique solution determined by $b_0$ and $b_n$ so that $\det(\boldsymbol{T}_n) \neq 0$. Thus, there exists a quasidefinite moment functional $\mu$ such that $\{\Psi_n\}_{n\geq 0}$ is the corresponding sequence of POPUC. This proves the first part of the theorem.

Let us now assume that $\{\Psi_n\}_{n\geq 0}$ is a sequence of polynomials with all its zeros on $\partial\mathbb{D}$. By Cohn's theorem [Rahman and Schmeisser 2002], we know that a polynomial with all its zeros on $\partial\mathbb{D}$ must be invariant. Furthermore, by Chen's theorem [1995, Theorem 1], we also know that a necessary and sufficient condition for all the zeros of $\Psi_n$ to lie on $\partial\mathbb{D}$ is that there exists a polynomial $\pi_{n-l} \in \mathbb{P}_{n-l}\backslash\mathbb{P}_{n-l-1}$ with all its zeros in $\mathbb{D}$ or on $\partial\mathbb{D}$ such that

$$\Psi_n(z) = z^l \pi_{n-l}(z) - \zeta_n \pi^*_{n-l}(z) \tag{3-18}$$

for some nonnegative integer $l$ and $\zeta_n \in \partial\mathbb{D}$. By the first part of the theorem, $\{\Psi_n\}_{n\geq 0}$ is a sequence of POPUC with respect to a quasidefinite functional $\mu$. By [Jones et al. 1989, Theorem 6.1(B)], (3-18) holds for $l \equiv 1$, $\zeta_n = \bar{\tau}_n$ and $\pi_{n-1} = \Phi_{n-1}$. We recall that $\{\Phi_n\}_{n\geq 0}$ is the sequence of OPUC associated with the POPUC $\{\Psi_n\}_{n\geq 0}$ and the functional $\mu$. At this point, we only can guarantee that the zeros of $\Phi_{n-1}$ lie on $\mathbb{C} \setminus \partial\mathbb{D}$ [Marcellán and Godoy 1991, Proposition 3.1]. If the zeros of $\Psi_n$ lie on $\partial\mathbb{D}$, then by Chen's theorem the zeros of $\Phi_{n-1}$ lie in $\mathbb{D}$. Finally, by Geronimus' theorem [1946, Theorem I], under our hypothesis the functional $\mu$ is positive definite.

The uniqueness of $\mu$ is a consequence inherited from the associated OPUC. □

It is very well known that for any three contiguous hypergeometric functions there is a linear contiguous relation. So if we are looking for a sequence of polynomials satisfying (3-7), we can find examples if we consider hypergeometric polynomials [Andrews et al. 1999, (2.5.16)]. Notice that from the previous theorem,

$$\begin{aligned}(c+n)\,{}_2F_1(-n-1, b; c; 1-z)\\ = ((b+n)z+c-b+n)\,{}_2F_1(-n, b; c; 1-z) - n\,z\,{}_2F_1(-n+1, b; c; 1-z),\end{aligned} \tag{3-19}$$

where $b, c \in \mathbb{C}$ and $c \neq 0, -1, -2, \ldots$ gives a set of hypergeometric POPUC.

**Example 3.5** (Askey POPUC). An example of polynomials satisfying (3-10) are the hypergeometric polynomials

$$\frac{(2a)_n}{(a+bi)_n}\,{}_2F_1(-n, a+bi; 2a; 1-z), \quad a > 0, \ \ b \in \mathbb{R}, \ \ n \geq 1, \tag{3-20}$$

where $(a)_n$ denotes the Pochhammer symbol defined by $(a)_0 := 1$ and $(a)_n := a(a+1)\cdots(a+n-1)$, and the recurrence coefficients are particularly chosen as

$$\beta_n = \frac{a+n-bi}{a+n+bi} \in \partial\mathbb{D}, \quad \gamma_n = \frac{n(2a+n-1)}{(a+bi+n-1)(a+bi+n)} \in \mathbb{C}\setminus\{0\}. \tag{3-21}$$

It is easy to check that the polynomials (3-20) have all their zeros on $\partial\mathbb{D}$. Thus, in accordance with Theorem 3.4, there exists a unique nontrivial probability measure $d\sigma$ supported on $\partial\mathbb{D}$ such that the polynomials (3-20) are the corresponding POPUC.

These polynomials are a particular case of the so-called Askey POPUC; see [Castillo 2015; Castillo et al. 2014; Dimitrov and Sri Ranga 2013].

Now we can go a step further with respect to the above example in order to obtain of some known results as a direct consequence of our study.

**Example 3.6** (Delsarte–Genin POPUC). Let $\{b_n\}_{n\geq 0}$ be an arbitrary sequence of nonzero real numbers, let $\{a_n\}_{n\geq 1}$ be an arbitrary positive chain sequence [Chihara 1978], and let $\{\varphi_n\}_{n\geq 0}$ be the sequence of polynomials recursively defined by

$$\varphi_{n+1}(z) = \big((1+ib_n)z + (1-ib_n)\big)\varphi_n(z) - 4a_n z\varphi_{n-1}(z), \tag{3-22}$$

with initial conditions $\varphi_0 := 1$ and $\varphi_1(z) := (1+ib_0)z + (1-ib_0)$. It is worth mentioning that (3-22) is the recurrence relation studied by Delsarte and Genin [1988; 1991a; 1991b; 1990], among others. The interlacing property on $\partial\mathbb{D}$ of the zeros of $\{\varphi_n\}_{n\geq 0}$ was recently proved in [Dimitrov and Sri Ranga 2013, Theorem 1.1]; see also [Castillo et al. 2014], although it was first proved in [Delsarte and Genin 1988, Section 5]. In any case, an easy computation shows that these polynomials satisfy the conditions of Theorem 3.4 with all their zeros on $\partial\mathbb{D}$. So, the interlacing property of the zeros of $\{\varphi_n\}_{n\geq 0}$ is also a direct consequence of the fact that from Theorem 3.4, these polynomials are POPUC associated with some nontrivial probability measure $d\sigma$ supported on $\partial\mathbb{D}$. Actually, we can say much more about the behavior of their zeros using the known results for the zeros of POPUC; see, e.g., [Cantero et al. 2002; Golinskii 2002; Simon 2005a; 2005b; 2007; 2011; Wong 2007].

An interesting and nontrivial extension of the results of this section is the connection with those obtained in [Lamblém et al. 2010] where a non-Hermitian linear functional $\tilde{\mu}$ on $\Lambda$ satisfying $c_n = c_{-n} \in \mathbb{C}$ is considered (compare with (2-1)) yielding the definition of Szegő-type polynomials. We recall that the case $c_n = c_{-n} \in \mathbb{R}$ was previously considered in [Delsarte and Genin 1986].

## 4. Applications

The aim of this last section is to establish two applications of the results presented in the previous section.

***Analytic theory of polynomials.*** The major role of POPUC, as it was pointed out in the Introduction is played by the zero behavior. The results in this direction can be divided into two sets, depending on the methodology used by the authors. The first one is composed by Cantero, Moral and Velázquez [Cantero et al. 2002], Golinskii [2002], and Wong [2007], whose basic tool is the Christoffel–Darboux formula. The second one is by Simon [2007] who used the theory of rank-one perturbations of unitary matrices.

Although many of the results of zeros of POPUC are well known, a natural question is still open: are two polynomials with simple and strictly interlacing zeros on $\partial\mathbb{D}$, with the possible exception of one of them which could be common, elements of a sequence of POPUC? After the previous section it is natural to conjecture that the answer will be yes. The following result for OPRL goes back to the work of Wendroff [1961]. We must emphasize that it was known by Geronimus [1946, pp. 744] also for OPUC.

**Theorem 4.1** (Geronimus–Wendroff theorem). *Let $\Psi_n$ and $\Psi_{n+1}$ be two monic polynomials whose zeros are simple and strictly interlacing on $\partial\mathbb{D}$. Then there exists a measure $d\sigma$ for which they are POPUC. All such measures have the same $\Psi_j$, $0 \le j \le n+1$. Moreover, if $\Psi_n$ and $\Psi_{n+1}$ have at most one zero in common, the statement of the theorem is also true.*

*Proof.* Let us assume that the zeros of $\Psi_n$ and $\Psi_{n+1}$ are strictly interlacing on $\partial\mathbb{D}$. The same idea can be used for the case that $\Psi_n$ and $\Psi_{n+1}$ have one zero in common. Let $\{e^{i\theta_{n,k}}\}_{k=1}^{n}$ be the zeros of $\Psi_n$. Set

$$\beta_n := -e^{i\sum_{k=1}^{n+1}\theta_{n+1,k}}e^{-i\sum_{k=1}^{n}\theta_{n,k}} \in \partial\mathbb{D}.$$

Notice that the polynomial

$$\prod_{k=1}^{n+1}(z-e^{i\theta_{n+1,k}}) - (z+\beta_n)\prod_{k=1}^{n}(z-e^{i\theta_{n,k}})$$

has a zero at $z=0$. Then,

$$\Psi_{n+1}(z) - (z+\beta_n)\Psi_n(z) = -\gamma_n z B_r(z),$$

where $B_r$ is a monic polynomial of degree at most $n-1$. Since $e^{i\theta_{n+1,k}} - \beta_n \neq 0$ and $\Psi_n(e^{i\theta_{n+1,k}}) \neq 0$, we have that $\gamma_n \neq 0$ and $B_r(e^{i\theta_{n+1,k}}) \neq 0$. Furthermore,

$$\Psi_{n+1}(e^{i\theta_{n,k}}) = -\gamma_n e^{i\theta_{n,k}} B_r(e^{i\theta_{n,k}}). \tag{4-23}$$

It is known that an arbitrary polynomial with simple zeros on $\partial\mathbb{D}$ is a POPUC with respect to some nontrivial probability measure supported on $\partial\mathbb{D}$ [Castillo et al. 2015]. Since $\Psi_n(\beta_n) \neq 0$ and $\Psi_{n+1}(\beta_n) \neq 0$, and we are interested in the zeros, there is no loss of generality if we assume that

$$\begin{aligned}(z+\beta_n)\Psi_n(z) &= \bar{\beta}P_{n+1}(z) - \beta P_{n+1}^*(z), \quad && \beta \in \mathbb{C}\setminus\{0\},\\ \Psi_{n+1}(z) &= \bar{\alpha}Q_{n+1}(z) - \alpha Q_{n+1}^*(z), && \alpha \in \mathbb{C}\setminus\{0\},\end{aligned}$$

where $P_{n+1}(z)$ and $Q_{n+1}(z)$ are the OPUC associated with $(z+\beta_n)\Psi_n(z)$ and $\Psi_{n+1}(z)$, respectively. Hence,

$$-\gamma_n z B_r(z) = (\bar{\alpha}P_{n+1}(z) - \bar{\beta}Q_{n+1}(z)) + (\alpha P_{n+1}^*(z) - \beta Q_{n+1}^*(z)).$$

Let us introduce two auxiliary functions

$$f_n(\theta) := \frac{\Psi_n(z)}{i z^{n/2}}, \quad g_r(\theta) := \frac{z B_r(z)}{i z^{(n+1)/2}},$$

where $(re^{i\theta})^{1/2} = \sqrt{r} e^{i\theta/2}$, $r > 0$, and $\theta \in (\tilde{\omega}, \tilde{\omega} + 2\pi)$. Clearly, $f_n(\theta)$ and $g_r(\theta)$ are real–valued $C^\infty$ functions defined on $(\tilde{\omega}, \tilde{\omega} + 2\pi)$ and, by definition they have the same number of zeros on $(\tilde{\omega}, \tilde{\omega} + 2\pi)$ as $\Psi_n$ and $B_r$ on $\partial\mathbb{D}$, respectively.

One denotes the zeros of $f_n$ by $x_{n,k}$, $k = 1, \dots, n$. As a consequence of the interlacing hypothesis, we have

$$f_{n+1}(x_{n,k+1}) f_{n+1}(x_{n,k}) < 0.$$

Therefore,

$$f_{n+1}(x_{n,k+1}) f_{n+1}(x_{n,k}) = |\gamma_n|^2 g_r(x_{n,k}) g_r(x_{n,k+1}),$$

from which

$$g_r(x_{n,k}) g_r(x_{n,k+1}) < 0.$$

This implies that $B_r$ has $n - 1$ zeros strictly interlacing on $\partial\mathbb{D}$ with the zeros of $\Psi_n$. If we define the polynomial $\Psi_{n-1} :\equiv B_r$, we can construct, just repeating the above procedure, a polynomial of degree $n - 2$ whose zeros interlace with those of $\Psi_{n-1}$. So we can find all the polynomials (uniquely determined) with degree less than $n-1$. By the above construction, the polynomials $\Psi_j$, $0 \le j \le n - 1$, are determined by $\Psi_n$ and $\Psi_{n+1}$. Finally, applying Theorem 3.4, the result is proved. □

***Computation of Szegő quadrature formulas.*** In some applications and theoretical problems, it is of interest to compute $\Phi_n(\,\cdot\,, \tau_n)$ for some $n \ge 1$ and a fixed $\tau_n \in \partial\mathbb{D}$. A motivation to this problem can be given when the estimation of the integral

$$I_\sigma(f) := \int f(z)\, d\sigma(z)$$

is considered by means of SQ formulas,

$$I_n(f) := \sum_{j=1}^{n} \lambda_j f(z_j), \quad z_j \in \partial\mathbb{D}, \quad j = 1, \dots, n, \quad z_j \ne z_k \text{ if } j \ne k, \quad n \ge 1. \tag{4-24}$$

Here, the nodes $\{z_j\}_{j=1}^{n}$ and weights $\{\lambda_j\}_{j=1}^{n}$ are determined so that $I_n(f) = I_\sigma(f)$ for all functions $f$ belonging to a subspace of $\Lambda$ whose dimension is as large as possible. The "optimal" subspace of exactness is $\Lambda_{-(n-1),n-1}$ (of dimension $2n-1$), and this one-parameter family of optimal SQ formulas can be characterized as:

(i) The nodes are the zeros of an $n$-th POPUC associated with the measure $d\sigma$.

(ii) The weights can be computed by

$$\lambda_j = -\frac{1}{2z_j} \frac{\Upsilon_n(z_j, \tau_n)}{\Phi_n'(z_j, \tau_n)} > 0, \quad j = 1, \dots, n, \tag{4-25}$$

where $\Upsilon_n(\,\cdot\,,\tau_n)$ represents the corresponding $n$-th second kind POPUC, which can be obtained from (2-6) with the same $\tau_n$ and Verblunsky coefficients $\{-\alpha_n\}_{n\geq 0}$ (see [Wong 2007]).

The positive character of the weights is of importance for stability and convergence reasons. If you fix one or two nodes in advance in (4-24) then you get an extension of the classical Gauss–Radau and Gauss–Lobatto quadrature formulas for measures supported on the real line. But the situation on the unit circle is completely different. Because of the dependence of the parameter $\tau_n$, Szegő–Radau quadrature formulas can be always constructed by taking an appropriate selection of the parameter $\tau_n$: if we want $\zeta \in \partial\mathbb{D}$ to be a node of the rule, then

$$\Phi_n(\zeta,\tau_n)=0 \quad\Longleftrightarrow\quad \zeta\,\Phi_{n-1}(\zeta)-\bar{\tau}_n\Phi^*_{n-1}(\zeta)=0 \quad\Longleftrightarrow\quad \tau_n=\zeta^{n-2}\frac{\Phi_{n-1}(\zeta)}{\overline{\Phi_{n-1}(\zeta)}},$$

and from Heine's formula [Simon 2005a, Theorem 1.5.11(a)], it is expressed in terms of $\zeta$ and the trigonometric moments of the measure $d\sigma$:

$$\tau_n=\zeta^{n-2}\frac{\Delta}{\overline{\Delta}},\qquad \Delta:=\det\begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1}\\ c_{-1} & c_0 & \cdots & c_{n-2}\\ \vdots & \vdots & & \vdots\\ c_{-n+2} & c_{-n+3} & \cdots & c_1\\ 1 & \zeta & \cdots & \zeta^{n-1}\end{bmatrix}.$$

The construction of Szegő–Lobatto quadrature formulas also requires the computation of an $n$-th POPUC with a specific value of the parameter $\tau_n$; see [Cruz-Barroso et al. 2015].

Another motivation to the same problem, concerning the construction of interpolatory quadrature formulas for the estimation of integrals with respect to measures supported on intervals of the real line can be found in [Bultheel et al. 2005]. In that paper, $n$-point positive interpolatory quadrature formulas on $[-1,1]$ are constructed by taking as nodes the real part of some of the zeros of certain POPUC, and it is also proved there that an appropriate selection of the paraorthogonality parameter makes it always possible to obtain "optimal" rules, exact in a subspace of algebraic polynomials of dimension $n+1$.

The results of Section 3 can be used to solve this problem in an alternative way, by computing directly a sequence of POPUC, instead of the associated OPUC. Indeed, let us consider first the initial conditions of the three-term recurrence for POPUC. For arbitrary $\tau_1\in\partial\mathbb{D}$, set

$$\Phi_{-1}(\,\cdot\,,\tau_{-1}):\equiv 0,\qquad \Phi_0(\,\cdot\,,\tau_0):\equiv 1,\qquad \Phi_1(z,\tau_1)=z-\bar{\tau}_1,$$

so if we define $\tau_0:=-1$, the formula for $\beta_n$ in Corollary 3.2 holds for $n=0$. Set now $\tau_2\in\partial\mathbb{D}$ also arbitrary, and $\beta_1=\tau_1\bar{\tau}_2$. From the combination of Theorem 3.1

and (3-11) with $n = 1$, and (2-6) with $n = 2$, it follows that

$$\Phi_2(z, \tau_2) = z^2 + (\bar{\tau}_2\alpha_0 - \bar{\alpha}_0)z - \bar{\tau}_2, \quad \gamma_1 = \bar{\tau}_2(\tau_1 - \alpha_0) - (\bar{\tau}_1 - \bar{\alpha}_0).$$

So, with these initial conditions, we are able to use (3-7) for the computation of a particular sequence of POPUC that depends on the free selection of $\tau_1, \tau_2 \in \partial\mathbb{D}$: for $n \geq 2$, we compute $\tau_{n+1}$ from (3-13), and hence $\beta_n$ and $\gamma_n$ are obtained from Corollary 3.2. Now, the key fact is that since the recurrence relation for the paraorthogonality parameters is invertible (see (3-14)), and both depend only on the measure $d\sigma$, the sequence $\{\tau_k\}_{k=0}^{n+1}$ can be also obtained starting from a fixed $\tau_{n+1}$ and an arbitrary $\tau_n$, until we get $\tau_2$ (from the initial conditions, $\tau_1$ will not be needed and notice that $\alpha_{-1} := -1$ always implies $\tau_0 = -1$). The remaining parameters $\beta_n$ and $\gamma_n$ in the recurrence are thus directly obtained from Corollary 3.2 for $n \geq 2$.

To end, let us illustrate the method. Despite what happens to OPRL, few measures on the unit circle provide families of POPUC that are explicitly given. A known family of measures of importance is the Jacobi-type weight functions

$$d\sigma_{\alpha,\beta}(\theta) = (1 - \cos\theta)^{\alpha+1/2}(1 + \cos\theta)^{\beta+1/2}d\theta, \quad \alpha, \beta > -1, \quad \theta \in [0, 2\pi),$$

with Verblunsky coefficients

$$\alpha_n = \frac{(-1)^n(\beta + \frac{1}{2}) - \alpha - \frac{1}{2}}{n + \alpha + \beta + 2},$$

but for which only for the four Chebyshev-type weight functions $\alpha, \beta \in \{\pm 1/2\}$ are there explicit expressions for POPUC (see [Daruis et al. 2002]). For other selections of $\alpha, \beta$, we can compute $\Phi_n(\,\cdot\,, \tau_n)$ for $d\sigma_{\alpha,\beta}$, $n \geq 2$, and a prescribed $\tau_n$ from our three-term recurrence. An example is given below.

**Example 4.2.** Notice that since $d\sigma_{\alpha,\beta}$ is a symmetric weight ($\alpha_n \in \mathbb{R}$), the polynomial $\Phi_n(\,\cdot\,, \pm 1)$ will have real coefficients, and the nonreal zeros will appear in complex conjugate pairs.

(i) Setting $\alpha = 0$, $\beta = 1$ and $\tau_1 = \tau_2 = 1$, we obtain from the forward recurrence (3-13) that $\tau_{15} = 1$ and that $\Phi_{15}(z, 1) = \sum_{j=0}^{7} a_j(z^j - z^{15-j})$ is given by $a_0 = -a_1 = -1$, $a_2 = -a_3 = -7/5$, $a_4 = -a_5 = -21/13$ and $a_6 = -a_7 = -245/143$.

(ii) Set $\alpha = \beta = 1/4$:

   (a) If $\tau_{10} = 1$ and $\tau_9 = i$, we can make use of the backward recurrence (3-14) to obtain $\tau_2 = 1$ and $\tau_1 = i$.

   (b) For the choice $\tau_{14} = (\sqrt{2}/2)(1+i)$, the corresponding POPUC and second kind POPUC have been computed from the three-term recurrence. The nodes and weights (obtained from (4-25)) of the 14-point SQ formula are displayed in Table 1.

| nodes | weights |
|---|---|
| $0.953562637438023+0.405929829869375i$ | 0.010335586974718 |
| $0.719435950626376+0.750717509685024i$ | 0.022827404435354 |
| $0.366191713059248+0.974873401744596i$ | 0.032048528041628 |
| $-0.047525423154063+1.041505278940548i$ | 0.034116841058949 |
| $-0.454499431518169+0.939759403074777i$ | 0.028200751436574 |
| $-0.789255597539979+0.686392494558016i$ | 0.016744575125821 |
| $-1.000962132835750+0.324853472172218i$ | 0.005014789370671 |
| $1.000962132835750-0.324853472172218i$ | 0.005014789370670 |
| $-0.953562637438025-0.405929829869376i$ | 0.010335586974718 |
| $0.789255597539980-0.686392494558019i$ | 0.016744575125820 |
| $-0.719435950626374-0.750717509685023i$ | 0.022827404435355 |
| $0.454499431518170-0.939759403074778i$ | 0.028200751436573 |
| $-0.366191713059248-0.974873401744598i$ | 0.032048528041627 |
| $0.047525423154063-1.041505278940550i$ | 0.034116841058948 |

**Table 1.** Nodes and weights of the 14-point SQ formula for $d\sigma_{\alpha,\beta}$ with $\alpha=\beta=1/4$, computed from the three-term recurrence relation for the corresponding POPUC and second kind POPUC.

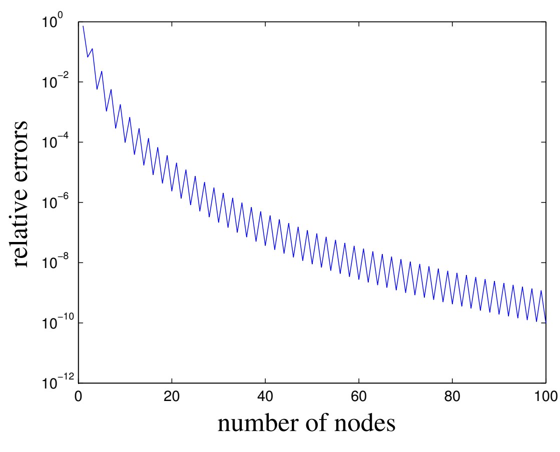


**Figure 1.** The relative errors in the estimation of $I_{\sigma_{\alpha,\beta}}(f)$ with $\alpha=\beta=1$ and $f(z)=\cos^2\theta$, with $z=e^{i\theta}$, by SQ formulas obtained from the three-term recursion for POPUC and second kind POPUC by taking $\tau_1=\tau_2=1$.

(iii) Setting $\alpha=\beta=1$, and $f(z)=\cos^2\theta$, with $z=e^{i\theta}$, the relative errors obtained in the estimation of $I_{\sigma_{\alpha,\beta}}(f)$ by using SQ formulas computed via three-term recursion for POPUC and second kind POPUC by taking $\tau_1=\tau_2=1$ are displayed in Figure 1.

For numerical reasons, the computation of the zeros of POPUC is preferably done from an eigenvalue problem of certain structured matrices (Hessenberg, CMV, snake-shaped) in a very fast and accurate way. The computations of our method can be arranged so that the nodal polynomial can be determined in only $o(n)$ arithmetic floating point operations. So, it should be said that this alternative procedure is competitive with respect to other procedures already known in the literature, but it is not really an improved algorithm. In any case, our work gives a new perspective to be considered in more detail.

## Acknowledgements

The authors thank the referee for her/his valuable suggestions and comments which have contributed to improve the final form of this paper. The research of the first author is supported by the Portuguese Government through the Fundação para a Ciência e a Tecnologia (FCT) under the grant SFRH/BPD/101139/2014 and partially supported by the Brazilian Government through the CNPq under the project 470019/2013-1 and the Dirección General de Investigación Científica y Técnica, Ministerio de Economía y Competitividad of Spain under the project MTM2012–36732–C03–01. The work of the second and third authors is partially supported by Dirección General de Programas y Transferencia de Conocimiento, Ministerio de Ciencia e Innovación of Spain under the project MTM2011–28781.

## References

[Andrews et al. 1999] G. E. Andrews, R. Askey, and R. Roy, *Special functions*, Encyclopedia of Mathematics and its Applications **71**, Cambridge Univ. Press, 1999. MR 2000g:33001 Zbl 0920.33001

[Atkinson 1964] F. V. Atkinson, *Discrete and continuous boundary problems*, Mathematics in Science and Engineering **8**, Academic Press, New York, 1964. MR 31 #416 Zbl 0117.05806

[Bultheel et al. 1999] A. Bultheel, P. González-Vera, E. Hendriksen, and O. Njåstad, *Orthogonal rational functions*, Cambridge Monographs on Applied and Computational Mathematics **5**, Cambridge Univ. Press, 1999. MR 2000c:33001 Zbl 0923.42017

[Bultheel et al. 2005] A. Bultheel, L. Daruis, and P. González-Vera, "Positive interpolatory quadrature formulas and para-orthogonal polynomials", *J. Comput. Appl. Math.* **179**:1-2 (2005), 97–119. MR 2005m:41061 Zbl 1068.41048

[Cantero et al. 2002] M. J. Cantero, L. Moral, and L. Velázquez, "Measures and para-orthogonal polynomials on the unit circle", *East J. Approx.* **8**:4 (2002), 447–464. MR 2004i:42021 Zbl 05199971

[Castillo 2015] K. Castillo, "Monotonicity of zeros for a class of polynomials including hypergeometric polynomials", *Appl. Math. Comput.* **266** (2015), 183–193. MR 3377550

[Castillo et al. 2014] K. Castillo, M. S. Costa, A. Sri Ranga, and D. O. Veronese, "A Favard type theorem for orthogonal polynomials on the unit circle from a three term recurrence formula", *J. Approx. Theory* **184** (2014), 146–162. MR 3218796 Zbl 1291.42021

[Castillo et al. 2015] K. Castillo, F. Marcellán, and M. N. Rebocho, "Zeros of para-orthogonal polynomials and linear spectral transformations on the unit circle", *Numer. Algorithms* (online publication June 2015).

[Chen 1995] W. Chen, "On the polynomials with all their zeros on the unit circle", *J. Math. Anal. Appl.* **190**:3 (1995), 714–724. MR 96b:30011 Zbl 0822.30008

[Chihara 1978] T. S. Chihara, *An introduction to orthogonal polynomials*, Mathematics and its Applications **13**, Gordon and Breach, New York, 1978. MR 58 #1979 Zbl 0389.33008

[Cruz-Barroso et al. 2007] R. Cruz-Barroso, L. Daruis, P. González-Vera, and O. Njåstad, "Sequences of orthogonal Laurent polynomials, bi-orthogonality and quadrature formulas on the unit circle", *J. Comput. Appl. Math.* **200**:1 (2007), 424–440. MR 2007j:65032 Zbl 1136.42020

[Cruz-Barroso et al. 2015] R. Cruz-Barroso, C. Díaz Mendoza, and F. Perdomo-Pío, "A connection between Szegő–Lobatto and quasi Gauss-type quadrature formulas", *J. Comput. Appl. Math.* **284** (2015), 133–143. MR 3319499 Zbl 1314.65039

[Daruis et al. 2002] L. Daruis, P. González-Vera, and O. Njåstad, "Szegö quadrature formulas for certain Jacobi-type weight functions", *Math. Comp.* **71**:238 (2002), 683–701. MR 2002k:41043 Zbl 0989.41015

[Daruis et al. 2003] L. Daruis, O. Njåstad, and W. Van Assche, "Para-orthogonal polynomials in frequency analysis", *Rocky Mountain J. Math.* **33**:2 (2003), 629–645. MR 2004j:42019 Zbl 1042.94001

[Delsarte and Genin 1986] P. Delsarte and Y. V. Genin, "The split Levinson algorithm", *IEEE Trans. Acoust. Speech Signal Process.* **34**:3 (1986), 470–478. MR 87f:94007

[Delsarte and Genin 1988] P. Delsarte and Y. Genin, "The tridiagonal approach to Szegő's orthogonal polynomials, Toeplitz linear systems, and related interpolation problems", *SIAM J. Math. Anal.* **19**:3 (1988), 718–735. MR 89d:30040 Zbl 0638.30037

[Delsarte and Genin 1990] P. Delsarte and Y. Genin, "On the role of orthogonal polynomials on the unit circle in digital signal processing applications", pp. 115–133 in *Orthogonal polynomials* (Columbus, OH, 1989), edited by P. Nevai, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **294**, Kluwer Acad. Publ., Dordrecht, 1990. MR 92b:94002 Zbl 0704.42026

[Delsarte and Genin 1991a] P. Delsarte and Y. Genin, "Tridiagonal approach to the algebraic environment of Toeplitz matrices, I: Basic results", *SIAM J. Matrix Anal. Appl.* **12**:2 (1991), 220–238. MR 92d:65071 Zbl 0728.65020

[Delsarte and Genin 1991b] P. Delsarte and Y. Genin, "Tridiagonal approach to the algebraic environment of Toeplitz matrices, II: Zero and eigenvalue problems", *SIAM J. Matrix Anal. Appl.* **12**:3 (1991), 432–448. MR 92d:65072 Zbl 0753.15015

[Delsarte et al. 1978] P. Delsarte, Y. V. Genin, and Y. G. Kamp, "Orthogonal polynomial matrices on the unit circle", *IEEE Trans. Circuits and Systems* **CAS-25**:3 (1978), 149–160. MR 58 #1981 Zbl 0408.15018

[Dimitrov and Sri Ranga 2013] D. K. Dimitrov and A. Sri Ranga, "Zeros of a family of hypergeometric para-orthogonal polynomials on the unit circle", *Math. Nachr.* **286**:17-18 (2013), 1778–1791. MR 3145170 Zbl 1290.33008

[Erdélyi et al. 1991] T. Erdélyi, P. Nevai, J. Zhang, and J. S. Geronimo, "A simple proof of "Favard's theorem" on the unit circle", *Atti Sem. Mat. Fis. Univ. Modena* **39**:2 (1991), 551–556. MR 93a:42010 Zbl 0753.33004

[Favard 1935] J. Favard, "Sur les polynômes de Tchebycheff", *C. R. Acad. Sci. Paris* **200** (1935), 2052–2055. Zbl 0012.06205

[Geronimus 1946] J. Geronimus, "On the trigonometric moment problem", *Ann. of Math.* (2) **47** (1946), 742–761. MR 8,265d Zbl 0060.16903

[Golinskii 2002] L. Golinskii, "Quadrature formula and zeros of para-orthogonal polynomials on the unit circle", *Acta Math. Hungar.* **96**:3 (2002), 169–186. MR 2003e:41048 Zbl 1017.42014

[Helsen et al. 2005] S. Helsen, A. B. J. Kuijlaars, and M. Van Barel, "Convergence of the isometric Arnoldi process", *SIAM J. Matrix Anal. Appl.* **26**:3 (2005), 782–809. MR 2006d:65037 Zbl 1082.65038

[Ismail 2005] M. E. H. Ismail, *Classical and quantum orthogonal polynomials in one variable*, Encyclopedia of Mathematics and its Applications **98**, Cambridge Univ. Press, 2005. MR 2007f:33001 Zbl 1082.42016

[Jones et al. 1989] W. B. Jones, O. Njåstad, and W. J. Thron, "Moment theory, orthogonal polynomials, quadrature, and continued fractions associated with the unit circle", *Bull. London Math. Soc.* **21**:2 (1989), 113–152. MR 90e:42027 Zbl 0637.30035

[Lamblém et al. 2010] R. L. Lamblém, J. H. McCabe, M. A. Piñar, and A. Sri Ranga, "Szegő type polynomials and para-orthogonal polynomials", *J. Math. Anal. Appl.* **370**:1 (2010), 30–41. MR 2011g:42066 Zbl 1213.42092

[Marcellán and Álvarez-Nodarse 2001] F. Marcellán and R. Álvarez-Nodarse, "On the "Favard theorem" and its extensions", *J. Comput. Appl. Math.* **127**:1-2 (2001), 231–254. MR 2001m:42048 Zbl 0970.33008

[Marcellán and Godoy 1991] F. Marcellán and E. Godoy, "Orthogonal polynomials on the unit circle: Distribution of zeros", *J. Comput. Appl. Math.* **37**:1-3 (1991), 195–208. MR 92m:42027 Zbl 0745.42013

[Peherstorfer 2011] F. Peherstorfer, "Positive trigonometric quadrature formulas and quadrature on the unit circle", *Math. Comp.* **80**:275 (2011), 1685–1701. MR 2012d:65051 Zbl 1221.42003

[Rahman and Schmeisser 2002] Q. I. Rahman and G. Schmeisser, *Analytic theory of polynomials*, London Mathematical Society Monographs. New Series **26**, Clarendon Press, Oxford, 2002. MR 2004b:30015 Zbl 1072.30006

[Simanek 2015] B. Simanek, "An electrostatic interpretation of the zeros of paraorthogonal polynomials on the unit circle", preprint, 2015. arXiv 1501.05672

[Simon 2005a] B. Simon, *Orthogonal polynomials on the unit circle, I: Classical theory*, American Mathematical Society Colloquium Publications **54**, Amer. Math. Soc., Providence, RI, 2005. MR 2006a:42002a Zbl 1082.42020

[Simon 2005b] B. Simon, *Orthogonal polynomials on the unit circle, II: Spectral theory*, American Mathematical Society Colloquium Publications **54**, Amer. Math. Soc., Providence, RI, 2005. MR 2006a:42002b Zbl 1082.42021

[Simon 2007] B. Simon, "Rank one perturbations and the zeros of paraorthogonal polynomials on the unit circle", *J. Math. Anal. Appl.* **329**:1 (2007), 376–382. MR 2008c:42027 Zbl 1110.33004

[Simon 2011] B. Simon, *Szegő's theorem and its descendants: Spectral theory for $L^2$ perturbations of orthogonal polynomials*, Princeton Univ. Press, 2011. MR 2012b:47080 Zbl 1230.33001

[Stieltjes 1894] T.-J. Stieltjes, "Recherches sur les fractions continues", *Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys.* **8**:4 (1894), J1–J122. MR 1508159 JFM 25.0326.01

[Stieltjes 1895] T.-J. Stieltjes, "Recherches sur les fractions continues [Suite et fin]", *Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys.* **9**:1 (1895), A5–A47. MR 1508160 JFM 25.0326.01

[Stone 1932] M. H. Stone, *Linear transformations in Hilbert space*, American Mathematical Society Colloquium Publications **15**, Amer. Math. Soc., Providence, RI, 1932. Zbl 0005.40003

[Szegő 1975] G. Szegő, *Orthogonal polynomials*, 4th ed., American Mathematical Society Colloquium Publications **23**, Amer. Math. Soc., Providence, RI, 1975. MR 51 #8724 Zbl 0305.42011

[Wendroff 1961] B. Wendroff, "On orthogonal polynomials", *Proc. Amer. Math. Soc.* **12** (1961), 554–555. MR 24 #A974 Zbl 0099.05601

[Wong 2007] M. L. Wong, "First and second kind paraorthogonal polynomials and their zeros", *J. Approx. Theory* **146**:2 (2007), 282–293. MR 2008f:42028 Zbl 1116.33012

Received February 9, 2015. Revised July 3, 2015.

KENIER CASTILLO
CMUC, DEPARTMENT OF MATHEMATICS
UNIVERSITY OF COIMBRA
3001-501 COIMBRA
PORTUGAL
kcastill@math.uc3m.3s
kenier@mat.uc.pt

RUYMÁN CRUZ-BARROSO
DEPARTAMENTO DE ANÁLISIS MATEMÁTICO
UNIVERSIDAD DE LA LAGUNA
38271 LA LAGUNA
TENERIFE, CANARY ISLANDS
SPAIN
rcruzb@ull.es

FRANCISCO PERDOMO-PÍO
DEPARTAMENTO DE ANÁLISIS MATEMÁTICO
UNIVERSIDAD DE LA LAGUNA
38271 LA LAGUNA
TENERIFE, CANARY ISLANDS
SPAIN
fjppio@ull.es

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.349

# SIGMA THEORY AND TWISTED CONJUGACY, II: HOUGHTON GROUPS AND PURE SYMMETRIC AUTOMORPHISM GROUPS

DACIBERG L. GONÇALVES AND PARAMESWARAN SANKARAN

**Let $\phi : \Gamma \to \Gamma$ be an automorphism of a group $\Gamma$. We say that $x, y \in \Gamma$ are in the same $\phi$-twisted conjugacy class and write $x \sim_\phi y$ if there exists an element $\gamma \in \Gamma$ such that $y = \gamma x \phi(\gamma^{-1})$. This is an equivalence relation on $\Gamma$ called the $\phi$-twisted conjugacy. Let $R(\phi)$ denote the number of $\phi$-twisted conjugacy classes in $\Gamma$. If $R(\phi)$ is infinite for all $\phi \in \mathrm{Aut}(\Gamma)$, we say that $\Gamma$ has the $R_\infty$-property.**

**The purpose of this note is to show that the symmetric group $S_\infty$, the Houghton groups and the pure symmetric automorphism groups have the $R_\infty$-property. We show, also, that the Richard Thompson group $T$ has the $R_\infty$-property. We obtain a general result establishing the $R_\infty$-property of the finite direct product of finitely generated groups.**

**This is a sequel to an earlier work by Gonçalves and Kochloukova, in which it was shown using the sigma theory of Bieri, Neumann and Strebel that, for most of the groups $\Gamma$ considered here, $R(\phi) = \infty$ where $\phi$ varies in a finite index subgroup of the automorphisms of $\Gamma$.**

## 1. Introduction

Let $\Gamma$ be a group and let $\phi : \Gamma \to \Gamma$ be an endomorphism. Then $\phi$ determines an action $\Phi$ of $\Gamma$ on itself where, for $\gamma \in \Gamma$ and $x \in \Gamma$, we have $\Phi_{\gamma(x)} = \gamma x \phi(\gamma^{-1})$. The orbits of this action are called the $\phi$-twisted conjugacy classes. We write $x \sim_\phi y$ if $x$ and $y$ are in the same $\phi$-twisted conjugacy class. Note that when $\phi$ is the identity automorphism, the orbits are the usual conjugacy classes of $\Gamma$. We denote by $\mathcal{R}(\phi)$ the set of all $\phi$-twisted conjugacy classes and by $R(\phi)$ the cardinality $\#\mathcal{R}(\phi)$ of $\mathcal{R}(\phi)$. We say that $\Gamma$ has the $R_\infty$-property if $R(\phi) = \infty$, that is, if $\mathcal{R}(\phi)$ is infinite, for every automorphism $\phi$ of $\Gamma$.

The problem of determining which groups have the $R_\infty$-property — more briefly the $R_\infty$-problem — has attracted the attention of many researchers since it was discovered that all nonelementary Gromov-hyperbolic groups have the $R_\infty$-property.

*MSC2010:* primary 20E45; secondary 20E36.

*Keywords:* twisted conjugacy, Reidemeister number, sigma theory, Houghton groups, infinite symmetric group, pure symmetric automorphism groups.

See [Levitt and Lustig 2000; Felshtyn 2001]. It is particularly interesting when the group in question is finitely generated or countable. The notion of twisted conjugacy arises naturally in fixed point theory, representation theory, algebraic geometry and number theory. In recent years the $R_\infty$-problem has emerged as an active research area.

Recall that Houghton introduced a family of groups $H_n$, $n \geq 2$, defined as follows: let $M_n := \{1, 2, \ldots, n\} \times \mathbb{N}$. Then the group $H_n$ consists of all bijections $f : M_n \to M_n$ for which there exist integers $t_1, \ldots, t_n$ such that $f(j, s) = (j, s + t_j)$ for all $s \in \mathbb{N}$ sufficiently large and all $j \leq n$. Note that necessarily $\sum_{1 \leq j \leq n} t_j = 0$. Let $Z = \left\{(t_1, \ldots, t_n) \mid \sum_{1 \leq j \leq n} t_j = 0\right\} \subset \mathbb{Z}^n \cong \mathbb{Z}^{n-1}$. One has a surjective homomorphism $\tau : H_n \to Z \cong \mathbb{Z}^{n-1}$ sending $f$ to its *translation part* $(t_1, \ldots, t_n)$ (with notation as above). It is easily verified that $\tau$ is surjective with kernel the group of all *finitary* permutations of $M_n$. K. S. Brown [1987a] showed that $H_n$ is finitely presented for $n \geq 3$ and that it is $\mathrm{FP}_{n-1}$ but not $\mathrm{FP}_n$. Note that the above definition of $H_n$ makes sense even for $n = 1$ and that we have $H_1 \cong S_\infty$. However, we treat the group $S_\infty$ separately and we shall always assume that $n \geq 2$ while considering the family $H_n$.

Next we recall the group $G_n$, the group of *pure symmetric automorphisms* of the free group $F_n$ of rank $n \geq 2$. Fix a basis $x_k$, $1 \leq k \leq n$, of $F_n$. Denote by $\alpha_{ij} \in \mathrm{Aut}(F_n)$, $1 \leq i \neq j \leq n$, the automorphism defined as $x_i \mapsto x_j x_i x_j^{-1}$, $x_k \to x_k$, $1 \leq k \leq n$, $k \neq i$. The group $G_n$ is the subgroup of $\mathrm{Aut}(F_n)$ generated by $\alpha_{ij}$, $1 \leq i \neq j \leq n$. McCool [1986] showed that $G_n$ is finitely presented where the generating relations are:

(i) $[\alpha_{ij}, \alpha_{kl}] = 1$ whenever $i, j, k, l$ are all different;

(ii) $[\alpha_{ik}, \alpha_{jk}] = 1$ and $[\alpha_{ij}\alpha_{kj}, \alpha_{ik}] = 1$ whenever $i, j, k$ are all different.

It was shown by Gonçalves and Kochloukova [2010] that $R(\phi) = \infty$ for all $\phi$ in a finite index subgroup of the group of all automorphisms of $\Gamma$ where $\Gamma = H_n, G_n$. Our main result is the following theorem. We give two proofs for the case of Houghton groups, neither of which use $\Sigma$-theory. However, we still need to use the results of [Gonçalves and Kochloukova 2010] in the case of $G_n$.

**Theorem 1.1.** *The following groups have the $R_\infty$-property*:

(i) *the group $S_\infty$ of finitary permutations of* $\mathbb{N}$,

(ii) *the Houghton groups $H_n$, $n \geq 2$, and*

(iii) *the group $G_n$, $n \geq 2$, of pure symmetric automorphisms of a free group of rank $n$.*

Recall that Richard Thompson constructed three finitely presented infinite groups $F \subset T \subset V$ around 1965 and showed that $T$ and $V$ are simple. The groups $F$, $T$, and $V$ arise as certain homeomorphism groups of the reals, the circle, and the

Cantor set respectively. Since then these constructions have been generalized by G. Higman [1974]. See also Brown [1987a], R. Bieri and R. Strebel [2014], and M. Stein [1992]. For an introduction to the Thompson groups $F$, $T$, $V$ see [Cannon et al. 1996].

**Theorem 1.2.** *The Richard Thompson group $T$ has the $R_\infty$-property.*

As the group $T$ is simple, $\Sigma$-theory yields no information about the $R_\infty$-property. The above theorem was first proved by Burillo, Matucci, and Ventura [Burillo et al. 2013]. Shortly thereafter, Gonçalves and Sankaran [2013] also independently obtained the same result.

In Section 2 we make some preliminary observations concerning the $R_\infty$-property which will be needed for our purposes. Theorem 1.1 will be established in Section 3. The $R_\infty$-property of the group $T$ will be proved in Section 4. In Section 5 we consider the $R_\infty$-property of finite direct products of groups and obtain a strengthening of a result of Gonçalves and Kochloukova [2010].

This is a sequel to the paper [Gonçalves and Kochloukova 2010]. We reassure the reader that this paper can be read independently of it. Although results from [Gonçalves and Kochloukova 2010] are used, we develop our own proof techniques to go forward.

*Note.* Just after this paper was submitted, J. H. Jo, J. B. Lee, and S. R. Lee [Jo et al. 2015] have announced almost simultaneously a proof of the $R_\infty$-property for the Houghton groups.

*If $f : X \to Y$ is a map of sets, we shall always write the argument to the right of $f$; thus $f(x)$ denotes the image of $x \in X$ under $f$.*

## 2. Preliminaries

We begin by recalling some general results concerning twisted conjugacy classes of an automorphism of a group and that of its restriction to a normal subgroup. We obtain a criterion for a periodic automorphism to have infinitely many twisted conjugacy classes. We shall also briefly recall the notion of the Bieri–Neumann–Strebel invariant and give its known description in the case of Houghton groups and the pure symmetric automorphism groups.

**2A. *Addition formula.*** The following addition formula is found in [Gonçalves and Wong 2003, Lemma 2.1]. This is a special case of a more general formula proved in [Gonçalves and Wong 2005, §2]. For any element $g \in G$, we shall denote by $\iota_g$ the inner automorphism $x \mapsto gxg^{-1}$ of $G$. When $N$ is a normal subgroup of $G$, we shall abuse notation and denote by the same symbol $\iota_g$ the automorphism of $N$ obtained by restriction of $\iota_g$ to $N$.

**Lemma 2.1.** *Suppose that we have a commutative diagram of homomorphisms of groups where the vertical arrows are isomorphisms and horizontal rows are short exact sequences*:

$$\begin{array}{ccccccccc} 1 & \to & N & \xrightarrow{i} & G & \xrightarrow{p} & G/N & \to & 1 \\ & & \downarrow \theta' & & \downarrow \theta & & \downarrow \bar{\theta} & & \\ 1 & \to & N & \xrightarrow{i} & G & \xrightarrow{p} & G/N & \to & 1 \end{array}$$

*Then*:

(i) *One has an exact sequence of* (*pointed*) *sets* $\mathcal{R}(\theta') \xrightarrow{i_*} \mathcal{R}(\theta) \xrightarrow{p_*} \mathcal{R}(\bar{\theta}) \to \{0\}$. *That is*, $p_*$ *is surjective and* $\mathrm{Im}(i_*)$ *equals* $p_*^{-1}(\{N\})$.

(ii) (*Addition formula*) *Suppose* $R(\bar{\theta}) < \infty$ *and* $\mathrm{Fix}(\iota_{\alpha N} \circ \bar{\theta}) = \{N\}$ *for all* $\alpha \in G$. *Then* $R(\theta) < \infty$ *if and only if* $R(\iota_\alpha \theta') < \infty$ *for all* $\alpha \in G$. *Moreover*, *the following addition formula holds if* $R(\theta) < \infty$: $R(\theta) = \sum_{[\alpha N] \in \mathcal{R}(\bar{\theta})} R(\iota_\alpha \theta')$. $\square$

We omit the proof. Part (i) is trivial. As mentioned above, the addition formula is also a known result. In any case, it can be proved in a straightforward manner. It can also be proved easily using the fixed point version of the following six-term exact sequence of sets due to P. R. Heath [2015, equation (2), p. 4] (cf. [Heath 1985, Theorem 1.8]), where $\bar{\alpha}$ denotes $\alpha N \in G/N$:

$$1 \to \mathrm{Fix}(\iota_\alpha \theta') \to \mathrm{Fix}(\iota_\alpha \theta') \to \mathrm{Fix}(\iota_{\bar{\alpha}} \bar{\theta}) \to \mathcal{R}(\iota_\alpha \theta') \to \mathcal{R}(\iota_\alpha \theta) \to \mathcal{R}(\iota_{\bar{\alpha}} \bar{\theta}) \to 1.$$

**Remark 2.2.** Note that if $G/N \cong \mathbb{Z}^n$, $n < \infty$, and if 1 is not an eigenvalue of the matrix of $\bar{\theta}$ with respect to a basis of $G/N$, then, for any $\alpha \in G$, we know that $\mathrm{Fix}(\iota_{\alpha N} \circ \bar{\theta}) = \mathrm{Fix}(\bar{\theta})$ consists only of the trivial element. So the lemma implies that if $R(\theta') = \infty$, then $R(\theta) = \infty$.

**2B.** ***Periodic outer automorphisms.*** Let $\Gamma$ be a group with infinitely many conjugacy classes. Then, for any automorphism $\phi : \Gamma \to \Gamma$ and any $g \in G$, we have $R(\phi) = R(\iota_g \circ \phi)$ where $\iota_g$ denotes the inner automorphism $x \mapsto gxg^{-1}$. Indeed, it is readily seen that the $\phi$-twisted conjugacy classes are the same as the left translation by $g$ of the $\iota_g \circ \phi$-twisted conjugacy classes. Thus $\Gamma$ has the $R_\infty$-property if and only if $R(\phi) = \infty$ for a set of coset representatives of $\mathrm{Out}(\Gamma) = \mathrm{Aut}(\Gamma)/\mathrm{Inn}(\Gamma)$. We have the following lemma. Compare with [Gonçalves and Sankaran 2013, Lemma 3.4].

**Lemma 2.3.** *Let* $\theta \in \mathrm{Aut}(\Gamma)$ *and let* $n \geq 1$. *Suppose that* $\{x^n \mid x \in \mathrm{Fix}(\theta)\}$ *is not contained in the union of finitely many* $\theta^n$*-twisted conjugacy classes of* $\Gamma$. *Then* $R(\theta) = \infty$.

*Proof.* Let $x \sim_\theta y$ in $\Gamma$ where $x, y \in \mathrm{Fix}(\theta)$. Thus there exists a $z \in \Gamma$ such that $y = z^{-1} x \theta(z)$. Applying $\theta^i$ to both sides, we obtain $y = \theta^i(z^{-1}) x \theta^{i+1}(z)$, since $x, y \in \mathrm{Fix}(\theta)$. Write $\phi := \theta^n$. Multiplying these equations successively for $0 \leq i < n$, we obtain

$$y^n = \prod_{0 \le i < n} \theta^i(z^{-1}) x \theta^{i+1}(z) = z^{-1} x^n \theta^n(z) = z^{-1} x^n \phi(z).$$

That is, $y^n \sim_\phi x^n$. Our hypothesis says that there are infinitely many elements $x_k \in \mathrm{Fix}(\theta)$, $k \geq 1$, such that the $x_k^n$ are in pairwise distinct $\phi$-twisted conjugacy classes of $\Gamma$. Hence we conclude that $R(\theta) = \infty$. □

**Remark 2.4.** When $\theta^n = \iota_\gamma$ is an inner automorphism, we see from the above lemma that $R(\theta) = \infty$ if $\{x^n \gamma \mid x \in \mathrm{Fix}(\theta)\}$ is not contained in a finite union of conjugacy classes of $\Gamma$. When $\theta^n = \mathrm{id}$, we see that $R(\theta) = \infty$ if $\mathrm{Fix}(\theta)$ contains elements of order $k$ for arbitrarily large values of $k \in \mathbb{N}$.

**2C. *Σ-theory of $H_n$ and $G_n$.*** Bieri, Neumann, and Strebel [Bieri et al. 1987] introduced, for any finitely generated group $\Gamma$, an invariant $\Sigma(\Gamma)$ which is a certain open subset — possibly empty — of the character sphere $S(\Gamma) := \mathrm{Hom}(\Gamma, \mathbb{R}) \setminus \{0\}/\mathbb{R}_{>0}$ where the action of the multiplicative group of positive reals is via scalar multiplication. The automorphism group $\mathrm{Aut}(\Gamma)$ acts on $S(\Gamma)$ where $\phi^* : S(\Gamma) \to S(\Gamma)$ is defined as $[\chi] \mapsto [\chi \circ \phi]$, $[\chi] \in S(\Gamma)$, for $\phi \in \mathrm{Aut}(\Gamma)$. This action preserves the subspace $\Sigma(\Gamma)$ and hence also its complement $\Sigma^c(\Gamma)$. If the image of the antihomomorphism $\eta : \mathrm{Aut}(\Gamma) \to \mathrm{Homeo}(\Sigma^c(\Gamma))$ defined as $\phi \mapsto \phi^*$ is a finite group, then $K = \ker(\eta)$ is a finite index subgroup of $\mathrm{Aut}(\Gamma)$ which fixes every character class in $\Sigma^c(\Gamma)$. This happens, for example, if $\Sigma^c(\Gamma)$ is a nonempty finite set. If $\Sigma^c(\Gamma)$ contains a *discrete* character class $[\chi]$, that is, a class represented by a character $\chi$ whose image $\chi(\Gamma) \subset \mathbb{R}$ is infinite cyclic, then it was observed by Gonçalves and Kochloukova [2010] that the character $\chi$ itself is fixed by the action of $K$ on $\mathrm{Hom}(\Gamma, \mathbb{R})$. That is, $\chi \circ \phi = \chi$ for all $\phi \in K \subset \mathrm{Aut}(\Gamma)$. This easily implies that $R(\phi) = \infty$ by Lemma 2.1(i), taking $G = \Gamma$, $N = \ker \chi$, $\theta = \phi$ in the notation of that lemma, so that $\bar{\theta} = \mathrm{id}$.

When $\Gamma$ is $G_n$, $n \geq 3$, the group of pure symmetric automorphisms of $F_n$, L. Orlandi-Korner [2000] has determined $\Sigma^c(\Gamma)$. When $\Gamma$ is $H_n$, the Houghton group, Brown [1987b] computed the set $\Sigma^c(\Gamma)$. Using these results, Gonçalves and Kochloukova [2010], showed that if $\Gamma$ is any one of the groups $H_n$, $n \geq 2$, $G_m$, $m \geq 3$, then the image of $\eta : \mathrm{Aut}(\Gamma) \to \mathrm{Homeo}(\Sigma^c(\Gamma))$ is finite.

In the case of the Houghton group $H_n$, $n \geq 2$, it turns out that $\Sigma^c(H_n)$ is a finite set of discrete character classes $[\chi_j]$, $1 \leq j \leq n$. Explicitly, $\chi_j : H_n \to \mathbb{Z}$ may be taken to be $-\pi_i \circ \tau$ where $\tau : H_n \to Z$ is the translation part (see Section 1) and $\pi_i : Z \to \mathbb{Z}$ is the restriction to $Z \subset \mathbb{Z}^n$ of the $i$-th projection (see [Brown 1987b]). (Recall from Section 1 that $Z = \{(t_1, \ldots, t_n) \in \mathbb{Z}^n \mid \sum_{1 \leq j \leq n} t_j = 0\}$.) Thus $\mathrm{Homeo}(\Sigma^c(H_n)) \cong S_n$ is finite and so is the image of $\eta : \mathrm{Aut}(H_n) \to \mathrm{Homeo}(\Sigma^c(H_n))$. As already remarked,

$R(\phi) = \infty$ for all $\phi \in \ker(\eta)$. The lemma below will not be used in this paper but is included here for illustrative purposes.

**Lemma 2.5.** *Suppose $\eta(\phi) : \Sigma^c(H_n) \to \Sigma^c(H_n)$ is not an $n$-cycle. Then $R(\phi) = \infty$.*

*Proof.* Since $\eta(\phi)$ is not an $n$-cycle, the orbit of $[\chi_1]$ under $\eta(\phi)$ consists of at most $n-1$ elements. Since $\chi_1$ is discrete, the orbit of $\chi_1 \in \mathrm{Hom}(H_n, \mathbb{R})$ consists of discrete elements. In fact, the orbit of $\chi_1$ is a subset of $\{\chi_j \mid 1 \le j \le n\}$. Now the orbit sum $\lambda := \sum_{1\le j\le k} \chi_1\phi^j$ is a *nonzero* character since any $n-1$ elements of $\chi_j$, $1 \le j \le n$, form a *basis* of $\mathrm{Hom}(H_n, \mathbb{R})$. It follows, since $\phi^*(\lambda) = \lambda$, that $R(\phi) = \infty$. □

If $\phi^* : \Sigma^c(H_n) \to \Sigma^c(H_n)$ is an $n$-cycle, the orbit sum is zero and the above argument fails. In fact, it is easily seen that every possible permutation of $\Sigma^c(H_n)$ may be realized as $\eta(\phi)$ for some $\phi \in \mathrm{Aut}(H_n)$; that is, $\eta : \mathrm{Aut}(H_n) \to \mathrm{Homeo}(\Sigma^c(H_n)) \cong S_n$ is surjective.

## 3. Proof of Theorem 1.1

Let $X$ be an infinite set. We will only be concerned with the case when $X$ is countably infinite. We shall denote by $S_\infty(X)$ the group of all finitary permutations of $X$, that is, those permutations which fix all but finitely many elements of $X$. The group of *all* permutations of $X$ will be denoted by $S(X)$. We shall denote $S(X)$ (resp. $S_\infty(X)$) simply by $S_\omega$ (resp. $S_\infty$) when $X$ is clear from the context. If $x = (x_k)_{k\in\mathbb{Z}}$ is a doubly infinite sequence in $X$ of pairwise distinct elements, we regard it as an element of $S(X)$ where $x(x_k) = x_{k+1}$ and $x(a) = a$ if $a \ne x_k$ for all $k \in \mathbb{Z}$. Two such sequences $x = (x_k)$ and $y = (y_k)$ define the same permutation if and only if $y$ is a shift of $x$, that is, there exists an $n$ such that $x_k = y_{k+n}$ for all $k \in \mathbb{Z}$. Thus, the sequence $x = (x_k)_{k\in\mathbb{Z}}$ is just the infinite cycle $x \in S(X)$. Any $f \in S(X)$ is uniquely expressible as a product of disjoint cycles. Such an expression of $f$ is its *cycle decomposition*. The *cycle type* of an $f \in S(X)$ is the function $c(f) : \mathbb{N} \cup \{\infty\} \to \mathbb{Z}_{\ge 0} \cup \{\infty\}$ where $c(f)(\alpha)$ is the number of $\alpha$-cycles in the cycle decomposition of $f$ if that number is finite; otherwise it is $\infty$ for $\alpha \in \mathbb{N} \cup \{\infty\}$. As in the case $S_\infty(X)$, if $f$ and $g$ have the same cycle type, then they are conjugate in $S(X)$. We need a criterion for $f$ and $g$ to be conjugate by an element of $S_\infty(X)$.

**Lemma 3.1.** *Let $x = (x_k)_{k\in\mathbb{Z}}$, $y = (y_k)_{k\in\mathbb{Z}} \in S_\omega(X)$ be two disjoint infinite cycles and let $(a, b) \in S_\infty$.*

(i) *If $a = x_0$, $b = x_k$, $k > 0$, then $(a, b)x = uv$, where $u = (u_j)_{j\in\mathbb{Z}} \in S_\omega$, $v \in S_\infty$ are disjoint cycles defined by*

$$u_j = \begin{cases} x_j & j < 0, \\ x_{j+k} & j \ge 0, \end{cases}$$

*and $v = (x_0, \dots, x_{k-1}) \in S_\infty$.*

(ii) *If* $a = x_0,\ b = y_0$, *then* $(a, b)xy = uv$, *where* $u = (u_j)_{j\in\mathbb{Z}},\ v = (v_j)_{j\in\mathbb{Z}}$ *are disjoint infinite cycles defined by*

$$u_j = \begin{cases} x_j & j < 0, \\ y_j & j \geq 0, \end{cases} \quad \textit{and} \quad v_j = \begin{cases} y_j & j < 0, \\ x_j & j \geq 0. \end{cases} \qquad \square$$

If $k \in \mathbb{N}$, we denote by $\mathbb{N}_{>k}$ the set of all integers greater than $k$. Note that $S_\infty = \bigcup_{k\geq 2} S_k$ where $S_k$ is the subgroup consisting of permutations of $\mathbb{N}$ which fix all $n > k$. In particular, the group $S_\infty$ is generated by transpositions $(i, i+1),\ i \geq 1$. The alternating group $A_\infty$ equals the commutator subgroup $[S_\infty, S_\infty]$, has index 2 in $S_\infty$ and is simple. The conjugacy class of any element of $S_\infty$ is determined by its cycle type, as in the case of finite symmetric groups. The group $S_\infty$ is a normal subgroup of $S_\omega = S(\mathbb{N})$. In particular, any bijection $f : \mathbb{N} \to \mathbb{N}$ defines an automorphism $\iota_f \in \mathrm{Aut}(S_\infty)$ by restricting the inner automorphism determined by $f \in S_\omega$. Moreover $\iota_f$ is the identity automorphism only if $f$ equals the identity map. The following result is well-known. See [Scott 1987, §11.4].

**Theorem 3.2.** *The homomorphism* $\iota : S_\omega \to \mathrm{Aut}(S_\infty)$ *is an isomorphism of groups.*

The following corollary is a special case of a more general result established in [Dixon and Mortimer 1996, Theorem 8.2A]. We include a proof, which is simpler in our special case.

**Corollary 3.3.** *Suppose that* $S_\infty$ *is a characteristic subgroup of a group* $H$ *contained in* $S_\omega$. *Then the automorphism group of* $H$ *is isomorphic to the normalizer* $N(H)$ *of* $H$ *in* $S_\omega$. *In particular*, *every automorphism of* $H$ *is the restriction to* $H$ *of a unique inner automorphism of* $S_\omega$.

*Proof.* We shall use the same symbol $\iota_f$ to denote the conjugation by $f \in S_\omega$ or its restriction to any subgroup normalized by $f$.

It is evident that $\iota : N(H) \to \mathrm{Aut}(H)$ defined as $f \mapsto \iota_f$ defines an homomorphism. (Here $\iota_f(h) = fhf^{-1}$ for all $h \in H$.) This is a monomorphism since $\iota_f$ is nontrivial on $S_\infty \subset H$ if $f$ is not the identity.

Let $\phi : H \to H$ be any automorphism and let $f \in S_\omega$ be the element such that $\phi|_{S_\infty} = \iota_f$. We claim that $\phi = \iota_f$. Suppose that $u := \phi(h),\ \iota_f(h) = fhf^{-1} =: v$ for some $h \in H$. We must show that $u(i) = v(i)$ for all $i \in \mathbb{N}$. It suffices to show that $\{u(i), u(j)\} = \{v(i), v(j)\}$ for all $i, j \in \mathbb{N},\ i \neq j$. Let $i, j \in \mathbb{N},\ i \neq j$. Now consider the transposition $(a, b) \in S_\infty$ such that $\iota_f(a, b) = \phi(a, b) = (i, j)$. We have

$$\phi(h(a, b)h^{-1}) = \phi(h)\phi(a, b)\phi(h^{-1}) = u(i, j)u^{-1} = (u(i), u(j)),$$

while

$$\iota_f(h(a, b)h^{-1}) = \iota_f(h)\iota_f(a, b)\iota_f(h^{-1}) = v(i, j)v^{-1} = (v(i), v(j)).$$

Therefore $(u(i), u(j)) = (v(i), v(j)) \in S_\infty$ since $\iota_f$ and $\phi$ agree on $S_\infty$. This implies that $\{u(i), u(j)\} = \{v(i), v(j)\}$, completing the proof. $\square$

**3A.** ***$S_\infty$ has the $R_\infty$-property.*** Let $\theta \in \mathrm{Aut}(S_\infty)$. In view of Theorem 3.2, $\theta = \iota_f$ for some $f \in S_\omega$. Let $x, y \in S_\infty$ and suppose that $y = zx\theta(z^{-1}) = zxfz^{-1}f^{-1}$ for some $z \in S_\infty$. Then we have $yf = z(xf)z^{-1}$ in $S_\omega$ for some $z \in S_\infty$. For any cycle (finite or infinite) $u = (u_j)$, we have that $zuz^{-1}$ is the cycle $(z(u_j))$. Any $z \in S_\infty$ moves only finitely many elements of $\mathbb{N}$. Hence when $u$ is an infinite cycle we have $z(u_j) = u_j$ for all but finitely many $j \in \mathbb{Z}$. For an arbitrary element $u$ expressed as a product of pairwise disjoint cycles, $u(\alpha) = (u(\alpha)_j)$, the element $zuz^{-1}$ being a product of $zu(\alpha)z^{-1}$, we see that $zu(\alpha)z^{-1} = u(\alpha)$ for all but a finitely many $\alpha$, and, moreover, if $u(\alpha) = (u(\alpha)_j)_{j\in\mathbb{Z}})$ is an infinite cycle, then $z(u(\alpha)_j) = u(\alpha)_j$ for all but finitely many $j \in \mathbb{Z}$.[1]

**Lemma 3.4.** *If $f \in S_\omega$ has an infinite cycle $u$, then there exist infinitely many transpositions $\tau_k \in S_\infty$ such that $\tau_j f \neq z\tau_k fz^{-1}$ for any $z \in S_\infty$.*

*Proof.* Fix an infinite cycle $u = (u_\alpha)_{\alpha\in\mathbb{Z}}$ that occurs in the cycle decomposition of $f$. Let $\tau_\alpha = (u_0, u_\alpha)$, $\alpha \geq 1$. Then we claim that $\tau_\alpha f$ and $\tau_\beta f$ are not conjugates if $\alpha \neq \beta$. To see this, we apply Lemma 3.1 to compute $\tau_\alpha u$, $\alpha \geq 1$. Note that the cycles that occur in $\tau_\alpha u$ also occur in the cycle decomposition of $\tau_\alpha f$. This is true in particular of the infinite cycle, denoted $v(\alpha)$, that occurs in $\tau_\alpha u$.

Now $v(\alpha)_p = v(\beta)_p = u_p$ for all $p < 0$ and $\alpha, \beta \geq 1$, and, when $\alpha \neq \beta$, we have $u_{p+\alpha} = v(\alpha)_p \neq v(\beta)_p = u_{p+\beta}$, $p \geq 0$. This implies that the $zv(\beta)z^{-1}$ cannot occur in $\tau_\alpha f$ for any $z \in S_\infty$ if $\alpha \neq \beta$ in its cycle decomposition, by the assertion made in the paragraph above the statement of the lemma. Hence $\tau_\alpha f \neq z\tau_\beta fz^{-1}$ for any $z \in S_\infty$. $\square$

We are now ready to prove part (i) of Theorem 1.1, restated below:

**Theorem 3.5.** *The group $S_\infty$ has the $R_\infty$-property.*

*Proof.* Let $\theta = \iota_f \in \mathrm{Aut}(S_\infty)$ where $f \in S_\omega$. We need to show that there exist pairwise distinct elements $\tau_j \in S_\infty$, $j \in \mathbb{N}$, such that $\tau_j f \neq z\tau_k fz^{-1}$ for any $z \in S_\infty$ if $j \neq k$. Since $S_\infty$ has infinitely many conjugacy classes, the assertion holds for $f \in S_\infty$; thus we need only consider the case $f \notin S_\infty$. In the cycle decomposition of $f$, either (i) there exists an infinite cycle, or (ii) all the cycles are finite and there are infinitely many of them.

*Case* (i). In this case the assertion has already been established in Lemma 3.4.

*Case* (ii). Suppose that $f = \prod_{\alpha\in\mathbb{N}} u(\alpha)$ where the $u(\alpha)$ are all finite cycles having length $\ell(\alpha)$ at least 2 for every $\alpha \in \mathbb{N}$. Let $J := \{\alpha \in \mathbb{N} \mid \ell(\alpha) \geq 3\}$. We break up the proof into two subcases depending on whether $J$ is infinite or not.

---

[1] There is a mild abuse of notation here; $u(\alpha)$ is not to be confused with the value of $u$ at $\alpha$. We will use Greek letters as labels in such situations.

*Subcase* (a). $J$ is infinite. Let $J_k \subset J$ be the set consisting of the first $k$ elements of $J$ (with respect to the usual ordering on $J \subset \mathbb{N}$). Write $u(\alpha) = (u(\alpha)_1, \ldots, u(\alpha)_{\ell(\alpha)})$ and set $U(\alpha) := \{u(\alpha)_i \mid 1 \leq i \leq \ell(\alpha)\}$, $\alpha \in \mathbb{N}$. Consider the collection of pairwise disjoint transpositions $\lambda_\alpha = (u(\alpha)_1, u(\alpha)_2)$, $\alpha \in J$, and let $\tau_k = \prod_{\alpha \in J_k} \lambda_\alpha$. Note that

$$\lambda_\alpha u(\alpha) = (u(\alpha)_1) \cdot (u(\alpha)_2, \ldots, u(\alpha)_{\ell(\alpha)}) = (u(\alpha)_2, \ldots, u(\alpha)_{\ell(\alpha)})$$

fixes only $u(\alpha)_1$ in the set $U(\alpha)$, as $\ell(\alpha) \geq 3$. Then $\tau_k \cdot \prod_{\alpha \in J_k} u(\alpha)$ fixes only $u(\alpha)_1 \in \mathbb{N}$, $\alpha \in J_k$, in the set $\bigcup_{\alpha \in J_k} U(\alpha)$. Let $F_0 = \mathrm{Fix}(f)$. Then $\mathrm{Fix}(\tau_k f) = F_0 \cup \{u(\alpha)_1 \mid \alpha \in J_k\} =: F_k$.

Suppose that $\tau_j f = z \tau_k f z^{-1}$ with $z \in S_\infty$ and $j \neq k$. Then $z$ defines a bijection $\zeta : F_j \to F_k$ between the fixed sets of $\tau_j f$ and $\tau_k f$. Clearly this is a contradiction if $\mathrm{Fix}(f) = F_0$ is finite. Assume that $F_0 \subset \mathbb{N}$ is infinite. Since $z$ is in $S_\infty$, it fixes all but finitely many elements of $F_0$. Let $L := \{m \in F_0 \mid z(m) \neq m\}$. Note that $\zeta$ restricts to the identity on $F_0 \setminus L$. Therefore $\zeta$ restricts to a bijection between $L \cup \{u(\beta)_1 \mid \beta \in J_j\}$ and $L \cup \{u(\beta)_1 \mid \beta \in J_k\}$. Since $j \neq k$, we have that $L$ is finite and $L \subset F_0$ is disjoint from $\{u(\beta)_1 \mid \beta \in J_n\}$, $n = j, k$, which is a contradiction.

*Subcase* (b). The set $J$ is finite; we set $K = \mathbb{N} \setminus J$ and define $K_j$, $j \in \mathbb{N}$, to be the set of first $\alpha$ elements of $K$. Again we set $\lambda_\alpha = (u(\alpha)_1), u(\alpha)_2) = u(\alpha)$, $\alpha \in K$. Now, if $\alpha \in K$, we have $\lambda_\alpha u(\alpha) = \mathrm{id}$; that is, $\lambda_\alpha u(j)$ fixes both points of $U(\alpha)$. We set $\tau_j := \prod_{\alpha \in K_j} \lambda_\alpha$ and $F_j := \mathrm{Fix}(\tau_j f) = F_0 \bigcup_{\alpha \in K_j} U(\alpha)$. Arguing exactly as above, for any $z \in S_\infty$, we see that $\tau_j f = z \tau_k f z^{-1}$ implies $j = k$, completing the proof. □

**3B.** ***Houghton groups.*** As in the introduction, $H_n$, $n \geq 2$, denotes the Houghton group. We first describe the group of outer automorphisms of $H_n$. Recall from Section 1 that one has an exact sequence

$$1 \to S_\infty(M_n) \hookrightarrow H_n \xrightarrow{\tau} Z \to 1$$

where $\tau : H_n \to Z$ sends $f \in H_n$ to the translation part $(t_1, \ldots, t_n) \in Z$ of $f$. The group $S_\infty(M_n)$ is the commutator subgroup of $H_n$ if $n \geq 3$. When $n = 2$, the commutator subgroup is the alternating group $A_\infty(M_2)$ which has index 2 in $S_\infty(M_2)$. In any case, $S_\infty = S_\infty(M)$ is characteristic in $H_n$ as $H_n/S_\infty$ is the maximal *torsion-free* abelian quotient of $H_n$.

**Lemma 3.6.** *Let $\phi : H_n \to H_n$, $n \geq 2$, be an automorphism. Then $\phi$ is inner if and only if $\bar{\phi} : Z \to Z$ is the identity automorphism.*

*Proof.* It is trivial to see that any inner automorphism of $H_n$ induces the identity automorphism of $Z$. For the converse, suppose that $\phi : H_n \to H_n$ induces the identity automorphism of $Z$.

Let $f \in S(M_n)$ be such that $\iota_f(H_n) = H_n$.

Consider the element $h_p : M_n \to M_n,\ 1 \le p < n$, in $H_n$ defined as follows:[2]

$$h_p(i,k) = \begin{cases} (p, k+1) & \text{if } i = p,\ k \ge 1, \\ (n, k-1) & \text{if } i = n,\ k > 1, \\ (p, 1) & \text{if } i = n,\ k = 1, \\ (i,k) & \text{if } i \ne p, n. \end{cases}$$

Thus $h_p$ permutes $\{p, n\} \times \mathbb{N}$ in a single cycle,

$$h_p = (\dots, (n,2), (n,1), (p,1), (p,2), \dots, (p,k), \dots),$$

and so $f h_p f^{-1}$ is the cycle

$$f h_p f^{-1} = (\dots, f(n,2), f(n,1), f(p,1), f(p,2), \dots, f(p,k), \dots) \in H_n.$$

The only infinite cycles in $H_n$ are those whose terms, except for a finite part of the cycle, are *consecutive* numbers along two rays, say $\{i_n\} \times \mathbb{N}$ and $\{i_p\} \times \mathbb{N}$, in the negative and positive directions respectively of the cycle $f h_p f^{-1}$. Therefore we have $\tau(f h_p f^{-1}) = e_{i_p} - e_{i_n}$. Moreover, there exist integers $t_n$, $t_p$ such that $f(n,k) = (i_n, k + t_n)$ and $f(p,k) = (i_p, k + t_p)$ for sufficiently large $k$. Clearly $i_n$ and $t_n$ are independent of $p$. Since $f$ is a bijection, the association $p \mapsto i_p$ is a permutation $\pi_f \in S_n$, and consequently $\sum_{1 \le q \le n} t_q = 0$. Note that $\pi_f = \mathrm{id}$ if and only if $f \in H_n$.

Since $S_\infty$ is characteristic in $H_n$, by Corollary 3.3, $\phi = \iota_g$ for a unique $g \in S(M_n)$. *We claim that* $g \in H_n$. Since $\tau(g h g^{-1}) = \tau(\phi(h)) = \tau(h)$ for all $h \in H_n$, we have $\pi_g(q) = q$ for all $q \le n$ and so we have $g \in H_n$. □

The group $S_n$ acts on the set $M_n = \{1, \dots, n\} \times \mathbb{N}$ in the obvious manner, by acting via the identity on $\mathbb{N}$. This defines an action $\psi$ of $S_n$ on the group $S(M_n)$ defined as $f \mapsto \sigma \circ f \circ \sigma^{-1}$ which preserves the subgroup $H_n$. Thus we obtain a homomorphism $\psi : S_n \to \mathrm{Aut}(H_n)$. It is readily seen that $\tau(\psi_\sigma(h)) = \sigma(\tau(h))$ for all $h \in H_n$, where $\sigma$ acts on $Z \subset \mathbb{Z}^n$ by permuting the standard basis elements $e_1, \dots, e_n$. In particular $\psi$ is a monomorphism. Let $\bar{\psi} : S_n \to \mathrm{Out}(H_n)$ be the composition of $\psi$ with the projection $\mathrm{Aut}(H_n) \to \mathrm{Out}(H_n)$.

**Proposition 3.7.** *The homomorphism* $\bar{\psi} : S_n \to \mathrm{Out}(H_n)$ *is an isomorphism and so* $\mathrm{Aut}(H_n) = \mathrm{Inn}(H_n) \rtimes S_n \cong H_n \rtimes S_n$.

*Proof.* Lemma 3.6 shows that $\bar{\psi}$ is a monomorphism. We shall show that it is surjective.

Let $\phi \in \mathrm{Aut}(H_n)$. Write $\phi = \iota_f$ for a (unique) $f \in S(M_n)$. With notation as in the proof of Lemma 3.6, let $\pi := \pi_f \in S_n$.

[2]The element $(p,k) \in M_n$ should not be confused with the transposition in $S(\mathbb{N})$.

Consider the automorphism $\psi_\pi^{-1}\phi =: \theta$. We have

$$\tau(\theta(h_p)) = \pi^{-1}(\tau(\phi(h_p))) = \pi^{-1}(\tau(f h_p f^{-1})) = \pi^{-1}(e_{\pi(p)} - e_{\pi(n)}) = e_p - e_n = \tau(h_p)$$

for $1 \leq p < n$. Since the group $Z$ is generated by $\tau(h_p)$, $1 \leq p < n$, it follows by Lemma 3.6 that $\theta$ is inner. Hence $\bar{\psi}(\pi) = \phi \pmod{\mathrm{Inn}(H_n)}$.

Finally, note that $\mathrm{Inn}(H_n) \cong H_n$ since the center of $H_n$ is trivial. □

The above description of $\mathrm{Aut}(H_n)$ has been obtained by Burillo, Cleary, Martino, and Röver [Burillo et al. 2014, Theorem 2.2] and also by Cox [2014, §2.2]. All the proofs make essential use of Theorem 3.2 and Corollary 3.3. The proof given by Burillo et al. and our proof seem to be based on the same idea, although conceived of independently.

**Theorem 3.8.** *The Houghton group $H_n$ has the $R_\infty$-property for any $n \geq 2$.*

We shall give two proofs for part (ii) of Theorem 1.1, restated above. The first one uses the structure of the automorphism group of $H_n$ and is more direct. The second one uses the result of Theorem 3.5 and the addition formula (Lemma 2.1).

*First proof.* Observe that there are infinitely many conjugacy classes in $H_n$ since two elements in $S_\infty = S_\infty(M_n) \subset H_n$ are conjugates in $H_n$ only if they have the same cycle type. It follows that $R(\phi) = \infty$ for any inner automorphism $\phi$ of $H_n$. Therefore, to show that $R(\phi) = \infty$ for an *arbitrary* $\phi \in \mathrm{Aut}(H_n)$, it suffices to show that $R(\phi) = \infty$ for all $\phi$ in a set of coset representatives of elements of $\mathrm{Out}(H_n)$. Thus we need only show that $R(\psi_\sigma) = \infty$ for any $\sigma \in S_n$, where $\psi : S_n \to \mathrm{Aut}(H_n)$ is as defined in the paragraph above Proposition 3.7. We shall use Lemma 2.3 and Remark 2.4 to achieve this.

For $k \geq 1$, consider the element $\xi_k$ which is defined as the product of $k$-cycles $((i,1), \ldots, (i,k)) \in H_n$, $1 \leq i \leq n$. Explicitly,

$$\xi_k(i,j) = \begin{cases} (i, j+1) & \text{if } 1 \leq j < k, \\ (i,1) & \text{if } j = k, \\ (i,j) & \text{if } j > k, \end{cases}$$

for all $i \leq n$. Then $\xi_k$ is fixed by $\psi_\sigma$ for every $\sigma \in S_n$. Thus, $\{\xi_k^n \mid k \geq 1\}$ contains elements of arbitrarily large orders and so by Remark 2.4 it follows that $R(\psi_\sigma) = \infty$ for all $\sigma \in S_n$, completing the proof. □

*Second proof.* Consider the exact sequence $1 \to S_\infty(M_n) \to H_n \to Z \to 0$. As remarked already, $S_\infty(M_n)$ is characteristic in $H_n$ and we have $Z \cong \mathbb{Z}^{n-1}$. Thus any automorphism $\theta$ of $H_n$ restricts to an automorphism $\theta'$ of $S_\infty(M_n)$ and induces an automorphism $\bar{\theta}$ of $Z$. If $R(\bar{\theta}) = \infty$ then, by Lemma 2.1(i), we have $R(\theta) = R(\bar{\theta}) = \infty$. Now suppose that $R(\bar{\theta}) < \infty$. Then $\mathrm{Fix}(\bar{\theta}) = 0$. Since $Z$ is abelian and since $R(\theta') = \infty$ by Theorem 3.5, the addition formula (Lemma 2.1(ii)) yields $R(\theta) = R(\theta') = \infty$, completing the proof. □

**3C.** ***The group of pure symmetric automorphisms.*** Recall that $G_n \subset \mathrm{Aut}(F_n)$, $n \geq 2$, denotes the group of pure symmetric automorphisms of the free group $F_n$ of rank $n$. A presentation for $G_n$, obtained by McCool [1986], was recalled in Section 1. It is immediate from this presentation that the abelianization $G_n^{\mathrm{ab}} = G_n/[G_n, G_n]$ is isomorphic to $\mathbb{Z}^{n^2-n}$ with basis the images $\bar{\alpha}_{ij}$, $1 \leq i \neq j \leq n$. We denote by $\{\chi_{ij} \mid 1 \leq i \neq j \leq n\}$ the basis of $\mathrm{Hom}(G_n^{\mathrm{ab}}, \mathbb{Z})$ dual to the basis $\{\bar{\alpha}_{ij} \mid 1 \leq i \neq j \leq n\}$. We shall denote by the same symbol $\chi_{ij}$ the composition $G_n \to G_n^{\mathrm{ab}} \xrightarrow{\chi_{ij}} \mathbb{Z} \hookrightarrow \mathbb{R}$. We will assume that $n \geq 3$, leaving out $G_2$ which is isomorphic to a free group of rank 2 and which is known to have the $R_\infty$-property.

We begin by recalling the explicit description of $\Sigma^c(G_n)$ due to Orlandi-Korner [2000].

Let $A_{ij} := \mathbb{R}\chi_{ij} + \mathbb{R}\chi_{ji}$ and $B_{ijk} := \mathbb{R}(\chi_{ij} - \chi_{kj}) + \mathbb{R}(\chi_{jk} - \chi_{ik}) + \mathbb{R}(\chi_{ki} - \chi_{ji})$, with $i, j, k$ pairwise distinct. Note that $A_{ij} = A_{ji}$ and $B_{ijk} = B_{pqr}$ if $\{i, j, k\} = \{p, q, r\}$. Let $S$ be the union of vector subspaces $S = \bigcup A_{pq} \cup \bigcup B_{ijk} \subset \mathrm{Hom}(G_n, \mathbb{R})$ where the unions are over all pairs of distinct numbers $p, q \leq n$ and all pairwise distinct numbers $i, j, k \leq n$. It was shown by Orlandi-Korner [2000] that $\Sigma^c(G_n)$ is the image of $S \setminus \{0\} \subset \mathrm{Hom}(G_n, \mathbb{R}) \setminus \{0\}$.

Let $\mathcal{S}_n$ denote the semidirect product $C_2^n \rtimes S_n$ where $S_n$ acts on $C_2^n$ by permuting the coordinates. Here $C_2 = \{1, -1\}$. The group $\mathcal{S}_n$ acts effectively on $F_n$, the free group with basis $\{x_1, \dots, x_n\}$ where $\pi \in S_n$ permutes the generators: we have the equality $\pi(x_j) = x_{\pi(j)}$, $1 \leq j \leq n$, and the action of the $k$-th factor of $C_2^n$ is given by the automorphism $t_k(x_k) = x_k^{-1}$, $t_k(x_j) = x_j$, $j \neq k$. Thus $\mathcal{S}_n$ is a subgroup of $\mathrm{Aut}(F_n)$. It is readily verified that $\mathcal{S}_n$ normalizes $G_n$: $t_k \alpha_{i,j} t_k^{-1} = \alpha_{i,j}^{-1}$ if $k = j$ and equals $\alpha_{i,j}$ otherwise; if $\pi \in S_n$, then $\pi \alpha_{i,j} \pi^{-1} = \alpha_{\pi(i),\pi(j)}$ for all $i, j$. In particular, $\pi^*(A_{ij}) = A_{\pi(i)\pi(j)}$ and $\pi^*(B_{ijk}) = B_{\pi(i)\pi(j)\pi(k)}$ for all $\pi \in S_n$. Thus we have the following lemma:

**Lemma 3.9.** *Let $n \geq 3$. The action of the group $\mathcal{S}_n \subset \mathrm{Aut}(F_n)$ on $\mathrm{Hom}(G_n, \mathbb{R})$ and on $\Sigma^c(G_n)$ is defined by $\pi^*(\chi_{i,j}) = \chi_{\pi(i),\pi(j)}$, $t^*(\chi_{i,j}) = t_i t_j \chi_{i,j}$, for all $\pi \in S_n$, $t = (t_1, \dots, t_n) \in C_2^n$.* □

The following proposition is a refinement of a statement in the proof of [Gonçalves and Kochloukova 2010, Theorem 4.11].

**Proposition 3.10.** *There exists a surjective homomorphism $\eta : \mathrm{Aut}(G_n) \to \mathcal{S}_n$ such that $\phi^*(\chi_{i,j}) = \epsilon_{i,j} \chi_{\sigma(i),\sigma(j)}$, $1 \leq i \neq j \leq n$, where $\epsilon_{i,j} \in \{1, -1\}$ and $\sigma = \eta(\phi) \in S_n$. In particular, $\mathrm{Aut}(G_n) \cong K \rtimes \mathcal{S}_n$ where $K = \ker(\eta)$.*

*Proof.* We see that $\phi^*$ preserves the collections of subspaces $\mathcal{A} := \{A_{ij} \mid 1 \leq i < j \leq n\}$ and $\mathcal{B} := \{B_{ijk} \mid 1 < i < j < k \leq n\}$, since $\phi^*$ is a linear isomorphism of $\mathrm{Hom}(G_n, \mathbb{R})$ and since $\phi^* : \Sigma^c(G_n) \to \Sigma^c(G_n)$ is a homeomorphism. Note that $\mathcal{B}$ is nonempty since $n \geq 3$. *In our notation $A_{pq}$, $B_{pqr}$, it is not assumed that $p < q < r$.*

It is readily seen that $(A_{pq} + A_{rs}) \cap B_{ijk} = 0$ unless $\{p, q, r, s\} = \{i, j, k\}$. On the other hand $(A_{ij} + A_{ik}) \cap B_{ijk} = \mathbb{R}(\chi_{k,i} - \chi_{j,i})$. It follows that $\phi^*$ preserves the collection of 1-dimensional spaces $\mathcal{C} := \{\mathbb{R}(\chi_{k,i} - \chi_{j,i}) \mid i, j, k \text{ pairwise distinct}\}$.

Let $\phi^*(A_{ij}) = A_{pq}$, $\phi^*(A_{ik}) = A_{rs}$, where $i, j, k$ are pairwise distinct. Then $\{p, q\} \cap \{r, s\}$ is a singleton, say $s = p$ — so that $\phi^*(A_{ik}) = A_{pr}$ — and we have $\phi^*(B_{ijk}) = B_{pqr}$. For, otherwise, $(A_{ij} + A_{ik}) \cap B_{ijk}$ is one-dimensional, whereas $\phi^*((A_{ij} + A_{ik}) \cap B_{ijk}) = (A_{pq} + A_{pr}) \cap \phi^*(B_{ijk}) = 0$.

In view of the fact that $\phi^*$ stabilizes $\mathcal{C}$, we have

$$\phi^*(\chi_{k,i} - \chi_{j,i}) = a(\chi_{r,p} - \chi_{q,p}). \tag{$*$}$$

On the other hand, we have $\chi_{k,i} \in A_{ik}$ and so $\phi^*(\chi_{k,i}) \in \phi^*(A_{ik}) = A_{pr}$ and so $\phi^*(\chi_{k,i}) = b\chi_{p,r} + c\chi_{r,p}$ for some $b, c \in \mathbb{R}$; similarly, $\phi^*(\chi_{j,i}) = b'\chi_{q,p} + c'\chi_{p,q}$ for some $b', c' \in \mathbb{R}$. Therefore,

$$\phi^*(\chi_{k,i} - \chi_{j,i}) = b\chi_{p,r} + c\chi_{r,p} - b'\chi_{q,p} - c'\chi_{p,q}. \tag{$**$}$$

Comparing $(*)$ and $(**)$ we see that $b = 0 = c'$, that is, $\phi^*(\chi_{k,i}) = c\chi_{r,p}$ and $\phi^*(\chi_{j,i}) = b'\chi_{q,p}$. Since $\phi^* : \mathrm{Hom}(G_n; \mathbb{R}) \to \mathrm{Hom}(G_n, \mathbb{R})$ preserves the lattice $\mathrm{Hom}(G_n, \mathbb{Z})$ and since $\chi_{k,i}, \chi_{j,i}$ are part of a $\mathbb{Z}$-basis of $\mathrm{Hom}(G_n, \mathbb{Z})$, we see that $c, b' = \pm 1$.

To complete the proof, we define the permutation $\sigma \in S_n$ associated to $\phi \in \mathrm{Aut}(G_n)$ as $\sigma(i) = p$ (with notation as above). Note that $\sigma$ is indeed a bijection since $\phi^*$ is an isomorphism. We define $\eta : \mathrm{Aut}(G_n) \to S_n$ by $\eta(\phi) = \sigma$. Then $\eta$ is a homomorphism of groups. It is surjective since its restriction to $S_n \subset \mathcal{S}_n$ is the identity by Lemma 3.9. This also shows that $\eta$ splits, completing the proof. $\square$

**Remark 3.11.** It seems plausible that there exists a surjective homomorphism $\tau : \mathrm{Aut}(G_n) \to \mathcal{S}_n$ that satisfies $\phi^*(\chi_{i,j}) = t_i t_j \chi_{\sigma(i),\sigma(j)}$, $1 \le i \neq j \le n$, where $\tau(\phi) = (t_1, \ldots, t_n) \in C_2^n$, $\sigma = \eta(\phi) \in S_n$. This would imply that $\mathrm{Aut}(G_n) \cong N \rtimes \mathcal{S}_n$ for a suitable subgroup $N \subset \mathrm{Aut}(G_n)$.

The above proposition says that the matrix of $\phi^*$, with respect to the basis $\{\chi_{i,j} \mid 1 \le i \neq j\}$ (ordered by, say, the lexicographic ordering of the indices $i, j$), is of the form $\phi^* = DP$ where $D$ is a diagonal matrix with eigenvalues $\pm 1$ and $P$ is a permutation matrix.

**Lemma 3.12.** *Let $T = DP$ where $D, P \in M_m(\mathbb{R})$ are such that $D$ is a diagonal matrix and $P$ is a permutation matrix. If $P = P_1 \cdots P_k$ is a cycle decomposition then there exist eigenvectors $v_1, \ldots, v_k$ which are linearly independent.*

*Proof.* The cycle decomposition allows us to express $\mathbb{R}^n$ as a direct sum $V_1 \oplus \cdots \oplus V_k$ where $V_j$ is spanned by $\{e_i \mid P_j(i) \neq i\}$. Specifically, if $P_j = (i_1, \ldots, i_k)$. Then $v_j := e_{i_1} + d_{i_1} e_{i_2} + \cdots + d_{i_1} \cdots d_{i_{k-1}} e_{i_k}$, which is the sum of the vectors in the $DP$-orbit of $e_{i_1}$, is an eigenvector of $T$ with eigenvalue $d_{i_1}\cdots d_{i_k}$. Evidently $v_1, \ldots, v_k$ are linearly independent. $\square$

We will use the above lemma to construct two linearly independent eigenvectors of $\phi^*$ (with further properties that are relevant for our purposes). Let $\sigma = \eta(\phi) \neq \mathrm{id}$ and $\phi^* = DP$ with $D$ diagonal and $P$ a permutation transformation (with respect to the basis $\{\chi_{i,j}\}$). Suppose that $\sigma$ has a $k$-cycle in its cycle decomposition, where $k > 2$. Choose any $i$ that occurs in the $k$-cycle and let $j := \sigma(i)$. Then $\chi_{i,j}$ and $\chi_{j,i}$ do not occur in the same orbit of $DP$ and therefore $v_{i,j} := \sum_{0\le r<k}(DP)^r(\chi_{i,j})$ and $v_{j,i} := \sum_{0\le j<k}(DP)^r\chi_{j,i}$ are eigenvectors of the same eigenvalue $\epsilon \in \{1, -1\}$. Without loss of generality, we assume that $i = 1,\ j = 2$ and define $v_{1,2} =: u$, $v_{2,1} =: v$. Suppose there is no such $k$-cycle in $\sigma$. Then $\sigma$ is a product of disjoint transpositions. Without loss of generality, suppose that the transposition $(1, 3)$ occurs in the decomposition. Since $n > 2$, either $\sigma$ has a fixed point, say 2, or $n > 3$ and, say, the transposition $(2, 4)$ occurs in the decomposition. In the first case, $u := \chi_{1,2} + d_{1,2}\chi_{3,2}$ and $v := \chi_{2,1} + d_{2,1}\chi_{2,3}$ are eigenvectors of $P$ and in the latter case, $u := \chi_{1,2} + d_{1,2}\chi_{3,4}$ and $v := \chi_{2,1} + d_{2,1}\chi_{4,3}$ are eigenvectors of $P$. Thus in *all* cases, $\chi_{1,2}$ occurs in $u$ and $\chi_{2,1}$ occurs in $v$ where $u$, $v$ are eigenvectors of $\phi^*$. If 1 is an eigenvalue of $\phi^*$, then $\bar{\phi}$ has a nonzero fixed element and so $R(\phi) = \infty$. Assume that $\phi^*(u) = -u,\ \phi^*(v) = -v$. Then there exist elements $\beta, \gamma \in G_n$ such that $\bar{\phi}(\bar{\beta}) = -\bar{\beta},\ \bar{\phi}(\bar{\gamma}) = -\bar{\gamma}$, where $\bar{\alpha}_{1,2}, \bar{\alpha}_{2,1}$ occur in $\bar{\beta}, \bar{\gamma}$ respectively, with coefficient 1.

Denote by $\Gamma_2 := \Gamma_2(G_n)$ the commutator subgroup of $G_n$ and by $\Gamma_3 := \Gamma_3(G_n)$ the subgroup $[G_n, \Gamma_2] \subset \Gamma_2$. Thus $G_n/\Gamma_3$ is a two-step nilpotent group and we have the following exact sequences:

$$1 \to \Gamma_3 \to G_n \to G_n/\Gamma_3 \to 1,$$
$$1 \to \Gamma_2/\Gamma_3 \to G_n/\Gamma_3 \to G_n/\Gamma_2 \to 1.$$

Since $\Gamma_2$ and $\Gamma_3$ are characteristic in $G_n$, any automorphism of $G_n$ restricts to automorphisms of $\Gamma_2$ and $\Gamma_3$ and hence induces automorphisms of the quotients $G/\Gamma_3$, $\Gamma_2/\Gamma_3$ and $G_n/\Gamma_2 = G_n^{\mathrm{ab}}$.

Let $\theta \in \mathrm{Aut}(G_n/\Gamma_3)$ be the automorphism defined by $\phi$ and $\theta'$, the restriction of $\theta$ to $\Gamma_2/\Gamma_3$. With notation as above, $[\beta, \gamma]\Gamma_3 \in \Gamma_2/\Gamma_3$ satisfies $\theta'([\beta, \gamma]\Gamma_3) = [\beta, \gamma]\Gamma_3$. By using the addition formula (Lemma 2.1), we conclude that $R(\theta) = \infty$, *provided* $[\beta, \gamma]/\Gamma_3$ is of infinite order. Granting this for the moment, by the first part of the same lemma we conclude that $R(\phi) = \infty$ using the first exact sequence above. Since $\phi \in \mathrm{Aut}(G_n)$ was arbitrary, we conclude that $G_n$ has the $R_\infty$-property. So all that remains is to show that $[\beta, \gamma]\Gamma_3$ is not a torsion element.

We use the fact that, under the surjection $\psi : G_n \to G_2$ that maps $\alpha_{i,j}$ to $\alpha_{i,j}$ when $\{i, j\} = \{1, 2\}$ and the remaining $\alpha_{i,j}$ to 1, we have that $\Gamma_k$ maps onto $\Gamma_k(G_2)$,

$k = 2, 3$. Let $\beta_2, \gamma_2 \in G_2$ be the images of $\beta$, $\gamma$ respectively under $\psi$. Then $\bar{\beta}_2 = \bar{\alpha}_{1,2}, \bar{\gamma}_2 = \bar{\alpha}_{2,1} \in G_2^{\mathrm{ab}}$. Therefore, $[\beta_2, \gamma_2]\Gamma_3(G_2) = [\alpha_{1,2}, \alpha_{2,1}]\Gamma_3(G_2)$. Since $G_2$ is a free group with basis $\{\alpha_{1,2}, \alpha_{2,1}\}$ we see that $[\alpha_{1,2}, \alpha_{2,1}]\Gamma_3(G_2)$ generates an infinite cyclic group. Hence the same is true of $[\beta, \gamma]\Gamma_3$. This completes the proof of part (iii) of Theorem 1.1, which is restated below:

**Theorem 3.13.** *The group $G_n$, $n \geq 2$, has the $R_\infty$-property.* □

## 4. The Thompson group $T$

Recall from Section 1 the description of the Richard Thompson group $T$ as the group of all orientation-preserving piecewise linear homeomorphisms of $\mathbb{S} = I/\{0, 1\}$ with slopes in the multiplicative group generated by $2 \in \mathbb{R}_{>0}$ and break points in $\mathbb{Z}[1/2]$. We regard the Thompson group $F$ as the subgroup of $T$ consisting of elements which fix the element $1 \in \mathbb{S}^1$. In this section we prove the following result.

**Theorem 4.1** [Burillo et al. 2013; Gonçalves and Sankaran 2013]. *The Richard Thompson group $T$ has the $R_\infty$-property.*

The fact that $T$ has the $R_\infty$-property was proved first by Burillo, Matucci, and Ventura [Burillo et al. 2013] (see also [Gonçalves and Sankaran 2013]). The crucial point in the proofs of the result above is the same in both of these papers and both the proofs rely on the description of the outer automorphism of $T$ (recalled in Theorem 4.2 below). However, since the approaches before getting to the main point are slightly different, we provide our proof here which may contain some features that are useful for other situations (such as in Remark 4.7 below).

It is readily seen that the reflection map $r$ defined as $r(x) = 1 - x$, $x \in [0, 1]$, induces an automorphism $\rho : T \to T$ defined as $\rho(f) = r \circ f \circ r^{-1} = r \circ f \circ r$. We now state the following result of Brin.

**Theorem 4.2** [Brin 1996]. *The group of inner automorphisms of $T$ is of index* 2 *in* $\mathrm{Aut}(T)$ *and the quotient group* $\mathrm{Out}(T)$ *is generated by $\rho$.*

As observed in Section 2B, for any group $\Gamma$ and any automorphism $\phi \in \mathrm{Aut}(\Gamma)$, and any $g \in \Gamma$, it is true that $R(\phi) = \infty$ if and only if $R(\phi \circ \iota_g) = \infty$. Therefore, to establish the $R_\infty$-property for $\Gamma$, it is enough to show that $R(\phi) = \infty$ for a set of coset representatives of $\mathrm{Out}(\Gamma)$. In the case $\Gamma = T$, in view of Theorem 4.2 due to Brin, we need only show that $R(\rho) = \infty$ and $R(\mathrm{id}) = \infty$. The latter equality is established in Proposition 4.5 as an easy consequence of Lemma 4.4 below. Since $\rho^2 = \mathrm{id}$, we may apply Remark 2.4 to show that $R(\rho) = \infty$. The main idea is to make use of homeomorphisms in $\mathrm{Fix}(\rho)$, whose supports have an arbitrarily large number of disjoint intervals in $\mathbb{S}^1$. (This was also the idea used in the proof by Burillo, Matucci, and Ventura [Burillo et al. 2013].)

**Definition 4.3.** Let $X$ be a Hausdorff topological space.

(i) The *support of* $f \in \mathrm{Homeo}(X)$ is the open set $\mathrm{supp}(f) := \{x \in X \mid f(x) \neq x\}$.

(ii) Let $\sigma : \mathrm{Homeo}(X) \to \mathbb{N} \cup \{\infty\}$ be defined as follows: $\sigma(\mathrm{id}) = 0$, if $f \neq \mathrm{id}$; $\sigma(f)$ is the number of connected components of $\mathrm{supp}(f)$, if that number is finite; otherwise $\sigma(f) = \infty$.

**Lemma 4.4.** *Let* $\Gamma \subset \mathrm{Homeo}(X)$ *and let* $\sigma$ *be as defined above. Suppose that* $\theta \in \mathrm{Homeo}(X)$ *normalizes* $\Gamma$. *Then* $\sigma(f) = \sigma(\theta f \theta^{-1})$.

*Proof.* It is clear that the number of connected components of an open set $U \subset X$ remains unchanged under a homeomorphism of $X$. The lemma follows immediately from the observation that $\mathrm{supp}(\theta f \theta^{-1}) = \theta(\mathrm{supp}(f))$. □

**Proposition 4.5.** *The groups $F$ and $T$ have infinitely many conjugacy classes.*

*Proof.* This follows from Lemma 4.4 on observing that $F$ has elements $f$ for which $\sigma(f)$ is any prescribed positive integer. Since $F \subset T$, the same is true of $T$. □

**Lemma 4.6.** *Suppose that* $h : \mathbb{R} \to \mathbb{R}$ *is an orientation-preserving homeomorphism. Then* $\mathrm{supp}(h) = \mathrm{supp}(h^k)$ *for any nonzero integer* $k$.

*Proof.* Since $\mathrm{supp}(h) = \mathrm{supp}(h^{-1})$ we may assume that $k > 0$. Since $h$ is orientation-preserving, it is order-preserving. Suppose $x \in \mathrm{supp}(h)$ so that $h(x) \neq x$, and suppose $x < h(x)$. Then applying $h$ to the inequality we obtain $h(x) < h^2(x)$ so that $x < h(x) < h^2(x)$. Repeating this argument yields $x < h(x) < \cdots < h^k(x)$ and so $x \in \mathrm{supp}(h^k)$. The case when $x > h(x)$ is analogous. Thus $\mathrm{supp}(h) \subset \mathrm{supp}(h^k)$. On the other hand, if $x \notin \mathrm{supp}(h)$, then $h(x) = x$ and so $h^k(x) = x$ for all $k$. Therefore, equality should hold, completing the proof. □

*Proof of Theorem 4.1.* By Theorem 4.2, $\mathrm{Out}(T) \cong \mathbb{Z}/2\mathbb{Z}$ is generated by $\rho$. By Proposition 4.5, $R(\mathrm{id}) = \infty$. It only remains to verify that $R(\rho) = \infty$. We apply Remark 2.4 with $\theta = \rho$, $n = 2$, $\gamma = 1$. It remains to show that $\mathrm{Fix}(\rho)$ has infinitely many elements $h$ such that the $h^2$ are pairwise nonconjugate.

Let $k \geq 1$. Let $f_k \in F \subset T$ be such that $\mathrm{supp}(f_k)$ is a subset of $(0, 1/2)$ which has exactly $k$ components. Thus, $\sigma(f_k) = k$. (It is easy to construct such an element.) Then $\mathrm{supp}(\rho(f_k)) = \mathrm{supp}(r f_k r^{-1}) = r(\mathrm{supp}(f_k)) \subset (1/2, 1)$ is disjoint from $\mathrm{supp}(f_k) \subset (0, 1/2)$. In particular, we have $f_k \cdot \rho(f_k) = \rho(f_k) \cdot f_k =: h_k$ and $\mathrm{supp}(h_k) = \mathrm{supp}(f_k) \cup r(\mathrm{supp}(f_k))$ and so $\sigma(h_k) = 2k$. Moreover, since $\rho^2 = 1$, we see that $h_k \in \mathrm{Fix}(\rho)$. By Lemma 4.6, we have $\sigma(h_k^2) = \sigma(h_k) = 2k$. It follows that $h_k^2$ are pairwise nonconjugate in $T$, completing the proof. □

**Remark 4.7.** In the case of the generalized Thompson groups $T_{n,r}$, suppose that $\theta \in \mathrm{Aut}(T_{n,r})$ is a torsion element, say of order $m$. Then our method of proof of Theorem 4.1 can be applied to show that $R(\theta) = \infty$. In fact, applying a theorem of McCleary and Rubin [2005] to the group $T_{n,r}$, we obtain that the automorphism group of $T_{n,r}$ equals its normalizer in the group of all homeomorphisms of the circle $\mathbb{S}^1 = [0, r]/\{0, r\}$. Let $\theta \in \mathrm{Aut}(T_{n,r})$ and $f \in \mathbb{S}^1$ such that $\theta(x) = fxf^{-1}$ with $f \in \mathrm{Homeo}(\mathbb{R}/r\mathbb{Z})$. Suppose $f^m = \gamma \in T_{n,r}$ so that $\theta$ represents a torsion element of $\mathrm{Out}(T_{n,r})$. If $\gamma = 1$, our method of proof of Theorem 4.1 can be applied to show that $R(\theta) = \infty$. See [Gonçalves and Sankaran 2013] for details. However, when $\gamma \neq 1$, it is not clear to us how to find elements of $\mathrm{Fix}(\theta)$ satisfying the hypotheses of Lemma 2.3. Our approach yields no information about automorphisms which represent nontorsion elements in the outer automorphism group. The study of the $R_\infty$-property for the groups $T_{n,r}$ is a work in progress.

## 5. Direct product of groups

It was shown in [Gonçalves and Kochloukova 2010, Theorem 4.8] that if we have $G = G_1 \times \cdots \times G_n$, where each $G_i$ is a finitely generated group with the property that $\Sigma^c(G_i)$ is a finite set of discrete character classes, not all of them empty, then there exists a finite index subgroup $H$ of $\mathrm{Aut}(G)$ such that $R(\phi) = \infty$ for all $\phi \in H$. Further, when each $G_i$ is a generalized Richard Thompson group $F_{n_i,\infty}$, $n_i \geq 2$, then $G$ itself has the $R_\infty$-property.

We shall strengthen the above result here. We make use (as did Gonçalves and Kochloukova [2010]) of a result of Meinert, recalled below, that describes the $\Sigma$-invariant of a direct product. (Meinert's theorem describes the $\Sigma$-invariant in the more general setting of a graph product of groups.)

Let $G = G_1 \times \cdots \times G_n$ and $r_j = rk(G_j^{\mathrm{ab}})$ so that $S(G_j) \cong \mathbb{S}^{r_j-1}$. We assume that $r_1 \geq 1$. Then $S(G) = \prod_{1\leq j\leq n} \mathrm{Hom}(G_j, \mathbb{R}) \setminus \{0\}/\sim \cong \mathbb{S}^{r-1}$ and so $S(G) \cong \mathbb{S}^{r-1}$, where $r := \sum_{1\leq j\leq n} r_j$. It is understood that $S(G_j) = \varnothing$ if $r_j = 0$. The sphere $S(G_i)$ is identified with the subspace of $S(G)$ comprising the set of points with $j$-th coordinate equal to zero for all $j \neq i$. Observe that $S(G_i) \cap S(G_j) = \varnothing$ if $i \neq j$. In order to emphasize this, we shall write $S(G_i) \sqcup S(G_j)$ to denote their union, where $S(G_i)$ and $S(G_j)$ are thought of as subspaces of $S(G)$.

Recall that $\Sigma^c(G)$ denotes the complement of $\Sigma^1(G) \subset S(G)$.

**Theorem 5.1** [Meinert 1995]. *Let $G = G_1 \times \cdots \times G_n$ be finitely generated and let $r_1 = rk(G_1^{\mathrm{ab}})$ be positive. With the above notation,* $\Sigma^c(G) = \bigsqcup_{1\leq j\leq n} \Sigma^c(G_j)$. □

We will exploit the fact that any $\phi \in \mathrm{Aut}(G)$ induces a homeomorphism of the character sphere $S(G)$ which preserves its rational structure. Recall that an element $[\chi] \in S(G)$ is called *discrete* (or *rational*) if $\mathrm{Im}(\chi) \subset \mathbb{R}$ is infinite cyclic; equivalently, $\chi$ may be chosen to take values in $\mathbb{Q} \subset \mathbb{R}$. The set of rational points in $S(G)$ is denoted by $S_{\mathbb{Q}}(G)$. We denote by $D_{\mathbb{Q}}(G)$ the set of isolated rational points in $\Sigma^c(G)$. The set of all limit points of $D_{\mathbb{Q}}(G)$ which are contained in $S_{\mathbb{Q}}(G)$ is denoted by $L_{\mathbb{Q}}(G)$. Also, we denote by $L(G)$ the set of all limit points of $\Sigma^c(G)$. Since $\Sigma^c(G)$ is closed, $L_{\mathbb{Q}}(G)$ and $L(G)$ are subsets of $\Sigma^c(G)$.

Any homeomorphism of $\Sigma^c(G)$ induced by an automorphism of $G$ maps $D_{\mathbb{Q}}(G)$, $L_{\mathbb{Q}}(G)$, $L(G)$ respectively onto itself.

We are now ready to prove the following theorem. The proof is essentially the same in spirit as that of [Gonçalves and Kochloukova 2010, Theorem 3.3]. See also [Gonçalves and Kochloukova 2010, §4c].

**Theorem 5.2.** *Suppose that $G = G_1 \times \cdots \times G_n,\ n \geq 1$, is finitely generated and that any one of the following holds*:

(i) *the set $D_{\mathbb{Q}}(G_1)$ is nonempty, finite, and contained in an open hemisphere and $D_{\mathbb{Q}}(G_j)$ is finite (possibly empty) for $2 \leq j \leq n$;*

(ii) *the set $L_{\mathbb{Q}}(G_1)$ is nonempty, finite, and contained in an open hemisphere and $L_{\mathbb{Q}}(G_j)$ is finite (possibly empty) for $2 \leq j \leq n$;*

(iii) *the set $L(G_1) \cap S_{\mathbb{Q}}(G_1)$ is nonempty, finite, and contained in an open hemisphere and $L(G_j) \cap S_{\mathbb{Q}}(G_j)$ is finite (possibly empty) for $2 \leq j \leq n$.*

*Then $G$ has the $R_\infty$-property.*

*Proof.* Suppose $\phi \in \mathrm{Aut}(G)$. We shall show that there exists a discrete character $\lambda \in \mathrm{Hom}(G, \mathbb{R})$ such that $\lambda \circ \phi = \lambda$. By the discussion in Section 2C, it follows that $R(\phi) = \infty$ and it follows that $G$ has the $R_\infty$-property.

First we suppose that $n = 1$. The theorem, then, is essentially due to Gonçalves and Kochloukova [2010]. Let $\phi^* : \Sigma^c(G) \to \Sigma^c(G)$ be the induced map, defined as $\phi^*([\chi]) = [\chi \circ \phi]$. Since $\phi^*$ is a homeomorphism, it maps isolated points to isolated points. Moreover, $\phi^*$ preserves the set of all rational points in $\Sigma^c(G)$. It follows that $\phi^*(W) = W$, where $W$ is one of the sets $D_{\mathbb{Q}}(G)$, $L_{\mathbb{Q}}(G)$ or $L(G) \cap S_{\mathbb{Q}}(G)$.

In each of the cases (i)–(iii), we see that there is a nonempty finite set of rational character classes $W(G) \subset S_{\mathbb{Q}}(G)$ that is contained in an open hemisphere and that is mapped to itself by $\phi^*$. Suppose that $[\chi] \in W(G)$ and that the orbit of $[\chi]$ under $\phi^*$, namely the set $\{(\phi^*)^j([\chi] = [\chi \circ \phi^j] \mid j \in \mathbb{N}\}$, has $k$ elements. Then the orbit sum $\lambda := \sum_{0 \leq j < k} \chi \circ \phi^j \in \mathrm{Hom}(G, \mathbb{R})$ is a nonzero discrete character invariant under $\phi^*$, as was to be shown.

Now let $n = 2$. By Meinert's theorem (Theorem 5.1) $D_{\mathbb{Q}}(G) = D_{\mathbb{Q}}(G_1) \sqcup D_{\mathbb{Q}}(G_2)$, $L_{\mathbb{Q}}(G) = L_{\mathbb{Q}}(G_1) \sqcup L_{\mathbb{Q}}(G_2)$ and $L(G) = L(G_1) \sqcup L(G_2)$.

*Case* (i). Suppose $[\chi] \in D_{\mathbb{Q}}(G_1)$, and consider the $\phi^*$-orbit of $[\chi]$, namely, the set $\{(\phi^k)^*([\chi]) = [\chi \circ \phi^k] \mid k \in \mathbb{Z}\}$. This set is finite since it is contained in $D_{\mathbb{Q}}(G) = D_{\mathbb{Q}}(G_1) \sqcup D_{\mathbb{Q}}(G_2)$, which is finite. Suppose that $[\chi \circ \phi^j],\ 0 \leq j < q$, are the distinct rational points in the orbit. Then we claim that the orbit sum $\lambda := \sum_{0 \leq j < q} \chi \circ \phi^j$ is a *nonzero* character such that $\lambda \circ \phi = \lambda$. To see that $\lambda \in \mathrm{Hom}(G, \mathbb{R})$ is nonzero, we note that its restriction to $G_1$ is the character $\lambda_J = \sum_{j \in J} \chi \circ \phi^j$ where $J := \{j < q \mid [\chi \circ \phi^j] \in D_{\mathbb{Q}}(G_1)\}$. Since $D_{\mathbb{Q}}(G_1)$ is contained in an open hemisphere, the characters $\chi \circ \phi^j$, $j \in J$, are in an open half-space of $\mathrm{Hom}(G_1, \mathbb{R})$. Therefore the same is true of their sum, $\lambda_1$, and we conclude that $\lambda \neq 0$. It is clear that $\lambda \circ \phi = \lambda$ since $[\lambda \circ \phi] = [\lambda]$ and since $\lambda$ is rational. As observed in the first paragraph of Section 2C, this implies that $R(\phi) = \infty$.

*Case* (ii). The proof in this case is almost identical, starting with $[\chi] \in L_{\mathbb{Q}}(G_1)$. We need only note that $\phi^*(L_{\mathbb{Q}}(G))$ equals $L_{\mathbb{Q}}(G)$ and that $L_{\mathbb{Q}}(G) = L_{\mathbb{Q}}(G_1) \sqcup L_{\mathbb{Q}}(G_2)$ is finite, as in case (i). The orbit sum $\lambda := \sum_{0 \leq j < q} \chi \circ \phi^j$ is again a nonzero character which is discrete and satisfies $\lambda \circ \phi = \lambda$. Again we conclude that $R(\phi) = \infty$.

*Case* (iii). Again we start with $\chi \in L(G_1) \cap S_{\mathbb{Q}}(G_1)$ and proceed as in case (ii). We leave the details to the reader.

Finally, let $n \geq 3$ be arbitrary, and let $H = G_2 \times \cdots \times G_n$. Again by Meinert's theorem, we have $D_{\mathbb{Q}}(H) = \bigsqcup_{2 \leq j \leq n} D_{\mathbb{Q}}(G_j)$; similar expressions hold for $L_{\mathbb{Q}}(H)$ and $L(H) \cap S_{\mathbb{Q}}(H)$. Our hypotheses on $G_j$ imply that one of the sets $D_{\mathbb{Q}}(H)$, $L_{\mathbb{Q}}(H)$, or $L(G) \cap S_{\mathbb{Q}}(G)$ is finite depending on case (i), (ii), and (iii), respectively. Since $G = G_1 \times H$, we are now reduced to the situation where $n = 2$, which has just been established. This completes the proof. $\square$

We conclude the paper with the following examples.

**Examples 5.3.** (i) Examples of groups with $D_{\mathbb{Q}}(G)$ nonempty, finite, and contained in an open hemisphere are known. These include nonpolycyclic nilpotent-by-finite groups of type $\mathrm{FP}_\infty$, the generalized Richard Thompson groups $F_{n,\infty}$, the double of a knot group $K$ with nonfinitely generated commutator subgroup (thus $G \cong K \star_{\mathbb{Z}^2} K$). For details see [Gonçalves and Kochloukova 2010, §4].

(ii) Examples of groups with $D_{\mathbb{Q}}(G)$ and $L_{\mathbb{Q}}(G)$ being finite sets are finite groups, the Houghton groups [Brown 1987a], the pure symmetric automorphism groups [Orlandi-Korner 2000], finitely generated infinite groups with finite abelianization (which include the generalized Richard Thompson groups $T_{n,r}$; see [Brown 1987a, p. 64]), $\mathbb{Z}^n$, $n \geq 1$, and the free groups of rank $n \geq 2$. Another class of such groups is provided by [Bieri et al. 1987, Theorem 8.1]. Consider a finitely generated group $G$ which is a subgroup of the group of all orientation-preserving PL-homeomorphisms of the interval [0, 1]. The group $G$ is said to be *irreducible* if there is no $G$ fixed point in (0, 1). The logarithms of the slopes near the end points 0, 1, define characters $\chi_0, \chi_1 : G \to \mathbb{R}$ respectively. We recall that two characters $\lambda$, $\chi$ are independent if $\lambda(\ker(\chi)) = \lambda(G)$ and $\chi(\ker(\lambda)) = \chi(G)$. It was shown in [Bieri et al. 1987, Theorem 8.1] that $\Sigma^c(G) = \{[\chi_0], [\chi_1]\}$ if $G$ is irreducible and $\chi_0$, $\chi_1$ are independent. (These points may not be in $S_{\mathbb{Q}}(G)$; see [Bieri et al. 1987, p. 470].)

(iii) Let $G = G_1 \times G_2$ where $G_1$ is a finite product of groups (with $G_1$ nontrivial) as in example (i), and where $G_2$ is a finite product of groups as in example (ii) above. Then $G$ has the $R_\infty$-property. Since there are continuously many pairwise nonisomorphic 2-generated infinite simple groups, taking $G_2$ to be any one of them, we obtain a continuous family of groups with $R_\infty$-property.

## Acknowledgments

We thank Dessislava Kochloukova for pointing out to us the papers [Cox 2014] and [Burillo et al. 2014] concerning the Houghton groups. We thank the referee for a careful reading of our paper and for pointing out to us the paper [Heath 2015] and references to Theorem 3.2, Corollary 3.3, namely [Scott 1987] and [Dixon and Mortimer 1996]. Gonçalves was partially supported by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP), Projeto Temático Topologia Algébrica, Geométrica e Diferencial 2012/24454-8. Sankaran was partially supported by Department of Atomic Energy, Government of India, under a XII Plan Project.

## References

[Bieri and Strebel 2014] R. Bieri and R. Strebel, "On groups of PL-homeomorphisms of the real line", preprint, 2014. arXiv 1411.2868

[Bieri et al. 1987] R. Bieri, W. D. Neumann, and R. Strebel, "A geometric invariant of discrete groups", *Invent. Math.* **90**:3 (1987), 451–477. MR 89b:20108 Zbl 0642.57002

[Brin 1996] M. G. Brin, "The chameleon groups of Richard J. Thompson: automorphisms and dynamics", *Inst. Hautes Études Sci. Publ. Math.* **84** (1996), 5–33. MR 99e:57003 Zbl 0891.57037

[Brown 1987a] K. S. Brown, "Finiteness properties of groups", *J. Pure Appl. Algebra* **44**:1-3 (1987), 45–75. MR 88m:20110 Zbl 0613.20033

[Brown 1987b] K. S. Brown, "Trees, valuations, and the Bieri–Neumann–Strebel invariant", *Invent. Math.* **90**:3 (1987), 479–504. MR 89e:20060 Zbl 0663.20033

[Burillo et al. 2013] J. Burillo, F. Matucci, and E. Ventura, "The conjugacy problem in extensions of Thompson's group $F$", preprint, 2013. arXiv 1307.6750v2

[Burillo et al. 2014] J. Burillo, S. Cleary, A. Martino, and C. E. Röver, "Commensurations and metric properties of Houghton's groups", preprint, 2014. arXiv 1403.0026

[Cannon et al. 1996] J. W. Cannon, W. J. Floyd, and W. R. Parry, "Introductory notes on Richard Thompson's groups", *Enseign. Math.* (2) **42**:3-4 (1996), 215–256. MR 98g:20058 Zbl 0880.20027

[Cox 2014] C. Cox, "Twisted conjugacy in Houghton's groups", preprint, 2014. arXiv 1410.7051v1

[Dixon and Mortimer 1996] J. D. Dixon and B. Mortimer, *Permutation groups*, Graduate Texts in Mathematics **163**, Springer, New York, 1996. MR 98m:20003 Zbl 0951.20001

[Felshtyn 2001] A. L. Felshtyn, "Число Райдемайстера любого автоморфизма громовской гиперболической группы бесконечно", *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*) **279**:6 (2001), 229–240. Translated as "The Reidemeister number of any automorphism of a Gromov hyperbolic group is infinite" in *Journal of Mathematical Sciences* (*New York*), 119:1 (2004) 117–123. MR 2002e:20081 Zbl 1070.20050

[Gonçalves and Kochloukova 2010] D. L. Gonçalves and D. H. Kochloukova, "Sigma theory and twisted conjugacy classes", *Pacific J. Math.* **247**:2 (2010), 335–352. MR 2011m:20099 Zbl 1207.20037

[Gonçalves and Sankaran 2013] D. L. Gonçalves and P. Sankaran, "Twisted conjugacy in Richard Thompson's group $T$", preprint, 2013. arXiv 1309.2875v2

[Gonçalves and Wong 2003] D. L. Gonçalves and P. N.-S. Wong, "Twisted conjugacy classes in exponential growth groups", *Bull. London Math. Soc.* **35**:2 (2003), 261–268. MR 2003j:20054 Zbl 1030.20020

[Gonçalves and Wong 2005] D. L. Gonçalves and P. N.-S. Wong, "Homogeneous spaces in coincidence theory II", *Forum Math.* **17**:2 (2005), 297–313. MR 2006e:55004 Zbl 1103.55002

[Heath 1985] P. R. Heath, "Product formulae for Nielsen numbers of fibre maps", *Pacific J. Math.* **117**:2 (1985), 267–289. MR 86h:55002 Zbl 0571.55002

[Heath 2015] P. R. Heath, "Groupoids and relations among Reidemeister and among Nielsen numbers", *Topology Appl.* **181** (2015), 3–33. MR 3303931 Zbl 1308.55001

[Higman 1974] G. Higman, *Finitely presented infinite simple groups*, Notes on Pure Mathematics **8**, Department of Pure Mathematics, Australian National University, Canberra, 1974. MR 51 #13049

[Jo et al. 2015] J. H. Jo, J. B. Lee, and S. R. Lee, "The $R_\infty$ property for Houghton's groups", preprint, 2015. arXiv 1412.8767

[Levitt and Lustig 2000] G. Levitt and M. Lustig, "Most automorphisms of a hyperbolic group have very simple dynamics", *Ann. Sci. École Norm. Sup.* (4) **33**:4 (2000), 507–517. MR 2002d:20066 Zbl 0997.20043

[McCleary and Rubin 2005] S. McCleary and M. Rubin, "Locally moving groups and the reconstruction problem for chains and circles", preprint, 2005. arXiv math/0510122

[McCool 1986] J. McCool, "On basis-conjugating automorphisms of free groups", *Canad. J. Math.* **38** (1986), 1525–1529. Zbl 0613.20024

[Meinert 1995] H. Meinert, "The Bieri–Neumann–Strebel invariant for graph products of groups", *J. Pure Appl. Algebra* **103**:2 (1995), 205–210. MR 96i:20047 Zbl 0838.20027

[Orlandi-Korner 2000] L. A. Orlandi-Korner, "The Bieri–Neumann–Strebel invariant for basis-conjugating automorphisms of free groups", *Proc. Amer. Math. Soc.* **128**:5 (2000), 1257–1262. MR 2000k:20046 Zbl 0963.20016

[Scott 1987] W. R. Scott, *Group theory*, 2nd ed., Dover, New York, 1987. MR 88d:20001 Zbl 0641.20001

[Stein 1992] M. Stein, "Groups of piecewise linear homeomorphisms", *Trans. Amer. Math. Soc.* **332**:2 (1992), 477–514. MR 92k:20075 Zbl 0798.20025

Received December 27, 2014. Revised June 15, 2015.

DACIBERG L. GONÇALVES
DEPARTMENT OF MATHEMATICS
INSTITUTE OF MATHEMATICS AND STATISTICS
UNIVERSITY OF SÃO PAULO
CAIXA POSTAL 66281
05314-970 SÃO PAULO-SP
BRAZIL
dlgoncal@ime.usp.br

PARAMESWARAN SANKARAN
THE INSTITUTE OF MATHEMATICAL SCIENCES
CIT CAMPUS
TARAMANI, CHENNAI 600113
INDIA
sankaran@imsc.res.in

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.371

# THE SECOND CR YAMABE INVARIANT

PAK TUNG HO

**Let $(M, \theta)$ be a compact strictly pseudoconvex CR manifold of real dimension $2n+1$ with a contact form $\theta$. Motivated by the work of Ammann and Humbert, we define the second CR Yamabe invariant, which is a natural generalization of the CR Yamabe invariant, and study its properties in this paper.**

## 1. Introduction

Let $(M, g)$ be an $n$-dimensional compact Riemannian manifold where $n \geq 3$. The Yamabe problem is to find a Riemannian metric $\tilde{g}$ conformal to $g$ such that the scalar curvature of $\tilde{g}$ is constant. Yamabe [1960] claimed to solve it. However, Trudinger [1968] realized that Yamabe's proof was incomplete, and he was able to solve the Yamabe problem when the scalar curvature of $g$ is nonpositive. When the scalar curvature of $g$ is positive, Aubin [1976] solved the case when $n \geq 6$ and $M$ is not locally conformally flat, and Schoen [1984] solved the remaining cases by using the positive mass theorem.

The method to solve the Yamabe problem was the following. If $\tilde{g} = u^{\frac{4}{n-2}} g$, where $u \in C^\infty(M)$ and $u > 0$, then

$$L_g(u) = R_{\tilde{g}} u^{\frac{n+2}{n-2}}, \tag{1-1}$$

where

$$L_g = -\frac{4(n-1)}{n-2} \Delta_g + R_g.$$

Here $\Delta_g$ is the Laplacian of $g$, and $R_g$ and $R_{\tilde{g}}$ are the scalar curvatures of $g$ and $\tilde{g}$. The Yamabe problem is to solve (1-1) with $R_{\tilde{g}}$ being constant. The Yamabe invariant $Y(M, g)$ of $(M, g)$ is defined as

$$Y(M, g) = \inf_{u \not\equiv 0,\, u \in C^\infty(M)} E(u),$$

where

$$E(u) = \frac{\int_M u L_g(u)\, dV_g}{\left(\int_M |u|^{\frac{2n}{n-2}}\, dV_g\right)^{\frac{n-2}{n}}}.$$

*MSC2010:* primary 32V20, 32V05; secondary 53C17.
*Keywords:* CR manifold, CR Yamabe problem, CR Yamabe invariant.

The key point of the resolution of the Yamabe problem is the following theorem due to Aubin [1976].

**Theorem 1.1.** *Let $(M, g)$ be a compact Riemannian manifold of dimension $n \geq 3$. If $Y(M, g) < Y(\mathbb{S}^n)$, then there exists a positive smooth function $u$ satisfying* (1-1). *Here $Y(\mathbb{S}^n)$ is the Yamabe invariant of the sphere $\mathbb{S}^n$ with respect to the standard metric.*

The strict inequality was used to show that a minimizing sequence does not concentrate at any point. Aubin [1976] and Schoen [1984] proved the following:

**Theorem 1.2.** *Let $(M, g)$ be a compact Riemannian manifold of dimension $n \geq 3$. Then $Y(M, g) \leq Y(\mathbb{S}^n)$. Moreover, the equality holds if and only if $(M, g)$ is conformally diffeomorphic to the sphere.*

These theorems solve the Yamabe problem. See also [Brendle 2005; 2007a; 2007b; Chow 1992; Schwetlick and Struwe 2003; Ye 1994] for using the flow approach to solve the Yamabe problem.

Ammann and Humbert [2006] defined the $k$-th Yamabe invariant as a generalization of the Yamabe invariant. More precisely, let

$$\lambda_1(g) < \lambda_2(g) \leq \lambda_3(g) \leq \cdots \leq \lambda_k(g) \cdots \to \infty$$

be the eigenvalues of $L_g$ appearing with multiplicities. Let $[g]$ be the conformal class of $g$. For any positive integer $k$, the $k$-th Yamabe invariant $Y_k(M, g)$ is defined by

$$Y_k(M, g) = \inf_{\tilde{g} \in [g]} \lambda_k(\tilde{g}) \operatorname{Vol}(M, \tilde{g})^{\frac{2}{n}}.$$

In particular, $Y_1(M, g) = Y(M, g)$ when the Yamabe invariant $Y(M, g)$ is nonnegative.

One can consider the following CR analogue of the Yamabe problem, the CR Yamabe problem. Suppose that $(M, \theta)$ is a compact strictly pseudoconvex CR manifold of real dimension $2n+1$ with a contact form $\theta$. The CR Yamabe problem is to find a contact form $\tilde{\theta}$ conformal to $\theta$ such that the Webster scalar curvature of $\tilde{\theta}$ is constant. Jerison and Lee [1987; 1988; 1989] solved the CR Yamabe problem when $n \geq 2$ and $M$ is not locally CR equivalent to the sphere. The remaining cases, namely when $n = 1$ or $M$ is locally CR equivalent to the sphere, were studied respectively by Gamara and Yacoub [2001] and by Gamara [2001]. See also the recent work of Cheng, Chiu and Yang [Cheng et al. 2014] and Cheng, Malchiodi and Yang [Cheng et al. 2013]. See also [Chang and Cheng 2002; Chang et al. 2010; Ho 2012; Zhang 2009] for using the flow approach to solve the Yamabe problem.

Motivated by the result of Ammann and Humbert [2006], we study the $k$-th CR Yamabe invariant in this paper. In Section 2, we define the $k$-th CR Yamabe invariant and the generalized contact form. In Section 3, we give the variational characterization of $Y_k(M,\theta)$. In Section 4, we derive the Euler–Lagrange equation for $Y_2(M,\theta)$. Sections 5 and 6 will be devoted to proving a lower bound and an upper bound for $Y_2(M,\theta)$ respectively. In Section 7, we study whether $Y_2(M,\theta)$ is attained by some contact form or generalized contact form. Finally, in Section 8, we study the properties of the $k$-th CR Yamabe invariant $Y_k(M,\theta)$.

## 2. Definitions

Suppose that $(M,\theta)$ is a compact strongly pseudoconvex CR manifold of real dimension $2n+1$ with a given contact form $\theta$. Let $u \in C^\infty(M)$, $u>0$. Then $\tilde{\theta} = u^{\frac{2}{n}}\theta$ is a contact form conformal to $\theta$, and the Webster scalar curvature $R_{\tilde{\theta}}$ of $\tilde{\theta}$ is given by

$$L_\theta(u) = R_{\tilde{\theta}} u^{1+\frac{2}{n}}. \tag{2-1}$$

Here

$$L_\theta = -\Big(2+\frac{2}{n}\Big)\Delta_\theta + R_\theta, \tag{2-2}$$

where $\Delta_\theta$ is the sub-Laplacian of $\theta$ and $R_\theta$ is the Webster scalar curvature of $\theta$. The CR Yamabe invariant is defined as

$$Y(M,\theta) = \inf_{u\not\equiv 0,\, u\in C^\infty(M)} E(u),$$

where

$$E(u) = \frac{\int_M \big(2+\frac{2}{n}\big)|\nabla_\theta u|_\theta^2 + R_\theta u^2 \, dV_\theta}{\big(\int_M |u|^{2+\frac{2}{n}} \, dV_\theta\big)^{\frac{n}{n+1}}}.$$

It is well known that $L_\theta$ has discrete spectrum

$$\operatorname{Spec}(L_\theta) = \{\lambda_1(\theta), \lambda_2(\theta), \dots\},$$

where the eigenvalues

$$\lambda_1(\theta) < \lambda_2(\theta) \le \lambda_3(\theta) \le \cdots \le \lambda_k(\theta) \cdots \to \infty$$

appear with multiplicities. The variational characterization of $\lambda_1(\theta)$ is given by

$$\lambda_1(\theta) = \inf_{u\neq 0,\, u\in C^\infty(M)} \frac{\int_M \big(2+\frac{2}{n}\big)|\nabla_\theta u|_\theta^2 + R_\theta u^2 \, dV_\theta}{\int_M u^2 \, dV_\theta}.$$

Let $[\theta]$ be the conformal class of $\theta$, i.e.,

$$[\theta] = \{\tilde{\theta} = u^{\frac{2}{n}}\theta \mid u \in C^\infty(M), u>0\}.$$

If $Y(M,\theta) \ge 0$, then it is easy to check that

$$Y(M,\theta) = \inf_{\tilde{\theta}\in[\theta]} \lambda_1(\tilde{\theta}) \operatorname{Vol}(M,\tilde{\theta})^{\frac{1}{n+1}}. \tag{2-3}$$

Following the definition of the $k$-th Yamabe invariant in [Ammann and Humbert 2006], we have the following:

**Definition.** For any positive integer $k$, the $k$-th CR Yamabe invariant is defined by

$$Y_k(M,\theta) = \inf_{\tilde{\theta}\in[\theta]} \lambda_k(\tilde{\theta}) \operatorname{Vol}(M,\tilde{\theta})^{\frac{1}{n+1}}. \tag{2-4}$$

Then it follows from (2-3) and Theorem 8.2 that

$$Y_1(M,\theta) = \begin{cases} Y(M,\theta) & \text{if } Y(M,\theta) \geq 0, \\ -\infty & \text{if } Y(M,\theta) < 0. \end{cases}$$

We write $L_+^{2+\frac{2}{n}}(M) = \{u \in L^{2+\frac{2}{n}}(M) \,|\, u \geq 0, u \not\equiv 0\}$. For $u \in L_+^{2+\frac{2}{n}}(M)$, we define $\operatorname{Gr}_k^u(C^\infty(M))$ to be the set of all $k$-dimensional subspaces of $C^\infty(M)$ such that the restriction operator to $M \setminus u^{-1}(0)$ is injective. More precisely, we have

$$\begin{aligned} &\operatorname{span}(v_1,\dots,v_k) \in \operatorname{Gr}_k^u(C^\infty(M)) \\ &\qquad\qquad \Longleftrightarrow v_1|_{M\setminus u^{-1}(0)},\dots,v_k|_{M\setminus u^{-1}(0)} \text{ are linearly independent} \\ &\qquad\qquad \Longleftrightarrow u^{\frac{1}{n}}v_1,\dots,u^{\frac{1}{n}}v_k \text{ are linearly independent.} \end{aligned}$$

Similarly, replacing $C^\infty(M)$ by $S_1^2(M)$, we obtain the definition of $\operatorname{Gr}_k^u(S_1^2(M))$. Hereafter, $S_1^2(M)$ denotes the Folland–Stein space, which is the completion of $C^1(M)$ with respect to the norm

$$\|u\|_{S_1^2(M)} = \left(\int_M (|\nabla_\theta u|_\theta^2 + u^2)\, dV_\theta\right)^{\frac{1}{2}}.$$

(For more properties about the Folland–Stein space, see [Folland and Stein 1974].)

**Proposition 2.1.** *Suppose $\tilde{\theta}$ is a contact form conformal to $\theta$. Then we have*

$$\lambda_k(\tilde{\theta}) = \inf_{V\in \operatorname{Gr}_k^u(S_1^2(M))} \sup_{v\in V\setminus\{0\}} \frac{\int_M vL_\theta v\, dV_\theta}{\int_M u^{\frac{2}{n}}v^2\, dV_\theta}. \tag{2-5}$$

*Proof.* Let $u \in C^\infty(M)$, $u > 0$. For all $f \in C^\infty(M)$, $f \not\equiv 0$, we set $\tilde{\theta} = u^{\frac{2}{n}}\theta$ and

$$F'(u,f) = \frac{\int_M fL_{\tilde{\theta}}f\, dV_{\tilde{\theta}}}{\int_M f^2\, dV_{\tilde{\theta}}}.$$

The operator $L_\theta$ is conformally invariant in the following sense:

$$u^{1+\frac{2}{n}}L_{\tilde{\theta}}(u^{-1}f) = L_\theta(f), \tag{2-6}$$

because

$$\begin{aligned}
u^{1+\frac{2}{n}} L_{\tilde\theta}(u^{-1} f) &= -\Big(2+\frac{2}{n}\Big) u^{1+\frac{2}{n}} \Delta_{\tilde\theta}(u^{-1} f) + R_{\tilde\theta} u^{1+\frac{2}{n}}(u^{-1} f) \\
&= -\Big(2+\frac{2}{n}\Big)\big(u\Delta_\theta(u^{-1} f) + 2\langle \nabla_\theta u, \nabla_\theta(u^{-1} f)\rangle_\theta\big) \\
&\qquad + \Big(-\Big(2+\frac{2}{n}\Big)\Delta_\theta u + R_\theta u\Big)(u^{-1} f) \\
&= -\Big(2+\frac{2}{n}\Big)\Delta_\theta f + R_\theta f = L_\theta(f),
\end{aligned}$$

where we have used (2-1) and (2-2). Combining (2-6) with the fact that

$$dV_{\tilde\theta} = u^{2+\frac{2}{n}}\, dV_\theta, \tag{2-7}$$

we get

$$\begin{aligned}
F'(u, f) &= \frac{\int_M f L_{\tilde\theta} f\, dV_{\tilde\theta}}{\int_M f^2\, dV_{\tilde\theta}} \\
&= \frac{\int_M f u^{-(1+\frac{2}{n})} L_\theta(uf) u^{2+\frac{2}{n}}\, dV_\theta}{\int_M f^2 u^{2+\frac{2}{n}}\, dV_\theta} = \frac{\int_M (uf) L_\theta(uf)\, dV_\theta}{\int_M u^{\frac{2}{n}} (uf)^2\, dV_\theta}.
\end{aligned} \tag{2-8}$$

Using the min-max principle, we have

$$\lambda_k(\tilde\theta) = \inf_{V\in \mathrm{Gr}_k(S_1^2(M))} \sup_{v\in V\setminus\{0\}} \frac{\int_M v L_{\tilde\theta} v\, dV_{\tilde\theta}}{\int_M v^2\, dV_{\tilde\theta}}. \tag{2-9}$$

Since $u > 0$, we have $\mathrm{Gr}_k(S_1^2(M)) = \mathrm{Gr}_k^u(S_1^2(M))$. Therefore, it follows from (2-8) and (2-9) that

$$\lambda_k(\tilde\theta) = \inf_{V\in \mathrm{Gr}_k(S_1^2(M))} \sup_{f\in V\setminus\{0\}} F'(u, f).$$

Now replacing $uf$ by $v$, we obtain (2-5) by (2-8). □

Now we can define the generalized contact form:

**Definition.** The generalized contact form $\tilde\theta$ is defined as $\tilde\theta = u^{\frac{2}{n}}\theta$, where $u$ is no longer necessarily positive or smooth, but $u \in L_+^{2+\frac{2}{n}}(M)$.

We enlarge the conformal class $[\theta]$ of $\theta$ by including all the generalized contact forms conformal to $\theta$, as follows:

$$[\theta] = \{\tilde\theta = u^{\frac{2}{n}}\theta \mid u \in L_+^{2+\frac{2}{n}}(M)\}.$$

In view of Proposition 2.1, for a generalized contact form $\tilde\theta = u^{\frac{2}{n}}\theta$, $u \in L_+^{2+\frac{2}{n}}(M)$, conformal to $\theta$, we define

$$\lambda_k(\tilde\theta) = \inf_{V\in \mathrm{Gr}_k^u(S_1^2(M))} \sup_{v\in V\setminus\{0\}} \frac{\int_M v L_\theta v\, dV_\theta}{\int_M u^{\frac{2}{n}} v^2\, dV_\theta}. \tag{2-10}$$

Using (2-10), we can generalize the definition of $k$-th CR Yamabe invariant to the generalized contact form by using (2-4).

## 3. Variational characterization of $Y_k(M,\theta)$

For all $u \in L_+^{2+\frac{2}{n}}(M)$, $v \in S_1^2(M)$ such that $u^{\frac{1}{n}}v \not\equiv 0$, we set

$$F(u,v) = \frac{\int_M \left(2+\frac{2}{n}\right)|\nabla_\theta v|_\theta^2 + R_\theta v^2\, dV_\theta}{\int_M u^{\frac{2}{n}} v^2\, dV_\theta}\left(\int_M u^{2+\frac{2}{n}}\, dV_\theta\right)^{\frac{1}{n+1}}.$$

**Proposition 3.1.** *If $[\theta]$ contains all the contact forms conformal to $\theta$, then*

$$Y_k(M,\theta) = \inf_{\substack{u\in C^\infty(M)\\ V\in \mathrm{Gr}_k^u(S_1^2(M))}} \sup_{v\in V\setminus\{0\}} F(u,v). \tag{3-1}$$

*Similarly, if $[\theta]$ contains all the generalized contact forms conformal to $\theta$, then*

$$Y_k(M,\theta) = \inf_{\substack{u\in L_+^{2+\frac{2}{n}}(M)\\ V\in \mathrm{Gr}_k^u(S_1^2(M))}} \sup_{v\in V\setminus\{0\}} F(u,v). \tag{3-2}$$

*Proof.* Using the definition of $Y_k(M,\theta)$ and the fact that $\mathrm{Vol}(M,\tilde\theta) = \int_M u^{2+\frac{2}{n}}\, dV_\theta$, we obtain from (2-5) that

$$\begin{aligned} Y_k(M,\theta) &= \inf_{\tilde\theta\in[\theta]} \lambda_k(\tilde\theta)\,\mathrm{Vol}(M,\tilde\theta)^{\frac{1}{n+1}} \\ &= \inf_{u\in C^\infty(M),u>0} \lambda_k(\tilde\theta)\left(\int_M u^{2+\frac{2}{n}}\, dV_\theta\right)^{\frac{1}{n+1}} \\ &= \inf_{\substack{u\in C^\infty(M),u>0\\ V\in \mathrm{Gr}_k^u(S_1^2(M))}} \sup_{v\in V\setminus\{0\}} F(u,v), \end{aligned}$$

which proves (3-1). Similarly, we can prove (3-2) by using the same arguments as above, except we need to replace $C^\infty(M)$ by $L_+^{2+\frac{2}{n}}(M)$. □

## 4. Generalized contact form and the Euler–Lagrange equation

We will need the following:

**Lemma 4.1.** *Let $u \in L^{2+\frac{2}{n}}(M)$ and $v \in S_1^2(M)$. We assume that*

$$L_\theta v = u^{\frac{2}{n}} v \tag{4-1}$$

*holds in the sense of distributions. Then $v \in L^{2+\frac{2}{n}+\varepsilon}(M)$ for some $\varepsilon > 0$.*

*Proof.* Without loss of generality, suppose $v \not\equiv 0$. We define $v_+ = \sup(v, 0)$. We let $q \in (1, (n+1)/n]$ be a fixed number and $l > 0$ be a large real number which will tend to $+\infty$. We let $\beta = 2q - 1$. We then define for $x \in \mathbb{R}$,

$$G_l(x) = \begin{cases} 0 & \text{if } x < 0, \\ x^\beta & \text{if } 0 \le x < l, \\ l^{q-1}(ql^{q-1}x - (q-1)l^q) & \text{if } x \ge l, \end{cases}$$

$$F_l(x) = \begin{cases} 0 & \text{if } x < 0, \\ x^q & \text{if } 0 \le x < l, \\ ql^{q-1}x - (q-1)l^q & \text{if } x \ge l. \end{cases}$$

It is easy to check that for all $x \in \mathbb{R}$,

$$(F_l'(x))^2 \le qG_l'(x), \tag{4-2}$$

$$(F_l(x))^2 \ge xG_l(x), \tag{4-3}$$

$$xG_l'(x) \le \beta G_l(x). \tag{4-4}$$

Since $F_l$ and $G_l$ are uniformly Lipschitz continuous functions, $F_l(v_+)$ and $G_l(v_+)$ belong to $S_1^2(M)$. Let $x_0 \in M$. Denote by $\eta$ a $C^2$ nonnegative function supported in $B(x_0, 2\delta)$, where $\delta > 0$ is a small fixed number such that $0 \le \eta \le 1$ and $\eta(B(x_0, \delta)) = \{1\}$. Multiply (4-1) by $\eta^2 G_l(v_+)$ and integrate over $M$. Since the supports of $v_+$ and $G_l(v_+)$ coincide, we get

$$\begin{aligned}\left(2 + \frac{2}{n}\right) \int_M \langle \nabla_\theta v_+, \nabla_\theta \eta^2 G_l(v_+)\rangle_\theta \, dV_\theta + \int_M R_\theta v_+ \eta^2 G_l(v_+) \, dV_\theta \\ = \int_M u^{\frac{2}{n}} v_+ \eta^2 G_l(v_+) \, dV_\theta.\end{aligned} \tag{4-5}$$

We are going to estimate the terms in (4-5). In the following, $C$ will denote a positive constant depending possibly on $\eta$, $q$, $\beta$, $\delta$, but not on $l$. Note that

$$\begin{aligned}&\int_M \langle \nabla_\theta v_+, \nabla_\theta \eta^2 G_l(v_+)\rangle_\theta \, dV_\theta \\ &= \int_M G_l(v_+)\langle \nabla_\theta v_+, \nabla_\theta \eta^2\rangle_\theta \, dV_\theta + \int_M G_l'(v_+)\eta^2 |\nabla_\theta v_+|_\theta^2 \, dV_\theta \\ &= -\int_M G_l(v_+) v_+ \Delta_\theta(\eta^2) \, dV_\theta - 2\int_M v_+ G_l'(v_+)\eta \langle \nabla_\theta v_+, \nabla_\theta \eta\rangle_\theta \, dV_\theta \\ &\qquad + \int_M G_l'(v_+)\eta^2 |\nabla_\theta v_+|_\theta^2 \, dV_\theta \\ &\ge -C\int_M v_+ G_l(v_+) \, dV_\theta - 2\int_M v_+^2 G_l'(v_+) |\nabla_\theta \eta|_\theta^2 \, dV_\theta \\ &\qquad + \frac{1}{2}\int_M G_l'(v_+)\eta^2 |\nabla_\theta v_+|_\theta^2 \, dV_\theta,\end{aligned} \tag{4-6}$$

where the last inequality follows from $|\langle\nabla_\theta v_+, \nabla_\theta\eta\rangle_\theta| \le |\nabla_\theta\eta|_\theta^2 + \frac{1}{4}|\nabla_\theta v_+|_\theta^2$. Hence, we have

$$
\begin{aligned}
(4\text{-}7)\quad \int_M \langle\nabla_\theta v_+, \nabla_\theta\eta^2 G_l(v_+)\rangle_\theta \, dV_\theta \\
&\ge -C\int_M v_+ G_l(v_+)\, dV_\theta - 2\int_M v_+^2 G_l'(v_+)|\nabla_\theta\eta|_\theta^2\, dV_\theta \\
&\qquad + \frac{1}{2}\int_M G_l'(v_+)\eta^2|\nabla_\theta v_+|_\theta^2\, dV_\theta \\
&\ge -C\int_M v_+ G_l(v_+)\, dV_\theta - 2\beta\int_M v_+ G_l(v_+)|\nabla_\theta\eta|_\theta^2\, dV_\theta \\
&\qquad + \frac{1}{2}\int_M G_l'(v_+)\eta^2|\nabla_\theta v_+|_\theta^2\, dV_\theta \\
&\ge -C\int_M (F_l(v_+))^2\, dV_\theta + \frac{1}{2q}\int_M (F_l'(v_+))^2\eta^2|\nabla_\theta v_+|_\theta^2\, dV_\theta \\
&= -C\int_M (F_l(v_+))^2\, dV_\theta + \frac{1}{2q}\int_M \eta^2|\nabla_\theta F_l(v_+)|_\theta^2\, dV_\theta \\
&\ge -C\int_M (F_l(v_+))^2\, dV_\theta + \frac{1}{4q}\int_M |\nabla_\theta(\eta F_l(v_+))|_\theta^2\, dV_\theta \\
&\qquad - \frac{1}{2q}\int_M |\nabla_\theta\eta|_\theta^2 (F_l(v_+))^2\, dV_\theta \\
&\ge -C\int_M (F_l(v_+))^2\, dV_\theta + \frac{1}{4q}\int_M |\nabla_\theta(\eta F_l(v_+))|_\theta^2\, dV_\theta,
\end{aligned}
$$

where the first inequality follows from (4-6), the second inequality follows from (4-4), the third inequality follows from (4-2) and (4-3), and the fourth inequality follows from

$$
\begin{aligned}
|\nabla_\theta(\eta F_l(v_+))|_\theta^2 &= |F_l(v_+)\nabla_\theta\eta + \eta\nabla_\theta F_l(v_+)|_\theta^2 \\
&\le 2\eta^2|\nabla_\theta F_l(v_+)|_\theta^2 + 2|\nabla_\theta\eta|_\theta^2 (F_l(v_+))^2.
\end{aligned}
$$

By the Folland–Stein embedding from $S_1^2(M)$ into $L^{2+\frac{2}{n}}(M)$, there exists a constant $A > 0$ depending only on $(M, \theta)$ such that

$$
\int_M |\nabla_\theta(\eta F_l(v_+))|_\theta^2\, dV_\theta \ge A\biggl(\int_M (\eta F_l(v_+))^{2+\frac{2}{n}}\, dV_\theta\biggr)^{\frac{n}{n+1}} - \int_M (\eta F_l(v_+))^2\, dV_\theta.
$$

From this, together with (4-7), we obtain

$$
\begin{aligned}
(4\text{-}8)\quad \int_M \langle\nabla_\theta v_+, \nabla_\theta\eta^2 G_l(v_+)\rangle_\theta \, dV_\theta \\
&\ge -C\int_M (F_l(v_+))^2\, dV_\theta + \frac{A}{4q}\biggl(\int_M (\eta F_l(v_+))^{2+\frac{2}{n}}\, dV_\theta\biggr)^{\frac{n}{n+1}}.
\end{aligned}
$$

Independently, we choose $\delta > 0$ small enough such that

$$\int_{B(x_0,2\delta)} u^{2+\frac{2}{n}}\, dV_\theta \le \left(\left(2+\frac{2}{n}\right)\frac{A}{8q}\right)^{n+1}. \tag{4-9}$$

Then it follows from (4-3), (4-9) and Hölder's inequality that

$$\begin{aligned} &\int_M u^{\frac{2}{n}} v_+ \eta^2 G_l(v_+)\, dV_\theta \\ &\quad \le \int_M u^{\frac{2}{n}} \eta^2 (F_l(v_+))^2\, dV_\theta \\ &\quad \le \left(\int_{B(x_0,2\delta)} u^{2+\frac{2}{n}}\, dV_\theta\right)^{\frac{1}{n+1}} \left(\int_M (\eta F_l(v_+))^{2+\frac{2}{n}}\, dV_\theta\right)^{\frac{n}{n+1}} \\ &\quad \le \left(2+\frac{2}{n}\right)\frac{A}{8q}\left(\int_M (\eta F_l(v_+))^{2+\frac{2}{n}}\, dV_\theta\right)^{\frac{n}{n+1}}. \end{aligned} \tag{4-10}$$

On the other hand, it follows from (4-3) that

$$\begin{aligned} \int_M R_\theta v_+ \eta^2 G_l(v_+)\, dV_\theta &\ge -(\max_M |R_\theta|) \int_M v_+ \eta^2 G_l(v_+)\, dV_\theta \\ &\ge -(\max_M |R_\theta|) \int_M \eta^2 (F_l(v_+))^2\, dV_\theta \\ &\ge -C \int_M (F_l(v_+))^2\, dV_\theta. \end{aligned} \tag{4-11}$$

Substituting (4-8), (4-10), (4-11) into (4-5), we obtain

$$\left(2+\frac{2}{n}\right)\frac{A}{8q}\left(\int_M (\eta F_l(v_+))^{2+\frac{2}{n}}\, dV_\theta\right)^{\frac{n}{n+1}} \le C \int_M (F_l(v_+))^2\, dV_\theta.$$

Now, by the Folland–Stein embedding, $v_+ \in L^{2+\frac{2}{n}}(M)$. Since $2q \le 2+\frac{2}{n}$ and $C$ does not depend on $l$, the right-hand side of the inequality is bounded when $l \to \infty$, and we obtain

$$\limsup_{l\to\infty} \int_M (\eta F_l(v_+))^{2+\frac{2}{n}}\, dV_\theta < \infty.$$

This proves that $v_+ \in L^{q(2+\frac{2}{n})}(B(x_0,\delta))$. Since $x_0$ is arbitrary, we get that $v_+ \in L^{q(2+\frac{2}{n})}(M)$. Doing the same with $v_- = \sup(-v, 0)$ instead of $v_+$, we get that $v \in L^{q(2+\frac{2}{n})}(M)$. This proves Lemma 4.1. □

**Proposition 4.2.** *For any generalized contact form $\tilde\theta = u^{\frac{2}{n}}\theta$, $u \in L_+^{2+\frac{2}{n}}(M)$, conformal to $\theta$, there exist two functions $v, w \in S_1^2(M)$ with $v \ge 0$ such that in the* *sense of distributions*

$$L_\theta v = \lambda_1(\tilde{\theta}) u^{\frac{2}{n}} v, \tag{4-12}$$

$$L_\theta w = \lambda_2(\tilde{\theta}) u^{\frac{2}{n}} w. \tag{4-13}$$

*Moreover, we can normalize $v$ and $w$ such that*

$$\int_M u^{\frac{2}{n}} v^2 \, dV_\theta = \int_M u^{\frac{2}{n}} w^2 \, dV_\theta = 1 \quad \textit{and} \quad \int_M u^{\frac{2}{n}} vw \, dV_\theta = 0. \tag{4-14}$$

*Proof.* Let $(v_m)_m$ be a minimizing sequence for $\lambda_1(\tilde{\theta})$, i.e., a sequence $v_m \in S_1^2(M)$ such that

$$\lim_{m\to\infty} \frac{\int_M \left(2+\frac{2}{n}\right)|\nabla_\theta v_m|_\theta^2 + R_\theta v_m^2 \, dV_\theta}{\int_M u^{\frac{2}{n}} v_m^2 \, dV_\theta} = \lambda_1(\tilde{\theta}).$$

It is well known that $(|v_m|)_m$ is also a minimizing sequence. Hence we can assume that $v_m \geq 0$. If we normalize $v_m$ by $\int_M u^{\frac{2}{n}} v_m^2 \, dV_\theta = 1$, then $(v_m)_m$ is bounded in $S_1^2(M)$ and after passing to a subsequence, we may assume that there exists $v \in S_1^2(M)$, $v \geq 0$ such that $v_m \to v$ weakly in $S_1^2(M)$ and strongly in $L^2(M)$ almost everywhere. If $u$ is smooth, then

$$\int_M u^{\frac{2}{n}} v^2 \, dV_\theta = \lim_{m\to\infty} \int_M u^{\frac{2}{n}} v_m^2 \, dV_\theta = 1, \tag{4-15}$$

and by standard arguments, $v$ is nonnegative minimizer of the functional associated to $\lambda_1(\tilde{\theta})$.

We must show that (4-15) still holds if $u \in L_+^{2+\frac{2}{n}}(M)$. Let $A > 0$ be a large real number and set $u_A = \inf(u, A)$. Then

$$\begin{aligned} &\left| \int_M u^{\frac{2}{n}} (v_m^2 - v^2) \, dV_\theta \right| \\ &\qquad \leq \int_M u_A^{\frac{2}{n}} |v_m^2 - v^2| \, dV_\theta + \int_M (u^{\frac{2}{n}} - u_A^{\frac{2}{n}})(|v_m| + |v|)^2 \, dV_\theta \\ &\qquad \leq A^{\frac{2}{n}} \int_M |v_m^2 - v^2| \, dV_\theta \\ &\qquad\quad + \left( \int_M (u^{\frac{2}{n}} - u_A^{\frac{2}{n}})^{n+1} \, dV_\theta \right)^{\frac{1}{n+1}} \left( \int_M (|v_m| + |v|)^{2+\frac{2}{n}} \, dV_\theta \right)^{\frac{n}{n+1}}, \end{aligned} \tag{4-16}$$

where we have used Hölder's inequality in the last inequality. Since

$$|u^{\frac{2}{n}} - u_A^{\frac{2}{n}}|^{n+1} \leq u^{2+\frac{2}{n}} \in L^1(M),$$

by Lebesgue's dominated convergence theorem we have

$$\lim_{A\to\infty} \int_M (u^{\frac{2}{n}} - u_A^{\frac{2}{n}})^{n+1} \, dV_\theta = \int_M \lim_{A\to\infty} (u^{\frac{2}{n}} - u_A^{\frac{2}{n}})^{n+1} \, dV_\theta = 0. \tag{4-17}$$

Since $(v_m)_m$ is bounded in $S_1^2(M)$, it is bounded in $L^{2+\frac{2}{n}}(M)$, and hence there exists $C > 0$ such that

$$\int_M (|v_m| + |v|)^{2+\frac{2}{n}}\, dV_\theta \le C. \tag{4-18}$$

By strong convergence in $L^2(M)$,

$$\lim_{m\to\infty} \int_M |v_m^2 - v^2|\, dV_\theta = 0. \tag{4-19}$$

Combining (4-16)–(4-19), we obtain (4-15). Therefore $v$ is a nonnegative minimizer of the functional associated to $\lambda_1(\tilde{\theta})$. Writing the Euler–Lagrange equation of $v$, we find that $v$ satisfies (4-12).

Now we define

$$\lambda_1'(\tilde{\theta}) = \inf \frac{\int_M \bigl(2+\frac{2}{n}\bigr)|\nabla_\theta w|_\theta^2 + R_\theta w^2\, dV_\theta}{\int_M u^{\frac{2}{n}} |w|^2\, dV_\theta},$$

where the infimum is taken over smooth functions $w$ such that $u^{\frac{1}{n}} w \not\equiv 0$ and such that

$$\int_M u^{\frac{2}{n}} v w\, dV_\theta = 0.$$

With the same method, we find a minimizer $w$ of this problem that satisfies (4-13) with $\lambda_2'(\tilde{\theta})$ instead of $\lambda_2(\tilde{\theta})$. However, it is not difficult to see that $\lambda_2'(\tilde{\theta}) = \lambda_2(\tilde{\theta})$ and Proposition 4.2 easily follows. □

**Lemma 4.3.** *Let $u \in L_+^{2+\frac{2}{n}}(M)$ with $\int_M u^{2+\frac{2}{n}}\, dV_\theta = 1$. Suppose that $w_1, w_2 \in S_1^2(M) \setminus \{0\}$, $w_1, w_2 \ge 0$ satisfy*

$$\int_M \Bigl(\Bigl(2+\frac{2}{n}\Bigr)|\nabla_\theta w_1|_\theta^2 + R_\theta w_1^2\Bigr)\, dV_\theta \le Y_2(M,\theta) \int_M u^{\frac{2}{n}} w_1^2\, dV_\theta, \tag{4-20}$$

$$\int_M \Bigl(\Bigl(2+\frac{2}{n}\Bigr)|\nabla_\theta w_2|_\theta^2 + R_\theta w_2^2\Bigr)\, dV_\theta \le Y_2(M,\theta) \int_M u^{\frac{2}{n}} w_2^2\, dV_\theta, \tag{4-21}$$

*and suppose that $(M \setminus w_1^{-1}(0)) \cap (M \setminus w_2^{-1}(0))$ has measure zero. Then $u$ is a linear combination of $w_1$ and $w_2$, and we have equality in* (4-20) *and* (4-21).

*Proof.* We let $\bar{u} = a w_1 + b w_2$, where $a, b > 0$ are chosen such that

$$\frac{b^{\frac{2}{n}} \int_M u^{\frac{2}{n}} w_1^2\, dV_\theta}{a^{\frac{2}{n}} \int_M u^{\frac{2}{n}} w_2^2\, dV_\theta} = \frac{\int_M w_1^{2+\frac{2}{n}}\, dV_\theta}{\int_M w_2^{2+\frac{2}{n}}\, dV_\theta}, \tag{4-22}$$

$$\int_M \bar{u}^{2+\frac{2}{n}}\, dV_\theta = a^{2+\frac{2}{n}} \int_M w_1^{2+\frac{2}{n}}\, dV_\theta + b^{2+\frac{2}{n}} \int_M w_2^{2+\frac{2}{n}}\, dV_\theta = 1. \tag{4-23}$$

Because of the variational characterization of $Y_2(M,\theta)$ in Proposition 3.1, we have

$$Y_2(M,\theta) \le \sup_{(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}} F(\bar{u},\lambda w_1+\mu w_2). \tag{4-24}$$

By (4-20), (4-21), (4-23), and since $(M\setminus w_1^{-1}(0))\cap(M\setminus w_2^{-1}(0))$ has measure zero, we obtain

$$\begin{aligned}
&F(\bar{u},\lambda w_1+\mu w_2)\\
&\quad=\frac{\int_M\bigl(2+\frac{2}{n}\bigr)|\nabla_\theta(\lambda w_1+\mu w_2)|_\theta^2+R_\theta(\lambda w_1+\mu w_2)^2\,dV_\theta}{\int_M \bar{u}^{\frac{2}{n}}(\lambda w_1+\mu w_2)^2\,dV_\theta}\\
&\quad=\frac{\lambda^2\int_M\bigl(2+\frac{2}{n}\bigr)|\nabla_\theta w_1|_\theta^2+R_\theta w_1^2\,dV_\theta+\mu^2\int_M\bigl(2+\frac{2}{n}\bigr)|\nabla_\theta w_2|_\theta^2+R_\theta w_2^2\,dV_\theta}{\lambda^2\int_M\bar{u}^{\frac{2}{n}}w_1^2\,dV_\theta+\mu^2\int_M\bar{u}^{\frac{2}{n}}w_2^2\,dV_\theta}\\
&\quad\le Y_2(M,\theta)\frac{\lambda^2\int_M u^{\frac{2}{n}}w_1^2\,dV_\theta+\mu^2\int_M u^{\frac{2}{n}}w_2^2\,dV_\theta}{\lambda^2a^{\frac{2}{n}}\int_M w_1^{2+\frac{2}{n}}\,dV_\theta+\mu^2b^{\frac{2}{n}}\int_M w_2^{2+\frac{2}{n}}\,dV_\theta}.
\end{aligned}\tag{4-25}$$

By (4-22), the right-hand side of (4-25) does not depend on $\lambda$ and $\mu$. Hence we can choose $\lambda=a$ and $\mu=b$ on the right-hand side of (4-25) to get

$$\begin{aligned}
&\sup_{(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}} F(\bar{u},\lambda w_1+\mu w_2)\\
&\qquad\le Y_2(M,\theta)\frac{a^2\int_M u^{\frac{2}{n}}w_1^2\,dV_\theta+b^2\int_M u^{\frac{2}{n}}w_2^2\,dV_\theta}{a^{2+\frac{2}{n}}\int_M w_1^{2+\frac{2}{n}}\,dV_\theta+b^{2+\frac{2}{n}}\int_M w_2^{2+\frac{2}{n}}\,dV_\theta}\\
&\qquad= Y_2(M,\theta)\int_M u^{\frac{2}{n}}(a^2w_1^2+b^2w_2^2)\,dV_\theta\\
&\qquad= Y_2(M,\theta)\int_M u^{\frac{2}{n}}\bar{u}^2\,dV_\theta\\
&\qquad\le Y_2(M,\theta)\Bigl(\int_M u^{2+\frac{2}{n}}\,dV_\theta\Bigr)^{\frac{1}{n+1}}\Bigl(\int_M \bar{u}^{2+\frac{2}{n}}\,dV_\theta\Bigr)^{\frac{n}{n+1}}\\
&\qquad= Y_2(M,\theta),
\end{aligned}\tag{4-26}$$

where we have used (4-23) in the first equality, the assumption that $(M\setminus w_1^{-1}(0))\cap(M\setminus w_2^{-1}(0))$ has measure zero in the second equality, Hölder's inequality in the second inequality, and the assumption $\int_M u^{2+\frac{2}{n}}\,dV_\theta=1$ and (4-23) in the last equality.

Combining (4-24) and (4-26), we have

$$\sup_{(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}} F(\bar{u},\lambda w_1+\mu w_2)=Y_2(M,\theta).$$

This implies the equality in Holder's inequality in (4-26), which implies that there exists a constant $c > 0$ such that $u = c\bar{u}$ almost everywhere. Since $\int_M u^{2+\frac{2}{n}}\, dV_\theta = \int_M \bar{u}^{2+\frac{2}{n}}\, dV_\theta = 1$ by (4-23), we have $c = 1$, i.e., $u = \bar{u} = aw_1 + bw_2$. Also, equality in (4-25) implies equality in (4-20) and (4-21). This proves the assertion. □

**Theorem 4.4** (Euler–Lagrange equation). *Assume $Y_2(M,\theta) \neq 0$ and that $Y_2(M,\theta)$ is attained by a generalized contact form $\tilde{\theta} = u^{\frac{2}{n}}\theta$ with $u \in L_+^{2+\frac{2}{n}}(M)$. Let $v$ and $w$ be as in Proposition 4.2. Then $u = |w|$. In particular,*

$$L_\theta w = Y_2(M,\theta)|w|^{\frac{2}{n}} w. \tag{4-27}$$

*Moreover, $w$ has alternating sign and $w \in C^{2,\alpha}(M)$ for all $\alpha \in \left[0, \frac{2}{n}\right]$.*

*Proof.* Without loss of generality, we can assume that $\int_M u^{2+\frac{2}{n}}\, dV_\theta = 1$. By assumption and by Proposition 3.1, we have $\lambda_2(\tilde{\theta}) = Y_2(M,\theta)$. Let $v, w \in S_1^2(M)$ be the functions satisfying (4-12), (4-13), and (4-14).

**Step 1.** We have $\lambda_1(\tilde{\theta}) < \lambda_2(\tilde{\theta})$.

We prove this by contradiction. Suppose that $\lambda_1(\tilde{\theta}) = \lambda_2(\tilde{\theta})$. After possibly replacing $w$ by a linear combination of $v$ and $w$, we can assume that the function $u^{\frac{1}{n}} w$ changes sign. If we define $w_1 = \sup(w, 0)$ and $w_2 = \sup(-w, 0)$, then they satisfy the assumption of Lemma 4.3 since $w$ satisfies (4-13) and $\lambda_2(\tilde{\theta}) = Y_2(M,\theta)$. Applying Lemma 4.3, we find $a, b > 0$ such that $u = aw_1 + bw_2$. Now, by Lemma 4.1, $w \in L^{2+\frac{2}{n}+\varepsilon}(M)$. By a standard bootstrap argument, (4-13) shows that $w \in C^{2,\alpha}(M)$ for all $\alpha \in (0,1)$. Since $u = aw_1 + bw_2 = a\sup(w,0) + b\sup(-w,0)$, we have $u \in C^{0,\alpha}(M)$ for all $\alpha \in (0,1)$.

Since $\lambda_1(\tilde{\theta}) = \lambda_2(\tilde{\theta})$ and by the definition of $\lambda_1(\tilde{\theta})$, $w$ is a minimizer of the functional $\bar{w} \mapsto F(u, \bar{w})$ among the functions in $S_1^2(M)$ with $u^{\frac{1}{n}}\bar{w} \not\equiv 0$ by Proposition 3.1. Since $F(u, w) = F(u, |w|)$, we have that $|w|$ is a minimizer for the functional associated to $\lambda_1(\tilde{\theta})$, and $|w|$ satisfies same equation as $w$. As a consequence, $|w|$ is $C^2$. By the maximum principle, we have $|w| > 0$ everywhere, which is false since $u^{\frac{1}{n}} w$ changes sign.

**Step 2.** The function $w$ changes sign.

Assume $w$ does not change sign. Then after possibly replacing $w$ by $-w$, we can assume that $w \geq 0$. Setting $w_1 = v$ and $w_2 = w$, we have (4-20) and (4-21). Using (4-14), we can conclude that $(M \setminus w_1^{-1}(0)) \cap (M \setminus w_2^{-1}(0))$ has measure zero. Applying Lemma 4.3, we have equality in (4-20). On the other hand, Step 1 implies that inequality (4-20) is strict since $\lambda_1(\tilde{\theta}) < \lambda_2(\tilde{\theta}) = Y_2(M,\theta)$. This contradiction shows that $w$ changes sign.

**Step 3.** There exist $a, b > 0$ such that $u = a\sup(w,0) + b\sup(-w,0)$. Moreover, $w \in C^{2,\alpha}(M)$ and $u \in C^{0,\alpha}(M)$ for all $\alpha \in (0,1)$.

As in the proof of Step 1, we apply Lemma 4.3 with $w_1 = \sup(w, 0)$ and $w_2 = \sup(-w, 0)$. We get $a, b > 0$ such that $u = aw_1 + bw_2$. As in Step 1, we get $w \in C^{2,\alpha}(M)$ and $u \in C^{0,\alpha}(M)$ for all $\alpha \in (0, 1)$.

**Step 4.** Conclusion.

Let $h \in C^\infty(M)$ such that $\operatorname{supp}(h) \subseteq M \setminus u^{-1}(0)$. For $t$ close to 0, set $u_t = |u + th|$. Since $u > 0$ on the support of $h$, and since $u$ is continuous, we have for $t$ close to 0, $u_t = u + th$. As $\operatorname{span}(v, w) \in \mathrm{Gr}_2^u(S_1^2(M))$, by Proposition 3.1 we have

$$Y_2(M, \theta) \leq \sup_{(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}} F(u_t, \lambda v + \mu w).$$

Note that

$$\begin{aligned} F(u_t, \lambda v + \mu w) &= \frac{\int_M \left(2 + \frac{2}{n}\right)|\nabla_\theta(\lambda v + \mu w)|_\theta^2 + R_\theta(\lambda v + \mu w)^2 \, dV_\theta}{\int_M u_t^{\frac{2}{n}}(\lambda v + \mu w)^2 \, dV_\theta} \left(\int_M u_t^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \\ &= \frac{\lambda^2\lambda_1(\tilde\theta)\int_M u^{\frac{2}{n}} v^2 \, dV_\theta + \mu^2\lambda_2(\tilde\theta)\int_M u^{\frac{2}{n}} w^2 \, dV_\theta}{\lambda^2 a_t + \lambda\mu b_t + \mu^2 c_t} \left(\int_M u_t^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \\ &= \frac{\lambda^2\lambda_1(\tilde\theta) + \mu^2\lambda_2(\tilde\theta)}{\lambda^2 a_t + \lambda\mu b_t + \mu^2 c_t} \left(\int_M u_t^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}}, \end{aligned} \tag{4-28}$$

where we have used (4-12), (4-13), and (4-14). Here

$$a_t = \int_M u_t^{\frac{2}{n}} v^2 \, dV_\theta, \quad b_t = 2\int_M u_t^{\frac{2}{n}} vw \, dV_\theta \quad \text{and} \quad c_t = \int_M u_t^{\frac{2}{n}} w^2 \, dV_\theta.$$

Note also that the functions $a_t$, $b_t$, and $c_t$ are smooth for $t$ close to 0. Furthermore, $a_0 = c_0 = 1$ and $b_0 = 0$ by (4-14). Define $f(t, \alpha) = F(u_t, \sin(\alpha)v + \cos(\alpha)w)$, which is smooth for small $t$. By (4-28), we have

$$\begin{aligned} f(t, \alpha) &= F(u_t, \sin(\alpha)v + \cos(\alpha)w) \\ &= \frac{\sin^2(\alpha)\lambda_1(\tilde\theta) + \cos^2(\alpha)\lambda_2(\tilde\theta)}{\sin^2(\alpha)a_t + \sin(\alpha)\cos(\alpha)b_t + \cos^2(\alpha)c_t} \left(\int_M u_t^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}}. \end{aligned} \tag{4-29}$$

Hence, using $\lambda_1(\tilde\theta) < \lambda_2(\tilde\theta)$, we can see that $f\left(0, \left(n + \frac{1}{2}\right)\pi\right)$ is minimum and $f(0, n\pi)$ is maximum for any integer $n$. This implies that

$$\frac{\partial}{\partial\alpha} f(0, \alpha) = 0 \text{ if and only if } \alpha \in \frac{\pi}{2}\mathbb{Z},$$

$$\frac{\partial^2}{\partial\alpha^2} f(0, \alpha) < 0 \text{ if } \alpha \in \pi\mathbb{Z} \quad \text{and} \quad \frac{\partial^2}{\partial\alpha^2} f(0, \alpha) > 0 \text{ if } \alpha \in \pi\mathbb{Z} + \frac{\pi}{2}.$$

Applying the implicit function theorem to $\partial f/\partial\alpha$ at the point $(0, 0)$, we see that there exists a smooth function $t \mapsto \alpha(t)$, defined on a neighborhood of 0 with

$\alpha(0)=0$ such that

$$f(t,\alpha(t))=\sup_{\alpha\in\mathbb{R}} f(t,\alpha)=\sup_{(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}} F(u_t,\lambda v+\mu w),$$

where the last equality follows from the fact that

$$F(u_t,c\lambda v+c\mu w)=F(u_t,\lambda v+\mu w)$$

for any nonzero constant $c$ by (4-28). Since $\alpha(0)=0$, we have

$$\begin{aligned}\frac{d}{dt}\sin^2\alpha(t)\Big|_{t=0}=\frac{d}{dt}\cos^2\alpha(t)\Big|_{t=0}&=\frac{d}{dt}(a_t\sin^2\alpha(t))\Big|_{t=0}\\&=\frac{d}{dt}(b_t\sin\alpha(t)\cos\alpha(t))\Big|_{t=0}=0.\end{aligned}$$

Hence, by (4-29), we have

(4-30)
$$\begin{aligned}&\frac{d}{dt}f(t,\alpha(t))\Big|_{t=0}\\&=\frac{d}{dt}\Bigg(\frac{\sin^2(\alpha(t))\lambda_1(\tilde{\theta})+\cos^2(\alpha(t))\lambda_2(\tilde{\theta})}{\sin^2(\alpha(t))a_t+\sin(\alpha(t))\cos(\alpha(t))b_t+\cos^2(\alpha(t))c_t}\\&\qquad\qquad\times\Big(\int_M u_t^{2+\frac{2}{n}}\,dV_\theta\Big)^{\frac{1}{n+1}}\Bigg)\Bigg|_{t=0}\\&=\lambda_2(\tilde{\theta})\Bigg(\Big(-\frac{d}{dt}c_t\Big|_{t=0}\Big)\Big(\int_M u^{2+\frac{2}{n}}\,dV_\theta\Big)^{\frac{1}{n+1}}+\frac{d}{dt}\Big(\int_M u_t^{2+\frac{2}{n}}\,dV_\theta\Big)^{\frac{1}{n+1}}\Big|_{t=0}\Bigg)\\&=\lambda_2(\tilde{\theta})\frac{2}{n}\Big(-\int_M u^{-1+\frac{2}{n}}hw^2\,dV_\theta+\int_M u^{1+\frac{2}{n}}h\,dV_\theta\Big).\end{aligned}$$

By the definition of $Y_2(M,\theta)$ and $\lambda_2(\tilde{\theta})=Y_2(M,\theta)$, $f$ admits a minimum at $t=0$ because

$$f(0,\alpha(0))=f(0,0)=F(u,w)$$

and $w$ satisfies (4-13). Since $\lambda_2(\tilde{\theta})=Y_2(M,\theta)\neq 0$, it follows from (4-30) that

$$\int_M u^{-1+\frac{2}{n}}hw^2\,dV_\theta=\int_M u^{1+\frac{2}{n}}h\,dV_\theta.$$

Since $h$ is arbitrary (we just have to ensure that its support is contained in $M\setminus u^{-1}(0)$), we get

$$u^{-1+\frac{2}{n}}w^2=u^{1+\frac{2}{n}}$$

and hence $u=|w|$ on $M\setminus u^{-1}(0)$. Together with Step 3, we have $u=|w|$ everywhere. □

## 5. Lower bound for $Y_2(M,\theta)$

For any compact CR manifold $(M,\theta)$ of the real dimension $2n+1$, by the definition of the CR Yamabe invariant $Y_1(M,\theta)$, we have

$$Y_1(M,\theta)=\inf_{u\in S_1^2(M)\setminus\{0\}}\frac{\int_M\big(2+\frac{2}{n}\big)|\nabla_\theta u|_\theta^2+R_\theta u^2\,dV_\theta}{\big(\int_M|u|^{2+\frac{2}{n}}\,dV_\theta\big)^{\frac{n}{n+1}}}. \tag{5-1}$$

**Theorem 5.1.** *We have*

$$Y_2(M,\theta)\geq 2^{\frac{1}{n+1}}Y_1(M,\theta). \tag{5-2}$$

*Moreover, if $M$ is connected and if $Y_2(M,\theta)$ is attained by a generalized contact form, then this inequality is strict.*

*Proof.* The functional

$$F(u,v)=\frac{\int_M\big(2+\frac{2}{n}\big)|\nabla_\theta v|_\theta^2+R_\theta v^2\,dV_\theta}{\int_M u^{\frac{2}{n}}v^2\,dV_\theta}\bigg(\int_M u^{2+\frac{2}{n}}\,dV_\theta\bigg)^{\frac{1}{n+1}}$$

is continuous on $L_+^{2+\frac{2}{n}}(M)\times(S_1^2(M)\setminus\{0\})$. As a consequence, $I(u,V):=\sup_{v\in V\setminus\{0\}}F(u,v)$ depends continuously on $u\in L_+^{2+\frac{2}{n}}(M)$ and $V\in\mathrm{Gr}_2^u(S_1^2(M))$. To prove Theorem 5.1, it suffices to show that $I(u,V)\geq 2^{\frac{1}{n+1}}Y_1(M,\theta)$ for all smooth $u>0$ and $V\in\mathrm{Gr}_2^u(S_1^2(M))$ thanks to Proposition 3.1. Without loss of generality, we can assume that

$$\int_M u^{2+\frac{2}{n}}\,dV_\theta=1. \tag{5-3}$$

The operator

$$v\mapsto P(v):=-\Big(2+\frac{2}{n}\Big)u^{-\frac{1}{n}}\Delta_\theta(u^{-\frac{1}{n}}v)+R_\theta u^{-\frac{2}{n}}v$$

is self-adjoint with respect to the $L^2$-scalar product and elliptic. Hence, $P$ has discrete spectrum $\lambda_1\leq\lambda_2\leq\cdots$ and the corresponding eigenfunctions $\varphi_1,\varphi_2,\dots$ are smooth. Setting $v_i=u^{-\frac{1}{n}}\varphi_i$, we obtain

$$\begin{aligned}\Big(-\Big(2+\frac{2}{n}\Big)\Delta_\theta+R_\theta\Big)(v_i)&=-\Big(2+\frac{2}{n}\Big)\Delta_\theta(u^{-\frac{1}{n}}\varphi_i)+R_\theta u^{-\frac{1}{n}}\varphi_i\\&=u^{\frac{1}{n}}P(\varphi_i)=\lambda_i u^{\frac{1}{n}}\varphi_i=\lambda_i u^{\frac{2}{n}}v_i\end{aligned} \tag{5-4}$$

and

$$\int_M u^{\frac{2}{n}}v_iv_j\,dV_\theta=\int_M\varphi_i\varphi_j\,dV_\theta=0 \text{ if } i\neq j.$$

The maximum principle implies that an eigenfunction to the smallest eigenvalue $\lambda_1$ has no zeros. Hence, $\lambda_1<\lambda_2$ and we can assume that $v_1>0$.

We define $w_+ = a_+ \sup(v_2, 0)$ and $w_- = a_- \sup(-v_2, 0)$, where $a_+, a_- > 0$ are chosen such that

$$\int_M u^{\frac{2}{n}} w_+^2 \, dV_\theta = \int_M u^{\frac{2}{n}} w_-^2 \, dV_\theta = 1. \tag{5-5}$$

We let $\Omega_- = \{v_2 < 0\}$ and $\Omega_+ = \{v_2 \geq 0\}$. By Hölder's inequality, we have

$$\begin{aligned} 2 &= \int_M u^{\frac{2}{n}} w_+^2 \, dV_\theta + \int_M u^{\frac{2}{n}} w_-^2 \, dV_\theta \\ &\leq \left(\int_{\Omega_+} u^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \left(\int_M w_+^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{n}{n+1}} \\ &\qquad + \left(\int_{\Omega_-} u^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \left(\int_M w_-^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{n}{n+1}}. \end{aligned} \tag{5-6}$$

Using the inequality (5-1), we get

$$\int_M u^{\frac{1}{n}} w_+ P(u^{\frac{1}{n}} w_+) \, dV_\theta \geq Y_1(M, \theta) \left(\int_M w_+^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{n}{n+1}},$$

which implies that

$$\begin{aligned} &\left(\int_{\Omega_+} u^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \left(\int_M u^{\frac{1}{n}} w_+ P(u^{\frac{1}{n}} w_+) \, dV_\theta\right) \\ &\qquad \geq Y_1(M, \theta) \left(\int_M w_+^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{n}{n+1}} \left(\int_{\Omega_+} u^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \\ &\qquad \geq Y_1(M, \theta) \int_M u^{\frac{2}{n}} w_+^2 \, dV_\theta = Y_1(M, \theta), \end{aligned} \tag{5-7}$$

where we have used Hölder's inequality in the last inequality, and (5-5) in the last equality. Similarly, we have

$$\left(\int_{\Omega_-} u^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \left(\int_M u^{\frac{1}{n}} w_- P(u^{\frac{1}{n}} w_-) \, dV_\theta\right) \geq Y_1(M, \theta). \tag{5-8}$$

Adding (5-7) and (5-8) together, we obtain

$$\begin{aligned} 2Y_1(M, \theta) \leq &\left(\int_{\Omega_+} u^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \left(\int_M u^{\frac{1}{n}} w_+ P(u^{\frac{1}{n}} w_+) \, dV_\theta\right) \\ &+ \left(\int_{\Omega_-} u^{2+\frac{2}{n}} \, dV_\theta\right)^{\frac{1}{n+1}} \left(\int_M u^{\frac{1}{n}} w_- P(u^{\frac{1}{n}} w_-) \, dV_\theta\right). \end{aligned} \tag{5-9}$$

Since $w_-$, respectively $w_+$, are multiples of $v_2$ on $\Omega_-$, respectively $\Omega_+$, they satisfy the same equation as $v_2$. Hence, we obtain from (5-4) and (5-9) that

$$\begin{aligned}(5\text{-}10)\quad 2Y_1(M,\theta) &\le \left(\int_{\Omega_+} u^{2+\frac{2}{n}}\,dV_\theta\right)^{\frac{1}{n+1}}\left(\int_M \lambda_2 u^{\frac{2}{n}} w_+^2\,dV_\theta\right)\\ &\qquad + \left(\int_{\Omega_-} u^{2+\frac{2}{n}}\,dV_\theta\right)^{\frac{1}{n+1}}\left(\int_M \lambda_2 u^{\frac{2}{n}} w_-^2\,dV_\theta\right)\\ &= \lambda_2\left(\left(\int_{\Omega_+} u^{2+\frac{2}{n}}\,dV_\theta\right)^{\frac{1}{n+1}} + \left(\int_{\Omega_-} u^{2+\frac{2}{n}}\,dV_\theta\right)^{\frac{1}{n+1}}\right),\end{aligned}$$

where the last equality follows from (5-5). Now, for any nonnegative numbers $a, b \ge 0$, Hölder's inequality yields

$$a+b \le 2^{\frac{n}{n+1}}(a^{n+1}+b^{n+1})^{\frac{1}{n+1}}.$$

Applying this inequality with

$$a = \left(\int_{\Omega_+} u^{2+\frac{2}{n}}\,dV_\theta\right)^{\frac{1}{n+1}} \quad \text{and} \quad b = \left(\int_{\Omega_-} u^{2+\frac{2}{n}}\,dV_\theta\right)^{\frac{1}{n+1}},$$

we derive from (5-10) that

$$\begin{aligned}2Y_1(M,\theta) &\le \lambda_2 2^{\frac{n}{n+1}}\left(\left(\int_{\Omega_+} u^{2+\frac{2}{n}}\,dV_\theta\right) + \left(\int_{\Omega_-} u^{2+\frac{2}{n}}\,dV_\theta\right)\right)^{\frac{1}{n+1}}\\ &= \lambda_2 2^{\frac{n}{n+1}}\left(\int_M u^{2+\frac{2}{n}}\,dV_\theta\right)^{\frac{1}{n+1}} = \lambda_2 2^{\frac{n}{n+1}},\end{aligned}$$

where the last equality follows from (5-3). This implies that $\lambda_2 \ge 2^{\frac{1}{n+1}} Y_1(M,\theta)$. Since $\lambda_2 = I(u, \operatorname{span}(v_1, v_2))$, this finishes the proof of the first part of Theorem 5.1.

Moreover, if $M$ were connected and if $Y_2(M,\theta)$ were attained by a generalized contact form, then inequality (5-9) would be an equality and we would have that $w_+$ or $w_-$ is a function for which equality in (5-1) is attained. By the maximum principle, we would get that $w_+$ or $w_-$ is positive on $M$, which is impossible. □

## 6. Upper bound for $Y_2(M,\theta)$

Hereafter, we denote $Y_k(\mathbb{S}^{2n+1})$ the $k$-th Yamabe invariant of $(\mathbb{S}^{2n+1}, \theta_{\mathbb{S}^{2n+1}})$, where $\theta_{\mathbb{S}^{2n+1}}$ is the standard contact form on $\mathbb{S}^{2n+1}$ given by

$$\theta_{\mathbb{S}^{2n+1}} = \sqrt{-1}\sum_{j=1}^{n+1}(z_j\,d\bar{z}_j - \bar{z}_j\,dz_j),$$

where $(z_1,\dots,z_{n+1}) \in \mathbb{S}^{2n+1} \subset \mathbb{C}^{n+1}$.

**Theorem 6.1.** *Suppose $(M,\theta)$ is a compact CR manifold of real dimension $2n+1$ with $Y_1(M,\theta)\geq 0$. Then*

$$Y_2(M,\theta)\leq \big(Y_1(M,\theta)^{n+1}+Y_1(\mathbb{S}^{2n+1})^{n+1}\big)^{\frac{1}{n+1}} \tag{6-1}$$

*when $Y_1(M,\theta)>0$ and $n\geq 3$, or $Y_1(M,\theta)=0$ and $n\geq 4$. On the other hand, the inequality in* (6-1) *is strict when*

(i) $Y_1(M,\theta)>0$, $n\geq 7$ *and $M$ is not locally CR equivalent to $\mathbb{S}^{2n+1}$, or*

(ii) $Y_1(M,\theta)=0$, $n\geq 4$ *and $M$ is not locally CR equivalent to $\mathbb{S}^{2n+1}$.*

To prove Theorem 5.4, we have the following:

**Lemma 6.2.** *For any $\alpha>2$, there exists a constant $C>0$ such that*

$$|a+b|^\alpha\leq a^\alpha+b^\alpha+C(a^{\alpha-1}b+ab^{\alpha-1})$$

*for all $a,b>0$.*

*Proof.* Dividing both sides by $a$, without loss of generality, we can assume that $a=1$. Then we set for $x>0$,

$$f(x)=\frac{|1+x|^\alpha-(1+x^\alpha)}{x^{\alpha-1}+x}.$$

By L'Hôpital's rule, we have

$$\lim_{x\to 0^+}f(x)=\lim_{x\to 0^+}\frac{\alpha(1+x)^{\alpha-1}-\alpha x^{\alpha-1}}{(\alpha-1)x^{\alpha-2}+1}=\alpha,$$
$$\lim_{x\to\infty}f(x)=\lim_{x\to\infty}\frac{\alpha(1+x)^{\alpha-1}-\alpha x^{\alpha-1}}{(\alpha-1)x^{\alpha-2}+1}=\alpha.$$

Since $f$ is continuous, $f$ is bounded by a constant $C$ on $(0,\infty)$. Clearly, this constant is the desired $C$ is the inequality of Lemma 6.2. □

*Proof of Theorem 6.1.* For $u\in S_1^2(M)\setminus\{0\}$, let

$$E(u)=\frac{\int_M\big(2+\frac{2}{n}\big)|\nabla_\theta u|_\theta^2+R_\theta u^2\,dV_\theta}{\big(\int_M|u|^{2+\frac{2}{n}}\,dV_\theta\big)^{\frac{n}{n+1}}}.$$

The solution of the CR Yamabe problem provides the existence of a smooth positive minimizer $v$ of $E$, and we can assume

$$\int_M v^{2+\frac{2}{n}}\,dV_\theta=1. \tag{6-2}$$

Then $v$ satisfies the CR Yamabe equation

$$L_\theta(v)=Y_1(M,\theta)v^{1+\frac{2}{n}}. \tag{6-3}$$

Let $x_0 \in M$ be fixed and choose pseudohermitian normal coordinates $(z, t)$ near $x_0$. Let $\delta > 0$ be a fixed number. If $\theta$ is well chosen in the conformal class and if $x_0$ is well chosen in $M$, it was proved by Jerison and Lee [1989, Theorem 4.1] that when $n \geq 3$, there exists a function $v_\varepsilon \geq 0$ with $\text{supp}(v_\varepsilon) \subseteq B(x_0, 2\delta)$ such that

$$E(v_\varepsilon) = Y_1(\mathbb{S}^{2n+1}) - c(M)\varepsilon^4 + O(\varepsilon^5), \tag{6-4}$$

where $c(M) \geq 0$ is a positive constant. In fact, $c(M)$ is the square of the norm of the Chern tensor at $x_0$ up to a dimensional constant. Therefore, we can assume that the constant $c(M)$ in (6-4) satisfies

$$c(M) > 0 \tag{6-5}$$

if $(M, \theta)$ is not locally CR equivalent to $\mathbb{S}^{2n+1}$. It follows from (6-4) that

$$\lim_{\varepsilon \to 0} E(v_\varepsilon) = Y_1(\mathbb{S}^{2n+1}). \tag{6-6}$$

More precisely, $v_\varepsilon$ is given by (see [Jerison and Lee 1989, p. 326])

$$v_\varepsilon = C_\varepsilon \eta \left( \frac{\varepsilon^2}{t^2 + (|z|^2 + \varepsilon^2)^2} \right)^{\frac{n}{2}},$$

where $\eta$ is a smooth cut-off function such that

$$0 \leq \eta \leq 1, \quad \eta(x) = \begin{cases} 1 & \text{if } x \in B(x_0, \delta), \\ 0 & \text{if } x \notin B(x_0, 2\delta), \end{cases}$$

and $C_\varepsilon > 0$ is a constant chosen such that

$$\int_M v_\varepsilon^{2+\frac{2}{n}} \, dV_\theta = 1. \tag{6-7}$$

It follows from [Jerison and Lee 1989, Proposition 4.2] that

$$C_\varepsilon = c(n) + O(\varepsilon^4) \tag{6-8}$$

for some positive constant $c(n)$ depending only on $n$. In the following, $C$ will denote a positive constant depending possibly on $\delta$, $n$, but not on $\varepsilon$. Let

$$\delta_\varepsilon(z, t) = (\varepsilon z, \varepsilon^2 t).$$

Note that

$$\delta_\varepsilon^* \left( \frac{1}{t^2 + (\varepsilon^2 + |z|^2)^2} \right) = \varepsilon^{-4} \left( \frac{1}{t^2 + (1 + |z|^2)^2} \right)$$

and $\delta_\varepsilon^* \, dz\, dt = \varepsilon^{2n+2}\, dz\, dt$. Hence,

$$\begin{aligned}
\text{(6-9)}\quad \int_M |v_\varepsilon|^p \, dV_\theta &\le C_\varepsilon^p \int_{\left\{\sqrt[4]{t^2+|z|^4}\le 2\delta\right\}} \frac{\varepsilon^{np}\, dz\, dt}{(t^2+(\varepsilon^2+|z|^2)^2)^{\frac{np}{2}}}\\
&= C_\varepsilon^p \int_{\left\{\sqrt[4]{t^2+|z|^4}\le 2\delta/\varepsilon\right\}} \frac{\varepsilon^{2n+2-np}\, dz\, dt}{(t^2+(1+|z|^2)^2)^{\frac{np}{2}}}\\
&\le C_\varepsilon^p \varepsilon^{2n+2-np} \int_{\{|z|\le 2\delta/\varepsilon\}} \left(\int_{-\infty}^{\infty} \frac{dt}{1+t^2}\right) \frac{dz}{(1+|z|^2)^{np-2}}\\
&= C_\varepsilon^p \pi \varepsilon^{2n+2-np} \int_{\{|z|\le 2\delta/\varepsilon\}} \frac{dz}{(1+|z|^2)^{np-2}}\\
&= C\varepsilon^{2n+2-np} \int_0^{2\delta/\varepsilon} \frac{r^{2n-1}\, dr}{(1+r^2)^{np-2}},
\end{aligned}$$

where we have used (6-8). Note that for $\varepsilon \ll 1$,

$$\int_0^{2\delta/\varepsilon} \frac{r^{2n-1}\, dr}{(1+r^2)^{np-2}} \le \int_0^{2\delta/\varepsilon} r^{2n+3-2np}\, dr \le \frac{C}{\varepsilon^{2n+4-2np}}$$

if $p \le 1+\frac{3}{2n}$, and

$$\begin{aligned}
\int_0^{2\delta/\varepsilon} \frac{r^{2n-1}\, dr}{(1+r^2)^{np-2}} &\le \int_0^1 r^{2n-1}\, dr + \int_1^{2\delta/\varepsilon} \frac{dr}{r^{2np-2n-3}}\\
&= \int_0^1 r^{2n-1}\, dr + \int_1^{2\delta/\varepsilon} \frac{dr}{r} = \frac{1}{2n} + \log\varepsilon
\end{aligned}$$

if $p = 1+\frac{2}{n}$. Combining these with (6-9), we obtain

$$\text{(6-10)}\qquad \int_M |v_\varepsilon|^p \, dV_\theta \le \begin{cases} C\varepsilon^{np-2} & \text{if } p \le 1+\frac{3}{2n},\\ C\varepsilon^n \log\varepsilon & \text{if } p = 1+\frac{2}{n}.\end{cases}$$

Similarly, for $\varepsilon \ll 1$, we have

$$\begin{aligned}
\text{(6-11)}\quad \int_M |v_\varepsilon|^p \, dV_\theta &\ge C_\varepsilon^p \int_{\left\{\sqrt[4]{t^2+|z|^4}\le \delta\right\}} \frac{\varepsilon^{np}\, dz\, dt}{(t^2+(\varepsilon^2+|z|^2)^2)^{\frac{np}{2}}}\\
&= C_\varepsilon^p \int_{\left\{\sqrt[4]{t^2+|z|^4}\le \delta/\varepsilon\right\}} \frac{\varepsilon^{2n+2-np}\, dz\, dt}{(t^2+(1+|z|^2)^2)^{\frac{np}{2}}}\\
&\ge C_\varepsilon^p \varepsilon^{2n+2-np} \int_{\{|z|\le \delta/2\varepsilon\}} \left(\int_{-\delta/2\varepsilon}^{\delta/2\varepsilon} \frac{dt}{1+t^2}\right) \frac{dz}{(1+|z|^2)^{np}}\\
&\ge 2C_\varepsilon^p \tan^{-1}(\delta/2)\varepsilon^{2n+2-np} \int_{\{|z|\le \delta/2\varepsilon\}} \frac{dz}{(1+|z|^2)^{np}}\\
&= C\varepsilon^{2n+2-np} \int_0^{\delta/2\varepsilon} \frac{r^{2n-1}\, dr}{(1+r^2)^{np}},
\end{aligned}$$

where we have used

$$t^2+(1+|z|^2)^2\le(1+t^2)(1+|z|^2)^2$$

and

$$\{|z|\le\delta/2\varepsilon\}\cap\{|t|\le\delta/2\varepsilon\}\subset\left\{\sqrt[4]{t^2+|z|^4}\le\delta/\varepsilon\right\}$$

in the second inequality, and (6-8) in the last equality. Note that for $\varepsilon\ll 1$,

$$\int_0^{\delta/2\varepsilon}\frac{r^{2n-1}\,dr}{(1+r^2)^{np}}\ge\int_0^1\frac{r^{2n-1}\,dr}{2^{np}}+\int_1^{\delta/2\varepsilon}\frac{r^{2n-1}\,dr}{(2r^2)^{np}}=C+\frac{C}{\varepsilon^{2n-2np}}$$

if $\le 1-\frac{1}{2n}$, and

$$\begin{aligned}\int_0^{\delta/2\varepsilon}\frac{r^{2n-1}\,dr}{(1+r^2)^{np}}&\ge\int_0^1\frac{r^{2n-1}\,dr}{2^{np}}+\int_1^{\delta/2\varepsilon}\frac{r^{2n-1}\,dr}{(2r^2)^{np}}\\&\ge\frac{1}{2^{np}}\left(\int_0^1 r^{2n-1}\,dr+\int_1^{\delta/2\varepsilon}\frac{dr}{r^{2np-2n+1}}\right)=C+C\varepsilon^{2np-2n}\end{aligned}$$

if $p>1$. Combining these with (6-11), we obtain

$$\int_M|v_\varepsilon|^p\,dV_\theta\ge\begin{cases}C\varepsilon^{np+2} & \text{if } p\le 1-\frac{1}{2n},\\ C\varepsilon^{2n+2-np} & \text{if } p>1.\end{cases}\tag{6-12}$$

First we assume that $Y_1(M,\theta)>0$. We set

$$u_\varepsilon=E(v_\varepsilon)^{\frac{n}{2}}v_\varepsilon+Y_1(M,\theta)^{\frac{n}{2}}v.$$

Let us find estimates for $F(u_\varepsilon,\lambda v_\varepsilon+\mu v)$. Let $(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}$. Then

$$\begin{aligned}&F(u_\varepsilon,\lambda v_\varepsilon+\mu v)\\&=\frac{\lambda^2\int_M v_\varepsilon L_\theta v_\varepsilon\,dV_\theta+\mu^2\int_M vL_\theta v\,dV_\theta+2\lambda\mu\int_M v_\varepsilon L_\theta v\,dV_\theta}{\int_M|u_\varepsilon|^{\frac{2}{n}}(\lambda v_\varepsilon+\mu v)^2\,dV_\theta}\cdot U\\&=\frac{\lambda^2E(v_\varepsilon)+\mu^2Y_1(M,\theta)+2\lambda\mu Y_1(M,\theta)\int_M v^{1+\frac{2}{n}}v_\varepsilon\,dV_\theta}{\lambda^2\int_M|u_\varepsilon|^{\frac{2}{n}}v_\varepsilon^2\,dV_\theta+\mu^2\int_M|u_\varepsilon|^{\frac{2}{n}}v^2\,dV_\theta+2\lambda\mu\int_M|u_\varepsilon|^{\frac{2}{n}}v_\varepsilon v\,dV_\theta}\cdot U,\end{aligned}\tag{6-13}$$

where $U=\left(\int_M u_\varepsilon^{2+2/n}\,dV_\theta\right)^{1/(n+1)}$ and where we have used (6-2), (6-3) and (6-7). Using the definition of $u_\varepsilon$, we have

$$u_\varepsilon\ge E(v_\varepsilon)^{\frac{n}{2}}v_\varepsilon\quad\text{and}\quad u_\varepsilon\ge Y_1(M,\theta)^{\frac{n}{2}}v,\tag{6-14}$$

which implies that

$$\begin{aligned}\text{(6-15)}\quad \lambda^2\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon^2\,dV_\theta+\mu^2\int_M |u_\varepsilon|^{\frac{2}{n}} v^2\,dV_\theta+2\lambda\mu\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon v\,dV_\theta \\ \geq \lambda^2 E(v_\varepsilon)\int_M v_\varepsilon^{2+\frac{2}{n}}\,dV_\theta+\mu^2 Y_1(M,\theta)\int_M v^{2+\frac{2}{n}}\,dV_\theta \\ +2\lambda\mu\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon v\,dV_\theta \\ =\lambda^2 E(v_\varepsilon)+\mu^2 Y_1(M,\theta)+2\lambda\mu\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon v\,dV_\theta,\end{aligned}$$

where the last equality follows from (6-2) and (6-7).

If $\lambda\mu\geq 0$, then we have

$$\text{(6-16)}\qquad 2\lambda\mu\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon v\,dV_\theta \geq 2\lambda\mu Y_1(M,\theta)\int_M v^{1+\frac{2}{n}} v_\varepsilon\,dV_\theta$$

by (6-14). Therefore, (6-15) and (6-16) imply that

$$\frac{\lambda^2 E(v_\varepsilon)+\mu^2 Y_1(M,\theta)+2\lambda\mu Y_1(M,\theta)\int_M v^{1+\frac{2}{n}} v_\varepsilon\,dV_\theta}{\lambda^2\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon^2\,dV_\theta+\mu^2\int_M |u_\varepsilon|^{\frac{2}{n}} v^2\,dV_\theta+2\lambda\mu\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon v\,dV_\theta}\leq 1.$$

If $\lambda\mu<0$, then

$$|u_\varepsilon|^{\frac{2}{n}}\leq \big(E(v_\varepsilon)^{\frac{n}{2}} v_\varepsilon+Y_1(M,\theta)^{\frac{n}{2}} v\big)^{\frac{2}{n}}\leq E(v_\varepsilon) v_\varepsilon^{\frac{2}{n}}+Y_1(M,\theta) v^{\frac{2}{n}}$$

when $n\geq 2$. Combining this with (6-14) and (6-15), we get

$$\begin{aligned}\lambda^2\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon^2\,dV_\theta+\mu^2\int_M |u_\varepsilon|^{\frac{2}{n}} v^2\,dV_\theta+2\lambda\mu\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon v\,dV_\theta \\ \geq \lambda^2 E(v_\varepsilon)+\mu^2 Y_1(M,\theta)-C\bigg(\int_M v_\varepsilon^{1+\frac{2}{n}} v\,dV_\theta+\int_M v^{1+\frac{2}{n}} v_\varepsilon\,dV_\theta\bigg) \\ \geq \lambda^2 E(v_\varepsilon)+\mu^2 Y_1(M,\theta)-C\bigg(\int_M v_\varepsilon^{1+\frac{2}{n}}\,dV_\theta+\int_M v_\varepsilon\,dV_\theta\bigg),\end{aligned}$$

where $C>0$ is a positive real number independent of $\varepsilon$. This, together with (6-10), gives

$$\begin{aligned}\lambda^2\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon^2\,dV_\theta+\mu^2\int_M |u_\varepsilon|^{\frac{2}{n}} v^2\,dV_\theta+2\lambda\mu\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon v\,dV_\theta \\ \geq \lambda^2 E(v_\varepsilon)+\mu^2 Y_1(M,\theta)-O(\varepsilon^n\log\varepsilon)-O(\varepsilon^{n-2}).\end{aligned}$$

This, together with the assumption that $\lambda\mu<0$, implies that

$$\frac{\lambda^2 E(v_\varepsilon)+\mu^2 Y_1(M,\theta)+2\lambda\mu Y_1(M,\theta)\int_M v^{1+\frac{2}{n}} v_\varepsilon\,dV_\theta}{\lambda^2\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon^2\,dV_\theta+\mu^2\int_M |u_\varepsilon|^{\frac{2}{n}} v^2\,dV_\theta+2\lambda\mu\int_M |u_\varepsilon|^{\frac{2}{n}} v_\varepsilon v\,dV_\theta}\leq 1+O(\varepsilon^{n-2}).$$

In any case, we have

(6-17)
$$\sup_{(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}}\frac{\lambda^2E(v_\varepsilon)+\mu^2Y_1(M,\theta)+2\lambda\mu Y_1(M,\theta)\int_M v^{1+\frac{2}{n}}v_\varepsilon\,dV_\theta}{\lambda^2\int_M|u_\varepsilon|^{\frac{2}{n}}v_\varepsilon^2\,dV_\theta+\mu^2\int_M|u_\varepsilon|^{\frac{2}{n}}v^2\,dV_\theta+2\lambda\mu\int_M|u_\varepsilon|^{\frac{2}{n}}v_\varepsilon v\,dV_\theta}\leq 1+O(\varepsilon^{n-2}).$$

On the other hand,

$$\begin{aligned}\int_M u_\varepsilon^{2+\frac{2}{n}}\,dV_\theta&=\int_M(E(v_\varepsilon)^{\frac{n}{2}}v_\varepsilon+Y_1(M,\theta)^{\frac{n}{2}}v)^{2+\frac{2}{n}}\,dV_\theta\\&\leq E(v_\varepsilon)^{n+1}\int_M v_\varepsilon^{2+\frac{2}{n}}\,dV_\theta+Y_1(M,\theta)^{n+1}\int_M v^{2+\frac{2}{n}}\,dV_\theta\\&\qquad+C\left(\int_M v_\varepsilon^{1+\frac{2}{n}}v\,dV_\theta+\int_M v^{1+\frac{2}{n}}v_\varepsilon\,dV_\theta\right)\\&=E(v_\varepsilon)^{n+1}+Y_1(M,\theta)^{n+1}+C\left(\int_M v_\varepsilon^{1+\frac{2}{n}}v\,dV_\theta+\int_M v^{1+\frac{2}{n}}v_\varepsilon\,dV_\theta\right),\end{aligned}$$

where the first inequality follows from Lemma 6.2 with

$$a=E(v_\varepsilon)^{\frac{n}{2}}v_\varepsilon\quad\text{and}\quad b=Y_1(M,\theta)^{\frac{n}{2}}v,$$

and the last equality follows from (6-2) and (6-7). This, together with (6-4) and (6-10), implies that

(6-18)
$$\left(\int_M u_\varepsilon^{2+\frac{2}{n}}\,dV_\theta\right)^{\frac{1}{n+1}}\leq(Y_1(\mathbb{S}^{2n+1})^{n+1}+Y_1(M,\theta)^{n+1})^{\frac{1}{n+1}}-c(M)\varepsilon^4+o(\varepsilon^4)+O(\varepsilon^{n-2}).$$

If $\varepsilon>0$ is small enough, it follows from (6-13), (6-17), and (6-18) that

(6-19)
$$\begin{aligned}Y_2(M,\theta)&\leq\sup_{(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}}F(u_\varepsilon,\lambda v_\varepsilon+\mu v)\\&\leq\big(Y_1(\mathbb{S}^{2n+1})^{n+1}+Y_1(M,\theta)^{n+1}\big)^{\frac{1}{n+1}}-c(M)\varepsilon^4+o(\varepsilon^4)+O(\varepsilon^{n-2}).\end{aligned}$$

Since $n\geq 3$, (6-1) follows from (6-19) by letting $\varepsilon$ go to zero. On the other hand, if $(M,\theta)$ is not locally CR equivalent to $\mathbb{S}^{2n+1}$, then (6-5) holds. Hence, if $n\geq 7$, the strict inequality in (6-1) follows from (6-19) by letting $\varepsilon$ go to zero.

Now we assume that $Y_1(M,\theta)=0$. We set $u_\varepsilon=v_\varepsilon$. Then we obtain for $(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}$,

$$\text{(6-20)}\quad F(u_\varepsilon,\lambda v_\varepsilon+\mu v)$$
$$=\frac{\lambda^2E(v_\varepsilon)\big(\int_M v_\varepsilon^{2+\frac{2}{n}}\,dV_\theta\big)^{\frac{1}{n+1}}}{\lambda^2\int_M|v_\varepsilon|^{\frac{2}{n}}v_\varepsilon^2\,dV_\theta+\mu^2\int_M v_\varepsilon^{\frac{2}{n}}v^2\,dV_\theta+2\lambda\mu\int_M|v_\varepsilon|^{\frac{2}{n}}v_\varepsilon v\,dV_\theta}$$
$$=\frac{\lambda^2E(v_\varepsilon)}{\lambda^2+\mu^2\int_M v_\varepsilon^{\frac{2}{n}}v^2\,dV_\theta+2\lambda\mu\int_M v_\varepsilon^{1+\frac{2}{n}}v\,dV_\theta}$$

by (6-7) and (6-13). Let $\lambda_\varepsilon$, $\mu_\varepsilon$ such that $\lambda_\varepsilon^2+\mu_\varepsilon^2=1$ and

$$F(u_\varepsilon,\lambda_\varepsilon v_\varepsilon+\mu_\varepsilon v)=\sup_{(\lambda,\mu)\in\mathbb{R}^2\setminus\{(0,0)\}}F(u_\varepsilon,\lambda v_\varepsilon+\mu v).$$

If $\lambda_\varepsilon=0$, we obtain that $F(u_\varepsilon,\lambda_\varepsilon v_\varepsilon+\mu_\varepsilon v)=0$ and the theorem would be proved. Then we assume that $\lambda_\varepsilon\neq 0$ and we can write

$$F(u_\varepsilon,\lambda_\varepsilon v_\varepsilon+\mu_\varepsilon v)=\frac{E(v_\varepsilon)}{1+2x_\varepsilon b_\varepsilon+x_\varepsilon^2a_\varepsilon},$$

where $x_\varepsilon=\mu_\varepsilon/\lambda_\varepsilon$ and

$$C\varepsilon^n\le b_\varepsilon=\int_M v_\varepsilon^{1+\frac{2}{n}}v\,dV_\theta\le C\varepsilon^{n-1}\log\varepsilon\quad\text{as }\varepsilon\to 0,$$
$$a_\varepsilon=\int_M v_\varepsilon^{\frac{2}{n}}v^2\,dV_\theta\ge C\varepsilon^4\quad\text{as }\varepsilon\to 0$$

by (6-10) and (6-12). Maximizing this expression in $x_\varepsilon$ and using (6-4), we obtain

$$\text{(6-21)}\quad F(u_\varepsilon,\lambda_\varepsilon v_\varepsilon+\mu_\varepsilon v)\le\frac{Y_1(\mathbb{S}^{2n+1})-c(M)\varepsilon^4+o(\varepsilon^4)}{1-b_\varepsilon^2/a_\varepsilon}=\frac{Y_1(\mathbb{S}^{2n+1})-c(M)\varepsilon^4+o(\varepsilon^4)}{1-C\varepsilon^{2n-6}\log^2\varepsilon},$$

since $\varepsilon\log\varepsilon\to 0$ as $\varepsilon\to 0$. For $n\ge 4$, it follows from (6-21) that

$$F(u_\varepsilon,\lambda_\varepsilon v_\varepsilon+\mu_\varepsilon v)\le Y_1(\mathbb{S}^{2n+1}),$$

which proves (6-1) for the case $Y_1(M,\theta)=0$. On the other hand, if $(M,\theta)$ is not locally CR equivalent to $\mathbb{S}^{2n+1}$, then (6-5) holds. Hence, the strictly inequality in (6-1) follows from (6-21) by letting $\varepsilon$ go to zero. This proves Theorem 6.1. □

## 7. Some properties of $Y_2(M,\theta)$

We have the following questions:

(1) Is $Y_2(M,\theta)$ attained by a contact form?

(2) Is $Y_2(M,\theta)$ attained by a generalized contact form?

For question 1, we have the following:

**Proposition 7.1.** *Let $\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1}$ be the disjoint union of two copies of the sphere equipped with the standard contact form induced from $\theta_{\mathbb{S}^{2n+1}}$. Then $Y_2(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1})=2^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1})$ and it is attained by the standard contact form.*

*Proof.* Let $\tilde\theta$ be an arbitrary smooth contact form on $\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1}$. We write $\mathbb{S}_1^{2n+1}$ for the first $\mathbb{S}^{2n+1}$ and $\mathbb{S}_2^{2n+1}$ for the second $\mathbb{S}^{2n+1}$. Then we have

$$\begin{aligned}\text{(7-1)}\quad &\lambda_2(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1},\tilde\theta)\\ &=\min\Big\{\lambda_2(\mathbb{S}_1^{2n+1},\tilde\theta),\lambda_2(\mathbb{S}_2^{2n+1},\tilde\theta),\max\{\lambda_1(\mathbb{S}_1^{2n+1},\tilde\theta),\lambda_1(\mathbb{S}_2^{2n+1},\tilde\theta)\}\Big\}.\end{aligned}$$

Therefore,

$$\begin{aligned}\text{(7-2)}\quad Y_2(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1})&\le\lambda_2(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1})\operatorname{Vol}(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1})^{\frac{1}{n+1}}\\ &=\lambda_2(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1})(2\operatorname{Vol}(\mathbb{S}^{2n+1}))^{\frac{1}{n+1}}\\ &=2^{\frac{1}{n+1}}\lambda_1(\mathbb{S}^{2n+1})\operatorname{Vol}(\mathbb{S}^{2n+1})^{\frac{1}{n+1}}\\ &=2^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1}),\end{aligned}$$

where we have used (7-1) in the second equality.

On the other hand, we have

$$\begin{aligned}\text{(7-3)}\quad \lambda_2(\mathbb{S}_1^{2n+1},\tilde\theta)\operatorname{Vol}(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1},\tilde\theta)^{\frac{1}{n+1}}&\ge\lambda_2(\mathbb{S}_1^{2n+1},\tilde\theta)\operatorname{Vol}(\mathbb{S}_1^{2n+1},\tilde\theta)^{\frac{1}{n+1}}\\ &\ge Y_2(\mathbb{S}_1^{2n+1})\\ &=2^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1}),\end{aligned}$$

where the last equality follows from Corollary 7.3. Similarly, we have

$$\text{(7-4)}\quad \lambda_2(\mathbb{S}_2^{2n+1},\tilde\theta)\operatorname{Vol}(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1},\tilde\theta)^{\frac{1}{n+1}}\ge 2^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1}).$$

By the definition of $Y_1(\mathbb{S}^{2n+1})$, we have

$$\lambda_1(\mathbb{S}_i^{2n+1},\tilde\theta)\operatorname{Vol}(\mathbb{S}^{2n+1},\tilde\theta)^{\frac{1}{n+1}}\ge Y_1(\mathbb{S}^{2n+1})\quad\text{for } i=1,2,$$

which implies

$$\begin{aligned}&2Y_1(\mathbb{S}^{2n+1})^{n+1}\\ &\quad\le\sum_{i=1}^{2}\lambda_1(\mathbb{S}_i^{2n+1},\tilde\theta)^{n+1}\operatorname{Vol}(\mathbb{S}_i^{2n+1},\tilde\theta)\\ &\quad\le\max\{\lambda_1(\mathbb{S}_1^{2n+1},\tilde\theta)^{n+1},\lambda_1(\mathbb{S}_2^{2n+1},\tilde\theta)^{n+1}\}\sum_{i=1}^{2}\operatorname{Vol}(\mathbb{S}_i^{2n+1},\tilde\theta)\\ &\quad=\max\{\lambda_1(\mathbb{S}_1^{2n+1},\tilde\theta)^{n+1},\lambda_1(\mathbb{S}_2^{2n+1},\tilde\theta)^{n+1}\}\operatorname{Vol}(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1},\tilde\theta),\end{aligned}$$

which gives

$$2^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1}) \le \max\{\lambda_1(\mathbb{S}_1^{2n+1},\tilde{\theta}),\lambda_1(\mathbb{S}_2^{2n+1},\tilde{\theta})\}\operatorname{Vol}(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1},\tilde{\theta})^{\frac{1}{n+1}}. \tag{7-5}$$

Combining (7-3), (7-4), and (7-5), we can derive from (7-1) that

$$2^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1}) \le \lambda_2(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1},\tilde{\theta})\operatorname{Vol}(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1},\tilde{\theta})^{\frac{1}{n+1}}.$$

Since $\tilde{\theta}$ is an arbitrary smooth contact form on $\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1}$, we have

$$2^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1}) \le Y_2(\mathbb{S}^{2n+1}\cup\mathbb{S}^{2n+1}). \tag{7-6}$$

Now Proposition 7.1 follows from combining (7-2) and (7-6). □

On the other hand, we have the following:

**Proposition 7.2.** *If $M$ is connected, then $Y_2(M,\theta)$ cannot be attained by a contact form.*

*Proof.* Otherwise, if $Y_2(M,\theta)$ were attained by a contact form $\tilde{\theta}=u^{\frac{2}{n}}\theta$, then by Theorem 4.4, we would have $u=|w|$, and hence $u$ cannot be positive since $w$ has alternating sign. □

For question 2, we have the following:

**Corollary 7.3.** *We have*

$$Y_2(\mathbb{S}^{2n+1}) = 2^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1}).$$

*Proof.* This follows from (6-1) and Theorem 5.1. □

**Corollary 7.4.** *$Y_2(\mathbb{S}^{2n+1})$ is not attained by a generalized contact form.*

*Proof.* This follows from Theorem 5.1 and Corollary 7.3. □

## 8. The $k$-th CR Yamabe invariant $Y_k(M,\theta)$

In view of Corollary 7.3, it is natural to conjecture that

$$Y_k(\mathbb{S}^{2n+1}) = k^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1})$$

for all $k$. However, the following result shows that it is false.

**Proposition 8.1.** *For $n \ge 3$, we have*

$$Y_{2n+3}(\mathbb{S}^{2n+1}) < (2n+3)^{\frac{1}{n+1}}Y_1(\mathbb{S}^{2n+1}).$$

*Proof.* Consider $\mathbb{S}^{2n+1}\subseteq\mathbb{C}^{n+1}$. Let $z_i$, where $i=1,2,\dots,n+1$, be the coordinates of $\mathbb{C}^{n+1}$. Since $-\Delta_{\theta_{\mathbb{S}^{2n+1}}}z_i=\frac{n}{2}z_i$ and $-\Delta_{\theta_{\mathbb{S}^{2n+1}}}\bar{z}_i=\frac{n}{2}\bar{z}_i$,

$$L_{\theta_{\mathbb{S}^{2n+1}}}(z_i)=\frac{(n+2)(n+1)}{2}z_i \quad\text{and}\quad L_{\theta_{\mathbb{S}^{2n+1}}}(\bar{z}_i)=\frac{(n+2)(n+1)}{2}\bar{z}_i$$

for $i = 1, 2, \dots, n+1$, and hence

$$\lambda_{2n+3}(\mathbb{S}^{2n+1}, \theta_{\mathbb{S}^{2n+1}}) \le \frac{(n+2)(n+1)}{2}.$$

This shows by the definition of $Y_{2n+3}$ that

$$\begin{aligned} Y_{2n+3}(\mathbb{S}^{2n+1}) &\le \lambda_{2n+3}(\mathbb{S}^{2n+1}, \theta_{\mathbb{S}^{2n+1}}) \operatorname{Vol}(\mathbb{S}^{2n+1}, \theta_{\mathbb{S}^{2n+1}})^{\frac{1}{n+1}} \\ &\le \frac{(n+2)(n+1)}{2} \operatorname{Vol}(\mathbb{S}^{2n+1}, \theta_{\mathbb{S}^{2n+1}})^{\frac{1}{n+1}}. \end{aligned} \tag{8-1}$$

Since

$$\begin{aligned} &\frac{(n+2)(n+1)}{2} \operatorname{Vol}(\mathbb{S}^{2n+1}, \theta_{\mathbb{S}^{2n+1}})^{\frac{1}{n+1}} \\ &\qquad < (2n+3)^{\frac{1}{n+1}} \frac{n(n+1)}{2} \operatorname{Vol}(\mathbb{S}^{2n+1}, \theta_{\mathbb{S}^{2n+1}})^{\frac{1}{n+1}} \\ &\qquad = (2n+3)^{\frac{1}{n+1}} Y_1(\mathbb{S}^{2n+1}) \end{aligned}$$

when $n \ge 3$, Proposition 8.1 follows from (8-1) . ☐

For the case when the $k$-th CR Yamabe invariant is negative, we have this:

**Theorem 8.2.** *Let $k$ be an positive integer. Assume that $Y_k(M, \theta) < 0$. Then $Y_k(M, \theta) = -\infty$.*

*Proof.* After a possible change of contact form in the conformal class, we can assume that $\lambda_k(\theta) < 0$. This implies that we can find smooth functions $v_1, \dots, v_k$ satisfying

$$L_\theta(v_i) = \lambda_i(\theta) v_i \quad \text{for all } i = 1, 2, \dots, k$$

and such that

$$\int_M v_i v_j \, dV_\theta = 0 \quad \text{for all } i, j = 1, 2, \dots, k \text{ and } i \ne j.$$

Let $v_k$ be defined as in the proof of Theorem 6.1. We define $u_\varepsilon = v_\varepsilon + \varepsilon$. We set $V = \operatorname{span}\{v_1, \dots, v_k\}$. For $v \in V$, we have

$$\begin{aligned} \int_M u_\varepsilon^{\frac{2}{n}} v^2 \, dV_\theta &\le \varepsilon^{\frac{2}{n}} \int_M v^2 \, dV_\theta + \int_M v_\varepsilon^{\frac{2}{n}} v^2 \, dV_\theta \\ &\le C\varepsilon^{\frac{2}{n}} + C \int_M v_\varepsilon^{\frac{2}{n}} \, dV_\theta \\ &\le \begin{cases} C\varepsilon^{\frac{2}{n}} + C\left(\displaystyle\int_M v_\varepsilon^{\frac{3}{n}} \, dV_\theta\right)^{\frac{2}{3}} \operatorname{Vol}(M, \theta)^{\frac{1}{3}} = C\varepsilon^{\frac{2}{n}} + C\varepsilon^{\frac{2}{3}} & \text{if } n \ge 2, \\ C\varepsilon^2 + C\left(\displaystyle\int_M v_\varepsilon^{\frac{5}{2}} \, dV_\theta\right)^{\frac{1}{5}} \operatorname{Vol}(M, \theta)^{\frac{4}{5}} = C\varepsilon^2 + C\varepsilon^{\frac{1}{10}} & \text{if } n = 1 \end{cases} \end{aligned}$$

by (6-10) and Hölder's inequality. From this, we have

$$\lim_{\varepsilon\to 0}\int_M u_\varepsilon^{\frac{2}{n}} v^2\, dV_\theta = 0$$

uniformly in $v \in V$. Since $\lambda_k(\theta) < 0$, it is then easy to see that

$$\sup_{v\in V} F(u_\varepsilon, v) = -\infty.$$

Together with the variational characterization of $Y_k(M,\theta)$ in Proposition 3.1, we get that $Y_k(M,\theta) = -\infty$. $\square$

## References

[Ammann and Humbert 2006] B. Ammann and E. Humbert, "The second Yamabe invariant", *J. Funct. Anal.* **235**:2 (2006), 377–412. MR 2007a:53073 Zbl 1142.53026

[Aubin 1976] T. Aubin, "Équations différentielles non linéaires et problème de Yamabe concernant la courbure scalaire", *J. Math. Pures Appl.* (9) **55**:3 (1976), 269–296. MR 55 #4288 Zbl 0336.53033

[Brendle 2005] S. Brendle, "Convergence of the Yamabe flow for arbitrary initial energy", *J. Differential Geom.* **69**:2 (2005), 217–278. MR 2006e:53119 Zbl 1085.53028

[Brendle 2007a] S. Brendle, "Convergence of the Yamabe flow in dimension 6 and higher", *Invent. Math.* **170**:3 (2007), 541–576. MR 2008k:53136 Zbl 1130.53044

[Brendle 2007b] S. Brendle, "A short proof for the convergence of the Yamabe flow on $S^n$", *Pure Appl. Math. Q.* **3**:2 (2007), 499–512. MR 2008i:53088 Zbl 1152.53051

[Chang and Cheng 2002] S.-C. Chang and J.-H. Cheng, "The Harnack estimate for the Yamabe flow on CR manifolds of dimension 3", *Ann. Global Anal. Geom.* **21**:2 (2002), 111–121. MR 2002m:53055 Zbl 1007.53034

[Chang et al. 2010] S.-C. Chang, H.-L. Chiu, and C.-T. Wu, "The Li–Yau–Hamilton inequality for Yamabe flow on a closed CR 3-manifold", *Trans. Amer. Math. Soc.* **362**:4 (2010), 1681–1698. MR 2011m:53118 Zbl 1192.32020

[Cheng et al. 2013] J. H. Cheng, A. Malchiodi, and P. Yang, "A positive mass theorem in three dimensional Cauchy–Riemann geometry", preprint, 2013. arXiv 1312.7764

[Cheng et al. 2014] J.-H. Cheng, H.-L. Chiu, and P. Yang, "Uniformization of spherical CR manifolds", *Adv. Math.* **255** (2014), 182–216. MR 3167481 Zbl 1288.32051

[Chow 1992] B. Chow, "The Yamabe flow on locally conformally flat manifolds with positive Ricci curvature", *Comm. Pure Appl. Math.* **45**:8 (1992), 1003–1014. MR 93d:53045 Zbl 0785.53027

[Folland and Stein 1974] G. B. Folland and E. M. Stein, "Estimates for the $\bar\partial_b$ complex and analysis on the Heisenberg group", *Comm. Pure Appl. Math.* **27** (1974), 429–522. MR 51 #3719 Zbl 0293.35012

[Gamara 2001] N. Gamara, "The CR Yamabe conjecture—the case $n = 1$", *J. Eur. Math. Soc.* (*JEMS*) **3**:2 (2001), 105–137. MR 2003d:32040a Zbl 0988.53013

[Gamara and Yacoub 2001] N. Gamara and R. Yacoub, "CR Yamabe conjecture—the conformally flat case", *Pacific J. Math.* **201**:1 (2001), 121–175. MR 2003d:32040b Zbl 1054.32020

[Ho 2012] P. T. Ho, "The long-time existence and convergence of the CR Yamabe flow", *Commun. Contemp. Math.* **14**:2 (2012), 1250014. MR 2901057 Zbl 1246.53087

[Jerison and Lee 1987] D. Jerison and J. M. Lee, "The Yamabe problem on CR manifolds", *J. Differential Geom.* **25**:2 (1987), 167–197. MR 88i:58162 Zbl 0661.32026

[Jerison and Lee 1988] D. Jerison and J. M. Lee, "Extremals for the Sobolev inequality on the Heisenberg group and the CR Yamabe problem", *J. Amer. Math. Soc.* **1**:1 (1988), 1–13. MR 89b:53063 Zbl 0634.32016

[Jerison and Lee 1989] D. Jerison and J. M. Lee, "Intrinsic CR normal coordinates and the CR Yamabe problem", *J. Differential Geom.* **29**:2 (1989), 303–343. MR 90h:58083 Zbl 0671.32016

[Schoen 1984] R. Schoen, "Conformal deformation of a Riemannian metric to constant scalar curvature", *J. Differential Geom.* **20**:2 (1984), 479–495. MR 86i:58137 Zbl 0576.53028

[Schwetlick and Struwe 2003] H. Schwetlick and M. Struwe, "Convergence of the Yamabe flow for "large" energies", *J. Reine Angew. Math.* **562** (2003), 59–100. MR 2004h:53097 Zbl 1079.53100

[Trudinger 1968] N. S. Trudinger, "Remarks concerning the conformal deformation of Riemannian structures on compact manifolds", *Ann. Scuola Norm. Sup. Pisa* (3) **22** (1968), 265–274. MR 39 #2093 Zbl 0159.23801

[Yamabe 1960] H. Yamabe, "On a deformation of Riemannian structures on compact manifolds", *Osaka Math. J.* **12** (1960), 21–37. MR 23 #A2847 Zbl 0096.37201

[Ye 1994] R. Ye, "Global existence and convergence of Yamabe flow", *J. Differential Geom.* **39**:1 (1994), 35–50. MR 95d:53044 Zbl 0846.53027

[Zhang 2009] Y. Zhang, "The contact Yamabe flow on $K$-contact manifolds", *Sci. China Ser. A* **52**:8 (2009), 1723–1732. MR 2010g:53126 Zbl 1189.53074

Received June 3, 2014. Revised May 18, 2015.

PAK TUNG HO
DEPARTMENT OF MATHEMATICS
SOGANG UNIVERSITY
SEOUL 121-742
SOUTH KOREA

ptho@sogang.ac.kr

paktungho@yahoo.com.hk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.401

# NO HYPERBOLIC PANTS FOR THE 4-BODY PROBLEM WITH STRONG POTENTIAL

CONNOR JACKMAN AND RICHARD MONTGOMERY

**The $N$-body problem with a $1/r^2$ potential has, in addition to translation and rotational symmetry, an effective scale symmetry which allows its zero energy flow to be reduced to a geodesic flow on complex projective $(N-2)$-space, minus a hyperplane arrangement. When $N = 3$ we get a geodesic flow on the 2-sphere minus three points. If, in addition we assume that the three masses are equal, then it was proved in a previous paper that the corresponding metric is hyperbolic: its Gaussian curvature is negative except at two points. Does the negative curvature property persist for $N = 4$, that is, in the equal mass $1/r^2$ potential 4-body problem? Here we prove that it does not by computing that the corresponding Riemannian metric in this $N = 4$ case has positive sectional curvature at some 2-planes. This curvature computation underlines an essential difference between the 3- and 4-body problem, a difference whose consequences remain to be explored.**

## 1. Introduction

In [Montgomery 2005] it was shown that the reduced Jacobi–Maupertuis metric for a certain 3-body problem had negative Gaussian curvature (except at two points where it is zero). This hyperbolicity led to deep dynamical consequences. Does hyperbolicity, i.e., curvature negativity, persist for the analogous $N$-body problem with $N > 3$? No. We show that the analogous reduced 4-body problem with its metric has 2-planes at which the sectional curvature is positive.

The $N$-body problem in question has equal masses and the inverse *cube* law attractive force between bodies.

## 2. Setup

Identify the complex numbers $\mathbb{C}$ with the Euclidean plane $\mathbb{R}^2$. Then the planar $N$-body problem has configuration space $\mathbb{C}^N \setminus \Delta$. Here $\Delta$ is the “fat diagonal”

*MSC2010:* 37N05, 70F10, 70G45.
*Keywords:* $N$-body problems, dynamical systems, differential geometry, hyperbolic flows.

consisting of all collisions:

$$\Delta = \{q = (q_1, q_2, \dots, q_N) \in \mathbb{C}^N : q_i = q_j \text{ for some pair } i \neq j\}.$$

The quotient of $\mathbb{C}^N \setminus \Delta$ by translations and rotations is the "reduced $N$-body configuration space":

$$C_N = Y_N \times \mathbb{R}^+, \qquad Y_N = \mathbb{CP}^{N-2} \setminus \mathbb{P}\Delta,$$

where $\mathbb{CP}^{N-2}$ is the projectivization of the center of mass subspace

$$\mathbb{C}^{N-1} = \left\{q \in \mathbb{C}^N : \sum m_i q_i = 0\right\}$$

and $\mathbb{P}\Delta \subset \mathbb{CP}^{N-2}$ is the projectivization of $\Delta \cap \mathbb{C}^{N-1}$. The $\mathbb{R}^+$ factor records the overall scale of the planar $N$-gon and is coordinatized by $\sqrt{I}$ with $I = \Sigma m_i |q_i|^2$ being the total moment of inertia about the center of mass. $Y_N$ is the moduli space of oriented similarity classes of noncollision $N$-gons and will be called "shape space".

The following considerations reduce the zero angular momentum, zero energy $N$-body problem to a geodesic flow on shape space $Y_N$, *provided* the potential $V$ is homogeneous of degree $-2$. If $V$ is homogeneous of degree $-\alpha$ then the virial identity, also known as the Lagrange–Jacobi identity, asserts that along solutions of energy $H$ we have $\ddot{I} = 4H - (4-2\alpha)V$, which implies that the only case in which we can generally guarantee that $\ddot{I} = 0$ is when $\alpha = 2$ and $H = 0$. If in addition $\dot{I} = 0$ then solutions lie on constant levels of $I$.

Now we recall the Jacobi–Maupertuis (JM) reformulation of mechanics, which asserts that the solutions to Newton's equations at energy $H$ are, after a time reparametrization, precisely the geodesic equations for the *Jacobi–Maupertuis metric*

$$ds^2_{\mathrm{JM}} = 2(H - V)\, ds^2$$

on the *Hill region* $\{H - V \geq 0\} \subset \mathbb{C}^N \setminus \Delta$ with $ds^2$ the mass metric. We are interested in the case $H = 0$, $-V > 0$ with $V$ homogeneous of degree $-2$, in which case the Hill region is all of $\mathbb{C}^N \setminus \Delta$ and

$$ds^2_{\mathrm{JM}} = U\, ds^2, \qquad U = -V.$$

The case of prime interest to us is

$$U = -V = \sum_{i \neq j} m_i m_j / r_{ij}^2. \tag{1}$$

This $U$, and hence the JM metric, is invariant under rotations and translations. Quotienting first by translations, we take representatives in the totally geodesic center-of-mass-zero subspace $\mathbb{C}^{N-1}$, which reduces the dynamics to geodesics of the metric $ds^2_{\mathrm{JM}}|_{\mathbb{C}^{N-1}}$ on $\mathbb{C}^{N-1}$. Moreover, $ds^2_{\mathrm{JM}}|_{\mathbb{C}^{N-1}}$ is also invariant under scaling since the homogeneities of $U$ and the Euclidean mass metric $ds^2$ on $\mathbb{C}^{N-1}$ cancel. Thus the JM metric admits the group $G = \mathbb{C}^*$ of rotations and scalings as an isometry group.

Now $Y_N$ is the quotient space: $Y_N = (\mathbb{C}^{N-1} \setminus \Delta)/G = \mathbb{CP}^{N-2} \setminus \Delta$. (By abuse of notation, we continue to use the symbol $\Delta$ to denote the image of the collision locus $\Delta$ under projectivization and intersection.) Insisting that the quotient map $\pi : \mathbb{C}^{N-1} \setminus \Delta \to Y_N$ is a Riemannian submersion induces a metric on $Y_N$. Recall that this means that we can define the metric on $Y_N$ by *isometrically* identifying the tangent space to $Y_N$ at a point $p$ with the orthogonal complement (relative to $ds^2_{\mathrm{JM}}$ or $ds^2$, and at any point lying over $p$ in $\mathbb{C}^{N-1}$) to the $G$-orbit that corresponds to that point. These orthogonality conditions are equivalent to the conditions that the linear momentum, angular momentum, and "scale momentum" $\dot{I}$ are all zero. To summarize, by using the JM metric and forming the Riemannian quotient, the zero angular momentum, zero energy $1/r^2$ potential $N$-body problem becomes equivalent to the problem of finding geodesics for the metric defined by Riemannian submersion on $Y_N$.

**Remark.** The metric quotient procedure just described realizes the Marsden–Weinstein symplectic reduced space of $T^*(\mathbb{C}^N \setminus \Delta)$ by the action of translations, rotations and scalings, $\mathbb{C} \rtimes \mathbb{C}^*$, at momentum values 0, together with the $N$-body reduced Hamiltonian flow, but valid only at zero energy.

**Remark.** This metric on $Y_N$ can be expressed as $U\, ds^2_{\mathrm{FS}}$ where $ds^2_{\mathrm{FS}}$ is the usual Fubini–Study metric on $\mathbb{CP}^{N-2}$.

**Remark.** For the standard $1/r^2$ potential of (1), this metric on $Y_N$ is complete, with infinite volume.

The collinear $N$-body problem defines a totally geodesic submanifold

$$\mathbb{RP}^{N-2} \setminus \Delta \subset \mathbb{CP}^{N-2} \setminus \Delta.$$

We obtain this submanifold by placing the $N$-masses anywhere along the real axis $\mathbb{R} \subset \mathbb{C}$, arranged so their center of mass is zero and so that there are no collisions, and then taking the quotient. In other words, $\mathbb{RP}^{N-2} \setminus \Delta$ is the quotient of $\mathbb{R}^{N-1} \subset \mathbb{C}^{N-1}$ by dilations and real reflections.

## 3. Main result

In case $N = 3$, with the potential (1) above, $Y_3$ is a pair of pants — a sphere minus three points. The point of [Montgomery 2005] was to show that the metric on $Y_3$ just described is hyperbolic provided $m_1 = m_2 = m_3$. Specifically, in this equal mass case the Gaussian curvature of the metric on the surface $Y_3$ is negative everywhere except at two points (these being the "Lagrange points" corresponding to equilateral triangles.) What about $Y_4$?

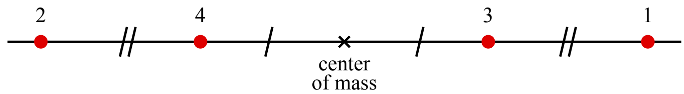


**Figure 1.** The collinear configurations $p$ which we consider.

**Theorem 1.** *Consider the Jacobi–Maupertuis metric on $Y_4$ induced as above for the case of* 4 *equal masses under the strong force* $1/r^2$ *potential* (1). *Then there are* 2*-planes $\sigma$ tangent to $Y_4$ at which the Riemannian sectional curvature $\mathcal{K}(\sigma)$ is positive.*

**Remark.** The 2-planes $\sigma$ of the theorem pass through special points $p \in \mathbb{RP}^2 \subset \mathbb{CP}^2$ which represent certain special collinear configurations. See Figure 1. The 2-plane $\sigma$ at $p$ will be the orthogonal complement to $T_p\mathbb{RP}^2$, the normal 2-plane, and is realized as $\sigma = iT_p\mathbb{RP}^2$, using the standard complex structure on $\mathbb{CP}^2$.

**Remark** (negative curvatures). The $\mathbb{RP}^2$ of the previous remark is a totally geodesic surface fixed by an isometric involution. There are other such totally geodesic surfaces defined as fixed loci of symmetries, and computer experiments suggest that these all have negative Gaussian curvature everywhere while their normal 2-planes can have positive sectional curvature at some points, like our special case $\mathbb{RP}^2$. Computer experiments also indicate that in the direction of the normal plane there is positive sectional curvature over all collinear configurations of $\mathbb{RP}^2$ and not just the special configurations verified in the theorem. An analytic proof of these claims beyond our special case, however, looks frightening.

**Remark** (uniqueness of free homotopy classes). The work in [Montgomery 2005] was chiefly meant as a route for proving the uniqueness (mod symmetries) of the $N = 3$ strong force figure-eight solution. For $N = 4$, hyperbolicity fails and we have no direct "hyperbolic" path for establishing uniqueness of various 4-body choreographies or free homotopy class representatives.

**Open Question.** A geodesic flow can still be hyperbolic as a flow, without the underlying metric having all sectional curvatures negative. Is geodesic flow on $Y_4$ hyperbolic as a flow? Is it even partially hyperbolic?

## 4. Proof of the theorem

We take the case $N = 4$ in the above considerations. When all the masses are equal to 1, the mass metric, used to compute the kinetic energy and moment of inertia, is the standard Hermitian metric in coordinates $(q_1, q_2, q_3, q_4) \in \mathbb{C}^4$, where the $q_i$ represent the positions of the $i$-th body. We reduce by translations by going to the center-of-mass-zero space, which is a 3-dimensional subspace $\mathbb{C}^3 \subset \mathbb{C}^4$ having

Jacobi coordinates as Hermitian orthonormal coordinates:

$$\mathbb{C}^3 \xrightarrow{L} \mathbb{C}^4 \text{ given by the matrix } \begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix} \text{ in standard bases.}$$

As is well-known, if we start tangent to the center-of-mass-zero subspace $L(\mathbb{C}^3)$, we stay tangent to it. Hence we can restrict the dynamics, potential, metric, etc. to the center-of-mass-zero subspace. We denote the potential restricted to the center-of-mass-zero subspace in Jacobi coordinates as $U_L = U \circ L$ and still write $ds^2_{\mathrm{JM}} = U_L ds^2$ for the restricted JM metric on $\mathbb{C}^3 \setminus \Delta$ where $ds^2$ is the standard metric on $\mathbb{C}^3$.

Continuing along the outline above, we now quotient by scaling and rotation isometries $\mathbb{C}^*$ of $ds^2_{\mathrm{JM}}$ to obtain the “shape space” $Y_4$ and we label the quotient map $\pi : \mathbb{C}^3 \setminus \Delta \to Y_4$, which takes a configuration $q$ to its orbit $\mathbb{C}^* q$. We denote the vertical and horizontal distributions as

$$\mathcal{V}_p = \ker d_p\pi = \mathbb{C}p \quad \text{and} \quad \mathcal{H}_p = \mathcal{V}_p^{\perp} \overset{d\pi}{\cong} T_{\pi(p)} Y_4.$$

Requiring $d\pi|_{\mathcal{H}, ds^2_{\mathrm{JM}}|_{\mathcal{H}}}$ to be an isometry defines our induced metric on $Y_4$ whose geodesics correspond to $N$-body motions in “shape space”. Under this induced metric on $Y_4$ we denote sectional curvature through the plane $\sigma \in T_{\pi(p)} Y_4$ by $\mathcal{K}(\sigma)$.

Suppressing the notation of evaluating at a representative $p \in \pi(q)$, our main tool in the computation of $\mathcal{K}(\sigma)$, the $ds^2_{\mathrm{JM}}$ curvature, is the equation

$$\begin{aligned} (2)\quad & U_L^3 \mathcal{K}(\sigma) \\ &= \tfrac{3}{4}\big((\partial_1 U_L)^2 + (\partial_2 U_L)^2\big) - \left\| \frac{\nabla U}{2} \right\|^2 - \frac{U_L}{2}(\partial_1^2 U_L + \partial_2^2 U_L) + 3\frac{U_L^2}{\|p\|^2}(v_1 \cdot i v_2)^2 \end{aligned}$$

Here $\partial_a f$ denotes $df(v_a)$ where $f \in C^\infty(\mathbb{C}^3)$ and where $a = 1, 2$ with $v_1, v_2 \in \mathcal{H}$ being $ds^2$-orthonormal vectors whose pushforwards $d\pi\, v_a$ span $\sigma$. The $\cdot$, $\| \, \|$, and $\nabla$ refer to the norm, metric, and Levi-Civita connection for the Euclidean metric $ds^2$. For the derivation of (2) see the Appendix.

The collinear configurations form a totally geodesic projective plane $\mathbb{RP}^2 \subset \mathbb{CP}^2$, the image under $\pi$ of the real 2-sphere in $\mathbb{C}^3$, which we parametrize by

$$p = (\cos\phi \cos\theta, \cos\phi \sin\theta, \sin\phi).$$

We evaluate (2) and find positive sectional curvature over the configurations with $\theta = \pi/2$ (see Figure 1) in the direction of the $iT\mathbb{RP}^2$ plane. This plane is spanned by the pushforwards of

$$v_1 = -i\frac{\partial p}{\partial \phi} = i(\sin\phi\cos\theta, \sin\phi\sin\theta, -\cos\phi),$$
$$v_2 = \frac{i}{\cos\phi}\frac{\partial p}{\partial \theta} = i(-\sin\theta, \cos\theta, 0).$$

**Terms 1.** *Over* $\mathbb{RP}^2$ *in the* $iT\mathbb{RP}^2$ *direction, the last and first summands on the second line of* (2) *vanish*:

$$v_1 \cdot iv_2 = 0, \qquad \partial_a U_L = 0.$$

*Proof.* That $v_1 \cdot iv_2 = 0$ is clear: $i$ rotates $v_2$ into purely real coordinates. To evaluate the first partials, note that $Lp$ has purely real coordinates and $\nabla U$ has $k$-th component $\sum_{j\neq k}(q_j - q_k)/r_{jk}^4$, so $\nabla|_{Lp}U$ has purely real coordinates. Now since $Lv_a$ has purely complex coordinates,

$$\partial_a U_L = \nabla|_{Lp}U \cdot Lv_a = 0. \qquad \square$$

**Terms 2.** *With the notation* $Lp = (q_1, q_2, q_3, q_4)$, $Lv_a = i(v_a^1, v_a^2, v_a^3, v_a^4)$, *and*

$$\rho_{jk} = 1/(q_j - q_k),\ \alpha_{jk} = (v_1^j - v_1^k)^2 + (v_2^j - v_2^k)^2 \in \mathbb{R},$$

*the sum of second partials in* (2) *is given by*

$$\partial_1^2 U_L + \partial_2^2 U_L = -2\sum_{j>k}\alpha_{jk}\rho_{jk}^4.$$

*Proof.* We write our standard coordinates on $\mathbb{C}^4$ as $q_j = x_j + iy_j$. Then since $Lv_a$ is purely imaginary, we have

$$\partial_a^2 U_L = \nabla|_{Lp}(\nabla U \cdot Lv_a)\cdot Lv_a = \left(\nabla|_{Lp}\frac{\partial U}{\partial y_k}v_a^k\right)\cdot Lv_a = \frac{\partial^2 U}{\partial y_j \partial y_k}\bigg|_{Lp} v_a^k v_a^j.$$

Next we compute $\displaystyle\frac{\partial^2 U}{\partial y_j \partial y_k}\bigg|_{Lp} = 2\rho_{jk}^4$ for $j \neq k$ and $\displaystyle\frac{\partial^2 U}{\partial y_k^2}\bigg|_{Lp} = -2\sum_{j\neq k}\rho_{jk}^4$, so now

$$\begin{aligned}\partial_a^2 U_L &= -2\sum_{j\neq k}\rho_{jk}^4\big((v_a^k)^2 - v_a^j v_a^k\big)\\ &= -2\sum_{j>k}\rho_{jk}^4\big((v_a^k)^2 - 2v_a^k v_a^j + (v_a^j)^2\big)\\ &= -2\sum_{j>k}\rho_{jk}^4(v_a^k - v_a^j)^2. \qquad \square\end{aligned}$$

**Result.** *Over the circle* $\theta = \pi/2$, $\mathcal{K}(iT\mathbb{RP}^2)$ *is positive.*

Now, substituting Terms 1 and 2 into formula (2), we see that

$$0<\mathcal{K}\Longleftrightarrow 0<U_L^3\mathcal{K}=-\|\nabla U/2\|^2+U_L\sum_{j>k}\alpha_{jk}\rho_{jk}^4$$

$$\Longleftrightarrow \sum_k\Big(\sum_{j\neq k}\rho_{jk}^3\Big)^2<\Big(\sum_{j>k}\rho_{jk}^2\Big)\Big(\sum_{j>k}\alpha_{jk}\rho_{jk}^4\Big). \tag{3}$$

Taking $\theta=\pi/2$ and with the notation introduced in Terms 2, we find the relations

$$\rho_{12}=\frac{1}{\sqrt{2}\cos\phi},\quad \rho_{34}=\frac{1}{\sqrt{2}\sin\phi}\qquad \alpha_{12}=\frac{1}{\rho_{34}^2},\quad \alpha_{34}=\frac{1}{\rho_{12}^2}$$

$$\rho_{13}=\frac{\sqrt{2}}{\cos\phi-\sin\phi}=-\rho_{24}\qquad \alpha_{13}=\frac{1}{\rho_{14}^2}+1=\alpha_{24}$$

$$\rho_{14}=\frac{\sqrt{2}}{\cos\phi+\sin\phi}=-\rho_{23}\qquad \alpha_{14}=\frac{1}{\rho_{13}^2}+1=\alpha_{23}.$$

Now the left side of (3) works out to

$$\begin{aligned}
&2((\rho_{12}^3+\rho_{13}^3+\rho_{14}^3)^2+(\rho_{13}^3-\rho_{14}^3-\rho_{34}^3)^2)\\
&\qquad=2\Big(\sum_{k>j}\rho_{jk}^6+2\rho_{12}^3(\rho_{13}^3+\rho_{14}^3)+2\rho_{34}^3(\rho_{14}^3-\rho_{13}^3)\Big)\\
&\qquad=2\sum_{k>j}\rho_{jk}^6-96\frac{1}{\sin^2 2\phi\cos^2 2\phi}=2\sum_{k>j}\rho_{jk}^6+\text{negative term},
\end{aligned}$$

and the right side of (3) works out to

$$\begin{aligned}
&\big(\rho_{12}^2+\rho_{34}^2+2(\rho_{13}^2+\rho_{14}^2)\big)\Big(\frac{\rho_{12}^4}{\rho_{34}^2}+\frac{\rho_{34}^4}{\rho_{12}^2}+2\Big(\rho_{13}^4+\rho_{14}^4+\frac{\rho_{13}^4}{\rho_{14}^2}+\frac{\rho_{14}^4}{\rho_{13}^2}\Big)\Big)\\
&\quad=\Big(\frac{2}{\sin^2 2\phi}+\frac{8}{\cos^2 2\phi}\Big)\Big(\sin^2 2\phi(\rho_{12}^6+\rho_{34}^6)\\
&\qquad\qquad+\frac{\cos^2 2\phi}{2}(\rho_{13}^6+\rho_{14}^6)+2(\rho_{13}^4+\rho_{14}^4)\Big)\\
&\quad=2\sum_{k>j}\rho_{jk}^6+\cot^2 2\phi(\rho_{13}^6+\rho_{14}^6)+8\tan^2 2\phi(\rho_{12}^6+\rho_{34}^6)\\
&\qquad\qquad+(\rho_{13}^4+\rho_{14}^4)\Big(\frac{4}{\sin^2 2\phi}+\frac{16}{\cos^2 2\phi}\Big)\\
&\quad=2\sum_{k>j}\rho_{jk}^6+\text{positive term}.
\end{aligned}$$

Therefore the inequality (3) holds! □

## Appendix: Derivation of (2)

Take a $ds^2$-orthonormal basis $\{v_a\}$ for $\mathbb{C}^3$ with $v_1, v_2 \in \mathcal{H}_p$.

The Kulkarni–Nomizu product formula for conformal curvatures [Sakai 1996, p. 51] reads:

$$\bar{R}_{abcd} - U_L R_{abcd} = -\left\{ds^2_{\mathrm{JM}} \owedge \left(\nabla du - du \otimes du + \tfrac{1}{2}\|du\|^2 ds^2\right)\right\}_{abcd}$$

where $u := \frac{1}{2}\log U_L$, the overbars denote curvature with respect to the $ds^2_{\mathrm{JM}}$-metric, and all other quantities (without overbars) are with respect to the $ds^2$-metric. Then $R_{abcd} = 0$ since $ds^2$ is the flat Euclidean metric of $\mathbb{C}^3 = \mathbb{R}^6$. Taking $cd = ab$, we have

$$\begin{aligned} U_L^2 \bar{K}_{ab} = \bar{R}_{abab} &= -U_L\left(\nabla du_{bb} + \nabla du_{aa} - du_b \otimes du_b - du_a \otimes du_a + \|du\|^2\right) \\ &= -U_L\left(\partial_a^2 u + \partial_b^2 u - (\partial_a u)^2 - (\partial_b u)^2 + \|\nabla u\|^2\right). \end{aligned}$$

Next, O'Neill's formula [1983, p. 213] gives

$$\mathcal{K}(d\pi v_1, d\pi v_2) = \bar{K}_{12} + \tfrac{3}{4}\left|[V_1, V_2]^{\mathcal{V}}\right|^2_{ds^2_{\mathrm{JM}}},$$

where $V_a = v_a/\sqrt{U_L(p)}$ and $X^{\mathcal{V}}$ denotes $ds^2_{\mathrm{JM}}$ projection of $X$ onto $\mathcal{V}$.

We then compute

$$\partial_a u = \frac{\partial_a U_L}{2U_L} = \frac{\nabla|_{Lp} U \cdot L v_a}{2U_L(p)}$$

and

$$\begin{aligned} \partial_a^2 u &= \frac{\partial_a^2 U_L}{2U_L} - \frac{(\partial_a U_L)^2}{2U_L^2} \\ &= \frac{\nabla|_{Lp}(\nabla U \cdot L v_a)\cdot L v_a}{2U_L(p)} - \frac{(\partial_a U_L)^2}{2U_L(p)^2}. \end{aligned}$$

Note that $\nabla U \in \left\{q \in \mathbb{C}^4 : \sum q_j = 0\right\}$ and $L v_a$ is a $ds^2$ orthonormal basis for this center-of-mass-zero subspace, hence

$$\begin{aligned} \|\nabla U\|^2 &= \textstyle\sum (\nabla U \cdot L v_a)^2 \\ &= \textstyle\sum (\partial_a U_L)^2 = 4U_L^2 \|\nabla u\|^2. \end{aligned}$$

Substitution into the Kulkarni–Nomizu formula gives

$$\bar{K}_{12} = -\frac{1}{U_L^3}\left(\frac{U_L}{2}(\partial_1^2 U_L + \partial_2^2 U_L) - \frac{3}{4}(\partial_1 U_L^2 + \partial_2 U_L^2) + \left\|\frac{\nabla U}{2}\right\|^2\right). \tag{4}$$

To compute O'Neill's Lie bracket term we write our standard coordinates on $\mathbb{C}^3$ as $(x^1 + ix^2, \ldots, x^5 + ix^6)$.

Let $H_1 = X^j \partial_{x^j}$, $H_2 = Y^j \partial_{x^j} \in \mathcal{H}$ be any horizontal vector fields. The vertical vector fields are spanned by the Euler vector field $E = x^j \partial_{x^j}$ and $iE$. Then $H_j \cdot E = H_j \cdot iE = 0$ and

$$\begin{aligned}
[H_1, H_2] \cdot E &= \sum_k X^j x^k \partial_{x^j} Y^k - Y^j x^k \partial_{x^j} X^k \\
&= \sum_k X^j (\partial_{x^j}(x^k Y^k) - \delta^k_j Y^k) - Y^j (\partial_{x^j}(x^k X^k) - \delta^k_j X^k) \\
&= \sum_k X^k Y^k - Y^k X^k = 0,
\end{aligned}$$

and likewise,

$$\begin{aligned}
[H_1, H_2] \cdot iE &= \sum_{k \text{ odd}} (Y^j \partial_{x^j} X^k - X^j \partial_{x^j} Y^k) x^{k+1} + (X^j \partial_{x^j} Y^{k+1} - Y^j \partial_{x^j} X^{k+1}) x^k \\
&= 2 \sum_{k \text{ odd}} -X^k Y^{k+1} + X^{k+1} Y^k = 2 H_1 \cdot i H_2.
\end{aligned}$$

Then

$$\begin{aligned}
\big|[V_1, V_2]^{\mathcal{V}_p}\big|^2 &= ds^2_{\mathrm{JM}}\biggl([V_1, V_2], \frac{E_p}{|p|\sqrt{U_L(p)}}\biggr)^2 + ds^2_{\mathrm{JM}}\biggl([V_1, V_2], \frac{iE_p}{|p|\sqrt{U_L(p)}}\biggr)^2 \\
&= \frac{U_L^2}{|p|^2 U_L}\big(([V_1, V_2] \cdot E)^2 + ([V_1, V_2] \cdot iE)^2\big) \\
&= \frac{4U_L(p)(V_1 \cdot i V_2)^2}{|p|^2} = \frac{4}{U_L(p)|p|^2}(v_1 \cdot i v_2)^2.
\end{aligned}$$

Now substitution of this Lie bracket expression and (4) into O'Neill's formula and multiplying by $U_L^3$ yields (2). □

## Acknowledgments

We would like to thank Slobodan Simic of San Jose State University for discussions on hyperbolicity, and Prof. Jie Qing and Wei Yuan of UC Santa Cruz for conversations regarding curvature computations. We thankfully acknowledge support by NSF grant DMS-1305844.

## References

[Montgomery 2005] R. Montgomery, "Fitting hyperbolic pants to a three-body problem", *Ergodic Theory Dynam. Systems* **25**:3 (2005), 921–947. MR 2006b:70013 Zbl 1069.70006

[O'Neill 1983] B. O'Neill, *Semi–Riemannian geometry: With applications to relativity*, 1st ed., Pure and Applied Mathematics **103**, Academic Press, New York, 1983. MR 85f:53002 Zbl 0531.53051

[Sakai 1996] T. Sakai, *Riemannian geometry*, 1st ed., Translations of Mathematical Monographs **149**, Amer. Math. Soc., Providence, RI, 1996. Zbl 0886.53002

Received February 5, 2015. Revised May 8, 2015.

CONNOR JACKMAN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
SANTA CRUZ, CA 95064
UNITED STATES
cfjackma@ucsc.edu

RICHARD MONTGOMERY
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
SANTA CRUZ, CA 95064
UNITED STATES
rmont@ucsc.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.411

# UNIONS OF LEBESGUE SPACES AND $A_1$ MAJORANTS

GREG KNESE, JOHN E. M^C CARTHY AND KABE MOEN

**We study two questions. When does a function belong to the union of Lebesgue spaces, and when does a function have an $A_1$ majorant? We provide a systematic study of these questions and show that they are fundamentally related. We show that the union of $L^p_w(\mathbb{R}^n)$ spaces with $w \in A_p$ is equal to the union of all Banach function spaces for which the Hardy–Littlewood maximal function is bounded on the space itself and its associate space.**

## 1. Introduction and statement of the main results

While the $L^p$ spaces are considered fundamental spaces of interest in analysis, the weighted $L^p$ spaces and the related study of $A_p$ weights are perhaps part of a more specialized area of analysis. It is the goal of this article to show that the $L^p$ spaces considered in aggregate are intimately linked to these latter topics and to the notion of an $A_1$ majorant. By recent developments our results indicate that weighted Lebesgue spaces with $A_p$ weights may be good candidates for ambient spaces for operators in harmonic analysis.

We begin with the following question.

**Question 1.1.** When does a function belong to the union of $L^p$ spaces?

Question 1.1 is vaguely stated on purpose. By union, we mean either the union of $L^p$ as $p$ varies or the union of $L^p_w$ as $w$ varies with $p$ fixed. The union of $L^p$ spaces often arises when considering a general domain to define operators in harmonic analysis. Several such operators are bounded on $L^p$ for all $1 < p < \infty$, and hence take functions from $\bigcup_{p>1} L^p$ into itself.

It turns out Question 1.1 is closely related to the theory of weighted Lebesgue spaces and the action of the Hardy–Littlewood maximal operator on these spaces.

Knese partially supported by NSF grants DMS-1419034, DMS-1363239.
M^c Carthy partially supported by NSF grant DMS-1300280.
Moen partially supported by NSF grant DMS-1201504.
*MSC2010:* primary 42B25, 42B35, 46E30; secondary 30H10, 30H15.
*Keywords:* maximal functions, $L^p$ spaces, Hardy spaces, $A_p$ spaces, weighted $L^p$ spaces.

For our purposes, a weight is a positive locally integrable function. An $A_1$ weight is one that satisfies

$$Mw \le Cw \quad \text{a.e.}$$

Here $M$ denotes the Hardy–Littlewood maximal operator

$$Mf(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f| \, dx.$$

We exclude the weight $w \equiv 0$ from belonging to $A_1$, and in this case we see that if $w \in A_1$ then $w > 0$ a.e. The $A_1$ class of weights characterizes when $M$ maps $L^1_w$ into $L^{1,\infty}_w$. When $1 < p < \infty$, $M$ is bounded on $L^p_w$ exactly when $w \in A_p$:

$$\left(\frac{1}{|Q|}\int_Q w\,dx\right)\left(\frac{1}{|Q|}\int_Q w^{-1/(p-1)}\,dx\right)^{p-1} \le C$$

for all cubes $Q$. At the other endpoint the $A_\infty$ class is defined to be the union of all $A_p$ for $p \ge 1$. We now come to our second question.

**Question 1.2.** Given a measurable function $f$, when does there exist an $A_1$ weight $w$ such that

$$|f| \le w? \tag{1}$$

We call a weight satisfying (1) an $A_1$ majorant of $f$ and write $\mathcal{M}_{A_1}$ for the set of measurable functions possessing an $A_1$ majorant. As stated, Question 1.2 does not seem to have been considered before. As far as we can tell, the first notion of an $A_1$ majorant appeared in an article by Rutsky [2011]. In Rutsky's paper, however, a different definition of an $A_1$ majorant is given — one which requires the function and the weight to a priori belong to a more restrictive class of functions.

If we examine weights locally, say on the interval [0, 1], then our problem has a remarkably simple answer which reveals a close connection between traditional $L^p$ spaces, weighted $L^p$ spaces, and $A_1$ majorants:

$$\mathcal{M}_{A_1}([0,1]) = \bigcup_{p>1} L^p([0,1]) = \bigcup_{w \in A_2} L^2_w([0,1]). \tag{2}$$

The proof of (2) is a synthesis of known important results for Muckenhoupt weights. This equivalence reinforces the saying attributed to Antonio Córdoba, "There are no $L^p$ spaces, only weighted $L^2$ spaces."

The local theory has several extensions including an application to Hardy spaces on the unit disk. In [M^c^Carthy 1990], while studying the range of Toeplitz operators, the second author showed that the Smirnov class, $N^+$, can be realized as a union of weighted Hardy spaces:

$$N^+ = \bigcup_{w \in \mathcal{W}} H^2_w$$

where $\mathscr{W}$ is the Szegő class of weights (see Section 2 for relevant definitions). The class $A_\infty(\mathbb{T})$ is a proper subset of $\mathscr{W}$ (as $\bigcup_{p>0} H^p$ is a proper subspace of $N^+$). Using our techniques we are able to give a characterization of $\bigcup_{p>0} H^p$ in terms of weighted $H^2$ spaces:

$$\bigcup_{p>0} H^p = \bigcup_{w\in A_\infty} H^2_w. \tag{3}$$

We refer the reader to Section 4 for more on the local case.

For functions on $\mathbb{R}^n$, the theory is not as nice. In the local case the $L^p([0,1])$ spaces are nested in $p$, whereas the $L^p(\mathbb{R}^n)$ spaces are not. We are not able to obtain equality of $\bigcup_{p>1} L^p(\mathbb{R}^n)$ and $\mathscr{M}_{A_1}(\mathbb{R}^n)$. Remarkably, even the much larger union over weak-$L^p(\mathbb{R}^n)$ spaces is not equal to $\mathscr{M}_{A_1}(\mathbb{R}^n)$. As a consequence of our results, if $p_0$ is any exponent satisfying $1 < p_0 < \infty$ then

$$\bigcup_{p>1} L^{p,\infty}(\mathbb{R}^n) \subsetneqq \bigcup_{w\in A_{p_0}} L^{p_0}_w(\mathbb{R}^n) \subsetneqq \mathscr{M}_{A_1}(\mathbb{R}^n). \tag{4}$$

The class $\mathscr{M}_{A_1}(\mathbb{R}^n)$ can be thought of as a generalization of $L^\infty(\mathbb{R}^n)$ — i.e., functions that are majorized by constants, which are $A_1$ weights — while $\bigcup_{w\in A_1} L^1_w(\mathbb{R}^n)$ is a generalization of $L^1(\mathbb{R}^n)$. With this in mind we obtain the following theorem.

**Theorem 1.3.** *Suppose* $1 < p < \infty$. *Then*

$$\bigcup_{w\in A_p} L^p_w(\mathbb{R}^n) = \mathscr{M}_{A_1}(\mathbb{R}^n) \cap \Bigg( \bigcup_{w\in A_1} L^1_w(\mathbb{R}^n) \Bigg).$$

Considering the basic fact

$$L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) \subset \bigcap_{1<p<\infty} L^p(\mathbb{R}^n),$$

Theorem 1.3 shows that if we enlarge both $L^\infty(\mathbb{R}^n)$ to $\mathscr{M}_{A_1}(\mathbb{R}^n)$ and $L^1(\mathbb{R}^n)$ to $\bigcup_{w\in A_1} L^1_w(\mathbb{R}^n)$ and intersect the two, then we pick up an even bigger class of functions, one that by (4) properly contains the union of all $L^p(\mathbb{R}^n)$ for $p > 1$. As a consequence to Theorem 1.3, we see that for all $1 < p, q < \infty$,

$$\bigcup_{w\in A_p} L^p_w(\mathbb{R}^n) = \bigcup_{u\in A_q} L^q_u(\mathbb{R}^n).$$

The proof of Theorem 1.3 uses the extrapolation theory of Rubio de Francia [1984; 1987] (see also the book [Cruz-Uribe et al. 2011]).

The union $\bigcup_{p>1} L^p$ is a good candidate for a natural collection of functions on which to iterate the Hardy–Littlewood maximal function. Rutsky [2014, Theorem 1] showed that Banach function spaces $\mathscr{X}$ on $\mathbb{R}^n$ (see Section 2) for which the Hardy–Littlewood maximal function is bounded on both the space $\mathscr{X}$ and the associate space $\mathscr{X}'$ act as a natural domain for the set of all Calderón–Zygmund operators. We end the introduction with our main result which says a function belongs to a function space $\mathscr{X}$ for which the Hardy–Littlewood maximal function is bounded on $\mathscr{X}$ and $\mathscr{X}'$ if and only if $f \in L^p_w(\mathbb{R}^n)$ for some $p > 1$ and $w \in A_p(\mathbb{R}^n)$.

**Theorem 1.4.** *Suppose* $1 < p < \infty$. *Then*

$$\bigcup_{w \in A_p} L^p_w(\mathbb{R}^n) = \bigcup \{\mathscr{X} : M \in \mathscr{B}(\mathscr{X}) \cap \mathscr{B}(\mathscr{X}')\},$$

*where the second union is over all Banach function spaces such that the Hardy–Littlewood maximal operator is bounded on* $\mathscr{X}$ *and* $\mathscr{X}'$.

Banach function spaces for which $M \in \mathscr{B}(\mathscr{X}) \cap \mathscr{B}(\mathscr{X}')$ are also related to the Fefferman–Stein inequality. Define the sharp maximal function $M^\#$ by

$$M^\# f(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f - f_Q| \, dx,$$

where $f_Q = \frac{1}{|Q|} \int_Q f \, dx$. Lerner [2010] proved that if $M \in \mathscr{B}(\mathscr{X})$, then the Fefferman–Stein inequality

$$\|f\|_{\mathscr{X}} \le c \|M^\# f\|_{\mathscr{X}} \tag{5}$$

holds for all nice functions in $\mathscr{X}$ if and only if $M \in \mathscr{B}(\mathscr{X}')$. In particular, Theorem 1.4 shows that if $f$ belongs to a Banach function space for which $M \in \mathscr{B}(\mathscr{X})$ and the Fefferman–Stein inequality (5) holds on $\mathscr{X}$, then for any $1 < p < \infty$, there exists $w \in A_p$ for which $f \in L^p_w(\mathbb{R}^n)$.

The outline of this paper is as follows. In Section 2 we state preliminary results that are necessary for the rest of the paper. In Section 3 we study the classes of functions with $A_1$ and $A_p$ majorants. In Section 4 we give a treatise of local theory with applications to Hardy spaces on the unit disk. Section 5 is devoted to the theory on $\mathbb{R}^n$, in particular the proofs of Theorems 1.3 and 1.4. We finish the article with some open questions in Section 6.

## 2. Preliminaries

In this section, $\Omega$ denotes either $\mathbb{R}^n$ or a cube $Q$ with sides parallel to the coordinate planes in $\mathbb{R}^n$. For $0 < p < \infty$, $L^p(\Omega)$ is the set of measurable functions such that

$$\|f\|^p_{L^p} = \int_\Omega |f|^p \, dx < \infty.$$

Given $p$ with $1 \le p \le \infty$, we use $p'$ to denote the dual exponent defined by the equation $1/p + 1/p' = 1$. A weight defined on a cube $Q$ is a positive function in $L^1(Q)$. A weight on $\mathbb{R}^n$ is a positive function in $L^1_{\text{loc}}(\mathbb{R}^n)$. Given a weight, $w$, define $L^p_w(\Omega)$ to be the collection of functions satisfying

$$\|f\|^p_{L^p_w} = \int_\Omega |f|^p w\,dx < \infty.$$

We define $L^\infty_w(\Omega)$ to be the space of functions for which $f/w \in L^\infty(\Omega)$. This space is normed by

$$\|f\|_{L^\infty_w} = \|f/w\|_\infty = \operatorname*{ess\,sup}_{x\in\Omega} \frac{|f(x)|}{w(x)}.$$

If $\mathbb{T}$ is the unit circle in the complex plane, then $L^p(\mathbb{T})$ and $L^p_w(\mathbb{T})$ are identified as the space of $2\pi$ periodic functions that belong to $L^p([0, 2\pi])$ and $L^p_w([0, 2\pi])$, respectively.

We also examine the "complex analyst's Hardy space", as opposed to the real analyst's Hardy space defined in terms of maximal functions. Let $\mathbb{D}$ denote the unit disk in the plane with boundary $\mathbb{T}$. Given $p$ with $0 < p < \infty$, let $H^p = H^p(\mathbb{D})$ be the space of analytic functions "normed" by

$$\|f\|_{H^p} = \sup_{0<r<1} \left( \int_0^{2\pi} |f(re^{i\theta})|^p \,\frac{d\theta}{2\pi} \right)^{1/p}.$$

"Norm" is in quotes since this is not a norm for $0 < p < 1$, but we use norm notation $\|\cdot\|$ nonetheless. The Nevanlinna class, denoted $N$, is the collection of analytic functions on $\mathbb{D}$ such that

$$\|f\|_N = \sup_{0<r<1} \int_0^{2\pi} \log^+ |f(re^{i\theta})| \,\frac{d\theta}{2\pi} < \infty.$$

Functions in $N$ have nontangential limits almost everywhere on the boundary, so we may treat them as functions on the disk or the circle. The Smirnov class $N^+$ consists of functions $f \in N$ such that

$$\lim_{r\to 1} \int_0^{2\pi} \log^+ |f(re^{i\theta})| \,\frac{d\theta}{2\pi} = \int_0^{2\pi} \log^+ |f(e^{i\theta})| \,\frac{d\theta}{2\pi}.$$

It is well known that

$$\bigcup_{p>0} H^p \subsetneq N^+ \subsetneq N$$

(see, e.g., the books by Duren [1970] or Rudin [1964]). The Smirnov class is often considered a natural limit of $H^p$ as $p \to 0$.

The weighted Hardy space $H^p_w = H^p_w(\mathbb{D})$ is the closure of analytic polynomials in $L^p_w(\mathbb{T})$. While there are real variable definitions of weighted Hardy spaces, this classical definition has an intuitive appeal.

Let $M_\Omega$ be the Hardy–Littlewood maximal operator restricted to $\Omega$, i.e.,

$$M_\Omega f(x) = \sup_{\substack{Q\subset\Omega \\ x\in Q}} \frac{1}{|Q|}\int_Q |f|\,dy.$$

When $\Omega = \mathbb{R}^n$ we write $M_{\mathbb{R}^n} f = Mf$.

We define $A_1(\Omega)$ to be the class of all weights on $\Omega$ such that $M_\Omega w(x) \le C w(x)$ a.e. $x \in \Omega$. For $p > 1$, $A_p(\Omega)$ is the class of all weights on $\Omega$ such that

$$\sup_{Q\subset\Omega}\left(\frac{1}{|Q|}\int_Q w\,dx\right)\left(\frac{1}{|Q|}\int_Q w^{1-p'}\,dx\right)^{p-1} < \infty.$$

Given an $A_p$ weight $w$ we refer to the weight $\sigma = w^{1-p'}$ as the dual weight. For the endpoint, $p = \infty$, we use the definition

$$A_\infty(\Omega) = \bigcup_{p\ge 1} A_p(\Omega).$$

There are several other definitions of $A_\infty$, e.g., weights satisfying a reverse Jensen inequality, a reverse Hölder inequality, or a fairness condition with respect to Lebesgue measure [Duoandikoetxea 2001; Grafakos 2008].

A weight on the torus is a positive function in $L^1(\mathbb{T})$. The classes $A_1(\mathbb{T})$, $A_p(\mathbb{T})$, and $A_\infty(\mathbb{T})$ are defined analogously on $\mathbb{T}$. The Szegő class of weights, denoted $\mathcal{W}$, are weights on $\mathbb{T}$ satisfying

$$\int_{\mathbb{T}} \log w\,d\theta > -\infty.$$

We notice that if $w \in A_\infty(\mathbb{T})$, then we have

$$\left(\int_{\mathbb{T}} w\,\frac{d\theta}{2\pi}\right)\exp\left(-\int_{\mathbb{T}}\log w\,\frac{d\theta}{2\pi}\right) < \infty.$$

In particular, $A_\infty(\mathbb{T}) \subset \mathcal{W}$.

**Example 2.1.** Let $x_0 \in \Omega$, $1 \le p \le \infty$, and $w_{x_0}(x) = |x - x_0|^\alpha$. Then $w_{x_0} \in A_p(\Omega)$ if and only if $-n < \alpha < n(p-1)$.

We will need some elementary properties of $A_p$ weights, most of which follow from the definition (see [Duoandikoetxea 2001, Proposition 7.2]).

**Theorem 2.2.** *The following hold*:

(i) $A_1 \subset A_p \subset A_q \subset A_\infty$ *if* $1 < p < q < \infty$.

(ii) *For* $1 < p < \infty$, $w \in A_p$ *if and only if* $\sigma = w^{1-p'} \in A_{p'}$.

(iii) *If* $0 < s \le 1$ *and* $w \in A_p$, *then* $w^s \in A_p$.

(iv) *If* $u, v \in A_1$, *then* $uv^{1-p} \in A_p$.

It is interesting to note that the converse of (iv) also holds, but the proof is much more intricate. This was shown by Jones [1980] and later by Rubio de Francia [1982]. We emphasize that we do not need this converse statement, only the statement (iv).

We also need the following deeper property of $A_\infty$ weights known as the reverse Hölder inequality. See [Hytönen et al. 2012] for a simple proof with nice constants.

**Theorem 2.3.** *If $w \in A_\infty(\Omega)$, then there exists $s > 1$ such that for every cube $Q \subset \Omega$,*

$$\frac{1}{|Q|}\int_Q w^s\,dx \le \left(\frac{2}{|Q|}\int_Q w\,dx\right)^s.$$

As a corollary to Theorem 2.3 we have the following openness properties of $A_p$ classes.

**Theorem 2.4.** *Let $1 \le p \le \infty$. The following hold*:

(i) *If $p > 1$ then $A_p(\Omega) = \bigcup_{1\le q<p} A_q(\Omega)$.*

(ii) *If $w \in A_p(\Omega)$ then $w^s \in A_p(\Omega)$ for some $s > 1$.*

For the results on $\mathbb{R}^n$ we need the notion of a Banach function space. We refer the reader to the book by Bennett and Sharpley [1988, Chapter 1] for an excellent reference on the subject. A mapping $\rho$, defined on the set of nonnegative $\mathbb{R}^n$-measurable functions and taking values in $[0, \infty]$, is said to be a Banach function norm if it satisfies the following properties:

(i) $\rho(f) = 0 \Leftrightarrow f = 0$ a.e., $\rho(af) = a\rho(f)$ for $a > 0$, $\rho(f+g) \le \rho(f) + \rho(g)$;

(ii) if $0 \le f \le g$ a.e., then $\rho(g) \le \rho(f)$;

(iii) if $f_n \uparrow f$ a.e., then $\rho(f_n) \uparrow \rho(f)$;

(iv) if $B \subset \mathbb{R}^n$ is bounded, then $\rho(\chi_B) < \infty$;

(v) if $B \subset \mathbb{R}^n$ is bounded, then

$$\int_B f\,dx \le C_B\,\rho(f)$$

for some constant $C_B$ with $0 < C_B < \infty$.

We note that our definition of a Banach function space is slightly different from that found in [Bennett and Sharpley 1988]. In particular, in the axioms (iv) and (v) we assume that the set $B$ is a bounded set, whereas it is sometimes assumed that $B$ merely satisfy $|B| < \infty$. We do this so that the spaces $L^p_w(\mathbb{R}^n)$ with $w \in A_p$ satisfy items (iv) and (v). (See also the discussion at the beginning of Chapter 1 on page 2 of [Bennett and Sharpley 1988].)

Given Banach function norm $\rho$, $\mathscr{X} = \mathscr{X}(\mathbb{R}^n, \rho)$ is the collection of measurable functions such that $\rho(|f|) < \infty$. In this case we may equip $\mathscr{X}$ with the norm

$$\|f\|_{\mathscr{X}} = \rho(|f|).$$

The associate space $\mathscr{X}'$ is the set of all measurable functions $g$ such that $fg \in L^1(\mathbb{R}^n)$ for all $f \in \mathscr{X}$. This space is normed by

$$\|g\|_{\mathscr{X}'} = \sup\left\{\int_{\mathbb{R}^n} |fg|\,dx : \|f\|_{\mathscr{X}} \le 1\right\}. \tag{6}$$

Equipped with this norm $\mathscr{X}'$ is also a Banach function space and

$$\int_{\mathbb{R}^n} |fg|\,dx \le \|f\|_{\mathscr{X}}\|g\|_{\mathscr{X}'}.$$

Typical examples of Banach function spaces are $L^p(\mathbb{R}^n)$ for $1 \le p \le \infty$, whose associate spaces are $L^{p'}(\mathbb{R}^n)$. Other Banach spaces include weak type spaces $L^{p,\infty}(\mathbb{R}^n)$, the Lorentz space $L^{p,q}(\mathbb{R}^n)$, and Orlicz spaces $L^{\Phi}(\mathbb{R}^n)$ defined for a Young function $\Phi$ (see [Bennett and Sharpley 1988; Cruz-Uribe et al. 2011]). When $w \in A_p(\mathbb{R}^n)$ and $1 \le p \le \infty$, the spaces $L^p_w(\mathbb{R}^n)$ are also Banach function spaces with respect to Lebesgue measure. To see this, it suffices to check property (v). Suppose $f \ge 0$, $1 < p < \infty$, and $B$ is bounded. Then $B \subset Q$ for some cube $Q$ so $\sigma(B) < \infty$, and Hölder's inequality implies

$$\int_B f\,dx = \int_B f w^{1/p} w^{-1/p}\,dx \le \sigma(B)^{1/p'}\left(\int_B f^p w\,dx\right)^{1/p} \le \sigma(B)^{1/p'}\|f\|_{L^p_w}.$$

To see that $L^1_w(\mathbb{R}^n)$ is a Banach function space when $w \in A_1(\mathbb{R}^n)$, note that

$$\int_B f\,dx = \int_B f w w^{-1}\,dx \le (\inf_B w)^{-1}\|f\|_{L^1_w}. \tag{7}$$

Finally, if $f \in L^\infty_w$, then

$$\int_B f\,dx = \int_B (f/w) w\,dx \le w(B)\|f\|_{L^\infty_w},$$

showing $L^\infty_w$ is a Banach function space.

When $1 < p < \infty$ and $w \in A_p$, the associate space of $L^p_w(\mathbb{R}^n)$ defined by the pairing in (6) is given not by $L^{p'}_w(\mathbb{R}^n)$ but by $L^{p'}_\sigma(\mathbb{R}^n)$ for $\sigma = w^{1-p'}$. When $p = 1$ and $w \in A_1$, the associate space of $L^1_w$ is given by $L^\infty_w(\mathbb{R}^n)$. We are particularly interested in Banach function spaces $\mathscr{X}$ for which

$$\|Mf\|_{\mathscr{X}} \le C\|f\|_{\mathscr{X}},$$

in which case we write $M \in \mathscr{B}(\mathscr{X})$.

We end this section with the classical result of Coifman and Rochberg [1980] (see also [García-Cuerva and Rubio de Francia 1985, Theorem 3.4, p. 158]). This result requires a definition.

**Definition 2.5.** We say that a function $f(x)$ belongs to $\mathcal{M}_F(\Omega)$ if

$$M_\Omega f(x) < \infty \quad \text{for a.e. } x \in \Omega.$$

If $f$ belongs to a Banach function space for which $M \in \mathcal{B}(\mathcal{X})$, then $f \in \mathcal{M}_F$.

**Theorem 2.6.** *If $f \in \mathcal{M}_F(\Omega)$ and $0 < \delta < 1$, then $(M_\Omega f)^\delta \in A_1(\Omega)$.*

We leave the reader with a table of the notation used throughout the article.

| | |
|---|---|
| $\Omega$ | Domain of interest, either $\mathbb{R}^n$ or a cube $Q \subset \mathbb{R}$; |
| $M_\Omega$ | Hardy–Littlewood maximal operator restricted to $\Omega$; |
| $A_p(\Omega)$ | class of $A_p$ weights on $\Omega$; |
| $\mathcal{M}^r_{A_p}(\Omega)$ | functions on $\Omega$ with $\lvert f\rvert^r$ majorized by an $A_p$ weight; |
| $\mathcal{M}_F(\Omega)$ | functions on $\Omega$ such that $M_\Omega f < \infty$ a.e.; |
| $A^F_p(\Omega)$ | $A_p(\Omega) \cap \mathcal{M}_F(\Omega)$; |
| $\mathcal{M}_{A^F_p}(\Omega)$ | functions majorized by $A^F_p(\Omega)$ weights. |

## 3. The classes $\mathcal{M}^r_{A_p}$

Let us now define a general class of functions majorized by $A_p$ weights and establish some properties of such classes. We remind the reader that a domain $\Omega$ will denote throughout either all of $\mathbb{R}^n$ or a cube $Q$ in $\mathbb{R}^n$.

**Definition 3.1.** Let $r$ and $p$ satisfy $0 < r < \infty$ and $1 \le p \le \infty$. Define $\mathcal{M}^r_{A_p}(\Omega)$ to be the collection of all measurable functions $f$ on $\Omega$ such that

$$|f(x)|^r \le w(x) \quad \text{for a.e. } x \in \Omega$$

for some $w \in A_p(\Omega)$. When $r = 1$ we simply write $\mathcal{M}_{A_p}(\Omega)$.

Theorem 2.4 implies the following general facts about the $\mathcal{M}^r_{A_p}$ classes.

**Theorem 3.2.** *Suppose $r$ and $p$ satisfy $0 < r < \infty$ and $1 \le p \le \infty$. Then*

$$\mathcal{M}^r_{A_p}(\Omega) = \bigcup_{s>r} \mathcal{M}^s_{A_p}(\Omega) \tag{8}$$

*and if $p > 1$,*

$$\mathcal{M}^r_{A_p}(\Omega) = \bigcup_{1\le q<p} \mathcal{M}^r_{A_q}(\Omega). \tag{9}$$

*Proof.* We first prove (8). It is clear from (iii) of Theorem 2.2 that the union $\bigcup_{r<s} \mathcal{M}^s_{A_p}(\Omega) \subset \mathcal{M}^r_{A_p}(\Omega)$. On the other hand, if $f \in \mathcal{M}^r_{A_p}(\Omega)$ then $|f|^r \le w \in A_p$. By (ii) of Theorem 2.4, there exists $t > 1$ such that $w^t \in A_p(\Omega)$. But then, taking $s = rt > r$ and $u = w^t$, we have $|f|^s \le u \in A_p$, so $f \in \bigcup_{r<s} \mathcal{M}^s_{A_p}(\Omega)$. The proof of equality (9) follows directly from (i) of Theorem 2.4. □

Our next theorem shows that for a function to have an $A_1$ majorant it is equivalent for its maximal function to have an $A_1$ majorant.

**Theorem 3.3.** *We have $f \in \mathcal{M}_{A_1}(\Omega)$ if and only if $M_\Omega f \in \mathcal{M}_{A_1}(\Omega)$.*

*Proof.* If $f \in \mathcal{M}_{A_1}(\Omega)$, then $M_\Omega f \le M_\Omega w \le Cw$ since $w \in A_1(\Omega)$, which is to say $M_\Omega f \in \mathcal{M}_{A_1}(\Omega)$. The converse statement follows from the fact that $|f| \le M_\Omega f$. □

Using the exact same reasoning it is easy to prove that $f \in \mathcal{M}^r_{A_1}(\Omega)$ if and only if $M_\Omega(|f|^r) \in \mathcal{M}_{A_1}(\Omega)$. However, there is a better result when $r \ge 1$.

**Theorem 3.4.** *If $r \ge 1$ then the following are equivalent*:

(i) $f \in \mathcal{M}^r_{A_1}(\Omega)$.

(ii) $M_\Omega(|f|^r) \in \mathcal{M}_{A_1}(\Omega)$.

(iii) $M_\Omega f \in \mathcal{M}^r_{A_1}(\Omega)$.

*Proof.* The equivalence (i) ⇔ (ii) follows from Theorem 3.3. We will prove (ii) ⇒ (iii) and (iii) ⇒ (i).

Suppose that $w \in A_1(\Omega)$ and $M_\Omega(|f|^r) \le w$. Since $r \ge 1$, we know that $(M_\Omega f)^r \le M_\Omega(|f|^r) \le w$, which is to say that $M_\Omega f \in \mathcal{M}^r_{A_1}$.

On the other hand if $(M_\Omega f)^r \le w \in A_1(\Omega)$, then $M_\Omega f < \infty$ a.e., and hence $f$ is locally integrable on $\Omega$. By the Lebesgue differentiation theorem we have

$$|f|^r \le (M_\Omega f)^r \le w. \qquad \square$$

In the case $0 < r < 1$, we still have $f \in \mathcal{M}^r_{A_1}(\Omega)$ if and only if $M_\Omega(|f|^r) \in \mathcal{M}_{A_1}(\Omega)$. However, it is not true that this is equivalent to $(M_\Omega f)^r \in \mathcal{M}_{A_1}(\Omega)$. Consider the following simple example.

**Example 3.5.** Let $f(x) = |x|^{-n}$ on $Q = [-1, 1]^n$. If $0 < r < 1$, then $f \in M^r_{A_1}(Q)$ but $M_Q f \equiv \infty$.

Of course, if $0 < r < 1$ and $M_\Omega f < \infty$ a.e., then $(M_\Omega f)^r \in A_1(\Omega)$ (and hence $M_\Omega f \in \mathcal{M}^r_{A_1}(\Omega)$) automatically by Theorem 2.6.

We now study the class $\mathcal{M}_{A_p}$. Since the $A_p$ classes are nested, we have

$$\mathcal{M}_{A_1} \subset \mathcal{M}_{A_p} \subset \mathcal{M}_{A_q} \subset \mathcal{M}_{A_\infty}$$

for $1 \le p \le q \le \infty$. In the local case we have the following characterization.

**Theorem 3.6.** *If $Q$ is a cube in $\mathbb{R}^n$ then*

$$\mathcal{M}_{A_1}(Q) = \mathcal{M}_{A_\infty}(Q).$$

*Proof.* It suffices to show $\mathcal{M}_{A_\infty}(Q) \subset \mathcal{M}_{A_1}(Q)$. Suppose that $f \in \mathcal{M}_{A_\infty}(Q)$, so that there exists $w \in A_\infty(Q)$ with

$$|f| \leq w.$$

Since $w \in A_\infty(Q)$, the reverse Hölder inequality implies that there exists $s > 1$ such that

$$(M_Q w^s)^{1/s} \leq 2M_Q w \leq 2(M_Q w^s)^{1/s}.$$

Moreover, since $w \in L^1(Q)$, we have $M_Q w < \infty$ a.e. By Theorem 2.6, $M_Q w$ is bounded above and below by an $A_1(Q)$ weight, and hence is in $A_1(Q)$ itself. $\square$

In the global case we have $\mathcal{M}_{A_1}(\mathbb{R}^n) \subsetneq \mathcal{M}_{A_p}(\mathbb{R}^n)$ for any $p > 1$, as the following example indicates.

**Example 3.7.** Let $p > 1$ and $0 < \alpha < n(p-1)$. Now consider the function $f(x) = |x|^\alpha$. Then $f \in A_p(\mathbb{R}^n) \subset \mathcal{M}_{A_p}(\mathbb{R}^n)$, but $f \notin \mathcal{M}_F(\mathbb{R}^n)$ so in particular, $f \notin \mathcal{M}_{A_1}(\mathbb{R}^n)$. To see this, notice that for every $x \in \mathbb{R}^n$ and $r > |x|$,

$$Mf(x) \geq \frac{c}{r^n} \int_{|x| \leq r} |x|^\alpha \, dx \simeq r^\alpha$$

so $Mf \equiv \infty$.

To obtain positive results on $\mathbb{R}^n$ for the classes $\mathcal{M}_{A_p}(\mathbb{R}^n)$ and $\mathcal{M}_{A_\infty}(\mathbb{R}^n)$ similar to Theorem 3.6, we must restrict to $A_p$ majorants whose maximal function is finite. Given $w \in A_\infty$, a simple way to create a weight in $A_\infty^F$ is to take a truncation: let $w_\lambda = \min(w, \lambda)$ for $\lambda > 0$. Then $w_\lambda \in A_\infty \cap L^\infty \subset A_\infty^F$. We end our study of the class $\mathcal{M}_{A_1}$ with the following characterizations.

**Theorem 3.8.** $\mathcal{M}_{A_1}(\mathbb{R}^n) = \mathcal{M}_{A_\infty^F}(\mathbb{R}^n)$.

*Proof.* Since $A_1(\mathbb{R}^n) \subset A_\infty(\mathbb{R}^n)$ and $A_1(\mathbb{R}^n) \subset \mathcal{M}_F(\mathbb{R}^n)$, we have the inclusion $\mathcal{M}_{A_1}(\mathbb{R}^n) \subset \mathcal{M}_{A_\infty^F}(\mathbb{R}^n)$. On the other hand, if $f$ is dominated by a weight $w$ in $A_\infty^F(\mathbb{R}^n) = A_\infty(\mathbb{R}^n) \cap \mathcal{M}_F(\mathbb{R}^n)$, then by Theorem 2.3 we have

$$M(w^s)^{1/s} \leq 2Mw < \infty \quad \text{a.e.}$$

for some $s > 1$. So in particular, $|f| \leq M(|f|^s)^{1/s} \leq M(w^s)^{1/s} \in A_1(\mathbb{R}^n)$. $\square$

**Theorem 3.9.** *A function $f$ belongs to $\mathcal{M}_{A_1}(\mathbb{R}^n)$ if and only if there is an $s > 1$ such that $|f|^s \in \mathcal{M}_F(\mathbb{R}^n)$.*

**Remark 3.10.** Given $r > 0$, if one defines $\mathcal{M}_F^r(\mathbb{R}^n)$ to be the class of functions such that $M(|f|^r) < \infty$ a.e. (equivalently $|f|^r \in \mathcal{M}_F(\mathbb{R}^n)$), then Theorem 3.9 can be stated as

$$\mathcal{M}_{A_1}(\mathbb{R}^n) = \bigcup_{r>1} \mathcal{M}_F^r(\mathbb{R}^n).$$

*Proof of Theorem 3.9.* Let $w$ be an $A_1(\mathbb{R}^n)$ majorant of $f$. Since $w \in A_1(\mathbb{R}^n)$, $w^s \in A_1(\mathbb{R}^n)$ for some $s > 1$, which implies $|f|^s \in \mathcal{M}_{A_1}(\mathbb{R}^n)$. By Theorem 3.4 we have $M(|f|^s) \in \mathcal{M}_{A_1}(\mathbb{R}^n) \subset L^1_{\text{loc}}(\mathbb{R}^n)$. On the other hand, if there exists $s > 1$ such that $M(|f|^s) < \infty$ a.e., then $M(|f|^s)^{1/s} \in A_1(\mathbb{R}^n)$ by Theorem 2.6, and $|f| \le M(|f|^s)^{1/s}$. □

## 4. The local case

For this section $Q$ will be a fixed cube in $\mathbb{R}^n$. We begin with the following extension of the equivalences in (2).

**Theorem 4.1.** *Let $Q$ be a cube in $\mathbb{R}^n$ and $r$, $p_0$ satisfy $0 < r < p_0 < \infty$. Then*

$$\mathcal{M}_{A_1}^r(Q) = \bigcup_{p>r} L^p(Q) = \bigcup_{w \in A_{p_0/r}} L_w^{p_0}(Q).$$

*Proof.* We will prove the chain of containments

$$\bigcup_{w \in A_{p_0/r}} L_w^{p_0}(Q) \subset \bigcup_{p>r} L^p(Q) \subset \mathcal{M}_{A_1}^r(Q) \subset \bigcup_{w \in A_{p_0/r}} L_w^{p_0}(Q).$$

- $\big(\bigcup_{w \in A_{p_0/r}} L_w^{p_0}(Q) \subset \bigcup_{p>r} L^p(Q)\big)$: Suppose we have $f \in L_w^{p_0}(Q)$ for some $w \in A_{p_0/r}(Q)$. Set $q_0 = p_0/r$. By (ii) of Theorem 2.2, $\sigma = w^{1-q_0'} \in A_{q_0'}(Q)$. By Theorem 2.3, $\sigma$ satisfies a reverse Hölder inequality:

$$\left(\frac{1}{|Q'|}\int_{Q'} \sigma^s\,dx\right)^{1/s} \le \frac{2}{|Q'|}\int_{Q'} \sigma\,dx$$

  for some $s > 1$ and all $Q' \subseteq Q$. This implies that $\sigma \in L^s(Q)$. Define $\frac{1}{q} = \frac{1}{q_0} + \frac{1}{sq_0'}$ so that $q > 1$, and let $p = rq > r$. Then

$$\begin{aligned}\left(\int_Q |f|^p\,dx\right)^{1/p} &= \left(\int_Q |f|^{rq} w^{q/q_0} w^{-q/q_0}\,dx\right)^{1/p} \\ &\le \left(\int_Q |f|^{p_0} w\,dx\right)^{1/p_0} \left(\int_Q \sigma^s\,dx\right)^{1/(sq_0')}.\end{aligned}$$

- $\big(\bigcup_{p>r} L^p(Q) \subset \mathcal{M}_{A_1}^r(Q)\big)$: If $f \in L^p(Q)$ for some $p > r$, then Theorem 2.6 implies $|f|^r \le M_Q(|f|^p)^{r/p} \in A_1(\Omega)$.

- $\big(\mathcal{M}^r_{A_1}(Q) \subset \bigcup_{w\in A_{p_0/r}} L^{p_0}_w(Q)\big)$: Set $q_0 = p_0/r > 1$ and suppose we have $g = |f|^r \le w \in A_1(Q)$. Then $w^{1-q_0} \in A_{q_0}(Q)$ by (iv) of Theorem 2.2 and
$$\int_Q |f|^{p_0} w^{1-q_0}\,dx = \int_Q g^{q_0} w^{1-q_0}\,dx \le \int_Q w\,dx < \infty. \qquad \square$$

Next, we extend Theorem 4.1 to $A_\infty$ weights.

**Theorem 4.2.** *Let $Q$ be a cube in $\mathbb{R}^n$ and $p_0$ be an exponent with $0 < p_0 < \infty$. Then*
$$\bigcup_{r>0} \mathcal{M}^r_{A_1}(Q) = \bigcup_{p>0} L^p(Q) = \bigcup_{w\in A_\infty} L^{p_0}_w(Q).$$

*Proof.* We first prove
$$\bigcup_{r>0} \mathcal{M}^r_{A_1}(Q) = \bigcup_{p>0} L^p(Q).$$

- ($\subset$): If $f \in \mathcal{M}^r_{A_1}(Q)$ for some $r > 0$, and $w \in A_1(Q)$ is such that $|f|^r \le w$, then $f \in L^r(Q) \subset \bigcup_{p>0} L^p(Q)$.
- ($\supset$): If $f \in L^p(Q)$ for some $p > 0$, let $r$ be such that $0 < r < p$. Then $|f|^r \le M_Q(|f|^p)^{r/p} \in A_1(Q)$.

Next we show
$$\bigcup_{p>0} L^p(Q) = \bigcup_{w\in A_\infty} L^{p_0}_w(Q).$$

- ($\subset$): Suppose $f \in L^p(Q)$ for some $0 < p < \infty$. Then if $r < \min(p, p_0)$ we have
$$f \in L^p(Q) \subset \bigcup_{r<p} L^p(Q) = \bigcup_{w\in A_{p_0/r}} L^{p_0}_w(Q) \subset \bigcup_{w\in A_\infty} L^{p_0}_w(Q).$$
- ($\supset$): Suppose $f \in L^{p_0}_w(Q)$ for some $w \in A_\infty$. Then $w \in A_q$ for some $q > 1$. Set $p = p_0/q$ and notice that $p < p_0$. Then

$$\int_Q |f|^p\,dx = \int_Q |f|^p w^{1/q} w^{-1/q}\,dx \le \left(\int_Q |f|^{p_0} w\,dx\right)^{1/q} \left(\int_Q w^{1-q'}\,dx\right)^{1/q'}. \quad \square$$

**Example 4.3.** The function

$$f(x) = x^{-1}(\log x)^{-2}\chi_{(0,1/2)}(x) \tag{10}$$

does not belong to $\mathcal{M}_{A_1}([0,1])$. This follows from Theorem 4.1 since it can be readily checked that
$$f \in L^1([0,1]) \backslash \Big(\bigcup_{p>1} L^p([0,1])\Big).$$
However, $f \in \mathcal{M}_F([0,1])$ since $f \in L^1([0,1])$.

**Remark 4.4.** Suppose $0 < p < \infty$. Then

$$L^p(Q) = \bigcup_{w \in A_1} L^p_w(Q).$$

The proof of the equality in Remark 4.4 follows from the fact that $1 \in A_1$ and from inequality (7) with $B = Q$.

We define $H_{A_1}(\mathbb{T})$ as the set of functions in $N^+$ whose boundary function is majorized by an $A_1(\mathbb{T})$ weight. Since we may identify the torus $\mathbb{T}$ with $Q = [0, 2\pi]$, it is obvious that Theorems 4.1 and 4.2 hold for $L^p(\mathbb{T})$ and $L^p_w(\mathbb{T})$ spaces. We have the following analogs for Hardy spaces.

**Theorem 4.5.** *If $p_0$ is an exponent satisfying $1 < p_0 < \infty$, then*

$$H_{A_1}(\mathbb{T}) = \bigcup_{p>1} H^p = \bigcup_{w \in A_{p_0}} H^{p_0}_w.$$

**Theorem 4.6.** *If $p_0$ is an exponent satisfying $0 < p_0 < \infty$, then*

$$\bigcup_{p>0} H^p = \bigcup_{w \in A_\infty} H^{p_0}_w.$$

*Proof of Theorems 4.5 and 4.6.* Since $N^+ \cap L^p(\mathbb{T}) = H^p$ for $p > 0$ [Duren 1970, Theorem 2.11], we see that

$$H_{A_1}(\mathbb{T}) = N^+ \cap \mathcal{M}_{\mathcal{A}_1}(\mathbb{T}) = N^+ \cap \bigcup_{p>1} L^p(\mathbb{T}) = \bigcup_{p>1} H^p.$$

This is the first part of Theorem 4.5.

To go from equality of the analogous $L^p$ spaces to the Hardy spaces is a matter of using two facts for $0 < p_0 < \infty$:

(a) $\int_{\mathbb{T}} \log w \, d\theta > -\infty$ and $w \in L^1(\mathbb{T})$ implies that $w = |h|^{p_0}$ for some outer function $h \in H^{p_0}$.

(b) If $h \in H^{p_0}$ is outer, then the set $h\mathbb{C}[z] = \vee\{z^j h : j \geq 0\}$ is dense in $H^{p_0}$.

Item (a) comes from the standard construction of an outer function [Duren 1970, Section 2.5]. As for item (b), when $1 \leq p_0 < \infty$ this is a standard generalization of Beurling's theorem [Duren 1970, Theorem 7.4]. When $0 < p_0 < 1$, this is a less well known result that can be found in Gamelin [1966, Theorem 4].

For Theorem 4.5 we must show for $1 < p_0 < \infty$ that

$$\bigcup_{p>1} H^p = \bigcup_{w \in A_{p_0}} H^{p_0}_w.$$

Now, for $f \in H^p \subset L^p$, we know there exists $w \in A_{p_0}(\mathbb{T})$ such that $f \in L^{p_0}_w(\mathbb{T})$ by (2). Factor $w = |h|^{p_0}$ with outer $h \in H^{p_0}$. Then, $fh \in N^+ \cap L^{p_0}(\mathbb{T}) = H^{p_0}$

while $h\mathbb{C}[z]$ is dense in $H^{p_0}$ so that there exist polynomials $Q_n$ satisfying

$$\int |fh - Q_n h|^{p_0}\, d\theta = \int |f - Q_n|^{p_0} w\, d\theta \to 0$$

as $n \to \infty$. This shows $f \in H^{p_0}_w$ (since it is initially defined as the closure of the analytic polynomials in $L^{p_0}_w(\mathbb{T})$).

Conversely, we have seen that if $f \in H^{p_0}_w$, then $f \in L^p(\mathbb{T})$ for some $p > 1$. Factor $w = |h|^{p_0}$ as before. Then, $fh \in H^{p_0}$ and $1/h$ is outer, so that $f = fh(1/h) \in N^+$. Since $f \in L^p(\mathbb{T})$, we can then conclude that $f \in H^p$.

The proof of Theorem 4.6, which claims for $0 < p_0 < \infty$ that

$$\bigcup_{p>0} H^p = \bigcup_{w \in A_\infty} H^{p_0}_w,$$

is similar once we know the corresponding fact for $L^p(\mathbb{T})$ spaces. Indeed, take $f \in H^p$ for some $p > 0$. There exists $w \in A_\infty$ such that $f \in L^{p_0}_w(\mathbb{T})$ by Theorem 4.2. Factor $w = |h|^{p_0}$ with outer $h \in H^{p_0}$. Then, $f \in H^{p_0}_w$ as above using Gamelin's result. The converse is similar to the previous proof. □

## 5. The global case

In this section we address the case when our functions are defined on all of $\mathbb{R}^n$. Let us first prove Theorem 1.3, which states that for any $1 < p < \infty$,

$$\bigcup_{w \in A_p} L^p_w(\mathbb{R}^n) = \mathcal{M}_{A_1}(\mathbb{R}^n) \cap \bigcup_{w \in A_1} L^1_w(\mathbb{R}^n).$$

*Proof of Theorem 1.3.* First we show

$$\mathcal{M}_{A_1}(\mathbb{R}^n) \cap \bigcup_{w \in A_1} L^1_w(\mathbb{R}^n) \subset \bigcup_{w \in A_p} L^p_w(\mathbb{R}^n).$$

Suppose $w$ is an $A_1$ majorant of $f$ and $f \in L^1_u(\mathbb{R}^n)$ for some $u \in A_1(\mathbb{R}^n)$. By Theorem 2.2, $uw^{1-p} \in A_p(\mathbb{R}^n)$ and

$$\int_{\mathbb{R}^n} |f|^p w^{1-p} u\, dx \le \int_{\mathbb{R}^n} |f| u\, dx.$$

To see the reverse containment suppose that $f \not\equiv 0$ belongs to $L^p_w(\mathbb{R}^n)$ for some $w \in A_p(\mathbb{R}^n)$. We will use the fact that $w \in A_p(\mathbb{R}^n)$ implies $M \in \mathcal{B}(L^p_w)$ to apply the Rubio de Francia algorithm:

$$Rf = \sum_{k=0}^{\infty} \frac{M^k f}{2^k \|M\|^k_{\mathcal{B}(L^p_w)}}.$$

Then $Rf$ is an $A_1$ majorant of $f$ so $f \in \mathcal{M}_{A_1}(\mathbb{R}^n)$. Also let $g$ be any function in $L^{p'}_\sigma(\mathbb{R}^n)$ where $\sigma = w^{1-p'}$ satisfying $\|g\|_{L^{p'}_\sigma(\mathbb{R}^n)} = 1$. Again, since $\sigma \in A_{p'}(\mathbb{R}^n)$, we apply the Rubio de Francia algorithm

$$Rg=\sum_{k=0}^{\infty}\frac{M^k g}{2^k\|M\|^k_{\mathscr{B}(L^{p'}_\sigma)}},$$

so that $Rg$ is in $A_1(\mathbb{R}^n)$ and $\|Rg\|_{L^{p'}_\sigma(\mathbb{R}^n)}\le 2$. Hence

$$\int_{\mathbb{R}^n}|f|Rg\,dx=\int_{\mathbb{R}^n}|f|w^{1/p}Rgw^{-1/p}\,dx\le\|f\|_{L^p_w(\mathbb{R}^n)}\|Rg\|_{L^{p'}_\sigma(\mathbb{R}^n)}\le 2\|f\|_{L^p_w(\mathbb{R}^n)},$$

showing that $f\in\bigcup_{w\in A_1}L^1_w(\mathbb{R}^n)$ as well. □

Before moving on, we remark that the intersection of $\mathcal{M}_{A_1}(\mathbb{R}^n)$ and $\bigcup_{w\in A_1}L^1_w(\mathbb{R}^n)$ is necessary for the result on $\mathbb{R}^n$. We did not encounter this phenomenon in the local case since for a fixed cube, $\mathcal{M}_{A_1}(Q)\subset L^1(Q)$. To see that the intersection is necessary, notice that the function in Example 4.3 viewed as a function on $\mathbb{R}$ belongs to $L^1(\mathbb{R})\subset\bigcup_{w\in A_1}L^1_w(\mathbb{R})$, but does not belong to $L^p_w(\mathbb{R})$ for any $p>1$ and $w\in A_p(\mathbb{R})$ since it is not in $L^p_{\mathrm{loc}}(\mathbb{R})$ for any $p>1$. Theorem 1.3 shows that for $1<p<\infty$,

$$\bigcup_{w\in A_p}L^p_w(\mathbb{R}^n)\subset\mathcal{M}_{A_1}(\mathbb{R}^n).$$

Below we will show this containment is proper (see Example 5.2).

We now prove Theorem 1.4.

*Proof of Theorem 1.4.* By Theorem 1.3 it suffices to show

$$\bigcup_{w\in A_p}L^p_w(\mathbb{R}^n)\subset\bigcup\{\mathscr{X}:M\in\mathscr{B}(\mathscr{X})\cap\mathscr{B}(\mathscr{X}')\}\tag{11}$$

and

$$\bigcup\{\mathscr{X}:M\in\mathscr{B}(\mathscr{X})\cap\mathscr{B}(\mathscr{X}')\}\subset\mathcal{M}_{A_1}(\mathbb{R}^n)\cap\bigcup_{w\in A_1}L^1_w(\mathbb{R}^n).\tag{12}$$

However, the containment (11) is immediate, since

$$M\in\mathscr{B}(L^p_w(\mathbb{R}^n))\Leftrightarrow w\in A_p(\mathbb{R}^n)\Leftrightarrow\sigma\in A_{p'}(\mathbb{R}^n)\Leftrightarrow M\in\mathscr{B}(L^{p'}_\sigma(\mathbb{R}^n)).$$

On the other hand, for containment (12), if $f\not\equiv 0$, then $f\in\mathscr{X}$ for some Banach function space $\mathscr{X}$ such that $M\in\mathscr{B}(\mathscr{X})\cap\mathscr{B}(\mathscr{X}')$. Then we may use the Rubio de Francia algorithm to construct an $A_1(\mathbb{R}^n)$ majorant:

$$Rf=\sum_{k=0}^{\infty}\frac{M^k f}{2^k\|M\|^k_{\mathscr{B}(\mathscr{X})}}.$$

Then $Rf \in A_1$ and $|f| \le Rf$, so $f \in \mathcal{M}_{A_1}(\mathbb{R}^n)$. Given $g \in \mathcal{X}'$ let

$$Rg = \sum_{k=0}^{\infty} \frac{M^k g}{2^k \|M\|^k_{\mathcal{B}(\mathcal{X}')}},$$

so that $Rg \in A_1(\mathbb{R}^n) \cap \mathcal{X}'$ and $\|Rg\|_{\mathcal{X}'} \le 2\|g\|_{\mathcal{X}'}$. Then

$$\int_{\mathbb{R}^n} |f| Rg \, dx \le \|f\|_{\mathcal{X}} \|Rg\|_{\mathcal{X}'} \le 2\|f\|_{\mathcal{X}} \|g\|_{\mathcal{X}'},$$

which yields $f \in \bigcup_{w \in A_1} L^1_w(\mathbb{R}^n)$. □

When $p > 1$, $L^{p,\infty}(\mathbb{R}^n)$ is a Banach function space on which $M$ is bounded (see [Grafakos 2008]), and likewise, its associate $(L^{p,\infty}(\mathbb{R}^n))' = L^{p',1}(\mathbb{R}^n)$, the Lorentz space with exponents $p'$ and 1, is also a Banach function space on which $M$ is bounded (see [Ariño and Muckenhoupt 1990]).

**Corollary 5.1.** *Suppose* $1 < p_0 < \infty$. *Then*

$$\bigcup_{p>1} L^{p,\infty}(\mathbb{R}^n) \subset \bigcup_{w \in A_{p_0}} L^{p_0}_w(\mathbb{R}^n).$$

From Corollary 5.1 we see that the analogous version of the equivalences in (2) are not true on $\mathbb{R}^n$. This follows since

$$\bigcup_{p>1} L^p(\mathbb{R}^n) \subsetneqq \bigcup_{p>1} L^{p,\infty}(\mathbb{R}^n).$$

For example, $f(x) = |x|^{-n/2} \in L^{2,\infty}(\mathbb{R}^n)$ but $f \notin \bigcup_{p>0} L^p(\mathbb{R}^n)$.

We also remark that the techniques required for $\mathbb{R}^n$ are completely different than the local case. For example, to prove the containment

$$\bigcup_{p>1} L^{p,\infty}(\mathbb{R}^n) \subset \mathcal{M}_{A_1}(\mathbb{R}^n)$$

it is not enough to simply dominate $|f|$ by $M(|f|^p)^{1/p}$. However, for $f \in L^{p,\infty}(\mathbb{R}^n)$, $M(|f|^p)$ may not be finite (take $f(x) = |x|^{-n/p}$, in which case $M(|f|^p) \equiv \infty$). Instead we must refine our construction of an $A_1$ majorant using the techniques of Rubio de Francia [1984].

We now provide examples to show that the inclusions in (4) are proper. We first show that the second inclusion is proper, i.e.,

$$\bigcup_{w \in A_p} L^p_w(\mathbb{R}^n) \subsetneqq \mathcal{M}_{A_1}(\mathbb{R}^n).$$

Since

$$\bigcup_{w \in A_p} L^p_w(\mathbb{R}^n) = \mathcal{M}_{A_1}(\mathbb{R}^n) \cap \bigcup_{w \in A_1} L^1_w(\mathbb{R}^n),$$

it suffices to find a function in $\mathcal{M}_{A_1}(\mathbb{R}^n) \setminus \big(\bigcup_{w \in A_1} L^1_w(\mathbb{R}^n)\big)$.

**Example 5.2.** The function $f(x) = 1$ belongs to $\mathcal{M}_{A_1}(\mathbb{R}^n) \setminus \bigcup_{w\in A_1} L^1_w(\mathbb{R}^n)$. To prove this we need the fact that if $w \in A_\infty$ then $w \notin L^1(\mathbb{R}^n)$. One way to see this (pointed out by the referee) is to notice that $A_\infty$ weights are doubling, and doubling measures have infinite total mass. We can also give an ad hoc argument using the reverse Hölder inequality. If $w$ satisfies

$$\left(\frac{1}{|Q|}\int_Q w^s\,dx\right)^{1/s} \le \frac{2}{|Q|}\int_Q w\,dx$$

for some $s > 1$ and all cubes $Q$, then by taking $Q_N = [-N, N]^n$, we have

$$\left(\frac{1}{|Q_N|}\int_{Q_1} w^s\,dx\right)^{1/s} \le \left(\frac{1}{|Q_N|}\int_{Q_N} w^s\,dx\right)^{1/s} \le \frac{2}{|Q_N|}\int_{Q_N} w\,dx \le \frac{2}{|Q_N|}\|w\|_{L^1(\mathbb{R}^n)}.$$

Letting $N \to \infty$ we arrive at a contradiction. Finally, to see $1 \notin \bigcup_{w\in A_1} L^1_w(\mathbb{R}^n)$, notice that $1 \in L^1_w(\mathbb{R}^n)$ if and only if $w \in L^1(\mathbb{R}^n)$.

Next we show that

$$\bigcup_{p>1} L^{p,\infty}(\mathbb{R}^n) \subsetneq \bigcup_{w\in A_p} L^p_w(\mathbb{R}^n).$$

For this example we need the following lemma.

**Lemma 5.3.** *Suppose $u, v \in A_1(\mathbb{R}^n)$. Then*

$$\max(u, v) \in A_1(\mathbb{R}^n) \quad \textit{and} \quad \min(u, v) \in A_1(\mathbb{R}^n).$$

*Proof.* To see that $\max(u, v)$ is in $A_1(\mathbb{R}^n)$ note that $\max(u, v) \le u+v \le 2\max(u, v)$, and hence

$$M(\max(u, v)) \le Mu + Mv \le C(u + v) \le 2C \max(u, v).$$

To prove $\min(u, v) \in A_1(\mathbb{R}^n)$ we use the equivalent definition of $A_1(\mathbb{R}^n)$:

$$w \in A_1(\mathbb{R}^n) \Leftrightarrow \frac{1}{|Q|}\int_Q w\,dx \le C \inf_Q w \quad \forall Q \subset \mathbb{R}^n$$

where the infimum is the essential infimum of $w$ over the cube $Q$. Set $w = \min(u, v)$ and let $Q$ be a cube. Notice that $\inf_Q u > \inf_Q v$ implies $\inf_Q w = \inf_Q v$ and hence

$$\frac{1}{|Q|}\int_Q w\,dx \le \frac{1}{|Q|}\int_Q v\,dx \le C \inf_Q v = C \inf_Q w.$$

On the other hand, if $\inf_Q u \le \inf_Q v$ then $\inf_Q w = \inf_Q u$ and so

$$\frac{1}{|Q|}\int_Q w\,dx \le \frac{1}{|Q|}\int_Q u\,dx \le C \inf_Q u = C \inf_Q w.$$

So $w \in A_1(\mathbb{R}^n)$. □

**Example 5.4.** Consider $f(x) = \max(|x|^{-\alpha n}, |x|^{-\beta n})$. If $0 < \alpha < \beta < 1$ then $f \notin \bigcup_{p>0} L^{p,\infty}(\mathbb{R}^n)$. However,

$$|f(x)| \le w(x)$$

where $w(x) = \max(|x|^{-\beta n}, 1)$, and $f \in L^1_u(\mathbb{R}^n)$ where $u(x) = \min(|x|^{-\gamma n}, 1)$ when $1 - \alpha < \gamma < 1$. By Lemma 5.3, both $u$ and $w$ belong to $A_1(\mathbb{R}^n)$. Thus

$$f \in \mathcal{M}_{A_1}(\mathbb{R}^n) \cap \bigcup_{w \in A_1} L^1_w(\mathbb{R}^n) = \bigcup_{w \in A_p} L^p_w(\mathbb{R}^n).$$

Finally, we end with brief descriptions of $\bigcup_{w\in A_1} L^1_w(\mathbb{R}^n)$ and $\mathcal{M}_{A_1}(\mathbb{R}^n)$ in terms of Banach function spaces.

**Theorem 5.5.** $\displaystyle\bigcup_{w\in A_1} L^1_w(\mathbb{R}^n) = \bigcup \{\mathscr{X} : M \in \mathscr{B}(\mathscr{X}')\} = \bigcup \{\mathscr{X} : \mathscr{X}' \cap A_1(\mathbb{R}^n) \neq \varnothing\}.$

*Proof.* It is clear that

$$\bigcup_{w\in A_1} L^1_w(\mathbb{R}^n) \subset \bigcup \{\mathscr{X} : M \in \mathscr{B}(\mathscr{X}')\},$$

since the associate space of $L^1_w(\mathbb{R}^n)$ is

$$L^\infty_w(\mathbb{R}^n) = \{f : f/w \in L^\infty\}$$

with norm $\|f\|_{L^\infty_w} = \|f/w\|_{L^\infty}$. For any cube $Q$,

$$\frac{1}{|Q|}\int_Q |f|\,dx \le \|f/w\|_{L^\infty} \frac{1}{|Q|}\int_Q w\,dx.$$

Hence if $w \in A_1$, then

$$Mf \le \|f\|_{L^\infty_w} Mw \le C\|f\|_{L^\infty_w} w,$$

and dividing through by $w$ we obtain $M \in \mathscr{B}(L^\infty_w)$.

The associate space is always a closed subspace of the dual space [Bennett and Sharpley 1988; Rubio de Francia 1987]. Suppose $\mathscr{X}$ is such that $M \in \mathscr{B}(\mathscr{X}')$. Given $g \in \mathscr{X}'$ with $g \not\equiv 0$ (notice Banach function spaces always contain nonzero functions by property (iv) of Banach function norms), let

$$w = \sum_{k=1}^{\infty} \frac{M^k g}{2^k \|M\|^k_{\mathscr{B}(\mathscr{X}')}}$$

so that $w \in A_1(\mathbb{R}^n)$ and $\|w\|_{\mathscr{X}'} \le \|g\|_{\mathscr{X}'}$. Thus $w \in \mathscr{X}' \cap A_1(\mathbb{R}^n)$, showing that

$$M \in \mathscr{B}(\mathscr{X}') \Rightarrow \mathscr{X} \cap A_1(\mathbb{R}^n) \neq \varnothing.$$

Finally, suppose $f \in \mathscr{X}$ for some $\mathscr{X}$ such that $\mathscr{X}'$ contains an $A_1$ weight. Let $w \in \mathscr{X}' \cap A_1(\mathbb{R}^n)$. Then

$$\int_{\mathbb{R}^n} |f| w \, dx \le \|f\|_{\mathscr{X}} \|w\|_{\mathscr{X}'},$$

so that $f \in L^1_w(\mathbb{R}^n)$. □

Finally we refer to a result of Chu [2013] which gives the final characterization of $\mathcal{M}_{A_1}(\mathbb{R}^n)$.

**Theorem 5.6** [Chu 2013]. $\mathcal{M}_{A_1}(\mathbb{R}^n) = \bigcup \{\mathscr{X} : M \in \mathscr{B}(\mathscr{X})\}$.

## 6. Questions

We leave the reader with some open questions.

**1.** Let $A_p^* = \bigcap_{q>p} A_q$. Is there a characterization of the union

$$\bigcup_{w \in A_p^*} L^p_w?$$

In general $A_p \subsetneqq A_p^*$. For example $w(x) = \max((\log |x|^{-1})^{-1}, 1)$ belongs to $A_1^*$ but not $A_1$. Moreover,

$$\{w : w, 1/w \in A_1^*\} = \mathrm{clos}_{BMO} L^\infty$$

(see [García-Cuerva and Rubio de Francia 1985; Johnson and Neugebauer 1987]). In the local case we have

$$\bigcup_{w \in A_p^*} L^p_w(Q) \subset \bigcap_{s<p} \bigcup_{r>s} L^r(Q) = \limsup_{r \to p^-} L^r(Q).$$

Are these two sets equal?

**2.** It is well known that

$$L^1 \cap L^\infty \subset \bigcap_{1<p<\infty} L^p \subset \bigcup_{1<p<\infty} L^p \subset L^1 + L^\infty.$$

When can we write a function as the sum of a function in $\mathcal{M}_{A_1}$ and $\bigcup_{w \in A_1} L^1_w$? That is, what conditions on a function guarantee it belongs to $\mathcal{M}_{A_1} + \bigcup_{w \in A_1} L^1_w$?

**3.** What can one say about

$$\bigcup_{w \in A_p} L^{p,\infty}_w?$$

If $w \in A_1$ and $p > 1$ then $M \in \mathscr{B}(L^{p,\infty}_w)$, so for $p > 1$,

$$\bigcup_{w \in A_1} L^{p,\infty}_w \subset \mathcal{M}_{A_1}.$$

**4.** Do these results transfer to more general domains? It is possible to consider a general open set $\Omega$ as our domain of interest. We may define the $A_p(\Omega)$ classes, $\mathcal{M}_{A_1}(\Omega)$, and the Hardy–Littlewood maximal operator $M_\Omega$ exactly as before. However, the openness results, Theorems 2.3 and 2.4, may not hold for $\Omega$, even if it is bounded [Cruz-Uribe et al. 2011]. In the local case we assume that weights belong to $L^1(\Omega)$. What happens if we only assume $L^1_{\mathrm{loc}}(\Omega)$?


## Acknowledgements

We would like to thank Carlos Pérez and Javier Duoandikoetxea for comments that helped improve the quality of this article. We would like to thank David Cruz-Uribe for pointing out Example 4.3. We would also like to offer many thanks to the referee for insightful suggestions and detailed corrections.


## References


[Ariño and Muckenhoupt 1990] M. A. Ariño and B. Muckenhoupt, "Maximal functions on classical Lorentz spaces and Hardy's inequality with weights for nonincreasing functions", *Trans. Amer. Math. Soc.* **320**:2 (1990), 727–735. MR 90k:42034 Zbl 0716.42016

[Bennett and Sharpley 1988] C. Bennett and R. Sharpley, *Interpolation of operators*, Pure and Applied Mathematics **129**, Academic Press, Boston, 1988. MR 89e:46001 Zbl 0647.46057

[Chu 2013] C. Chu, "A characterization of $\mathcal{M}_{A_1}(\mathbb{R}^n)$", preprint, 2013. arXiv 1310.2136

[Coifman and Rochberg 1980] R. R. Coifman and R. Rochberg, "Another characterization of BMO", *Proc. Amer. Math. Soc.* **79**:2 (1980), 249–254. MR 81b:42067 Zbl 0432.42016

[Cruz-Uribe et al. 2011] D. V. Cruz-Uribe, J. M. Martell, and C. Pérez, *Weights, extrapolation and the theory of Rubio de Francia*, Operator Theory: Advances and Applications **215**, Birkhäuser, Basel, 2011. MR 2012f:42001 Zbl 1234.46003

[Duoandikoetxea 2001] J. Duoandikoetxea, *Fourier analysis*, Graduate Studies in Mathematics **29**, American Mathematical Society, Providence, RI, 2001. MR 2001k:42001 Zbl 0969.42001

[Duren 1970] P. L. Duren, *Theory of $H^p$ spaces*, Pure and Applied Mathematics **38**, Academic Press, New York, 1970. MR 42 #3552 Zbl 0215.20203

[Gamelin 1966] T. W. Gamelin, "$H^p$ spaces and extremal functions in $H^1$", *Trans. Amer. Math. Soc.* **124** (1966), 158–167. MR 35 #4731 Zbl 0146.37202

[García-Cuerva and Rubio de Francia 1985] J. García-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies **116**, North-Holland, Amsterdam, 1985. MR 87d:42023 Zbl 0578.46046

[Grafakos 2008] L. Grafakos, *Classical Fourier analysis*, 2nd ed., Graduate Texts in Mathematics **249**, Springer, New York, 2008. MR 2011c:42001 Zbl 1220.42001

[Hytönen et al. 2012] T. P. Hytönen, C. Pérez, and E. Rela, "Sharp reverse Hölder property for $A_\infty$ weights on spaces of homogeneous type", *J. Funct. Anal.* **263**:12 (2012), 3883–3899. MR 2990061 Zbl 1266.42045

[Johnson and Neugebauer 1987] R. Johnson and C. J. Neugebauer, "Homeomorphisms preserving $A_p$", *Rev. Mat. Iberoamericana* **3**:2 (1987), 249–273. MR 90d:42013 Zbl 0677.42019

[Jones 1980] P. W. Jones, "Factorization of $A_p$ weights", *Ann. of Math.* (2) **111**:3 (1980), 511–530. MR 82b:46035 Zbl 0493.42030

[Lerner 2010] A. K. Lerner, "Some remarks on the Fefferman–Stein inequality", *J. Anal. Math.* **112** (2010), 329–349. MR 2011m:42032 Zbl 1214.42031

[McCarthy 1990] J. E. McCarthy, "Common range of co-analytic Toeplitz operators", *J. Amer. Math. Soc.* **3**:4 (1990), 793–799. MR 91f:47041 Zbl 0733.47028

[Rubio de Francia 1982] J. L. Rubio de Francia, "Factorization and extrapolation of weights", *Bull. Amer. Math. Soc.* (*N.S.*) **7**:2 (1982), 393–395. MR 83i:42016 Zbl 0491.42020

[Rubio de Francia 1984] J. L. Rubio de Francia, "Factorization theory and $A_p$ weights", *Amer. J. Math.* **106**:3 (1984), 533–547. MR 86a:47028a Zbl 0558.42012

[Rubio de Francia 1987] J. L. Rubio de Francia, "Linear operators in Banach lattices and weighted $L^2$ inequalities", *Math. Nachr.* **133** (1987), 197–209. MR 89d:47070 Zbl 0648.47029

[Rudin 1964] W. Rudin, *Principles of mathematical analysis*, 2nd ed., McGraw-Hill, New York, 1964. MR 29 #3587 Zbl 0148.02903

[Rutsky 2011] D. V. Rutsky, "BMO-регулярность в решетках измеримых функций на пространствах однородного типа", *Algebra i Analiz* **23**:2 (2011), 248–295. Translated as "BMO-regularity in lattices of measurable functions on spaces of homogeneous type" in *St. Petersburg Math. J.* **23**:2 (2012), 381–412. MR 2012m:46027 Zbl 1245.46015

[Rutsky 2014] D. V. Rutsky, "$\mathrm{A}_1$-regularity and boundedness of Calderón–Zygmund operators", *Studia Math.* **221**:3 (2014), 231–247. MR 3208299 Zbl 1316.46020

Received August 22, 2014. Revised January 9, 2015.

GREG KNESE
DEPARTMENT OF MATHEMATICS
WASHINGTON UNIVERSITY IN ST. LOUIS
ONE BROOKINGS DRIVE
ST. LOUIS, MO 63130
UNITED STATES
geknese@math.wustl.edu

JOHN E. MCCARTHY
DEPARTMENT OF MATHEMATICS
WASHINGTON UNIVERSITY IN ST. LOUIS
ONE BROOKINGS DRIVE
ST. LOUIS, MO 63130
UNITED STATES
mccarthy@wustl.edu

KABE MOEN
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ALABAMA
TUSCALOOSA, AL 35487
UNITED STATES
kmoen@as.ua.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.433

# COMPLEX HYPERBOLIC (3, 3, $n$) TRIANGLE GROUPS

JOHN R. PARKER, JIEYAN WANG AND BAOHUA XIE

**Let $p, q, r$ be positive integers. Complex hyperbolic $(p, q, r)$ triangle groups are representations of the hyperbolic $(p, q, r)$ reflection triangle group to the holomorphic isometry group of complex hyperbolic space $H^2_{\mathbb{C}}$, where the generators fix complex lines. In this paper, we obtain all the discrete and faithful complex hyperbolic $(3, 3, n)$ triangle groups for $n \geq 4$. Our result solves a conjecture of Schwartz in the case when $p = q = 3$.**

## 1. Introduction

An abstract $(p, q, r)$ reflection triangle group for positive integers $p, q, r$ is the group

$$\Delta_{p,q,r} = \langle \sigma_1, \sigma_2, \sigma_3 \mid \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = (\sigma_2\sigma_3)^p = (\sigma_3\sigma_1)^q = (\sigma_1\sigma_2)^r = \mathrm{id} \rangle.$$

We sometimes take (at least) one of $p, q, r$ to be $\infty$, in which case the corresponding relation does not appear.

It is interesting to seek geometrical representations of $\Delta_{p,q,r}$. An extremely well-known fact is that $\Delta_{p,q,r}$ may be realised geometrically as the reflections in the side of a geodesic triangle with internal angles $\pi/p$, $\pi/q$, $\pi/r$. Furthermore, if $1/p + 1/q + 1/r > 1$, $= 1$ or $< 1$ then this triangle is spherical, Euclidean or hyperbolic respectively. Moreover, up to isometries (or similarities in the Euclidean case) there is a unique such triangle and the representation is rigid. In the case where (at least) one of $p, q, r$ is $\infty$, we omit the relevant term from $1/p+1/q+1/r$ and we insist that the sides of the triangle are asymptotic. Thus the $(\infty, \infty, \infty)$ triangle is a triangle in the hyperbolic plane with all three vertices on the boundary.

In contrast, if we choose a geometrical representation of $\Delta_{p,q,r}$ in a space of nonconstant curvature then more interesting things can happen; see, for example, [Brehm 1990]. In this paper, we consider representations of $\Delta_{p,q,r}$ to SU(2, 1), which is (a triple cover of) the group of holomorphic isometries of complex hyperbolic space $H^2_{\mathbb{C}}$. A convenient model of $H^2_{\mathbb{C}}$ is the unit ball in $\mathbb{C}^2$ with the

---

Wang and Xie were supported by NSFC (grant numbers 11071059 and 11371126). Xie was also supported by NSFC (grant number 11201134) and the Young Teachers Support program of Hunan University.
*MSC2010:* primary 20H10; secondary 22E40, 51M10.
*Keywords:* complex hyperbolic geometry, complex hyperbolic triangle groups.

Bergman metric, having constant holomorphic sectional curvature and 1/4-pinched real sectional curvatures.

A *complex hyperbolic triangle group* will be a representation of $\Delta_{p,q,r}$ to SU(2, 1) where the generators fix complex lines. Note we could have made other choices. For example, we could choose the generators to be antiholomorphic isometries, or we could choose reflections in three complex lines but with higher order. These choices lead to interesting results, but we will not consider them here. A crucial observation is that when $\min\{p, q, r\} \geq 3$, there is a one (real) dimensional representation space of complex hyperbolic triangle groups with $1/p+1/q+1/r < 1$ (either make a simple dimension count or see [Brehm 1990] for example). This means that the representation is determined up to conjugacy by $p$, $q$, $r$ and one extra variable. This variable is determined by certain traces; see, for example, [Pratoussevitch 2005].

In order to state our main results, we need a little terminology. Elements of SU(2, 1) act on complex hyperbolic space $\boldsymbol{H}^2_{\mathbb{C}}$ and its boundary (see below). An element $A \in$ SU(2, 1) is called *loxodromic* if it fixes two points, both of which lie on $\partial\boldsymbol{H}^2_{\mathbb{C}}$; *parabolic* if it fixes exactly one point, and this point lies on $\partial\boldsymbol{H}^2_{\mathbb{C}}$; *elliptic* if it fixes at least one point of $\boldsymbol{H}^2_{\mathbb{C}}$. Discrete groups cannot contain elliptic elements of infinite order. Therefore in a representation of an abstract group to SU(2, 1), if an element of infinite order in the abstract group is represented by an elliptic map then the representation is not discrete or not faithful (or both); compare with [Goldman and Parker 1992].

Complex hyperbolic triangle groups have a rich history; see Schwartz's ICM survey [2002] for an overview. In particular, he presented the following conjectural picture:

**Conjecture 1.1** [Schwartz 2002]. *Let $\Delta_{p,q,r}$ be a triangle group with $p \leq q \leq r$. Then any complex hyperbolic representation $\Gamma$ of $\Delta_{p,q,r}$ is discrete and faithful if and only if $W_A = I_1I_3I_2I_3$ and $W_B = I_1I_2I_3$ are not elliptic. Furthermore*:

(i) *If $p < 10$ then $\Gamma$ is discrete and faithful if and only if $W_A = I_1I_3I_2I_3$ is nonelliptic.*

(ii) *If $p > 13$ then $\Gamma$ is discrete and faithful if and only if $W_B = I_1I_2I_3$ is nonelliptic.*

The initial step towards solving this conjecture is the following result of Grossi.

**Proposition 1.2** [Grossi 2007]. *Let $\Delta_{p,q,r}$ be a triangle group with $p \leq q \leq r$. Define $W_A = I_1I_3I_2I_3$ and $W_B = I_1I_2I_3$. Then for complex hyperbolic representations of $\Delta_{p,q,r}$*:

(i) *If $p < 10$ and $W_A = I_1I_3I_2I_3$ is nonelliptic then $W_B$ is nonelliptic.*

(ii) *If $p > 13$ and $W_B = I_1I_2I_3$ is nonelliptic then $W_A$ is nonelliptic.*

A motivating example, initially considered by Goldman and Parker [1992] and completed by Schwartz [2001b; 2005], concerns complex hyperbolic ideal triangle groups, that is, representations of $\Delta_{\infty,\infty,\infty}$. This result may be summarised as follows:

**Theorem 1.3** [Goldman and Parker 1992; Schwartz 2001b; 2005]. *Let $\Gamma = \langle I_1, I_2, I_3\rangle$ be a complex hyperbolic $(\infty, \infty, \infty)$ triangle group. Then $\Gamma$ is a discrete and faithful representation of $\Delta_{\infty,\infty,\infty}$ if and only if $I_1 I_2 I_3$ is nonelliptic.*

Note that this gives a complete solution to Schwartz's conjecture in the case $p = q = r = \infty$. Furthermore, Schwartz [2001a] gives an elegant description of the group where $I_1 I_2 I_3$ is parabolic.

**Theorem 1.4** [Schwartz 2001a]. *Let $\Gamma = \langle I_1, I_2, I_3\rangle$ be the $(\infty, \infty, \infty)$ complex hyperbolic triangle group for which $I_1 I_2 I_3$ is parabolic. Let $\Gamma_2 = \langle I_1 I_2, I_1 I_3\rangle$ be the index-2 subgroup of $\Gamma$ with no complex reflections. Then $\boldsymbol{H}^2_{\mathbb{C}}/\,\Gamma_2$ is a complex hyperbolic orbifold whose boundary is a triple cover of the Whitehead link complement.*

Schwartz [2007] proves his conjecture for $\min\{p, q, r\}$ sufficiently large (but unfortunately with no effective bound on this minimum).

**Theorem 1.5** [Schwartz 2007]. *Let $\Gamma = \langle I_1, I_2, I_3\rangle$ be a complex hyperbolic $(p, q, r)$ triangle group with $p \le q \le r$. If $p$ is sufficiently large, then $\Gamma$ is a discrete and faithful representation of $\Delta_{p,q,r}$ if and only if $I_1 I_2 I_3$ is nonelliptic.*

Our main result solves Schwartz's conjecture in the case when $p = q = 3$.

**Theorem 1.6.** *Let $n$ be an integer at least 4. Let $\Gamma = \langle I_1, I_2, I_3\rangle$ be a complex hyperbolic $(3, 3, n)$ triangle group. Then $\Gamma$ is a discrete and faithful representation of $\Delta_{3,3,n}$ if and only if $I_1 I_3 I_2 I_3$ is nonelliptic.*

Note that the "only if" part is a consequence of our observation about elliptic elements above. The "if" part will follow from Corollary 4.4 below.

For the representation where $I_1 I_3 I_2 I_3$ is parabolic, when $n = 4$ and 5 we have the following description of the quotient orbifold from the census of Falbel, Koseleff and Rouillier [Falbel et al. 2015]. The case $n = 4$ combines work of Deraux, Falbel and Wang [Deraux and Falbel 2015; Falbel and Wang 2014]. The cleanest statement may be found in [Deraux 2015, Theorem 4.2], which also treats the case $n = 5$.

**Theorem 1.7** [Deraux 2015, Theorem 4.2]. (i) *Let $\Gamma = \langle I_1, I_2, I_3\rangle$ be the complex hyperbolic $(3, 3, 4)$ triangle group for which $I_1 I_3 I_2 I_3$ is parabolic. Let $\Gamma_2 = \langle I_1 I_2, I_1 I_3\rangle$ be the index-2 subgroup of $\Gamma$ with no complex reflections. Then $\Gamma_2$ is conjugate to both $\rho_{1-1}(\pi_1(M_4))$ and $\rho_{4-1}(\pi_1(M_4))$ from [Falbel et al. 2015]. In particular, $\boldsymbol{H}^2_{\mathbb{C}}/\,\Gamma_2$ is a complex hyperbolic orbifold whose boundary is the figure eight knot complement.*

(ii) *Let* $\Gamma = \langle I_1, I_2, I_3\rangle$ *be the complex hyperbolic* $(3, 3, 5)$ *triangle group for which* $I_1 I_3 I_2 I_3$ *is parabolic. Let* $\Gamma_2 = \langle I_1 I_2, I_1 I_3\rangle$ *be the index-*$2$ *subgroup of* $\Gamma$ *with no complex reflections. Then* $\Gamma_2$ *is conjugate to both* $\rho_{4-3}(\pi_1(M_9))$ *and* $\rho_{3-3}(\pi_1(M_{15}))$ *from [Falbel et al. 2015].*

It should be possible to give a similar description of the other complex hyperbolic $(3, 3, n)$ triangle groups for which $I_1 I_3 I_2 I_3$ is parabolic.

Note that Theorem 1.6 holds in the case $n = \infty$. This follows from recent work of Parker and Will [2015b] (see also [Parker and Will 2015a]). Furthermore, if as above $\Gamma_2 = \langle I_1 I_2, I_1 I_3\rangle$ is the index-2 subgroup of representation of the $(3, 3, \infty)$ triangle group for which $I_1 I_3 I_2 I_3$ is parabolic, then $\boldsymbol{H}^2_{\mathbb{C}}/\Gamma_2$ is a complex hyperbolic orbifold whose boundary is the Whitehead link complement. This is one of the representations in [Falbel et al. 2015].

Finally, we note some further interesting groups in this family.

**Theorem 1.8** [Thompson 2010]. *The complex hyperbolic* $(3, 3, 4)$ *triangle group with* $I_1 I_3 I_2 I_3$ *of order* $7$ *and the complex hyperbolic* $(3, 3, 5)$ *triangle group with* $I_1 I_3 I_2 I_3$ *of order* $5$ *are both lattices.*

Our method of proof will be to construct a Dirichlet domain based at the fixed point of the order-$n$ elliptic map $I_1 I_2$. Since this point has nontrivial stabiliser, this domain is not a fundamental domain for $\Gamma$, but it is a fundamental domain for the coset space of the stabiliser of this point in $\Gamma$. Of course, in order to prove directly that this is a Dirichlet domain, we would have to check infinitely many inequalities. Instead, we construct a candidate Dirichlet domain and then use the Poincaré polyhedron theorem for coset decompositions (see [Mostow 1980, Theorem 6.3.2] or [Deraux et al. 2015, Theorem 3.2], for example).

In the case of a Fuchsian $(3, 3, n)$ triangle group acting on the hyperbolic plane, a fundamental domain is a hyperbolic triangle with internal angles $\pi/3$, $\pi/3$ and $\pi/n$. The Dirichlet domain with centre the fixed point of an order-$n$ elliptic map is a regular hyperbolic $2n$-gon with internal angles $2\pi/3$. This $2n$-gon is made up of $2n$ copies of the triangular fundamental domain for the $(3, 3, n)$ group; see Figure 1. The stabiliser of the order-$n$ fixed point, which is a dihedral group of order $2n$, fixes the $2n$-gon and permutes the triangles.

For the complex hyperbolic $(3, 3, n)$ triangle groups, we will see that the combinatorial structure of the Dirichlet domain $D$ is the same as that in the Fuchsian case. Namely, $D$ has $2n$ sides, each of which is contained in a bisector. Each side meets exactly two other sides (in the case where $I_1 I_3 I_2 I_3$ is parabolic, there are some additional tangencies between sides on the ideal boundary). The sides are permuted by the dihedral group $\langle I_1, I_2\rangle$.

In Section 2 we give the necessary background on complex hyperbolic geometry and the Poincaré polyhedron theorem. In Section 3 we normalise the generators of $\Gamma$ and discus the parameters this involves. Finally, in Section 4 we consider the bisectors and their intersection properties. This is the heart of the paper.

## 2. Background

***Complex hyperbolic space.*** Let $\mathbb{C}^{2,1}$ be the three-dimensional complex vector space equipped with a Hermitian form $H$ of signature $(2, 1)$. In this paper we consider the diagonal Hermitian form $H = \mathrm{diag}(1, 1, -1)$. Thus if $\boldsymbol{u} = (u_1, u_2, u_3)^t$ and $\boldsymbol{v} = (v_1, v_2, v_3)^t$ then the Hermitian form is given by

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \boldsymbol{v}^* H \boldsymbol{u} = u_1\bar{v}_1 + u_2\bar{v}_2 - u_3\bar{v}_3.$$

Define

$$V_- = \{\boldsymbol{v} \in \mathbb{C}^{2,1} : \langle \boldsymbol{v}, \boldsymbol{v}\rangle < 0\}, \quad V_0 = \{\boldsymbol{v} \in \mathbb{C}^{2,1} - \{0\} : \langle \boldsymbol{v}, \boldsymbol{v}\rangle = 0\}.$$

There is a natural projection map $\mathbb{P}$ from $\mathbb{C}^{2,1} - \{0\}$ to $\mathbb{CP}^2$ that identifies all nonzero (complex) scalar multiples of a vector in $\mathbb{C}^{2,1}$. *Complex hyperbolic space* is defined to be $\boldsymbol{H}^2_{\mathbb{C}} = \mathbb{P}V_-$ and its boundary is $\partial \boldsymbol{H}^2_{\mathbb{C}} = \mathbb{P}V_0$. Clearly, if $\boldsymbol{v}$ lies in $V_-$ or $V_0$ then $v_3 \neq 0$ and so $\boldsymbol{H}^2_{\mathbb{C}} \cup \partial \boldsymbol{H}^2_{\mathbb{C}}$ is contained in the affine chart of $\mathbb{CP}^2$ with $v_3 \neq 0$. We canonically identify this chart with $\mathbb{C}^2$ by setting $z = v_1/v_3$ and $w = v_2/v_3$. Thus a vector $(z, w) \in \mathbb{C}^2$ corresponds to $[z : w : 1]^t$ in $\mathbb{CP}^2$. Evaluating the Hermitian form at this point gives $|z|^2 + |w|^2 - 1 = (|v_1|^2 + |v_2|^2 - |v_3|^2)/|v_3|^2$. Therefore

$$\boldsymbol{H}^2_{\mathbb{C}} = \{(z, w) \in \mathbb{C}^2 : |z|^2 + |w|^2 < 1\}, \quad \partial \boldsymbol{H}^2_{\mathbb{C}} = \{(z, w) \in \mathbb{C}^2 : |z|^2 + |w|^2 = 1\}.$$

In other words, $\boldsymbol{H}^2_{\mathbb{C}}$ is the unit ball in $\mathbb{C}^2$ and its boundary is the unit sphere $S^3$.

The Bergman metric on $\boldsymbol{H}^2_{\mathbb{C}}$ is given in terms of the Hermitian form. Let $u$ and $v$ be points in $\boldsymbol{H}^2_{\mathbb{C}}$ and let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors in $V_-$ so that $\mathbb{P}\boldsymbol{u} = u$ and $\mathbb{P}\boldsymbol{v} = v$. The Bergman metric is given as a Riemannian metric $ds^2$ or a distance function $\rho(u, v)$ by the formulae

$$ds^2 = \frac{-4}{\langle \boldsymbol{u}, \boldsymbol{u}\rangle^2} \det\begin{pmatrix} \langle \boldsymbol{u}, \boldsymbol{u}\rangle & \langle d\boldsymbol{u}, \boldsymbol{u}\rangle \\ \langle \boldsymbol{u}, d\boldsymbol{u}\rangle & \langle d\boldsymbol{u}, d\boldsymbol{u}\rangle \end{pmatrix}, \quad \cosh^2\left(\frac{\rho(u, v)}{2}\right) = \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle\langle \boldsymbol{v}, \boldsymbol{u}\rangle}{\langle \boldsymbol{u}, \boldsymbol{u}\rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle}.$$

The formulae for the Bergman metric are homogeneous and so the ambiguity in the choice of $\boldsymbol{u}$ and $\boldsymbol{v}$ does not matter.

Let $\mathrm{SU}(2, 1)$ be the group of unimodular matrices preserving the Hermitian form $H$. An element $A$ of $\mathrm{SU}(2, 1)$ acts on $\boldsymbol{H}^2_{\mathbb{C}}$ as $A(u) = \mathbb{P}(A\boldsymbol{u})$, where $\boldsymbol{u}$ is any vector in $V_-$ with $\mathbb{P}\boldsymbol{u} = u$. It is clear that scalar multiples of the identity act trivially. Since the determinant of $A$ is 1, such a scalar multiple must be a cube root of unity. Therefore, we define $\mathrm{PU}(2, 1) = \mathrm{SU}(2, 1)/\{\omega I : \omega^3 = 1\}$. Since the Bergman metric is given in terms of the Hermitian form, it is clear that elements of $\mathrm{SU}(2, 1)$ or $\mathrm{PU}(2, 1)$, act as isometries of $\boldsymbol{H}^2_{\mathbb{C}}$. Indeed, $\mathrm{PU}(2, 1)$ is the full group of holomorphic isometries of $\boldsymbol{H}^2_{\mathbb{C}}$. In what follows, we choose to work with matrices in $\mathrm{SU}(2, 1)$.

There are two kinds of totally geodesic two-dimensional submanifolds in $\boldsymbol{H}^2_{\mathbb{C}}$: complex lines and totally real totally geodesic subspaces. Let $\boldsymbol{c} \in \mathbb{C}^{2,1}$ be a vector with $\langle \boldsymbol{c}, \boldsymbol{c} \rangle > 0$. Then a *complex line* is the projection of the set $\{\boldsymbol{z} \in \mathbb{C}^{2,1} : \langle \boldsymbol{z}, \boldsymbol{c} \rangle = 0\}$. The vector $\boldsymbol{c}$ is then called a *polar vector* of the complex line. The *complex reflection* with polar vector $\boldsymbol{c}$ is defined to be

$$I_{\boldsymbol{c}}(\boldsymbol{z}) = -\boldsymbol{z} + \frac{2\langle \boldsymbol{z}, \boldsymbol{c} \rangle}{\langle \boldsymbol{c}, \boldsymbol{c} \rangle}\boldsymbol{c}.$$

***Bisectors and Dirichlet domains.*** We will consider subgroups of SU(2, 1) acting on $\boldsymbol{H}^2_{\mathbb{C}}$ and we want to show they are discrete. We will do this by constructing a fundamental polyhedron and using the Poincaré polyhedron theorem. There are no totally geodesic real hypersurfaces in $\boldsymbol{H}^2_{\mathbb{C}}$ and so we must choose hypersurfaces for the sides of our polyhedra. We choose to work with bisectors. A *bisector* in $\boldsymbol{H}^2_{\mathbb{C}}$ is the locus of points equidistant (with respect to the Bergman metric) from a given pair of points in $\boldsymbol{H}^2_{\mathbb{C}}$. Suppose that these points are $u$ and $v$. Choose lifts $\boldsymbol{u} = (u_1, u_2, u_3)^t$ and $\boldsymbol{v} = (v_1, v_2, v_3)^t$ to $V_-$ so that $\langle \boldsymbol{u}, \boldsymbol{u} \rangle = \langle \boldsymbol{v}, \boldsymbol{v} \rangle$. Then the bisector equidistant from $u$ and $v$ is

$$\begin{aligned} \mathcal{B} = \mathcal{B}(u, v) &= \{(z, w) \in \boldsymbol{H}^2_{\mathbb{C}} : \rho((z, w), u) = \rho((z, w), v)\} \\ &= \{(z, w) \in \boldsymbol{H}^2_{\mathbb{C}} : |z\bar{u}_1 + w\bar{u}_2 - \bar{u}_3| = |z\bar{v}_1 + w\bar{v}_2 - \bar{v}_3|\}. \end{aligned}$$

Suppose that we are given three points $u$, $v_1$ and $v_2$ in $\boldsymbol{H}^2_{\mathbb{C}}$. If the three corresponding vectors $\boldsymbol{u}$, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in $V_-$ form a basis for $\mathbb{C}^{2,1}$ then the intersection $\mathcal{B}(u, v_1) \cap \mathcal{B}(u, v_2)$ is called a Giraud disc. This is a particularly nice type of bisector intersection (see [Deraux et al. 2015, Section 2.5]).

Suppose that $\Gamma$ is a discrete subgroup of PU(2, 1). Let $u$ be a point of $\boldsymbol{H}^2_{\mathbb{C}}$ and write $\Gamma_u$ for the stabiliser of $u$ in $\Gamma$ (that is, the subgroup of $\Gamma$ comprising all elements fixing $u$). Then the *Dirichlet domain* $D_u(\Gamma)$ for $\Gamma$ with centre $u$ is defined to be

$$D_u(\Gamma) = \{v \in \boldsymbol{H}^2_{\mathbb{C}} : \rho(v, u) < \rho(v, A(u)) \text{ for all } A \in \Gamma - \Gamma_u\}.$$

Dirichlet domains for certain cyclic groups are particularly simple.

**Proposition 2.1.** *Let $A$ be a regular elliptic element of* PU(2, 1) *of order* 3. *Then for any point $u$ not fixed by $A$, the Dirichlet domain $D_u(\langle A \rangle)$ for the cyclic group $\langle A \rangle$ with centre $u$ has exactly two sides.*

*Proof.* Since there are only two nontrivial elements in $\langle A \rangle$, neither of which fix $u$, the Dirichlet domain $D_u(\langle A \rangle)$ is

$$D_u(\langle A \rangle) = \{v \in \boldsymbol{H}^2_{\mathbb{C}} : \rho(v, u) < \rho(v, A(u)),\ \rho(v, u) < \rho(v, A^{-1}(u))\}.$$

Its images under $A$ and $A^{-1}$ are

$$A(D_u(\langle A\rangle)) = \{v : \rho(v, A(u)) < \rho(v, u),\ \rho(v, A(u)) < \rho(v, A^{-1}(u))\},$$
$$A^{-1}(D_u(\langle A\rangle)) = \{v : \rho(v, A^{-1}(u)) < \rho(v, u),\ \rho(v, A^{-1}(u)) < \rho(v, A(u))\}.$$

By considering the minimum of $\rho(v, u)$, $\rho(v, A(u))$, $\rho(v, A^{-1}(u))$ as $v$ varies over $\boldsymbol{H}^2_{\mathbb{C}}$, it is clear these three domains are disjoint and their closures cover $\boldsymbol{H}^2_{\mathbb{C}}$. □

**Proposition 2.2** [Phillips 1992]. *Let $A \in \mathrm{SU}(2,1)$ have real trace which is at least* 3. *Then for any $u \in \boldsymbol{H}^2_{\mathbb{C}}$, the bisectors $\mathcal{B}(u, A(u))$ and $\mathcal{B}(u, A^{-1}(u))$ are disjoint. Thus, the Dirichlet domain $D_u(\langle A\rangle)$ has exactly two sides.*

***The Poincaré polyhedron theorem.*** Our goal is to construct the Dirichlet domain for a complex hyperbolic representation $\Gamma$ of the $(3, 3, n)$ triangle group with centre the fixed point of an order-$n$ elliptic map. If we use the definition of Dirichlet domain, then we need to check infinitely many inequalities. Thus, we need to use another method. This method is to construct a candidate Dirichlet domain and then use the Poincaré polyhedron theorem.

The main tool we use to show discreteness is the Poincaré polyhedron theorem. The version of this theorem that we use is for polyhedra $D$ with a finite stabiliser; see [Mostow 1980, Theorem 6.3.2] or [Deraux et al. 2015, Theorem 3.2]. Rather than give a general statement of this theorem, we will state it in the particular case we are interested in, namely Dirichlet polyhedra for reflection groups.

Let $u$ be a point in $\boldsymbol{H}^2_{\mathbb{C}}$ and let $\Upsilon$ be a finite subgroup of $\mathrm{PU}(2,1)$ fixing $u$. Let $A_1, \ldots, A_n$ be a finite collection of involutions in $\mathrm{PU}(2,1)$ (so $A_i^2$ is the identity for each $i$). Suppose that no $A_i$ fixes $u$. Suppose that the group $\Upsilon$ preserves this collection of involutions under conjugation. That is, for each $A_i$ with $1 \le i \le n$ and each $P \in \Upsilon$, we suppose that $PA_iP^{-1} = A_j$ for some $1 \le j \le n$. Let $\mathcal{B}_i = \mathcal{B}(u, A_i(u))$ be the bisector equidistant from $u$ and $A_i(u)$. If $P \in \Upsilon$ satisfies $PA_iP^{-1} = A_j$ then $PA_i(u) = A_j(u)$ (since $P(u) = u$) and so $P$ maps $\mathcal{B}_i$ to $\mathcal{B}_j$. We define $D$ to be the component of $\boldsymbol{H}^2_{\mathbb{C}} - \bigcup_{i=1}^n \mathcal{B}_i$ containing $u$, and we suppose that there are points from each of the $\mathcal{B}_i$ on the boundary of $D$ (that is, the $\mathcal{B}_i$ are not nested). This construction makes $D$ open. Note that, by construction, $\Upsilon$ maps $D$ to itself.

For each $1 \le i \le n$, let $s_i = \mathcal{B}_i \cap \overline{D}$. We call $s_i$ a *side* of $D$. Such a side can be given a cell structure based on how it intersects other sides. We suppose that the involutions $A_i$ for $1 \le i \le n$ satisfy the following conditions, and so form a *side pairing* of $D$:

(1) For each $1 \le i \le n$, the involution $A_i$ sends $s_i$ to itself, preserving the cell structure. The relation $A_i^2 = \mathrm{id}$ is called a *reflection relation*.

(2) For each $1 \le i \le n$, we have $\overline{D} \cap A_i(\overline{D}) = s_i$ and $D \cap A(D) = \varnothing$.

(3) If $v$ is a point in $s_i$ and in no other side (that is, $v$ lies in the relative interior of $s_i$) then there is an open neighbourhood $U_v$ of $v$ lying in $\overline{D} \cup A_i(\overline{D})$.

Note that, unlike the case of reflection groups in constant curvature, $A_i$ does not fix $s_i$ pointwise. Therefore, we could have subdivided $s_i$ into two sets (each of dimension 3) that are interchanged by $A_i$. In practice this would cause unnecessary complication.

Suppose that $s_i$ and $s_j$ are two sides with nonempty intersection. Their intersection $r = s_i \cap s_j$ is called a *ridge* of $D$. Since $A_i$ preserves the cell structure of $s_i$, we see that $A_i(r) = s_i \cap s_k$ is another ridge of $D$. Applying $A_k$ gives another ridge in $s_k$. Continuing in this way gives a *ridge cycle*

$$(r_1, s_{i_0}, s_{i_1}) \xrightarrow{A_{i_1}} (r_2, s_{i_1}, s_{i_2}) \xrightarrow{A_{i_2}} (r_3, s_{i_2}, s_{i_3}) \cdots .$$

Here $(r_j, s_{i_{j-1}}, s_{i_j})$ is an ordered triple with $r_j = s_{i_{j-1}} \cap s_{i_j}$. Since there are finitely many $\Upsilon$ orbits of $r_1$, eventually we find a ridge $r_{m+1} = s_{i_m} \cap s_{i_{m+1}}$ so that the corresponding ordered triple satisfies

$$(r_{m+1}, s_{i_m}, s_{i_{m+1}}) \xrightarrow{P} (r_1, s_{i_0}, s_{i_1})$$

for some $P \in \Upsilon$. We call $T_1 = PA_{i_m} \cdots A_{i_1}$ the *cycle transformation* associated to $r_1$. It means that the ridge cycle starts at $(r_1, s_{i_0}, s_{i_1})$ and ends to itself by $T_1$. Clearly $T_1$ maps $r_1$ to itself. Of course, $T_1$ may not act as the identity on $r_1$ and even if it does, it may not act as the identity on $\boldsymbol{H}^2_{\mathbb{C}}$. Nevertheless, we suppose $T_1$ has finite order $\ell$. The relation $T_i^{\ell} = \mathrm{id}$ is called a *cycle relation*.

In the example we are interested in, the ridge cycle is

$$(r_1, s_{i_0}, s_{i_1}) \xrightarrow{A_{i_1}} (r_2, s_{i_1}, s_{i_2}) \xrightarrow{P} (r_1, s_{i_0}, s_{i_1})$$

and, in fact, $s_{i_2} = s_{i_0}$ and so $r_2 = r_1$. Moreover, $P$ is an involution with $P(r_1) = r_1$ and $P(s_{i_1}) = s_{i_0}$. Hence the cycle transformation is $T_1 = PA_{i_1}$, which happens to have order 3. Thus, the cycle relation is $T_1^3 = (PA_{i_1})^3 = \mathrm{id}$.

We suppose that $D$ satisfies the *cycle condition* which means that copies of $D$ tessellate a neighbourhood for each ridge $r$. Furthermore, the relevant copies of $D$ are its preimages under suffix subwords of $T^{\ell}$. The full statement is explained in [Deraux et al. 2015]. For brevity, we state this condition only in the special case we are interested in. Let $r$ be a ridge and let $T = PA_i$ be its cycle transformation with cycle relation $(PA_i)^3 = \mathrm{id}$. Let $\mathcal{C} = \{\mathrm{id}, PA_i, (PA_i)^2\}$. Then the cycle condition states that

(1) $$r = \bigcap_{C \in \mathcal{C}} C^{-1}(\bar{D}).$$

(2) If $C_1, C_2 \in \mathcal{C}$ with $C_1 \neq C_2$ then $C_1^{-1}(D) \cap C_2^{-1}(D) = \varnothing$.

(3) If $v$ is a point in $r$ and in no other ridge (that is, $v$ lies in the relative interior of $r$) then there is an open neighbourhood $U_v$ of $v$ with

$$U_v \subset \bigcup_{C \in \mathcal{C}} C^{-1}(\bar{D}).$$

It means that there are exactly three copies of $D$ along each ridge $r$, which are $D$, $T(D)$ and $T^2(D)$. Observe that the stabiliser of $r$ is generated by $A_i$ and $P$. Hence it is a dihedral group of order 6. Since $A_i$, $P$ and $PA_iP^{-1}$ preserve one of the three copies and interchange the other two, the stabiliser preserves the three copies of $D$.

Finally, if two sides of $D$ are asymptotic at a point $v$ of $\partial \boldsymbol{H}^2_{\mathbb{C}}$ then there is a horoball $H_v$ so that $H_v$ intersects $\bar{D}$ only in facets of $D$ containing $v$ and $H_v$ is preserved by the stabiliser of $v$ in $\Gamma$. We say that $H_v$ is a *consistent horoball* at $v$. In particular, if $v$ is a fixed point of a parabolic element of $\Gamma$ then there exists a consistent horoball at $v$.

The Poincaré polyhedron theorem states:

**Theorem 2.3** [Mostow 1980, Theorem 6.3.2; Deraux et al. 2015, Theorem 3.2]. *Suppose that $D$ is a polyhedron on $\boldsymbol{H}^2_{\mathbb{C}}$ with sides contained in bisectors together with a side pairing. Let $\Upsilon < \mathrm{PU}(2,1)$ be a discrete group of automorphisms of $D$. Let $\Gamma$ be the group generated by $\Upsilon$ and the side pairing maps. Suppose that the cycle condition holds at all ridges of $D$ and that there is a consistent horoball at all points* (*if any*) *where sides of $D$ are asymptotic. Then*:

(1) $\Gamma$ *is discrete.*

(2) *The images of $D$ under the cosets of $\Upsilon$ in $\Gamma$ tessellate $\boldsymbol{H}^2_{\mathbb{C}}$.*

(3) *A fundamental domain for $\Gamma$ may be obtained by intersecting $D$ with a fundamental domain for $\Upsilon$.*

(4) *A presentation for $\Gamma$ is given as follows. The generators are a generating set for $\Upsilon$ together with all side pairing maps. The relations are generated by all relations in $\Upsilon$, all reflection relations and all cycle relations.*

## 3. The generators

Consider complex reflections $I_1$ and $I_2$ in $\mathrm{SU}(2,1)$ so that $I_1I_2$ has order $n$ and fixes the origin $o$. Writing $c = \cos(\pi/n)$ and $s = \sin(\pi/n)$, we may choose $I_1$ and $I_2$ to be

$$I_1 = \begin{bmatrix} -c & s & 0 \\ s & c & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad I_2 = \begin{bmatrix} -c & -s & 0 \\ -s & c & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{3-1}$$

Note that polar vectors of $I_1$ and $I_2$ are

$$\boldsymbol{n}_1 = \begin{bmatrix} s \\ 1+c \\ 0 \end{bmatrix}, \quad \boldsymbol{n}_2 = \begin{bmatrix} -s \\ 1+c \\ 0 \end{bmatrix}.$$

We want to find $I_3$ so that $I_1I_3$ and $I_2I_3$ both have order 3. Conjugating by a diagonal map $\mathrm{diag}(e^{i\psi}, e^{i\psi}, e^{-2i\psi})$ if necessary, we may suppose that the polar

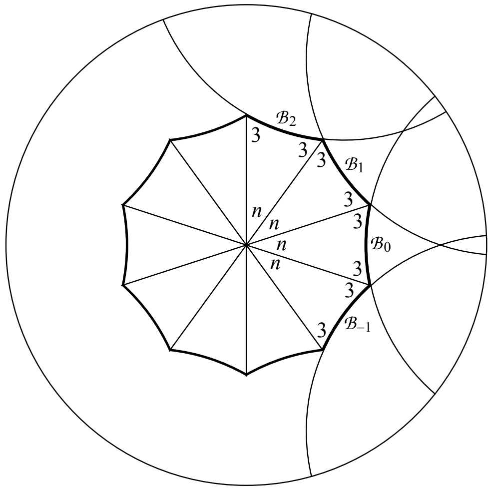


**Figure 1.** The $2n$-gon in the hyperbolic plane made up of $2n$ copies of a $(3, 3, n)$ triangle.

vector of $I_3$ is

$$\boldsymbol{n}_3 = \begin{bmatrix} a \\ be^{i\theta} \\ d-1 \end{bmatrix},$$

where $a$, $b$, $d$ are nonnegative real numbers satisfying $a^2+b^2-(d-1)^2 = 2(d-1)$, that is, $a^2+b^2-d^2 = -1$. Furthermore, complex conjugating if necessary, we may always assume $\theta \in [0, \pi]$. Then

$$I_3 = \begin{bmatrix} -1+a^2/(d-1) & abe^{-i\theta}/(d-1) & -a \\ abe^{i\theta}/(d-1) & -1+b^2/(d-1) & -be^{i\theta} \\ a & be^{-i\theta} & -d \end{bmatrix}. \tag{3-2}$$

It is easy to check that $I_3$ lies in SU(2, 1), has order 2 and polar vector $\boldsymbol{n}_3$.

**Lemma 3.1.** *Let $I_1$, $I_2$ and $I_3$ be given by* (3-1) *and* (3-2). *If $I_1I_3$ and $I_2I_3$ have order* 3 *then $\theta = \pi/2$ and*

$$c(a^2-b^2) = d(d-1). \tag{3-3}$$

*Proof.* The condition that $I_1I_3$ and $I_2I_3$ have order 3 is equivalent to $\mathrm{tr}(I_1I_3) = \mathrm{tr}(I_2I_3) = 0$. That is,

$$\frac{-c(a^2-b^2)+2sab\cos\theta}{d-1} + d = \frac{-c(a^2-b^2)-2sab\cos\theta}{d-1} + d = 0.$$

The result follows directly. $\square$

From now on, we write $\theta = \pi/2$ in (3-2). Since we know $a^2 + b^2 = d^2 - 1$ and $a^2 - b^2 = d(d-1)/c$, we immediately have

$$a^2 = (d-1)(1+d+d/c)/2, \quad b^2 = (d-1)(1+d-d/c)/2. \tag{3-4}$$

**Corollary 3.2.** *Let*

$$\iota : \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \longmapsto \begin{bmatrix} \bar{z}_1 \\ -\bar{z}_2 \\ \bar{z}_3 \end{bmatrix}.$$

*Then* $\iota$ *has order* 2 *and*

$$\iota I_1 \iota = I_2, \quad \iota I_2 \iota = I_1, \quad \iota I_3 \iota = I_3.$$

*Proof.* It is easy to see that $\iota^2$ is the identity. A simple calculation shows $\iota(\boldsymbol{n}_1) = \boldsymbol{n}_2$ and $\iota(\boldsymbol{n}_3) = \boldsymbol{n}_3$, using $e^{i\theta} = i$. □

**Lemma 3.3.** *The group* $\langle I_1, I_2, I_3 \rangle$ *is determined up to conjugacy by the variable* $d$, *which lies in the interval* $1 < d \le c/(1-c)$. *Moreover*, $\langle I_1, I_2, I_3 \rangle$ *lies in* $\mathrm{SO}(2,1)$ *when* $d = c/(1-c)$.

*Proof.* We have conjugated so that $I_1$ and $I_2$ have the form (3-1), and $I_3$ has the form (3-2) with $\theta = \pi/2$. After this conjugation, the only remaining parameters are the nonnegative real numbers $a$, $b$ and $d$. Using (3-4) these are completely determined by $d$. Moreover, again using (3-4) we see that $a^2$ and $b^2$ are nonnegative if and only if $d \ge 1$ and $d \le c/(1-c)$. We cannot have $d = 1$ or else $\boldsymbol{n}_3$ is the zero vector. Thus $1 < d \le c/(1-c)$. Finally, when $d = c/(1-c)$, we have $b = 0$ and the entries of $I_3$ are all real. □

**Lemma 3.4.** *Let* $I_1$, $I_2$ *and* $I_3$ *be given by* (3-1) *and* (3-2). *Suppose* $I_1 I_3$ *and* $I_2 I_3$ *have order* 3. *Then* $I_1 I_3 I_2 I_3$ *is elliptic if and only if* $d < 3/(4s^2)$.

*Proof.* Calculating directly, we see that

$$\begin{aligned} \mathrm{tr}(I_1 I_3 I_2 I_3) &= \frac{c^2(a^2-b^2)^2}{(d-1)^2} + \frac{2(c^2-s^2)(d-1-a^2-b^2)}{d-1} - 2c(a^2-b^2) + d^2 \\ &= 4s^2 d. \end{aligned}$$

(We could have derived this using the formulae in [Pratoussevitch 2005].) The condition that $I_1 I_3 I_2 I_3$ is elliptic is equivalent to $3 > \mathrm{tr}(I_1 I_3 I_2 I_3) = 4s^2 d$. □

Thus, our parameter space for $\langle I_1, I_2, I_3 \rangle$ with $I_1 I_3 I_2 I_3$ nonelliptic is given by

$$\frac{3}{4s^2} \le d \le \frac{c}{1-c}. \tag{3-5}$$

Note that the condition $n > 3$ implies both $3/(4s^2) > 1$ and $c/(1-c) > 1$. For example, when $n = 4$ we have $c = s = 1/\sqrt{2}$ and our range becomes

$$3/2 \le d \le \sqrt{2} + 1.$$

## 4. The bisectors

We define a polyhedron $D$ bounded by sides contained in $2n$ bisectors.

**Definition 4.1.** For $k \in \mathbb{Z}$, define the involution $A_k \in \langle I_1, I_2, I_3\rangle$ as follows:

(1) If $k = 2m$ is an even integer then $A_k = (I_2 I_1)^{k/2} I_3 (I_1 I_2)^{k/2}$.

(2) If $k = 2m+1$ is an odd integer then $A_k = (I_2 I_1)^{(k-1)/2} I_2 I_3 I_2 (I_1 I_2)^{(k-1)/2}$.

Let $o$ be the fixed point of $I_1 I_2$ in $\boldsymbol{H}^2_{\mathbb{C}}$. For all integers $k$, the bisector $\mathcal{B}_k$ is defined to be the bisector equidistant from $o$ and $A_k(o)$. Note that in both cases $A_{k+2n} = A_k$ and so $\mathcal{B}_{k+2n} = \mathcal{B}_k$. This gives $2n$ bisectors $\mathcal{B}_{-n+1}$ to $\mathcal{B}_n$ and we may take the index $k$ mod $2n$.

The following lemma follows immediately from the definition.

**Lemma 4.2.** *Let $\mathcal{B}_{-n+1}$ to $\mathcal{B}_n$ be as defined in Definition 4.1. Then for each $k$ mod $2n$ and each $m$ mod $n$:*

(1) *The map $(I_2 I_1)^m$ sends $\mathcal{B}_k$ to $\mathcal{B}_{2m+k}$.*

(2) *The map $(I_2 I_1)^m I_2$ sends $\mathcal{B}_k$ to $\mathcal{B}_{2m+1-k}$. In particular, the map $(I_2 I_1)^k I_2$ sends $\mathcal{B}_k$ to $\mathcal{B}_{k+1}$.*

(3) *The antiholomorphic involution $\iota$ defined in Corollary 3.2 sends $\mathcal{B}_k$ to $\mathcal{B}_{-k}$. In particular, the map $(I_2 I_1)^m I_2 \iota$ sends $\mathcal{B}_k$ to $\mathcal{B}_{2m+1+k}$.*

The main result of this section is that the combinatorial configuration of the bisectors does not change as $d$ decreases from $c/(1-c)$ to $3/(4s^2)$. More precisely:

**Theorem 4.3.** *Let $\mathcal{B}_{-n+1}$ to $\mathcal{B}_n$ be as defined in Definition 4.1. Suppose that $3/(4s^2) \le d \le c/(1-c)$. Then, taking the indices mod $2n$, for each $k$:*

(1) *The bisector $\mathcal{B}_k$ intersects $\mathcal{B}_{k\pm 1}$ in a Giraud disc. This Giraud disc is preserved by $A_k A_{k\pm 1}$, which has order 3.*

(2) *The intersection of $\mathcal{B}_k$ with $\mathcal{B}_{k\pm 2}$ is contained in the halfspace bounded by $\mathcal{B}_{k\pm 1}$ not containing $o$.*

(3) *The bisector $\mathcal{B}_k$ does not intersect $\mathcal{B}_{k\pm\ell}$ for $3 \le \ell \le n$. Moreover, the boundaries of these bisectors are disjoint except for when $\ell = 3$ and $d = 3/(4s^2)$, in which case the boundaries intersect in a single point, which is a parabolic fixed point.*

As a corollary to this theorem, we can use the Poincaré polyhedron theorem to prove the "if" part of Theorem 1.6.

**Corollary 4.4.** *Let $A_{-n+1}$ to $A_n$ and $\mathcal{B}_{-n+1}$ to $\mathcal{B}_n$ be as in Theorem 4.3. Suppose that $3/(4s^2) \le d \le c/(1-c)$. Let $D$ be the polyhedron in $\boldsymbol{H}^2_{\mathbb{C}}$ containing $o$ and bounded by $\mathcal{B}_{-n+1}$ to $\mathcal{B}_n$. Then the maps $A_{-n+1}$ to $A_n$ form a side paring for $D$ that satisfies the conditions of the Poincaré polyhedron theorem, Theorem 2.3. In particular, $\langle I_1, I_2, I_3\rangle$ is a discrete and faithful representation of $\Delta_{3,3,n}$.*

*Proof.* Since $A_k$ is an involution, it is clear that the $\{A_k\}$ form a side pairing for $D$. Now consider the ridge $r_k = \mathcal{B}_k \cap \mathcal{B}_{k+1}$. Applying either of the side pairing maps $A_k$ or $A_{k+1}$ sends this ridge to itself. We then apply $P_k = (I_2 I_1)^k I_2$ to obtain the cycle transformation $P_k A_k$. When $k$ is even,

$$P_k A_k = (I_2 I_1)^k I_2 (I_2 I_1)^{k/2} I_3 (I_1 I_2)^{k/2} = (I_2 I_1)^{k/2} I_2 I_3 (I_1 I_2)^{k/2},$$

and when $k$ is odd,

$$\begin{aligned} P_k A_k &= (I_2 I_1)^k I_2 (I_2 I_1)^{(k-1)/2} I_2 I_3 I_2 (I_1 I_2)^{(k-1)/2} \\ &= (I_2 I_1)^{(k+1)/2} I_3 I_1 (I_1 I_2)^{(k+1)/2}. \end{aligned}$$

In both cases, $P_k A_k$ is equal to $A_k A_{k+1}$, which has order 3. There is a neighbourhood $U_k$ of the ridge $r_k$ for which the intersection of $U_k$ with $D$ is the same as its intersection with the Dirichlet domain for $\langle P_k A_k \rangle$. Therefore, we have local tessellation around all the ridges of $D$ using the argument of Proposition 2.1.

All the other sides of $D$ are disjoint, apart from when $d = 3/(4s^2)$, in which case $\mathcal{B}_k$ and $\mathcal{B}_{k\pm 3}$ are asymptotic at a point of $\partial \boldsymbol{H}^2_{\mathbb{C}}$. This point is a parabolic fixed point, as required.

Finally, each side yields the reflection relation $A_k^2$, which is conjugate to $I_3^2$. The cycle relations give $(P_k A_k)^3$, which are conjugate to $(I_2 I_3)^3$ when $k$ is even and $(I_3 I_1)^3$ when $k$ is odd. In addition we have the relations from $\Upsilon = \langle I_1, I_2 \rangle$, which are $I_1^2$, $I_2^2$ and $(I_1 I_2)^n$. From the Poincaré theorem, all other relations may be deduced from these. Thus $\langle I_1, I_2, I_3 \rangle$ is a faithful representation of $\Delta_{3,3,n}$. $\square$

Write $c_k = \cos(k\pi/n)$ and $s_k = \sin(k\pi/n)$. Then

$$(I_2 I_1)^m = \begin{bmatrix} c_{2m} & -s_{2m} & 0 \\ s_{2m} & c_{2m} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (I_2 I_1)^m I_2 \begin{bmatrix} -c_{2m+1} & -s_{2m+1} & 0 \\ -s_{2m+1} & c_{2m+1} & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

We have

$$(I_2 I_1)^m I_3(o) = \begin{bmatrix} -c_{2m} a + s_{2m} bi \\ -s_{2m} a - c_{2m} bi \\ -d \end{bmatrix}, \quad (I_1 I_2)^m I_3(o) = \begin{bmatrix} -c_{2m} a - s_{2m} bi \\ s_{2m} a - c_{2m} bi \\ -d \end{bmatrix}.$$

Also

$$(I_2 I_1)^m I_2 I_3(o) = \begin{bmatrix} c_{2m+1} a + s_{2m+1} bi \\ s_{2m+1} a - c_{2m+1} bi \\ d \end{bmatrix}, \quad (I_1 I_2)^m I_1 I_3(o) = \begin{bmatrix} c_{2m+1} a - s_{2m+1} bi \\ -s_{2m+1} a - c_{2m+1} bi \\ d \end{bmatrix}.$$

We begin by proving Theorem 4.3(1).

**Proposition 4.5.** *For each $-n+1 \le k \le n$, the bisectors $\mathcal{B}_k$ and $\mathcal{B}_{k\pm 1}$ (with indices taken mod $2n$) intersect in $\boldsymbol{H}^2_{\mathbb{C}}$ in a Giraud disc. This Giraud disc is preserved by $(I_2 I_1)^{k/2}(I_2 I_3)(I_1 I_2)^{k/2}$ when $k$ is even and $(I_2 I_1)^{(k+1)/2}(I_3 I_1)(I_1 I_2)^{(k+1)/2}$ when $k$ is odd.*

*Proof.* Using Lemma 4.2 we need only consider $k = 0$ and $k = 1$. The bisectors $\mathcal{B}_0$ and $\mathcal{B}_1$ are equidistant from $o$ and from $I_3(o) = I_3 I_2(o)$ and from $I_2 I_3(o)$ respectively. Observe that $I_2 I_3$ does not fix $o$. Since the map $I_2 I_3$ has order 3, the Dirichlet domain with centre $o$ for the cyclic group $\langle I_2 I_3\rangle$ only contains faces contained in these two bisectors. The intersection is a Giraud disc invariant under powers of $I_2 I_3$ by construction. □

Next we prove Theorem 4.3(3) in the case where $\ell = 2m+1$ is odd.

**Proposition 4.6.** *Suppose that* $3/(4s^2) \le d \le c/(1-c)$. *For each* $-n+1 \le k \le n$ *and* $1 \le m \le (n-1)/2$, *the bisectors* $\mathcal{B}_k$ *and* $\mathcal{B}_{k\pm(2m+1)}$ (*with indices taken mod* $2n$) *do not intersect in* $\boldsymbol{H}^2_{\mathbb{C}}$. *Moreover, their closures intersect on* $\partial \boldsymbol{H}^2_{\mathbb{C}}$ *if and only if* $d = 3/(4s^2)$ *and* $m = 1$. *In the latter case, the closures intersect in a unique point, which is a parabolic fixed point.*

*Proof.* Using Lemma 4.2 we need only consider $\mathcal{B}_0$ and $\mathcal{B}_{2m+1}$. These bisectors are equidistant from $o$ and $I_3(o) = I_3 I_2 (I_1 I_2)^m(o)$ and from $(I_2 I_1)^m I_2 I_3(o)$ respectively. Consider the Dirichlet domain with centre $o$ for the cyclic group $\langle (I_2 I_1)^m I_2 I_3\rangle$. We claim that this Dirichlet domain has exactly two sides and these sides are disjoint. To do so, we use Phillips' theorem, Proposition 2.2.

A brief calculation shows that

$$\operatorname{tr}\bigl((I_2 I_1)^m I_2 I_3\bigr) = -c_{2m+1}\frac{a^2-b^2}{d-1} + d = \frac{d(c-c_{2m+1})}{c} = \frac{2d s_{m+1} s_m}{c}.$$

When $1 \le m \le (n-1)/2$, we have

$$s_m s_{m+1} \ge s s_2 = 2s^2 c$$

with equality if and only if $m = 1$. Therefore,

$$\operatorname{tr}\bigl((I_2 I_1)^m I_2 I_3\bigr) = 2d s_{m+1} s_m / c \ge 4ds^2$$

with equality if and only if $m = 1$. Hence, when $4ds^2 \ge 3$, we have $(I_2 I_1)^m I_2 I_3$ is nonelliptic with real trace, and is loxodromic unless $m = 1$ and $d = 3/(4s^2)$. By Phillips' theorem we see that any Dirichlet domain for $\langle (I_2 I_1)^m I_2 I_3\rangle$ has two faces and these faces do not intersect in $\boldsymbol{H}^2_{\mathbb{C}}$.

In fact, when $d = 3/(4s^2)$ and $m = 1$, the bisectors $\mathcal{B}_0$ and $\mathcal{B}_3$ are asymptotic on the boundary of $\boldsymbol{H}^2_{\mathbb{C}}$ at the (parabolic) fixed point of $I_2 I_1 I_2 I_3$. □

**Proposition 4.7.** (i) *Suppose* $p = [z, w, 1]^t$ *lies on* $\mathcal{B}_{2\ell} \cap \mathcal{B}_{-2\ell}$. *Then for some angles* $\theta, \phi$, *we have*

$$z = \frac{s_{2\ell} a(\cos\theta e^{i\phi} + d) - c_{2\ell} b \sin\theta e^{i\phi}}{c_{2\ell} s_{2\ell}(a^2-b^2)},$$

$$w = \frac{-s_{2\ell} bi(\cos\theta e^{i\phi} + d) + c_{2\ell} ai \sin\theta e^{i\phi}}{c_{2\ell} s_{2\ell}(a^2-b^2)}.$$

(ii) *Suppose* $p = [z, w, 1]^t$ *lies on* $\mathcal{B}_{2\ell+1} \cap \mathcal{B}_{-2\ell-1}$. *Then for some angles* $\theta, \phi$, *we have*

$$z = \frac{s_{2\ell+1}a(\cos\theta e^{i\phi} + d) - c_{2\ell+1}b \sin\theta e^{i\phi}}{c_{2\ell+1}s_{2\ell+1}(a^2 - b^2)},$$
$$w = \frac{s_{2\ell+1}bi(\cos\theta e^{i\phi} + d) - c_{2\ell+1}ai \sin\theta e^{i\phi}}{c_{2\ell+1}s_{2\ell+1}(a^2 - b^2)}.$$

*Proof.* First consider the bisector intersection from (i). Then $z$ and $w$ satisfy

$$1 = \bigl|z(-c_{2\ell}a + s_{2\ell}bi) + w(s_{2\ell}a + c_{2\ell}bi) + d\bigr|,$$
$$1 = \bigl|z(-c_{2\ell}a - s_{2\ell}bi) + w(-s_{2\ell}a + c_{2\ell}bi) + d\bigr|.$$

Expanding out, adding and subtracting yields

$$1 = \bigl|zc_{2\ell}a - wc_{2\ell}bi - d\bigr|^2 + \bigl|zs_{2\ell}bi + ws_{2\ell}a\bigr|^2,$$
$$0 = 2\,\mathrm{Re}\bigl((zc_{2\ell}a - wc_{2\ell}bi - d)(-\bar{z}s_{2\ell}bi + \bar{w}s_{2\ell}a)\bigr).$$

Thus we can write

$$zc_{2\ell}a - wc_{2\ell}bi - d = \cos\theta e^{i\phi},$$
$$zs_{2\ell}bi + ws_{2\ell}a = i\sin\theta e^{i\phi}.$$

Inverting these equations yields

$$z = \frac{s_{2\ell}a(\cos\theta e^{i\phi} + d) - c_{2\ell}b \sin\theta e^{i\phi}}{c_{2\ell}s_{2\ell}(a^2 - b^2)},$$
$$w = \frac{-s_{2\ell}bi(\cos\theta e^{i\phi} + d) + c_{2\ell}ai \sin\theta e^{i\phi}}{c_{2\ell}s_{2\ell}(a^2 - b^2)}.$$

For the second bisector intersection, we have

$$1 = \bigl|z(c_{2\ell+1}a + s_{2\ell+1}bi) + w(-s_{2\ell+1}a + c_{2\ell+1}bi) - d\bigr|^2,$$
$$1 = \bigl|z(c_{2\ell+1}a - s_{2\ell+1}bi) + w(s_{2\ell+1}a + c_{2\ell+1}bi) - d\bigr|^2.$$

Expanding out, adding and subtracting yields

$$1 = \bigl|zc_{2\ell+1}a + wc_{2\ell+1}bi - d\bigr|^2 + \bigl|-zs_{2\ell+1}bi + ws_{2\ell+1}a\bigr|^2,$$
$$0 = 2\,\mathrm{Re}\bigl((zc_{2\ell+1}a + wc_{2\ell+1}bi - d)(\bar{z}s_{2\ell+1}bi + \bar{w}s_{2\ell+1}a)\bigr).$$

So once again we have

$$zc_{2\ell+1}a + wc_{2\ell+1}bi - d = \cos\theta e^{i\phi},$$
$$-zs_{2\ell+1}bi + ws_{2\ell+1}a = -i\sin\theta e^{i\phi}.$$

Thus,

$$z = \frac{s_{2\ell+1}a(\cos\theta e^{i\phi}+d) - c_{2\ell+1}b\sin\theta e^{i\phi}}{c_{2\ell+1}s_{2\ell+1}(a^2-b^2)},$$
$$w = \frac{s_{2\ell+1}bi(\cos\theta e^{i\phi}+d) - c_{2\ell+1}ai\sin\theta e^{i\phi}}{c_{2\ell+1}s_{2\ell+1}(a^2-b^2)}. \qquad \square$$

We can now prove Theorem 4.3(3) in the case where $\ell = 2m$ is even.

**Proposition 4.8.** *Suppose that* $3/(4s^2) \le d \le c/(1-c)$. *For each* $-n+1 \le k \le n$ *and* $2 \le m \le n/2$, *the bisectors* $\mathcal{B}_k$ *and* $\mathcal{B}_{k\pm 2m}$ (*with indices taken mod* $2n$) *do not intersect in complex hyperbolic space.*

*Proof.* Using Lemma 4.2, we need only consider $\mathcal{B}_m$ and $\mathcal{B}_{-m}$ where $2 \le m \le n/2$.

Using Proposition 4.7 we see that an intersection point $p = [z, w, 1]^t$ of $\mathcal{B}_m$ and $\mathcal{B}_{-m}$ must satisfy

$$z = \frac{s_m a(\cos\theta e^{i\phi}+d) - c_m b\sin\theta e^{i\phi}}{c_m s_m(a^2-b^2)},$$
$$w = \pm\frac{-s_m bi(\cos\theta e^{i\phi}+d) + c_m ai\sin\theta e^{i\phi}}{c_m s_m(a^2-b^2)}.$$

We claim that $|z|^2+|w|^2 \ge 1$ and so such a point does not lie in $\boldsymbol{H}^2_{\mathbb{C}}$. We have

$$\begin{aligned}
&c_m^2 s_m^2(a^2-b^2)^2(|z|^2+|w|^2-1)\\
&\quad= \bigl|s_m a(\cos\theta e^{i\phi}+d) - c_m b\sin\theta e^{i\phi}\bigr|^2\\
&\qquad\quad + \bigl|-s_m bi(\cos\theta e^{i\phi}+d) + c_m ai\sin\theta e^{i\phi}\bigr|^2 - c_m^2 s_m^2(a^2-b^2)^2\\
&\quad= s_m^2(a^2+b^2)(\cos^2\theta + 2d\cos\theta\cos\phi + d^2)\\
&\qquad\quad - 2c_m s_m ab(2\cos\theta\sin\theta + 2d\sin\theta\cos\phi)\\
&\qquad\qquad + c_m^2(a^2+b^2)\sin^2\theta - c_m^2 s_m^2(a^2+b^2)^2 + 4c_m^2 s_m^2 a^2 b^2\\
&\quad= s_m^2(d^2-1)(\cos^2\theta + 2d\cos\theta\cos\phi + d^2)\\
&\qquad\quad - 4c_m s_m ab(\cos\theta\sin\theta + d\sin\theta\cos\phi)\\
&\qquad\qquad + c_m^2(d^2-1)\sin^2\theta - c_m^2 s_m^2(d^2-1)^2 + 4c_m^2 s_m^2 a^2 b^2\\
&\quad= \bigl(\cos\theta\sin\theta + d\sin\theta\cos\phi - 2c_m s_m ab\bigr)^2 + d^2\sin^2\theta\sin^2\phi\\
&\qquad\quad + (s_m^2(d^2-1) - \sin^2\theta)\bigl(\cos^2\theta + 2d\cos\theta\cos\phi + d^2 - c_m^2(d^2-1)\bigr)\\
&\quad\ge (s_m^2(d^2-1) - \sin^2\theta)\bigl(\cos^2\theta + 2d\cos\theta\cos\phi + d^2 - c_m^2(d^2-1)\bigr).
\end{aligned}$$

Therefore, it is sufficient to prove

$$0 < s_m^2(d^2-1) - \sin^2\theta, \tag{4-1}$$

$$0 < \cos^2\theta + 2d\cos\theta\cos\phi + d^2 - c_m^2(d^2-1). \tag{4-2}$$

In order to prove these inequalities, we need to use the lower bound on $d$. Using $m \geq 2$ and $d \geq 3/(4s^2)$, we have

$$(1-c_m)d \geq (1-c_2)d = 2s^2 d \geq 3/2. \tag{4-3}$$

We also use $s_m^2 = 1 - c_m^2 = (1-c_m)(1+c_m)$ and $c_m \geq 0$ (the latter uses $m \leq n/2$).

First, we consider (4-1):

$$\begin{aligned} s_m^2(d^2-1) - \sin^2\theta &= \frac{1+c_m}{1-c_m}((1-c_m)d)^2 - 2 + c_m^2 + \cos^2\theta \\ &\geq ((1-c_m)d)^2 - 2 \\ &\geq 1/4, \end{aligned}$$

where the last inequality follows from (4-3). This proves (4-1).

Now consider (4-2):

$$\begin{aligned} \cos^2\theta + 2d\cos\theta\cos\phi + d^2 - c_m^2(d^2-1) &= \frac{(d(1-c_m)+\cos\theta\cos\phi)^2 + \cos^2\theta\sin^2\phi}{1-c_m} \\ &\quad + \frac{c_m}{1-c_m}\big((d(1-c_m))^2 - \cos^2\theta\big) + c_m^2 \\ &\geq \frac{c_m}{1-c_m}(9/4 - \cos^2\theta) \\ &> 0. \end{aligned}$$

Again we used (4-3). This proves (4-2) and so establishes the result. □

Propositions 4.6 and 4.8 complete the proof of Theorem 4.3(3). It remains to prove Theorem 4.3(2). That is, we must consider the intersection of $\mathcal{B}_k$ and $\mathcal{B}_{k\pm 2}$.

Consider $\mathcal{B}_1 \cap \mathcal{B}_{-1}$. We claim that the fixed point of $I_3I_1I_2I_3$ (that is $I_3(o)$) lies on $\mathcal{B}_1 \cap \mathcal{B}_{-1}$. The bisector $\mathcal{B}_1$ consists of all points equidistant from $o$ and $A_1(o) = I_2I_3I_2(o) = I_2I_3(o)$. We have

$$\rho\big(I_3(o),\, I_2I_3(o)\big) = \rho\big(o,\, I_3I_2I_3(o)\big) = \rho\big(o,\, I_2I_3(o)\big).$$

The first equality follows since $I_3$ is an isometry and the second since $I_3I_2I_3 = I_2I_3I_2$ and $I_2(o) = o$. Thus $I_3(o)$ lies on $\mathcal{B}_1$. A similar argument shows

$$\rho\big(I_3(o),\, I_1I_3(o)\big) = \rho\big(o,\, I_1I_3(o)\big).$$

and so $I_3(o)$ lies on $\mathcal{B}_{-1}$ as well. Thus $\mathcal{B}_1 \cap \mathcal{B}_{-1}$ is nonempty, which can be seen in Figure 1. By symmetry, this comment also applies to the intersection of $\mathcal{B}_k$ and $\mathcal{B}_{k\pm 2}$. We must show that this intersection never contributes a ridge of $D$.

**Proposition 4.9.** *Suppose that $3/(4s^2) \leq d \leq c/(1-c)$. For each $-n+1 \leq k \leq n$, all points of $\mathcal{B}_k \cap \mathcal{B}_{k\pm 2}$ lie in the halfspace bounded by $\mathcal{B}_{k\pm 1}$ not containing $o$.*

*Proof.* Using Lemma 4.2 as before, it suffices to consider $\mathcal{B}_1$ and $\mathcal{B}_{-1}$. We need to show that all points of $\mathcal{B}_1 \cap \mathcal{B}_{-1}$ lie in the halfspace closer to $I_3(o)$ than to $o$.

Suppose that $p = [z, w, 1]^t$ lies on $\mathcal{B}_1 \cap \mathcal{B}_{-1}$. Using Proposition 4.7(ii) with $m = 0$, and using (3-3) to write $c(a^2 - b^2) = d(d-1)$, we find

$$z = \frac{sa(\cos\theta e^{i\phi} + d) - cb\sin\theta e^{i\phi}}{sd(d-1)}, \tag{4-4}$$

$$w = \frac{sbi(\cos\theta e^{i\phi} + d) - cai\sin\theta e^{i\phi}}{sd(d-1)}. \tag{4-5}$$

Note that we used (3-3) to simplify the denominator.

The point $p = [z, w, 1]^t$ lies in the halfspace closer to $I_3(o)$ than to $o$ if and only if $1 > |za - wbi - d|$. We want to give this inequality in terms of $\theta$, $\phi$ and $d$. Suppose $z$ and $w$ satisfy (4-4) and (4-5) and consider $za - wbi - d$:

$$\begin{aligned} za - wbi - d &= \frac{sa^2(\cos\theta e^{i\phi} + d) - cab\sin\theta e^{i\phi}}{sd(d-1)} \\ &\qquad\qquad + \frac{sb^2(\cos\theta e^{i\phi} + d) - cab\sin\theta e^{i\phi}}{sd(d-1)} - d \\ &= \frac{s(a^2+b^2)\cos\theta e^{i\phi}}{sd(d-1)} - \frac{2cab\sin\theta e^{i\phi}}{sd(d-1)} + \frac{s(a^2+b^2)d}{sd(d-1)} - d \\ &= \frac{s(d^2-1)\cos\theta e^{i\phi}}{sd(d-1)} - \frac{2cab\sin\theta e^{i\phi}}{sd(d-1)} + \frac{s(d^2-1)d}{sd(d-1)} - d \\ &= \frac{(d+1)\cos\theta e^{i\phi}}{d} - \frac{\sqrt{c^2(d+1)^2 - d^2}\,\sin\theta e^{i\phi}}{sd} + 1. \end{aligned}$$

Therefore,

$$\begin{aligned} |za-wbi-d|^2 - 1 &= \frac{(d+1)^2\cos^2\theta}{d^2} + \frac{c^2(d+1)^2\sin^2\theta}{s^2d^2} - \frac{\sin^2\theta}{s^2} \\ &\quad - \frac{2(d+1)\sqrt{c^2(d+1)^2-d^2}\,\cos\theta\sin\theta}{sd^2} \\ &\quad\quad + \frac{2(d+1)\cos\theta\cos\phi}{d} - \frac{2\sqrt{c^2(d+1)^2-d^2}\,\sin\theta\cos\phi}{sd}. \end{aligned}$$

Arguing as in the proof of Proposition 4.8, we have

$$\begin{aligned} &|z|^2 + |w|^2 - 1 \\ &\quad = \left|\frac{sa(\cos\theta e^{i\phi} + d) - cb\sin\theta e^{i\phi}}{sd(d-1)}\right|^2 + \left|\frac{sbi(\cos\theta e^{i\phi} + d) - cai\sin\theta e^{i\phi}}{sd(d-1)}\right|^2 - 1 \end{aligned}$$

$$
\begin{aligned}
&= \frac{s^2(a^2+b^2)|\cos\theta e^{i\phi}+d|^2}{s^2d^2(d-1)^2} + \frac{c^2(a^2+b^2)\sin^2\theta}{s^2d^2(d-1)^2} - 1 \\
&\qquad + \frac{isc(2abi)(2\cos\theta\sin\theta + 2d\sin\theta\cos\phi)}{s^2d^2(d-1)^2} \\
&= \frac{(d+1)\cos^2\theta}{d^2(d-1)} + \frac{2(d+1)\cos\theta\cos\phi}{d(d-1)} + \frac{d+1}{d-1} + \frac{c^2(d+1)\sin^2\theta}{s^2d^2(d-1)} - 1 \\
&\qquad - \frac{2\sqrt{c^2(d+1)^2-d^2}\,\cos\theta\sin\theta}{sd^2(d-1)} - \frac{2\sqrt{c^2(d+1)^2-d^2}\,\sin\theta\cos\phi}{sd(d-1)} \\
&= \frac{2}{d-1} + \frac{(d+1)\cos^2\theta}{d^2(d-1)} + \frac{c^2(d+1)\sin^2\theta}{s^2d^2(d-1)} - \frac{2\sqrt{c^2(d+1)^2-d^2}\,\cos\theta\sin\theta}{sd^2(d-1)} \\
&\qquad + \frac{2(d+1)\cos\theta\cos\phi}{d(d-1)} - \frac{2\sqrt{c^2(d+1)^2-d^2}\,\sin\theta\cos\phi}{sd(d-1)}.
\end{aligned}
$$

Now we eliminate $\cos\phi$ using the equation for $|za-wbi-d|^2$ derived above:

$$
\begin{aligned}
|z|^2+|w|^2-1 &= \frac{1}{d-1}(|za-wbi-d|^2-1) + \frac{2\cos^2\theta}{d-1} + \frac{2\sin^2\theta}{d-1} \\
&\qquad + \frac{(d+1)\cos^2\theta}{d^2(d-1)} + \frac{c^2(d+1)\sin^2\theta}{s^2d^2(d-1)} - \frac{2\sqrt{c^2(d+1)^2-d^2}\,\cos\theta\sin\theta}{sd^2(d-1)} \\
&\qquad - \frac{(d+1)^2\cos^2\theta}{d^2(d-1)} - \frac{c^2(d+1)^2\sin^2\theta}{s^2d^2(d-1)} + \frac{\sin^2\theta}{s^2(d-1)} \\
&\qquad + \frac{2(d+1)\sqrt{c^2(d+1)^2-d^2}\,\cos\theta\sin\theta}{sd^2(d-1)} \\
&= \frac{1}{d-1}(|za-wbi-d|^2-1) \\
&\qquad + \frac{1}{d}\left(\cos\theta + \frac{\sqrt{c^2(d+1)^2-d^2}\,\sin\theta}{s(d-1)}\right)^2 + \frac{(4s^2d-3)\sin^2\theta}{s^2(d-1)^2}.
\end{aligned}
$$

Since the last two terms are nonnegative, all points $p=[z,\,w,\,1]^t$ with $z$ and $w$ given by (4-4) and (4-5) and that satisfy $|z|^2+|w|^2<1$ must also satisfy $|za-wbi-d|<1$. Geometrically, this means that all points in $\boldsymbol{H}^2_{\mathbb{C}}$ that are on $\mathcal{B}_1\cap\mathcal{B}_{-1}$ are in the halfspace closer to $I_3(o)$ than to $o$. This proves the result. $\square$

This completes the proof of Theorem 4.3.

## Acknowledgements

This research was carried out while John Parker and Jieyan Wang were visiting the Department of Mathematics at Hunan University. We would like to thank Yueping Jiang for making this visit possible. We would like to thank the referee for his comments which improved a previous version of this paper.

## References

[Brehm 1990] U. Brehm, "The shape invariant of triangles and trigonometry in two-point homogeneous spaces", *Geom. Dedicata* **33**:1 (1990), 59–76. MR 91c:53048 Zbl 0695.53038

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

[Deraux and Falbel 2015] M. Deraux and E. Falbel, "Complex hyperbolic geometry of the figure-eight knot", *Geom. Topol.* **19**:1 (2015), 237–293. MR 3318751 Zbl 06413574

[Deraux et al. 2015] M. Deraux, J. R. Parker, and J. Paupert, "New non-arithmetic complex hyperbolic lattices", *Invent. Math.* (online publication May 2015).

[Falbel and Wang 2014] E. Falbel and J. Wang, "Branched spherical CR structures on the complement of the figure-eight knot", *Michigan Math. J.* **63**:3 (2014), 635–667. MR 3255694 Zbl 1300.32034

[Falbel et al. 2015] E. Falbel, P.-V. Koseleff, and F. Rouillier, "Representations of fundamental groups of 3-manifolds into PGL(3, $\mathbb{C}$): Exact computations in low complexity", *Geom. Dedicata* **177**:1 (2015), 229–255. MR 3370032 Zbl 06468790

[Goldman and Parker 1992] W. M. Goldman and J. R. Parker, "Complex hyperbolic ideal triangle groups", *J. Reine Angew. Math.* **425** (1992), 71–86. MR 93c:20076 Zbl 0739.53055

[Grossi 2007] C. H. Grossi, "On the type of triangle groups", *Geom. Dedicata* **130** (2007), 137–148. MR 2008k:20092 Zbl 1139.53005

[Mostow 1980] G. D. Mostow, "On a remarkable class of polyhedra in complex hyperbolic space", *Pacific J. Math.* **86**:1 (1980), 171–276. MR 82a:22011 Zbl 0456.22012

[Parker and Will 2015a] J. R. Parker and P. Will, "Complex hyperbolic free groups with many parabolic elements", pp. 327–348 in *Groups, geometry and dynamics*, edited by C. S. Aravinda et al., Contemp. Math. **639**, Amer. Math. Soc., Providence, RI, 2015.

[Parker and Will 2015b] J. R. Parker and P. Will, "A complex hyperbolic Riley slice", preprint, 2015. arXiv 1510.01505

[Phillips 1992] M. B. Phillips, "Dirichlet polyhedra for cyclic groups in complex hyperbolic space", *Proc. Amer. Math. Soc.* **115**:1 (1992), 221–228. MR 93a:32042 Zbl 0768.53033

[Pratoussevitch 2005] A. Pratoussevitch, "Traces in complex hyperbolic triangle groups", *Geom. Dedicata* **111**:1 (2005), 159–185. MR 2006d:32036 Zbl 1115.32015

[Schwartz 2001a] R. E. Schwartz, "Degenerating the complex hyperbolic ideal triangle groups", *Acta Math.* **186**:1 (2001), 105–154. MR 2002e:20079 Zbl 0998.53050

[Schwartz 2001b] R. E. Schwartz, "Ideal triangle groups, dented tori, and numerical analysis", *Ann. of Math.* (2) **153**:3 (2001), 533–598. MR 2002j:57031 Zbl 1055.20040

[Schwartz 2002] R. E. Schwartz, "Complex hyperbolic triangle groups", pp. 339–349 in *Proceedings of the international congress of mathematicians, II: Invited lectures* (Beijing, 2002), edited by T. T. Li et al., Higher Ed. Press, Beijing, 2002. MR 2004b:57002 Zbl 1022.53034

[Schwartz 2005] R. E. Schwartz, "A better proof of the Goldman–Parker conjecture", *Geom. Topol.* **9** (2005), 1539–1601. MR 2006j:20063 Zbl 1098.20034

[Schwartz 2007] R. E. Schwartz, *Spherical CR geometry and Dehn surgery*, Annals of Mathematics Studies **165**, Princeton Univ. Press, 2007. MR 2009a:32047 Zbl 1116.57016

[Thompson 2010] J. M. Thompson, *Complex hyperbolic triangle groups*, Ph.D. thesis, Durham University, 2010, available at http://core.ac.uk/download/pdf/85700.pdf.

Received January 27, 2015. Revised June 18, 2015.

JOHN R. PARKER
DEPARTMENT OF MATHEMATICAL SCIENCES
DURHAM UNIVERSITY
DURHAM
DH1 3LE
UNITED KINGDOM
j.r.parker@durham.ac.uk

JIEYAN WANG
ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE
CHINESE ACADEMY OF SCIENCES
BEIJING, 100190
CHINA
jywang@hnu.edu.cn

BAOHUA XIE
COLLEGE OF MATHEMATICS AND ECONOMETRICS
HUNAN UNIVERSITY
CHANGSHA, 410082
CHINA
xiexbh@hnu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.455

# TOPOLOGICAL ASPECTS OF HOLOMORPHIC MAPPINGS OF HYPERQUADRICS FROM $\mathbb{C}^2$ TO $\mathbb{C}^3$

MICHAEL REITER

**In this article we deduce some topological results concerning holomorphic mappings of hyperquadrics under biholomorphic equivalence. We study the class $\mathcal{F}$ of so-called nondegenerate and transversal holomorphic mappings locally sending the sphere in $\mathbb{C}^2$ to a Levi-nondegenerate hyperquadric in $\mathbb{C}^3$, which contains the most interesting mappings. We show that from a topological point of view there is a major difference when the target is the sphere or the hyperquadric with signature (2, 1). In the first case, $\mathcal{F}$ modulo the group of automorphisms is discrete, in contrast to the second case, where this property fails to hold. Furthermore, we study some basic properties such as freeness and properness of the action on $\mathcal{F}$ of automorphisms fixing a given point to obtain a structural result for a particularly interesting subset of $\mathcal{F}$.**

## 1. Introduction and results

We study holomorphic mappings between the sphere $\mathbb{S}^2 \subset \mathbb{C}^2$ and the hyperquadric $\mathbb{S}^3_\varepsilon \subset \mathbb{C}^3$, which for $\varepsilon = \pm 1$ is given by

$$\mathbb{S}^3_\pm := \big\{(z_1, z_2, z_3) \in \mathbb{C}^3 \mid |z_1|^2 + |z_2|^2 \pm |z_3|^2 = 1\big\},$$

so that $\mathbb{S}^3_+ = \mathbb{S}^3$ is the sphere in $\mathbb{C}^3$. Faran [1982] classified holomorphic mappings between spheres in $\mathbb{C}^2$ and $\mathbb{C}^3$ and Lebl [2011] classified mappings sending $\mathbb{S}^2$ to $\mathbb{S}^3_-$. In [Reiter 2015] we give a new CR-geometric approach to reprove Faran's and Lebl's results in a unified manner. Let us introduce the following equivalence relation. For $k = 1, 2$ let $H_k : U_k \to \mathbb{C}^3$ be a holomorphic mapping where $U_k$ is an open and connected neighborhood of $p_k \in \mathbb{S}^2$ and $H_k(U_k \cap \mathbb{S}^2) \subset \mathbb{S}^3_\varepsilon$. We say $H_1$ is *equivalent* to $H_2$ if there exist automorphisms $\phi$ of $\mathbb{S}^2$ and $\phi'$ of $\mathbb{S}^3_\varepsilon$ such that $H_2 = \phi' \circ H_1 \circ \phi^{-1}$.

The author was supported by the FWF, projects Y377 and I382, and QNRF, project NPRP 7-511-1-098.
*MSC2010:* 32H02, 32V30, 57S05, 57S25, 58D19.
*Keywords:* holomorphic mappings, hyperquadrics, nondegeneracy, transversality, automorphism groups, group actions, quotient space, principal fiber bundle.

**Theorem 1.1** [Reiter 2015, Theorem 1.3]. *Given $p \in \mathbb{S}^2$, let $U \subset \mathbb{C}^2$ be an open and connected neighborhood of $p$ and $H : U \to \mathbb{C}^3$ a nonconstant holomorphic mapping satisfying $H(U \cap \mathbb{S}^2) \subset \mathbb{S}^3_\varepsilon$. Then $H$ is equivalent to exactly one of the following maps*:

(i) $H_1^\varepsilon(z, w) = (z, w, 0)$,

(ii) $H_2^\varepsilon(z, w) = \left(z^2, \dfrac{(1-\varepsilon+z(1+\varepsilon))w}{\sqrt{2}}, w^2\right)$,

(iii) $H_3^\varepsilon(z, w) = \left(z, \dfrac{(1-\varepsilon+z^2(1+\varepsilon))w}{2z}, \dfrac{(1-\varepsilon+z(1+\varepsilon))w^2}{2z}\right)$,

(iv) $H_4^\varepsilon(z, w) = \dfrac{\left(4z^3, (3(1-\varepsilon)+(1+3\varepsilon)w^2)w, \sqrt{3}(1-\varepsilon+2(1+\varepsilon)w+(1-\varepsilon)w^2)z\right)}{1+3\varepsilon+3(1-\varepsilon)w^2}$.

*Additionally, for $\varepsilon = -1$, we have*

(v) $H_5(z, w) = \left(\dfrac{(2+\sqrt{2}z)z}{1+\sqrt{2}z+w}, w, \dfrac{(1+\sqrt{2}z-w)z}{1+\sqrt{2}z+w}\right)$,

(vi) $H_6(z, w) = \dfrac{\left((1-w)z, 1+w-w^2, (1+w)z\right)}{1-w-w^2}$,

(vii) $H_7(z, w) = \left(1, h(z, w), h(z, w)\right)$ *for some nonconstant holomorphic function* $h : U \to \mathbb{C}$.

In fact, we study holomorphic mappings between the Heisenberg hypersurface $\mathbb{H}^2 \subset \mathbb{C}^2$ and $\mathbb{H}^3_\varepsilon$, where $\mathbb{H}^3_+ = \mathbb{H}^3$ is the Heisenberg hypersurface in $\mathbb{C}^3$. The hypersurfaces $\mathbb{H}^2$ and $\mathbb{H}^3_\varepsilon$ are biholomorphic to $\mathbb{S}^2$ and $\mathbb{S}^3_\varepsilon$ respectively, except one point, and are given by

$$\mathbb{H}^2 = \left\{(z, w) \in \mathbb{C}^2 \mid \operatorname{Im} w = |z|^2\right\},$$
$$\mathbb{H}^3_\varepsilon = \left\{(z_1', z_2', w') \in \mathbb{C}^3 \mid \operatorname{Im} w' = |z_1'|^2 + \varepsilon|z_2'|^2\right\}.$$

We denote by $\mathcal{F}$ the class of germs of 2-nondegenerate transversal mappings sending a small piece of $\mathbb{H}^2$ to $\mathbb{H}^3_\varepsilon$. $\mathcal{F}$ is introduced in more detail in Definition 2.5 below. This is, in some sense, the most natural and interesting class of mappings when studying holomorphic mappings between $\mathbb{H}^2$ to $\mathbb{H}^3_\varepsilon$. From [Reiter 2015] we know that $\mathcal{F}$ consists of mappings belonging to the orbits of the maps listed in (ii)–(vi) of Theorem 1.1 with respect to the equivalence relation of automorphisms introduced above, after composing with an appropriate Cayley transform.

For a germ of a real-analytic CR-submanifold $(M, p)$ of $\mathbb{C}^N$, we write $\operatorname{Aut}_p(M, p)$ for germs of real-analytic CR-diffeomorphisms fixing $p$, which we refer to as *isotropies* of $(M, p)$. Let us denote by $G_0 := \operatorname{Aut}_0(\mathbb{H}^2, 0) \times \operatorname{Aut}_0(\mathbb{H}^3_\varepsilon, 0)$ the direct product of the groups of isotropies of $(\mathbb{H}^2, 0)$ and $(\mathbb{H}^3_\varepsilon, 0)$, which we introduce in Definition 2.3 below in more detail.

After showing that $\pi : \mathcal{F} \to \mathcal{F}/G_0$ is continuous, we obtain the following results.

**Theorem 1.2.** *The quotient topology $\tau_Q$ on $\mathcal{F}/G_0$ coincides with the induced topology $\tau_J$ of $\mathcal{F}$, which carries the topology induced by the jet space $J_0^3(\mathbb{H}^2, \mathbb{H}^3_\varepsilon)$.*

In the next theorem we equip $\mathcal{F}$ with the topology $\tau_J$ induced by the jet space $J_0^3(\mathbb{H}^2, \mathbb{H}^3_\varepsilon)$, the automorphism groups carry the topology of the 2-jet group (see [Baouendi et al. 1997] for more details), and the quotient space $X$ of $\mathcal{F}$ with respect to the equivalence relation of Theorem 1.1 carries the quotient topology. For more details on the different topologies we use, we refer to Section 2 below.

**Theorem 1.3.** *The quotient space $X$ of $\mathcal{F}$ with respect to the equivalence relation of automorphisms of $\mathbb{H}^2$ and $\mathbb{H}^3_\varepsilon$ is discrete for $\varepsilon = +1$ and not Hausdorff for $\varepsilon = -1$.*

The above result was not known before and shows one major difference between holomorphic mappings from the sphere in $\mathbb{C}^2$ to the sphere in $\mathbb{C}^3$ and to the hyperquadric with signature (2, 1) in $\mathbb{C}^3$. Furthermore, we study the action of $G_0$ on $\mathcal{F}$ given by $G_0 \times \mathcal{F} \to \mathcal{F}$, $(\phi, \phi', H) \mapsto \phi' \circ H \circ \phi^{-1}$. The action is called *proper* if the associated map $(\phi, \phi', H) \mapsto (H, \phi' \circ H \circ \phi^{-1})$ is a proper map, such that the following result holds:

**Theorem 1.4.** *The mapping $N : G_0 \times \mathcal{F} \to \mathcal{F}$ given by $N(\phi, \phi', H) := \phi' \circ H \circ \phi^{-1}$ is a proper action.*

We write $\mathfrak{F} \subset \mathcal{F}$ for the set of maps which have trivial stabilizers given below in Lemma 3.1. Based on the above result we obtain the following theorem concerning the real-analytic structure of $\mathfrak{F}$, where $\Pi : \mathfrak{F} \to \mathfrak{N}$ denotes the normalization map induced by the mapping $N$, and $\mathfrak{N}$ denotes a particular set of representatives of the quotient $\mathfrak{F}/G_0$ defined in Lemma 3.1 below.

**Theorem 1.5.** *If $\varepsilon = +1$ then $\Pi : \mathfrak{F} \to \mathfrak{F}/G_0$ is a real-analytic principal fiber bundle with structure group $G_0$. If $\varepsilon = -1$ then $\mathfrak{F}$ is locally mapped to $G_0 \times \mathfrak{N}$ via local real-analytic diffeomorphisms. In particular, $\mathfrak{F}$ is not a smooth manifold.*

Note that the second part of Theorem 1.5 stands in contrast to the case of maps in $\mathrm{Aut}_p(M, p)$. Assuming some nondegeneracy conditions for certain germs of real-analytic CR-submanifolds $(M, p)$, such as Levi-nondegeneracy, it is known that $\mathrm{Aut}_p(M, p)$ admits a manifold structure (see [Baouendi et al. 1997; 1999; 2004; Kowalski 2005; Kim and Zaitsev 2005; Lamel and Mir 2007; Lamel et al. 2008; Juhlin and Lamel 2013]). To prove Theorem 1.5 we use a real-analytic version of the so-called local slice theorem for free and proper actions. For proper smooth actions of noncompact Lie groups the first proof of the local slice theorem was given in [Palais 1961, 2.2.2 Proposition]. In the real-analytic setting a global slice theorem was proved by [Heinzner et al. 1996, Section VI] and [Illman and Kankaanrinta 2000, Theorem 0.6].

We organize this paper as follows. We introduce the necessary notations, tools and results in Section 2. In the following sections we study properties of the action of the group of isotropies on $\mathcal{F}$ and in Section 5 we investigate the connectedness of $\mathcal{F}$ and discreteness of the quotient space. Using these results, in Section 7 we obtain some structural and topological information of $\mathfrak{F}$ and $\mathfrak{F}/G_0$. Finally, in Section 8 we study different normal forms with respect to isotropies. This article is based on the author's thesis [Reiter 2014] at the University of Vienna. Some computations are carried out with *Mathematica 7.0.1.0* [Wolfram 2008].

## 2. Preliminaries

**Definition 2.1.** We fix coordinates $(z, w) = (z_1, \dots, z_n, w) \in \mathbb{C}^{n+1}$. For a germ $h : (\mathbb{C}^{n+1}, 0) \to (\mathbb{C}, 0)$ of a holomorphic function $h(z, w) = \sum_{\alpha,\beta} a_{\alpha\beta} z^\alpha w^\beta$, we write $\bar{h}(\bar{z}, \bar{w}) := \overline{h(z, w)} = \sum_{\alpha,\beta} \bar{a}_{\alpha\beta} \bar{z}^\alpha \bar{w}^\beta$ for the complex conjugate of $h$. Derivatives of $h$ with respect to $z$ or $w$ are denoted by

$$h_{z^\alpha w^\beta}(0) := \frac{\partial^{|\alpha|+|\beta|} h}{\partial z^\alpha \partial w^\beta}(0).$$

For $n \geq 1$ and a germ of a map $H : (\mathbb{C}^{n+1}, 0) \to (\mathbb{C}^{n'+1}, 0)$ with components $H = \big(f_1, \dots, f_{n'}, g\big)$, we write $H_{z^\alpha w^\beta}(0) = \big(f_{1 z^\alpha w^\beta}(0), \dots, f_{n' z^\alpha w^\beta}(0), g_{z^\alpha w^\beta}(0)\big)$.

***Classes of maps, automorphisms and equivalence relations.***

**Definition 2.2.** We write $\mathcal{H}(p; p') := \{H : (\mathbb{C}^N, p) \to (\mathbb{C}^{N'}, p') \mid H \text{ holomorphic}\}$ for the *set of germs of holomorphic mappings from* $(\mathbb{C}^N, p)$ *to* $(\mathbb{C}^{N'}, p')$. For germs of real-analytic hypersurfaces $(M, p) \subset \mathbb{C}^N$ and $(M', p') \subset \mathbb{C}^{N'}$, we denote by

$$\mathcal{H}(M, p; M', p') := \{H \in \mathcal{H}(p; p') \mid H(M \cap U) \subset M', \text{ for } U \text{ a neighborhood of } p\},$$

the *set of germs of holomorphic mappings from* $(M, p)$ *to* $(M', p')$.

**Definition 2.3.** (i) We denote the collection of germs of locally real-analytic CR-diffeomorphisms of $(M, p)$ by

$$\mathrm{Aut}(M, p) := \big\{H : (\mathbb{C}^N, p) \to \mathbb{C}^N \mid H \text{ holomorphic}, H(M) \subset M, \det(H'(p)) \neq 0\big\}$$

and the *group of isotropies of* $(M, p)$ *fixing* $p$ by

$$\mathrm{Aut}_p(M, p) := \{H \in \mathrm{Aut}(M, p) \mid H(p) = p\}.$$

We write $G_0 := \mathrm{Aut}_0(\mathbb{H}^2, 0) \times \mathrm{Aut}_0(\mathbb{H}^3_\varepsilon, 0)$ and refer to elements of $G_0$ as *isotropies of* $(\mathbb{H}^2, 0)$ *and* $(\mathbb{H}^3_\varepsilon, 0)$.

(ii) We write $\mathbb{R}^+ := \{x \in \mathbb{R} \mid x > 0\}$ for the positive real numbers, denote the unit circle in $\mathbb{C}$ by $\mathbb{S}^1 := \{e^{it} \mid 0 \leq t < 2\pi\}$ and set $\Gamma := \mathbb{R}^+ \times \mathbb{R} \times \mathbb{S}^1 \times \mathbb{C}$. For an element $\sigma_\gamma \in \mathrm{Aut}_0(\mathbb{H}^2, 0)$ we denote $\gamma = (\lambda, r, u, c) \in \Gamma$ and write

$$\sigma_\gamma(z, w) := \frac{(\lambda u(z + cw), \lambda^2 w)}{1 - 2i\bar{c}z + (r - i|c|^2)w}. \tag{2-1}$$

(iii) We define, for $\theta = \pm 1$ if $\varepsilon = -1$ and $\theta = +1$ if $\varepsilon = +1$,

$$\mathcal{S}^2_{\varepsilon,\theta} := \big\{a' = (a'_1, a'_2) \in \mathbb{C}^2 \mid |a'_1|^2 + \varepsilon |a'_2|^2 = \theta\big\}, \tag{2-2}$$

and let

$$U' := \begin{pmatrix} u'a'_1 & -\varepsilon u' a'_2 \\ \bar{a}'_2 & \bar{a}'_1 \end{pmatrix}, \quad u' \in \mathbb{S}^1,\ a' = (a'_1, a'_2) \in \mathcal{S}^2_{\varepsilon,\theta}. \tag{2-3}$$

We set $\Gamma' := \mathbb{R}^+ \times \mathbb{R} \times \mathbb{S}^1 \times \mathcal{S}^2_{\varepsilon,\theta} \times \mathbb{C}^2$ to denote elements $\sigma'_{\gamma'} \in \mathrm{Aut}_0(\mathbb{H}^3_\varepsilon, 0)$ via $\gamma' = (\lambda', r', u', a', c') \in \Gamma'$, where $c' = (c'_1, c'_2)$:

$$\sigma'_{\gamma'}(z', w') := \frac{(\lambda' U'\,{}^t(z' + c'w'), \theta\lambda'^2 w')}{1 - 2\mathrm{i}(\bar{c}'_1 z'_1 + \varepsilon \bar{c}'_2 z'_2) + \big(r' - \mathrm{i}(|c'_1|^2 + \varepsilon |c'_2|^2)\big) w'}. \tag{2-4}$$

(iv) We call elements of $\Gamma \times \Gamma'$ *standard parameters*. If the standard parameters $(\gamma, \gamma') \in \Gamma \times \Gamma'$ are chosen such that $(\sigma_\gamma, \sigma'_{\gamma'}) = (\mathrm{id}_{\mathbb{C}^2}, \mathrm{id}_{\mathbb{C}^3})$, we say the standard parameters are *trivial*.

**Definition 2.4.** For $G, H \in \mathcal{H}(M, p; M', p')$, we define an equivalence relation

$$G \sim H :\Leftrightarrow \exists (\phi, \phi') \in \mathrm{Aut}_p(M, p) \times \mathrm{Aut}_{p'}(M', p') : G = \phi' \circ H \circ \phi^{-1}.$$

The equivalence classes in $\mathcal{H}(M, p; M', p')/\!\sim$ are denoted by

$$[F] := \{G \in \mathcal{H}(M, p; M', p') \mid G \sim F\}.$$

In the case where $(p, p') = (0, 0)$ and $(M, M') = (\mathbb{H}^2, \mathbb{H}^3_\varepsilon)$, we call the above relation *isotropic equivalence* and write $O_0(H)$ for the orbit of a map $H$, called the *isotropic orbit of H*.

***The class $\mathcal{F}$, the normal form $\mathcal{N}$ and its classification.*** In [Reiter 2015] we introduced the following class of mappings, which are 2-nondegenerate and transversal. These mappings represent the immersive maps, which are not equivalent to the linear embedding (see [Reiter 2015, Proposition 2.16]).

**Definition 2.5.** For a neighborhood $U \subset \mathbb{C}^2$ of 0, define $\mathcal{F}(U)$ to be the set of holomorphic mappings $H = (f_1, f_2, g)$, with $H(U \cap \mathbb{H}^2) \subset \mathbb{H}^3_\varepsilon$, which satisfy $H(0) = 0$, $f_{1z}(0) f_{2z^2}(0) - f_{2z}(0) f_{1z^2}(0) \neq 0$ and $g_w(0) > 0$. Define $\mathcal{F}$ to be the set of germs $H$, such that $H \in \mathcal{F}(U)$ for some neighborhood $U \subset \mathbb{C}^2$ of 0.

**Proposition 2.6** [Reiter 2015, Proposition 3.1]. *Let $H \in \mathcal{F}$. Then there exist isotropies $(\sigma, \sigma') \in G_0$ such that $\widehat{H} := \sigma' \circ H \circ \sigma^{-1}$ satisfies $\widehat{H}(0) = 0$ and the following conditions*:

(i) $\widehat{H}_z(0) = (1, 0, 0)$, (iii) $\hat{f}_{2z^2}(0) = 2$, (vi) $\mathrm{Re}\big(\hat{g}_{w^2}(0)\big) = 0$,

(ii) $\widehat{H}_w(0) = (0, 0, 1)$, (iv) $\hat{f}_{2zw}(0) = 0$, (vii) $\mathrm{Re}\big(\hat{f}_{2z^2 w}(0)\big) = 0$.

(v) $\hat{f}_{1w^2}(0) = |\hat{f}_{1w^2}(0)| \geq 0$, *A holomorphic mapping of $\mathcal{F}$ satisfying the above conditions is called a normalized mapping. The set of normalized mappings is denoted by $\mathcal{N}$.*

**Remark 2.7.** A mapping $H \in \mathcal{N}$ necessarily satisfies the following conditions (see [Reiter 2015, Remark 3.4]):

(i) $H(0) = (0, 0, 0)$,
(ii) $H_z(0) = (1, 0, 0)$,
(iii) $H_w(0) = (0, 0, 1)$,
(iv) $H_{z^2}(0) = (0, 2, 0)$,
(v) $H_{zw}(0) = \left(\frac{\mathrm{i}\varepsilon}{2}, 0, 0\right)$,
(vi) $H_{w^2}(0) = (|f_{1w^2}(0)|, f_{2w^2}(0), 0)$,
(vii) $H_{z^2w}(0) = \left(4\mathrm{i}|f_{1w^2}(0)|, \mathrm{i}\,\mathrm{Im}(f_{2z^2w}(0)), 0\right)$.

We classify all mappings belonging to $\mathcal{N} \simeq \mathcal{F}/G_0$ in [Reiter 2015].

**Theorem 2.8** [Reiter 2015, Theorem 4.1]. *The set $\mathcal{N}$ consists of the following mappings, where $s \geq 0$:*

$$
\begin{aligned}
G_1^{\varepsilon}(z, w) &:= \left(2z(2+\mathrm{i}\varepsilon w), 4z^2, 4w\right)/(4-w^2),\\
G_{2,s}^{\varepsilon}(z, w) &:= \left(4z - 4\varepsilon s z^2 + \mathrm{i}(\varepsilon - s^2)zw + sw^2,\right.\\
&\qquad\qquad \left. 4z^2 + s^2w^2, w(4 - 4\varepsilon s z - \mathrm{i}(\varepsilon + s^2)w)\right)\\
&\qquad\qquad\qquad /\left(4 - 4\varepsilon s z - \mathrm{i}(\varepsilon + s^2)w - 2\mathrm{i}szw - \varepsilon s^2 w^2\right),\\
G_{3,s}^{\varepsilon}(z, w) &:= \left(256\varepsilon z + 96\mathrm{i}zw + 64\varepsilon s w^2 + 64z^3 + 64\mathrm{i}\varepsilon s z^2 w\right.\\
&\qquad\qquad\qquad - 3(3\varepsilon - 16s^2)zw^2 + 4\mathrm{i}sw^3,\\
&\qquad 256\varepsilon z^2 - 16w^2 + 256sz^3 + 16\mathrm{i}z^2w - 16\varepsilon szw^2 - \mathrm{i}\varepsilon w^3,\\
&\qquad\qquad \left. w(256\varepsilon - 32\mathrm{i}w + 64z^2 - 64\mathrm{i}\varepsilon szw - (\varepsilon + 16s^2)w^2)\right)\\
&\qquad /\left(256\varepsilon - 32\mathrm{i}w + 64z^2 - 192\mathrm{i}\varepsilon szw - (17\varepsilon + 144s^2)w^2\right.\\
&\qquad\qquad\qquad \left. + 32\mathrm{i}\varepsilon z^2 w + 24szw^2 + \mathrm{i}w^3\right).
\end{aligned}
$$

*Each mapping in $\mathcal{N}$ is not isotropically equivalent to any different mapping in $\mathcal{N}$.*

For $\varepsilon = \pm 1$, Figure 1 depicts $\mathcal{N}$ in the parameter space according to Theorem 2.8 (see [Reiter 2015, §4] for more details).

***Associated topologies.*** We deal with the following topologies (see, e.g., [Baouendi et al. 1997]).

**Definition 2.9.** For $K \subset \mathbb{C}^N$ a compact neighborhood of $p \in \mathbb{C}^N$, we denote by $\mathcal{H}_K(p; p')$ the space of holomorphic mappings, defined in a neighborhood of $K$, which map $p \in \mathbb{C}^N$ to $p' \in \mathbb{C}^{N'}$ equipped with the uniform norm on $K$. We equip $\mathcal{H}(p; p')$ with the inductive limit topology with respect to $\mathcal{H}_K(p; p')$, where $K$ is some compact neighborhood of $p$ in $\mathbb{C}^N$. Then for $H, H_n \in \mathcal{H}(p; p')$, we say that $H_n$ converges to $H$ if there exists $K \subset \mathbb{C}^N$ a compact neighborhood of $p$ such that

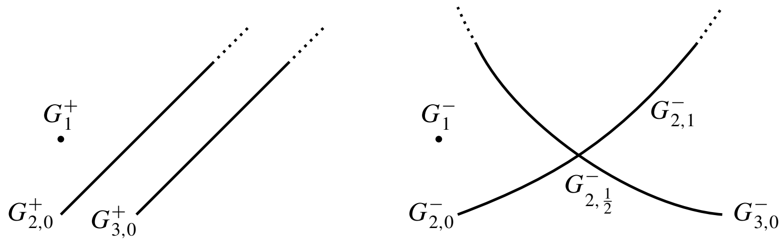


**Figure 1.** $\mathcal{N}$ for $\varepsilon = \pm 1$ in the parameter space.

each $H_n$ is holomorphic in a neighborhood of $K$ and $H_n$ converges uniformly to $H$ on $K$. For $\mathcal{H}(M, p; M', p') \subset \mathcal{H}(p; p')$, we consider the induced topology of $\mathcal{H}(p; p')$ denoted by $\tau_C$.

**Definition 2.10.** Let $Z \in \mathbb{C}^N$ be coordinates in $\mathbb{C}^N$, $H : \mathbb{C}^N \to \mathbb{C}^{N'}$ a holomorphic mapping defined at $p \in \mathbb{C}^N$ and $\alpha \in \mathbb{N}^N$. We denote by $j_p^k H$ the *k-jet of $H$ at $p$* defined as

$$j_p^k H := \Big( \frac{\partial^{|\alpha|} H}{\partial Z^\alpha}(p) : |\alpha| \leq k \Big),$$

and by $J_{p,p'}^k$ the collection of all $k$-jets at $p$ of germs of mappings from $(\mathbb{C}^N, p)$ to $(\mathbb{C}^{N'}, p')$. We set $J_p^k := J_{p,p}^k$ and denote the topology for $J_{p,p'}^k$ by $\tau_J$, which we refer to as the *topology of the jet space*. Let $(M, p) \subset (\mathbb{C}^N, p)$ and $(M', p') \subset (\mathbb{C}^{N'}, p')$ be germs of submanifolds. For $k \in \mathbb{N}$ we denote by $J_q^k(M, p; M', p')$ the *space of k-jets of $\mathcal{H}(M, p; M', p')$ at $q$*. We also define $J_q^k(M, p) := J_q^k(M, p; M, p)$ and $J_0^k(M; M') := J_0^k(M, 0; M', 0)$. We denote by $G_p^k(M, p) \subset J_p^k(M, p)$ the *space of k-jets of* $\mathrm{Aut}_p(M, p)$ *at $p$*.

Note that $J_p^k(M, p; M', p') \subset J_{p,p'}^k$. We identify $J_{p,p'}^k$ with the space of germs of holomorphic polynomial mappings, up to degree $k$, from $\mathbb{C}^N$ to $\mathbb{C}^{N'}$, which map $p \in \mathbb{C}^N$ to $p' \in \mathbb{C}^{N'}$. Thus $J_{p,p'}^k$ can be identified with some $\mathbb{C}^K$, where $K := N'\binom{N+k}{N}$, such that the topology $\tau_J$ for $J_{p,p'}^k$ is induced by the natural topology of $\mathbb{C}^K$.

**Definition 2.11.** We say $\mathcal{K} \subset \mathcal{H}(M, p; M', p')$ admits a *jet parametrization for $\mathcal{K}$ of order $k$* if there exists a mapping $\Psi : \mathbb{C}^N \times \mathbb{C}^K \supset U \to \mathbb{C}^{N'}$, with $K = N'\binom{N+k}{N}$, from above and $U$ an open neighborhood of $\{p\} \times J_p^k(M, p; M', p')$, which is holomorphic in the first $N$ variables and real-analytic in the remaining $K$ variables, such that $F(Z) = \Psi(Z, j_p^k F)$ for all $F \in \mathcal{K}$.

If $\mathcal{K} \subset \mathcal{H}(M, p; M', p')$ admits a jet parametrization of some order $k$, then $\tau_C = \tau_J$, which follows from the real-analyticity in the last $K$ variables. We need the following jet determination result which is an immediate consequence of the normalization and classification of maps in $\mathcal{F}$.

**Corollary 2.12** [Reiter 2015, Corollary 4.8]. *Let $U \subset \mathbb{C}^2$ be a neighborhood of $0$ and $H : U \to \mathbb{C}^3$ a holomorphic mapping. We denote the components of $H$ by $H = (f, g) = (f_1, f_2, g)$ and write $j_0(H) := \{j_0^2(H), f_{z^2 w}(0)\}$. If for $H_1, H_2 \in \mathcal{F}$ the coefficients belonging to $j_0(H_1)$ and $j_0(H_2)$ coincide, then we have $H_1 \equiv H_2$.*

**Remark 2.13.** Based on [Lamel 2001, Proposition 25, Corollary 26–27] we obtain a jet parametrization of order 4 for $\mathcal{K} = \mathcal{F}$ in [Reiter 2015, Lemma 4.3], and by Corollary 2.12 we have that $K = K_0 := 15$. Using Theorem 2.8 and the notation from Corollary 2.12, we identify $\mathcal{F}$ with a subset $\mathfrak{J} \subset \mathbb{C}^{K_0}$ given by $\mathfrak{J} := \{j_0(H) \mid H \in \mathcal{F}\}$, and the topology we use in the sequel for $\mathcal{F}$ is $\tau_J$.

**Definition 2.14.** Let $X$ be a topological space, $Y$ a set and $q : X \to Y$ a surjective mapping. We call the topology on $Y$ induced by $q$ the *quotient topology $\tau_Q$ on $Y$*, where a set $U \subset Y$ is open in $Y$ if $q^{-1}(U)$ is open in $X$.

## 3. The isotropic stabilizer and freeness of the group action on $\mathfrak{F}$

**Lemma 3.1.** *Set $\mathfrak{N} := \mathcal{N} \setminus \{G_1^{\varepsilon}, G_{2,0}^{\varepsilon}, G_{3,0}^{\varepsilon}\}$ and $\mathfrak{F} := \bigcup_{H \in \mathfrak{N}} O_0(H)$. The isotropic stabilizer* $\operatorname{stab}_0(H) := \{(\phi, \phi') \in G_0 \mid \phi' \circ H \circ \phi^{-1} = H\}$ *of $H$ is trivial for $H \in \mathfrak{N}$. Furthermore, we have that* $\operatorname{stab}_0(G_1^{\varepsilon}) = \operatorname{stab}_0(G_{2,0}^{\varepsilon})$ *is homeomorphic to $\mathbb{S}^1$ and* $\operatorname{stab}_0(G_{3,0}^{\varepsilon})$ *is homeomorphic to $\mathbb{Z}_2$.*

*Proof.* Let $H = (f, g) = (f_1, f_2, g) \in \mathcal{N}$ satisfy the conditions in Remark 2.7. We write $s := 2|f_{1w^2}(0)| \geq 0$, $x := f_{2w^2}(0) \in \mathbb{C}$ and $y := \operatorname{Im}(f_{2z^2w}(0)) \in \mathbb{R}$. By Corollary 2.12 we only need to consider coefficients in $j_0(H)$. We let $(\sigma, \sigma') \in G_0$ with the notation from (2-1), (2-3) and (2-4), and consider the equation

$$\sigma' \circ H \circ \sigma^{-1} = H, \tag{3-1}$$

where we parametrize $\sigma^{-1}$ as in (2-1). The coefficients of order 1, which are $f_z(0)$ and $H_w(0)$, are given by

$$U'^{\,t}(u\lambda\lambda', 0) = (1, 0) \quad \text{and} \quad (U'^{\,t}(uc + \lambda c_1', \lambda c_2'), \theta\lambda\lambda') = (0, 0, 1).$$

These equations imply $\theta = +1$, $\lambda' = 1/\lambda$, $a_2' = c_2' = 0$, $a_1' = 1/(uu')$ and $c_1' = -uc/\lambda$. Assuming these standard parameters we consider the coefficients of order 2, which are $f_{z^2}(0)$, $H_{zw}(0)$ and $H_{w^2}(0)$, given by

$$\big(0, 2u'u^3\lambda\big) = (0, 2), \tag{3-2}$$

$$\big(-r - \lambda^2 r' + i\varepsilon\lambda^2/2, 2u'u^3\lambda c, 0\big) = (i\varepsilon/2, 0, 0), \tag{3-3}$$

$$\big(\lambda^2(\lambda s + i\varepsilon uc)/u, uu'\lambda(\lambda^2 x + 2u^2c^2), -2(r + \lambda^2 r')\big) = (s, x, 0). \tag{3-4}$$

The second component of (3-3) implies $c = 0$. If we assume this value for $c$ we obtain for the third order terms $f_{z^2w}(0)$ the equation

$$\text{(3-5)} \qquad \big(2\mathrm{i}u\lambda^3 s,\, u'u^3\lambda(-4r - 2\lambda^2 r' + \mathrm{i}\lambda^2 y)\big) = (4\mathrm{i}s, \mathrm{i}y).$$

The second component of (3-2) shows $\lambda = 1$. Furthermore we obtain from the third component of (3-4) that $r' = -r$ and since from the second component of (3-2) we get $u'u^3 = 1$, which uniquely determines $u'$, we obtain from the second component of (3-5) that $r = 0$. The remaining equation from the first component of (3-4), which comes from the coefficient $f_{1w^2}(0)$, is $s/u = s$. If $s > 0$ we obtain that $u = 1$ and hence all standard parameters are trivial, which proves the first claim of the lemma.

If $s = 0$, then $H \in \{G_1^\varepsilon, G_{2,0}^\varepsilon, G_{3,0}^\varepsilon\}$, since these maps are precisely those satisfying $f_{1w^2}(0) = 0$ in the list of mappings from Theorem 2.8. It is easy to check that the isotropic stabilizers of the maps $G_1^\varepsilon$ and $G_{2,0}^\varepsilon$ are generated by the isotropies $(\sigma(z,w), \sigma'(z_1', z_2', w')) = (uz, w, z_1'/u, z_2'/u^2, w')$ with $|u| = 1$. If we consider $G_{3,0}^\varepsilon$ in (3-1), then we obtain that $(\sigma(z,w), \sigma'(z_1', z_2', w')) = (\delta z, w, \delta z_1', z_2', w')$, where $\delta = \pm 1$, are the only elements of $\mathrm{stab}_0(G_{3,0}^\varepsilon)$, which proves the last claim of the lemma. □

**Proposition 3.2.** *The map $N : G_0 \times \mathfrak{F} \to \mathfrak{F}$ given by $N(\phi, \phi', H) := \phi' \circ H \circ \phi^{-1}$ is a free action.*

*Proof.* Lemma 3.1 shows that $N$ restricted to $\mathfrak{N}$ is a free action. We assume the general case $H \in \mathfrak{F}$ and consider the equation $\phi' \circ H \circ \phi^{-1} = H$ for $(\phi, \phi') \in G_0$. We write $H = \widehat{\phi}' \circ \widehat{H} \circ \widehat{\phi}^{-1}$, where $\widehat{H} \in \mathfrak{N}$ and $(\widehat{\phi}, \widehat{\phi}') \in G_0$ are unique according to Lemma 3.1. After setting $(\psi, \psi') = (\widehat{\phi}^{-1} \circ \phi \circ \widehat{\phi}, \widehat{\phi}'^{-1} \circ \phi' \circ \widehat{\phi}')$, we rewrite $\phi' \circ H \circ \phi^{-1} = H$ as $\psi' \circ \widehat{H} \circ \psi^{-1} = \widehat{H}$. Since each map in $\mathfrak{N}$ admits a trivial stabilizer, we obtain that $(\psi, \psi') = (\mathrm{id}_{\mathbb{C}^2}, \mathrm{id}_{\mathbb{C}^3})$ and the freeness of the action. □

## 4. Continuity of the normalization map

**Remark 4.1.** For $F : (\mathbb{C}^2, 0) \to (\mathbb{C}^3, 0)$ a germ of a holomorphic mapping, for which we assume that $F \in \mathcal{F}$ and the jet $j_0(F) \subset j_0^3 F$ is of the form as in Remark 2.7, we write $F = (f^1, f^2, f^3)$ for the components and denote derivatives of $F$ at 0 by $f_{\ell m}^k := f_{z^\ell w^m}^k(0)$. We set $\Delta(F) := f_{10}^1 f_{20}^2 - f_{20}^1 f_{10}^2$.

**Lemma 4.2.** *For $n \in \mathbb{N}$, we let $(\phi_n, \phi_n') \in G_0$ and $H_n, H \in \mathcal{F}$ be such that $\phi_n' \circ H_n \circ \phi_n^{-1} \to H$ as $n \to \infty$, where $\mathcal{F}$ is equipped with the topology $\tau_J$. If we assume $H_n, H \in \mathcal{N}$, then $H_n \to H$, and if we assume $H_n, H \in \mathfrak{N}$, then $(\phi_n, \phi_n') \to (\mathrm{id}_{\mathbb{C}^2}, \mathrm{id}_{\mathbb{C}^3})$ as $n \to \infty$.*

*Proof.* We assume that $H_n = (h_n^1, h_n^2, h_n^3)$ and $H = (h^1, h^2, h^3)$ are given as in Remark 4.1, where the coefficients of $H_n$ depend on $n \in \mathbb{N}$. Let $s_n := 2|h_{n02}^1| \geq 0$,

$x_n := h^2_{n02} \in \mathbb{C}$ and $y_n := \operatorname{Im}\bigl(h^2_{n21}\bigr)$. To each $(\phi_n, \phi'_n) \in G_0$ we associate the standard parameters $(\gamma_n, \gamma'_n) \in \Gamma \times \Gamma'$, where we use the notation for the parametrization of $G_0$ from (2-1) and (2-4). According to Theorem 2.8, $H_n$ depends on $s_n \geq 0$. Let us denote $\Xi := \Gamma \times \Gamma' \times \mathbb{R}^+_0$ and write $\xi_n = (\gamma_n, \gamma'_n, s_n) \in \Xi$. We define $\Psi_n := \phi'_n \circ H_n \circ \phi_n^{-1}$, which depends on $\xi_n \in \Xi$. For components of $\Psi_n$, we write $\Psi_n = (\psi^1_n, \psi^2_n, \psi^3_n)$ and $\psi_n = (\psi^1_n, \psi^2_n)$. Limits are always considered when $n \to \infty$.

We start with the first order terms of $\Psi_n$. We let $U'_n$ be the $2 \times 2$-matrix from (2-3) with entries $u'_n$, $a'_{1n}$ and $a'_{2n}$ instead of $u'$, $a'_1$ and $a'_2$, so that we have

$$\psi_{nz}(0) = \lambda_n \lambda'_n U'_n{}^{t}(u_n, 0), \tag{4-1}$$

$$\Psi_{nw}(0) = \lambda_n \lambda'_n \bigl(U'_n{}^{t}(u_n c_n + \lambda_n c'_{1n}, \lambda_n c'_{2n}), \theta_n \lambda_n \lambda'_n\bigr). \tag{4-2}$$

Since $\psi^3_{nw}(0) \to 1$ we obtain $\theta_n = +1$, $\lambda_n \lambda'_n \to 1$. This implies that $u_n u'_n a'_{1n} \to 1$ and $a'_{2n} \to 0$, considering $\psi_{nz}(0) \to (1, 0)$ in (4-1). Because $a'_n = (a'_{1n}, a'_{2n}) \in \mathcal{S}^2_{\varepsilon,\theta}$ from (2-2), we have $|a'_{1n}| \to 1$. If we consider the first two components in (4-2), we obtain from $\psi_{nw}(0) \to (0, 0)$ and $(|a'_{1n}|, |a'_{2n}|) \to (1, 0)$ that $u_n c_n + \lambda_n c'_{1n} \to 0$ and $c'_{2n} \to 0$.

Next we consider the second order terms of $\Psi_n$ to obtain

$$\psi_{nz^2}(0) = 2u_n \lambda_n \lambda'_n U'_n{}^{t}\bigl(2\mathrm{i}(\bar{c}_n + u_n \lambda_n \bar{c}'_{1n}), u_n \lambda_n\bigr), \tag{4-3}$$

where the left-hand side of (4-3), $\psi_{nz^2}(0)$, must converge to $(0, 2)$. After applying $U'^{-1}_n$ we rewrite the second component of (4-3) as

$$2u_n^2 \lambda_n^2 \lambda'_n = a'_{1n}\bigl(-\bar{a}'_{2n} \psi^1_{nz^2}(0)/(u'_n a'_{1n}) + \psi^2_{nz^2}(0)\bigr). \tag{4-4}$$

Since $(|a'_{1n}|, |a'_{2n}|) \to (1, 0)$, the absolute value of the right-hand side of (4-4) converges to 2. Taking the absolute value of the left-hand side of (4-4) implies $\lambda_n \to 1$, which together with $\lambda_n \lambda'_n \to 1$ shows $\lambda'_n \to 1$. Next we consider

$$\psi_{nzw}(0) = \tfrac{\mathrm{i}}{2} \lambda_n \lambda'_n U'_n{}^{t}\bigl(T_1(\gamma_n, \gamma'_n), 4\lambda_n(c'_{2n}(\bar{c}_n + u_n \lambda_n \bar{c}'_{1n}) - \mathrm{i}u_n^2 c_n)\bigr), \tag{4-5}$$

where the real-analytic function $T_1 : \Gamma \times \Gamma' \to \mathbb{C}$ does not depend on $a'_n \in \mathcal{S}^2_{\varepsilon,\theta}$ and $u'_n$. The left-hand side of (4-5) has to converge to $(\mathrm{i}\varepsilon/2, 0)$ and we rewrite the second component of (4-5) as

$$\begin{aligned} &4\lambda_n\bigl(c'_{2n}(\bar{c}_n + u_n \lambda_n \bar{c}'_{1n}) - \mathrm{i}u_n^2 c_n\bigr) \\ &\qquad = -2\mathrm{i}\bigl(-\bar{a}'_{2n} \psi^1_{nzw}(0) + u'_n a'_{1n} \psi^2_{nzw}(0)\bigr)/(\lambda_n \lambda'_n u'_n). \end{aligned} \tag{4-6}$$

Taking the limit, we know, since $(|a'_{1n}|, |a'_{2n}|) \to (1, 0)$ and $(\lambda_n, \lambda'_n) \to (1, 1)$, that the right-hand side of (4-6) converges to 0 and if we also use $u_n c_n + \lambda_n c'_{1n} \to 0$ we obtain that $c_n \to 0$, such that $c'_{1n} \to 0$.

Next we compute

$$(4\text{-}7)\quad \psi^3_{nw^2}(0) = 2\lambda_n^2\lambda_n'^2\big(-(r_n+\lambda_n^2 r_n') + \mathrm{i}\big(|c_n|^2+\varepsilon\lambda_n^2|c_{2n}'|^2+\lambda_n\bar{c}_{1n}'(2u_nc_n+\lambda_n c_{1n}')\big)\big).$$

We take all of the previously obtained limits as $n\to\infty$ of the sequences $c_n' = (c_{1n}', c_{2n}')\in\mathbb{C}^2$, $c_n$ and $\lambda_n$, $\lambda_n'$. Then since $\psi^3_{nw^2}(0)\to 0$, we have that $r_n+\lambda_n^2 r_n'\to 0$. Next we compute

$$(4\text{-}8)\qquad \psi_{nw^2}(0) = \lambda_n\lambda_n' U_n'\,{}^t\big(\lambda_n^3 s_n + T_2(\gamma_n,\gamma_n'),\, \lambda_n^3 x_n + T_3(\gamma_n,\gamma_n')\big),$$

where $T_2, T_3 : \Gamma\times\Gamma'\to\mathbb{C}$ are real-analytic functions and $T_2$ is given by

$$\begin{aligned} T_2(\gamma_n,\gamma_n') = 2(u_nc_n+c_{1n}'\lambda_n)\big(\mathrm{i}|c_n|^2 - r_n - \lambda_n^2 r_n'\big)& \\ +2\mathrm{i}\lambda_n\bar{c}_{1n}'(u_nc_n+\lambda_nc_{1n}')(2u_nc_n+\lambda_nc_{1n}')& \\ +\mathrm{i}\varepsilon\lambda_n^2\big(u_nc_n(1+2|c_{2n}'|^2)+2\lambda_nc_{1n}'|c_{2n}'|\big)&, \end{aligned}$$

so that $T_2(\gamma_n,\gamma_n')\to 0$. Then the first component of (4-8) becomes

$$(4\text{-}9)\qquad \lambda_n^3 s_n + T_2(\gamma_n,\gamma_n') = \big(\bar{a}_{1n}'\psi^1_{nw^2}(0)+\varepsilon u_n' a_{2n}'\psi^2_{nw^2}(0)\big)/(\lambda_n\lambda_n'u_n').$$

Since $\big(\psi^1_{nw^2}(0), \psi^2_{nw^2}(0)\big)\to\big(2|h^1_{02}|, h^2_{02}\big)\in\mathbb{R}^+\times\mathbb{C}$, we obtain $s_n\to 2|h^1_{02}|$, and if $|h^1_{02}|\neq 0$ we have $\bar{a}_{1n}'/u_n'\to 1$. Then $u_nu_n'a_{1n}'\to 1$ implies that $u_n\to 1$ and further inspection of (4-4) gives $u_n^2/a_{1n}'\to 1$, which shows $a_{1n}'\to 1$ and $u_n'\to 1$. Note that if $|h^1_{02}|=0$ we have that $a_{1n}', u_n, u_n'\in\mathbb{S}^1$ for all $n\in\mathbb{N}$. Observe that the following considerations are independent of the value of $h^1_{02}$:

$$(4\text{-}10)\quad \psi_{nz^2w}(0) = \lambda_n\lambda_n'U_n'\begin{pmatrix} -4\mathrm{i}u_n^2\lambda_n^3 s_n + T_4(\gamma_n,\gamma_n') \\ -2\varepsilon u_n^2\lambda_n(2r_n+\lambda_n^2r_n') + \mathrm{i}\varepsilon u_n^2\lambda_n^3 y_n + 6u_n^3\lambda_n^2 c_n s_n + T_5(\gamma_n,\gamma_n')\end{pmatrix},$$

where $T_4, T_5 : \Gamma\times\Gamma'\to\mathbb{C}$ are real-analytic functions and $T_5$ is given by

$$\begin{aligned} T_5(\gamma_n,\gamma_n') = 2\mathrm{i}\varepsilon\lambda_n\big(4\mathrm{i}\bar{c}_nc_{2n}'(\bar{c}_n+2u_n\lambda_n\bar{c}_{1n}') + 2c_nu_n^2(5\bar{c}_n+3u_n\lambda_n\bar{c}_{1n}')& \\ +u_n^2\lambda_n^2\big(|c_{1n}'|^2+3\varepsilon|c_{2n}'|^2+4\mathrm{i}\bar{c}_{1n}'c_{2n}'\big)\big)&, \end{aligned}$$

hence $T_5(\gamma_n,\gamma_n')\to 0$. Since $\big(\psi^1_{nz^2w}(0), \psi^2_{nz^2w}(0)\big)\to\big(2\mathrm{i}|h^1_{02}|, \mathrm{i}h^2_{21}\big)\in\mathrm{i}\mathbb{R}\times\mathrm{i}\mathbb{R}$, considering the real part of the second component of (4-10) we obtain $2r_n+r_n'\to 0$, which together with $r_n+\lambda_n^2r_n'\to 0$ shows $(r_n, r_n')\to(0,0)$. To sum up, we obtain that $H_n\to H$, and moreover, if $|h^1_{02}|\neq 0$, we conclude $(\phi_n,\phi_n')\to(\mathrm{id}_{\mathbb{C}^2},\mathrm{id}_{\mathbb{C}^3})$, which completes the proof. □

**Proposition 4.3.** *The map $\pi:\mathcal{F}\to\mathcal{N}$ given by $\pi(H):=\phi'\circ H\circ\phi^{-1}$ for $(\phi,\phi')\in G_0$, according to Proposition 2.6, is continuous with respect to $\tau_J$.*

*Proof.* Let $(H_n)_{n\in\mathbb{N}}$ be a sequence of mappings in $\mathcal{F}$ and $H \in \mathcal{F}$, such that $H_n \to H$. Assuming without loss of generality that $H \in \mathcal{N}$, we need to show $\widetilde{H}_n := \pi(H_n) \to H$. We have $\widetilde{H}_n = \phi'_n \circ H_n \circ \phi_n^{-1} \in \mathcal{N}$, where $(\phi_n, \phi'_n) \in G_0$ are the isotropies according to Proposition 2.6. Since $H_n = {\phi'}_n^{-1} \circ \widetilde{H}_n \circ \phi_n \to H$, we conclude by Lemma 4.2 that $\widetilde{H}_n \to H$. □

Using Proposition 4.3 we are able to prove Theorem 1.2.

*Proof of Theorem 1.2.* We show that $\pi : \mathcal{F} \to \mathcal{N}$ is a surjective, continuous and closed mapping with respect to $\tau_J$. Surjectivity is clear from Proposition 2.6 and Theorem 2.8 and continuity we have shown in Proposition 4.3. It remains to prove that $\pi$ is closed with respect to $\tau_J$. Let $C \subset \mathcal{F}$ be a closed subset. We need to show that $\pi(C) \subset \mathcal{N}$ is a closed subset. To prove this statement we let $H_n \in \pi(C)$ for $n \in \mathbb{N}$, forming a sequence of mappings in $\mathcal{N}$ such that $H_n \to H_0$, where $H_0 \in \mathcal{N}$. To show that $\pi(C)$ is closed we need to conclude that $H_0 \in \pi(C)$. By Theorem 2.8 we can write $H_n = G^{\varepsilon}_{k_n,s_n}$ and $H_0 = G^{\varepsilon}_{k_0,s_0}$ for $k_n, k_0 \in \{2, 3\}$. Note that since $H_n \to H_0$ in $\mathcal{N}$ we have $s_n \to s_0$. This implies that for any convergent sequence $G_n \in \pi^{-1}(H_n)$ the map $G_0 := \lim_{n\to\infty} G_n$ belongs to $\pi^{-1}(H_0)$. Since $C$ is closed, an arbitrary convergent sequence $F_n \in \pi^{-1}(H_n) \cap C$ with $F_n \to F_0$ thus satisfies $F_0 \in \pi^{-1}(H_0) \cap C$, which implies $H_0 = \pi(F_0) \in \pi(C)$. □

## 5. A topological property of the quotient space of $\mathcal{F}$

**Lemma 5.1.** *The class $\mathcal{F}$ consists of $\frac{5+\varepsilon}{2}$ connected components.*

*Proof.* According to Proposition 2.6 and Proposition 4.3, we denote by $\pi : \mathcal{F} \to \mathcal{N}$ the normalization map, which is continuous with respect to $\tau_J$. By Theorem 1.2, we equip $\mathcal{F}$ and $\mathcal{N}$ with $\tau_J$. For $k \in \{2, 3\}$, we set $C_k := \{G^{\varepsilon}_{k,s} \mid s \geq 0\}$ and $\mathcal{N}^* := C_2 \cup C_3$. The space of standard parameters $\Gamma \times \Gamma'$ is path-connected, since as defined in Definition 2.5 for maps $H = (f_1, f_2, g) \in \mathcal{F}$, we assumed $g_w(0) > 0$, which implies that for isotropies as in (2-4) we require $\theta = +1$ for $\varepsilon = \pm 1$. Thus for any $H \in \mathcal{N}$ the isotropic orbit $O_0(H)$ is path-connected. First we treat the case $\varepsilon = -1$. We observe that $\mathcal{F}^* := \bigcup_{H\in\mathcal{N}^*} O_0(H)$ is path-connected. If $\mathcal{F}$ were connected then $\pi(\mathcal{F}) = \mathcal{N}$ would be connected, which is not possible, since $\mathcal{N}$ consists of 2 connected components $G_1^-$ and $\mathcal{N}^*$. Thus $\mathcal{F}$ has 2 connected components $O_0(G_1^-)$ and $\mathcal{F}^*$. For $\varepsilon = +1$ we note that the set $O_0(C_k) := \bigcup_{H\in C_k} O_0(H)$ for $k \in \{2, 3\}$ is path-connected and $\mathcal{N}$ consists of 3 connected components. Thus $\mathcal{F}$ admits at most 3 connected components. $\mathcal{F}$ is not connected since then $\pi(\mathcal{F}) = \mathcal{N}$ would be connected. If $\mathcal{F}$ consists of 2 connected components $\mathcal{F}_1, \mathcal{F}_2$ such that $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$, we need to distinguish several cases. Either $\mathcal{F}_1 = O_0(G_1^+) \cup O_0(C_k)$, and $\mathcal{F}_2 = O_0(C_\ell)$, where $k \neq \ell$ and $k, \ell \in \{2, 3\}$, or $\mathcal{F}_1 = O_0(C_2) \cup O_0(C_3)$ and $\mathcal{F}_2 = O_0(G_1^+)$. In all cases we have, by the continuity of $\pi$, that $\pi(\mathcal{F}_1)$ is connected, which is not possible. □

*Proof of Theorem 1.3.* The quotient space $X$ consists of elements denoted by $[F]$ for $F \in \mathcal{F}$. We equip $X$ with the quotient topology such that the canonical projection $p : \mathcal{F} \to X$ is continuous. For $\varepsilon = +1$ we have $X = \{G_1^+, G_{2,0}^+, G_{3,0}^+\}$ by our classification. By Lemma 5.1 we obtain that $p^{-1}(H)$ for $H \in X$ is a connected component of $\mathcal{F}$, hence open. Thus $X$ carries the discrete topology. To prove the statement for $\varepsilon = -1$ we write $H_0 := G_{2,1/2}^- \in \mathcal{N}$ and $H_1 := G_{3,0}^- \in \mathcal{N}$. For $k \in \{0, 1\}$, let $U_k \in X$ be an open neighborhood of $[H_k]$. Then $V_k := p^{-1}(U_k)$ is an open neighborhood of the orbit of $H_k$ in $\mathcal{F}$. According to our classification there exists a sequence $(G_n)_{n\in\mathbb{N}}$ of mappings in $\mathcal{F}$, where each $G_n \in [H_1]$ and $G_n \to H_0$ in $\mathcal{F}$ as $n \to \infty$. Thus there exists $N \in \mathbb{N}$ such that $G_n \in V_0 \cap V_1$ for all $n \geq N$, which shows $[H_1] \in U_0 \cap U_1$ and completes the proof. □

## 6. Properness of the group action

*Proof of Theorem 1.4.* For $n \in \mathbb{N}$ we let $G_n = (g_n^1, g_n^2, g_n^3)$, $H'_n = (h'^1_n, h'^2_n, h'^3_n) \in \mathcal{F}$ with $G_n = \varphi'_n \circ H'_n \circ \varphi_n^{-1}$, where $(\varphi_n, \varphi'_n) \in G_0$. Equipping $J_0^3$ with a suitable norm $\|\cdot\|$, we need to show that if we let $N > 1$ such that $\|j_0(G_n)\|, \|j_0(H'_n)\| \leq N$ and $|g^3_{n01}|, |h'^3_{n01}| \geq 1/N$ as well as $|\Delta(G_n)|, |\Delta(H'_n)| \geq 1/N$, then we have that $\{(\varphi_n, \varphi'_n) \mid n \in \mathbb{N}\}$ is relatively compact in $G_0$. For a simplification, we write $H'_n = \phi'_n \circ H_n \circ \phi_n^{-1}$, where $H_n \in \mathcal{N}$ and $(\phi_n, \phi'_n) \in G_0$ according to Proposition 2.6. Since we have shown in Proposition 4.3 that the map $\pi : \mathcal{F} \to \mathcal{N}$ is continuous, it follows that the sequence $H_n$ is relatively compact, and we assume that each $H_n$ satisfies all conditions we assumed for $H'_n$. Further we assume that $H_n$ is given as described in Remark 2.7, where we set $s_n := 2|h^1_{n02}| \geq 0$, $x_n := h^2_{n02} \in \mathbb{C}$ and $y_n := \operatorname{Im}(h^2_{n21})$.

In the proof of Proposition 2.6 given in [Reiter 2015, Proposition 3.1], we give explicit formulas for $(\phi_n, \phi'_n)$, which shows that $\{(\phi_n, \phi'_n) \mid n \in \mathbb{N}\}$ is bounded, since the sequence $H'_n$ is relatively compact. We set $\psi_n := \varphi_n \circ \phi_n$ and $\psi'_n := \varphi'_n \circ \phi'_n$. Hence we need to prove that $\{(\psi_n, \psi'_n) \mid n \in \mathbb{N}\}$ is bounded in $G_0$. If we use the parametrization of $(\psi_n, \psi'_n)$ from (2-1) and (2-4), we show that $\{(\gamma_n, \gamma'_n) \mid n \in \mathbb{N}\}$ is bounded in $\Gamma \times \Gamma'$. More precisely, we need to show the boundedness of the sequences $\lambda_n, c_n, r_n, a'_{1n}, a'_{2n}, \lambda'_n, c'_n, r'_n$ in $\Gamma \times \Gamma'$. We use the equations from the proof of Lemma 4.2, where $\Psi_n$ plays the role of $G_n$. We start considering the third component of (4-2), which gives $1/\sqrt{N} \leq \lambda_n\lambda'_n \leq \sqrt{N}$. Then we rewrite (4-1) to obtain, for $k = 1, 2$, that $|a'_{kn}| = |g^k_{n10}|/(\lambda_n\lambda'_n) \leq N\sqrt{N}$. After rewriting the first two components of (4-2) we obtain that $|u_n c_n + \lambda_n c'_{1n}|, |\lambda_n c'_{2n}| \leq 2N^3$. Then, using (4-2) and (4-3), we compute

$$\Delta(G_n) = \left|\lambda_n\lambda'_n U'_n \begin{pmatrix} u_n & 4\mathrm{i}u_n(\bar{c}_n + u_n\lambda_n\bar{c}'_{1n}) \\ 0 & u_n\lambda_n \end{pmatrix}\right| = u_n^2\lambda_n^3\lambda'^2_n,$$

such that the boundedness of $\Delta(G_n)$ from below implies that $1/N^2 \leq \lambda_n \leq N^2$. This gives $1/(\sqrt{N}N^2) \leq \lambda'_n \leq \sqrt{N}N^2$, and from (4-2) we derive boundedness of the sequence $|c'_{2n}|$. Then from (4-5) we obtain that the sequence $|c_n|$ is bounded, such that (4-2) shows the boundedness of $|c'_{1n}|$.

Finally, using all the previous bounds, we get from (4-7) and the second component of (4-10) that the sequences $|r_n+\lambda_n^2 r'_n|$ and $|2r_n+\lambda_n^2 r'_n|$ are both bounded, which implies that $|r_n|$ and $|r'_n|$ are bounded from above. Thus the sequence $(\psi_n, \psi'_n)$ is relatively compact. Since $(\phi_n, \phi'_n)$ is relatively compact, this implies that $(\varphi_n, \varphi'_n)$ is also relatively compact, completing the proof. □

## 7. On the real-analytic structure of $\mathfrak{F}$

**Lemma 7.1.** *Let* $\Pi : \mathfrak{F} \to \mathfrak{N}$ *be given by* $\Pi(H) := \phi' \circ H \circ \phi^{-1}$, *where* $(\phi, \phi') \in G_0$ *are the unique isotropies according to Proposition 2.6 and Lemma 3.1. For* $k = 2, 3$ *we write* $M_{k,\varepsilon} := \{\Pi^{-1}(G^{\varepsilon}_{k,s}) \mid s > 0\}$. *Then* $M_{k,\varepsilon}$ *is a real-analytic real submanifold of* $\mathfrak{F}$ *of real dimension* 16.

*Proof.* For fixed $k \in \{2, 3\}$, $s > 0$ and $\delta > 0$, we write $G_{\delta,s} := \{G^{\varepsilon}_{k,t} \mid t \in B_\delta(s) \cap \mathbb{R}^+\}$, where $B_\delta(s) := \{t \in \mathbb{R}^+ \mid |t - s| < \delta\}$. To prove the lemma we show that for every $s_0 \in \mathbb{R}^+$ and sufficiently small $\delta_0 > 0$, there exists a locally real-analytic parametrization for $M := \Pi^{-1}(G_{\delta_0,s_0})$. As noted in Remark 2.13, we identify $\mathcal{F}$ with the set $\mathfrak{J} \subset \mathbb{C}^{K_0}$.

Theorem 2.8 implies that for each $H \in M$ there exist $(\phi, \phi') \in G_0$, $k \in \{2, 3\}$ and $s_1 \in B_{\delta_0}(s_0) \cap \mathbb{R}^+$, such that $H = \phi' \circ G^{\varepsilon}_{k,s_1} \circ \phi^{-1}$. This fact is used to describe $M$ locally via parametrizations as follows. For $s > 0$ sufficiently near $s_0$, let $F_s$ be a mapping as in Remark 4.1, which depends real-analytically on $s := 2|f^1_{02}|$. For the remaining coefficients in $j_0(F_s)$ we write $x := f^2_{02}$ and $y := \operatorname{Im}\big(f^2_{21}\big)$, where we suppress the dependence on $s$ notationally. We use the real version of the notation for the parametrization of $G_0$ as in (2-1) and (2-4). Here we denote the set of real parameters of $G_0$ by $\Gamma \times \Gamma'$. Let us write $\Xi := \Gamma \times \Gamma' \times \mathbb{R}^+ \subset \mathbb{R}^{N_0}$, where $N_0 := 16$. For $\xi = (\gamma, \gamma', s) \in \Xi$, we define the mapping

$$\Psi : \Xi \to \mathfrak{J}, \quad \Psi(\xi) := j_0(\phi'_{\gamma'} \circ F_s \circ \phi_\gamma^{-1}), \tag{7-1}$$

where we use the notation as in (2-1) and (2-4) for $\phi_\gamma$ and $\phi'_{\gamma'}$ respectively and suppress the dependence on $\varepsilon$.

We set $\check{\Psi}(z, w) := (\phi'_{\gamma'} \circ F_s \circ \phi_\gamma^{-1})(z, w)$ with components $\check{\Psi} = (\check{\psi}^1, \check{\psi}^2, \check{\psi}^3)$ and $\check{\psi} := \big(\check{\psi}^1, \check{\psi}^2\big)$. The holomorphic mapping $\check{\Psi}$ is defined in a small neighborhood $U \subset \mathbb{C}^2$ of 0 and satisfies $\check{\Psi}(\mathbb{H}^2 \cap U) \subset \mathbb{H}^3_\varepsilon$. By Theorem 2.8 and the real-analytic dependence of the isotropies on the standard parameters, which can be observed by inspecting the proof of Proposition 2.6 in [Reiter 2015, Proposition 3.1], we note that $\Psi$ and $\check{\Psi}$ are real-analytic in $\xi \in \Xi$. We assume without loss of generality that $\xi_0$ is chosen in such a way that $(\phi_\gamma, \phi'_{\gamma'}) = (\mathrm{id}_{\mathbb{C}^2}, \mathrm{id}_{\mathbb{C}^3})$. Consequently we write $O(2)$ for terms involving standard parameters of the isotropies which vanish to second order at $\xi_0$, and we consider $a_1' \in \mathbb{C}$ near 1 such that we substitute $\bar{a}_1' = (1-\varepsilon|a_2'|^2)/a_1'$ and take $\theta = +1$ in $\Psi$, which is then given by the following expressions:

$$
\begin{aligned}
\check{\psi}_z(0) &= \bigl(uu'\lambda\lambda' a_1', u\lambda\lambda'\bar{a}_2'\bigr),\\
\check{\Psi}_w(0) &= \bigl(u'\lambda\lambda' a_1'(uc+\lambda c_1'), \lambda^2\lambda' c_2'/a_1', \lambda^2\lambda'^2\bigr) + O(2),\\
\check{\psi}_{z^2}(0) &= \bigl(2\mathrm{i}uu'\lambda\lambda'(\mathrm{i}\varepsilon u\lambda a_2' + 2(\bar{c}+u\lambda\bar{c}_1')a_1'), 2u^2\lambda^2\lambda'/a_1'\bigr) + O(2),\\
\check{\Psi}_{zw}(0) &= \Bigl(-\frac{1}{2}uu'\lambda\lambda' a_1'(2(r+\lambda^2 r') - \mathrm{i}\varepsilon\lambda^2),\\
&\qquad\qquad u\lambda^2\lambda'\Bigl(\frac{\mathrm{i}\varepsilon}{2}\lambda\bar{a}_2' + \frac{2uc}{a_1'}\Bigr), 2\mathrm{i}\lambda^2\lambda'^2(\bar{c}+u\lambda\bar{c}_1')\Bigr) + O(2),\\
\check{\Psi}_{w^2}(0) &= \bigl(u'\lambda^3\lambda'\bigl(a_1'(\mathrm{i}\varepsilon uc+\lambda s) - \varepsilon\lambda a_2' x\bigr),\\
&\qquad\qquad \lambda^4\lambda'\bigl(x/a_1' + \bar{a}_2' s\bigr), -2\lambda^2\lambda'^2(r+\lambda^2 r')\bigr) + O(2),\\
\check{\psi}_{z^2w}(0) &= \Bigl(-uu'\lambda^3\lambda'\bigl(4a_1'\bigl(-\mathrm{i}u\lambda s + \varepsilon(\bar{c}+u\lambda\bar{c}_1')\bigr) + \mathrm{i}\varepsilon u\lambda a_2' y\bigr),\\
&\qquad\qquad u^2\lambda^2\lambda'\bigl(\bigl(-2(2r+\lambda^2 r') + 6\varepsilon u\lambda cs + \mathrm{i}\lambda^2 y\bigr)/a_1' + 2\mathrm{i}\lambda^2\bar{a}_2' s\bigr)\Bigr) + O(2).
\end{aligned}
$$

As a first step we show that for given $\xi_0 \in \Xi$ the Jacobian of $\Psi$ with respect to $\xi$ evaluated at $\xi_0$, denoted by $\Psi_\xi(\xi_0)$, is of full rank $N_0$. Instead of considering the real equations of $\Psi$, however, we conjugate $\Psi$ and compute the Jacobian of the system $\Phi := (\Psi, \overline{\Psi}) \in \mathbb{C}^{2K_0}$ with respect to

$$\xi = (u, \lambda, c, r, u', a_1', a_2', \lambda', c_1', c_2', r', s; \bar{c}, \bar{a}_2', \bar{c}_1', \bar{c}_2') \in \mathbb{C}^{N_0}$$

and evaluate at

$$\xi_0 = (1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 0, s_0; 0, 0, 0, 0) \in \mathbb{R}^{N_0}, \tag{7-2}$$

denoted by $\Phi_\xi(\xi_0)$. We bring the transpose of $\Phi_\xi(\xi_0)$ into echelon form, denoting the resulting matrix by $\varphi = (\varphi^1, \dots, \varphi^{N_0})$, where $\varphi^j = (\varphi_1^j, \dots, \varphi_{2K_0}^j) \in \mathbb{C}^{2K_0}$ for $1 \le j \le N_0$, such that $\operatorname{rank}\bigl(\Phi_\xi(\xi_0)\bigr) = \operatorname{rank}(\varphi)$. In the following we suppress the evaluation of $\Phi$ at $\xi_0$ notationally and perform elementary row operations. The matrix given by

$$
\begin{aligned}
(\varphi^1, \dots, \varphi^{11}) := \Bigl(&\Phi_u, \Phi_{\bar{a}_2'}, \Phi_{c_1'}, \Phi_{c_2'}, \Phi_\lambda, \Phi_{\bar{c}}, \Phi_{a_1'}, \Phi_{r'}, \Phi_c, \Phi_{a_2'}, \Phi_s\Bigr)\\
&-\Bigl(0, 0, 0, \Phi_u, \Phi_u, 0, \Phi_u, 0, \Phi_{c_1'}, \frac{\mathrm{i}\varepsilon}{2}\Phi_{\bar{c}}, 0\Bigr),
\end{aligned}
$$

is in row echelon form, with constant nonzero entries on the main diagonal. Each 0 above represents $0 \in \mathbb{C}^{2K_0}$. Next we define

$$
\begin{aligned}
\varphi^{12} &:= \Phi_{\lambda'} + \frac{1}{3}\Phi_u - \Phi_\lambda - \frac{1}{3}\Phi_{a_1'} - \frac{\mathrm{i}\varepsilon}{8}\Phi_{r'} + \frac{10s_0}{3}\Phi_s,\\
\varphi^{13} &:= \Phi_{u'} - \frac{1}{3}\Phi_u - \frac{2}{3}\Phi_{a_1'} - \frac{2}{3}\Phi_s,
\end{aligned}
$$

which are of the following form, where we denote by $h'$ derivatives of a function $h$ depending on $s$ with respect to $s$:

$$
\begin{aligned}
\varphi^{12} &= (0,\dots,0,\varphi^{12}_{12},\dots,\varphi^{12}_{2K_0})\\
&= \Bigl(0,\dots,0,\frac{-2(4x-5s_0x')}{3},2\mathrm{i}\varepsilon,\frac{8\mathrm{i}s_0}{3},\frac{2\mathrm{i}(3\varepsilon-3y+5s_0y')}{3},-\frac{1}{3},\varphi^{12}_{17},\dots,\varphi^{12}_{2K_0}\Bigr)\\
\varphi^{13} &= (0,\dots,0,\varphi^{13}_{12},\dots,\varphi^{13}_{2K_0})\\
&= \Bigl(0,\dots,0,\frac{2x-s_0x'}{3},0,-\frac{8\mathrm{i}s_0}{3},-\frac{\mathrm{i}s_0y'}{3},-\frac{2}{3},\varphi^{13}_{17},\dots,\varphi^{13}_{2K_0}\Bigr).
\end{aligned}
$$

Then we define $\varphi^{14} := \Phi_r - \Phi_{r'}$, $\varphi^{15} := \Phi_{\tilde c'_2}$ and $\varphi^{16} := \Phi_{\tilde c'_1}$, from which we compute $\varphi^{14} = -2(e_{15}+e_{2K_0})$, $\varphi^{15} = e_{19}$ and $\varphi^{16} = -2e_{24} + \mathrm{i}\varepsilon e_{26} - 12\varepsilon s e_{2K_0}$, where for $j \in \mathbb{N}$ we denote by $e_j$ the $j$-th unit vector in $\mathbb{R}^{2K_0}$.

We have to consider several cases. First, in case $\varphi^{12}_{12} \neq 0$, then we consider $\tilde\varphi^{13} := \varphi^{13} - \varphi^{13}_{12}\varphi^{12}/\varphi^{12}_{12}$, such that $\tilde\varphi^{13}_{13}$ is a multiple of $-2x+s_0x'$. If $\tilde\varphi^{13}_{13} \neq 0$, then $\varphi = (\varphi^1,\dots,\varphi^{12},\tilde\varphi^{13},\varphi^{14},\varphi^{15},\varphi^{16})$ is in echelon form. If $\tilde\varphi^{13}_{13} = 0$, then $x = Cs^2$, where $C \in \mathbb{C}\setminus\{0\}$ and we have $\tilde\varphi^{13}_{14} \neq 0$, which again implies that $\varphi = (\varphi^1,\dots,\varphi^{12},\tilde\varphi^{13},\varphi^{14},\varphi^{15},\varphi^{16})$ is in echelon form.

Next we treat the case $\varphi^{12}_{12} = 0$. First we consider the trivial case. If $x = 0$, then since $s_0 > 0$, we have $x' = 0$ and so $\varphi = (\varphi^1,\dots,\varphi^{16})$ is in echelon form. Now we assume $x \neq 0$, which implies $x' \neq 0$, and solve $\varphi^{12}_{12} = 0$. The solution is given by $x = Cs^{4/5}$, where $C \in \mathbb{C}\setminus\{0\}$ and $\varphi = (\varphi^1,\dots,\varphi^{11},\varphi^{13},\varphi^{12},\varphi^{14},\varphi^{15},\varphi^{16})$ is in echelon form.

To sum up, we conclude that in all cases the Jacobian $\Phi_\xi(\xi_0)$ of the system $\Phi$ evaluated at $\xi_0$ is of full rank $N_0$, and hence that $\Psi$ from (7-1) is a real-analytic locally regular mapping if we choose $\delta_0 > 0$ sufficiently small in $M$. For $\Psi$ to be a local parametrization of $M$ it remains to show that for each sufficiently small neighborhood $U \subset \Xi \subset \mathbb{R}^{N_0}$ of $\xi_0$, there exists a neighborhood $W \subset \mathbb{C}^{K_0}$ of $\Psi(\xi_0) = F_{s_0}$, such that $\Psi(U) = W \cap M$. We have

$$\Psi(U) = \{ j_0(H) \mid \exists \xi = (\gamma,\gamma',t) \in U : H = {\phi'}^{-1}_{\gamma'} \circ F_t \circ \phi_\gamma \}$$

and with the notation from the beginning of this proof for $\delta > 0$ we have

$$
\begin{aligned}
M &= \Pi^{-1}(F_{\delta,s_0})\\
&= \{ H \in \mathfrak{F} \mid \exists (\gamma,\gamma',s) \in \Gamma \times \Gamma' \times B_\delta(s_0) \cap \mathbb{R}^+ : {\phi'}_{\gamma'} \circ H \circ \phi^{-1}_\gamma = F_s \}.
\end{aligned}
$$

Remark 2.13, together with the fact that for each $H \in M$ we can write $H = {\phi'}^{-1}_{\gamma'} \circ F_s \circ \phi_\gamma$, shows $\Psi(U) \subset M$. We assume that there exists a neighborhood $U \subset \Xi$ of $\xi_0$, such that for any neighborhood $W$ of $\Psi(\xi_0) = F_{s_0}$ we have $\Psi(U) \neq W \cap M$. We choose open, connected neighborhoods $(W_n)_{n\in\mathbb{N}}$ of $F_{s_0}$ with $\bigcap_n W_n = \{F_{s_0}\}$ and $\Psi(U) \neq W_n \cap M$ for all $n \in \mathbb{N}$. There exists a sequence of mappings $(H_n)_{n\in\mathbb{N}} \in \mathfrak{F}$

such that $H_n \in W_n \cap M$ and $H_n \notin \Psi(U)$. We write $H_n = {\phi'}^{-1}_{\gamma'_n} \circ F_{s_n} \circ \phi_{\gamma_n}$ and conclude by Lemma 4.2 that $(\gamma_n, \gamma'_n, s_n) \to \xi_0$ in $\Xi$. Thus eventually $H_n \in \Psi(U)$ for large enough $n \in \mathbb{N}$, which completes the proof of the lemma. □

We need the following theorem concerning free and proper group actions on manifolds.

**Theorem 7.2** [Duistermaat and Kolk 2000, Theorem 1.11.4]. *Let $k \in \mathbb{N} \cup \{\infty, \omega\}$ be nonzero and $M$ a $C^k$ manifold equipped with a $C^k$ action $G \times M \to M$, where $G$ is a $C^k$ Lie group. Assume that the action is free and proper. Then $M/G$ has the unique structure of a $C^k$ manifold of real dimension* $\dim_{\mathbb{R}} M - \dim_{\mathbb{R}} G$ *and the topology of $M/G$ is the quotient topology $\tau_Q$. We denote by $\varphi : M \to M/G$ the canonical projection given by $\varphi(m) = G \cdot m := \{g \cdot m \mid g \in G\}$ for $m \in M$. For every $s \in M/G$ there is an open neighborhood $S \subset M/G$ of $s$ and a $C^k$ diffeomorphism $\psi : \varphi^{-1}(S) \to G \times S$, $\psi : m \mapsto (\psi_1(m), \psi_2(m))$, such that for $m \in \varphi^{-1}(S)$, $g \in G$ we have $\varphi(m) = \psi_2(m)$ and $\psi(g \cdot m) = (g \cdot \psi_1(m), \psi_2(m))$. We say the triple $(\varphi, M, M/G)$ is a $C^k$* principal fiber bundle with structure group *$G$.*

*Proof of Theorem 1.5.* By [Baouendi et al. 1997, Corollary 1.2] the group $G_0$ is a totally real, closed, real-analytic submanifold of

$$G_0^2(\mathbb{H}^2, 0) \times G_0^2(\mathbb{H}^3_\varepsilon, 0) \subset J_0^2(\mathbb{H}^2, 0) \times J_0^2(\mathbb{H}^3_\varepsilon, 0).$$

Hence $G_0$ is a real-analytic real Lie group. With the notation of Lemma 7.1 we define for $(\gamma, \gamma') \in \Gamma \times \Gamma'$ the map $N_{\gamma,\gamma'} : M_{k,\varepsilon} \to M_{k,\varepsilon}$, $N_{\gamma,\gamma'}(H) := \phi'_{\gamma'} \circ H \circ \phi_\gamma^{-1}$, where $(\phi_\gamma, \phi'_{\gamma'}) \in G_0$ according to (2-1) and (2-4). We would like to conclude that for each fixed $(\gamma, \gamma') \in \Gamma \times \Gamma'$, the map $N_{\gamma,\gamma'}$ is real-analytic. By Remark 2.13, instead of $N_{\gamma,\gamma'}$ it suffices to consider $N'_{\gamma,\gamma'} : \mathfrak{J}_{k,\varepsilon} \to \mathfrak{J}_{k,\varepsilon}$, where $\mathfrak{J}_{k,\varepsilon} := \{j_0(H) \mid H \in M_{k,\varepsilon}\}$, and $N'_{\gamma,\gamma'}(j_0(H)) := j_0(\phi'_{\gamma'} \circ H \circ \phi_\gamma^{-1})$ is a restriction of $N_{\gamma,\gamma'}$. By considering the components of $N'_{\gamma,\gamma'}(j_0(H))$ for $H \in M_{k,\varepsilon}$, we obtain that $N'_{\gamma,\gamma'}(j_0(H))$ is a polynomial in $j_0(H)$, thus by [Bochner and Montgomery 1945, Theorem 4] the action of $G_0$ on $M_{k,\varepsilon}$ is real-analytic.

By Proposition 3.2 and Theorem 1.4 the map $N : \mathfrak{F} \times G_0 \to \mathfrak{F}$ defined by $N(\phi, \phi', H) = \phi' \circ H \circ \phi^{-1}$ is a free and proper action. For $\varepsilon = +1$ we note that by Lemma 5.1 and Lemma 7.1 the set $\mathfrak{F}$ is a real-analytic manifold, thus from Theorem 7.2 the conclusion for $\varepsilon = +1$ follows.

Next we show the claim for $\varepsilon = -1$. According to Lemma 7.1, for $k = 1, 2$ we set $N_k := \{G^-_{k+1,s} \mid s > 0\}$ and $N_0 := N_1 \cap N_2 = \{G^-_{2,1/2}\}$. The corresponding preimages are denoted by $M_k := \Pi^{-1}(N_k) \subset \mathfrak{F}$, so that $M_0 := M_1 \cap M_2 = \Pi^{-1}(N_0)$. Now set $M := M_1 \cup M_2$. By Lemma 7.1 for $k = 1, 2$ we have that $M_k$ is a real-analytic submanifold of $\mathfrak{F}$. We obtain by Theorem 7.2 that locally $M_k$ is real-analytically diffeomorphic to $G_0 \times S_k$, where $S_k$ is a real submanifold with $\dim_{\mathbb{R}}(S_k) = \dim_{\mathbb{R}}(M_k) - \dim_{\mathbb{R}}(G_0) = 1$, by Lemma 7.1. By Proposition 2.6 it is possible to normalize any element in $S_k$ with unique isotropies which depend real-analytically on elements of $S_k$. Thus, since $\dim_{\mathbb{R}}(N_k) = 1$, we map $S_k$ to $N_k$ via real-analytic diffeomorphisms. We obtain that for $k = 1, 2$ there exists an open neighborhood $U_k \subset \mathfrak{F}$ of $N_0$ and a real-analytic diffeomorphism $\phi_k : U_k \to V_k$ such that $\phi_k(U_k \cap M_k) = (G_0 \times N_k) \cap V_k$, where $V_k$ is an open neighborhood of $N_0' := \{\mathrm{id}\} \times N_0 \subset G_0 \times M$, with $\mathrm{id} = (\mathrm{id}_{\mathbb{C}^2}, \mathrm{id}_{\mathbb{C}^3})$. Moreover, we have that $\phi_k(U_k \cap N_k) = (\{\mathrm{id}\} \times N_k) \cap V_k$ and $\phi_k$ satisfies the properties given in Theorem 7.2. We define $\phi : U_0 \to V_0$, $\phi(x) := \phi_k(x)$ for $x \in U_0 \cap U_k$, where $k = 1, 2$ and $V_0 = V_1 \cup V_2$ is an open neighborhood of $N_0'$. Write $\widetilde{U} := U_1 \cap U_2 \cap U_0 \subset \mathfrak{F}$ for an open neighborhood of $N_0$. Then we have $\phi|_{\widetilde{U}} = \phi_1|_{\widetilde{U}} = \phi_2|_{\widetilde{U}}$, which implies that $\phi$ is a real-analytic diffeomorphism. Furthermore, since

$$\mathrm{image}(\phi_1|_{\widetilde{U}\cap M}) = \mathrm{image}(\phi_2|_{\widetilde{U}\cap M}) = (G_0 \times N_0) \cap \widetilde{V},$$

where $\widetilde{V}$ is an open neighborhood of $N_0' \subset G_0 \times M$, the mapping $\phi$ locally maps $M_0$ real-analytic diffeomorphically to $G_0 \times N_0$.

Finally the last statement follows from Theorem 7.2, since if $\mathfrak{F}$ were a smooth manifold, then by the smooth version of Theorem 7.2, the quotient $\mathfrak{N}$ would have to be a smooth manifold, which is not the case. □

## 8. Homeomorphic variations of normal forms

In the following we use the notation from Definition 2.4.

**Definition 8.1.** Let $\mathcal{H}$ be a subset of $\mathcal{H}(M, p; M', p')$. A proper subset $\mathcal{K} \subsetneq \mathcal{H}$ is called a *normal form for* $\mathcal{H}$ if for each $[F] \in \mathcal{H}/{\sim}$, there exists a unique representative $G \in \mathcal{K} \cap [F]$. We denote the mapping which assigns to each $H \in \mathcal{H}$ the representative $G \in \mathcal{K} \cap [H]$ as $\pi : \mathcal{H} \to \mathcal{K}$. A normal form $\mathcal{K}$ for $\mathcal{H}$ is called *admissible* if $\pi : \mathcal{H} \to \mathcal{K}$ is continuous.

The uniqueness of the representative $F \in \mathcal{K} \cap [F]$ in Definition 8.1 is not a restriction. Assume we have another representative $F \neq G \in \mathcal{K}$ in the class $[F]$, then $G$ is equivalent to $F$, hence it suffices to choose exactly one element from the set of all representatives which belong to $\mathcal{K} \cap [F]$. If there exists an admissible normal form $\mathcal{K}$ for $\mathcal{H}$ we observe that in each orbit of any not necessarily admissible normal form $\mathcal{K}'$ for $\mathcal{H}$, there exists an element of $\mathcal{K}$.

**Theorem 8.2.** *Let $\mathcal{N}'$ be an admissible normal form for $\mathcal{F}$. Then $\mathcal{N}'$ is homeomorphic to $\mathcal{N}$, where we equip $\mathcal{N}'$ and $\mathcal{N}$ with $\tau_J$.*

*Proof.* Let us denote by $\pi' : \mathcal{F} \to \mathcal{N}'$ the continuous mapping as in Definition 8.1. We note that the class $\mathcal{N}$ introduced in Proposition 2.6 is an admissible normal form for $\mathcal{F}$ as in Definition 2.5. If we equip $\mathcal{F}$ with $\tau_J$, we obtain by Proposition 4.3 that the mapping $\pi : \mathcal{F} \to \mathcal{N}$, $H \mapsto \sigma' \circ H \circ \sigma^{-1}$ is continuous.

$$\begin{tikzcd}
 & \mathcal{F} \arrow[dl, shift right, "\pi'"'] \arrow[dr, shift left, "\pi"] & \\
\mathcal{N}' \arrow[ur, shift right, "\mathrm{incl}'"'] \arrow[rr, "\psi"'] & & \mathcal{N} \arrow[ul, shift left, "\mathrm{incl}"]
\end{tikzcd}$$

**Figure 2.** Diagram for admissible normal forms.

In Figure 2, the mapping $\mathrm{incl}' : \mathcal{N}' \to \mathcal{F}$ is the inclusion mapping, which is given by $\mathrm{incl}'(H) := H$ for all $H \in \mathcal{N}'$, and similarly for $\mathrm{incl} : \mathcal{N} \to \mathcal{F}$. The map $\psi : \mathcal{N}' \to \mathcal{N}$ is given by $\psi(H) := F$ for $H \in \mathcal{N}'$ and $F \in \mathcal{N} \cap [H]$. Since $\mathcal{N}'$ and $\mathcal{N}$ are normal forms, we obtain that $\psi$ is a bijective mapping. Furthermore, since $\psi = \pi \circ \mathrm{incl}'$ and $\psi^{-1} = \pi' \circ \mathrm{incl}$ are compositions of continuous mappings, we obtain that $\psi$ is a homeomorphism. $\square$

**Example 8.3.** Beginning with $\mathcal{N}$, we can construct different admissible normal forms $\mathcal{N}'$ as follows. We fix a pair of isotropies $(\phi_0, \phi_0') \in G_0$ and consider the isotropies $(\tilde{\phi}, \tilde{\phi}') \in G_0$ from Proposition 2.6, such that $\pi : \mathcal{F} \to \mathcal{N}$ is given by $\pi(H) := \tilde{\phi}' \circ H \circ \tilde{\phi}^{-1}$, denoted by $\widehat{H}$. We define $\phi := \phi_0 \circ \tilde{\phi}$ and $\phi' := \phi_0' \circ \tilde{\phi}'$, to obtain for $F \in \mathcal{F}$ that

$$\phi' \circ F \circ \phi^{-1} = \phi_0' \circ \tilde{\phi}' \circ F \circ \tilde{\phi}^{-1} \circ \phi_0^{-1} = \phi_0' \circ \widehat{F} \circ \phi_0^{-1},$$

where $\widehat{F} \in \mathcal{N}$. We define $\mathcal{N}' := \{\phi_0' \circ \widehat{F} \circ \phi_0^{-1} \mid \widehat{F} \in \mathcal{N}\}$. As observed above $\pi$ induces an admissible normal form, which implies that the mapping $\pi' : \mathcal{F} \to \mathcal{N}'$ given by $\pi'(F) := \phi' \circ F \circ \phi^{-1}$ is continuous and $\mathcal{N}'$ is an admissible normal form.

## Acknowledgments

We would like to thank Bernhard Lamel, Giuseppe della Sala and Nordine Mir for their interest in this work and many discussions related to the topic of this article.

## References

[Baouendi et al. 1997] M. S. Baouendi, P. Ebenfelt, and L. P. Rothschild, “Parametrization of local biholomorphisms of real analytic hypersurfaces”, *Asian J. Math.* **1**:1 (1997), 1–16. MR 99b:32022 Zbl 0943.32021

[Baouendi et al. 1999] M. S. Baouendi, P. Ebenfelt, and L. P. Rothschild, “Rational dependence of smooth and analytic CR mappings on their jets”, *Math. Ann.* **315**:2 (1999), 205–249. MR 2001b:32075 Zbl 0942.32027

[Baouendi et al. 2004] M. S. Baouendi, L. P. Rothschild, J. Winkelmann, and D. Zaitsev, “Lie group structures on groups of diffeomorphisms and applications to CR manifolds”, *Ann. Inst. Fourier (Grenoble)* **54**:5 (2004), 1279–1303. MR 2005k:32048 Zbl 1062.22046

[Bochner and Montgomery 1945] S. Bochner and D. Montgomery, "Groups of differentiable and real or complex analytic transformations", *Ann. of Math.* (2) **46** (1945), 685–694. MR 7,241d Zbl 0061.04406

[Duistermaat and Kolk 2000] J. J. Duistermaat and J. A. C. Kolk, *Lie groups*, Springer, Berlin, 2000. MR 2001j:22008 Zbl 0955.22001

[Faran 1982] J. J. Faran, "Maps from the two-ball to the three-ball", *Invent. Math.* **68**:3 (1982), 441–475. MR 83k:32038 Zbl 0519.32016

[Heinzner et al. 1996] P. Heinzner, A. T. Huckleberry, and F. Kutzschebauch, "A real analytic version of Abels' theorem and complexifications of proper Lie group actions", pp. 229–273 in *Complex analysis and geometry* (Trento, 1993), edited by V. Ancona et al., Lecture Notes in Pure and Appl. Math. **173**, Dekker, New York, 1996. MR 96j:57047 Zbl 0861.32011

[Illman and Kankaanrinta 2000] S. Illman and M. Kankaanrinta, "Three basic results for real analytic proper $G$-manifolds", *Math. Ann.* **316**:1 (2000), 169–183. MR 2001f:58025 Zbl 0945.57016

[Juhlin and Lamel 2013] R. Juhlin and B. Lamel, "Automorphism groups of minimal real-analytic CR manifolds", *J. Eur. Math. Soc.* (*JEMS*) **15**:2 (2013), 509–537. MR 3017044 Zbl 1267.32013

[Kim and Zaitsev 2005] S.-Y. Kim and D. Zaitsev, "Equivalence and embedding problems for CR-structures of any codimension", *Topology* **44**:3 (2005), 557–584. MR 2005j:32037 Zbl 1079.32022

[Kowalski 2005] R. T. Kowalski, "Rational jet dependence of formal equivalences between real-analytic hypersurfaces in $\mathbb{C}^2$", *Pacific J. Math.* **220**:1 (2005), 107–139. MR 2007h:32054 Zbl 1106.32025

[Lamel 2001] B. Lamel, "Holomorphic maps of real submanifolds in complex spaces of different dimensions", *Pacific J. Math.* **201**:2 (2001), 357–387. MR 2003e:32066 Zbl 1061.32027

[Lamel and Mir 2007] B. Lamel and N. Mir, "Parametrization of local CR automorphisms by finite jets and applications", *J. Amer. Math. Soc.* **20**:2 (2007), 519–572. MR 2008b:32025 Zbl 1112.32017

[Lamel et al. 2008] B. Lamel, N. Mir, and D. Zaitsev, "Lie group structures on automorphism groups of real-analytic CR manifolds", *Amer. J. Math.* **130**:6 (2008), 1709–1726. MR 2009k:32044 Zbl 1165.32017

[Lebl 2011] J. Lebl, "Normal forms, Hermitian operators, and CR maps of spheres and hyperquadrics", *Michigan Math. J.* **60**:3 (2011), 603–628. MR 2861091 Zbl 1237.32005

[Palais 1961] R. S. Palais, "On the existence of slices for actions of non-compact Lie groups", *Ann. of Math.* (2) **73** (1961), 295–323. MR 23 #A3802 Zbl 0103.01802

[Reiter 2014] M. Reiter, *Holomorphic mappings of hyperquadrics from $\mathbb{C}^2$ to $\mathbb{C}^3$*, thesis, University of Vienna, 2014, Available at http://othes.univie.ac.at/33603.

[Reiter 2015] M. Reiter, "Classification of holomorphic mappings of hyperquadrics from $\mathbb{C}^2$ to $\mathbb{C}^3$", *J. Geom. Anal.* (online publication March 2015).

[Wolfram 2008] Wolfram Research, "Mathematica", 2008, Available at http://www.wolfram.com/mathematica. version 7.0.1.0, Champaign, IL.

Received November 6, 2014. Revised May 20, 2015.

MICHAEL REITER
FACULTY OF MATHEMATICS
UNIVERSITY OF VIENNA
OSKAR-MORGENSTERN-PLATZ 1
1090 VIENNA
AUSTRIA
m.reiter@univie.ac.at

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.475

# 2-BLOCKS WITH MINIMAL NONABELIAN DEFECT GROUPS III

BENJAMIN SAMBALE

**We prove that two 2-blocks of (possibly different) finite groups with a common minimal nonabelian defect group and the same fusion system are isotypic (and therefore perfectly isometric) in the sense of Broué. This continues former work by Cabanes and Picaronny (*J. Fac. Sci. Univ. Tokyo Sect. IA Math.* 39:1 (1992), 141–161), Sambale (*J. Algebra* 337 (2011), 261–284) and Eaton et al. (*J. Group Theory* 15:3 (2012), 311–321).**

## 1. Introduction

Since its appearance in 1990, Broué's abelian defect conjecture gained much attention among representation theorists. On the level of characters it predicts the existence of a perfect isometry between a block with abelian defect group and its Brauer correspondent. These blocks have a common defect group and the same fusion system. Although Broué's conjecture is false for nonabelian defect groups (see [Cliff 2000]), one can still ask if perfect isometries or even isotypies exist. We affirmatively answer this question for $p = 2$ and minimal nonabelian defect groups (see Theorem 9 below). These are the nonabelian defect groups such that any proper subgroup is abelian. Doing so, we verify the character-theoretic version of Rouquier's conjecture [2001, A.2] in this special case (see Corollary 10 below). At the same time we provide a new infinite family of defect groups supporting a blockwise $Z^*$-Theorem.

By Rédei's classification of minimal nonabelian $p$-groups, one has to consider three distinct families of defect groups. For two of these families the result already appeared in the literature (see [Cabanes and Picaronny 1992; Sambale 2011; Eaton et al. 2012]). Hence, it suffices to handle the remaining family which we will do in the next section. The proof of the main result is an application of [Horimoto and Watanabe 2012, Theorem 2]. The last section of the present paper also contains a related result for the nonabelian defect group of order 27 and exponent 9.

Our notation is fairly standard. We consider blocks $B$ of finite groups with respect to a $p$-modular system $(K, \mathcal{O}, F)$ where $\mathcal{O}$ is a complete discrete valuation

*MSC2010:* 20C15, 20C20.
*Keywords:* minimal nonabelian defect groups, perfect isometries, isotypies.

ring with quotient field $K$ of characteristic 0 and field of fractions $F$ of characteristic $p$. As usual, we assume that $K$ is "large" enough and $F$ is algebraically closed. The number of irreducible ordinary characters (resp. Brauer characters) of $B$ is denoted by $k(B)$ (resp. $l(B)$). Moreover, $k_i(B)$ is the number of those irreducible characters of $B$ which have height $i \geq 0$. For other results on block invariants and fusion systems we often refer to [Sambale 2014]. Moreover, for the definition and construction of perfect isometries we follow [Broué and Puig 1980a; Cabanes and Picaronny 1992]. A cyclic group of order $n \in \mathbb{N}$ is denoted by $C_n$.

## 2. A class of minimal nonabelian defect groups

Let $B$ be a non-nilpotent 2-block of a finite group $G$ with defect group

(1) $$D = \langle x, y \mid x^{2^r} = y^2 = [x, y]^2 = [x, x, y] = [y, x, y] = 1\rangle \cong C_2^2 \rtimes C_{2^r}$$

where $r \geq 2$, $[x, y] := xyx^{-1}y^{-1}$ and $[x, x, y] := [x, [x, y]]$.

We have already investigated some properties of $B$ in [Sambale 2011], and later gave simplified proofs in [Sambale 2014, Chapter 12]. For the convenience of the reader we restate some of these results.

**Lemma 1** [Sambale 2014, Lemma 12.3]. *Let* $z := [x, y]$. *Then*:

(i) $\Phi(D) = \mathrm{Z}(D) = \langle x^2, z\rangle \cong C_{2^{r-1}} \times C_2$.

(ii) $D' = \langle z\rangle \cong C_2$.

(iii) $|\mathrm{Irr}(D)| = 5 \cdot 2^{r-1}$.

Recall that a (saturated) fusion system $\mathcal{F}$ on a $p$-group $P$ determines the following subgroups:

$$\mathrm{Z}(\mathcal{F}) := \{x \in P : x \text{ is fixed by every morphism in } \mathcal{F}\},$$
$$\mathfrak{foc}(\mathcal{F}) := \langle f(x)x^{-1} : x \in Q \leq P,\ f \in \mathrm{Aut}_{\mathcal{F}}(Q)\rangle,$$
$$\mathfrak{hyp}(\mathcal{F}) := \langle f(x)x^{-1} : x \in Q \leq P,\ f \in \mathrm{O}^p(\mathrm{Aut}_{\mathcal{F}}(Q))\rangle.$$

**Lemma 2.** *The fusion system $\mathcal{F}$ of $B$ is the constrained fusion system of the finite group $A_4 \rtimes C_{2^r}$ where $C_{2^r}$ acts as a transposition in* $\mathrm{Aut}(A_4) \cong S_4$. *In particular, $B$ has inertial index* 1 *and $Q := \langle x^2, y, z\rangle \cong C_{2^{r-1}} \times C_2^2$ is the only $\mathcal{F}$-essential subgroup of $D$. Moreover*, $\mathrm{Aut}_{\mathcal{F}}(Q) \cong S_3$. *Without loss of generality*, $\mathrm{Z}(\mathcal{F}) = \langle x^2\rangle$ *and* $\mathfrak{hyp}(B) = \mathfrak{foc}(B) = \mathfrak{foc}(\mathcal{F}) = \langle y, z\rangle$.

*Proof.* We have seen in [Sambale 2014, Proposition 12.7] that $\mathcal{F}$ is constrained and coincides with the fusion system of $A_4 \rtimes C_{2^r}$. The construction of the semidirect product $A_4 \rtimes C_{2^r}$ is slightly different in [Sambale 2014], but it is easy to see that both constructions give isomorphic groups. The remaining claims follow from the proof of [Sambale 2014, Proposition 12.7]. □

By a result of Watanabe [2014, Theorem 3 and Lemma 3], the hyperfocal subgroup of a 2-block is trivial or noncyclic. Hence, our situation with a Klein-four (hyper)focal subgroup represents the first nontrivial example in some sense. Recall that a $B$-*subsection* is a pair $(u, b_u)$ such that $u \in D$ and $b_u$ is a Brauer correspondent of $B$ in $\mathrm{C}_G(u)$.

**Lemma 3.** *The set* $\mathcal{R} := \mathrm{Z}(D) \cup \{x^i y^j : i, j \in \mathbb{Z},\ i \text{ odd}\}$ *is a set of representatives for the* $\mathcal{F}$*-conjugacy classes of* $D$ *with* $|\mathcal{R}| = 2^{r+1}$. *For* $u \in \mathcal{R}$ *let* $(u, b_u)$ *be a* $B$*-subsection. Then* $b_u$ *has defect group* $\mathrm{C}_D(u)$. *Moreover*, $l(b_u) = 1$ *whenever* $u \in \mathcal{R} \setminus \langle x^2 \rangle$.

*Proof.* By Lemma 2, it is easy to see that $\mathcal{R}$ is in fact a set of representatives for the $\mathcal{F}$-conjugacy classes of $D$. Observe that $\langle u \rangle$ is fully $\mathcal{F}$-normalized for all $u \in \mathcal{R}$. Hence, by [Sambale 2014, Lemma 1.34], $b_u$ has defect group $\mathrm{C}_D(u)$ and fusion system $\mathrm{C}_{\mathcal{F}}(\langle u \rangle)$. It is easy to see that $\mathrm{C}_{\mathcal{F}}(\langle u \rangle)$ is trivial unless $u \in \mathrm{Z}(\mathcal{F}) = \langle x^2 \rangle$. This shows $l(b_u) = 1$ for $u \in \mathcal{R} \setminus \langle x^2 \rangle$. □

**Theorem 4** [Sambale 2014, Theorem 12.4]. *We have* $k(B) = 5 \cdot 2^{r-1}$, $k_0(B) = 2^{r+1}$, $k_1(B) = 2^{r-1}$ *and* $l(B) = 2$.

*Proof.* By Lemma 2, we have $|D : \mathfrak{foc}(B)| = 2^r$. In particular, $2^r \mid k_0(B)$ by [Robinson 2008, Theorem 1]. Moreover, [Kessar et al. 2015, Theorem 1.1] implies $2^{r+1} \le k_0(B)$. By Lemma 3 we have $l(b_x) = 1$. Thus, we obtain $k_0(B) = 2^{r+1}$ by a result of Robinson (see [Sambale 2014, Theorem 4.12]). In order to determine $l(B)$, we use induction on $r$. Let $u := x^2$. Then $b_u$ dominates a block $\overline{b_u}$ of $\mathrm{C}_G(u)/\langle u \rangle$ with defect group $\bar{D} := D/\langle u \rangle \cong D_8$ and fusion system $\bar{\mathcal{F}} := \mathcal{F}/\langle u \rangle$. By [Linckelmann 2007, Theorem 6.3], $\langle x^2, y, z \rangle/\langle u \rangle \cong C_2^2$ is the only $\bar{\mathcal{F}}$-essential subgroup of $\bar{D}$. Therefore, a result of Brauer (see [Sambale 2014, Theorem 8.1]) shows that $l(b_u) = l(\overline{b_u}) = 2$. By Lemma 3 and [Sambale 2014, Theorem 1.35] it follows that $k(B) > k_0(B)$. Since $|\mathrm{Z}(D) : \mathrm{Z}(D) \cap \mathfrak{foc}(B)| = 2^{r-1}$, we have $2^{r-1} \mid k_i(B)$ for $i \ge 1$ by [Robinson 2008, Theorem 2]. Thus, by [Robinson 1991, Theorem 3.4] we obtain

$$2^{r+2} \le k_0(B) + 4(k(B) - k_0(B)) \le \sum_{i=0}^{\infty} k_i(B) 2^{2i} \le |D| = 2^{r+2}.$$

This gives $k_1(B) = 2^{r-1}$ and $k(B) = k_0(B) + k_1(B) = 5 \cdot 2^{r-1}$. In case $r = 2$, [Sambale 2014, Theorem 1.35] implies

$$l(B) = k(B) - \sum_{1 \ne u \in \mathcal{R}} l(b_u) = 10 - 8 = 2.$$

Now let $r \ge 3$ and $1 \ne \langle u \rangle < \langle x^2 \rangle$. Then $\overline{b_u}$ as above has the same type of defect group as $B$ except that $r$ is smaller. Hence, induction gives $l(b_u) = l(\overline{b_u}) = 2$. Now the claim $l(B) = 2$ follows again by [Sambale 2014, Theorem 1.35]. □

In the following results we denote the set of irreducible characters of $B$ of height $i$ by $\mathrm{Irr}_i(B)$.

**Proposition 5** [Sambale 2014, Proposition 12.9]. *The set* $\mathrm{Irr}_0(B)$ *contains four* 2*-rational characters and two families of* 2*-conjugate characters of size* $2^i$ *for every* $i = 1, \dots, r-1$. *The characters of height* 1 *split into two* 2*-rational characters and one family of* 2*-conjugate characters of size* $2^i$ *for every* $i = 2, \dots, r-2$.

**Proposition 6.** *There are* 2*-rational characters* $\chi_i \in \mathrm{Irr}(B)$ *for* $i = 1, 2, 3$ *such that*

$$\mathrm{Irr}_0(B) = \{\chi_i * \lambda : i = 1, 2,\ \lambda \in \mathrm{Irr}(D/\mathfrak{foc}(B))\},$$
$$\mathrm{Irr}_1(B) = \{\chi_3 * \lambda : \lambda \in \mathrm{Irr}(\mathrm{Z}(D)\mathfrak{foc}(B)/\mathfrak{foc}(B))\}.$$

*In particular, the characters of height* 1 *have the same degree and*

$$|\{\chi(1) : \chi \in \mathrm{Irr}_0(B)\}| \le 2.$$

*Proof.* We have already seen in the proof of Theorem 4 that the action of $D/\mathfrak{foc}(B)$ on $\mathrm{Irr}_0(B)$ via the $*$-construction has two orbits, and the action of $\mathrm{Z}(D)\mathfrak{foc}(B)/\mathfrak{foc}(B)$ on $\mathrm{Irr}_1(B)$ is regular. By Proposition 5 we can choose 2-rational representatives for these orbits, having identified the sets $\mathrm{Irr}(D/\mathfrak{foc}(B))$ and $\mathrm{Irr}(\mathrm{Z}(D)\mathfrak{foc}(B)/\mathfrak{foc}(B))$ with subsets of $\mathrm{Irr}(D)$ in an obvious manner. $\square$

In the situation of Proposition 6 it is conjectured that $\chi_1(1) \neq \chi_2(1)$ (see [Malle and Navarro 2011]).

**Proposition 7** [Sambale 2014, Proposition 12.8]. *The Cartan matrix of $B$ is given by*

$$2^{r-1}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$$

*up to basic sets.*

Observe that Proposition 7 also gives the Cartan matrix for the defect group $D_8$ and the corresponding fusion system (this would be the case $r = 1$).

Now we are in a position to obtain the generalized decomposition matrix of $B$. This completes partial results in [Sambale 2011, Section 3.3].

**Proposition 8.** *Let $\mathcal{R}$ and $\chi_i$ be as in Lemma 3 and Proposition 6 respectively. Then there are basic sets for $b_u$ ($u \in \mathcal{R}$) and signs $\epsilon, \sigma \in \{\pm 1\}$ such that the generalized decomposition numbers of $B$ have the following form*:

| $u$ | $x^{2i}$ | $x^{2i}z$ | $x^{2i+1}$ | $x^{2i+1}y$ |
|---|---|---|---|---|
| $d^u_{\chi_1\varphi}$ | $(1, 0)$ | $1$ | $1$ | $1$ |
| $d^u_{\chi_2\varphi}$ | $(0, \epsilon)$ | $\epsilon$ | $\epsilon$ | $-\epsilon$ |
| $d^u_{\chi_3\varphi}$ | $(\sigma, \sigma)$ | $-2\sigma$ | $0$ | $0$ |

*Proof.* Since the Galois group of $\mathbb{Q}(e^{2\pi i/2^r})$ over $\mathbb{Q}$ acts on the columns of the generalized decomposition matrix (see Proposition 5), we only need to determine the numbers $d^u_{\chi_i\varphi}$ for $u \in \{x, xy, x^{2^j}, x^{2^j}z\}$ ($i = 1, 2, 3,\ j = 1, \dots, r$). First let $u = x$. Then the orthogonality relations show that

$$2^r|d^x_{\chi_1\varphi}|^2 + 2^r|d^x_{\chi_2\varphi}|^2 + 2^{r-1}|d^x_{\chi_3\varphi}|^2 = 2^{r+1}.$$

Since $\chi_1$ and $\chi_2$ have height 0, we have $d^x_{\chi_1\varphi} \neq 0 \neq d^x_{\chi_2\varphi}$ (see [Sambale 2014, Proposition 1.36]). It follows that $d^x_{\chi_i\varphi} = \pm 1$ for $i = 1, 2$ and $d^x_{\chi_3\varphi} = 0$, because $\chi_i$ is 2-rational. By replacing $\varphi$ with $-\varphi$ if necessary (i.e., changing the basic set for $b_x$), we may assume that $d^x_{\chi_1\varphi} = 1$. We set $d^x_{\chi_2\varphi} =: \epsilon_0$. Similarly, we obtain $d^{xy}_{\chi_1\varphi} = 1$, $d^{xy}_{\chi_2\varphi} = \pm 1$ and $d^{xy}_{\chi_3\varphi} = 0$. Now since the columns $d^x$ and $d^{xy}$ of the generalized decomposition matrix are orthogonal, we obtain $d^{xy}_{\chi_2\varphi} = -\epsilon_0$.

Now let $u := x^{2^j}$ for some $j \in \{1, \dots, r\}$. Let $\mathrm{IBr}(b_u) = \{\varphi_1, \varphi_2\}$ (see the proof of Theorem 4). Then by Proposition 7 we get

$$\begin{aligned} 2^r|d^u_{\chi_1\varphi_1}|^2 + 2^r|d^u_{\chi_2\varphi_1}|^2 + 2^{r-1}|d^u_{\chi_3\varphi_1}|^2 &= 3 \cdot 2^{r-1}, \\ 2^r|d^u_{\chi_1\varphi_2}|^2 + 2^r|d^u_{\chi_2\varphi_2}|^2 + 2^{r-1}|d^u_{\chi_3\varphi_2}|^2 &= 3 \cdot 2^{r-1}, \\ 2^r d^u_{\chi_1\varphi_1}\overline{d^u_{\chi_1\varphi_2}} + 2^r d^u_{\chi_2\varphi_1}\overline{d^u_{\chi_2\varphi_2}} + 2^{r-1} d^u_{\chi_3\varphi_1}\overline{d^u_{\chi_3\varphi_2}} &= 2^{r-1}. \end{aligned}$$

Obviously, $d^u_{\chi_1\varphi_1}d^u_{\chi_2\varphi_1} = 0$ and we may assume that $(d^u_{\chi_1\varphi_1}, d^u_{\chi_1\varphi_2}) = (1, 0)$ and $(d^u_{\chi_2\varphi_1}, d^u_{\chi_2\varphi_2}) = (0, \epsilon_j)$ for a sign $\epsilon_j \in \{\pm 1\}$. Moreover, $d^u_{\chi_3\varphi_1} = d^u_{\chi_3\varphi_2} =: \sigma_j \in \{\pm 1\}$. Now let $u := x^{2^j}z$. Then we have

$$2^r|d^u_{\chi_1\varphi}|^2 + 2^r|d^u_{\chi_2\varphi}|^2 + 2^{r-1}|d^u_{\chi_3\varphi}|^2 = 2^{r+2}.$$

It is known that $2 \mid d^u_{\chi_3\varphi} \neq 0$, since $b_u$ is major (see [Sambale 2014, Proposition 1.36]). This gives $d^u_{\chi_1\varphi} = 1$, $d^u_{\chi_2\varphi} = \pm 1$ and $d^u_{\chi_3\varphi} = \pm 2$. By the orthogonality to $d^{x^{2^j}}$ we obtain that $d^u_{\chi_3\varphi} = -2\sigma_j$ and $d^u_{\chi_2\varphi} = \epsilon_j$.

It remains to show that the signs $\epsilon_j$ and $\sigma_j$ do not depend on $j$. For this we consider characters $\lambda, \psi \in \mathrm{Irr}(D)$ whose values are given as follows:

| | $x^{2^j}$ | $x^{2^j}z$ | $x$ | $xy$ |
|---|---|---|---|---|
| $\lambda$ | 1 | 1 | 1 | $-1$ |
| $\psi$ | 2 | $-2$ | 0 | 0 |

Observe that $\psi$ is the inflation of the irreducible character of $D/\langle x^2\rangle \cong D_8$ of degree 2. It is easy to see that $(\lambda + \psi)(x^{2k}y) = -1 = 1 - 2 = (\lambda + \psi)(x^{2k}z)$ for every $k \in \mathbb{Z}$. It follows that $\lambda + \psi$ is $\mathcal{F}$-stable, i.e., $(\lambda + \psi)(u) = (\lambda + \psi)(v)$ whenever $u$ and $v$ are $\mathcal{F}$-conjugate. By [Broué and Puig 1980a], $\chi_1 * (\lambda + \psi)$ is a generalized character of $B$. In particular, the scalar product $(\chi_1 * (\lambda + \psi), \chi_3)_G$ is an integer. This number can be computed by using the so-called contribution numbers $m^u_{\chi_1\chi_3} := d^u_{\chi_1} C_u^{-1} \overline{d^u_{\chi_3}}^{\mathrm{T}}$ where $C_u$ is the Cartan matrix of $b_u$ and $d^u_{\chi_i}$ is the row of the generalized decomposition matrix corresponding to $(u, b_u)$ and $\chi_i$. For $u = x^{2^j}$ we have

$$C_u^{-1} = 2^{-r-2}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

by Proposition 7. This gives $m^u_{\chi_1\chi_3} = 2^{-r-1}\sigma_j$. Similarly, $m^u_{\chi_1\chi_3} = -2^{-r-1}\sigma_j$ for $u = x^{2^j}z$. Thus, we obtain

$$\begin{aligned}(\chi_1 * (\lambda+\psi), \chi_3)_G &= \sum_{u\in\mathcal{R}} (\lambda+\psi)(u) m^u_{\chi_1\chi_3} = \sum_{u\in \mathrm{Z}(D)} (\lambda+\psi)(u) m^u_{\chi_1\chi_3} \\ &= (3+1)\Bigl(2^{-r-1}\sigma_r + 2^{-r-1}\sum_{j=1}^{r-1}\sigma_j 2^{r-j-1}\Bigr) \\ &= 2^{-r+1}\sigma_r + \sum_{j=1}^{r-1}\sigma_j 2^{-j}.\end{aligned}$$

If $\sigma_1 = \sigma_j$ for some $j \neq 1$, then it follows immediately that $\sigma_1 = \cdots = \sigma_r$ (otherwise the scalar product above is not an integer). Now suppose that $-\sigma_1 = \sigma_2 = \cdots = \sigma_r$. In this case we replace $\chi_3$ by the 2-rational character $\chi_3 * \tau$ where $\tau \in \mathrm{Irr}(\mathrm{Z}(D)\mathfrak{foc}(B)/\mathfrak{foc}(B))$ such that $\tau(x^2) = -1$. This changes $\sigma_1$, but does not affect $\sigma_j$ for $j > 1$.

A similar argument with the scalar product $(\chi_2 * (\lambda+\psi), \chi_3)_G$ implies that $\epsilon_1 = \cdots = \epsilon_r$. In case $\epsilon_0 = -\epsilon_1$, we replace $\chi_2$ by $\chi_2 * \tau$ where $\tau \in \mathrm{Irr}(D/\mathfrak{foc}(B))$ such that $\tau(x) = -1$. Observe again that this changes $\epsilon_0$, but keeps $\epsilon_j$ for $j > 0$. This completes the proof. □

## 3. The main result

**Theorem 9.** *Let $B$ and $\tilde{B}$ be 2-blocks of (possibly different) finite groups with a common minimal nonabelian defect group and the same fusion system. Then $B$ and $\tilde{B}$ are isotypic (and therefore perfectly isometric).*

*Proof.* We may assume that $B$ is not nilpotent by [Broué and Puig 1980b]. Let $D$ be a defect group of $B$ and $\tilde{B}$. If $|D| = 8$, then the claim follows from [Cabanes and Picaronny 1992]. Now suppose that $D$ is given as in (1). We will attach a tilde to everything associated with $\tilde{B}$. By Proposition 8 and [Horimoto and Watanabe 2012, Theorem 2] there is a perfect isometry $I : \mathrm{CF}(G, B) \to \mathrm{CF}(\tilde{G}, \tilde{B})$ where $\mathrm{CF}(G, B)$ denotes the space of class functions with basis $\mathrm{Irr}(B)$ over $K$. It remains to show that $I$ is also an isotypy. In order to do so, we follow [Cabanes and Picaronny 1992, Section V.2]. For each $u \in D$ let $\mathrm{CF}(\mathrm{C}_G(u)_{2'}, b_u)$ be the space of class functions on $\mathrm{C}_G(u)$ which vanish on the $p$-singular classes and are spanned by $\mathrm{IBr}(b_u)$. The decomposition map $d^u_G : \mathrm{CF}(G, B) \to \mathrm{CF}(\mathrm{C}_G(u)_{2'}, b_u)$ is defined by

$$d^u_G(\chi)(s) := \chi(e_{b_u} us) = \sum_{\varphi \in \mathrm{IBr}(b_u)} d^u_{\chi\varphi}\varphi(s)$$

for $\chi \in \mathrm{Irr}(B)$ and $s \in \mathrm{C}_G(u)_{2'}$ where $e_{b_u}$ is the block idempotent of $b_u$ over $\mathcal{O}$. Then $I$ determines isometries

$$I^u : \mathrm{CF}(\mathrm{C}_G(u)_{2'}, b_u) \to \mathrm{CF}(\mathrm{C}_{\tilde{G}}(u)_{2'}, \tilde{b}_u)$$

by the equation $d^u_{\tilde{G}} \circ I = I^u \circ d^u_G$. Note that $I^1$ is the restriction of $I$. We need to show that $I^u$ can be extended to a perfect isometry $\widehat{I^u} : \mathrm{CF}(\mathrm{C}_G(u), b_u) \to \mathrm{CF}(\mathrm{C}_{\tilde{G}}(u), \tilde{b}_u)$. Suppose first that $b_u$ is nilpotent. Then by Proposition 8, $d^u_G(\chi_1) = \epsilon\varphi$ and $d^u_{\tilde{G}}(I(\chi_1)) = \tilde{\epsilon}\tilde{\varphi}$ where $\mathrm{IBr}(b_u) = \{\varphi\}$ and $\mathrm{IBr}(\tilde{b}_u) = \{\tilde{\varphi}\}$ for some signs $\epsilon, \tilde{\epsilon} \in \{\pm 1\}$. It follows that $I^u(\varphi) = \epsilon\tilde{\epsilon}\tilde{\varphi}$. Let $\psi \in \mathrm{Irr}_0(b_u)$ and $\tilde{\psi} \in \mathrm{Irr}_0(\tilde{b}_u)$ be 2-rational characters. Then it is well known that $\varphi = d^1_{\mathrm{C}_G(u)}(\psi)$ and $\mathrm{Irr}(b_u) = \{\psi * \lambda : \lambda \in \mathrm{Irr}(D)\}$ (see [Broué and Puig 1980b]). Therefore, we may define $\widehat{I^u}$ by $\widehat{I^u}(\psi * \lambda) := \epsilon\tilde{\epsilon}\tilde{\psi} * \lambda$ for $\lambda \in \mathrm{Irr}(D)$. Then $\widehat{I^u}$ is a perfect isometry and

$$\widehat{I^u}(\varphi) = \widehat{I^u}(d^1_{\mathrm{C}_G(u)}(\psi)) = d^1_{\mathrm{C}_{\tilde{G}}(u)}(\widehat{I^u}(\psi)) = \epsilon\tilde{\epsilon} d^1_{\mathrm{C}_{\tilde{G}}(u)}(\tilde{\psi}) = \epsilon\tilde{\epsilon}\tilde{\varphi} = I^u(\varphi).$$

Hence, $\widehat{I^u}$ extends $I^u$. Moreover, $\widehat{I^u}$ does not depend on the generator of $\langle u \rangle$, since the signs $\epsilon$ and $\tilde{\epsilon}$ were defined by means of 2-rational characters.

Assume next that $b_u$ is non-nilpotent. Then $u \in \langle x^2 \rangle$ and $b_u$ has defect group $D$. By Proposition 8, we can choose basic sets $\varphi_1, \varphi_2$ (resp. $\tilde{\varphi}_1, \tilde{\varphi}_2$) for $b_u$ (resp. $\tilde{b}_u$) such that $\varphi_i = d^u_G(\chi_i)$ and $\tilde{\varphi}_i = d^u_{\tilde{G}}(I(\chi_i))$ for $i = 1, 2$. Then $I^u(\varphi_i) = \tilde{\varphi}_i$ for $i = 1, 2$. Since the Cartan matrix of $b_u$ with respect to the basic set $\varphi_1, \varphi_2$ is already fixed (and given by Proposition 7), we find 2-rational characters $\psi_i \in \mathrm{Irr}_0(b_u)$ such that $d^1_{\mathrm{C}_G(u)}(\psi_i) = \epsilon_i\varphi_i$ with $\epsilon_i \in \{\pm 1\}$ for $i = 1, 2$ (see the proof of Proposition 8). Similarly, one has $\tilde{\psi}_i \in \mathrm{Irr}_0(\tilde{b}_u)$ such that $d^1_{\mathrm{C}_{\tilde{G}}(u)}(\tilde{\psi}_i) = \tilde{\epsilon}_i\tilde{\varphi}_i$. Then, by what we have already shown, there exists a perfect isometry

$$\widehat{I^u} : \mathrm{CF}(\mathrm{C}_G(u), b_u) \to \mathrm{CF}(\mathrm{C}_{\tilde{G}}(u), \tilde{b}_u)$$

sending $\psi_i$ to $\epsilon_i\tilde{\epsilon}_i\tilde{\psi}_i$ for $i = 1, 2$. We have

$$\widehat{I^u}(\varphi_i) = \epsilon_i\widehat{I^u}(d^1_{\mathrm{C}_G(u)}(\psi_i)) = \epsilon_i d^1_{\mathrm{C}_{\tilde{G}}(u)}(\widehat{I^u}(\psi_i)) = \tilde{\epsilon}_i d^1_{\mathrm{C}_{\tilde{G}}(u)}(\tilde{\psi}_i) = \tilde{\varphi}_i = I^u(\varphi_i)$$

for $i = 1, 2$. This shows that $\widehat{I^u}$ extends $I^u$. Moreover, it is easy to see that $\widehat{I^u}$ does not depend on the generator of $\langle u \rangle$.

Altogether we have proved the theorem if $D$ is given as in (1). By [Sambale 2014, Theorem 12.4] it remains to handle the case

$$D \cong \langle x, y \mid x^{2^r} = y^{2^r} = [x, y]^2 = [x, x, y] = [y, x, y] = 1 \rangle$$

where $r \geq 2$. Here $B$ and $\tilde{B}$ are Morita equivalent and therefore perfectly isometric. However, a Morita equivalence does not automatically provide an isotypy. Nevertheless, in this special case the Morita equivalence is a composition of various "natural" equivalences (namely Fong reductions, Külshammer–Puig reduction and Külshammer's reduction for blocks with normal defect groups, see [Eaton et al. 2012, proof of Theorem 1]). In particular, the generalized decomposition matrices of $B$ and $\tilde{B}$ coincide up to signs (see [Watanabe 1985]). Now we can use the same methods as above in order to construct an isotypy. In fact, for every $B$-subsection $(u, b_u)$ one has that $b_u$ is nilpotent or $u = [x, y]$ and $b_u$ is Morita equivalent to $B$ (see the proof of [Sambale 2011, Proposition 4.3]). We omit the details. □

**Corollary 10.** *Let $B$ be a 2-block of a finite group $G$ with minimal nonabelian defect group $D \ncong D_8$. Then $B$ is isotypic to a Brauer correspondent in* $\mathrm{N}_G(\mathfrak{hyp}(B))$.

*Proof.* Let $b_D$ be a Brauer correspondent of $B$ in $D\,\mathrm{C}_G(D)$. Since $D\,\mathrm{C}_G(D) \subseteq \mathrm{N}_G(\mathfrak{hyp}(B))$, the Brauer correspondent $b := b_D^{\mathrm{N}_G(\mathfrak{hyp}(B))}$ of $B$ has defect group $D$. By Theorem 9, it suffices to show that $B$ and $b$ have the same fusion system. Observe that $\mathrm{N}_G(D, b_D) \subseteq \mathrm{N}_G(\mathfrak{hyp}(B))$. In particular, $B$ and $b$ have the same inertial quotient. If there is only the trivial fusion system on $D$, then we are done (this applies if $D$ is metacyclic of order at least 16). In case $D \cong Q_8$, $B$ is a controlled block (see, e.g., [Cabanes and Picaronny 1992]). Since $B$ and $b$ have the same inertial quotient, it follows that these blocks also have the same fusion system. It remains to consider the two other families of defect groups (see [Sambale 2014, Theorem 12.4]). For one of these families the fusion system is again controlled (see [Sambale 2014, Proposition 12.7]). Finally, if $D$ is given as in (1), then the fusion system is constrained and the automorphisms of the essential subgroup (if it exists) also act on $\mathfrak{hyp}(B)$. Hence, $B$ is nilpotent if and only if $b$ is nilpotent. Again the claim follows from Theorem 9. □

We remark that Corollary 10 would be false in case $D \cong D_8$. The principal 2-block of $\mathrm{GL}(3, 2)$ gives a counterexample. If $B$ is a block of a finite group $G$ with defect group as given in (1), then $B$ is also isotypic to a Brauer correspondent in $\mathrm{C}_G(u)$ where $u \in \mathrm{Z}(\mathcal{F})$. This resembles Glauberman's $\mathrm{Z}^*$-theorem.

In the situation of Theorem 9 (or Corollary 10) it is desirable to extend the isotypies to Morita equivalences (as we did in [Eaton et al. 2012]). This is not always possible if $|D| = 8$, since for example the principal 2-blocks of the symmetric groups $S_4$ and $S_5$ are not Morita equivalent. Nevertheless, the possible Morita equivalence classes in case $|D| = 8$ are known by Erdmann's classification of tame algebra [Erdmann 1990] (at least over $F$, see [Holm 2001]). In view of [Eaton et al. 2012] one may still ask if two non-nilpotent 2-blocks with isomorphic defect groups as in Section 2 are Morita equivalent. We will see that the answer is again negative.

Consider the groups $G_1 := A_4 \rtimes C_{2^r}$ and $G_2 := A_5 \rtimes C_{2^r}$ constructed similarly as in Lemma 2. Then $G_1/\mathrm{Z}(G_1) \cong S_4$ and $G_2/\mathrm{Z}(G_2) \cong S_5$. Let $B_i$ be the principal 2-block of $G_i$, and let $\overline{B_i}$ be the principal 2-block of $G_i/\mathrm{Z}(G_i)$ for $i = 1, 2$. Then the Cartan matrix of $B_i$ is just the Cartan matrix of $\overline{B_i}$ multiplied by $|\mathrm{Z}(G_i)| = 2^{r-1}$. It is known that the Cartan matrices of $\overline{B_1}$ and $\overline{B_2}$ do not coincide (regardless of the labeling of the simple modules). Therefore, $B_1$ and $B_2$ are not Morita equivalent.

Nevertheless, the structure of a finite group $G$ with a minimal nonabelian Sylow 2-subgroup $P$ as given in (1) is fairly restricted. More precisely, Glauberman's $\mathrm{Z}^*$-theorem implies $x^2 \in \mathrm{Z}^*(G)$, and the structure of $G/\mathrm{Z}^*(G)$ follows from the Gorenstein–Walter theorem [1965]. In particular, $G$ has at most one nonabelian composition factor by Feit–Thompson.

We use the opportunity to present a related result for $p = 3$ which extends [Sambale 2014, Theorem 8.15].

**Theorem 11.** *Let $B$ and $\tilde{B}$ be non-nilpotent blocks of (possibly different) finite groups both with defect group $C_9 \rtimes C_3$. Then $B$ and $\tilde{B}$ are isotypic.*

*Proof.* As in the proof of Theorem 9, we will make use of [Horimoto and Watanabe 2012, Theorem 2]. Let

$$D := \langle x, y \mid x^9 = y^3 = 1,\ yxy^{-1} = x^4\rangle$$

be a defect group of $B$, and let $\mathcal{F}$ be the fusion system of $B$. By [Stancu 2006], $B$ is controlled with inertial index 2, and we may assume that $x$ and $x^{-1}$ are $\mathcal{F}$-conjugate (see the proof of [Sambale 2014, Theorem 8.8]). Then $\mathcal{R} := \{1, x, x^3, y, y^2, xy, xy^2\}$ is a set of representatives for the $\mathcal{F}$-conjugacy classes of $D$ (see the proof of [Sambale 2014, Theorem 8.15]). It suffices to show that the generalized decomposition numbers of $B$ are essentially unique (up to basic sets and signs and permutations of rows). Since the Galois group of $\mathbb{Q}(e^{2\pi i/9})$ over $\mathbb{Q}$ acts on the columns of the generalized decomposition matrix, we only need to determine the numbers $d^u_{\chi\varphi}$ for $u \in \{x, x^3, y, xy\}$. By [Sambale 2014, Theorem 8.15] there are four 3-rational characters $\chi_i \in \mathrm{Irr}(B)$ $(i = 1, \dots, 4)$ such that $\chi_1$, $\chi_2$ and $\chi_3$ have height 0 and $\chi_4$ has height 1. Since $\mathfrak{foc}(B) = \langle x\rangle$, we see that

$$\mathrm{Irr}(B) = \{\chi_i * \lambda : i = 1, 2, 3,\ \lambda \in \mathrm{Irr}(D/\mathfrak{foc}(B))\} \cup \{\chi_4\}.$$

Let $u := x^3$. Then $\mathrm{IBr}(b_u) = \{\varphi\}$ and $d^u_{\chi_i\varphi}$ are nonzero (rational) integers. Moreover, $d^u_{\chi_4\varphi} \equiv 0 \pmod 3$. After permuting $\chi_1$, $\chi_2$ and $\chi_3$ and changing the basic set for $b_u$ if necessary, we may assume that $d^u_{\chi_1\varphi} = 2$, $d^u_{\chi_2\varphi} =: \epsilon_1 \in \{\pm 1\}$, $d^u_{\chi_3\varphi} =: \epsilon_2 \in \{\pm 1\}$ and $d^u_{\chi_4\varphi} = 3\epsilon_3 \in \{\pm 3\}$. Now let $u := x$. Then $d^u_{\chi_i\varphi} = \pm 1$ for $i = 1, 2, 3$ and $d^u_{\chi_4\varphi} = 0$. We may choose a basic set for $b_u$ such that $d^u_{\chi_1\varphi} = 1$. Then by the orthogonality relations, $d^u_{\chi_2\varphi} = -\epsilon_1$ and $d^u_{\chi_3\varphi} = -\epsilon_2$. Next let $u := y$. Then $b_u$ dominates a block of $\mathrm{C}_G(u)/\langle u\rangle$ with cyclic defect group $\mathrm{C}_D(u)/\langle u\rangle \cong C_3$ and inertial index 2. This yields $\mathrm{IBr}(b_u) = \{\varphi_1, \varphi_2\}$ and the Cartan matrix of $b_u$ is given by

$$3\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

(not only up to basic sets, but this is not important here). We can choose a basic set such that $(d^u_{\chi_1\varphi_1}, d^u_{\chi_1\varphi_2}) = (1, 1)$, $(d^u_{\chi_2\varphi_1}, d^u_{\chi_2\varphi_2}) = (\sigma_1, 0)$, $(d^u_{\chi_3\varphi_1}, d^u_{\chi_3\varphi_2}) = (0, \sigma_2)$ and $(d^u_{\chi_4\varphi_1}, d^u_{\chi_4\varphi_2}) = (0, 0)$ for some signs $\sigma_1, \sigma_2 \in \{\pm 1\}$. Finally for $u := xy$ we obtain $d^u_{\chi_1\varphi} = 1$, $d^u_{\chi_i\varphi} = -\sigma_{i-1}$ for $i = 2, 3$ and $d^u_{\chi_4\varphi} = 0$ after changing the basic set if necessary. The following table summarizes the results:

| $u$ | $x^3$ | $x$ | $y$ | $xy$ |
|---|---|---|---|---|
| $d^u_{\chi_1\varphi}$ | $2$ | $1$ | $(1, 1)$ | $1$ |
| $d^u_{\chi_2\varphi}$ | $\epsilon_1$ | $-\epsilon_1$ | $(\sigma_1, 0)$ | $-\sigma_1$ |
| $d^u_{\chi_3\varphi}$ | $\epsilon_2$ | $-\epsilon_2$ | $(0, \sigma_2)$ | $-\sigma_2$ |
| $d^u_{\chi_4\varphi}$ | $3\epsilon_3$ | $0$ | $(0, 0)$ | $0$ |

It suffices to show that $\epsilon_i = \sigma_i$ for $i = 1, 2$ (observe that we do not need the ordinary decomposition numbers in order to apply [Horimoto and Watanabe 2012, Theorem 2]). For this, let $\lambda \in \mathrm{Irr}(D/\langle x^3\rangle)$ such that $\lambda(x) = e^{2\pi i/3}$ and $\lambda(y) = 1$. Then the generalized character $\psi := \lambda + \overline{\lambda} - 2 \cdot 1_D$ of $D$ is constant on $\langle x\rangle \setminus \langle x^3\rangle$ and thus $\mathcal{F}$-stable. By [Broué and Puig 1980a], $\chi_1 * \psi$ is a generalized character of $B$ and $(\chi_1 * \psi, \chi_2)_G \in \mathbb{Z}$. As in the proof of Theorem 9, we compute

$$\begin{aligned}(\chi_1 * \psi, \chi_2)_G &= \sum_{u\in\mathcal{R}} \psi(u) m^u_{\chi_1\chi_2} = \psi(x) m^x_{\chi_1\chi_2} + \psi(xy) m^{xy}_{\chi_1\chi_2} + \psi(xy^2) m^{xy^2}_{\chi_1\chi_2} \\ &= \tfrac{1}{3}\epsilon_1 + \tfrac{2}{3}\sigma_1.\end{aligned}$$

This shows $\epsilon_1 = \sigma_1$. Similarly, one gets $\epsilon_2 = \sigma_2$ by computing $(\chi_1 * \psi, \chi_3)_G$. Hence, [Horimoto and Watanabe 2012, Theorem 2] gives a perfect isometry $I : \mathrm{CF}(G, B) \to \mathrm{CF}(\tilde{G}, \tilde{B})$. In order to show that $I$ is also an isotypy, we make use of the notation introduced in the proof of Theorem 9. Let $u \in D$ such that $b_u$ is nilpotent. Then by the table above, we have $\mathrm{IBr}(b_u) = \{\pm d^u_G(\chi_2)\}$. Thus, one can extend $I^u$ just as in Theorem 9. Now suppose that $b_u$ is non-nilpotent and thus $u = y$ (up to inversion). We choose a basic set $\varphi_1, \varphi_2$ for $b_u$ as above such that $d^u_G(\chi_i) = \varphi_{i-1}$ for $i = 2, 3$. Now we have to determine the ordinary decomposition numbers of $b_u$ with respect to $\varphi_1, \varphi_2$. The defect group of $b_u$ is $\mathrm{C}_D(y) = \langle x^3, y\rangle \cong C_3 \times C_3$ and $\mathfrak{foc}(b_u) = \langle x^3\rangle$. By [Kiyota 1984], $k(b_u) = 9$. Therefore, there are 3-rational characters $\psi_i \in \mathrm{Irr}(b_u)$ such that

$$\mathrm{Irr}(b_u) = \{\psi_i * \lambda : i = 1, 2, 3,\ \lambda \in \mathrm{Irr}(\langle x^3, y\rangle/\langle x^3\rangle)\}.$$

By the Cartan matrix of $b_u$ given above (with respect to $\varphi_1, \varphi_2$), it follows immediately that $d^1_{\mathrm{C}_G(u)}(\psi_i) = \epsilon_i\varphi_i$ with $\epsilon_i \in \{\pm 1\}$ for $i = 1, 2$ after a suitable permutation of

$\psi_1, \psi_2, \psi_3$. Similarly, $d^1_{C_{\tilde{G}}(u)}(\tilde{\psi}_i) = \tilde{\epsilon}_i \tilde{\varphi}_i$. By a result of Usami [1988], there is a perfect isometry $\mathrm{CF}(C_G(u), b_u) \to \mathrm{CF}(C_{\tilde{G}}(u), \tilde{b}_u)$. However, we need the additional information that $\psi_i$ is mapped to $\pm\tilde{\psi}_i$. In order to show this, we use [Horimoto and Watanabe 2012, Theorem 2] again. Observe that $d^u_{C_G(u)}(\psi_i) = \zeta_i d^1_{C_G(u)}(\psi_i) = \zeta_i \epsilon_i \varphi_i$ for a cube root of unity $\zeta_i$. But since $d^u_{\psi_i \varphi_i}$ is rational, we have $\zeta_i = 1$. Now an elementary application of the orthogonality relations shows that the generalized decomposition matrix of $b_u$ (in $C_G(u)$) is determined by

| $v$ | $1$ | $y$ | $x^3$ | $x^3 y$ |
|---|---|---|---|---|
| $d^v_{\psi_1 \varphi}$ | $(\epsilon_1, 0)$ | $(\epsilon_1, 0)$ | $\epsilon_1$ | $\epsilon_1$ |
| $d^v_{\psi_2 \varphi}$ | $(0, \epsilon_2)$ | $(0, \epsilon_2)$ | $\epsilon_2$ | $\epsilon_2$ |
| $d^v_{\psi_3 \varphi}$ | $(\epsilon_3, \epsilon_3)$ | $(\epsilon_3, \epsilon_3)$ | $-\epsilon_3$ | $-\epsilon_3$ |

It follows that there is a perfect isometry $\widehat{I^u} : \mathrm{CF}(C_G(u), b_u) \to \mathrm{CF}(C_{\tilde{G}}(u), \tilde{b}_u)$ such that $\widehat{I^u}(\psi_i) = \epsilon_i \tilde{\epsilon}_i \tilde{\psi}_i$ for $i = 1, 2$. Therefore $\widehat{I^u}$ extends $I^u$. As in the proof of Theorem 9, it is also clear that $\widehat{I^u}$ is independent of the choice of the generator of $\langle u \rangle$. This finishes the proof. $\square$

The proof method of Theorem 11 also works for other defect groups. In fact, Watanabe [2015] showed independently (using more complicated methods) that two $p$-blocks ($p > 2$) with a common metacyclic, minimal nonabelian defect group and the same fusion system are perfectly isometric. Again, this gives evidence for the character-theoretic version of Rouquier's conjecture (see [Watanabe 2014, Theorem 2]). As another remark, Holloway, Koshitani and Kunugi [2010, Example 4.3] constructed a perfect isometry between the principal 3-block of $G := \mathrm{Aut}(\mathrm{SL}(2, 8)) \cong {}^2G_2(3)$ and its Brauer correspondent. Since $G$ has a Sylow 3-subgroup isomorphic to $C_9 \rtimes C_3$, this is a special case of Theorem 11. Note that in the introduction of [Ruengrot 2011] it is erroneously stated that these blocks are *not* perfectly isometric.

## Acknowledgment

This work is supported by the German Research Foundation and the Daimler and Benz Foundation. The author thanks Atumi Watanabe for providing a copy of [Watanabe 2015]. Moreover, the author thanks Burkhard Külshammer for answering some questions.

## References

[Broué and Puig 1980a] M. Broué and L. Puig, "Characters and local structure in $G$-algebras", *J. Algebra* **63**:2 (1980), 306–317. MR 81j:20021 Zbl 0428.20005

[Broué and Puig 1980b] M. Broué and L. Puig, "A Frobenius theorem for blocks", *Invent. Math.* **56**:2 (1980), 117–128. MR 81d:20011 Zbl 0425.20008

[Cabanes and Picaronny 1992] M. Cabanes and C. Picaronny, "Types of blocks with dihedral or quaternion defect groups", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **39**:1 (1992), 141–161. MR 93f:20014 Zbl 0781.20006

[Cliff 2000] G. Cliff, "On centers of 2-blocks of Suzuki groups", *J. Algebra* **226**:1 (2000), 74–90. MR 2001d:20013 Zbl 0953.20003

[Eaton et al. 2012] C. W. Eaton, B. Külshammer, and B. Sambale, "2-blocks with minimal nonabelian defect groups, II", *J. Group Theory* **15**:3 (2012), 311–321. MR 2920888 Zbl 1253.20008

[Erdmann 1990] K. Erdmann, *Blocks of tame representation type and related algebras*, Lecture Notes in Mathematics **1428**, Springer, Berlin, 1990. MR 91c:20016 Zbl 0696.20001

[Gorenstein and Walter 1965] D. Gorenstein and J. H. Walter, "The characterization of finite groups with dihedral Sylow 2-subgroups, I", *J. Algebra* **2** (1965), 85–151. MR 31 #1297a Zbl 0192.11902

[Holloway et al. 2010] M. Holloway, S. Koshitani, and N. Kunugi, "Blocks with nonabelian defect groups which have cyclic subgroups of index $p$", *Arch. Math.* (*Basel*) **94**:2 (2010), 101–116. MR 2011c:20013 Zbl 1195.20010

[Holm 2001] T. Holm, "Notes on Donovan's Conjecture for blocks of tame representation type", 2001, http://www2.iazd.uni-hannover.de/~tholm/ARTIKEL/donovan.ps. Unpublished notes.

[Horimoto and Watanabe 2012] H. Horimoto and A. Watanabe, "On a perfect isometry between principal p-blocks of finite groups with cyclic p-hyperfocal subgroups", preprint, 2012, http://hdl.handle.net/2433/194497. In Japanese.

[Kessar et al. 2015] R. Kessar, M. Linckelmann, and G. Navarro, "A characterisation of nilpotent blocks", *Proc. Amer. Math. Soc.* (online publication June 2015).

[Kiyota 1984] M. Kiyota, "On 3-blocks with an elementary abelian defect group of order 9", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **31**:1 (1984), 33–58. MR 85k:20036 Zbl 0546.20013

[Linckelmann 2007] M. Linckelmann, "Introduction to fusion systems", pp. 79–113 in *Group representation theory*, edited by M. Geck et al., EPFL Press, Lausanne, 2007. MR 2008f:20021 Zbl 1161.20007

[Malle and Navarro 2011] G. Malle and G. Navarro, "Blocks with equal height zero degrees", *Trans. Amer. Math. Soc.* **363**:12 (2011), 6647–6669. MR 2012g:20016 Zbl 1277.20013

[Robinson 1991] G. R. Robinson, "On the number of characters in a block", *J. Algebra* **138**:2 (1991), 515–521. MR 92h:20022a Zbl 0727.20010

[Robinson 2008] G. R. Robinson, "On the focal defect group of a block, characters of height zero, and lower defect group multiplicities", *J. Algebra* **320**:6 (2008), 2624–2628. MR 2009f:20010 Zbl 1153.20006

[Rouquier 2001] R. Rouquier, "Block theory via stable and Rickard equivalences", pp. 101–146 in *Modular representation theory of finite groups* (Charlottesville, VA, 1998), edited by M. J. Collins et al., de Gruyter, Berlin, 2001. MR 2003g:20018 Zbl 0998.20006

[Ruengrot 2011] P. Ruengrot, *Perfect isometry groups for blocks of finite groups*, Ph.D. thesis, University of Manchester, 2011, https://www.escholar.manchester.ac.uk/uk-ac-man-scw:142297.

[Sambale 2011] B. Sambale, "2-blocks with minimal nonabelian defect groups", *J. Algebra* **337** (2011), 261–284. MR 2012d:20019 Zbl 1247.20010

[Sambale 2014] B. Sambale, *Blocks of finite groups and their invariants*, Lecture Notes in Mathematics **2127**, Springer, Berlin, 2014. MR 3289382 Zbl 1315.20009

[Stancu 2006] R. Stancu, "Control of fusion in fusion systems", *J. Algebra Appl.* **5**:6 (2006), 817–837. MR 2007j:20025 Zbl 1118.20020

[Usami 1988] Y. Usami, "On $p$-blocks with abelian defect groups and inertial index 2 or 3, I", *J. Algebra* **119**:1 (1988), 123–146. MR 89i:20024 Zbl 0659.20008

[Watanabe 1985] A. Watanabe, "On generalized decomposition numbers and Fong's reductions", *Osaka J. Math.* **22**:2 (1985), 393–400. MR 86i:20018 Zbl 0575.20011

[Watanabe 2014] A. Watanabe, "The number of irreducible Brauer characters in a $p$-block of a finite group with cyclic hyperfocal subgroup", *J. Algebra* **416** (2014), 167–183. MR 3232798 Zbl 06339539

[Watanabe 2015] A. Watanabe, "On blocks of finite groups with metacyclic, minimal non-abelian defect groups", preprint, 2015.

Received April 26, 2015. Revised June 6, 2015.

BENJAMIN SAMBALE
FACHBEREICH MATHEMATIK
TU KAISERSLAUTERN
67653 KAISERSLAUTERN
GERMANY

sambale@mathematik.uni-kl.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 280, No. 2, 2016
dx.doi.org/10.2140/pjm.2016.280.489

# NUMBER OF SINGULARITIES OF STABLE MAPS ON SURFACES

TAKAHIRO YAMAMOTO

**Let $N$ denote the plane $\mathbb{R}^2$ or the 2-sphere $S^2$. In this paper, we determine the 5-tuples of integers $(g, d, i, c, n)$ such that there exists a degree $d$ stable map $\Sigma_g \to N$ whose singular point set consists of $i$ connected components, $c$ cusps, and $n$ nodes, where $\Sigma_g$ is the standard genus $g$ surface.**

## 1. Introduction

Throughout this paper, all surfaces and manifolds are connected and of class $C^\infty$ (i.e., smooth), and all maps are of class $C^\infty$. Let $M$ be a closed surface and $N$ be a surface. For a $C^\infty$ map $\varphi : M \to N$, denote by $S(\varphi)$ the set of singular points of $\varphi$. Call $\varphi(S(\varphi))$ the *apparent contour* (*contour* for short), and denote it by $\gamma(\varphi)$. In this paper, all $C^\infty$ maps $M \to N$ have nonempty singular point sets.

A $C^\infty$ map $\varphi : M \to N$ is said to be *stable* if it satisfies the following two properties.

(1) For each $p \in M$, the map germ of $\varphi$ at $p$ is $C^\infty$ right-left equivalent to one of the map germs at $0 \in \mathbb{R}^2$ as follows:

$$(a, x) \mapsto \begin{cases} (a, x), & p \text{ is a regular point,} \\ (a, x^2), & p \text{ is a fold point,} \\ (a, x^3 + ax), & p \text{ is a cusp point.} \end{cases}$$

Hence, $S(\varphi)$ is a disjoint union of finitely many circles.

(2) For each $q \in \gamma(\varphi)$, the map germ $(\varphi|_{S(\varphi)}, \varphi^{-1}(q) \cap S(\varphi))$ is right-left equivalent to one of the three multigerms as depicted in Figure 1.

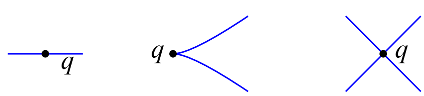


**Figure 1.** The multigerms of $\varphi|_{S(\varphi)}$.

*MSC2010:* primary 57R45; secondary 57M20, 58K15.
*Keywords:* stable map, cusp, node.

According to a classical result of Whitney [1955], stable maps form an open dense subset of the space of all $C^\infty$ maps $M \to N$ with respect to the Whitney $C^\infty$ topology.

For a stable map $\varphi : M \to N$, denote by $c(\varphi)$, $n(\varphi)$ and $i(\varphi)$ the numbers of cusps, nodes and connected components of $S(\varphi)$, respectively.

For a nonnegative integer $g$, the closed and oriented surface of genus $g$, which is the connected sum of $g$ copies of the 2-dimensional torus $T^2$, is denoted by $\Sigma_g$. The 2-dimensional sphere and the plane are denoted by $S^2$ and $\mathbb{R}^2$, respectively.

For any stable map $f : M \to S^2$ (or $\mathbb{R}^2$) of a closed and oriented surface $M$, one can associate the 5-tuple of integers $(g, d(f), i(f), c(f), n(f))$, where $g$ is the genus of $M$ and $d(f)$ is the mapping degree of $f$. This paper studies the following question: which 5-tuples $(g, d, i, c, n)$ can occur in this way?

Some necessary conditions have been obtained by Pignoni [1993], Kamenosono and Yamamoto [2009] (see also Proposition 3.4), Eliashberg [1970], and Quine [1978] (see also Theorem 3.11 of the present paper). M. Yamamoto [2009] studied the numbers $i(f)$ of fold maps $f : \Sigma_g \to \Sigma_h$.

András Szűcs posed the following question at the International Workshop on Real and Complex Singularities, held in São Carlos in 2012: whether these conditions form a complete set of restrictions.

The answer is No. There is a geometrical condition for the number of nodes. More precisely, there is the minimal number of nodes for a given 4-tuple $(g, d, i, c)$. The main results of this paper are the following two theorems.

Let $\mathrm{v}_1 = (2, 1)$ and $\mathrm{v}_2 = (0, 2)$ be vectors in $\mathbb{R}^2$. For given integers $k, \ell \geq 0$, denote by $L_{k,\ell}$ the affine lattice $\{\mu_1 \mathrm{v}_1 + \mu_2 \mathrm{v}_2 + (2, 0) \mid \mu_1, \mu_2 \in \mathbb{Z}\}$ if $k \equiv \ell \bmod 2$, and the lattice $\{\mu_1 \mathrm{v}_1 + \mu_2 \mathrm{v}_2 \mid \mu_1, \mu_2 \in \mathbb{Z}\}$ otherwise. For given integers $k, \ell \geq 0$, set $\delta_{k,\ell} = 2$ if $k \equiv \ell \bmod 2$, and $\delta_{k,\ell} = 0$ otherwise.

**Theorem 1.1.** *Let $g \geq 0$ and $i \geq 1$. If $f : \Sigma_g \to \mathbb{R}^2$ is a stable map whose singular point set consists of $i$ components, then the pair $(c(f), n(f))$ is in $L_{i,g} \cap D$, where $D$ denotes the subset of $\mathbb{R}^2$ (expressed by coordinates $(x, y)$) defined by the following*:

$$D = \begin{cases} \left\{(x, y) \;\middle|\; x \geq \delta_{i,g},\ y \geq 0,\ y \geq -\frac{1}{2}x + g - i + 3,\ y \geq \frac{1}{2}x - g - i + 1\right\} & \\ & \text{if } 1 \leq i \leq g \\ \left\{(x, y) \;\middle|\; x \geq \delta_{i,g},\ y \geq 0,\ y \geq \frac{1}{2}x - g - i + 1\right\} & \text{if } i > g. \end{cases}$$

*Furthermore, for any pair $(c, n)$ in $L_{i,g} \cap D$, there is a stable map $f : \Sigma_g \to \mathbb{R}^2$ with $S(f)$ consisting of $i$ components, $c$ cusps, and $n$ nodes.*

**Theorem 1.2.** *Let $g, d \geq 0$ and $i \geq 1$. If $f : \Sigma_g \to S^2$ is a degree $d$ stable map whose singular point set consists of $i$ components, then the pair $(c(f), n(f))$ is in $L_{i,g+d} \cap D$, where $D$ denotes the subset of $\mathbb{R}^2$ (expressed by coordinates $(x, y)$) defined by the following*:

$g=0$:

$$D=\begin{cases}\{(x,y)\mid x\geq 2(d+1-i),\ y\geq 0\} & \textit{if } 1\leq i\leq d,\\ \{(x,y)\mid x\geq \delta_{i,d},\ y\geq 0\} & \textit{if } i\geq d.\end{cases}$$

$g\geq 1$:

$$D=\begin{cases}(1)\quad \{(x,y)\mid x\geq \delta_{i,g},\ y\geq 0,\ y\geq -\frac{1}{2}x+g+3-i\} \\ \qquad\qquad\qquad\qquad \textit{if } d=0 \textit{ and } 1\leq i\leq g,\\ (2)\quad \{(x,y)\mid x\geq \delta_{i,g},\ y\geq 0\} \quad \textit{if } d=0 \textit{ and } i>g,\\ (3)\quad \{(x,y)\mid x\geq 2(d+1-g-i),\ y\geq 0,\ y\geq -\frac{1}{2}x+d+g+3-i\} \\ \qquad\quad \cup\{(x,y)\mid x\geq 2(d+g+1-i),\ y\geq 0\} \\ \qquad\qquad\qquad\qquad \textit{if } d\geq 1 \textit{ and } 1\leq i\leq d-g+1,\\ (4)\quad \{(x,y)\mid x\geq \delta_{i,g+d},\ y\geq 0,\ y\geq -\frac{1}{2}x+d+g+3-i\} \\ \qquad\quad \cup\{(x,y)\mid x\geq 2(d+g+1-i),\ y\geq 0\} \\ \qquad\qquad\qquad\qquad \textit{if } d\geq 1 \textit{ and } d-g+1\leq i\leq d+g-1,\\ (5)\quad \{(x,y)\mid x\geq \delta_{i,g+d},\ y\geq 0\} \quad \textit{if } d\geq 1 \textit{ and } i\geq d+g.\end{cases}$$

*Furthermore, for any pair* $(c,n)$ *in* $L_{i,g+d}\cap D$, *there is a degree* $d$ *stable map* $f:\Sigma_g\to S^2$ *with* $S(f)$ *consisting of* $i$ *components*, $c$ *cusps and* $n$ *nodes.*

In Theorems 1.1 and 1.2, generators $v_1$ and $v_2$ correspond to two modifications for stable maps between surfaces: $v_1$ is passing through the swallow-tail singularity (Figure 2), while $v_2$ is passing through the tangency singularity (Figure 3).

In order to prove Theorems 1.1 and 1.2, we will construct maps for any 5-tuples in the list. The constructions go as follows. There are ten modifications that can apply to any map in order to obtain a map with a new 5-tuple.

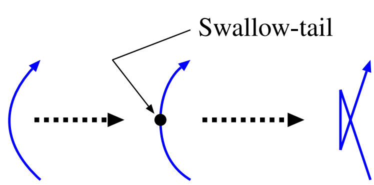


**Figure 2.** Swallow-tail singularity.

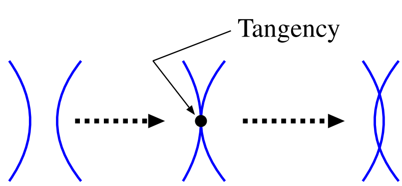


**Figure 3.** Tangency singularity.

(1) Passing through the swallow-tail singularity (Figure 2):

$$(g, d, i, c, n) \to (g, d, i, c+2, n+1),$$

(2) Passing through the tangency singularity of the singular curve (Figure 3):

$$(g, d, i, c, n) \to (g, d, i, c, n+2),$$

(3) Attaching two spheres (Figure 4):

$$(g, d, i, c, n) \to (g, d, i, c+4, n),$$

(4) Attaching a handle vertically (Figure 5, left):

$$(g, d, i, c, n) \to (g+1, d, i+1, c, n),$$

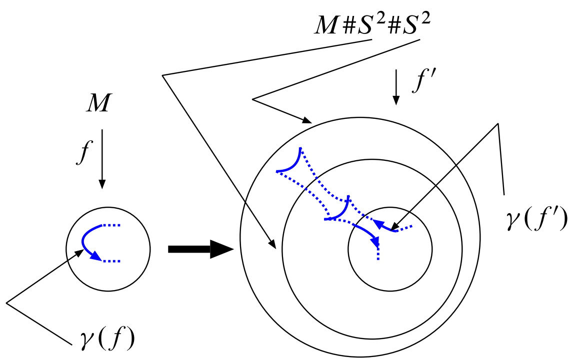


**Figure 4.** Attaching two spheres: by attaching two maps $\mathrm{id}_{S^2}$ and $-\mathrm{id}_{S^2}$ to $f : M \to S^2$, we obtain a stable map $M \to S^2$, where $\mathrm{id}_{S^2}$ denotes the identity map on $S^2$ and $-\mathrm{id}_{S^2}$ the $C^\infty$ map of $S^2$ into $S^2$ defined by $x \mapsto -x$.

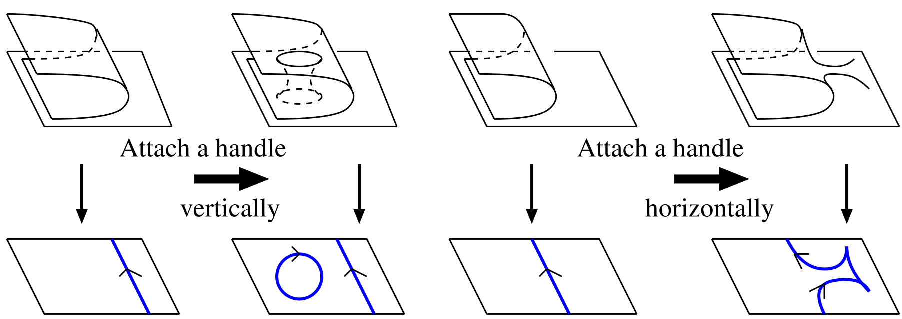


**Figure 5.** Left: attaching a handle vertically. Right: attaching a handle horizontally. The map is obtained when one projects these surfaces to the horizontal plane.

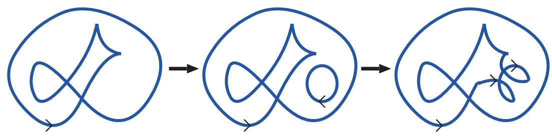

**Figure 6.** Attaching a pair of handles: attaching a handle vertically and then attaching a handle horizontally.

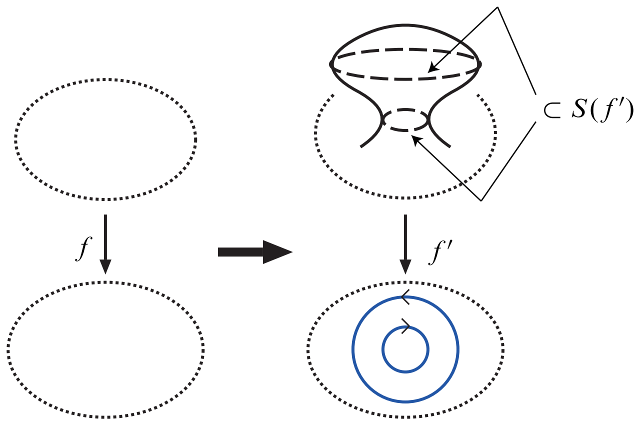


**Figure 7.** Attaching a balloon.

(5) Attaching a handle horizontally (Figure 5, right):

$$(g, d, i, c, n) \to (g+1, d, i, c+2, n),$$

(6) Attaching a pair of handles. More precisely, attaching a handle vertically and then attaching a handle horizontally (Figure 6):

$$(g, d, i, c, n) \to (g+2, d, i, c, n+2),$$

(7) Attaching a balloon (Figure 7):

$$(g, d, i, c, n) \to (g, d, i+2, c, n),$$

(8) Making a wrinkle (Figure 8):

$$(g, d, i, c, n) \to (g, d, i+1, c+2, n),$$

(9) Attaching a sphere horizontally (Figure 9):

$$(g, d, i, c, n) \to (g, d+1, i, c+2, n),$$

(10) Attaching a sphere vertically (Figure 10):

$$(g, d, i, c, n) \to (g, d+1, i+1, c, n).$$

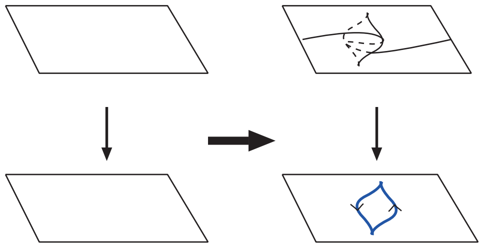

**Figure 8.** Making a wrinkle.

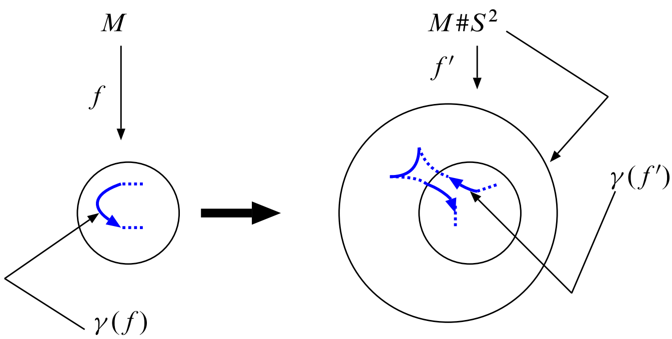


**Figure 9.** Attaching a sphere horizontally.

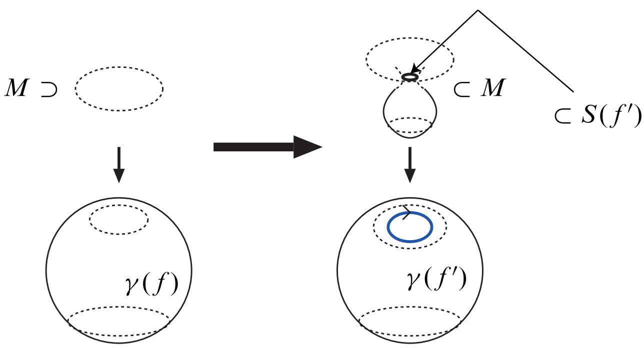


**Figure 10.** Attaching a sphere vertically.

(11) Attaching a pair of a sphere and a handle. More precisely, attaching a sphere vertically and then, attaching a handle horizontally (Figure 11):

$$(g, d, i, c, n) \to (g+1, d+1, i, c, n+2),$$

We remark that we can apply all modifications $(i)$ with $1 \leq i \leq 11$ to stable maps of surfaces into the sphere. But we can apply modifications (1) and (2), (4), (5), (6), (7), (8) to stable maps of surfaces into the plane.

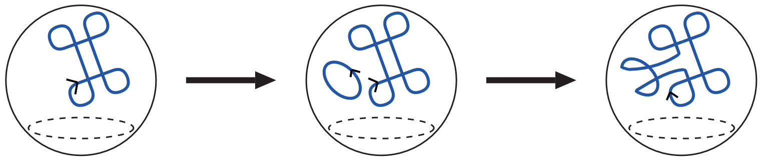

**Figure 11.** Attaching a pair of a sphere and a handle: attaching a sphere vertically and then, attaching a handle horizontally.

These modifications never decrease the number of 5-tuple $(g, d, i, c, n)$, but they increase some of them. Hence it is enough to construct maps providing the minimal 5-tuples. These constructions can be found in [Demoto 2005; Fukuda and Yamamoto 2011; Kamenosono and Yamamoto 2009; Yamamoto 2009; Yamamoto 2010]. We will sketch their descriptions in Section 2.

**Remark 1.3.** Theorems 1.1 and 1.2 together with the previous results [Fukuda and Yamamoto 2011; Kamenosono and Yamamoto 2009; Pignoni 1993; Yamamoto 2009; 2010] make the very first step toward classifying generic $C^\infty$ maps of closed surfaces into the plane or the sphere up to right-left equivalence.

**Remark 1.4.** Let $M$ be a closed surface and $N$ a surface. Let $\mathbb{A}$ be an element, an ordered pair, or triple consisting of some elements in $\{c, i, n, c+n\}$. For a stable map $\varphi : M \to N$, denote by $\mathbb{A}(\varphi)$ the element, the ordered pair, or triple consisting of the corresponding elements in $\{c(\varphi), i(\varphi), n(\varphi), c(\varphi)+n(\varphi)\}$. For a $C^\infty$ map $\varphi_0 : M \to N$, we say that a stable map $\varphi : M \to N$ has an $\mathbb{A}$-*minimal contour for* $\varphi_0$ if $\mathbb{A}(\varphi)$ is minimal with respect to the lexicographic order among those stable maps which are homotopic to $\varphi_0$.

Let $\mathbb{A} = (i, c, n)$. The $(i, c, n)$-minimal contours were studied in [Demoto 2005; Kamenosono and Yamamoto 2009; Pignoni 1993]. The $(i, c, n)$-minimal contours of a $C^\infty$ map $\Sigma_g \to S^2$ of degree $d$ correspond to the bottom left corner of the lattice $L_{1,g+d} \cap D$. Note that for a $C^\infty$ map $M \to N$, there is a stable map with $S(f)$ consisting of one component.

This paper is organized as follows. In Section 2, we prepare some stable maps $M \to \mathbb{R}^2$ and $M \to S^2$ which we employ in the following section. In Section 3, we prove Theorems 1.1 and 1.2. In Section 4, we pose two problems which concern the apparent contours of stable maps between surfaces. In the Appendix, we study $i$-$(c, n)$-minimal contours and $i$-$(n, c)$-minimal contours of stable maps $\Sigma_g \to S^2$.

## 2. Stable maps $\Sigma_g \to N$ (for $N = \mathbb{R}^2$ or $N = S^2$)

In this section, we show that there exist stable maps $\Sigma_g \to \mathbb{R}^2$ and $\Sigma_g \to S^2$ whose triples $(i, c, n)$ are in the lists of Theorems 1.1 and 1.2, respectively. For two integers $k$ and $\ell$, set $\delta_{k,\ell} = 2$ if $k \equiv \ell$, and $\delta_{k,\ell} = 0$ otherwise.

**2A.** ***Stable maps $\Sigma_g \to \mathbb{R}^2$.*** Let $p_{S^2} : S^2 \to \mathbb{R}^2$ be the standard projection defined by $(x, y, z) \mapsto (x, y)$. Then, the contour $\gamma(p_{S^2})$ is an embedded circle in $\mathbb{R}^2$, namely, the triple $(i, c, n)$ is equal to $(1, 0, 0)$. Then, by applying modifications (1) and (2) to $p_{S^2}$ inductively, for each $(c, n)$ in

$$L_{1,0} \cap \left\{(x, y) \;\middle|\; x \geq 0,\ y \geq \tfrac{1}{2}x\right\},$$

we obtain a stable map $f : S^2 \to \mathbb{R}^2$ with $S(f)$ consisting of one component, $c$ cusps, and $n$ nodes. Furthermore, for a given integer $i \geq 1$, by applying modifications (7) and (8) $i - 1$ times to stable maps $S^2 \to \mathbb{R}^2$ whose pairs $(c, n)$ are in

$$L_{1,0} \cap \left\{(x, y) \;\middle|\; x \geq 0,\ y \geq \tfrac{1}{2}x\right\},$$

for each $(c, n)$ in

$$L_{i,0} \cap \left\{(x, y) \;\middle|\; x \geq \delta_{i,0},\ y \geq 0,\ y \geq \tfrac{1}{2}x - i + 1\right\},$$

we obtain a stable map $f : S^2 \to \mathbb{R}^2$ with $S(f)$ consisting of $i$ components, $c$ cusps, and $n$ nodes.

Then, for given integers $g \geq 1$ and $i \geq 1$, with $i > g$, let $k$ and $\ell$ be nonnegative integers satisfying $k + \ell = g$. By applying modifications (4) $k$ times and (5) $\ell$ times to stable maps $S^2 \to \mathbb{R}^2$ whose pairs $(c, n)$ are in

$$L_{i,0} \cap \left\{(x, y) \;\middle|\; x \geq \delta_{i,0},\ y \geq 0,\ y \geq \tfrac{1}{2}x - i + 1\right\},$$

for each $(c, n)$ in

$$L_{i,g} \cap \left\{(x, y) \;\middle|\; x \geq \delta_{i,g},\ y \geq 0,\ y \geq \tfrac{1}{2}x - g - i + 1\right\},$$

we obtain a stable map $f : \Sigma_g \to \mathbb{R}^2$ with $S(f)$ consisting of $i$ components, $c$ cusps, and $n$ nodes.

Thus, we obtain stable maps $\Sigma_g \to \mathbb{R}^2$ whose pairs $(c, n)$ are in the list of Theorem 1.1 with $i > g$.

**Proposition 2.1.** *Let $g \geq 1$. For each pair $(c, n)$ in*

$$L_{1,g} \cap \left\{(x, y) \;\middle|\; y = -\tfrac{1}{2}x + g + 2,\ y \geq 1\right\},$$

*there is a stable map $\Sigma_g \to \mathbb{R}^2$ with $S(f)$ consisting of one component, $c$ cusps, and $n$ nodes.*

*Proof.* There exist stable maps $T^2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(1, 2, 2)$ and $(1, 4, 1)$. There also exist stable maps $\Sigma_2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(1, 0, 4)$ and $(1, 2, 3)$. See [Pignoni 1993; Yamamoto 2010] for the details.

By applying modifications (5) and (6) to the above four stable maps $T^2 \to \mathbb{R}^2$, $\Sigma_2 \to \mathbb{R}^2$ inductively, we obtain the desired stable maps $\Sigma_g \to \mathbb{R}^2$. For example, let us consider the case $g = 2$. By applying modification (5) to stable maps $T^2 \to \mathbb{R}^2$

whose triples $(i, c, n)$ are equal to $(1, 2, 2)$ and $(1, 4, 1)$, we obtain stable maps $\Sigma_2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(1, 4, 2)$ and $(1, 6, 1)$, respectively. Furthermore, let us consider the case $g = 3$. By applying modification (6) to stable maps $T^2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(1, 2, 2)$ and $(1, 4, 1)$, we obtain stable maps $\Sigma_3 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(1, 2, 4)$ and $(1, 4, 3)$, respectively. By applying modification (5) to stable maps $\Sigma_2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(1, 4, 2)$ and $(1, 6, 1)$, we obtain stable maps $\Sigma_3 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(1, 6, 2)$ and $(1, 8, 1)$, respectively. □

Then, by applying modifications (1) and (2) inductively to stable maps in Proposition 2.1, for each $(c, n)$ in

$$L_{1,g} \cap \left\{(c, n) \,\middle|\, x \geq \delta_{1,g},\ y \geq 1,\ y \geq -\tfrac{1}{2}x + g + 2,\ y \geq \tfrac{1}{2}x - g - 2\right\},$$

we obtain a stable map $\Sigma_g \to \mathbb{R}^2$ with $S(f)$ consisting of one component, $c$ cusps, and $n$ nodes.

For given integers $g \geq 1$ and $i \geq 1$ with $1 \leq i \leq g$, by applying modification (4) $i - 1$ times to stable maps $\Sigma_{g-i+1} \to \mathbb{R}^2$ whose pairs $(c, n)$ are in

$$L_{1,g-i+1} \cap \left\{(c, n) \,\middle|\, x \geq \delta_{1,g-i+1},\ y \geq 1,\ y \geq -\tfrac{1}{2}x + (g - i + 1) + 2,\ y \geq \tfrac{1}{2}x - (g - i + 1) - 2\right\},$$

for each $(c, n)$ in

$$L_{i,g} \cap \left\{(c, n) \,\middle|\, x \geq \delta_{i,g},\ y \geq 1,\ y \geq -\tfrac{1}{2}x + g + 2,\ y \geq \tfrac{1}{2}x - g - 2\right\},$$

we obtain a stable map $\Sigma_g \to \mathbb{R}^2$ with $S(f)$ consisting of $i$ components, $c$ cusps, and $n$ nodes.

Thus, we obtain all stable maps $\Sigma_g \to \mathbb{R}^2$ whose triples $(i, c, n)$ are in the list of Theorem 1.1.

**2B. *Stable maps $\Sigma_g \to S^2$.*** Note that stable maps $\Sigma_g \to \mathbb{R}^2$ obtained in Section 2A induce degree zero stable maps $\Sigma_g \to S^2$.

Let us consider stable maps $S^2 \to S^2$. Denote by $f^{S^2,S^2}_{(1,2,0)}$ a degree one stable map $S^2 \to S^2$ whose contour is shown in Figure 12.

Let $d \geq 1$ and $i \geq 1$ be integers with $i \leq d$. By applying modifications (1), (2), and (3) inductively to a degree $d$ stable map $S^2 \to S^2$ whose triple $(i, c, n)$ is equal to $(1, 2d, 0)$, for each $(c, n)$ in

$$L_{1,d} \cap \{(x, y) \mid x \geq 2d,\ y \geq 0\},$$

we obtain a degree $d$ stable map $f : S^2 \to S^2$ with $S(f)$ consisting of one component, $c$ cusps, and $n$ nodes. Then, by applying modification (10) $i - 1$ times inductively

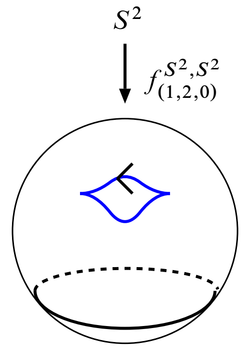


**Figure 12.** Stable map $f^{S^2,S^2}_{(1,2,0)} : S^2 \to S^2$.

to these degree $d-i+1$ stable maps $S^2 \to S^2$ whose pairs $(c, n)$ are in

$$L_{1,d-i+1} \cap \{(x, y) \mid x \geq 2(d+1-i),\ y \geq 0\},$$

for each $(c, n)$ in

$$L_{i,d} \cap \{(x, y) \mid x \geq 2(d+1-i),\ y \geq 0\},$$

we obtain a degree $d$ stable map $f : S^2 \to S^2$ with $S(f)$ consisting of $i$ components, $c$ cusps, and $n$ nodes.

Let $d \geq 0$ and $i \geq 1$ be integers with $i \geq d$. By applying modification (10) $d$ times and $d-1$ times inductively to degree zero stable maps $S^2 \to S^2$ and degree one stable maps $S^2 \to S^2$ whose pairs $(c, n)$ are in

$$L_{i,0} \cap \{(x, y) \mid x \geq \delta_{i,0},\ y \geq 0\} \quad \text{and} \quad L_{i,1} \cap \{(x, y) \mid x \geq \delta_{i,1},\ y \geq 0\}$$

respectively, for each $(c, n)$ in

$$L_{i,d} \cap \{(x, y) \mid x \geq \delta_{i,d},\ y \geq 0\},$$

we obtain a degree $d$ stable map $S^2 \to S^2$ with $S(f)$ consisting of $i$ components, $c$ cusps and $n$ nodes.

Thus, we obtain all stable maps $S^2 \to S^2$ whose pairs $(c, n)$ are in the lists of Theorem 1.2 with $g = 0$.

In the following, assume $g \geq 1$. Let us consider degree zero stable maps $\Sigma_g \to S^2$ which are not induced from stable maps $\Sigma_g \to \mathbb{R}^2$.

**Proposition 2.2.** *Let $g \geq 1$ and $i \geq 1$ with $i \leq g$. For each pair $(c, n)$ in*

$$L_{i,g} \cap \left\{(x, y) \,\middle|\, x \geq \delta_{i,g},\ y \geq 0,\ y = -\tfrac{1}{2}x + g + 3 - i\right\},$$

*there is a degree zero stable map $\Sigma_g \to S^2$ with $S(f)$ consisting of $i$ components, $c$ cusps, and $n$ nodes.*

*Proof.* For each $(c, n)$ in

$$L_{i,g} \cap \left\{(x, y) \,\middle|\, x \geq \delta_{i,g},\ y \geq 1,\ y = -\tfrac{1}{2}x + g + 3 - i\right\},$$

we already obtained a degree zero stable map $f : \Sigma_g \to S^2$ with $S(f)$ consisting of $i$ components, $c$ cusps and $n$ nodes in Section 2A.

By attaching a sphere which is mapped by orientation reversely to $f^{S^2,S^2}_{(1,2,0)}$, we obtain a degree zero stable map $S^2 \to S^2$ whose triple $(i, c, n)$ is equal to $(1, 4, 0)$. Then, for each integers $g \geq 1$ and $i \geq 1$ with $i \leq g$, by applying modifications (4) $i - 1$ times and (5) $g - i + 1$ times to this degree zero stable map $S^2 \to S^2$, we obtain a degree zero stable map $f : \Sigma_g \to S^2$ with $S(f)$ consisting of $i$ components, $2(g + 3 - i)$ cusps, and no nodes. □

Let $g \geq 1$ and $i \geq 1$ with $i \leq g$. By applying modifications (1), (2), and (3) to stable maps obtained in the above subsection and in Proposition 2.2, for each pair $(c, n)$ in

$$L_{i,g} \cap \left\{(x, y) \,\middle|\, x \geq \delta_{i,g},\ y \geq 0,\ y \geq -\tfrac{1}{2}x + g + 3 - i\right\},$$

we obtain a stable map $\Sigma_g \to S^2$ with $S(f)$ consisting of $i$ components, $c$ cusps, and $n$ nodes.

Let $g \geq 1$ and $i \geq 1$ with $i > g$. By applying modifications (1), (2), and (3) inductively to a degree zero stable map $\Sigma_g \to S^2$ whose triple $(i, c, n)$ is equal to $(i, \delta_{i,g}, 0)$, for each $(c, n)$ in

$$L_{i,g} \cap \{(x, y) \mid x \geq \delta_{i,g},\ y \geq 0\},$$

we obtain a degree zero stable map $f : \Sigma_g \to S^2$ with $S(f)$ consisting of $i$ components, $c$ cusps and $n$ nodes.

Thus, we obtain degree zero stable maps $\Sigma_g \to S^2$ in the lists of Theorem 1.2 (1) and (2).

In the following, assume $g \geq 1$ and $d \geq 1$.

**Proposition 2.3.** (1) *If $g \leq d$, then for each $(c, n)$ in*

$$L_{1,g+d} \cap \left\{(x, y) \,\middle|\, x \geq 2(d - g),\ y = -\tfrac{1}{2}x + g + d + 2\right\},$$

*there is a degree $d$ stable map $\Sigma_g \to S^2$ with $S(f)$ consisting one component, $c$ cusps, and $n$ nodes.*

(2) *If $d \leq g$, then for each $(c, n)$ in*

$$L_{1,g+d} \cap \left\{(x, y) \,\middle|\, y = -\tfrac{1}{2}x + g + d + 2,\ y \geq 3\right\},$$

*there is a degree $d$ stable map $\Sigma_g \to S^2$ with $S(f)$ consisting one component, $c$ cusps, and $n$ nodes.*

*Proof.* Stable maps $\Sigma_g \to \mathbb{R}^2$ and $\Sigma_g \to S^2$ whose triples $(i, c, n)$ are one of the following lists (1) and (2) were obtained in [Kamenosono and Yamamoto 2009; Yamamoto 2010] respectively:

(1) Stable maps $f : \Sigma_g \to \mathbb{R}^2$ whose triples $(i, c, n)$ are

$$(i, c, n) = \begin{cases} (1, 0, 0) & \text{if } g = 0, \\ (1, 2, g+1) & \text{if } g \text{ is odd}, \\ (1, 0, g+2) & \text{otherwise.} \end{cases}$$

Degree $d \geq 0$ stable maps $f : \Sigma_g \to S^2$ whose triples $(i, c, n)$ are

$$(i, c, n) = \begin{cases} (1, 2d, 0) & \text{if } g = 0, \\ (1, 2(d-g), 2g+2) & \text{if } d \geq g \geq 1, \\ (1, 0, d+g+2) & \text{if } d \leq g \text{ and } d \equiv g, \\ (1, 2, d+g+1) & \text{if } d < g \text{ and } d \not\equiv g. \end{cases}$$

(2) Stable maps $f : \Sigma_g \to \mathbb{R}^2$ whose triples $(i, c, n)$ are

$$(i, c, n) = \begin{cases} (1, 0, 0) & \text{if } g = 0, \\ (1, 2g+2, 1) & \text{otherwise.} \end{cases}$$

Degree $d \geq 0$ stable maps $f : \Sigma_g \to S^2$ whose triples $(i, c, n)$ are

$$(i, c, n) = \begin{cases} (1, 2d, 0) & \text{if } g = 0, \\ (1, 2(g+2), 0) & \text{if } d = 0 \text{ and } g \geq 1, \\ (1, 2(d+g), 0) & \text{otherwise.} \end{cases}$$

On the other hand, there exists a degree one stable map $T^2 \to S^2$ whose triple $(i, c, n)$ is equal to $(1, 2, 3)$; see [Kamenosono and Yamamoto 2009] for the details.

By applying modification (5) or (6), (9), (11) for stable maps $\Sigma_g \to \mathbb{R}^2$ and $\Sigma_g \to S^2$ in these lists, and a stable map $T^2 \to S^2$ whose triple $(i, c, n)$ is equal to $(1, 2, 3)$, we obtain the desired stable maps $\Sigma_g \to S^2$. □

Then, by applying modifications (1) and (2), (3), (10) inductively to stable maps $\Sigma_g \to S^2$ in Proposition 2.3, we obtain each stable map $\Sigma_g \to S^2$ in the list of Theorem 1.2(3) and (4).

Let $i \geq 1$ with $i \geq g+d$. By applying modifications (1) and (2), (3), (10) to degree zero stable maps $\Sigma_g \to S^2$ whose triples $(i, c, n)$ are $(g+1, 0, 0)$ and $(g+2, 2, 0)$, we obtain each stable map $\Sigma_g \to S^2$ in the list of Theorem 1.2(5).

Thus, we obtain all stable maps $\Sigma_g \to S^2$ whose triples $(i, c, n)$ are in the lists of Theorem 1.2.

## 3. Proof of Theorems 1.1 and 1.2

**3A.** ***Preparation.*** In this subsection, some notions concerning the apparent contour of a stable map $M \to S^2$ of a closed surface are introduced, where $M$ is a closed surface and $S^2$ is oriented.

Let $\varphi : M \to S^2$ be a stable map whose contour is nonempty. Let $S(\varphi) = S_1 \cup \cdots \cup S_\ell$ be the decomposition of $S(\varphi)$ into the connected components and set $\gamma_i = \varphi(S_i)$ $(i = 1, \ldots, \ell)$. Note that $\gamma(\varphi) = \gamma_1 \cup \cdots \cup \gamma_\ell$. Let $m(\varphi)$ be the smallest number of elements in the set $\varphi^{-1}(y)$, where $y \in S^2$ runs over all regular values of $\varphi$. Fix a regular value $\infty$ such that $\varphi^{-1}(\infty)$ consists of $m(\varphi)$ points. For each $\gamma_i$, denote by $U_i$ the component of $S^2 \setminus \gamma_i$ which contains $\infty$. Note that $\partial U_i \subset \gamma_i$.

Orient $\gamma_i$ so that at each fold point image, the surface is "folded to the left hand side." More precisely, for a point $y \in \gamma_i$ which is not a cusp or a node, choose a normal vector $v$ of $\gamma_i$ at $y$ such that $\varphi^{-1}(y')$ contains more elements than $\varphi^{-1}(y)$, where $y'$ is a regular value of $\varphi$ close to $y$ in the direction of $v$. Let $\tau$ be a tangent vector of $\gamma_i$ at $y$ such that the ordered pair $(\tau, v)$ is compatible with the given orientation of $S^2$. It is easy to see that $\tau$ gives a well-defined orientation for $\gamma_i$.

**Definition 3.1.** A point $y \in \partial U_i \setminus \{\text{cusps, nodes}\}$ is said to be *positive* if the normal orientation $v$ at $y$ points toward $U_i$. Otherwise, it is said to be *negative*.

A component $\gamma_i$ is said to be *positive* if all points of $\partial U_i \setminus \{\text{cusps, nodes}\}$ are positive; otherwise, $\gamma_i$ is said to be *negative*. The number of positive and negative components is denoted by $i^+$ and $i^-$, respectively. Note that there is at least one negative component unless $S(f) = \varnothing$.

**Definition 3.2.** A point $y \in \partial U_i \setminus \{\text{cusps, nodes}\}$ is called an *admissible starting point* if $y$ is a positive point of a positive component $\gamma_i$ (or a negative point of a negative component). Note that for each $i$, there always exists an admissible starting point on $\gamma_i$.

**Definition 3.3.** Suppose that $y \in \gamma_i$ is an admissible starting point and $Q \in \gamma_i$ is a node. Let $\alpha : [0, 1] \to \gamma_i$ be a parametrization consistent with the orientation, singular only when the image is a cusp such that $\alpha^{-1}(y) = \{0, 1\}$. Then, there are two numbers $0 < t_1 < t_2 < 1$ satisfying $\alpha(t_1) = \alpha(t_2) = Q$.

We say that $Q$ is *positive* if the orientation of $S^2$ at $Q$ defined by the ordered pair $(\alpha'(t_1), \alpha'(t_2))$ coincides with that of $S^2$ at $Q$; *negative*, otherwise.

The number of positive nodes on $\gamma_i$ is denoted by $N_i^+$ (and negative nodes by $N_i^-$). The definition of a positive or negative node on $\gamma_i$ depends on the choice of an admissible starting point $y$. However, it is known that the difference $N_i^+ - N_i^-$ does not depend on the choice of $y$; see [Whitney 1941] for details. Thus, the number $N^+ - N^- = \sum_{i=1}^{k}(N_i^+ - N_i^-)$ is well defined. Note that nodes arising from $\gamma_i \cap \gamma_j$ $(i \neq j)$ play no role in the computation.

Then, we obtain the following as an easy application of Pignoni's formula.

**Proposition 3.4** [Kamenosono and Yamamoto 2009; Pignoni 1993]. *For a stable map $\varphi : M \to S^2$ of a closed surface of genus $g$, we have*

$$g = \varepsilon(M)\big((N^+ - N^-) + \tfrac{1}{2}c(\varphi) + (1 + i^+ - i^-) - m(\varphi)\big), \tag{3-1}$$

*where $\varepsilon(M)$ is equal to* 1 *if $M$ is orientable, and* 2 *otherwise.*

Note that even if a 5-tuple $(N^+, N^-, c, i^+, i^-)$ satisfies formula (3-1), there may not be a stable map $f : M \to S^2$ with $S(f)$ consisting of $i^+ + i^-$ components, $c$ cusps, and $N^+ + N^-$ nodes.

In the following of this section, we assume that $\gamma_i \cap \gamma_j = \varnothing$ if $i \neq j$ because we study the minimal number of nodes. Denote by $U_\infty \subset S^2 \setminus \gamma(\varphi)$ the component which contains $\infty$. Denote by $\gamma_1$ the component of $\gamma(\varphi)$ which contains $\partial U_\infty$. Note that $\gamma_1$ is a negative component of $\varphi$. Then, the following lemmas and corollary were obtained by Fukuda and Yamamoto.

**Lemma 3.5** [Yamamoto 2010]. *If $\gamma_1$ has a node, then it has a negative node.*

**Lemma 3.6** [Yamamoto 2010]. *If a positive component $\gamma_i$ has a node, then it has a positive node.*

**Corollary 3.7** [Fukuda and Yamamoto 2011]. *If the number of negative components of $\gamma(\varphi)$ is equal to one and $\gamma(\varphi)$ has a node, then it has a negative node.*

Corollary 3.7 implies the following corollary.

**Corollary 3.8.** *If the number of negative components of $\gamma(\varphi)$ is equal to one and $\gamma_1$ has no node, then it has no node.*

Formula (3-1) and Lemma 3.6 imply the following lemma.

**Lemma 3.9.** *Suppose that $g \geq 1$ and $f : \Sigma_g \to \mathbb{R}^2$ is a stable map with $2 \leq i(f) \leq g$. If $\gamma_1$ has no node, then $\gamma(f)$ has at least two negative components.*

*Proof.* Assume that $\gamma(f)$ has only one negative component. Then, Lemma 3.6 and the assumption imply that $\gamma(f)$ has no node. Then, by the geometrical condition for a cusp, $\gamma(f)$ has no cusps. Thus, the formula (3-1) implies the contradiction

$$0 \leq g - i(f) = -1. \qquad \square$$

By formula (3-1) and the three modifications (1), (2), and (3), in order to prove Theorem 1.1, for a given triple $(g, i, c)$, we only have to study the minimal number of nodes among stable maps $f : \Sigma_g \to \mathbb{R}^2$ with $S(f)$ consisting of $i$ components and $c$ cusps. Analogously, in order to prove Theorem 1.2, for a given 4-tuple $(g, d, i, c)$, we only have to study the minimal number of nodes among degree $d$ stable maps $f : \Sigma_g \to S^2$ with $S(f)$ consisting of $i$ components and $c$ cusps.

Note that for a fixed pair $(i^+, i^-)$, if we increase the number of negative node by one, there are two ways to satisfy (3-1). One way is to increase the number of cusps by two. This corresponds to modification (1). The other way is to increase the number of positive node by one. This corresponds to modification (2).

**Lemma 3.10.** *Let $g \geq 0$ and $d \geq 0$. If a degree $d$ stable map $f : \Sigma_g \to S^2$ satisfies*

$$\tfrac{1}{2}c(f) \equiv g + d + i(f) \mod 2,$$

*then $f$ has at least one node.*

*Proof.* If $f$ has no node, then (3-1) implies that

$$g + m(f) + 2i^- = \tfrac{1}{2}c(f) + 1 + i(f). \qquad \square$$

In particular, Lemma 3.10 implies that if a stable map $f : \Sigma_g \to \mathbb{R}^2$ satisfies $\frac{1}{2}c(f) \equiv g + i(f) \bmod 2$, then $f$ has at least one node.

We recall a formula obtained by Eliasberg and Quine.

**Theorem 3.11** [Eliashberg 1970; Quine 1978]. *For a stable map $f : M \to N$ between closed connected oriented surfaces, we have*

$$\chi(M) - 2\chi(M_-) + \sum_{q_k \textit{ is a cusp}} \operatorname{sign}(q_k) = (\deg f)\,\chi(N) \tag{3-2}$$

*where $\chi$ denotes the Euler characteristic,* $\deg f$ *denotes the mapping degree of $f$, $M_-$ is the closure of the set of regular points whose neighborhoods are mapped by $f$ in an orientation reversing way, and* $\operatorname{sign}(q_k) = \pm 1$ *is the sign of a cusp $q_k$ defined as the local mapping degree.*

Then, Theorem 3.11 implies the following lemma.

**Lemma 3.12.** *Let $g \geq 0$, $d \geq 0$, and $f : \Sigma_g \to S^2$ be a degree $d$ stable maps. Then, $\gamma(f)$ has at least $2(d + 1 - g - i)$ cusps.*

*Proof.* Formula (3-2) implies that

$$\sum_{q_k \text{ is a cusp}} \operatorname{sign}(q_k) = 2(d + g - 1 + \chi((\Sigma_g)_-).$$

On the other hand, we have $\chi(\Sigma_g) - i \leq \chi((\Sigma_g)_-) \leq i$. $\square$

In particular, for a stable map $f : \Sigma_g \to \mathbb{R}^2$, $\gamma(f)$ has at least $2(g + 1 - i)$ cusps. Theorem 3.11 also implies the following lemma.

**Lemma 3.13** [Fukuda and Yamamoto 2011]. *Let $d \geq 0$ and $f : \Sigma_g \to S^2$ be a degree $d$ stable map. If $i(f) \equiv d + g \bmod 2$, then $\gamma(f)$ has at least two cusps.*

In particular, for a stable map $f : M \to \mathbb{R}^2$, if $i(f) \equiv g \bmod 2$, then $\gamma(f)$ has at least two cusps.

Furthermore, if the contour has no nodes, then we have the following lemma.

**Lemma 3.14.** (1) *Let $g \geq 0$ and $f : \Sigma_g \to \mathbb{R}^2$ be a stable map with $2 \leq i(f) \leq g$. If $\gamma(f)$ has no nodes, then $c(f) \geq 2(g+3-i)$.*

(2) *Let $g \geq 0$ and $d \geq 0$, $f : \Sigma_g \to S^2$ be a degree $d$ stable map with $2 \leq i(f) \leq g$. If $\gamma(f)$ has no nodes, then $c(f) \geq 2(g+d+1-i)$.*

*Proof.* Let us consider (1). Then, formula (3-1) implies that

$$g + 2i^- - 1 - i(f) = \tfrac{1}{2}c(f).$$

Then, Lemma 3.9 yields the conclusion.

The case (2) is also proved in a similar way. □

### 3B. *Proof of Theorem 1.1.*

**Lemma 3.15.** *Let $g \geq 1$ and $f : \Sigma_g \to \mathbb{R}^2$ be a stable map with $1 \leq i(f) \leq g$. If $c(f) \leq 2(g+2-i(f))$, then $n(f) \geq -\frac{1}{2}c(f) + g + 3 - i(f)$.*

*Proof.* Assume $i(f) = 1$. In this case, for a stable map $f : \Sigma_g \to \mathbb{R}^2$ with $i(f) = 1$, formula (3-1) implies that

$$g - \tfrac{1}{2}c(f) = (N^+ - N^-). \tag{3-3}$$

Then, Lemma 3.5 implies that

$$n(f) = N^+ + N^- = -\tfrac{1}{2}c(f) + g + 2N^- \geq -\tfrac{1}{2}c(f) + g + 2.$$

Assume $i(f) \geq 2$. If the negative component $\gamma_1$ has a node, then formula (3-1) and Lemma 3.5 imply that

$$n(f) = N^+ + N^- \geq -\tfrac{1}{2}c(f) + g + 3 - i(f).$$

If the negative component $\gamma_1$ has no node, then formula (3-1) and Lemma 3.9 also imply that $n(f) \geq -\frac{1}{2}c(f) + g + 3 - i(f)$. □

**Lemma 3.16.** *Let $g \geq 0$ and $f : \Sigma_g \to \mathbb{R}^2$ be a stable map. If $c(f) \geq 2(g + i(f))$, then $N^- \geq \frac{1}{2}c(f) - g - i(f) + 1$.*

*Proof.* Formula (3-1) and the inequality $i^+ - i^- \geq -i(f)$ imply that

$$g \geq (N^+ - N^-) + \tfrac{1}{2}c(f) + (1 - i(f)).$$ □

Lemmas 3.10, 3.12, 3.13, 3.14(1), 3.15, and 3.16 prove Theorem 1.1 with $1 \leq i \leq g$. Lemmas 3.12, 3.13, 3.14(1) and 3.16 prove Theorem 1.1 with $i > g$.

This complete the proof of Theorem 1.1.

**3C.** ***Proof of Theorem 1.2.*** Lemmas 3.12 and 3.13 prove Theorem 1.2 with $g = 0$. Lemma 3.13 proves Theorem 1.2(2) and (5).

**Lemma 3.17.** *Let $g \geq 1$. For a degree zero stable map $f : \Sigma_g \to S^2$, if $S(f)$ consists of one component and $\gamma(f)$ has no nodes, then $m(f) \geq 2$.*

*Proof.* Under the assumption, formula (3-1) implies that

$$g = \tfrac{1}{2}c(f) - m(f).$$

By the geometrical condition for a cusp, if $n(f) = 0$ and $m(f) = 0$, then $f$ has no cusps. Then, we have $g = 0$, which is a contradiction. □

Lemma 3.17 and formula (3-1) imply the following lemma.

**Lemma 3.18.** *Let $g \geq 1$ and $f : \Sigma_g \to S^2$ be a degree zero stable map with $1 \leq i(f) \leq g$. If $c(f) \leq 2(g + 3 - i(f))$, then $n(f) \geq -\frac{1}{2}c(f) + g + 3 - i(f)$.*

*Proof.* Formula (3-1) implies that

$$n(f) = N^+ + N^- = g + 2i^- - i(f) - \tfrac{1}{2}c(f) - 1 + 2N^-.$$

Consider the case that $i(f) = 1$. Then, by Lemma 3.5, $n(f) \geq -\frac{1}{2}c(f) + g + 2$. Note that there is no degree zero stable map $f : \Sigma_g \to S^2$ with $S(f)$ consisting of one component and no nodes unless $g = 0$.

Now consider the case that $2 \leq i(f) \leq g$. If $\gamma_1$ has a node, then Lemma 3.5 implies that

$$n(f) \geq -\tfrac{1}{2}c(f) + g + 3 - i(f).$$

If $\gamma_1$ has no node, then Lemma 3.9 also implies the same inequality. □

Let $f : \Sigma_g \to S^2$ be a degree zero stable map with no nodes and $2 \leq i(f) \leq g$.

If $m(f) = 0$, then $f$ induces a stable map $\Sigma_g \to \mathbb{R}^2$ whose triple $(i, c, n)$ is equal to that of $f$. Then, Lemma 3.9 and formula (3-1) imply that

$$g + 4 \leq g + 2i^- = \tfrac{1}{2}c(f) + (1 + i(f)).$$

This inequality shows that $c(f) \geq 2(g + 3 - i(f))$.

If $m(f) \neq 0$, then Lemma 3.5 and formula (3-1) imply that

$$g + 4 \leq g + m(f) + 2i^- = \tfrac{1}{2}c(f) + (1 + i(f)).$$

This inequality also shows that $c(f) \geq 2(g + 3 - i(f))$.

Thus, Lemma 3.10 and Lemmas 3.12, 3.18 prove Theorem 1.2(1).

Lemma 3.17 and formula (3-1) imply the following lemma.

**Lemma 3.19.** *Let $g \geq 1$ and $d \geq 1$, $f : \Sigma_g \to S^2$ be a degree $d$ stable map with $1 \leq i(f) \leq g + d - 1$. If $c(f) \leq 2(g + d - i)$, then $n(f) \geq -\frac{1}{2}c(f) + g + d + 3 - i(f)$.*

*Proof.* Let us consider the case $i(f) = 1$. In this case, the formula (3-1) implies that

$$n(f) = N^+ + N^- \geq -\tfrac{1}{2}c(f) + g + d + 2N^-$$

Then, Lemma 3.5 yields the conclusion.

Let us consider the case $2 \leq i(f) \leq g + d - 1$. If $i^- = 1$, then the formula (3-1) implies that

$$n(f) = N^+ + N^- \geq -\tfrac{1}{2}c(f) + g + d + 1 - i(f) + 2N^-.$$

Thus, Lemma 3.5 and Corollary 3.8 yield the conclusion. If $i^- \geq 2$, then the formula (3-1) implies that

$$n(f) = N^+ + N^- \geq -\tfrac{1}{2}c(f) + g + d + 3 - i(f) + 2N^-. \qquad \square$$

Thus, Lemma 3.10 and Lemmas 3.12, 3.19 prove Theorem 1.2(3) and (4). It completes the proof of Theorem 1.2.

## 4. Problems

In this section, we pose some problems which concern the number of the singularities of stable maps between surfaces.

**Problem 4.1.** Study the triples $(i, c, n)$ of stable maps $M \to N$ ($N = \mathbb{R}^2$ or $N = S^2$) of closed and nonorientable surfaces.

Pignoni [1993] (see also [Kamenosono and Yamamoto 2009]) observed that there are differences between $(i, c+n)$-minimal contours and $(i, c, n)$-minimal contours — see Remark 1.4 for the definitions — of $C^\infty$ maps of the real projective plane into $\mathbb{R}^2$ and $S^2$.

Figure 13 shows that the contours of stable maps $S^2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(2, 2, 2)$. Figure 14 also shows that the contours of stable maps $T^2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are equal to $(2, 0, 4)$.

**Problem 4.2.** Introduce notions which distinguish two contours in Figure 13 (or 14). Then, study contours of stable maps between surfaces under the notions.

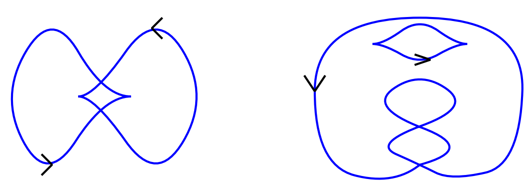

**Figure 13.** Contours of stable maps $S^2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are $(2, 2, 2)$.

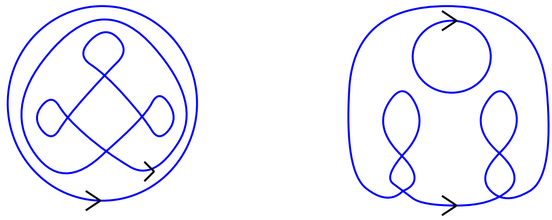

**Figure 14.** Contours of stable maps $T^2 \to \mathbb{R}^2$ whose triples $(i, c, n)$ are $(2, 0, 4)$.

Problem 4.2 makes the second step toward classifying generic $C^\infty$ maps of closed surfaces into $\mathbb{R}^2$ or $S^2$ up to right-left equivalence.

## Appendix

In this section, we introduce the notions of an $i$-$(c, n)$-minimal contour and an $i$-$(n, c)$-minimal contour for a $C^\infty$ map $M \to N$ between surfaces. We study such minimal contours.

Taishi Fukuda and the author [Fukuda and Yamamoto 2011] studied $(c+n)$-minimal contours among stable maps $f : \Sigma_g \to S^2$ homotopic to a given $C^\infty$ map $\Sigma_g \to S^2$ such that $i(f) = i$, for each integer $i \geq 2$. Let us call such a minimal contour an $i$-$(c+n)$-*minimal contour*. Note that the case $g = 2$ of [Fukuda and Yamamoto 2011, Theorem 1.2] has one error. The correct table of $i$-$(c+n)$-minimal contours for degree $d \geq 0$ stable maps $\Sigma_2 \to S^2$ is the following:

$$(c, n) = \begin{cases} (2(d-i-1), 6) & \text{if } 1 \leq i \leq d-1, \\ (2, 4) \text{ or } (6, 0) & \text{if } i = d, \\ (0, 4) \text{ or } (4, 0) & \text{if } i = d+1, \\ (2, 2) & \text{if } (d, i) = (0, 2), \\ (2, 0) & \text{if } i \geq d+2,\ i \equiv d \bmod 2, \text{ except } (d, i) = (0, 2), \\ (0, 0) & \text{if } i \geq d+2,\ i \not\equiv d \bmod 2, \end{cases}$$

For a nonnegative integer $i$, let us consider $(c, n)$-minimal contours among stable maps $f : \Sigma_g \to S^2$ homotopic to a given $C^\infty$ map $\Sigma_g \to S^2$ such that $i(f) = i$. Let us call such a minimal contour an $i$-$(c, n)$-*minimal contour*. Then, Theorems 1.1 and 1.2 imply the following proposition.

**Proposition A.1.** (1) *The contour $\gamma(f)$ of a stable map $f : \Sigma_g \to \mathbb{R}^2$ is $i$-$(c, n)$-minimal if and only if the pair $(c, n)$ is one of the following*:

$$(c,n)=\begin{cases}(2,g+2-i) & \textit{if } g\geq i \textit{ and } g\equiv i \mod 2,\\ (0,g+3-i) & \textit{if } g\geq i \textit{ and } g\not\equiv i \mod 2,\\ (2,0) & \textit{if } g< i \textit{ and } g\equiv i \mod 2,\\ (0,0) & \textit{if } g< i \textit{ and } g\not\equiv i \mod 2.\end{cases}$$

(2) *Let $f:\Sigma_g\to S^2$ be a degree $d\geq 0$ stable map such that $S(f)$ consists of $i$ components. Then, the contour $\gamma(f)$ is $i$-$(c,n)$-minimal if and only if the pair $(c,n)$ for $\gamma(f)$ is one of the following*:

$g=0$:

$$(c,n)=\begin{cases}(2(d-i+1),0) & \textit{if } 1\leq i\leq d+1,\\ (2,0) & \textit{if } i\geq d+2,\ i\equiv d \mod 2,\\ (0,0) & \textit{if } i\geq d+2,\ i\not\equiv d \mod 2,\end{cases}$$

$g\geq 1$:

$$(c,n)=\begin{cases}(2(d-g-i+1),2+2g) & \textit{if } 1\leq i\leq d-g+1,\\ (2,d+g-i+2) & \textit{if } d-g+2\leq i< d+g-1,\\ & \quad \textit{and } i\equiv d+g \mod 2,\\ (0,d+g-i+3) & \textit{if } d-g+2\leq i\leq d+g-1,\\ & \quad \textit{and } i\not\equiv d+g \mod 2,\\ (2,2) & \textit{if } (d,i)=(0,g),\\ (2,0) & \textit{if } i\geq d+g,\ i\equiv d+g \mod 2,\\ & \quad \textit{except } (d,i)=(0,g),\\ (0,0) & \textit{if } i\geq d+g,\ i\not\equiv d+g \mod 2.\end{cases}$$

Let us study $(n,c)$-minimal contours among stable maps $f:\Sigma_g\to S^2$ homotopic to a given $C^\infty$ map $\Sigma_g\to S^2$ such that $i(f)=i$, for each integer $i\geq 1$. Let us call such a minimal contour an *$i$-$(n,c)$-minimal contour*. Then, Theorems 1.1 and 1.2 also imply the following proposition.

**Proposition A.2.** (1) *The contour $\gamma(f)$ of a stable map $f:\Sigma_g\to\mathbb{R}^2$ is $i$-$(n,c)$-minimal if and only if the pair $(c,n)$ is one of the following*:

$i=1$:

$$(c,n)=\begin{cases}(0,0) & \textit{if } g=0,\\ (2g+2,1) & \textit{otherwise},\end{cases}$$

$i\geq 2$:

$$(c,n)=\begin{cases}(2(g+3-i),0) & \textit{if } g\geq i,\\ (2,0) & \textit{if } g< i \textit{ and } g\equiv i \mod 2,\\ (0,0) & \textit{otherwise}.\end{cases}$$

(2) *Let* $f : \Sigma_g \to S^2$ *be a degree* $d \geq 0$ *stable map such that* $S(f)$ *consists of* $i$ *components. Then, the contour* $\gamma(f)$ *is* $i$*-*$(n, c)$*-minimal if and only if the pair* $(c, n)$ *for* $\gamma(f)$ *is one of the items below*:

$$(c, n) = \begin{cases} (2(g+3-i), 0) & \text{if } d = 0 \text{ and } 1 \leq i \leq g, \\ (2(d+g+1-i), 0) & \text{if } d \neq 0 \text{ and } 1 \leq i \leq d+g-1, \\ (2, 0) & \text{if } i \geq d+g \text{ and } i \equiv d+g \bmod 2, \\ (0, 0) & \text{if } i > d+g \text{ and } i \not\equiv d+g \bmod 2. \end{cases}$$

## Acknowledgments

The author would like to express his gratitude to Andras Szűcs for the fruitful questions and comments at the International Workshop on Real and Complex Singularities held in São Carlos in 2012. He would like to thank Osamu Saeki for useful comments and suggestions which improved the paper. He would like to thank the referee for useful suggestions that helped him to improve the original manuscript. The author also expresses special thanks to Akiko Neriugawa for the useful advice in English grammar and for encouraging support. This work was supported by JSPS KAKENHI Grant Numbers 21740056 and 23654028.

## References

[Demoto 2005] S.-I. Demoto, "Stable maps between 2-spheres with a connected fold curve", *Hiroshima Math. J.* **35**:1 (2005), 93–113. MR 2006b:57041 Zbl 1080.57029

[Eliashberg 1970] Y. M. Eliashberg, "Об особенностях типа складки", *Izv. Akad. Nauk SSSR Ser. Mat.* **34**:5 (1970), 1110–1126. Translated as "On singularities of folding type" in *Math. USSR Izv.* **4**:5 (1970), 1119–1134. MR 43 #4051 Zbl 0226.57012

[Fukuda and Yamamoto 2011] T. Fukuda and T. Yamamoto, "Apparent contours of stable maps into the sphere", *J. Singul.* **3** (2011), 113–125. MR 2012k:58064 Zbl 1292.57022

[Kamenosono and Yamamoto 2009] A. Kamenosono and T. Yamamoto, "The minimal number of singularities of stable maps between surfaces", *Topology Appl.* **156**:14 (2009), 2390–2405. MR 2010m:57041 Zbl 1207.57039

[Pignoni 1993] R. Pignoni, "Projections of surfaces with a connected fold curve", *Topology Appl.* **49**:1 (1993), 55–74. MR 93m:57030 Zbl 0767.57020

[Quine 1978] J. R. Quine, "A global theorem for singularities of maps between oriented 2-manifolds", *Trans. Amer. Math. Soc.* **236** (1978), 307–314. MR 57 #14020 Zbl 0379.57006

[Whitney 1941] H. Whitney, "On regular families of curves", *Bull. Amer. Math. Soc.* **47** (1941), 145–147. MR 2,322b Zbl 0025.23602

[Whitney 1955] H. Whitney, "On singularities of mappings of euclidean spaces, I: Mappings of the plane into the plane", *Ann. of Math.* (2) **62** (1955), 374–410. MR 17,518d Zbl 0068.37101

[Yamamoto 2009] M. Yamamoto, "The number of singular set components of fold maps between oriented surfaces", *Houston J. Math.* **35**:4 (2009), 1051–1069. MR 2011e:58064 Zbl 1186.57021

[Yamamoto 2010] T. Yamamoto, "Apparent contours with minimal number of singularities", *Kyushu J. Math.* **64**:1 (2010), 1–16. MR 2011m:57038 Zbl 1195.57064

Received August 9, 2014. Revised March 25, 2015.

TAKAHIRO YAMAMOTO
FACULTY OF ENGINEERING
KYUSHU SANGYO UNIVERSITY
3-1 MATSUKADAI 2-CHOME
HIGASHIKU
FUKUOKA 813-8503
JAPAN
yama.t@ip.kyusan-u.ac.jp

# CONTENTS

Volume 280, no. 1 and no. 2

## Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

# PACIFIC JOURNAL OF MATHEMATICS

Volume 280 No. 2 February 2016





0030-8730(2016)280:2;1-5