Pacific Journal of Mathematics

Volume 281 No. 1 March 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. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.1

COMPATIBLE SYSTEMS OF SYMPLECTIC GALOIS REPRESENTATIONS AND THE INVERSE GALOIS PROBLEM, II: TRANSVECTIONS AND HUGE IMAGE

SARA ARIAS-DE-REYNA, LUIS DIEULEFAIT AND GABOR WIESE

This article is the second part of a series of three articles about compatible systems of symplectic Galois representations and applications to the inverse Galois problem.

This part is concerned with symplectic Galois representations having a huge residual image, by which we mean that a symplectic group of full dimension over the prime field is contained up to conjugation. A key ingredient is a classification of symplectic representations whose image contains a nontrivial transvection: these fall into three very simply describable classes, the reducible ones, the induced ones and those with huge image. Using the idea of an (n, p)-group of Khare, Larsen and Savin, we give simple conditions under which a symplectic Galois representation with coefficients in a finite field has a huge image. Finally, we combine this classification result with the main result of the first part to obtain a strengthened application to the inverse Galois problem.

1. Introduction

This article is the second of a series of three about compatible systems of symplectic Galois representations and applications to the inverse Galois problem.

This part is concerned with symplectic Galois representations having a *huge image*: for a prime ℓ, a finite subgroup $G \subseteq \mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$ is called *huge* if it contains a conjugate (in $\mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$) of $\mathrm{Sp}_n(\mathbb{F}_\ell)$. By Corollary 1.3 below, this notion is the same as the one introduced in Part I [Arias-de-Reyna et al. 2013].

Whereas the classification of the finite subgroups of $\mathrm{Sp}_n(\overline{\mathbb{F}}_\ell)$ appears very complicated to us, it turns out that the finite subgroups containing a nontrivial transvection can be very cleanly classified into three classes, one of which is that of huge subgroups (see Theorem 1.1 below). Translating this group theoretic result into the language of symplectic representations whose image contains a nontrivial

MSC2010: 11F80, 20G40, 12F12.
Keywords: compatible systems of symplectic Galois representations, inverse Galois problem.

transvection, these also fall into three very simply describable classes: the reducible ones, the induced ones and those with huge image (see Corollary 1.2).

Using the idea of an (n, p)-group of [Khare et al. 2008] (i.e., of a maximally induced place of order p, in the terminology of Part I), some number theory allows us to give very simple conditions under which a symplectic Galois representation with coefficients in $\overline{\mathbb{F}}_\ell$ has huge image (see Theorem 1.5 below).

This second part is independent of the first, except for Corollary 1.6, which combines the main results of Part I [Arias-de-Reyna et al. 2013] and the present Part II. In Part III [Arias-de-Reyna et al. 2015], written in collaboration with Sug Woo Shin, a compatible system satisfying the assumptions of Corollary 1.6 is constructed.

Statement of results. To fix terminology, we recall some standard definitions. Let K be a field. An n-dimensional K-vector space V equipped with a symplectic form (i.e., nonsingular and alternating), denoted by $\langle v, w\rangle = v \bullet w$ for $v, w \in V$, is called a *symplectic K-space*. A K-subspace $W \subseteq V$ is called a *symplectic K-subspace* if the restriction of $\langle\cdot,\cdot\rangle$ to $W \times W$ is nonsingular (hence, symplectic). The *general symplectic group* $\mathrm{GSp}(V, \langle\cdot,\cdot\rangle) =: \mathrm{GSp}(V)$ consists of those $A \in \mathrm{GL}(V)$ such that there is $\alpha \in K^\times$, the *multiplier* (or *similitude factor*) of A, such that we have $(Av)\bullet(Aw) = \alpha(v \bullet w)$ for all $v, w \in V$. The *symplectic group* $\mathrm{Sp}(V, \langle\cdot,\cdot\rangle) =: \mathrm{Sp}(V)$ is the subgroup of $\mathrm{GSp}(V)$ of elements with multiplier 1. An element $\tau \in \mathrm{GL}(V)$ is a *transvection* if $\tau - \mathrm{id}_V$ has rank 1, i.e., if τ fixes a hyperplane pointwisely, and there is a line U such that $\tau(v) - v \in U$ for all $v \in V$. We will consider the identity as a "trivial transvection". Any transvection has determinant and multiplier 1. A *symplectic transvection* is a transvection in $\mathrm{Sp}(V)$. Any symplectic transvection has the form

$$T_v[\lambda] \in \mathrm{Sp}(V) : u \mapsto u + \lambda(u \bullet v)v$$

with *direction vector* $v \in V$ and *parameter* $\lambda \in K$; see, e.g., [Artin 1957, pp. 137–138].

The classification result on subgroups of general symplectic groups containing a nontrivial transvection which plays the key role in our approach is the following.

Theorem 1.1. *Let K be a finite field of characteristic at least 5 and V a symplectic K-vector space of dimension n. Then any subgroup G of $\mathrm{GSp}(V)$ which contains a nontrivial symplectic transvection satisfies one of the following assertions*:

(1) *There is a proper K-subspace $S \subset V$ such that $G(S) = S$.*

(2) *There are mutually orthogonal nonsingular symplectic K-subspaces $S_i \subset V$ with $i = 1, \dots, h$ of dimension m for some $m < n$ such that $V = \bigoplus_{i=1}^h S_i$ and for all $g \in G$, there is a permutation $\sigma_g \in \mathrm{Sym}_h$ (the symmetric group on $\{1, \dots, h\}$) with $g(S_i) = S_{\sigma_g(i)}$. Moreover, the action of G on the set $\{S_1, \dots, S_h\}$ thus defined is transitive.*

(3) *There is a subfield L of K such that the subgroup generated by the symplectic transvections of G is conjugated (in $\mathrm{GSp}(V)$) to $\mathrm{Sp}_n(L)$.*

In Section 2 we show how this theorem can be deduced from results of Kantor [1979]. In a previous version of this article, we gave a self-contained proof, which is still available on arXiv. For our application to Galois representations, we provide the following representation theoretic reformulation of Theorem 1.1.

Corollary 1.2. *Let ℓ be a prime at least 5, let Γ be a compact topological group and*

$$\rho : \Gamma \to \mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$$

a continuous representation (for the discrete topology on $\overline{\mathbb{F}}_\ell$). Assume that the image of ρ contains a nontrivial transvection. Then one of the following assertions holds:

(1) *ρ is reducible.*

(2) *There is a closed subgroup $\Gamma' \subsetneq \Gamma$ of finite index $h \mid n$ and a representation $\rho' : \Gamma' \to \mathrm{GSp}_{n/h}(\overline{\mathbb{F}}_\ell)$ such that $\rho \cong \mathrm{Ind}_{\Gamma'}^{\Gamma}(\rho')$.*

(3) *There is a finite field L of characteristic ℓ such that the subgroup generated by the symplectic transvections in the image of ρ is conjugated (in $\mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$) to $\mathrm{Sp}_n(L)$; in particular, the image is huge.*

The following corollary shows that the definition of a huge subgroup of $\mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$, which we give in Part I [Arias-de-Reyna et al. 2013], coincides with the simpler definition stated above.

Corollary 1.3. *Let K be a finite field of characteristic $\ell \geq 5$, V a symplectic K-vector space of dimension n, and G a subgroup of $\mathrm{GSp}(V)$ which contains a symplectic transvection. Then the following are equivalent:*

(i) *G is huge.*

(ii) *There is a subfield L of K such that the subgroup generated by the symplectic transvections of G is conjugated (in $\mathrm{GSp}(V)$) to $\mathrm{Sp}_n(L)$.*

Combining the group theoretic results above with (n, p)-groups, introduced by [Khare et al. 2008], some number theory allows us to prove the following theorem. Before stating it, let us collect some notation.

Set-up 1.4. Let n, N be positive integers with n even and $N = N_1 \cdot N_2$ with $\gcd(N_1, N_2) = 1$. Let L_0 be the compositum of all number fields of degree $\leq n/2$, which are ramified at most at the primes dividing N_2 (which is a number field). Let q be a prime which is completely split in L_0, and let p be a prime dividing $q^n - 1$ but not dividing $q^{n/2} - 1$, and $p \equiv 1 \pmod{n}$.

Theorem 1.5. *Assume Set-up 1.4. Let $k \in \mathbb{N}$, $\ell \neq p, q$ be a prime such that $\ell > kn! + 1$ and $\ell \nmid N$. Let $\chi_q : G_{\mathbb{Q}_{q^n}} \to \overline{\mathbb{Q}}_\ell^\times$ be a character satisfying the assumptions of Lemma 3.1, and $\bar{\chi}_q$ the composition of χ_q with the reduction map $\overline{\mathbb{Z}}_\ell \to \overline{\mathbb{F}}_\ell$. Let $\bar{\alpha} : G_{\mathbb{Q}_q} \to \overline{\mathbb{F}}_\ell^\times$ be an unramified character.*

Let

$$\rho : G_{\mathbb{Q}} \to \mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$$

be a Galois representation, ramified only at the primes dividing $Nq\ell$, satisfying that a twist by some power of the cyclotomic character is regular in the sense of Definition 3.2 with tame inertia weights at most k, and such that

(1) $\mathrm{Res}^{G_{\mathbb{Q}}}_{G_{\mathbb{Q}_q}}(\rho) = \mathrm{Ind}^{G_{\mathbb{Q}_q}}_{G_{\mathbb{Q}_{q^n}}}(\bar{\chi}_q) \otimes \bar{\alpha}$,

(2) *the image of ρ contains a nontrivial transvection and*

(3) *for all primes ℓ_1 dividing N_1, the image under ρ of I_{ℓ_1}, the inertia group at ℓ_1, has order prime to $n!$.*

Then the image of ρ is a huge subgroup of $\mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$.

Combining Theorem 1.5 with the results of Part I [Arias-de-Reyna et al. 2013] of this series yields the following corollary.

Corollary 1.6. *Assume Set-up 1.4. Let $\rho_\bullet = (\rho_\lambda)_\lambda$ (where λ runs through the finite places of a number field L) be an n-dimensional a. e. absolutely irreducible a. e. symplectic compatible system, as defined in Part I [Arias-de-Reyna et al. 2013], for the base field $\mathbb{Q}$, which satisfies the following assumptions*:

- *For all places λ, the representation ρ_λ is unramified outside $Nq\ell$, where ℓ is the rational prime below λ.*
- *There are $a \in \mathbb{Z}$ and $k \in \mathbb{N}$ such that, for all but possibly finitely many places λ of L, the reduction mod λ of $\chi_\ell^a \otimes \rho_\lambda$ is regular in the sense of Definition 3.2, with tame inertia weights at most k.*
- *The multiplier of the system is a finite order character times a power of the cyclotomic character.*
- *For all primes ℓ not belonging to a density zero set of rational primes, and for each $\lambda \mid \ell$, the residual representation $\bar{\rho}_\lambda$ contains a nontrivial transvection in its image.*
- *For all places λ not above q, one has*

$$\mathrm{Res}^{G_{\mathbb{Q}}}_{G_{\mathbb{Q}_q}}(\rho_\lambda) = \mathrm{Ind}^{G_{\mathbb{Q}_q}}_{G_{\mathbb{Q}_{q^n}}}(\chi_q) \otimes \alpha,$$

  where $\alpha : G_{\mathbb{Q}_q} \to \overline{L}_\lambda^\times$ is some unramified character and $\chi_q : G_{\mathbb{Q}_{q^n}} \to \overline{\mathbb{Z}}^\times$ is a character such that its composite with the embedding $\overline{\mathbb{Z}}^\times \hookrightarrow \overline{\mathbb{Q}}_\ell^\times$ given by λ satisfies the assumptions of Lemma 3.1 for all primes $\ell \nmid pq$. In the terminology of Part I, q is called a maximally induced place of order p.
- *For all primes ℓ_1 dividing N_1 and for all but possibly finitely many places λ, the group $\bar{\rho}_\lambda(I_{\ell_1})$ has order prime to $n!$ (where I_{ℓ_1} denotes the inertia group at ℓ_1).*

Then we obtain:

(a) *For all primes* ℓ *not belonging to a density zero set of rational primes*, *and for each* $\lambda \mid \ell$, *the image of the residual representation* $\bar{\rho}_\lambda$ *is a huge subgroup of* $\mathrm{GSp}_n(\bar{\mathbb{F}}_\ell)$.

(b) *For any* $d \mid \frac{p-1}{n}$ *there exists a set* $\mathcal{L}_d$ *of rational primes* ℓ *of positive density such that for all* $\ell \in \mathcal{L}_d$, *there is a place* λ *of* L *above* ℓ *satisfying that the image of* $\bar{\rho}_\lambda^{\mathrm{proj}}$ *is* $\mathrm{PGSp}_n(\mathbb{F}_{\ell^d})$ *or* $\mathrm{PSp}_n(\mathbb{F}_{\ell^d})$.

The proofs of Theorem 1.5 and Corollary 1.6 are given in Section 3.

Remark 1.7. It is natural to ask which of the two alternatives in Corollary 1.6(b) actually holds. It is very hard to give a general answer. The same indeterminacy occurs in [Khare et al. 2008] (see its arXiv version arXiv:math/0610860v3).

In the very special case, when there is no residual inner twist at a prime λ, the multiplier determines which case one is in. More precisely, by definition there is no residual inner twist at λ if the residue fields modulo λ of $E_{\rho_\bullet}$ (the field of definition of $\rho_\bullet$) and $K_{\rho_\bullet}$ (the projective field of definition of $\rho_\bullet$) coincide (see [Arias-de-Reyna et al. 2013, Section 4]); call it $\mathbb{F}$. In that case, if the image of $\bar{\rho}_\lambda$ is huge, up to conjugation we have $\mathrm{Sp}_n(\mathbb{F}) \subseteq \bar{\rho}_\lambda(G_\mathbb{Q}) \subseteq \mathrm{GSp}_n(\mathbb{F})$ and thus $\mathrm{PSp}_n(\mathbb{F}) \subseteq \bar{\rho}_\lambda^{\mathrm{proj}}(G_\mathbb{Q}) \subseteq \mathrm{PGSp}_n(\mathbb{F})$. The first inclusion is an equality if and only if the multiplier of $\bar{\rho}_\lambda$ is a square in $\mathbb{F}$; otherwise the second inclusion is an equality.

When $n = 2$, in [Dieulefait and Wiese 2011] the difference between the field of definition and the projective one could be controlled due to special choices of modular forms; this allowed distinguishing between the two possibilities and, for every $d \geq 1$, realising the simple group $\mathrm{PSL}_2(\mathbb{F}_{\ell^d})$ as a Galois group over $\mathbb{Q}$ for a positive density set of primes ℓ.

If the two residue fields do not coincide, the multiplier is not enough to distinguish between the two cases.

2. Symplectic representations containing a transvection

This section is devoted to Theorem 1.1. This theorem can be deduced from more general results, like those of [Guralnick and Saxl 2003]. We prefer to deduce it from the results of Kantor [1979], together with some representation theory of groups. We hope that the detailed and quite elementary proof we give on page 6 will be of value to the number theory community.

Throughout the section, our setting will be the following: $\ell \geq 5$ denotes a prime number, n an even positive integer and V a symplectic n-dimensional vector space over a finite field K of characteristic ℓ.

Kantor's classification result. Kantor [1979] classifies subgroups of classical linear groups which are generated by a conjugacy class of elements of long root subgroups.

In this paper, we are only concerned with subgroups of the symplectic group $\mathrm{Sp}(V)$. This case is addressed in [Kantor 1979, §11].

We need some notation in order to state his result. First of all, recall that in the symplectic case, the elements of long root subgroups are precisely the symplectic transvections. Given a subgroup $H \subseteq \mathrm{Sp}(V)$, denote by $O_\ell(H)$ the maximal normal ℓ-subgroup contained in H, denote by $[H, H]$ the commutator subgroup of H, and by $Z_{\mathrm{Sp}(V)}(H)$ the centraliser of H in $\mathrm{Sp}(V)$. Below we state the result of Kantor in the symplectic case (leaving aside the cases of characteristic 2 and 3).

Theorem 2.1 (Kantor). *Assume that $\ell \geq 5$, and let $H \subseteq \mathrm{Sp}(V)$ be a subgroup satisfying the following conditions*:

(1) *There exists a set $\mathfrak{X} \subseteq H$ consisting of transvections, closed under conjugation in H, which generates H.*

(2) $O_\ell(H) \leq [H, H] \cap Z_{\mathrm{Sp}(V)}(H)$.

(3) *H does not preserve any nonsingular subspace of V.*

Then there is a subfield L of K such that H is conjugated (*in* $\mathrm{Sp}(V)$) *to* $\mathrm{Sp}_n(L)$.

We will apply this result in the case when H is an irreducible subgroup. In this case, conditions (2) and (3) are satisfied. We elaborate on condition (2). Let $W \subseteq V$ be the subspace of elements that are left invariant by all elements in $O_\ell(H)$. Since $O_\ell(H)$ is an ℓ-group acting on a finite ℓ-group V, the cardinality of W is divisible by ℓ (see Lemma 1 of Chapter IX of [Serre 1979]); hence $W \neq \{0\}$. Moreover, since $O_\ell(H)$ is a normal subgroup of H, it follows that H stabilises W. But H is an irreducible group; hence $W = V$ and $O_\ell(H) = \{\mathrm{Id}\}$. Furthermore, if we take into account that the conjugate of a transvection is again a transvection, we can reformulate condition (1) as follows: "the transvections contained in H generate H", or simply "H is generated by transvections". This discussion proves the following corollary.

Corollary 2.2. *Assume that $\ell \geq 5$, and let $H \subseteq \mathrm{Sp}(V)$ be an irreducible subgroup which is generated by transvections. Then there is a subfield L of K such that H is conjugated* (*in* $\mathrm{Sp}(V)$) *to* $\mathrm{Sp}_n(L)$.

Proof of the group theoretic results. We will make use of the following facts about transvections, the simple proofs of which are omitted.

Lemma 2.3. *Let $T_u[\lambda] \in \mathrm{Sp}(V)$ be a symplectic transvection. Then*

(a) *For any $A \in \mathrm{GSp}(V)$ with multiplier $\alpha \in K^\times$, we have $AT_u[\lambda]A^{-1} = T_{Au}\left[\frac{\lambda}{\alpha}\right]$.*

(b) *Suppose $W \subseteq V$ is a K-vector subspace stabilised by $T_u[\lambda]$ with $\lambda \in K^\times$. Then we have*

  (1) *$u \in W$ or $u \in W^\perp$;*

  (2) *$u \in W^\perp \Leftrightarrow T_u[\lambda]|_W = \mathrm{id}_W$.*

Proof of Theorem 1.1. Let $G \subseteq \mathrm{GSp}(V)$ be a subgroup which contains a nontrivial transvection. If the action of G on V is reducible, we are in case (1) of the theorem. Assume that the action of G on V is irreducible, and define the subgroup $H := \langle \tau \in G : \tau \text{ is a transvection} \rangle$. Note that H is nontrivial. If the action of H on V is irreducible, we can apply Corollary 2.2 to the group H and conclude that H is conjugate in $\mathrm{GSp}(V)$ to $\mathrm{Sp}_n(L)$ for some subfield $L \subseteq K$. This is case (3) of the theorem.

Assume then that the action of H on V is reducible. Let $W \subset V$ be a K-vector subspace on which H acts irreducibly. By Lemma 2.3(a), the group H is a normal subgroup of G. Thus we can apply Clifford's theorem (see [Curtis and Reiner 1981, (11.1)]), to obtain $g_1, \dots, g_r \in G$ such that we have the equality of H-modules

$$V = \bigoplus_{i=1}^{r} g_i W. \tag{2-1}$$

We first remark that W is not the trivial H-module, as otherwise H would act trivially on V and thus H would be the trivial group. Now consider $W' = \langle u \in W : \exists \lambda \in K^\times : T_u[\lambda] \in H \rangle$. As W is a nontrivial H-module, $W' \neq 0$. Let $T_v[\mu] \in H$ and $u \in W'$. By Lemma 2.3(b), $v \in W'$ or $v \in W^\perp$. In both cases, we have $T_v[\mu](u) = u + \mu(u \bullet v)v \in W'$, showing that H preserves W', so that the irreducibility of W implies $W' = W$.

Let $\widetilde{W} = gW$ be a conjugate of W for which we assume $\widetilde{W} \neq W$, so that $\widetilde{W} \cap W = 0$ since W is irreducible. We have just seen that there are $w_1, \dots, w_m \in W$ spanning W and $\lambda_1, \dots, \lambda_m \in K^\times$ such that $T_{w_1}[\lambda_1], \dots, T_{w_m}[\lambda_m] \in H$. As H also preserves $\widetilde{W}$, Lemma 2.3(b) shows $w_i \in \widetilde{W}^\perp$ for $1 \leq i \leq m$. This proves two things. Firstly, $W \subseteq \widetilde{W}^\perp$ and this means that the decomposition (2-1) of V is into mutually orthogonal spaces. From this it follows these subspaces are also symplectic, i.e., that the pairing is nondegenerate on each subspace. Secondly, $T_{w_1}[\lambda_1]$ is the identity on $\widetilde{W}$, but it is nontrivial on W (e.g., by the nondegeneration of W, there is $u \in W$ such that $u \bullet w_1 \neq 0$, whence $T_{w_1}[\lambda_1](u) \neq u$). Hence, W and $\widetilde{W}$ are nonisomorphic as H-modules.

Considering the composite maps $gW \hookrightarrow V \xrightarrow{\text{projection}} g_i W$, in view of the irreducibility of the $g_i W$ and the fact that $g_i W \not\cong g_j W$ for $i \neq j$, it follows that gW is one of the $g_i W$. Thus, G acts on the set $\{g_1 W, \dots, g_r W\}$. If this action were not transitive, then the sum of the spaces in one orbit would be a proper nontrivial G-submodule of V, contradicting the irreducibility of V. Thus, all statements of case (2) of the theorem are proved. □

Proof of Corollary 1.2. Since Γ is compact and the topology on $\overline{\mathbb{F}}_\ell$ is discrete, the image of ρ is a subgroup of $\mathrm{GSp}_n(K)$ for a certain finite field K of characteristic ℓ. Therefore one of the three possibilities of Theorem 1.1 holds for $G := \mathrm{im}(\rho)$. If the first holds, then ρ is reducible, and if the third holds, then $\mathrm{im}(\rho)$ contains a group conjugate to $\mathrm{Sp}_n(L)$ for some subfield L of K.

Assume now that the second possibility holds. We use notation as in Theorem 1.1. Let Γ' be $\{g \in \Gamma \mid \sigma_g(1) = 1\}$, the stabiliser of the first subspace. This is a closed subgroup of Γ of finite index. Choose coset representatives and write $\Gamma = \bigsqcup_{i=1}^{h'} g_i \Gamma'$. The set $\{\gamma S_1 \mid \gamma \in \Gamma\}$ contains h' elements, namely precisely the $g_i S_1$ for $i = 1, \dots, h'$. As the action of G on the decomposition is transitive, this set is precisely $\{S_1, \dots, S_h\}$, whence $h = h'$. Define ρ' as the restriction of ρ to Γ' acting on S_1. Then as a Γ'-representation, we have the isomorphism

$$V \cong \bigoplus_{i=1}^{h} S_i \cong \bigoplus_{i=1}^{h} g_i S_1.$$

Proposition (10.5) of §10A of [Curtis and Reiner 1981] implies $\rho = \mathrm{Ind}_{\Gamma'}^{\Gamma}(\rho')$. □

Proof of Corollary 1.3. Assume that G contains a subgroup conjugate (in $\mathrm{GSp}(V)$) to $\mathrm{Sp}_n(\mathbb{F}_\ell)$. In particular, G does not fix any proper subspace $S \subset V$, nor any decomposition $V = \bigoplus_{i=1}^{h} S_i$ into mutually orthogonal nonsingular symplectic subspaces. Hence by Theorem 1.1, there is a subfield L of K such that the subgroup generated by the symplectic transvections of G is conjugated (in $\mathrm{GSp}(V)$) to $\mathrm{Sp}_n(L)$. The other implication is clear. □

3. Symplectic representations with huge image

In this section we establish Theorem 1.5.

(n, p)-groups. As a generalisation of dihedral groups, in [Khare et al. 2008], Khare, Larsen and Savin introduce so-called (n, p)-groups. We briefly recall some facts and some notation to be used. For the definition of (n, p)-groups, we refer to [loc. cit.]. Let q be a prime number, and let $\mathbb{Q}_{q^n}/\mathbb{Q}_q$ be the unique unramified extension of $\mathbb{Q}_q$ of degree n (inside a fixed algebraic closure $\overline{\mathbb{Q}}_q$). Assume p is a prime such that the order of q modulo p is n. Recall that $\mathbb{Q}_{q^n}^\times \simeq \mu_{q^n-1} \times U_1 \times q^{\mathbb{Z}}$, where μ_{q^n-1} is the group of (q^n-1)-th roots of unity and U_1 the group of 1-units. Let ℓ be a prime distinct from p and q. Assuming that $p, q > n$, in [loc. cit.], the authors construct a character $\chi_q : \mathbb{Q}_{q^n}^\times \to \overline{\mathbb{Q}}_\ell^\times$ that satisfies the three properties of the following lemma, which is proved in [loc. cit., Section 3.1].

Lemma 3.1. *Let $\chi_q : \mathbb{Q}_{q^n}^\times \to \overline{\mathbb{Q}}_\ell^\times$ be a character satisfying:*

- *χ_q has order $2p$.*
- *$\chi_q|_{\mu_{q^n-1} \times U_1}$ has order p.*
- *$\chi_q(q) = -1$.*

This character gives rise to a character (which by abuse of notation we call also χ_q) of $G_{\mathbb{Q}_{q^n}}$ by means of the reciprocity map of local class field theory.

Let

$$\rho_q = \mathrm{Ind}_{G_{\mathbb{Q}_{q^n}}}^{G_{\mathbb{Q}_q}}(\chi_q).$$

Then ρ_q is irreducible and symplectic, in the sense that it can be conjugated to take values in $\mathrm{Sp}_n(\overline{\mathbb{Q}}_\ell)$, and the image of the reduction $\bar\rho_q$ of ρ_q in $\mathrm{Sp}_n(\overline{\mathbb{F}}_\ell)$ is an (n, p)-group. Moreover, if $\bar\alpha : G_{\mathbb{Q}_q} \to \overline{\mathbb{F}}_\ell^\times$ is an unramified character, then $\bar\rho_q \otimes \bar\alpha$ is also irreducible.

Note that also the reduction of ρ_q is $\mathrm{Ind}_{G_{\mathbb{Q}_{q^n}}}^{G_{\mathbb{Q}_q}}(\bar\chi_q)$, which is an irreducible representation. Here $\bar\chi_q$ is the composite of χ_q and the projection $\overline{\mathbb{Z}}_\ell \twoheadrightarrow \overline{\mathbb{F}}_\ell$. To see why the last assertion is true, note that to see that

$$\bar\rho_q \otimes \bar\alpha = \mathrm{Ind}_{G_{\mathbb{Q}_{q^n}}}^{G_{\mathbb{Q}_q}}(\bar\chi_q \otimes (\bar\alpha|_{G_{\mathbb{Q}_{q^n}}}))$$

is irreducible, it suffices to prove that the n characters

$$\bar\chi_q \otimes (\bar\alpha|_{G_{\mathbb{Q}_{q^n}}}), (\bar\chi_q \otimes (\bar\alpha|_{G_{\mathbb{Q}_{q^n}}}))^q, \dots, (\bar\chi_q \otimes (\bar\alpha|_{G_{\mathbb{Q}_{q^n}}}))^{q^{n-1}}$$

are different (see [Serre 1977, Proposition 23, Chapter 7]). But the order of the restriction of $\bar\chi_q \otimes (\bar\alpha|_{G_{\mathbb{Q}_{q^n}}})$ to the inertia group at q is p (since $\bar\alpha$ is unramified), and the order of q mod p is n.

Regular Galois representations. In our result we assume that our representation ρ is regular, which is a condition on the tame inertia weights of ρ.

Definition 3.2 (regularity). Let ℓ be a prime number, n a natural number, V an n-dimensional vector space over $\overline{\mathbb{F}}_\ell$ and $\rho : G_{\mathbb{Q}_\ell} \to \mathrm{GL}(V)$ a Galois representation, and denote by I_ℓ the inertia group at ℓ. We say that ρ is *regular* if there exists an integer s between 1 and n, and for each $i = 1, \dots, s$, a set S_i of natural numbers in $\{0, 1, \dots, \ell-1\}$, of cardinality r_i, with $r_1 + \cdots + r_s = n$, say $S_i = \{a_{i,1}, \dots, a_{i,r_i}\}$, such that the cardinality of $S = S_1 \cup \cdots \cup S_s$ equals n (i.e., all the $a_{i,j}$ are distinct) and such that, if we denote by B_i the matrix

$$B_i \sim \begin{pmatrix} \psi_{r_i}^{b_i} & & & 0 \\ & \psi_{r_i}^{b_i \ell} & & \\ & & \ddots & \\ 0 & & & \psi_{r_i}^{b_i \ell^{r_i - 1}} \end{pmatrix}$$

with ψ_{r_i} our fixed choice of fundamental character of niveau r_i and $b_i = a_{i,1} + a_{i,2}\ell + \cdots + a_{i,r_i}\ell^{r_i - 1}$, then

$$\rho|_{I_\ell} \sim \begin{pmatrix} \boxed{B_1} & & * \\ & \ddots & \\ 0 & & \boxed{B_s} \end{pmatrix}.$$

The elements of S are called *tame inertia weights* of ρ. We will say that ρ has *tame inertia weights at most* k if $S \subseteq \{0, 1, \ldots, k\}$. We will say that a global representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}(V)$ is *regular* if $\rho|_{G_{\mathbb{Q}_\ell}}$ is regular.

Lemma 3.3. *Let $\rho : G_{\mathbb{Q}_\ell} \to \mathrm{GL}_n(\overline{\mathbb{F}}_\ell)$ be a Galois representation which is regular with tame inertia weights at most k. Assume that $\ell > kn! + 1$. Then all the $n!$-th powers of the characters on the diagonal of $\rho|_{I_\ell}$ are distinct.*

Proof. We use the notation of Definition 3.2. Assume we have that the $n!$-th powers of two characters of the diagonal coincide, say

$$\psi_{r_i}^{n!(c_0+c_1\ell+\cdots+c_{r_i-1}\ell^{r_i-1})} = \psi_{r_j}^{n!(d_0+d_1\ell+\cdots+d_{r_j-1}\ell^{r_j-1})}, \tag{3-1}$$

where $c_0, \ldots, c_{r_i-1}, d_0, \ldots, d_{r_j-1}$ are distinct elements of $S_1 \cup \cdots \cup S_s$.

Let $\psi_{r_i r_j}$ be a fundamental character of niveau $r_i r_j$ such that

$$\psi_{r_i r_j}^{\Phi_i^{(i,j)}} = \psi_{r_i} \quad \text{and} \quad \psi_{r_i r_j}^{\Phi_j^{(i,j)}} = \psi_{r_j},$$

where

$$\Phi_i^{(i,j)} = \frac{\ell^{r_i r_j} - 1}{\ell^{r_i} - 1} \quad \text{and} \quad \Phi_j^{(i,j)} = \frac{\ell^{r_i r_j} - 1}{\ell^{r_j} - 1}.$$

We can write (3-1) above as

$$\psi_{r_i r_j}^{\Phi_i^{(i,j)} n!(c_0+c_1\ell+\cdots+c_{r_i-1}\ell^{r_i-1})} = \psi_{r_i r_j}^{\Phi_j^{(i,j)} n!(d_0+d_1\ell+\cdots+d_{r_j-1}\ell^{r_j-1})}.$$

In other words, $\ell^{r_i r_j} - 1$ divides the quantity

$$C_0 = \left|\Phi_i^{(i,j)} n!(c_0+c_1\ell+\cdots+c_{r_i-1}\ell^{r_i-1}) - \Phi_j^{(i,j)} n!(d_0+d_1\ell+\cdots+d_{r_j-1}\ell^{r_j-1})\right|.$$

Note that C_0 is nonzero because modulo ℓ it is congruent to $n!(c_0 - d_0)$, and by assumption all elements in $S_1 \cup \cdots \cup S_s$ are in different congruence classes modulo ℓ. But $|c_0 + c_1\ell + \cdots + c_{r_i-1}\ell^{r_i-1}| \le k(1 + \ell + \cdots + \ell^{r_i-1}) = k(\ell^{r_i} - 1)/(\ell - 1)$. Analogously $|d_0 + d_1\ell + \cdots + d_{r_j-1}\ell^{r_j-1}| < k(\ell^{r_j} - 1)/(\ell - 1)$. Thus C_0 is bounded above by

$$\max\left\{\left|\Phi_i^{(i,j)} n!(c_0+c_1\ell+\cdots+c_{r_i-1}\ell^{r_i-1})\right|, \left|\Phi_j^{(i,j)} n!(d_0+d_1\ell+\cdots+d_{r_j-1}\ell^{r_j-1})\right|\right\}$$
$$\le n!k\left(\frac{\ell^{r_i r_j} - 1}{\ell - 1}\right) < n!k(\ell^{r_i r_j - 1} + 2\ell^{r_i r_j - 2}).$$

Since $\ell - 2 \ge n!k$, we have $\ell^2 - 1 > \ell^2 - 4 \ge n!k(\ell + 2)$ and thus

$$C_0 < n!k(\ell^{r_i r_j - 1} + 2\ell^{r_i r_j - 2}) = n!k(\ell + 2)\ell^{r_i r_j - 2} < \ell^{r_i r_j} - 1.$$

Hence $\ell^{r_i r_j} - 1$ cannot divide C_0. $\square$

We will now use these lemmas to study the ramification at ℓ of an induced representation under the assumption of regularity (possibly after a twist by a power of the cyclotomic character) and boundedness of tame inertia weights.

Proposition 3.4. *Let $n, m, k \in \mathbb{N}$, $a \in \mathbb{Z}$ and let $\ell > kn! + 1$ be a prime, $K/\mathbb{Q}$ a finite extension such that $[K:\mathbb{Q}]\cdot m = n$, $\rho : G_K \to \mathrm{GL}_m(\overline{\mathbb{F}}_\ell)$ a Galois representation and let $\beta = \mathrm{Ind}_{G_K}^{G_\mathbb{Q}} \rho$. If $\chi_\ell^a \otimes \beta$ is regular with tame inertia weights at most k, then $K/\mathbb{Q}$ does not ramify at ℓ.*

Proof. Assume that $K/\mathbb{Q}$ ramifies at ℓ; we will derive a contradiction. First of all, let us fix some notation: let $N/\mathbb{Q}$ be the Galois closure of $K/\mathbb{Q}$, and let us fix a prime λ of N above ℓ. Denote by $I_\ell \subset G_\mathbb{Q}$ the inertia group at ℓ, $I_{\ell,\mathrm{w}} \subset I_\ell$ the wild inertia group at ℓ and $I_N \subset G_N$ the inertia group at the prime λ. Let W be the $\overline{\mathbb{F}}_\ell$-vector space underlying ρ. For each $\gamma \in G_\mathbb{Q}$, let ${}^\gamma K = \gamma(K)$ and define ${}^\gamma\rho : G_{{}^\gamma K} \to \mathrm{GL}(W)$ by ${}^\gamma\rho(\sigma) = \rho(\gamma\sigma\gamma^{-1})$.

Let us now pick any $\gamma \in G_\mathbb{Q}$, $\sigma \in I_\ell$ and $\tau \in I_N$. Since $I_\ell/I_{\ell,\mathrm{w}}$ is cyclic, we have that the commutator $\sigma^{-1}\tau\sigma\tau^{-1}$ belongs to $I_{\ell,\mathrm{w}}$. Since $I_N \subset I_\ell$ is normal, $\sigma^{-1}\tau\sigma \in I_N \subset G_N \subset G_{{}^\gamma K}$, so we may apply ${}^\gamma\rho$ and conclude

$$ {}^\gamma\rho(\sigma^{-1}\tau\sigma)\,{}^\gamma\rho(\tau^{-1}) = {}^\gamma\rho(\sigma^{-1}\tau\sigma\tau^{-1}) \in {}^\gamma\rho(I_{\ell,\mathrm{w}}); $$

hence ${}^\gamma\rho(\sigma^{-1}\tau\sigma)$ and ${}^\gamma\rho(\tau)$ have exactly the same eigenvalues.

Since $N/\mathbb{Q}$ ramifies in ℓ, we may pick $\sigma \in I_\ell \setminus G_N$, and since $N = \prod_{\gamma \in G_\mathbb{Q}} {}^\gamma K$, there exists some $\gamma \in G_\mathbb{Q}$ such that $\sigma \notin G_{{}^\gamma K}$. This implies that $\beta(\sigma\gamma)(W) \cap \beta(\gamma)(W) = 0$. Choose now a set of left-coset representatives $\{\gamma_1 G_K, \dots, \gamma_d G_K\}$ of G_K in $G_\mathbb{Q}$ with $\gamma_1 = \gamma$ and $\gamma_2 = \sigma\gamma$; Mackey's formula [Curtis and Reiner 1981, 10.13] implies that

$$ \mathrm{Res}_{G_N}^{G_\mathbb{Q}} \mathrm{Ind}_{G_K}^{G_\mathbb{Q}} \rho = \bigoplus_{i=1}^{d} \mathrm{Res}_{G_N}^{G_{\gamma_i K}} {}^{\gamma_i}\rho. $$

Therefore $\beta(\tau)$ is a block-diagonal matrix, where one block is ${}^\gamma\rho(\tau)$ and another block is ${}^{\sigma\gamma}\rho(\tau) = {}^\gamma\rho(\sigma^{-1}\tau\sigma)$. But, by hypothesis, the tame inertia weights of $\chi_\ell^a \otimes \beta$ are bounded. By Lemma 3.3, we have that the $n!$-powers of the characters on the diagonal of $\chi_\ell^a \otimes \beta|_{I_\ell}$ are all different, which implies that the characters on the diagonal of $\beta|_{I_N}$ are all different. Thus ${}^\gamma\rho(\tau)$ and ${}^\gamma\rho(\sigma^{-1}\tau\sigma)$ cannot have the same eigenvalues for all $\tau \in I_N$. □

Representations induced in two ways. We need a proposition concerning representations induced from different subgroups of a certain group G.

Proposition 3.5. *Let G be a finite group, $N \trianglelefteq G$, $H \le G$. Assume $(G:N) = n$, and let $p > n$ be a prime. Let K be a field of characteristic coprime to $|G|$ containing all $|G|$-th roots of unity. Let S be a $K[H]$-module, $\chi : N \to K^\times$ a character, say $\chi = \chi_1 \otimes \chi_2$, where $\chi_1 : N \to K^\times$ (resp. $\chi_2 : N \to K^\times$) has order equal to a nontrivial power of p (resp. not divisible by p). Assume*

$$ \rho := \mathrm{Ind}_H^G(S) = \mathrm{Ind}_N^G(\chi), $$

and furthermore the n characters $\{\chi_1^\sigma : \sigma \in G/N\}$ are different. Then $N \le H$.

Following [Serre 1977, 7.2], if G is a finite group and we are given two G-modules V_1 and V_2, we will define $\langle V_1, V_2\rangle_G := \dim \mathrm{Hom}_G(V_1, V_2)$. It is known (Lemma 2 of Chapter 7 of [loc. cit.]) that, if φ_1 and φ_2 are the characters of V_1 and V_2, then

$$\langle V_1, V_2\rangle_G = \langle \varphi_1, \varphi_2\rangle_G := \frac{1}{|G|}\sum_{g\in G}\varphi_1(g^{-1})\varphi_2(g).$$

Before giving the proof, we will first prove a lemma.

Lemma 3.6. *Let G be a group, $N \trianglelefteq G$ and $H \leq G$ such that $(G : H) \leq n$. Let p be a prime such that $p > n$, let K be a field of characteristic coprime to $|G|$ containing all $|G|$-th roots of unity, and let $\chi : N \to K^\times$ be a character whose order is a nontrivial power of p. Then $\mathrm{Res}^N_{H\cap N}\chi$ is not trivial.*

Proof. Assume $\mathrm{Res}^N_{H\cap N}\chi$ is trivial. Then $H \cap N \leq \ker\chi$. But $\ker\chi \leq N$, and the index $(N : \ker\chi)$ is at least p. Therefore $(N : H \cap N) \geq p$. But on the other hand $p > n \geq (G : H) \geq (HN : H) = (N : N \cap H)$, a contradiction. □

Proof of Proposition 3.5. Observe that ρ is irreducible. Namely, there is a well-known criterion characterising when an induced representation is irreducible (see [Serre 1977, Proposition 23, Chapter 7]). In particular, since N is normal in G, we have that $\mathrm{Ind}^G_N\chi$ is irreducible if and only if χ is irreducible (which clearly holds) and, for all $g \in G/N$, $(\mathrm{Res}^G_N(\chi))^h$ is not isomorphic to $\mathrm{Res}^G_N(\chi)$. This last condition holds because the n characters $\{\chi_1^\sigma : \sigma \in G/N\}$ are different, and χ_2 has order prime to p.

Since ρ is irreducible, we have that

$$1 = \langle\rho,\rho\rangle_G = \langle \mathrm{Ind}^G_H(S), \mathrm{Ind}^G_N(\chi)\rangle_G = \langle S, \mathrm{Res}^G_H\mathrm{Ind}^G_N(\chi)\rangle_H = \cdots,$$

where in the last step we used Frobenius reciprocity. Now we apply Mackey's formula [Curtis and Reiner 1981, 10.13] on the right-hand side; note that, since N is normal, $H\backslash G/N \simeq G/(H\cdot N)$:

$$\cdots = \Big\langle S, \bigoplus_{\gamma\in G/(H\cdot N)} \mathrm{Ind}^H_{H\cap N}\mathrm{Res}^N_{H\cap N}(\chi^\gamma)\Big\rangle_H$$
$$= \sum_{\gamma\in G/(H\cdot N)} \Big\langle S, \mathrm{Ind}^H_{H\cap N}\mathrm{Res}^N_{H\cap N}(\chi^\gamma)\Big\rangle_H.$$

Hence there is a unique $\gamma \in G/(H\cdot N)$ such that

$$\langle S, \mathrm{Ind}^H_{H\cap N}\mathrm{Res}^N_{H\cap N}(\chi^\gamma)\rangle_H = 1.$$

If we prove that, for all γ, $\mathrm{Ind}^H_{H\cap N}\mathrm{Res}^N_{H\cap N}(\chi^\gamma)$ is irreducible, then we will have

$$S \simeq \mathrm{Ind}^H_{H\cap N}\mathrm{Res}^N_{H\cap N}(\chi^\gamma)$$

(for some γ); hence $\dim(S) = (H : H \cap N)$. But, on the other hand, since $\rho = \mathrm{Ind}_H^G(S) = \mathrm{Ind}_N^G(\chi)$, we have that $\dim(S) \cdot (G : H) = (G : N)$, so

$$\dim(S) = \frac{(G : HN)(HN : N)}{(G : HN)(HN : H)} = \frac{(H : N \cap H)}{(N : N \cap H)},$$

and therefore the conclusion is that $(N : N \cap H) = 1$; in other words, $N \leq H$.

Therefore to conclude, we only need to see that $\mathrm{Ind}_{H\cap N}^H \mathrm{Res}_{H\cap N}^N(\chi^\gamma)$ is irreducible. Since conjugation by γ plays no role here, let us just assume $\gamma = 1$. We apply again the criterion characterising when an induced representation is irreducible. In particular, since $H \cap N$ is normal in H, we have that $\mathrm{Ind}_{H\cap N}^H \mathrm{Res}_{H\cap N}^N(\chi)$ is irreducible if and only if $\mathrm{Res}_{H\cap N}^N(\chi)$ is irreducible (which clearly holds) and, for all $h \in H/N \cap H$, $(\mathrm{Res}_{H\cap N}^N(\chi))^h$ is not isomorphic to $\mathrm{Res}_{H\cap N}^N(\chi)$.

So pick $h \in H \setminus N$. We have

$$(\mathrm{Res}_{H\cap N}^N(\chi))^h = \mathrm{Res}_{H\cap N}^N(\chi^h).$$

Assume $\mathrm{Res}_{H\cap N}^N(\chi^h) = \mathrm{Res}_{H\cap N}^N(\chi)$. In particular, we obtain $\mathrm{Res}_{H\cap N}^N(\chi_1^h) = \mathrm{Res}_{H\cap N}^N(\chi_1)$. By Lemma 3.6, we have $\chi_1 = \chi_1^h$ as characters of N. But for all $\sigma \in G/N$, we know that $\chi_1^\sigma \neq \chi_1$. Now it suffices to note that $H/(H\cap N) \hookrightarrow G/N$. $\square$

Proofs. Finally we carry out the proof of Theorem 1.5.

Lemma 3.7. *Assume Set-up 1.4. Let $k \in \mathbb{N}$, $\ell \neq p, q$ be a prime such that $\ell > kn!+1$ and $\ell \nmid N$. Let $\chi_q : G_{\mathbb{Q}_{q^n}} \to \overline{\mathbb{Q}}_\ell^\times$ be a character satisfying the assumptions of Lemma 3.1, and $\bar{\chi}_q$ the composition of χ_q with the reduction map $\overline{\mathbb{Z}}_\ell \to \overline{\mathbb{F}}_\ell$. Let $\bar{\alpha} : G_{\mathbb{Q}_q} \to \overline{\mathbb{F}}_\ell^\times$ be an unramified character.*

Let $\rho : G_\mathbb{Q} \to \mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$ be a Galois representation, ramified only at the primes dividing $Nq\ell$, such that a twist by some power of the cyclotomic character is regular in the sense of Definition 3.2 with tame inertia weights at most k, and satisfying (1) *and* (3) *of Theorem 1.5. Then ρ is not induced from a representation of an open subgroup $H \subsetneq G_\mathbb{Q}$.*

Proof. Let $H \subset G_\mathbb{Q}$ be an open subgroup, say of index h, and $\rho' : H \to \mathrm{GL}_{n/h}(\overline{\mathbb{F}}_\ell)$ a representation such that

$$\rho \cong \mathrm{Ind}_H^{G_\mathbb{Q}}(\rho').$$

Call S_1 and V, with $S_1 \subseteq V$, the spaces underlying ρ' and ρ, respectively, so that $\rho = \mathrm{Ind}_H^{G_\mathbb{Q}}(S_1)$. Recall that by assumption

$$\mathrm{Res}_{G_{\mathbb{Q}_q}}^{G_\mathbb{Q}}(\rho) = \mathrm{Ind}_{G_{\mathbb{Q}_{q^n}}}^{G_{\mathbb{Q}_q}}(\bar{\chi}_q) \otimes \bar{\alpha}.$$

We want to compute $\mathrm{Res}_{G_{\mathbb{Q}_q}}^{G_\mathbb{Q}} \mathrm{Ind}_H^{G_\mathbb{Q}}(S_1)$. Let us apply Mackey's formula [Curtis and Reiner 1981, 10.13]. By Lemma 3.1 we know that

$$\mathrm{Res}_{G_{\mathbb{Q}_q}}^{G_\mathbb{Q}} \mathrm{Ind}_H^{G_\mathbb{Q}}(S_1) = \mathrm{Ind}_{G_{\mathbb{Q}_{q^n}}}^{G_{\mathbb{Q}_q}}(\bar{\chi}_q) \otimes \bar{\alpha}$$

is irreducible, so there can only be one summand in the formula; hence

$$\mathrm{Res}^{G_{\mathbb{Q}}}_{G_{\mathbb{Q}_q}} \mathrm{Ind}^{G_{\mathbb{Q}}}_{H}(S_1) = \mathrm{Ind}^{G_{\mathbb{Q}_q}}_{G_{\mathbb{Q}_q}\cap H} \mathrm{Res}^{H}_{G_{\mathbb{Q}_q}\cap H}(S_1),$$

and therefore

$$\mathrm{Ind}^{G_{\mathbb{Q}_q}}_{G_{\mathbb{Q}_q}\cap H} \mathrm{Res}^{H}_{G_{\mathbb{Q}_q}\cap H}(S_1) = \mathrm{Ind}^{G_{\mathbb{Q}_q}}_{G_{\mathbb{Q}_{q^n}}}(\bar{\chi}_q) \otimes \bar{\alpha}. \tag{3-2}$$

We now apply Proposition 3.5 to (3-2). Note that

$$\mathrm{Res}^{G_{\mathbb{Q}}}_{G_{\mathbb{Q}_q}} \rho = \mathrm{Ind}^{G_{\mathbb{Q}_q}}_{G_{\mathbb{Q}_{q^n}}}(\bar{\chi}_q) \otimes \bar{\alpha} = \mathrm{Ind}^{G_{\mathbb{Q}_q}}_{G_{\mathbb{Q}_{q^n}}}(\bar{\chi}_q \otimes (\bar{\alpha}|_{G_{\mathbb{Q}_{q^n}}})).$$

We can write $\bar{\chi}_q \otimes (\bar{\alpha}|_{G_{\mathbb{Q}_{q^n}}}) = \bar{\chi}_1 \otimes \bar{\chi}_2$, where $\bar{\chi}_1$ has order a power of p and $\bar{\chi}_2$ has order prime to p. Note that the restriction of $\bar{\chi}_q \otimes (\bar{\alpha}|_{G_{\mathbb{Q}_{q^n}}})$ to the inertia group I_q of $G_{\mathbb{Q}_q}$ coincides with the restriction of $\bar{\chi}_q$, which has order p. Thus $(\bar{\chi}_1 \otimes \bar{\chi}_2)|_{I_q} = \bar{\chi}_q|_{I_q} = \bar{\chi}_1|_{I_q}$. Since the order of q mod p is n, we know that the n characters $\bar{\chi}_1|_{I_q}, \bar{\chi}_1^q|_{I_q}, \dots, \bar{\chi}_1^{q^n}|_{I_q}$ are distinct. We can take $G = \rho(G_{\mathbb{Q}_q})$ in the statement of Proposition 3.5, whose order is a divisor of $2np \cdot \mathrm{ord}(\bar{\alpha})$ and, hence, prime to ℓ. It thus follows that $G_{\mathbb{Q}_{q^n}} \le (G_{\mathbb{Q}_q} \cap H)$.

Note that, on the one hand

$$n = \dim V = \dim(\mathrm{Ind}^{G_{\mathbb{Q}}}_{H} S_1) = (G_{\mathbb{Q}} : H)\dim(S_1).$$

On the other hand,

$$n = \dim(\mathrm{Ind}^{G_{\mathbb{Q}_q}}_{G_{\mathbb{Q}_q}\cap H} \mathrm{Res}^{H}_{G_{\mathbb{Q}_q}\cap H}(S_1)) = (G_{\mathbb{Q}_q} : G_{\mathbb{Q}_q} \cap H)\dim(S_1);$$

hence $(G_{\mathbb{Q}} : H) = (G_{\mathbb{Q}_q} : G_{\mathbb{Q}_q} \cap H)$.

Let L be the number field such that $H = \mathrm{Gal}(\overline{\mathbb{Q}}/L)$. Now $\mathrm{Gal}(\overline{\mathbb{Q}}/L) \cap G_{\mathbb{Q}_q} = \mathrm{Gal}(\overline{\mathbb{Q}}_q/L_{\mathfrak{q}})$, where $\mathfrak{q}$ is a certain prime of L above q and $L_{\mathfrak{q}}$ denotes the completion of L at $\mathfrak{q}$. The inclusion $G_{\mathbb{Q}_{q^n}} \le \mathrm{Gal}(\overline{\mathbb{Q}}_q/L_{\mathfrak{q}})$ means that we have the field inclusions

$$\mathbb{Q}_q \subseteq L_{\mathfrak{q}} \subseteq \mathbb{Q}_{q^n} \subseteq \overline{\mathbb{Q}}_q$$

and $[L_{\mathfrak{q}} : \mathbb{Q}_q] = (G_{\mathbb{Q}_q} : G_{\mathbb{Q}_q} \cap H) = (G_{\mathbb{Q}} : H) = [L : \mathbb{Q}]$; hence q is inert in $L/\mathbb{Q}$.

Let ℓ_1 be a prime dividing N_1, let $\widetilde{L}/\mathbb{Q}$ be a Galois closure of $L/\mathbb{Q}$, Λ_1 a prime of $\widetilde{L}$ above ℓ_1 and I_1 the inertia group of Λ_1 over $\mathbb{Q}$. Since $\gcd(|\rho(I_{\ell_1})|, n!) = 1$ and $\mathrm{Gal}(\widetilde{L}/\mathbb{Q})$ has order dividing $n!$, we get that the projection of $\rho(I_1) \subseteq \rho(I_{\ell_1})$ into $\rho(G_{\mathbb{Q}})/\rho(G_{\widetilde{L}})$ is trivial. Thus, $\rho(I_1) \subseteq \rho(G_{\widetilde{L}})$. Hence $\widetilde{L}/\mathbb{Q}$ is unramified at ℓ_1 and so is $L/\mathbb{Q}$.

To sum up, we know that L can only be ramified at the primes dividing $Nq\ell$. But L cannot ramify at q since $L_{\mathfrak{q}} \subseteq \mathbb{Q}_{q^n}$ (and $\mathbb{Q}_{q^n}$ is an unramified extension of $\mathbb{Q}_q$). We just saw that L cannot ramify at the primes dividing N_1. We also know that L cannot be ramified at ℓ (see Proposition 3.4). Hence L only ramifies at the primes dividing N_2. By the choice of q, it is completely split in L, and at the same time inert in L. This shows $L = \mathbb{Q}$ and $H = G_{\mathbb{Q}}$. ☐

Now we can easily prove the main group theoretic result.

Proof of Theorem 1.5. Let $G = \mathrm{Im}\rho$. Since G contains a transvection, one of the following three possibilities holds (cf. Corollary 1.2):

(1) ρ is reducible.

(2) There exists an open subgroup $H \subsetneq G_{\mathbb{Q}}$, say of index h with n/h even, and a representation $\rho' : H \to \mathrm{GSp}_{n/h}(\overline{\mathbb{F}}_\ell)$ such that $\rho \cong \mathrm{Ind}_H^{G_{\mathbb{Q}}} \rho'$.

(3) The group generated by the transvections in G is conjugated (in $\mathrm{GSp}_n(\overline{\mathbb{F}}_\ell)$) to $\mathrm{Sp}_n(\mathbb{F}_{\ell^r})$ for some exponent r.

By Lemma 3.1, G acts irreducibly on V; hence the first possibility cannot occur. By Lemma 3.7, the second possibility does not occur. Hence the third possibility holds, and this finishes the proof of the theorem. □

Proof of Corollary 1.6. This follows from the main theorem of Part I [Arias-de-Reyna et al. 2013] concerning the application to the inverse Galois problem. In order to be able to apply it, there are two things to check: Firstly, we note that $\rho_\bullet$ is maximally induced of order p at the prime q. Secondly, the existence of a transvection in the image of $\bar{\rho}_\lambda$ together with the special shape of the representation at q allow us to conclude from Theorem 1.5 that the image of $\bar{\rho}_\lambda$ is huge for all $\lambda \mid \ell$, where ℓ runs through the rational primes outside a density zero set. □

Acknowledgements

Arias-de-Reyna worked on this article as a fellow of the Alexander von Humboldt Foundation. She thanks the Université du Luxembourg for its hospitality during a long term visit in 2011. She was also partially supported by the project MTM2012-33830 of the Ministerio de Economía y Competitividad of Spain. Dieulefait was supported by the project MTM2012-33830 of the Ministerio de Economía y Competitividad of Spain and by an ICREA Academia Research Prize. Wiese was partially supported by the DFG collaborative research centre TRR 45, the DFG priority program 1489 and the Fonds National de la Recherche Luxembourg (INTER/DFG/12/10/COMFGREP). Arias-de-Reyna and Wiese thank the Centre de Recerca Matemàtica for its support and hospitality during a long term visit in 2010.

The authors thank the anonymous referee for his suggestions on Section 2. Arias-de-Reyna and Wiese thank Gunter Malle for his detailed explanations concerning [Kantor 1979].

References

[Arias-de-Reyna et al. 2013] S. Arias-de-Reyna, L. V. Dieulefait, and G. Wiese, "Compatible systems of symplectic Galois representations and the inverse Galois problem, I: Images of projective representations", preprint, 2013. To appear in *Trans. Amer. Math. Soc.* arXiv 1203.6546

[Arias-de-Reyna et al. 2015] S. Arias-de-Reyna, L. V. Dieulefait, S. W. Shin, and G. Wiese, "Compatible systems of symplectic Galois representations and the inverse Galois problem, III: Automorphic construction of compatible systems with suitable local properties", *Math. Ann.* **361**:3-4 (2015), 909–925. MR 3319552 Zbl 06421505

[Artin 1957] E. Artin, *Geometric algebra*, Interscience, New York, 1957. MR 18,553e Zbl 0077.02101

[Curtis and Reiner 1981] C. W. Curtis and I. Reiner, *Methods of representation theory, I: With applications to finite groups and orders*, Wiley, New York, 1981. MR 82i:20001 Zbl 0469.20001

[Dieulefait and Wiese 2011] L. V. Dieulefait and G. Wiese, "On modular forms and the inverse Galois problem", *Trans. Amer. Math. Soc.* **363**:9 (2011), 4569–4584. MR 2012k:11069 Zbl 1264.11045

[Guralnick and Saxl 2003] R. M. Guralnick and J. Saxl, "Generation of finite almost simple groups by conjugates", *J. Algebra* **268**:2 (2003), 519–571. MR 2005f:20057 Zbl 1037.20016

[Kantor 1979] W. M. Kantor, "Subgroups of classical groups generated by long root elements", *Trans. Amer. Math. Soc.* **248**:2 (1979), 347–379. MR 80g:20057 Zbl 0406.20040

[Khare et al. 2008] C. Khare, M. Larsen, and G. Savin, "Functoriality and the inverse Galois problem", *Compos. Math.* **144**:3 (2008), 541–564. MR 2009m:11076 Zbl 1194.11062

[Serre 1977] J.-P. Serre, *Linear representations of finite groups*, Graduate Texts in Mathematics **42**, Springer, New York, 1977. MR 56 #8675 Zbl 0355.20006

[Serre 1979] J.-P. Serre, *Local fields*, Graduate Texts in Mathematics **67**, Springer, New York, 1979. MR 82e:12016 Zbl 0423.12016

Received July 16, 2015.

SARA ARIAS-DE-REYNA
FACULTÉ DES SCIENCES, DE LA TECHNOLOGIE ET DE LA COMMUNICATION
UNIVERSITY OF LUXEMBOURG
6, RUE RICHARD COUDENHOVE-KALERGI
L-1359 LUXEMBOURG
LUXEMBOURG

sara.ariasdereyna@uni.lu

LUIS DIEULEFAIT
DEPARTAMENT D'ÁLGEBRA I GEOMETRIA, FACULTAT DE MATEMATIQUES
UNIVERSITAT DE BARCELONA
GRAN VIA DE LES CORTS CATALANES, 585
08007 BARCELONA
SPAIN

ldieulefait@ub.edu

GABOR WIESE
FACULTÉ DES SCIENCES, DE LA TECHNOLOGIE ET DE LA COMMUNICATION
UNIVERSITY OF LUXEMBOURG
6, RUE RICHARD COUDENHOVE-KALERGI
L-1359 LUXEMBOURG
LUXEMBOURG

gabor.wiese@uni.lu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.17

ON THE NUMBER OF LINES IN THE LIMIT SET FOR DISCRETE SUBGROUPS OF PSL(3, $\mathbb{C}$)

WALDEMAR BARRERA, ANGEL CANO AND JUÁN NAVARRETE

Given a discrete subgroup $G \subset \mathrm{PSL}(3, \mathbb{C})$, acting on the complex projective plane, $\mathbb{P}^2_{\mathbb{C}}$, in the canonical way, we list all possible values for the number of complex projective lines and for the maximum number of complex projective lines lying in the complement of each of: the equicontinuity set of G, the Kulkarni discontinuity region of G, and maximal open subsets of $\mathbb{P}^2_{\mathbb{C}}$ on which G acts properly discontinuously.

1. Introduction

A classical result in the theory of Kleinian groups states that the limit set of an infinite Kleinian group consists of one, two, or uncountably many points. If the number of points in the limit set is smaller or equal to two then the Kleinian group is called *elementary*. On the other hand, if the number of points in the limit set is greater than two then the group is called *nonelementary* and its limit set is a perfect set.

In this paper, we prove an analogous result for complex Kleinian groups acting on $\mathbb{P}^2_{\mathbb{C}}$. We recall that $G \subset \mathrm{PSL}(3, \mathbb{C})$ is a complex Kleinian group, whenever there exists a G-invariant nonempty open set $U \subset \mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously.

There is no standard definition of limit set in the theory of complex Kleinian groups, and we use the following three notions of limit set for a complex Kleinian group: The Myrberg limit set $\mathbb{P}^2_{\mathbb{C}} \setminus \mathrm{Eq}(G)$ (see Section 3A), the Kulkarni limit set $\mathbb{P}^2_{\mathbb{C}} \setminus \Omega(G)$ (see Section 3B), and the complement of a maximal G-invariant open subset of $\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously. In what follows, we denote by $U_{\max}(G)$ any maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously.

The main results in this paper are the following:

Research supported by CONACYT Project Number 176680, SEP grant P/PIFI-2011-31MSU0098J-15, project PAPIIT IN106814.
MSC2010: primary 32Q45, 37F30; secondary 37F45, 22E40.
Keywords: Kleinian groups, projective complex plane, discrete groups, limit set.

Theorem 1.1. *If $G \subset \mathrm{PSL}(3, \mathbb{C})$ is an infinite discrete subgroup and U is equal to one of* $\mathrm{Eq}(G)$, $\Omega(G)$, *or* $U_{\max}(G)$, *then the number of complex projective lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is equal to* 1, 2, 3, *or* ∞. *Moreover, if there are infinitely many complex projective lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$, then there exists a perfect set of complex projective lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.*

Theorem 1.2. *If $G \subset \mathrm{PSL}(3, \mathbb{C})$ is an infinite discrete subgroup and U is equal to one of* $\mathrm{Eq}(G)$, $\Omega(G)$, *or* $U_{\max}(G)$, *then the maximum number of complex projective lines in general position contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is equal to* 1, 2, 3, 4, *or* ∞.

We begin our exposition with a brief section on projective geometry. The material in this section is standard. Also, we set the notation we use throughout the paper.

In Section 3, we recall the definitions of equicontinuity set, Myrberg limit set, Kulkarni limit set, and Kulkarni discontinuity region. Also, we include some useful results such as Theorem 3.4 and Proposition 3.6. Finally, we recall the definition of complex Kleinian group.

In Section 4, we use Segre's embedding to prove that the set of effective lines is closed in $(\mathbb{P}^2_{\mathbb{C}})^*$. Consequently, the union of all effective lines for a discrete group $G \subset \mathrm{PSL}(3, \mathbb{C})$ is a closed set of $\mathbb{P}^2_{\mathbb{C}}$ and this union is equal to the complement of the equicontinuity set of G, except in one case; see Corollary 4.5. The existence of loxodromic elements, whenever the limit set contains at least three lines in general position, is proved in Proposition 4.10.

In Section 5, we include all results needed to prove the main Theorem 1.1. In order to give a sketch of the proof of this theorem, and for the reader's convenience, we use the notation $\lambda(U)$ and $\mu(U)$ to denote the number of complex projective lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ and the maximum number of complex projective lines in general position contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$, respectively. A sketch of the proof of Theorem 1.1 is as follows:

Since G is an infinite group, $\mu(U) \geq 1$ (see the proof of [Cano et al. 2013, Proposition 3.3.4]).

If $\mu(U) \leq 3$ and $\lambda(U) < \infty$, then $\lambda(U) = \mu(U)$ (see Propositions 5.4 and 5.6).

If $1 < \mu(U) \leq 3$ and $\lambda(U) = \infty$, then there exists a perfect set of lines contained in the complement of U (see Proposition 5.7).

If $\mu(U) \geq 4$, then the complement of U is the union of a perfect set of lines (see Proposition 5.15).

In Section 6, we prove Theorem 1.2. The sketch of the proof is the following: We assume that $4 < \mu(U) < \infty$ and we find precisely two points called *vertices* such that each one of these points lies in infinitely many lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ (see Proposition 6.5). Moreover, if the line ℓ is contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ and does not pass through one of these vertices then its orbit contains infinitely many lines. Since $\mu(U) < \infty$, we obtain another vertex, contradicting Proposition 6.5.

The last section contains examples showing all distinct possible values that $\lambda(U)$ and $\mu(U)$ can take.

2. Preliminaries and notation

We recall that the complex projective plane $\mathbb{P}^2_{\mathbb{C}}$ is defined as

$$\mathbb{P}^2_{\mathbb{C}} := (\mathbb{C}^3 \setminus \{\mathbf{0}\}) / \mathbb{C}^*,$$

where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ acts on $\mathbb{C}^3 \setminus \{\mathbf{0}\}$ by the usual scalar multiplication. This is a compact connected complex 2-dimensional manifold. Let $[\,] : \mathbb{C}^3 \setminus \{\mathbf{0}\} \to \mathbb{P}^2_{\mathbb{C}}$ be the quotient map. If $\beta = \{e_1, e_2, e_3\}$ is the standard basis of $\mathbb{C}^3$, we write $[e_j] = e_j$, for $j = 1, 2, 3$, and if $\boldsymbol{z} = (z_1, z_2, z_3) \in \mathbb{C}^3 \setminus \{\mathbf{0}\}$ then we write $[\boldsymbol{z}] = [z_1 : z_2 : z_3]$. Also, $\ell \subset \mathbb{P}^2_{\mathbb{C}}$ is said to be a complex line if $[\ell]^{-1} \cup \{\mathbf{0}\}$ is a complex linear subspace of dimension 2. Given two distinct points $[\boldsymbol{z}], [\boldsymbol{w}] \in \mathbb{P}^2_{\mathbb{C}}$, there is a unique complex projective line passing through $[\boldsymbol{z}]$ and $[\boldsymbol{w}]$; such a complex projective line is called a *line*, for short, and it is denoted by $\overleftrightarrow{[\boldsymbol{z}], [\boldsymbol{w}]}$. Consider the action of $\mathbb{C}^*$ on $\mathrm{GL}(3, \mathbb{C})$ given by the usual scalar multiplication. Then

$$\mathrm{PGL}(3, \mathbb{C}) = \mathrm{GL}(3, \mathbb{C}) / \mathbb{C}^*$$

is a Lie group whose elements are called projective transformations. Now let $[\![\,]\!] : \mathrm{GL}(3, \mathbb{C}) \to \mathrm{PGL}(3, \mathbb{C})$ be the quotient map, $g \in \mathrm{PGL}(3, \mathbb{C})$ and $\boldsymbol{g} \in \mathrm{GL}(3, \mathbb{C})$, we say that $\boldsymbol{g}$ is a lift of g if $[\![\boldsymbol{g}]\!] = g$. One can show that $\mathrm{PGL}(3, \mathbb{C})$ is a Lie group which acts transitively, effectively, and by biholomorphisms on $\mathbb{P}^2_{\mathbb{C}}$ via $[\![\boldsymbol{g}]\!]([\boldsymbol{w}]) = [\boldsymbol{g}(\boldsymbol{w})]$, where $\boldsymbol{w} \in \mathbb{C}^3 \setminus \{\mathbf{0}\}$ and $\boldsymbol{g} \in \mathrm{GL}(3, \mathbb{C})$.

We could have considered the action of the cube roots of unity $\{1, \omega, \omega^2\} \subset \mathbb{C}^*$ on $\mathrm{SL}(3, \mathbb{C})$ given by the usual scalar multiplication, in which case

$$\mathrm{PSL}(3, \mathbb{C}) = \mathrm{SL}(3, \mathbb{C}) / \{1, \omega, \omega^2\} \cong \mathrm{PGL}(3, \mathbb{C}).$$

We denote by $\mathrm{M}_{3\times 3}(\mathbb{C})$ the space of all 3×3 matrices with entries in $\mathbb{C}$ equipped with the standard topology. The quotient space

$$\mathrm{SP}(3, \mathbb{C}) := (\mathrm{M}_{3\times 3}(\mathbb{C}) \setminus \{\mathbf{0}\}) / \mathbb{C}^*$$

is called the space of *pseudoprojective maps of* $\mathbb{P}^2_{\mathbb{C}}$ and it is naturally identified with the projective space $\mathbb{P}^8_{\mathbb{C}}$. Since $\mathrm{GL}(3, \mathbb{C})$ is an open, dense, $\mathbb{C}^*$-invariant set of $\mathrm{M}_{3\times 3}(\mathbb{C}) \setminus \{\mathbf{0}\}$, we obtain that the space of pseudoprojective maps of $\mathbb{P}^2_{\mathbb{C}}$ is a compactification of $\mathrm{PGL}(3, \mathbb{C})$ (or $\mathrm{PSL}(3, \mathbb{C})$). As in the case of projective maps, if $\boldsymbol{s}$ is an element in $\mathrm{M}_{3\times 3}(\mathbb{C}) \setminus \{\mathbf{0}\}$, then $[\boldsymbol{s}]$ denotes the equivalence class of the matrix $\boldsymbol{s}$ in the space of pseudoprojective maps of $\mathbb{P}^2_{\mathbb{C}}$. Also, we say that $\boldsymbol{s} \in \mathrm{M}_{3\times 3}(\mathbb{C}) \setminus \{\mathbf{0}\}$ is a lift of the pseudoprojective map S whenever $[\boldsymbol{s}] = S$.

Let S be an element in $(\mathrm{M}_{3\times 3}(\mathbb{C})\setminus\{\mathbf{0}\})/\mathbb{C}^*$ and $\boldsymbol{s}$ a lift to $\mathrm{M}_{3\times 3}(\mathbb{C})\setminus\{\mathbf{0}\}$ of S. The matrix $\boldsymbol{s}$ induces a nonzero linear transformation $s:\mathbb{C}^3\to\mathbb{C}^3$, which is not necessarily invertible. Let $\mathrm{Ker}(s)\subsetneq\mathbb{C}^3$ be its kernel and let $\mathrm{Ker}(S)$ denote its projectivization to $\mathbb{P}^2_{\mathbb{C}}$, taking into account that $\mathrm{Ker}(S):=\varnothing$ whenever $\mathrm{Ker}(s)=\{(0,0,0)\}$.

3. Discontinuous actions on $\mathbb{P}^2_{\mathbb{C}}$

Definition 3.1. Let $G\subset\mathrm{PSL}(3,\mathbb{C})$ be a discrete group. We say that G acts *properly and discontinuously* on the open nonempty G-invariant set $U\subset\mathbb{P}^2_{\mathbb{C}}$ if and only if, for each pair of compact subsets $C, D\subset U$, the set

$$\{g\in G : g(C)\cap D\neq\varnothing\}$$

is *finite*.

3A. *The equicontinuity set.*

Definition 3.2. The *equicontinuity set* for a family $\mathcal{F}$ of endomorphisms of $\mathbb{P}^2_{\mathbb{C}}$, denoted $\mathrm{Eq}(\mathcal{F})$ is defined as the set of points $z\in\mathbb{P}^2_{\mathbb{C}}$ for which there is an open neighborhood U of z such that $\{f|_U : f\in\mathcal{F}\}$ is a normal family.

Definition 3.3. Let $G\subset\mathrm{PSL}(3,\mathbb{C})$ be a discrete group. If

$$G'=\{S \text{ is a pseudoprojective map of } \mathbb{P}^2_{\mathbb{C}} : S \text{ is a cluster point of } G\};$$

then the *Myrberg limit set* [1925] is defined as the set

$$\Lambda_{\mathrm{Myr}}(G)=\bigcup_{S\in G'}\mathrm{Ker}(S).$$

Myrberg [1925] shows that G acts properly and discontinuously on $\mathbb{P}^2_{\mathbb{C}}\setminus\Lambda_{\mathrm{Myr}}(G)$.

Theorem 3.4 [Barrera et al. 2011a]. *If* $G\subset\mathrm{PSL}(3,\mathbb{C})$ *is a discrete group, then*:

(i) *The group G acts properly and discontinuously on* $\mathrm{Eq}(G)$.

(ii) *The equicontinuity set of G satisfies*:

$$\mathrm{Eq}(G)=\mathbb{P}^2_{\mathbb{C}}\setminus\Lambda_{\mathrm{Myr}}(G)$$

(iii) *If U is an open G-invariant subset such that $\mathbb{P}^2_{\mathbb{C}}\setminus U$ contains at least three complex lines in general position, then* $U\subset\mathrm{Eq}(G)$.

3B. *The Kulkarni discontinuity region.*

Definition 3.5 [Kulkarni 1978]. If $G\subset\mathrm{PSL}(3,\mathbb{C})$ is a group, then:

- The set $L_0(G)$ is the closure of the set of points in $\mathbb{P}^2_{\mathbb{C}}$ with infinite isotropy group.
- The set $L_1(G)$ is the closure of the set of cluster points of the orbit Gz, where z runs over $\mathbb{P}^2_{\mathbb{C}}\setminus L_0(G)$.

- The set $L_2(G)$ is the closure of the set of cluster points of the family of compact sets $\{g(K) : g \in G\}$, where K runs over all the compact subsets of $\mathbb{P}^2_{\mathbb{C}} \setminus (L_0(G) \cup L_1(G))$.

The *Kulkarni limit set* of G is defined as

$$\Lambda_{\text{Kul}}(G) = L_0(G) \cup L_1(G) \cup L_2(G).$$

The *Kulkarni discontinuity region* of G is defined as

$$\Omega(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \Lambda_{\text{Kul}}(G).$$

Kulkarni [1978] proves that G acts properly and discontinuously on the set $\Omega(G)$. However, $\Omega(G)$ is not necessarily the maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously. It is proved in [Barrera et al. 2011a] that $\text{Eq}(G) \subset \Omega(G)$ whenever $G \subset \text{PSL}(3, \mathbb{C})$ is discrete.

Proposition 3.6. *If H is a finite index subgroup of $G \subset \text{PSL}(3, \mathbb{C})$, then*

(i) $L_0(H) = L_0(G)$,

(ii) $L_1(H) = L_1(G)$,

(iii) $L_2(H) = L_2(G)$,

(iv) $\Lambda_{\text{Kul}}(H) = \Lambda_{\text{Kul}}(G)$ *and* $\Omega(H) = \Omega(G)$.

Proof. Let us assume that $m = [G : H]$ and

$$G = \bigcup_{i=1}^{m} H\gamma_i.$$

(i) It is not hard to see that $L_0(H) \subset L_0(G)$. Now, if $x \in \mathbb{P}^2_{\mathbb{C}}$ and $|\text{Isot}(x, G)| = \infty$, then there exists a sequence of distinct elements $(g_n) \subset G$ such that

$$g_n(x) = x \quad \text{for all } n \in \mathbb{N}.$$

We can assume there exists $1 \le i_0 \le m$ such that

$$g_n = h_n\gamma_{i_0} \quad \text{for all } n \in \mathbb{N}.$$

Hence, $(\tilde{h}_n) \subset H$, where $\tilde{h}_n = h_n h_1^{-1}$, is a sequence of distinct elements in H such that $\tilde{h}_n(x) = x$ for all $n \in \mathbb{N}$. Therefore $x \in L_0(H)$.

(ii) It is not hard to check that $L_1(H) \subset L_1(G)$. Conversely, if $(g_n) \subset G$ is a sequence of distinct elements and $x \in \mathbb{P}^2_{\mathbb{C}} \setminus L_0(G) = \mathbb{P}^2_{\mathbb{C}} \setminus L_0(H)$, such that

$$g_n(x) \to z \quad \text{as } n \to \infty,$$

then we can assume that

$$g_n(x) = h_n(\gamma_{i_0}(x)) \to z \quad \text{as } n \to \infty,$$

where $\gamma_{i_0}(x) \in \mathbb{P}^2_{\mathbb{C}} \setminus L_0(G) = \mathbb{P}^2_{\mathbb{C}} \setminus L_0(H)$. It follows that $z \in L_1(H)$.

(iii) It is not hard to check that $L_2(H) \subset L_2(G)$. Conversely, let us assume z is a cluster point of the family

$$\{g_n(K) : n \in \mathbb{N}\},$$

where $(g_n) \subset G$ is a sequence of distinct elements and $K \subset \mathbb{P}^2_{\mathbb{C}} \setminus (L_0(G) \cup L_1(G)) = \mathbb{P}^2_{\mathbb{C}} \setminus (L_0(H) \cup L_1(H))$ is a compact set. We can assume there exists $1 \le i_0 \le m$ such that

$$g_n = h_n \gamma_{i_0} \quad \text{for all } n \in \mathbb{N}.$$

It follows that z is a cluster point of the family

$$\{h_n(\gamma_{i_0}(K)) : n \in \mathbb{N}\},$$

where $(h_n) \subset H$ is a sequence of distinct elements and $\gamma_{i_0}(K) \subset \mathbb{P}^2_{\mathbb{C}} \setminus (L_0(G) \cup L_1(G)) = \mathbb{P}^2_{\mathbb{C}} \setminus (L_0(H) \cup L_1(H))$ is a compact set. Therefore $z \in L_2(H)$.

(iv) It follows from (i), (ii), and (iii). □

Definition 3.7 [Cano et al. 2013]. We say that $G \subset \mathrm{PSL}(3, \mathbb{C})$ is a *complex Kleinian group* if there exists a G-invariant nonempty open subset of $\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously.

4. Some useful results

Definition 4.1. We say ℓ is an *effective line* for the discrete group $G \subset \mathrm{PSL}(3, \mathbb{C})$ if there exists a pseudoprojective transformation $S \in G'$ such that $\ell = \mathrm{Ker}(S)$. The set of effective lines for G is denoted by $\mathcal{E}(G)$, or simply $\mathcal{E}$ when there is no danger of confusion.

Proposition 4.2. *If $G \subset \mathrm{PSL}(3, \mathbb{C})$ is a discrete group then $\mathcal{E}$ is a closed subset of $(\mathbb{P}^2_{\mathbb{C}})^*$, where $(\mathbb{P}^2_{\mathbb{C}})^*$ denotes the space of complex projective lines in $\mathbb{P}^2_{\mathbb{C}}$.*

Proof. We assume that (ℓ_n) is a sequence in $\mathcal{E}$ such that $\ell_n \to \ell$ as $n \to \infty$. For each $n \in \mathbb{N}$, there exists $S_n \in G' \subset \mathrm{SP}(3, \mathbb{C})$ such that $\ell_n = \mathrm{Ker}(S_n)$. Since $\mathrm{SP}(3, \mathbb{C})$ is compact, we can assume that $S_n \to S \in \mathrm{SP}(3, \mathbb{C})$ as $n \to \infty$. Moreover, $S \in G'$ because G' is closed.

In order to prove that $\ell = \mathrm{Ker}(S)$ we use the Segre embedding:

$$\psi : \mathbb{P}^2_{\mathbb{C}} \times \mathbb{P}^2_{\mathbb{C}} \to \mathrm{SP}(3, \mathbb{C})$$

$$\psi([\boldsymbol{v}], [\boldsymbol{t}]) = \left[\begin{pmatrix} t_1 \boldsymbol{v} \\ t_2 \boldsymbol{v} \\ t_3 \boldsymbol{v} \end{pmatrix}\right] = \left[\begin{pmatrix} v_1 \boldsymbol{t} & v_2 \boldsymbol{t} & v_3 \boldsymbol{t} \end{pmatrix}\right],$$

where

$$\boldsymbol{v} = \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{t} = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}.$$

We notice that the image of ψ is precisely the set of pseudoprojective transformations in SP(3, $\mathbb{C}$) whose kernel is equal to one line. In fact, $[\boldsymbol{v}]$ can be identified with $\mathrm{Ker}(\psi([\boldsymbol{v}],[\boldsymbol{t}]))$. Since ψ is continuous, it follows that $\psi(\mathbb{P}^2_{\mathbb{C}}\times\mathbb{P}^2_{\mathbb{C}})$ is compact in SP(3, $\mathbb{C}$), so it is closed. Therefore, $\mathrm{Ker}(S)$ is equal to one line.

Set $\psi^{-1}(S_n)=([\boldsymbol{v}_n],[\boldsymbol{t}_n])$, for each $n\in\mathbb{N}$, and $\psi^{-1}(S)=([\boldsymbol{v}],[\boldsymbol{t}])$. Since $\psi^{-1}:\psi(\mathbb{P}^2_{\mathbb{C}}\times\mathbb{P}^2_{\mathbb{C}})\to\mathbb{P}^2_{\mathbb{C}}\times\mathbb{P}^2_{\mathbb{C}}$ is continuous, it follows that $\psi^{-1}(S_n)\to\psi^{-1}(S)$ as $n\to\infty$. Therefore $[\boldsymbol{v}_n]\to[\boldsymbol{v}]$ as $n\to\infty$. In other words, $\ell_n\to\mathrm{Ker}(S)$. □

Corollary 4.3. *If $G\subset\mathrm{PSL}(3,\mathbb{C})$ is a discrete subgroup then*

$$\bigcup_{\ell\in\mathcal{E}}\ell\subset\mathbb{P}^2_{\mathbb{C}}$$

is a closed set.

Proof. If (x_n) is a sequence of points in $\bigcup_{\ell\in\mathcal{E}}\ell$ such that $x_n\to x$ as $n\to\infty$, then for each $n\in\mathbb{N}$ there exists $\ell_n\in\mathcal{E}$ such that $x_n\in\ell_n$. Since $(\mathbb{P}^2_{\mathbb{C}})^*$ is compact and $\mathcal{E}$ is closed, we can assume that $\ell_n\to\ell\in\mathcal{E}$ as $n\to\infty$. It follows that $x\in\ell$ and

$$x\in\bigcup_{\ell\in\mathcal{E}}\ell.$$ □

The following lemma is a generalization of a classical result in Kleinian groups theory. See, for example, [Maskit 1988, Proposition II.C.6].

Lemma 4.4. *If $g\in\mathrm{PSL}(3,\mathbb{C})$ is a complex homothety such that $\Lambda_{\mathrm{Kul}}(g)=\ell\cup\{p\}$ and $h\in\mathrm{PSL}(3,\mathbb{C})$ is a transformation such that $h(\ell)=\ell$ and $h(p)\neq p$ then the subgroup $\langle g,h\rangle\subset\mathrm{PSL}(3,\mathbb{C})$ is not discrete.*

Proof. We can assume that $\ell=\overleftrightarrow{e_1,e_2}$ and $p=e_3$. Then

$$\boldsymbol{g}=\begin{pmatrix}a&0&0\\0&a&0\\0&0&a^{-2}\end{pmatrix},\quad\text{where } 0<|a|<1 \text{ and } \boldsymbol{h}=\begin{pmatrix}\boldsymbol{h}_{11}&\boldsymbol{h}_{12}&\boldsymbol{h}_{13}\\\boldsymbol{h}_{21}&\boldsymbol{h}_{22}&\boldsymbol{h}_{23}\\0&0&\boldsymbol{h}_{33}\end{pmatrix}$$

are lifts of g and h respectively. Since $h(p)\neq p$, either $\boldsymbol{h}_{13}\neq0$ or $\boldsymbol{h}_{23}\neq0$.

By straightforward computations $[g^n,h]$ is induced by the matrix:

$$\begin{pmatrix}1&0&(a^{3n}-1)\boldsymbol{h}_{13}/\boldsymbol{h}_{33}\\0&1&(a^{3n}-1)\boldsymbol{h}_{23}/\boldsymbol{h}_{33}\\0&0&1\end{pmatrix}.$$

It follows that the sequence of distinct elements $[g^n,h]\in\langle g,h\rangle$ converges to the transformation in PSL(3, $\mathbb{C}$) induced by the matrix

$$\begin{pmatrix}1&0&-\boldsymbol{h}_{13}/\boldsymbol{h}_{33}\\0&1&-\boldsymbol{h}_{23}/\boldsymbol{h}_{33}\\0&0&1\end{pmatrix}.$$

Hence, $\langle g,h\rangle$ is not discrete. □

Corollary 4.5. *If $G \subset \mathrm{PSL}(3, \mathbb{C})$ is a discrete subgroup then $\Lambda_{\mathrm{Myr}}(G)$ and $\bigcup_{\ell \in \mathcal{E}} \ell$ are equal except in the case when $\Lambda_{\mathrm{Myr}}(G)$ is equal to a disjoint union of one line and one point.*

Proof. Clearly $\bigcup_{\ell \in \mathcal{E}} \ell \subset \Lambda_{\mathrm{Myr}}(G)$.

If $\Lambda_{\mathrm{Myr}}(G)$ is equal to one line, then G contains a parabolic element, hence $\Lambda_{\mathrm{Myr}}(G) \subset \bigcup_{\ell \in \mathcal{E}} \ell$.

If we assume that $\Lambda_{\mathrm{Myr}}(G)$ is not equal to one line and $x \in \Lambda_{\mathrm{Myr}}(G) \setminus \bigcup_{\ell \in \mathcal{E}} \ell$ then $\{x\} = \mathrm{Ker}(S)$ for some $S \in G'$. It follows that $\mathrm{Im}(S)$ is an effective line by [Barrera et al. 2011a, Lemma 3.2(ii)]. Thus, x does not lie on the line $\mathrm{Im}(S)$. Hence, G contains a complex homothety with an isolated fixed point not lying in the closed set $\bigcup_{\ell \in \mathcal{E}} \ell$. To see this, consider a "round" closed neighborhood U of x disjoint from the closed set $\bigcup_{\ell \in \mathcal{E}} \ell$. Since $S \in G'$, there exists a sequence of distinct elements $g_n \in G$ such that $g_n \to S$ uniformly on compact subsets of $\mathbb{P}^2_{\mathbb{C}} \setminus \mathrm{Ker}(S)$. Now, by [loc. cit.] there exists a subsequence of g_n denoted the same, such that $g_n^{-1}(\,\cdot\,) \to x$ as $n \to \infty$ uniformly on compact subsets of $\mathbb{P}^2_{\mathbb{C}} \setminus \mathrm{Im}(S)$. For n large enough, g_n^{-1} sends U into its interior. It follows that g_n is loxodromic, with a fixed point in the interior of U. This g_n is necessarily a complex homothety, because otherwise U would intersect $\bigcup_{\ell \in \mathcal{E}} \ell$.

If $\bigcup_{\ell \in \mathcal{E}} \ell$ is not equal to one line then there is an effective line, ℓ_0, different from the fixed line of the complex homothety. We reach a contradiction because we can iterate ℓ_0 with respect to the complex homothety and obtain that its isolated fixed point is in the closed set $\bigcup_{\ell \in \mathcal{E}} \ell$.

If $\bigcup_{\ell \in \mathcal{E}} \ell$ is equal to one line then, by hypothesis, there exists points $y \neq x$ such that $y \notin \bigcup_{\ell \in \mathcal{E}} \ell$. It follows that there exist two distinct complex homotheties with one common fixed line, so G is not discrete by Lemma 4.4. □

Notation 4.6. Let $U \subset \mathbb{P}^2_{\mathbb{C}}$ be an open set.

- The *number of lines* contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is denoted by $\lambda(U)$.
- The *maximum number of lines in general position* contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is denoted by $\mu(U)$.

There are examples of discrete groups $G \subset \mathrm{PU}(2, 1) \subset \mathrm{PSL}(3, \mathbb{C})$ such that

$$\Omega(G) = \mathrm{Eq}(G) = \mathbb{H}^2_{\mathbb{C}}.$$

Thus, in this case, the set

$$\bigcup_{\ell \in \mathcal{E}} \ell = \mathbb{P}^2_{\mathbb{C}} \setminus \mathbb{H}^2_{\mathbb{C}}$$

is big enough as to contain infinitely many lines which are not effective lines. On the other hand, we have the following:

Remark 4.7. If $G \subset \mathrm{PSL}(3, \mathbb{C})$ is a discrete subgroup, then:

(i) $\mu(\mathrm{Eq}(G)) = 1$ if and only if there is only one effective line for G.

(ii) If we assume that $\Lambda_{\mathrm{Myr}}(G) \neq \mathbb{P}^2_{\mathbb{C}}$ then the maximum number of effective lines for G in general position is equal to two if and only if $\mu(\mathrm{Eq}(G)) = 2$. (It could happen that $\Lambda_{\mathrm{Myr}}(G) = \mathbb{P}^2_{\mathbb{C}}$ but the maximum number of effective lines in general position is equal to two, for example in the double suspension of a Picard group. See Example (iii) in Section 7B.)

(iii) If we assume that $\Lambda_{\mathrm{Myr}}(G) \neq \mathbb{P}^2_{\mathbb{C}}$ then the maximum number of effective lines for G in general position is equal to three if and only if $\mu(\mathrm{Eq}(G)) = 3$. (It could happen that $\Lambda_{\mathrm{Myr}}(G) = \mathbb{P}^2_{\mathbb{C}}$ but the maximum number of effective lines in general position is equal to three, for example, in the suspension of a Picard group extended by an infinite group. See Example (ii) in Section 7C)

Lemma 4.8. *If $G \subset \mathrm{PSL}(3, \mathbb{C})$ is a subgroup, and there exists $S \in G' \subset \mathrm{SP}(3, \mathbb{C})$ such that $\mathrm{Ker}(S)$ is a line and $\mathrm{Im}(S) \notin \mathrm{Ker}(S)$, then G contains a loxodromic element.*

Proof. There exists a sequence of distinct elements $g_n \in G$ such that $g_n \to S$ as $n \to \infty$, uniformly on compact subsets of $\mathbb{P}^2_{\mathbb{C}} \setminus \mathrm{Ker}(S)$. In particular, if $V \subset \mathbb{P}^2_{\mathbb{C}}$ is an open "ball" containing the point $\mathrm{Im}(S)$, such that $\overline{V} \cap \mathrm{Ker}(S) = \varnothing$ then there exists $N > 0$ such that $n \geq N$ implies that $g_n(\overline{V}) \subset V$. Therefore, g_n is loxodromic for every $n \geq N$; see [Navarrete 2008, Definition 6.1]. □

Lemma 4.9. *If $G \subset \mathrm{PSL}(3, \mathbb{C})$ is a subgroup and there exists $S, T \in G' \subset \mathrm{SP}(3, \mathbb{C})$ such that $\mathrm{Ker}(S)$ and $\mathrm{Ker}(T)$ are lines, $\mathrm{Im}(T) \notin \mathrm{Ker}(S)$ and $\mathrm{Im}(S) \notin \mathrm{Ker}(T)$, then G contains a loxodromic element.*

Proof. Let (g_n) and (h_n) be sequences of distinct elements in G such that $g_n \to S$ and $h_n \to T$ as $n \to \infty$. Then the sequence $f_n := g_n \circ h_n$ of elements of G satisfies that $f_n \to S \circ T$ as $n \to \infty$ and $\mathrm{Im}(S \circ T) = \mathrm{Im}(S) \notin \mathrm{Ker}(T) = \mathrm{Ker}(S \circ T)$. It follows from Lemma 4.8 that G contains a loxodromic element. □

Proposition 4.10. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $U \neq \varnothing$ be equal to $\mathrm{Eq}(G)$, $\Omega(G)$ or $U_{\max}(G)$. If $\mu(U) \geq 3$ then G contains a loxodromic element*

Proof. The hypothesis $\mu(U) \geq 3$ and Theorem 3.4(iii) imply that $U = \mathrm{Eq}(G)$.

By Corollary 4.5, $\Lambda_{\mathrm{Myr}}(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \mathrm{Eq}(G) \neq \mathbb{P}^2_{\mathbb{C}}$ is the union of effective lines for G. Therefore, there exist three pseudoprojective maps $S_1, S_2, S_3 \in G' \subset \mathrm{SP}(3, \mathbb{C})$ such that $\mathrm{Ker}(S_1), \mathrm{Ker}(S_2), \mathrm{Ker}(S_3)$ are three lines in general position. If $\mathrm{Im}(S_j) \notin \mathrm{Ker}(S_j)$ for some $1 \leq j \leq 3$ then Lemma 4.8 implies that G contains a loxodromic element. Hence, we can assume that

$$\mathrm{Im}(S_1) \in \mathrm{Ker}(S_1), \qquad \mathrm{Im}(S_2) \in \mathrm{Ker}(S_2), \qquad \mathrm{Im}(S_3) \in \mathrm{Ker}(S_3).$$

In this case, it is not hard to check that there exists $i \neq j$, $1 \leq i, j \leq 3$, such that $\mathrm{Im}(S_i) \notin \mathrm{Ker}(S_j)$ and $\mathrm{Im}(S_j) \notin \mathrm{Ker}(S_i)$. By Lemma 4.9, there exists a loxodromic element. $\square$

5. Counting lines

Definition 5.1. If p is a point and ℓ is a line such that $p \notin \ell$, then there is a *projection* from $\mathbb{P}^2_{\mathbb{C}} \setminus \{p\}$ to ℓ, denoted by

$$\pi = \pi_{p,\ell} : \mathbb{P}^2_{\mathbb{C}} \setminus \{p\} \to \ell,$$
$$\pi(z) = \overleftrightarrow{z, p} \cap \ell.$$

Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a group, and $p \in \mathbb{P}^2_{\mathbb{C}}$ a point such that $Gp = p$, then there is a group morphism given by

$$\Pi = \Pi_{p,\ell} : G \to \mathrm{Bihol}(\ell),$$
$$\Pi(g)(x) = \pi(g(x)).$$

Lemma 5.2. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup. If $V \subset \mathbb{P}^2_{\mathbb{C}}$ is an open G-invariant set such that $\mu(V) = 2$, then there is a point $p \in \mathbb{P}^2_{\mathbb{C}} \setminus V$ such that $Gp = p$.*

Proof. Let $\mathcal{L} = \{\ell \in (\mathbb{P}^2_{\mathbb{C}})^* : \ell \subset \mathbb{P}^2_{\mathbb{C}} \setminus V\}$. Since $\mu(V) = 2$, it follows that $\bigcap_{\ell \in \mathcal{L}} \ell$ is equal to one point denoted p. If $g \in G$ then $g(\mathcal{L}) = \mathcal{L}$, so

$$g(p) = g\left(\bigcap_{\ell \in \mathcal{L}} \ell\right) = \bigcap_{\ell \in \mathcal{L}} g(\ell) = \bigcap_{\ell \in \mathcal{L}} \ell = p. \qquad \square$$

Lemma 5.3. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $V \subset \mathbb{P}^2_{\mathbb{C}}$ a G-invariant open set such that $L_0(G) \subset \mathbb{P}^2_{\mathbb{C}} \setminus V$. If $3 \leq \lambda(V) < \infty$ and $\mu(V) = 2$, then:*

(i) *If ℓ is any line not containing the G fixed point, p, then $\Pi_{p,\ell}(G)$ is finite.*

(ii) *The normal subgroup $\mathrm{Ker}(\Pi)$ has finite index in G.*

(iii) *There exists $h_0 \in \mathrm{PSL}(3, \mathbb{C})$ such that every element in $h_0 \mathrm{Ker}(\Pi)(h_0)^{-1}$ of infinite order has a lift to $\mathrm{SL}(3, \mathbb{C})$ of the form*

$$\begin{pmatrix} 1 & b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(iv) *If h_0 is as in (iii), then the set A consisting of all $a^3 \in \mathbb{C}^*$, such that there exists $g \in h_0 \mathrm{Ker}(\Pi) h_0^{-1}$ with a lift of the form:*

$$\begin{pmatrix} a^{-2} & b & c \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix},$$

is a finite subgroup of $\mathbb{C}^*$.

(v) *There is a line* ℓ_0 *such that* $\mathrm{Eq}(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \ell_0$. *Moreover,* $L_0(G) = \ell_0$.

(vi) *The Kulkarni discontinuity region* $\Omega(G)$ *is equal to* $\mathbb{P}^2_{\mathbb{C}} \setminus \ell_0 = \mathrm{Eq}(G)$.

Proof. Set $\mathcal{L} = \{\ell \in (\mathbb{P}^2_{\mathbb{C}})^* : \ell \subset \mathbb{P}^2_{\mathbb{C}} \setminus V\}$ and $n_0 = |\mathcal{L}|$. Since $\mu(V) = 2$ then every line in $\mathcal{L}$ passes through a point denoted by p, and $Gp = p$.

(i) Since $\{g(\ell) : g \in G,\ \ell \in \mathcal{L}\} = \mathcal{L}$, it follows that $F := \pi\left(\bigcup_{\ell\in\mathcal{L}}(\ell \setminus \{p\})\right)$ is a $\Pi(G)$-invariant set whose cardinality is $n_0 \geq 3$. Thus

$$\Gamma = \bigcap_{x\in F} \mathrm{Isot}(x, \Pi(G))$$

is a normal subgroup of $\Pi(G)$ with finite index. Moreover, every element in F is fixed by Γ. Since F contains more than three elements we conclude that $\Gamma = \{\mathrm{Id}\}$. Therefore $\Pi(G)$ is finite.

(ii) The normal subgroup $\mathrm{Ker}(\Pi)$ has finite index because $G/\mathrm{Ker}(\Pi) \cong \Pi(G)$ is finite.

(iii) We can assume, by conjugating, that $Ge_1 = e_1$ and projection $\Pi = \Pi_{e_1, \overleftrightarrow{e_2, e_3}}$. If $g \in \mathrm{Ker}(\Pi)$ then any lift for g in $\mathrm{SL}(3, \mathbb{C})$ has the form

$$\boldsymbol{g} = \begin{pmatrix} a^{-2} & b & c \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix}, \quad \text{where } a \in \mathbb{C}^* \text{ and } b, c \in \mathbb{C}.$$

If $\boldsymbol{g}$ is diagonalizable, then there are $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{C}^3$ such that $\{e_1, \boldsymbol{v}_1, \boldsymbol{v}_2\}$ is an eigenbasis for $\boldsymbol{g}$ whose respective eigenvalues are $\{a^{-2}, a, a\}$. Consequently, $\overleftrightarrow{v_1, v_2} \subset L_0(g) \subset L_0(G) \subset \mathbb{P}^2_{\mathbb{C}} \setminus V$ and $e_1 \notin \overleftrightarrow{v_1, v_2}$, which contradicts the hypothesis that $\mu(V) = 2$. Therefore, $\boldsymbol{g}$ is not diagonalizable, which implies that $a^{-2} = a$, so g has a lift of the form

$$\boldsymbol{g} = \begin{pmatrix} 1 & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{where } b, c \in \mathbb{C}. \tag{1}$$

We can assume that there is an element $g_0 \in \mathrm{Ker}(\Pi)$ such that g_0 has a lift $\boldsymbol{g}_0 \in \mathrm{SL}(3, \mathbb{C})$ given by

$$\boldsymbol{g}_0 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2}$$

If we assume that there is an element $g_1 \in \mathrm{Ker}(\Pi)$ which has a lift $\boldsymbol{g}_1 \in \mathrm{SL}(3,\mathbb{C})$ given as in (1) with $c \neq 0$, then for every $n \in \mathbb{N}$,

$$\boldsymbol{g}_0^n \boldsymbol{g}_1 = \begin{pmatrix} 1 & b+n & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By straightforward computations, we see that

$$\ell_n = \overleftrightarrow{[e_1], [0 : -c : b+n]} \subset L_0(g_0^n g_1) \subset L_0(G) \subset \mathbb{P}^2_{\mathbb{C}} \setminus V.$$

Moreover, $\ell_n \neq \ell_m$ whenever $n \neq m$. Thus $\mathcal{L}$ contains infinitely many lines, which contradicts the hypothesis that $\lambda(V) < \infty$.

(iv) By straightforward computations, the set A is a subgroup of $\mathbb{C}^*$. By (iii), every element in $\mathrm{Ker}(\Pi)$ is elliptic or parabolic. It follows that $A \subset \mathbb{S}^1$. Assume that A is infinite; then there is a sequence $a_n^3 \subset A$ of distinct elements such that $a_n^{1/2} \to 1$ as $n \to \infty$. For each $n \in \mathbb{N}$, let $g_n \in \mathrm{Ker}(\Pi)$ with a lift $\boldsymbol{g}_n \in \mathrm{SL}(3,\mathbb{C})$ of the form

$$\boldsymbol{g}_n = \begin{pmatrix} a_n & b_n & c_n \\ 0 & a_n^{-1/2} & 0 \\ 0 & 0 & a_n^{-1/2} \end{pmatrix}, \quad \text{where } b_n, c_n \in \mathbb{C}.$$

If $\boldsymbol{g_0}$ is as in (2), then

$$\boldsymbol{g}_n^{-1} \boldsymbol{g}_0 \boldsymbol{g}_n = \begin{pmatrix} 1 & a_n^{-3/2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

as $n \to \infty$, a contradiction to the hypothesis that G is discrete.

(v) Let H denote the finite index subgroup of $h_0 \,\mathrm{Ker}(\Pi) h_0^{-1}$ consisting of all elements with a lift of the form

$$\begin{pmatrix} 1 & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We can assume, by (iii), that each element of infinite order, $h \in H$ has a lift $\boldsymbol{h} \in \mathrm{SL}(3,\mathbb{C})$ which is given by

$$\boldsymbol{h} = \begin{pmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It follows that $\mathrm{Eq}(H) = \mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_1, e_3}$. Since H has finite index in $h_0 \,\mathrm{Ker}(\Pi) h_0^{-1}$,

$$\mathrm{Eq}(h_0 G h_0^{-1}) = \mathrm{Eq}(h_0 \,\mathrm{Ker}(\Pi) h_0^{-1}) = \mathrm{Eq}(H) = \mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_1, e_3}.$$

Finally, $L_0(h_0 G h_0^{-1}) = L_0(h_0 \,\mathrm{Ker}(\Pi) h_0^{-1}) = L_0(H) = \overleftrightarrow{e_1, e_3}$.

(vi) G acts properly and discontinuously on $\mathrm{Eq}(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \ell_0$. Since $\mathrm{Ker}(\Pi)$ has finite index in G and every element of infinite order has canonical form as in (2), we notice that G does not contain loxoparabolic elements. It follows by [Barrera et al. 2014a, Theorem 1.2] that $\Omega(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \ell_0$. □

Proposition 5.4. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $U \subset \mathbb{P}^2_{\mathbb{C}}$ be one of $\mathrm{Eq}(G)$, $\Omega(G)$, or $U_{\max}(G)$. If $\mu(U) = 2$ and $\lambda(U) < \infty$, then $\lambda(U) = 2$.*

Proof. If U is either $\mathrm{Eq}(G)$ or $\Omega(G)$ and $2 < \lambda(U) < \infty$, then by Lemma 5.3(v) and (vi), $\mathrm{Eq}(G) = \Omega(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \ell$ for some line ℓ. Thus, $\mu(U) = 1 = \lambda(U)$, a contradiction.

If $U = U_{\max}(G)$ and $2 < \lambda(U) < \infty$ then $L_0(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \mathrm{Eq}(G)$ by Lemma 5.3(v). Since G acts properly and discontinuously on U, it follows that $L_0(G) \subset \mathbb{P}^2_{\mathbb{C}} \setminus U$. Thus, $U \subset \mathrm{Eq}(G)$, so $U = \mathrm{Eq}(G)$ is the complement of one line in $\mathbb{P}^2_{\mathbb{C}}$, a contradiction of the hypothesis that $\mu(U) = 2$. □

Lemma 5.5. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete group and $V \subset \mathbb{P}^2_{\mathbb{C}}$ be an open G-invariant set such that $L_0(G) \subset \mathbb{P}^2_{\mathbb{C}} \setminus V$. If $\mu(V) = 3$ and $3 < \lambda(V) < \infty$, then there is a line ℓ_1 and $p \in \mathbb{P}^2_{\mathbb{C}} \setminus \ell_1$ such that*

$$\mathrm{Eq}(G) = \mathbb{P}^2_{\mathbb{C}} \setminus (\ell_1 \cup \{p\}) = \mathbb{P}^2_{\mathbb{C}} \setminus L_0(G)$$

Proof. Let us assume that $n_0 = \lambda(V) > 3$. If we define

$$\mathcal{L} = \{\ell \in (\mathbb{P}^2_{\mathbb{C}})^* : \ell \subset \mathbb{P}^2_{\mathbb{C}} \setminus V\} = \{\ell_1, \ldots, \ell_{n_0}\},$$

then the group

$$G_0 = \bigcap_{i=1}^{n_0} \mathrm{Isot}(\ell_i, G)$$

is a finite index normal subgroup of G. Since $\mu(U) = 3$ then, conjugating by a projective transformation, we can assume that

$$\ell_1 = \overleftrightarrow{e_1, e_2}, \qquad \ell_2 = \overleftrightarrow{e_2, e_3}, \qquad \ell_3 = \overleftrightarrow{e_3, e_1}.$$

It follows that every element $g \in G_0$ has a lift $\boldsymbol{g} \in \mathrm{SL}(3, \mathbb{C})$ of the form

$$\boldsymbol{g} = \begin{pmatrix} \boldsymbol{g}_{11} & 0 & 0 \\ 0 & \boldsymbol{g}_{22} & 0 \\ 0 & 0 & \boldsymbol{g}_{33} \end{pmatrix}, \quad \text{where } \boldsymbol{g}_{11}\boldsymbol{g}_{22}\boldsymbol{g}_{33} = 1.$$

If ℓ_j is any line in $\mathcal{L} \setminus \{\ell_1, \ell_2, \ell_3\}$ then $e_1 \in \ell_j$ or $e_2 \in \ell_j$ or $e_3 \in \ell_j$, because $\mu(V) = 3$. We assume, without loss of generality, that $e_3 \in \ell_j$ for all lines ℓ_j in $\mathcal{L} \setminus \{\ell_1, \ell_2, \ell_3\}$. Set $\Pi = \Pi_{e_3, \ell_1}$ and $\pi = \pi_{e_3, \ell_1}$, and notice that $\pi(e_1)$, $\pi(e_2)$, $\pi(\ell_j \setminus e_3)$ are three distinct fixed points in ℓ_1 for the group $\Pi(G_0)$, so $\Pi(G_0) = \{\mathrm{Id}\}$. Therefore, for each $g \in G_0$, there is a nonzero complex number $\boldsymbol{g}_{33}^2$ such that $\boldsymbol{g} \in \mathrm{SL}(3, \mathbb{C})$ given by

$$\boldsymbol{g} = \begin{pmatrix} \boldsymbol{g}_{33} & 0 & 0 \\ 0 & \boldsymbol{g}_{33} & 0 \\ 0 & 0 & \boldsymbol{g}_{33}^{-2} \end{pmatrix}$$

is a lift of g. We conclude that

$$\mathrm{Eq}(G) = \mathrm{Eq}(G_0) = \mathbb{P}^2_{\mathbb{C}} \setminus (\ell_1 \cup \{e_1\}). \qquad \square$$

Proposition 5.6. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $U \subset \mathbb{P}^2_{\mathbb{C}}$ be one of* $\mathrm{Eq}(G)$, $\Omega(G)$, *or* $U_{\max}(G)$. *If $\mu(U) = 3$ and $\lambda(U) < \infty$ then $\lambda(U) = 3$.*

Proof. Given that $\mu(U) = 3$, Theorem 3.4(iii) implies that $U = \mathrm{Eq}(G)$. If $\lambda(U) > 3$ then applying Lemma 5.5, we obtain that $\mu(U) = 1$, a contradiction to the hypothesis that $\mu(U) = 3$. Therefore $\lambda(U) = 3$. $\square$

Proposition 5.7. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $\varnothing \neq U \subset \mathbb{P}^2_{\mathbb{C}}$ be equal to on of* $\mathrm{Eq}(G)$, $\Omega(G)$, *or* $U_{\max}(G)$. *If $\mu(U) \in \{2, 3\}$ and $\lambda(U) = \infty$, then there is a perfect set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$, so there are uncountably many lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.*

First, we consider the case when $\mu(U) = 2$. In this case, there exists a fixed point of G corresponding to the intersection point of any two distinct lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$. We can assume that e_3 is this fixed point. Hence every element $g \in G$ is induced by a unique matrix of the form

$$\boldsymbol{g} = \begin{pmatrix} \boldsymbol{g}_{11} & \boldsymbol{g}_{12} & 0 \\ \boldsymbol{g}_{21} & \boldsymbol{g}_{22} & 0 \\ \boldsymbol{g}_{31} & \boldsymbol{g}_{32} & 1 \end{pmatrix} \tag{3}$$

If we set $\Pi = \Pi_{e_3, \overleftrightarrow{e_1, e_2}} : G \to \mathrm{Bihol}(\overleftrightarrow{e_1, e_2})$, then we consider the subcases depending on whether $\Pi(G)$ is not elementary or not discrete, elementary of two limit points, elementary of one limit point, or finite. These subcases are considered in Lemmas 5.8, 5.12, 5.13, and 5.14.

The subgroup $\mathrm{Ker}(\Pi)$ also plays an important role in the proof of Proposition 5.7 in the case when $\mu(U) = 2$, and we prove in Lemma 5.9 that $\mathrm{Ker}(\Pi)$ contains a free abelian finite index subgroup H, consisting of all elements of infinite order and the identity. Moreover, we prove in Lemma 5.10 that necessarily the rank of H is smaller or equal to 2. This result is analogous to the first Bieberbach theorem with the difference that H does not have maximal rank.

Lemma 5.8. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $\varnothing \neq U \subset \mathbb{P}^2_{\mathbb{C}}$ be equal to one of* $\mathrm{Eq}(G)$, $\Omega(G)$, *or* $U_{\max}(G)$. *If $\mu(U) = 2$, $\lambda(U) = \infty$, and $\Pi(G)$ is not elementary or not discrete then there is a perfect set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.*

Proof. The action of $\Pi(G)$ over the line $\overleftrightarrow{e_1, e_2}$ represents the action of the group G on the full pencil of lines passing through the point e_3. Since $\Pi(G)$ is not elementary or not discrete, we have that for any line ℓ passing through e_3 (except for a finite number of lines), the closure of the set $\{g(\ell) : g \in G\}$ contains a perfect set of lines. Since $\lambda(U) = \infty$, there exists a line ℓ contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ such that

$$\overline{\{g(\ell) : g \in G\}} \subset \mathbb{P}^2_{\mathbb{C}} \setminus U$$

contains a perfect set of lines. □

Lemma 5.9. *The set $H \subset \operatorname{Ker}(\Pi)$, consisting of all elements of infinite order and the identity, is a free abelian finite index normal subgroup of* $\operatorname{Ker}(\Pi)$. *Moreover, every element in H is induced by a matrix of the form*

$$\boldsymbol{h} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \boldsymbol{h}_{31} & \boldsymbol{h}_{32} & 1 \end{pmatrix}. \tag{4}$$

Hence, H is isomorphic to a discrete subgroup of $\mathbb{C}^2$.

Proof. Let us assume that $h \in \operatorname{Ker}(\Pi)$ has infinite order and it is induced by a matrix of the form

$$\boldsymbol{h} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ \boldsymbol{h}_{31} & \boldsymbol{h}_{32} & 1 \end{pmatrix}. \tag{5}$$

If $|a| \neq 1$ then h is a complex homothety and $\{e_3\} \cup \ell_h$ is the set of fixed points of h, where ℓ_h is a line not passing through e_3. It is a contradiction of the hypothesis that $\mu(U) = 2$. Hence $|a| = 1$.

If we assume that $a \neq 1$ then h is an elliptic element of infinite order, so G is not discrete, contradiction. Therefore, $a = 1$ and every element in H is induced by a matrix of the form (4). It follows immediately that H is a free abelian normal subgroup of $\operatorname{Ker}(\Pi)$.

If $H = \{\mathrm{Id}\}$ then $\operatorname{Ker}(\Pi)$ is a discrete subgroup and every element in $\operatorname{Ker}(\Pi)$ has finite order. It follows that $\operatorname{Ker}(\Pi)$ is finite.

If $H \neq \{\mathrm{Id}\}$ then there exists $h \in H$ induced by a matrix of the form (4) where $(\boldsymbol{h}_{31}, \boldsymbol{h}_{32}) \neq (0, 0)$. If $g \in \operatorname{Ker}(\Pi)$ is induced by a matrix of the form

$$\boldsymbol{g} = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ \boldsymbol{g}_{31} & \boldsymbol{g}_{32} & 1 \end{pmatrix}, \tag{6}$$

then $g^{-n}hg^n \in H$ is induced by a matrix of the form

$$\boldsymbol{h} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a^n\boldsymbol{h}_{31} & a^n\boldsymbol{h}_{32} & 1 \end{pmatrix}. \tag{7}$$

It follows that H is a finite index subgroup of $\operatorname{Ker}(\Pi)$. Otherwise, there exists a sequence of distinct elements of H tending to an element in $\mathrm{PSL}(3, \mathbb{C})$. □

Lemma 5.10. *Let H be as in Lemma 5.9. If H has rank at least 3, then H is not a complex Kleinian group.*

Proof. Let us assume that H contains the free abelian group generated by

$$\boldsymbol{h}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_1 & b_1 & 1 \end{pmatrix}, \qquad \boldsymbol{h}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_2 & b_2 & 1 \end{pmatrix}, \qquad \boldsymbol{h}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_3 & b_3 & 1 \end{pmatrix},$$

where $\{(a_1, b_1), (a_2, b_2), (a_3, b_3)\}$ is an $\mathbb{R}$-linearly independent set of vectors, and $\{(a_1, b_1), (a_2, b_2)\}$ is a $\mathbb{C}$-linearly independent set of vectors.

Conjugating by an element in $\mathrm{PGL}(3, \mathbb{C})$, we can assume that $(a_1, b_1) = (1, 0)$, $(a_2, b_2) = (0, 1)$, and $(a_3, b_3) = (\lambda, \mu)$, where $\lambda \notin \mathbb{R}$.

If m, n, k are integers, not all zero, then the element

$$\boldsymbol{h}_1^n \boldsymbol{h}_2^m \boldsymbol{h}_3^k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ n + k\lambda & m + k\mu & 1 \end{pmatrix}$$

has the property that

$$\operatorname{Ker}(h_1^n h_2^m h_3^k) = L_0(h_1^n h_2^m h_3^k)$$

is the complex line in $\mathbb{P}^2_{\mathbb{C}}$, passing through e_3 and defined by

$$\{[z_1 : z_2 : z_3] \in \mathbb{P}^2_{\mathbb{C}} : (n + k\lambda)z_1 + (m + k\mu)z_2 = 0\}.$$

Moreover, this complex line can be identified via dual vector with the point

$$[n + k\lambda : m + k\mu : 0].$$

Now, the set

$$\{[n + k\lambda : m + k\mu : 0] : m, n, k \text{ are integers, not all zero}\}$$

is dense in the set $\{[A : B : 0] : A, B \in \mathbb{C} \text{ not both zero}\}$ which is identified with the set of all lines in $\mathbb{P}^2_{\mathbb{C}}$ passing through e_3.

Now, let us assume that H acts properly and discontinuously on the open set $U \subset \mathbb{P}^2_{\mathbb{C}}$. Since, $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is a closed set containing every line of the form $\operatorname{Ker}(h_1^n h_2^m h_3^k) = L_0(h_1^n h_2^m h_3^k)$, for $m, n, k \in \mathbb{Z}$, not all zero, it follows that $\mathbb{P}^2_{\mathbb{C}} \setminus U$ contains the complete pencil of lines passing through e_3. Therefore, $U = \varnothing$. □

Lemma 5.11. *Let H be a free abelian group of rank 2, as in Lemma 5.9, acting properly and discontinuously on the open set $U \subset \mathbb{P}^2_{\mathbb{C}}$. If H is generated by transformations h_1 and h_2 induced by the matrices*

$$\boldsymbol{h}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_1 & b_1 & 1 \end{pmatrix}, \quad \boldsymbol{h}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_2 & b_2 & 1 \end{pmatrix},$$

and $\{(a_1, b_1), (a_2, b_2)\}$ is a basis of $\mathbb{C}^2$ then $\mathbb{P}^2_{\mathbb{C}} \setminus U$ contains a perfect set of lines.

Proof. As in the proof of Lemma 5.10, conjugating if needed, we can assume that $(a_1, b_1) = (1, 0)$ and $(a_2, b_2) = (0, 1)$.

If m, n are integers, not both zero, then the element

$$\boldsymbol{h}_1^n \boldsymbol{h}_2^m = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ n & m & 1 \end{pmatrix}$$

has the property that

$$\operatorname{Ker}(h_1^n h_2^m) = L_0(h_1^n h_2^m)$$

is the complex line in $\mathbb{P}^2_{\mathbb{C}}$, passing through e_3 and defined by

$$\{[z_1 : z_2 : z_3] \in \mathbb{P}^2_{\mathbb{C}} : n z_1 + m z_2 = 0\}.$$

Moreover, this complex line can be identified with the point

$$[n : m : 0].$$

Now, the set

$$\{[n : m : 0] : m, n \text{ are integers, not both zero }\}$$

is dense in the set of lines $\{[A : B : 0] : A, B \in \mathbb{R} \text{ not both zero }\} \cong \mathbb{P}^1_{\mathbb{R}}$.

Since, $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is a closed set containing every line of the form $\operatorname{Ker}(h_1^n h_2^m) = L_0(h_1^n h_2^m)$, for $m, n \in \mathbb{Z}$, not both zero, it follows that $\mathbb{P}^2_{\mathbb{C}} \setminus U$ contains a circle of lines passing through e_3. $\square$

Lemma 5.12. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $\varnothing \neq U \subset \mathbb{P}^2_{\mathbb{C}}$ be equal to one of $\mathrm{Eq}(G)$, $\Omega(G)$, or $U_{\max}(G)$. If $\mu(U) = 2$ and $\Pi(G)$ is elementary with two limit points, then $\lambda(U) = 2$ or there exists a perfect set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.*

Proof. Since $\Pi(G)$ is elementary with two limit points, there exists a G-invariant set of two lines passing through e_3; we can assume that $\overleftrightarrow{e_1, e_3}$ and $\overleftrightarrow{e_2, e_3}$ are these two lines. Moreover, we can assume that each one of these lines is G-invariant (up to a finite index subgroup). It follows that every element in G is induced by a matrix of the form

$$\begin{pmatrix} \boldsymbol{g}_{11} & 0 & 0 \\ 0 & \boldsymbol{g}_{22} & 0 \\ \boldsymbol{g}_{31} & \boldsymbol{g}_{32} & 1 \end{pmatrix}. \tag{8}$$

By Lemma 5.10, H has rank at most 2, and we consider the cases according to the rank of this group. If the rank of H is 2 and we assume that the hypotheses of Lemma 5.11 are satisfied, then $\mathbb{P}^2_{\mathbb{C}} \setminus U$ contains a perfect set of lines. Hence, we can assume that the rank of H is 2 and H is generated by two elements h_1 and h_2, induced by the matrices

$$\boldsymbol{h}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a_1 & b_1 & 1 \end{pmatrix}, \quad \boldsymbol{h}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \omega a_1 & \omega b_1 & 1 \end{pmatrix},$$

where $\omega \in \mathbb{C} \setminus \mathbb{R}$.

Since $\Pi(G)$ is elementary with two limit points, there exists an element $g \in G$ induced by a matrix of the form (8) where $|\boldsymbol{g}_{11}| \neq |\boldsymbol{g}_{22}|$.

Since $g^{-1}hg \in H$, it follows that

$$(\boldsymbol{g}_{11}a_1, \boldsymbol{g}_{22}b_1) = m(a_1, b_1) + n\omega(a_1, b_1), \quad \text{for some } m, n \in \mathbb{Z}. \tag{9}$$

If $a_1 \neq 0 \neq b_1$ then, by (9),

$$\boldsymbol{g}_{11} = m + n\omega = \boldsymbol{g}_{22},$$

a contradiction. Hence, $a_1 = 0$ or $b_1 = 0$.

We can assume that $a_1 = 0$; then for every $g_1, g_2 \in G$ we have $[g_1, g_2] \in H$, and by a straightforward computation, we can conjugate G so that every element in this conjugate group is induced by a matrix of the form

$$\begin{pmatrix} \boldsymbol{g}_{11} & 0 & 0 \\ 0 & \boldsymbol{g}_{22} & 0 \\ 0 & \boldsymbol{g}_{32} & 1 \end{pmatrix}.$$

Moreover, G contains a finite index abelian subgroup, G_0, generated by the elements

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & b_1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \omega b_1 & 1 \end{pmatrix}, \quad \begin{pmatrix} \zeta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & y & 1 \end{pmatrix},$$

where $b_1 \in \mathbb{C} \setminus \{0\}$, $\omega \in \mathbb{C} \setminus \mathbb{R}$ and $|\zeta| < 1$. It follows that $\mathrm{Eq}(G) = \mathrm{Eq}(G_0) = \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_3} \cup \overleftrightarrow{e_2, e_3})$. Therefore, $\lambda(\mathrm{Eq}(G)) = \lambda(\mathrm{Eq}(G_0)) = 2$.

If U is an open set where G acts properly and discontinuously, then $\overleftrightarrow{e_1, e_3} = L_0(G_0) \subset \mathbb{P}^2_{\mathbb{C}} \setminus U$. If $\mu(U) = 2$, there exists another line ℓ passing through e_3 such that $\ell \subset \mathbb{P}^2_{\mathbb{C}} \setminus U$. By hypothesis, $\Pi(G)$ is an elementary group of two limit points, it follows that $\overleftrightarrow{e_2, e_3} \subset \overline{G \cdot \ell} \subset \mathbb{P}^2_{\mathbb{C}} \setminus U$. Hence,

$$U \subset \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_3} \cup \overleftrightarrow{e_2, e_3}) = \mathrm{Eq}(G). \tag{10}$$

In the case when $U = \Omega(G)$, it is known that $\mathrm{Eq}(G) \subset \Omega(G) = U$; hence

$$U = \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_3} \cup \overleftrightarrow{e_2, e_3}).$$

Therefore, $\lambda(U) = 2$.

In the case when $U = U_{\max}(G)$, it follows from (10) that $U = \mathrm{Eq}(G)$, and $\lambda(U) = 2$.

The case when H has rank one is analogous and we omit it. Finally, when $H = \{\mathrm{Id}\}$, the group is a finite extension of a cyclic group generated by a loxodromic element. It follows that $\mathrm{Eq}(G) = \Omega(G)$ is the complement of two lines in $\mathbb{P}^2_{\mathbb{C}}$. On the other hand, every maximal open set, U, where the action of G is properly discontinuous, is equal to the complement of one line and one point or the complement of one single line. □

Lemma 5.13. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $\varnothing \neq U \subset \mathbb{P}^2_{\mathbb{C}}$ be equal to one of* $\mathrm{Eq}(G)$, $\Omega(G)$, *or* $U_{\max}(G)$. *If $\mu(U) = 2$, $\lambda(U) = \infty$, and $\Pi(G)$ is elementary with one limit point, then there exists a perfect set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.*

Proof. Since $\Pi(G)$ is elementary with one fixed point, we can assume that $\overleftrightarrow{e_2, e_3}$ is a G-invariant line, and every element in G is induced by a unique matrix of the form

$$\begin{pmatrix} g_{11} & 0 & 0 \\ g_{21} & g_{11} & 0 \\ g_{31} & g_{32} & 1 \end{pmatrix}. \tag{11}$$

Now, the rank of the finite index abelian subgroup $H \subset \mathrm{Ker}(\Pi)$ (see Lemmas 5.9 and 5.10) is at most 2.

If the rank of H is at most 2 and the hypotheses of Lemma 5.11 are not satisfied, then we can assume, conjugating with the appropriate matrix, that every element of H is induced by a matrix of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ h_{31} & 0 & 1 \end{pmatrix}. \tag{12}$$

(Alternatively, it can be of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & h_{32} & 1 \end{pmatrix},$$

but the proof is analogous.) It follows that H acts properly and discontinuously on $\mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_2, e_3}$, so $\mathrm{Ker}(\Pi)$ acts properly and discontinuously on $\mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_2, e_3}$.

Now, we prove that G acts properly and discontinuously on $\mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_2, e_3}$. Let (g_n) be a sequence of distinct elements of G such that $g_n(C) \cap D \neq \varnothing$ for some compact subsets C, $D \subset \mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_2, e_3}$. Since $\Pi(G)$ is elementary with one limit point, we can assume that $\Pi(g_n) = \Pi(g_1)$ for every $n \in \mathbb{N}$, so $g_1^{-1} g_n \in \mathrm{Ker}(\Pi)$ for every $n \in \mathbb{N}$. It follows that

$$g_1^{-1} g_n(C) \cap g_1^{-1}(D) \neq \varnothing,$$

contradicting the fact that $\mathrm{Ker}(\Pi)$ acts properly and discontinuously on $\mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_2, e_3}$.

It follows, by [Barrera et al. 2014a, Theorem 1.2], that $\lambda(U) \leq 2$, a contradiction. Hence, the rank of H is equal to 2 and hypotheses of Lemma 5.11 are satisfied. Therefore, $\mathbb{P}^2_{\mathbb{C}} \setminus U$ contains a perfect set of lines. □

Lemma 5.14. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete subgroup and $\varnothing \neq U \subset \mathbb{P}^2_{\mathbb{C}}$ be equal to one of $\mathrm{Eq}(G)$, $\Omega(G)$, or $U_{\max}(G)$. If $\mu(U) = 2$ and $\Pi(G)$ is finite, then there exists a perfect set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.*

Proof. Since $\Pi(G)$ is finite, it follows that $\mathrm{Ker}(\Pi)$ has finite index in G. By Lemma 5.9, H has finite index in G. Thus,

$$\Omega(G) = \Omega(H) = \mathrm{Eq}(H) = \mathrm{Eq}(G)$$

By Lemma 5.10, H has rank at most 2. First, we assume that H has rank 1. Then $\mu(\Omega(H)) = \mu(\mathrm{Eq}(H)) = 1$, and $\Omega(H)$ is a maximal open subset where G acts properly and discontinuously, contradicting the hypothesis $\mu(U) = 2$. Therefore, H has rank 2.

In the case when H does not satisfy the hypotheses of Lemma 5.11, $\Omega(H) = \mathrm{Eq}(H)$ is the complement of one line in $\mathbb{P}^2_{\mathbb{C}}$, and again $\Omega(H)$ is a maximal open set where G acts properly and discontinuously, contradicting the hypothesis $\mu(U) = 2$.

It follows that H has rank 2, and it satisfies the hypotheses of Lemma 5.11. Therefore, $\mathbb{P}^2_{\mathbb{C}} \setminus U$ contains a perfect set of lines. □

Proof of Proposition 5.7. If $\mu(U) = 2$ then the result is obtained by applying Lemmas 5.8, 5.12, 5.13, and 5.14.

Now, we consider the case when $\mu(U) = 3$ and $\lambda(U) = \infty$; then there exists a point p and a line ℓ not passing through p such that $G \cdot p = p$ and $G \cdot \ell = \ell$. We can assume that $p = e_3$ and $\ell = \overleftrightarrow{e_1, e_2}$, so every element $g \in G$ is induced by a matrix of the form

$$g = \begin{pmatrix} g_{11} & g_{12} & 0 \\ g_{21} & g_{22} & 0 \\ 0 & 0 & g_{33} \end{pmatrix} \tag{13}$$

Since $\mu(U)=3$, by Theorem 3.4(iii), we obtain $U\subset \mathrm{Eq}(G)$. It follows that

$$U=\mathrm{Eq}(G),$$

whenever U is a maximal open set where G acts properly and discontinuously or U is Kulkarni domain of discontinuity. Hence, it suffices to prove Proposition 5.7 for $U=\mathrm{Eq}(G)$.

If we set, as before, $\Pi=\Pi_{e_3,\overleftrightarrow{e_1,e_2}}:G\to \mathrm{Bihol}(\overleftrightarrow{e_1,e_2})$, then we consider the subcases depending on whether $\Pi(G)$ is not elementary or not discrete, elementary with two limit points, elementary with one limit point, or finite.

In the case when $\Pi(G)$ is not elementary or not discrete, one can prove that there exists a perfect set of lines contained in $\mathbb{P}^2_{\mathbb{C}}\setminus U$ as in the proof of Lemma 5.8.

In the case when $\Pi(G)$ is finite, $\mathrm{Ker}(\Pi)$ is a finite index subgroup of G, then

$$\mathrm{Eq}(G)=\mathrm{Eq}(\mathrm{Ker}(\Pi)).$$

Since every element in $\mathrm{Ker}(\Pi)$ is induced by a matrix of the form

$$\begin{pmatrix} g_{11} & 0 & 0\\ 0 & g_{11} & 0\\ 0 & 0 & g_{33}\end{pmatrix},$$

it follows that $\mathbb{P}^2_{\mathbb{C}}\setminus \mathrm{Eq}(G)=\mathbb{P}^2_{\mathbb{C}}\setminus \mathrm{Eq}(\mathrm{Ker}(\Pi))$ is at most $\overleftrightarrow{e_1,e_2}\cup\{e_3\}$, contradicting the hypothesis. Therefore, $\Pi(G)$ cannot be finite.

If $\Pi(G)$ is Euclidean — i.e., elementary with one limit point — then there exists a finite index subgroup of G such that every element of this subgroup can be induced by a matrix of the form

$$\begin{pmatrix} 1 & g_{12} & 0\\ 0 & 1 & 0\\ 0 & 0 & g_{33}\end{pmatrix}.$$

It follows that $\lambda(\mathrm{Eq}(G))\leq 2$, contradicting the hypothesis. Therefore, $\Pi(G)$ cannot be Euclidean.

If $\Pi(G)$ is a two limit points elementary group, then there exists a finite index subgroup of G such that every element of this subgroup can be induced by a matrix of the form

$$\begin{pmatrix} g_{11} & 0 & 0\\ 0 & g_{22} & 0\\ 0 & 0 & g_{33}\end{pmatrix}.$$

It follows that $\lambda(\mathrm{Eq}(G))\leq 3$, contradicting the hypothesis. Therefore, $\Pi(G)$ cannot be elementary with two limit points. □

Some examples of groups as in the statement of Proposition 5.7 are given in Sections 7B and 7C.

Proposition 5.15. *Let $G \subset \mathrm{PSL}(3,\mathbb{C})$ be a discrete subgroup and $U \subset \mathbb{P}^2_{\mathbb{C}}$ be equal to one of $\mathrm{Eq}(G)$, $\Omega(G)$, or $U_{\max}(G)$. If $\mu(U) \geq 4$ then $\lambda(U) = \infty$. Moreover $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is the union of a perfect set of lines, so there are uncountably many lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.*

Proof. If $U = \Omega(G)$ then $U = \mathrm{Eq}(G)$, by [Barrera et al. 2011a, Theorem 3.6]. If $U = U_{\max}(G)$. Thus, $U = \mathrm{Eq}(G)$ by maximality. Hence, it suffices to prove the statement for $U = \mathrm{Eq}(G)$. If $\Lambda_{\mathrm{Myr}}(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \mathrm{Eq}(G) = \mathbb{P}^2_{\mathbb{C}}$ then there is nothing to prove, so we assume $\Lambda_{\mathrm{Myr}}(G) \neq \mathbb{P}^2_{\mathbb{C}}$.

Finally, we prove that $\mathcal{E}(G)$ is a perfect set. By Proposition 4.2, $\mathcal{E}(G)$ is closed. Moreover, if ℓ is an effective line for G, then $\ell = \mathrm{Ker}(S)$ for some $S \in G'$ so it follows from in [op. cit., Lemma 3.2(3)] that there is a sequence of distinct effective lines accumulating at ℓ (because the maximum number of effective lines for G is at least 4, by Remark 4.7). □

Proof of Theorem 1.1. First of all, $\lambda(U) \geq 1$ because the complement of an open subset of $\mathbb{P}^2_{\mathbb{C}}$, where the infinite discrete group $G \subset \mathrm{PSL}(3,\mathbb{C})$ acts properly and discontinuously, always contains a line.

- If $\mu(U) = 1$ then $\lambda(U) = 1$.
- If $\mu(U) = 2$ then there are two subcases depending on whether $\lambda(U) < \infty$ or $\lambda(U) = \infty$. If $\lambda(U) < \infty$ then $\lambda(U) = 2$, by Proposition 5.4. In the other case, $\lambda(U) = \infty$ and Proposition 5.7 implies that there exists a perfect set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.
- If $\mu(U) = 3$ then there are two subcases depending on whether $\lambda(U) < \infty$ or $\lambda(U) = \infty$. If $\lambda(U) < \infty$ then $\lambda(U) = 3$, by Proposition 5.6. In the other case, by Proposition 5.7, there is a perfect set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$.
- If $\mu(U) \geq 4$ then Proposition 5.15 implies that $\lambda(U) = \infty$ and $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is the union of a perfect set of lines. □

6. Proof of Theorem 1.2

Lemma 6.1. *Let $G \subset \mathrm{PSL}(3,\mathbb{C})$ be a discrete subgroup and ℓ a G-invariant line such that the action of G restricted to ℓ is trivial, then:*

(i) *If there is an element in G with infinite order and a diagonalizable lift, then G is conjugate to a subgroup of $\mathrm{PSL}(3,\mathbb{C})$ such that every element has a lift to $\mathrm{SL}(3,\mathbb{C})$ of the form*

$$\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a^{-2} \end{pmatrix}, \quad \textit{where } a \in \mathbb{C}^*.$$

(ii) *If G does not contain an element with infinite order and diagonalizable lift, then G is conjugate to a subgroup of* $\mathrm{PSL}(3,\mathbb{C})$ *such that every element has a lift to* $\mathrm{SL}(3,\mathbb{C})$ *of the form*

$$\begin{pmatrix} a & 0 & b \\ 0 & a & c \\ 0 & 0 & a^{-2} \end{pmatrix}, \quad \textit{where } |a| = 1.$$

Proof. (i) After conjugating with a projective transformation, we can assume that $\ell = \overleftrightarrow{e_1, e_2}$ and there exists $g_0 \in G$ with a lift $\boldsymbol{g}_0 \in \mathrm{SL}(3,\mathbb{C})$ of the form

$$\begin{pmatrix} a_0 & 0 & 0 \\ 0 & a_0 & 0 \\ 0 & 0 & a_0^{-2} \end{pmatrix}, \quad \text{where } |a_0| < 1.$$

On the other hand, each element $g \in G$ has a lift $\boldsymbol{g} \in \mathrm{SL}(3,\mathbb{C})$ of the form

$$\begin{pmatrix} \boldsymbol{g}_{11} & 0 & \boldsymbol{g}_{13} \\ 0 & \boldsymbol{g}_{11} & \boldsymbol{g}_{23} \\ 0 & 0 & \boldsymbol{g}_{11}^{-2} \end{pmatrix}. \tag{14}$$

If $\boldsymbol{g}_{13} \neq 0$ or $\boldsymbol{g}_{23} \neq 0$ for some $g \in G$, then Lemma 4.4 implies that G is not discrete. Therefore, $\boldsymbol{g}_{13} = 0 = \boldsymbol{g}_{23}$ for every $g \in G$.

(ii) As before, we can assume that every $g \in G$ has a lift $\boldsymbol{g} \in \mathrm{SL}(3,\mathbb{C})$ as in (14). If for some $g \in G$ we assume that $|\boldsymbol{g}_{11}| \neq 1$, then $\boldsymbol{g}_{11} \neq \boldsymbol{g}_{11}^{-2}$ and $\boldsymbol{g}$ is diagonalizable, so we have a contradiction. □

Proposition 6.2. *Let $G \subset$ $\mathrm{PSL}(3,\mathbb{C})$ be an infinite discrete subgroup and ℓ a G-invariant line such that G acts trivially on ℓ, then there exists a point p such that*

$$\mathrm{Eq}(G) = \ell \cup \{p\}.$$

Proof. By Lemma 6.1 we have two cases according to whether there is an element in G of infinite order with a diagonalizable lift or there is not such an element in G. In the first case, $\mathrm{Eq}(G) = \ell \cup \{p\}$ where p is the isolated fixed point of any element in G of infinite order. In the second case, $\mathrm{Eq}(G) = \ell$. □

Lemma 6.3. *Let $G \subset$ $\mathrm{PSL}(3,\mathbb{C})$ be a discrete subgroup such that each element $g \in G$ has a lift $\boldsymbol{g} \in$ $\mathrm{SL}(3,\mathbb{C})$ of the form*

$$\begin{pmatrix} \boldsymbol{g}_{11} & 0 & 0 \\ 0 & \boldsymbol{g}_{22} & \boldsymbol{g}_{23} \\ 0 & \boldsymbol{g}_{32} & \boldsymbol{g}_{33} \end{pmatrix}.$$

Then the maximum number of effective lines for G in general position is at most 3. *In particular, if $\Lambda_{\mathrm{Myr}}(G) \neq \mathbb{P}^2_{\mathbb{C}}$, then*

$$\mu(\mathrm{Eq}(G)) \leq 3.$$

Proof. If $(g_n) \subset G$ is a sequence of distinct elements in G such that $g_n \to S$ in $\mathrm{SP}(3, \mathbb{C})$ as $n \to \infty$, then S is induced by a matrix of the form

$$\mathbf{0} \neq \boldsymbol{s} = \begin{pmatrix} \boldsymbol{s}_{11} & 0 & 0 \\ 0 & \boldsymbol{s}_{22} & \boldsymbol{s}_{23} \\ 0 & \boldsymbol{s}_{32} & \boldsymbol{s}_{33} \end{pmatrix}.$$

Since G is discrete,

$$\boldsymbol{s}_{11}(\boldsymbol{s}_{22}\boldsymbol{s}_{33} - \boldsymbol{s}_{23}\boldsymbol{s}_{32}) = 0.$$

Hence, $\mathrm{Ker}(S)$ is equal to the point e_1, a line passing through e_1, the line $\overleftrightarrow{e_2, e_3}$, or a point in $\overleftrightarrow{e_2, e_3}$. □

Definition 6.4. If $U \subset \mathbb{P}^2_{\mathbb{C}}$ is an open set and $\mathcal{L}$ is a set of lines in general position contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$, then we say the point $v \in \mathbb{P}^2_{\mathbb{C}} \setminus U$ is a *vertex for U and $\mathcal{L}$* whenever:

- There are infinitely many lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ passing through v.
- There exist two distinct lines ℓ_1, ℓ_2 in $\mathcal{L}$ passing through v.

Proposition 6.5. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete group and $U \subset \mathbb{P}^2_{\mathbb{C}}$ be one of* $\mathrm{Eq}(G)$, $\Omega(G)$, *or* $U_{\max}(G)$. *If $4 \leq \mu(U) < \infty$ then, for each set $\mathcal{L}$ consisting of lines in general position contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ such that $|\mathcal{L}| = \mu(U)$:*

(i) *There are precisely two vertices for U and $\mathcal{L}$.*

(ii) *If ℓ is a line not containing vertices for U, then the G-orbit of ℓ is infinite.*

Proof. If U is equal to $\Omega(G)$ or $U_{\max}(G)$, then the hypothesis $\mu(U) \geq 4$ and Theorem 3.4(iii) imply that $U = \mathrm{Eq}(G)$. Hence it suffices to prove the lemma in the case $U = \mathrm{Eq}(G)$.

(i) Since $|\mathcal{L}| = \mu(U) < \infty$, every line contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ passes through an intersection point of lines in $\mathcal{L}$. By Proposition 5.15, $\lambda(U) = \infty$. Hence, *there exists at least one vertex for U and $\mathcal{L}$.*

Given that $\mu(U) < \infty$, the set of vertices for U and $\mathcal{L}$ is finite and it is G-invariant. It follows that *the isotropy subgroup of any vertex for U and $\mathcal{L}$ has finite index in G.*

Since $\mu(U) \geq 4$, $\mathbb{P}^2_{\mathbb{C}} \setminus U$ is a union of lines (see Corollary 4.5). If we assume that there is only one vertex for U and $\mathcal{L}$ then

$$\mathbb{P}^2_{\mathbb{C}} \setminus U = \Big(\bigcup_{\ell \in \mathcal{B}} \ell\Big) \cup \Big(\bigcup_{\ell \in \mathcal{A}} \ell\Big),$$

where $\mathcal{B}$ is the closed set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ passing through the vertex, and $\mathcal{A}$ is the set of lines contained in $\mathbb{P}^2_{\mathbb{C}} \setminus U$ not passing through the vertex.

Since there is only one vertex, the G-invariant set $\mathcal{A}$ contains finitely many lines. If $\ell_0 \in \mathcal{A}$ then the subgroup $G_0 = \mathrm{Isot}(\ell_0, G)$ has finite index in G, so the open set $U_0 = \mathbb{P}^2_{\mathbb{C}} \setminus \left(\left(\bigcup_{\ell \in \mathcal{B}} \ell\right) \cup \ell_0\right)$ is G_0-invariant. It follows from Theorem 3.4 iii) that

$$U_0 \subset \mathrm{Eq}(G_0) = \mathrm{Eq}(G).$$

Thus, $3 = \mu(U_0) = \mu(\mathrm{Eq}(G)) \geq 4$, a contradiction. Therefore, *there exist at least two vertices for U and $\mathcal{L}$.*

If we assume that the vertices for U and $\mathcal{L}$ do not lie on a line, then there are three vertices of U and $\mathcal{L}$ in general position. Moreover, we can assume that $\{e_1, e_2, e_3\}$ are those vertices. Thus every element in the finite index subgroup

$$G_1 = \bigcap_{j=1}^{3} \mathrm{Isot}(e_j, G) \subset G$$

is induced by a diagonal matrix. It follows from Lemma 6.3 that

$$3 \geq \mu(\mathrm{Eq}(G_1)) = \mu(\mathrm{Eq}(G)) \geq 4,$$

a contradiction. Therefore, *the vertices for U and $\mathcal{L}$ lie in a complex line.*

If we assume that there are more than two vertices for U and $\mathcal{L}$ then there exist three distinct vertices, v_1, v_2, v_3 for U and $\mathcal{L}$ contained in a line ℓ. The finite index subgroup

$$G_1 = \bigcap_{j=1}^{3} \mathrm{Isot}(v_j, G) \subset G$$

fixes three distinct points in the line ℓ, so it acts trivially on ℓ. It follows from Proposition 6.2 that

$$\mu(\mathrm{Eq}(G_1)) = 1,$$

contradicting the fact that $\mathrm{Eq}(G_1) = \mathrm{Eq}(G)$ and $\mu(\mathrm{Eq}(G)) \geq 4$. Therefore, *there are precisely two vertices for U and $\mathcal{L}$.*

(ii) If we assume there exists a line ℓ_0 with finite G-orbit and not passing through any vertex v_1 or v_2 for U, then

$$G_2 = \mathrm{Isot}(\ell_0, G) \cap \mathrm{Isot}(v_1, G) \cap \mathrm{Isot}(v_2, G)$$

is a finite index subgroup of G fixing the points

$$v_1,\ v_2,\ \ell_0 \cap \overleftrightarrow{v_1, v_2}.$$

Thus, G_2 acts trivially on $\overleftrightarrow{v_1, v_2}$, so $\mathrm{Eq}(G) = \mathrm{Eq}(G_2)$ is the complement of the union of a line and a point (by Proposition 6.2), contradicting the hypothesis that $\mu(\mathrm{Eq}(G)) \geq 4$. □

Proof of Theorem 1.2. On the contrary, let us assume that $4 < \mu(U) < \infty$. Then there is a finite set of lines in general position, $\mathcal{L}$, such that $\ell \subset \mathbb{P}^2_{\mathbb{C}} \setminus U$ for every $\ell \in \mathcal{L}$, and $|\mathcal{L}| = \mu(U)$. By Proposition 6.5(i), there are precisely two vertices for U and $\mathcal{L}$. Let us denote by v_1 and v_2 these two vertices. Since $\mu(U) > 4$, there is a line in $\mathcal{L}$ not passing through v_1 nor v_2. By Proposition 6.5(ii), this line has infinite G-orbit, then there is another vertex for U and $\mathcal{L}$ distinct from v_1, v_2, a contradiction. Therefore, $\mu(U)$ is equal to 1, 2, 3, 4, or ∞. In Sections 7A to 7E we give examples of infinite discrete subgroups $G \subset \mathrm{PSL}(3, \mathbb{C})$ with corresponding open sets U satisfying $\mu(U) \in \{1, 2, 3, 4, \infty\}$. □

7. Examples

7A. *One line complex Kleinian groups.* (i) Suppose that G is the cyclic subgroup of $\mathrm{PSL}(3, \mathbb{C})$ generated by a complex homothety g_0 induced by a matrix of the form

$$\begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a^{-2} \end{pmatrix}, \qquad \text{where } 0 < |a| < 1.$$

Then

$$\Omega(G) = \mathrm{Eq}(G) = \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_2} \cup \{e_3\})$$

is the maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously. Hence we have the table:

	$\Omega(G)$	$\mathrm{Eq}(G)$	$U_{\max}(G)$
λ	1	1	1
μ	1	1	1

(ii) If $G \subset \mathrm{PGL}(3, \mathbb{C})$ is the cyclic group generated by the loxoparabolic element induced by the matrix

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad 0 < |a| < 1,$$

then

$$\Omega(G) = \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_2} \cup \overleftrightarrow{e_1, e_3}) = \mathrm{Eq}(G)$$

However, G acts properly and discontinuously on $\mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_1, e_2}$; see [Barrera et al. 2014a, Example 2.3]. Moreover, $U_{\max} = \mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_1, e_2}$ is the maximal open subset of

$\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously. Hence we have the table:

	$\Omega(G)$	$\mathrm{Eq}(G)$	$U_{\max}(G)$
λ	2	2	1
μ	2	2	1

(iii) The abelian group G, generated by two projective transformations induced by the matrices in $\mathrm{GL}(3, \mathbb{C})$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & i \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{where } 0 < |a| < 1,$$

satisfies the property that

$$\Omega(G) = \mathbb{P}^2_{\mathbb{C}} \setminus \overleftrightarrow{e_1, e_2}$$

is the maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$, where G acts properly and discontinuously. On the other hand,

$$\mathrm{Eq}(G) = \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_2} \cup \overleftrightarrow{e_1, e_3}),$$

and we have the table:

	$\Omega(G)$	$\mathrm{Eq}(G)$	$U_{\max}(G)$
λ	1	2	1
μ	1	2	1

Those subgroups of $\mathrm{PGL}(3, \mathbb{C})$ whose Kulkarni limit set is equal to one line are classified in [Barrera et al. 2014a, Theorem 1.1].

If $\Lambda_{\mathrm{Kul}}(G)$ is equal to one line then $\Omega(G)$ is a maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously. (If an infinite subgroup $G \subset \mathrm{PGL}(3, \mathbb{C})$ acts properly and discontinuously on the open set $U \subset \mathbb{P}^2_{\mathbb{C}}$ then $\lambda(U) \geq 1$.)

Conversely, if $\mathbb{P}^2_{\mathbb{C}} \setminus \ell$ is maximal open set where G acts properly and discontinuously, then $\Lambda_{\mathrm{Kul}}(G)$ is equal to one line, except in the case when G contains a cyclic subgroup of finite index generated by a loxoparabolic element; see [op. cit., Theorem 1.2].

In the case when $\Lambda_{\mathrm{Myr}}(G)$ is equal to one line, ℓ, then G does not contain loxoparabolic elements and it acts properly and discontinuously on $\mathbb{P}^2_{\mathbb{C}} \setminus \ell$, then $\Lambda_{\mathrm{Kul}}(G) = \ell$ by the same theorem.

7B. ***Two line complex Kleinian groups.*** In this section we give some examples of complex Kleinian groups such that $\mu(\Omega(G)) = 2$, $\mu(\mathrm{Eq}(G)) = 2$, or $\mu(U_{\max}(G)) = 2$.

(i) If $g \in \mathrm{PGL}(3, \mathbb{C})$ is an element induced by a matrix of the form

$$\boldsymbol{g} = \begin{pmatrix} \boldsymbol{g}_{11} & 0 & 0 \\ 0 & \boldsymbol{g}_{22} & 0 \\ 0 & 0 & \boldsymbol{g}_{33} \end{pmatrix}, \quad \text{where } |\boldsymbol{g}_{11}| < |\boldsymbol{g}_{22}| < |\boldsymbol{g}_{33}|,$$

then the cyclic group $G = \langle g \rangle$ satisfies

$$\Omega(G) = \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_2} \cup \overleftrightarrow{e_2, e_3}) = \mathrm{Eq}(G).$$

On the other hand, $U_{\max}(G) = \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_2} \cup \{e_3\})$ is a maximal open set where G acts properly and discontinuously. Thus we have:

	$\Omega(G)$	$\mathrm{Eq}(G)$	$U_{\max}(G)$
λ	2	2	1
μ	2	2	1

(ii) Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be the group induced by matrices of the form:

$$\begin{pmatrix} 1 & m & n \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{where } m, n \in \mathbb{Z}.$$

This group contains only parabolic elements (except for the identity) and satisfies that $U = \Omega(G) = \mathrm{Eq}(G)$ is the maximal open set where G acts properly and discontinuously. Moreover, we have:

	$\Omega(G)$	$\mathrm{Eq}(G)$	$U_{\max}(G)$
λ	∞	∞	∞
μ	2	2	2

(iii) *The double suspension construction*. Given a subgroup $G \subset \mathrm{PSL}(2, \mathbb{C})$ we can construct a new group $\hat{G} \subset \mathrm{PSL}(3, \mathbb{C})$, called *the double suspension of* G, acting on $\mathbb{P}^2_{\mathbb{C}}$ in such way that the restriction of this action to the line at infinity is the action of G on $\mathbb{P}^1_{\mathbb{C}} \cong S^2$. See [Navarrete 2008; Seade and Verjovsky 2001]. The elements in $\hat{G}$ are represented by all matrices of the form

$$\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{where } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C}) \text{ induces an element in } G.$$

In other words, $\hat{G}$ is a double covering of G. Moreover, when G is a classical Kleinian group with limit set $L(G)$, then $\Lambda_{\mathrm{Kul}}(\hat{G}) = \Lambda_{\mathrm{Myr}}(\hat{G})$ is equal to the complex cone with vertex e_3 and base $L(G)$ (considered as a subset of the line at infinity $\overleftrightarrow{e_1, e_2}$). In symbols,

$$\Lambda_{\mathrm{Kul}}(\hat{G}) = \Lambda_{\mathrm{Myr}}(\hat{G}) = \bigcup_{x \in L(G)} \overleftrightarrow{e_3, x}.$$

If G is a nonelementary classical Kleinian group, then $\Omega(\hat{G})$ is the maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$ where $\hat{G}$ acts properly and discontinuously. Thus:

	$\Omega(\hat{G})$	$\mathrm{Eq}(\hat{G})$	$U_{\max}(\hat{G})$
λ	∞	∞	∞
μ	2	2	2

7C. ***Three line complex Kleinian groups.*** (i) Let G be the group generated by $A, B \in \mathrm{PSL}(3, \mathbb{C})$ where A and B are induced by the matrices

$$\boldsymbol{A} = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then

$$\Omega(G) = \mathbb{P}^2_{\mathbb{C}} \setminus (\overleftrightarrow{e_1, e_2} \cup \overleftrightarrow{e_2, e_3} \cup \overleftrightarrow{e_3, e_1}) = \mathrm{Eq}(G)$$

is a maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously; see [Barrera et al. 2011a, Example 4.3]. It follows that:

	$\Omega(G)$	$\mathrm{Eq}(G)$	$U_{\max}(G)$
λ	3	3	3
μ	3	3	3

(ii) [Cano et al. 2013, Subsection 5.5.1] If $G \subset \mathrm{PSL}(2, \mathbb{C})$ is a Kleinian group and $D \subset \mathbb{C}^*$ a discrete subgroup, then the *suspension of G extended by the group D*, denoted by $\mathrm{Susp}(G, D)$ is the group generated by the double suspension and all the elements in $\mathrm{PSL}(3, \mathbb{C})$ induced by diagonal matrices of the form:

$$\begin{pmatrix} d & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d^{-2} \end{pmatrix}, \quad \text{where } d \in D.$$

In the case when $G \subset \mathrm{PSL}(2, \mathbb{C})$ is a nonelementary Kleinian group and $D \subset \mathbb{C}^*$ is an infinite discrete subgroup,

$$\Lambda_{\mathrm{Kul}}(\mathrm{Susp}(G, D)) = \overleftrightarrow{e_1, e_2} \cup \left(\bigcup_{x \in L(G)} \overleftrightarrow{x, e_3} \right) = \Lambda_{\mathrm{Myr}}(\mathrm{Susp}(G, D)),$$

and $\Omega(\mathrm{Susp}(G,D)) = \mathrm{Eq}(\mathrm{Susp}(G,D))$ is the maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$ where $\mathrm{Susp}(G,D)$ acts properly discontinuously. Therefore:

	$\Omega(\mathrm{Susp}(G,D))$	$\mathrm{Eq}(\mathrm{Susp}(G,D))$	$U_{\max}(\mathrm{Susp}(G,D))$
λ	∞	∞	∞
μ	3	3	3

7D. ***Four line complex Kleinian groups.*** See [Barrera et al. 2011b].

An element $A \in \mathrm{SL}(2, \mathbb{Z})$, is called a *hyperbolic toral automorphism* if none of its eigenvalues lie on the unit circle. Any subgroup of $\mathrm{PSL}(3, \mathbb{C})$ conjugate to the group

$$G_A = \left\{ \begin{pmatrix} A^k & \boldsymbol{b} \\ \boldsymbol{0} & 1 \end{pmatrix} : \boldsymbol{b} \in M(2\times 1, \mathbb{Z}),\ k \in \mathbb{Z} \right\},$$

where $A \in \mathrm{SL}(2, \mathbb{Z})$ is a hyperbolic toral automorphism, is called a *hyperbolic toral group*.

Theorem 7.1 [Barrera et al. 2011b]. *Let $G \subset \mathrm{PSL}(3, \mathbb{C})$ be a discrete group. The maximum number of complex lines in general position contained in Kulkarni's limit set is equal to four if and only if G contains a hyperbolic toral group whose index is at most eight.*

Furthermore, it is proved that $\Omega(G) = \mathrm{Eq}(G)$. However $\Omega(G)$ is not the maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously. It can be shown that there exist two open maximal sets $U^{(1)}_{\max}$, $U^{(2)}_{\max} \subset \mathbb{P}^2_{\mathbb{C}}$ where G acts properly and discontinuously.

For the reader's convenience, we give a brief outline of the proof that the group G_A has the properties mentioned above. Let $A \in \mathrm{SL}(2, \mathbb{Z})$ be a hyperbolic toral automorphism. We define

$$T = \begin{pmatrix} \boldsymbol{t} & \boldsymbol{0} \\ 0 & 1 \end{pmatrix}, \quad \text{where } \boldsymbol{t} = \begin{pmatrix} 1 & u \\ v & 1 \end{pmatrix} \text{ with } u, v \in \mathbb{R} - \mathbb{Q}.$$

Hence the group $\hat{G}_A = T G_A T^{-1}$ is equal to

$$\begin{pmatrix} \alpha^n & 0 & ky_0 + lx_0 \\ 0 & \alpha^{-n} & kx_0 + lz_0 \\ 0 & 0 & 1 \end{pmatrix}$$

where $k, l, n \in \mathbb{Z}$ and

$$x_0 = \frac{-1}{uv-1}, \quad y_0 = \frac{u}{uv-1}, \quad z_0 = \frac{v}{uv-1}.$$

By straightforward computations, the Kulkarni limit set is

$$\overleftrightarrow{e_1, e_2} \cup \bigcup_{r\in\mathbb{R}} \overleftrightarrow{e_1, [0:r:1]} \cup \overleftrightarrow{e_2, [r:0:1]}$$

It is not hard to check that the set of vertices is $\{e_1, e_2\}$ and $\mu(\hat{G}_A) = 4$, which implies that $\Omega(\hat{G}_A) = \mathrm{Eq}(\hat{G}_A)$. The group $\hat{G}_A$ acts properly and discontinuously on $U^{(i)}_{\max}(\hat{G}_A) = \mathbb{P}^2_{\mathbb{C}} - \mathcal{C}_i$, $i = 1, 2$, where

$$\mathcal{C}_1 = \overleftrightarrow{e_1, e_2} \cup \bigcup_{r\in\mathbb{R}} \overleftrightarrow{e_1, [0:r:1]} \quad \text{and} \quad \mathcal{C}_2 = \overleftrightarrow{e_1, e_2} \cup \bigcup_{r\in\mathbb{R}} \overleftrightarrow{e_2, [r:0:1]}$$

are maximal regions where the group $\hat{G}_A$ acts properly and discontinuously. In summary, we have the following table:

	$\Omega(G)$	$\mathrm{Eq}(G)$	$U^{(1)}_{\max}(G)$	$U^{(2)}_{\max}(G)$
λ	∞	∞	∞	∞
μ	4	4	2	2

7E. ***Complex Kleinian groups with infinitely many lines.*** (i) If $G \subset \mathrm{PU}(2, 1)$ then it is proved in [Navarrete 2006] that

$$\Lambda_{\mathrm{Kul}}(G) = \bigcup_{z\in L(G)} \ell_z,$$

where $L(G) \subset \partial\mathbb{H}^2_{\mathbb{C}} = S^3$ denotes the Chen–Greenberg limit set of G considered as acting on $\mathbb{H}^2_{\mathbb{C}}$ by holomorphic isometries, and ℓ_z is the only tangent line to $\partial\mathbb{H}^2_{\mathbb{C}}$ at z. The Kulkarni region of discontinuity, $\Omega(G)$, is the maximal open subset where G acts properly and discontinuously, whenever $\lambda(\Omega(G)) > 2$. Moreover, it is proved in [Cano and Seade 2010] that

$$\Lambda_{\mathrm{Myr}}(G) = \bigcup_{z\in L(G)} \ell_z.$$

In the case when G satisfies $L(G) = \partial\mathbb{H}^2_{\mathbb{C}}$,

$$\Omega(G) = \mathbb{H}^2_{\mathbb{C}} = \mathrm{Eq}(G)$$

is the maximal open subset where G acts properly and discontinuously. Thus we have the table:

	$\Omega(G)$	$\mathrm{Eq}(G)$	$U_{\max}(G)$
λ	∞	∞	∞
μ	∞	∞	∞

If $G \subset \mathrm{PU}(2, 1)$ satisfies the property that $L(G)$ is an $\mathbb{R}$-circle, then we have the same table as above; see [Cano et al. $\geq$ 2016].

(ii) In [Barrera et al. 2014b], it is shown that a family of complex Kleinian groups $G_n \subset \mathrm{PSL}(3, \mathbb{R})$ exists, such that for all $n \in \mathbb{N}$, G_n is a free group, not conjugate in $\mathrm{PSL}(3, \mathbb{C})$ to any subgroup of $\mathrm{PU}(2, 1)$. G_n has no invariant lines nor fixed points.

The Kulkarni limit set $\Lambda_{\mathrm{Kul}}(G_n)$ contains at least five complex projective lines in general position. Hence it contains infinitely many complex projective lines in general position. Moreover, $\Omega(G_n) = \mathrm{Eq}(G_n)$ is the maximal open subset of $\mathbb{P}^2_{\mathbb{C}}$, where G_n acts properly and discontinuously. Thus we have the table:

	$\Omega(G_n)$	$\mathrm{Eq}(G_n)$	$U_{\max}(G_n)$
λ	∞	∞	∞
μ	∞	∞	∞

References

[Barrera et al. 2011a] W. Barrera, A. Cano, and J. P. Navarrete, "The limit set of discrete subgroups of $\mathrm{PSL}(3, \mathbb{C})$", *Math. Proc. Cambridge Philos. Soc.* **150**:1 (2011), 129–146. MR 2012b:32037 Zbl 1214.30032

[Barrera et al. 2011b] W. Barrera, A. Cano, and J. P. Navarrete, "Subgroups of $PSL(3, \mathbb{C})$ with four lines in general position in its limit set", *Conform. Geom. Dyn.* **15** (2011), 160–176. MR 2012i:37067 Zbl 1252.37038

[Barrera et al. 2014a] W. Barrera, A. Cano, and J. P. Navarrete, "One line complex Kleinian groups", *Pacific J. Math.* **272**:2 (2014), 275–303. MR 3284888 Zbl 06406046

[Barrera et al. 2014b] W. Barrera, A. Cano, and J. P. Navarrete, "Pappus' theorem and a construction of complex Kleinian groups with rich dynamics", *Bull. Braz. Math. Soc.* (*N.S.*) **45**:1 (2014), 25–52. MR 3194081 Zbl 06307360

[Cano and Seade 2010] A. Cano and J. Seade, "On the equicontinuity region of discrete subgroups of $\mathrm{PU}(1, n)$", *J. Geom. Anal.* **20**:2 (2010), 291–305. MR 2011c:32045 Zbl 1218.37059

[Cano et al. 2013] A. Cano, J. P. Navarrete, and J. Seade, *Complex Kleinian groups*, Progress in Mathematics **303**, Birkhäuser, Basel, 2013. MR 2985759 Zbl 1267.30001

[Cano et al. $\geq$ 2016] A. Cano, J. Parker, and J. Seade, "Actions of $\mathbb{R}$-Fuchsian groups on $\mathbb{CP}^2$", preprint. To appear in *Asian J. Math.*

[Kulkarni 1978] R. S. Kulkarni, "Groups with domains of discontinuity", *Math. Ann.* **237**:3 (1978), 253–272. MR 81m:30046 Zbl 0369.20028

[Maskit 1988] B. Maskit, *Kleinian groups*, Grundlehren der Mathematischen Wissenschaften **287**, Springer, Berlin, 1988. MR 90a:30132 Zbl 0627.30039

[Myrberg 1925] P. J. Myrberg, "Untersuchungen über die automorphen Funktionen beliebig vieler Variablen", *Acta Math.* **46**:3-4 (1925), 215–336. MR 1555203 JFM 51.0298.02

[Navarrete 2006] J. P. Navarrete, "On the limit set of discrete subgroups of $\mathrm{PU}(2, 1)$", *Geom. Dedicata* **122** (2006), 1–13. MR 2008i:32035 Zbl 1131.32013

[Navarrete 2008] J. P. Navarrete, "The trace function and complex Kleinian groups in $\mathbb{P}^2_{\mathbb{C}}$", *Internat. J. Math.* **19**:7 (2008), 865–890. MR 2009g:32056 Zbl 1167.30025

[Seade and Verjovsky 2001] J. Seade and A. Verjovsky, "Actions of discrete groups on complex projective spaces", pp. 155–178 in *Laminations and foliations in dynamics, geometry and topology* (Stony Brook, NY, 1998), edited by M. Lyubich et al., Contemporary Mathematics **269**, American Mathematical Society, Providence, RI, 2001. MR 2002d:32024 Zbl 1161.32301

Received December 19, 2014. Revised July 14, 2015.

Waldemar Barrera
Facultad de Matemáticas
Universidad Autónoma de Yucatán
Anillo Periférico Norte Tablaje Cat
13615 Mérida
Mexico
bvargas@uady.mx

Angel Cano
Instituto de Matemáticas
Universidad Nacional Autónoma de México
Av. Universidad S/N
Col. Lomas de Chamilpa, C.P.
62210 Cuernavaca
Mexico
angelcano@im.unam.mx
angel@matcuer.unam.mx

Juán Navarrete
Facultad de Matemáticas
Universidad Autónoma de Yucatán
Anillo Periférico Norte Tablaje Cat
13615 Mérida
Mexico
jp.navarrete@uady.mx

PACIFIC JOURNAL OF MATHEMATICS
Vol. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.51

GALOIS THEORY, FUNCTIONAL LINDEMANN–WEIERSTRASS, AND MANIN MAPS

DANIEL BERTRAND AND ANAND PILLAY

We prove several new results of Ax–Lindemann type for semiabelian varieties over the algebraic closure K of $\mathbb{C}(t)$, making heavy use of the Galois theory of logarithmic differential equations. Using related techniques, we also give a generalization of the theorem of the kernel for abelian varieties over K. This paper is a continuation of earlier work by Bertrand and Pillay (2010), as well as an elaboration on the methods of Galois descent introduced by Bertrand (2009, 2011) in the case of abelian varieties.

1. Introduction

This paper has three related themes, the common feature being differential Galois theory and its applications.

Firstly, given a semiabelian variety B over the algebraic closure K of $\mathbb{C}(t)$, a K-rational point a of the Lie algebra LG of its universal vectorial extension $G = \widetilde{B}$, and a solution $y \in G(K^{\text{diff}})$ of the logarithmic differential equation

$$\partial\ell n_G(y) = a, \quad a \in LG(K),$$

we want to describe $\text{tr.deg}\big(K_G^\sharp(y)/K_G^\sharp\big)$ in terms of gauge transformations *over K itself*. Here $K_G^\sharp$ is the differential field generated over K by solutions of $\partial\ell n_G(-) = 0$ in K^{diff}. Introducing this field as base presents both advantages and difficulties. On the one hand, it allows us to use the differential Galois theory developed by

---

Bertrand thanks the ANR-13-BS01-0006 for support, as well as MSRI, Berkeley.
Pillay thanks the EPSRC grant EP/I002294/1, MSRI, Berkeley, as well as NSF grant DMS-1360702.
MSC2010: 03C60, 11J95, 12H05, 14K05, 34M15.
Keywords: differential Galois theory, algebraic independence over function fields, semiabelian schemes, Manin maps.

Pillay [1998; 1997; 2004], thereby replacing the study of transcendence degrees by the computation of a Galois group. On the other hand, we have only a partial knowledge of the extension $K_G^\sharp/K$. However, it was observed by Bertrand [2009; 2011] that in the case of an abelian variety, what we do know essentially suffices to perform a Galois descent from $K_G^\sharp$ to the field K of the desired gauge transformation. In Sections 2B and 3 of the present paper, we extend this principle to semiabelian varieties B whose toric part is $\mathbb{G}_m$, and give a definitive description of $\mathrm{tr.deg}\big(K_G^\sharp(y)/K_G^\sharp\big)$ when B is an abelian variety.

The main application we have in mind of these Galois-theoretic results forms the second theme of our paper, and concerns Lindemann–Weierstrass statements for the semiabelian variety B over K, by which we mean the description of the transcendence degree of $\exp_B(x)$ where x is a K-rational point of the Lie algebra LB of B. The problem is covered in the above setting by choosing as data

$$a := \partial_{LG}(\tilde{x}) \in \partial_{LG}(LG(K)),$$

where $\tilde{x}$ is an arbitrary K-rational lift of x to $G = \widetilde{B}$. This study was initiated in our joint paper [2010], where the Galois approach was mentioned, but only under the hypothesis that $K_G^\sharp = K$, described as *K-largeness* of G. There are natural conjectures in analogy with the well-known "constant" case (where B is over $\mathbb{C}$), although as pointed out in [Bertrand and Pillay 2010], there are also counterexamples provided by nonconstant extensions of a constant elliptic curve by the multiplicative group. In Sections 2C and 4 of this paper, we extend the main result of [Bertrand and Pillay 2010] to the base $K_G^\sharp$, but assuming the toric part of B is at most 1-dimensional. Furthermore, we give in this case a full solution of the Lindemann–Weierstrass statement when the abelian quotient of B is also 1-dimensional. This uses results from [Bertrand et al. 2013] which deal with the "logarithmic" case. In this direction, we will also formulate an "Ax–Schanuel" type conjecture for abelian varieties over K.

The third theme of the paper concerns the "theorem of the kernel", which we generalize in Sections 2D and 5 by proving that linear independence with respect to $\mathrm{End}(A)$ of points $y_1, \ldots, y_n$ in $A(K)$ implies linear independence of

$$\mu_A(y_1), \ldots, \mu_A(y_n)$$

with respect to $\mathbb{C}$ (this answers a question posed to us by Hrushovski). Here A is an abelian variety over $K = \mathbb{C}(t)^{\mathrm{alg}}$ with $\mathbb{C}$-trace 0 and μ_A is the differential-algebraic Manin map. However, we will give an example showing that its $\mathbb{C}$-linear extension $\mu_A \otimes 1$ on $A(K) \otimes_{\mathbb{Z}} \mathbb{C}$ is not always injective. In contrast, we observe that the $\mathbb{C}$-linear extension $M_{K,A} \otimes 1$ of the classical (differential-arithmetic) Manin map $M_{K,A}$ is always injective. Differential Galois theory and the logarithmic case of nonconstant Ax–Schanuel are involved in the proofs.

2. Statements of results

2A. ***Preliminaries on logarithmic equations.*** We give here a quick background to the basic notions and objects so as to be able to state our main results in the next subsections. The remaining Sections 3, 4, and 5 of the paper are devoted to the proofs. We refer the reader to [Bertrand and Pillay 2010] for more details including differential algebraic preliminaries.

We fix a differential field (K, ∂) of characteristic 0 whose field of constants C_K is algebraically closed (the reader will lose nothing by taking $C_K = \mathbb{C}$). We usually assume that K is algebraically closed, and denote by K^{diff} the differential closure of K. We let $\mathcal{U}$ denote a universal differential field containing K, with constant field $\mathcal{C}$. If X is an algebraic variety over K we will identify X with its set $X(\mathcal{U})$ of $\mathcal{U}$ points.

We start with algebraic ∂-groups, which provide the habitat of the (generalized) differential Galois theory of [Pillay 1998; 1997; 2004] discussed later on. A (connected) *algebraic ∂-group* over K is a (connected) algebraic group G over K together with a lifting D of the derivation ∂ of K to a derivation of the structure sheaf $\mathbb{O}_G$ which respects the group structure. The derivation D may be identified with a section s, in the category of algebraic groups, of the projection map $T_\partial(G) \to G$, where $T_\partial(G)$ denotes the twisted tangent bundle of G. This $T_\partial(G)$ is a (connected) algebraic group over K, which is a torsor under the tangent bundle TG, and is locally defined by equations

$$\sum_{i=1}^{n} \frac{\partial P}{\partial x_i}(\bar{x})u_i + P^\partial(\bar{x}) = 0,$$

for polynomials P in the ideal of G, where P^∂ is obtained by applying the derivation ∂ of K to the coefficients of P. Notice for later use that for any differential extension L/K, there is a group homomorphism $G(L) \to T_\partial G(L)$, which is given in coordinates by $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n, \partial x_1, \ldots, \partial x_n)$ and will be denoted by ∂.

We write the algebraic ∂-group as (G, D) or (G, s). Not every algebraic group over K has a ∂-structure. But when G is defined over the constants C_K of K, there is a privileged ∂-structure s_0 on G which is precisely the 0-section of $TG = T_\partial G$. Given an algebraic ∂-group (G, s) over K we obtain an associated "logarithmic derivative" $\partial\ell n_{G,s}(-)$ from G to the Lie algebra LG of G defined by $\partial\ell n_{G,s}(y) = \partial(y)s(y)^{-1}$, where the product is computed in the algebraic group $T_\partial(G)$. This is a differential rational crossed homomorphism from G onto LG (at the level of $\mathcal{U}$-points or points in a differentially closed field) defined over K. Its kernel $\ker(\partial\ell n_{G,s})$ is a differential algebraic subgroup of G which we denote $(G, s)^\partial$, or simply G^∂ when the context is clear. Now s equips the Lie algebra LG of G with its own structure of a ∂-group (in this case a ∂-module) which we call ∂_{LG} (depending on (G, s)) and again the kernel is denoted $(LG)^\partial$.

In the case where G is defined over C_K and $s = s_0$, the map $\partial\ell n_{G,s}$ is precisely Kolchin's logarithmic derivative, taking $y \in G$ to $\partial(y)y^{-1}$. In general, as soon as s is understood, we will abbreviate $\partial\ell n_{G,s}$ by $\partial\ell n_G$.

By a *logarithmic differential equation* over K on the algebraic ∂-group (G, s), we mean a differential equation $\partial\ell n_{G,s}(y) = a$ for some $a \in LG(K)$. When $G = \mathrm{GL}_n$ and $s = s_0$ this is the equation for a fundamental system of solutions of a linear differential equation $Y' = aY$ in vector form. And more generally, for G an algebraic group over C_K and $s = s_0$, this is a logarithmic differential equation on G over K in the sense of Kolchin. There is a well-known Galois theory here. In the given differential closure K^{diff} of K, any two solutions y_1, y_2 of $\partial\ell n_G(-) = a$ in $G(K^{\mathrm{diff}})$ differ by an element in the kernel G^∂ of $\partial\ell n_G(-)$. But $G^\partial(K^{\mathrm{diff}})$ is precisely $G(C_K)$. Hence $K(y_1) = K(y_2)$. In particular, tr.deg$(K(y)/K)$ is the same for all solutions y in K^{diff}. Moreover, Aut$(K(y)/K)$ has the structure of an algebraic subgroup of $G(C_K)$: for any $\sigma \in \mathrm{Aut}(K(y)/K)$, let $\rho_\sigma \in G(C_K)$ be such that $\sigma(y) = y\rho_\sigma$. Then the map taking σ to ρ_σ is an isomorphism between Aut$(K(y)/K)$ and an algebraic subgroup $H(C_K)$ of $G(C_K)$, which we call the differential Galois group of $K(y)/K$. This depends on the choice of solution y, but another choice yields a conjugate of H. Of course when G is commutative, H is independent of the choice of y. In any case tr.deg$(K(y)/K) = \dim(H)$, so computing the differential Galois group gives us a transcendence estimate.

Continuing with this Kolchin situation, we have the following well-known fact, whose proof we present in the setting of the more general situation considered in Fact 2.2(i).

Fact 2.1 (for G/C_K). *Suppose K algebraically closed. Then*, tr.deg$(K(y)/K)$ *is the dimension of a minimal connected algebraic subgroup H of G, defined over C_K, such that for some $g \in G(K)$, $gag^{-1} + \partial\ell n_G(g) \in LH(K)$. Moreover*, $H(C_K)$ *is the differential Galois group of $K(y)/K$.*

Proof. Let H be a connected algebraic subgroup of G, defined over C_K such that $H^\partial(K^{\mathrm{diff}}) = H(C_K)$ is the differential Galois group of $K(y)$ over K. Now the $H^\partial(K^{\mathrm{diff}})$-orbit of y is defined over K in the differential algebraic sense, so the H-orbit of y is defined over K in the differential algebraic sense. A result of Kolchin on constrained cohomology (see Proposition 3.2 of [Pillay 1998], or Theorem 2.2 of [Bertrand 2011]) implies that this orbit has a K-rational point g^{-1}. So, there exists $z^{-1} \in H$ such that $g^{-1} = yz^{-1}$, and $z = gy$, which satisfies $K(y) = K(z)$, is a solution of $\partial\ell n_G(-) = a'$ where $a' = gag^{-1} + \partial\ell n_G(g)$. □

(Such a map $LG(K) \to LG(K)$ taking $a \in LG(K)$ to $gag^{-1} + \partial\ell n_G(g)$ for some $g \in G(K)$ is called a gauge transformation.)

Now in the case of an arbitrary algebraic ∂-group (G, s) over K, and logarithmic differential equation $\partial\ell n_{G,s}(-) = a$ over K, two solutions y_1, y_2 in $G(K^{\mathrm{diff}})$ differ by an element of $(G, s)^{\partial}(K^{\text{diff}})$ which in general may not be contained in $G(K)$. (For instance, if $(G = \mathbb{G}_a, s)$ is the ∂-module attached to $\partial y - y = 0$, and $a = 1 - t$, then $y_1 = t$ is rational over $K = \mathbb{C}(t)$, while $y_2 = t + e^t$ is transcendental over K.) So to obtain both a transcendence statement independent of the choice of solution, as well as a Galois theory, we should work over $K^{\sharp}_{G,s}$ which is the (automatically differential) field generated by K and $(G, s)^{\partial}(K^{\text{diff}})$. This field may be viewed as a field of "new constants", and its algebraic closure in K^{diff} will be denoted by $K^{\sharp\,\text{alg}}_{G,s}$. As with $\partial\ell n_G$ and G^{∂}, we will abbreviate $K^{\sharp}_{G,s}$ as $K^{\sharp}_{G}$, or even $K^{\sharp}$, when the context is clear, and similarly for its algebraic closure.

Fixing a solution $y \in G(K^{\text{diff}})$ of $\partial\ell n_G(-) = a$, for $\sigma \in \text{Aut}(K^{\sharp}(y)/K^{\sharp})$ we have $\sigma(y) = y\rho_\sigma$ for unique $\rho_\sigma \in G^{\partial}(K^{\text{diff}}) = G^{\partial}(K^{\sharp}) \subseteq G(K^{\sharp})$, and again the map $\sigma \mapsto \rho_\sigma$ defines an isomorphism between $\text{Aut}(K^{\sharp}(y)/K^{\sharp})$ and $(H, s)^{\partial}(K^{\text{diff}})$ for an algebraic ∂-subgroup H of (G, s), ostensibly defined over $K^{\sharp}$. The ∂-group H (or more properly H^{∂}, or $H^{\partial}(K^{\sharp})$) is called the *(differential) Galois group* of $K^{\sharp}(y)$ over $K^{\sharp}$, and when G is commutative does not depend on the choice of y, just on the data $a \in LG(K)$ of the logarithmic equation, and in fact only on the image of a in the cokernel $LG(K)/\partial\ell n_G G(K)$ of $\partial\ell n_G$. Again $\text{tr.deg}(K^{\sharp}(y)/K^{\sharp}) = \dim(H)$. In any case, Fact 2.1 extends to this context with essentially the same proof. This can also be extracted from Proposition 3.4 of [Pillay 1998] and the setup of [Pillay 2004]. For the commutative case (part (ii) below) see [Bertrand 2011, Theorem 3.2]. Note that in the present paper, it is this Fact 2.2(ii) we will use. Going to the algebraic closure of $K^{\sharp}$ as in Fact 2.2(i) would force us to consider profinite groups, for which our descent arguments may not work.

Fact 2.2 (for G/K). *Let y be a solution of $\partial\ell n_{G,s}(-) = a$ in $G(K^{\text{diff}})$, and let $K^{\sharp} = K(G^{\partial})$, with algebraic closure $K^{\sharp\,\text{alg}}$. Then the following hold*:

(i) *The transcendence degree* $\text{tr.deg}(K^{\sharp}(y)/K^{\sharp})$ *is the dimension of a minimal connected algebraic ∂-subgroup H of G, which is defined over $K^{\sharp\,\text{alg}}$ such that $gag^{-1} + \partial\ell n_{G,s}(g) \in LH(K^{\sharp\,\text{alg}})$ for some $g \in G(K^{\sharp\,\text{alg}})$. And $H^{\partial}(K^{\sharp\,\text{alg}})$ is the differential Galois group of $K^{\sharp\,\text{alg}}(y)/K^{\sharp\,\text{alg}}$.*

(ii) *Suppose that G is commutative. Then the identity component of the differential Galois group of $K^{\sharp}(y)/K^{\sharp}$ is $H^{\partial}(K^{\sharp})$, where H is the smallest algebraic ∂-subgroup of G defined over $K^{\sharp}$ such that $a \in LH + \mathbb{Q} \cdot \partial\ell n_{G,s} G(K^{\sharp})$.*

Remark. We point out that when G is commutative, then in Facts 2.1 and 2.2, the Galois group, say $\widetilde{H}$, of $K^{\sharp}(y)/K^{\sharp}$ is a unique subgroup of G, so its identity component H must indeed be the smallest algebraic subgroup of G with the required properties (see also [Bertrand 2011, Section 3.1]). Of course, $\widetilde{H}$ is automatically connected in Fact 2.2(i), where the base $K^{\sharp\,\text{alg}}$ is algebraically closed, but as just announced, our proofs in Section 3 will be based on 2.2(ii). Now, in this commutative case, the map $\sigma \mapsto \rho_\sigma$ described above depends $\mathbb{Z}$-linearly on a. So, if

$N=[\widetilde{H}:H]$ denotes the number of connected components of $\widetilde{H}$, then replacing a by Na turns the Galois group into a connected algebraic group, without modifying $K^\sharp$ nor $\text{tr.deg}(K^\sharp(y)/K^\sharp)=\text{tr.deg}(K^\sharp(Ny)/K^\sharp)$. Therefore, in the computations of Galois groups later on, we will tacitly replace y by Ny and determine the connected component H of $\widetilde{H}$. But it turns out that in our main Conjecture 2.3 and in all its cases under study here, we can then assume that y itself lies in H. Indeed, y appears only via its class modulo $G(K)$, and in particular, modulo its torsion subgroup (recall that K is algebraically closed). So, once we have proven that Ny lies in H, then a translate y' of y by an N-torsion point will lie in H. Replacing y by y' does not modify the Galois group $\widetilde{H}$ of $K^\sharp(y)$ over $K^\sharp$, so we may assume that y lies in H, in which case $\widetilde{H}$ coincides with H, and will in the end always be connected.[1]

2B. ***Galois-theoretic results.*** The question which we deal with in this paper is when and whether in Fact 2.2, it suffices to consider H defined over K and $g\in G(K)$. In fact it is not hard to see that the Galois group is defined over K, but the second point is problematic. The case where (G, s) is a ∂-module, namely G is a vector space V, and the logarithmic derivative $\partial\ell n_{G,s}(y)$ has the form $\nabla_V(y)=\partial y-By$ for some $n\times n$ matrix B over K, was considered in [Bertrand 2001], and shown to provide counterexamples, unless the ∂-module (V,∇_V) is semisimple. The rough idea is that the Galois group $\text{Gal}(K_V^\sharp/K)$ of ∇_V is then reductive, allowing an argument of Galois descent from $K_V^\sharp$ to K to construct a K-rational gauge transformation g. The argument was extended in [Bertrand 2009; 2011] to ∂-groups (G, s) attached to abelian varieties, which by Poincaré reducibility are in a sense again semisimple.

We will here focus on the *almost semiabelian* case namely certain ∂-groups attached to semiabelian varieties, which provide the main source of nonsemisimple situations. If B is a semiabelian variety over K, then $\widetilde{B}$, the universal vectorial extension of B, is a (commutative) algebraic group over K which has a *unique* algebraic ∂-group structure. Let U be any unipotent algebraic ∂-subgroup of $\widetilde{B}$. Then $\widetilde{B}/U$, which by [Bertrand and Pillay 2010, Lemma 3.4] also has a unique ∂-group structure, is what we mean by an almost semiabelian ∂-group over K. When B is an abelian variety A we call $\widetilde{A}/U$ an almost abelian algebraic ∂-group over K. If G is an almost semiabelian algebraic ∂-group over K, then because the ∂-group structure s on G is unique, the abbreviation $K_G^\sharp$ for $K_{G,s}^\sharp$ is now unambiguous. Under these conditions, we make the following conjecture.

[1]We take the opportunity of this remark to mention two errata in [Bertrand 2011]: in the proof of its Theorem 3.2, replace "of finite index" by "with quotient of finite exponent"; in the proof of Theorem 4.4, use the reduction process described above to justify that the Galois group is indeed connected.

Conjecture 2.3. *Let G be an almost semiabelian ∂-group over $K = \mathbb{C}(t)^{\mathrm{alg}}$. Let $a \in LG(K)$ and $y \in G(K^{\mathrm{diff}})$ be such that $\partial\ell n_G(y) = a$. Then* $\mathrm{tr.deg}\big(K^\sharp_G(y)/K^\sharp_G\big)$ *is the dimension of the smallest algebraic ∂-subgroup H of G defined over K such that $a \in LH + \partial\ell n_G(G(K))$, i.e., $a + \partial\ell n_G(g) \in LH(K)$ for some $g \in G(K)$; H is, equivalently, the smallest algebraic ∂-subgroup of G, defined over K, such that $y \in H + G(K) + G^\partial(K^{\mathrm{diff}})$. Moreover $H^\partial(K^{\mathrm{diff}})$ is the Galois group of $K^\sharp_G(y)$ over $K^\sharp_G$.*

The conjecture can be restated to say that there is a smallest algebraic ∂-subgroup H of (G, s) defined over K such that $a \in LH + \partial\ell n_G(G(K))$ and it coincides with the Galois group of $K^\sharp_G(y)$ over $K^\sharp_G$. In comparison with Fact 2.2(ii), notice that since K is algebraically closed, $\partial\ell n_G(G(K))$ is already a $\mathbb{Q}$-vector space, so we do not need to tensor with $\mathbb{Q}$ in the condition on a.

A corollary of Conjecture 2.3 is the following special *generic case*, where an additional assumption on nondegeneracy is made on a.

Conjecture 2.4. *Let G be an almost semiabelian ∂-group over $K = \mathbb{C}(t)^{\mathrm{alg}}$, and let $a \in LG(K)$ and $y \in G(K^{\mathrm{diff}})$ satisfy the equation $\partial\ell n_G(y) = a$. Assume that $a \notin LH + \partial\ell n_G G(K)$ for any proper algebraic ∂-subgroup H of G, defined over K (equivalently, $y \notin H + G(K) + G^\partial(K^{\mathrm{diff}})$ for any proper algebraic ∂-subgroup of G defined over K). Then* $\mathrm{tr.deg}\big(K^\sharp_G(y)/K^\sharp_G\big) = \dim(G)$.

We will prove the following results in the direction of Conjectures 2.3 and (the weaker) 2.4.

Proposition 2.5. *Conjecture 2.3 holds when G is "almost abelian".*

The truth of the weaker Conjecture 2.4 in the almost abelian case is already established in [Bertrand 2009, Section 8.1(i)]. This reference does not address Conjecture 2.3 itself, even if in this case, the ingredients for its proof are there (see also [Bertrand 2011]). So we take the liberty to give a reasonably self-contained proof of Proposition 2.5 in Section 3.

As announced above, one of the main points of the Galois-theoretic part of this paper is to try to extend Proposition 2.5 to the almost semiabelian case. Due to technical complications, which will be discussed later, we restrict our attention to the simplest possible extension of the almost abelian case, namely where the toric part of the semiabelian variety is 1-dimensional, and also we sometimes just consider the generic case. For simplicity we will state and prove our results for an almost semiabelian G of the form $\widetilde{B}$ for B semiabelian. So, the next theorem gives Conjecture 2.4 for an extension by $\mathbb{G}_m$ of the universal vectorial extension of an abelian variety.

Theorem 2.6. *Suppose that B is a semiabelian variety over $K = \mathbb{C}(t)^{\mathrm{alg}}$ with toric part of dimension ≤ 1. Let $G = \widetilde{B}$, $a \in LG(K)$ and $y \in G(K^{\mathrm{diff}})$ be a solution of*

$\partial\ell n_G(-) = a$. *Suppose that for no proper algebraic ∂-subgroup H of G defined over K is $y \in H + G(K)$. Then* $\mathrm{tr.deg}(K^{\sharp}_G(y)/K^{\sharp}_G) = \dim(G)$ *and $G^{\partial}(K^{\mathrm{diff}})$ is the differential Galois group.*

Note that the above hypothesis "$y \notin H + G(K)$ for any proper algebraic ∂-subgroup of G over K" is formally weaker than "$y \notin H + G(K) + G^{\partial}(K^{\mathrm{diff}})$ for any proper algebraic ∂-subgroup of G over K", but nevertheless suffices, as shown by the proof of Theorem 2.6 in Section 3B. More specifically, assume that $G = \widetilde{A}$ for a simple abelian variety A/K, that A is traceless (i.e., that there is no nonzero morphism from an abelian variety defined over $\mathbb{C}$ to A), that the maximal unipotent ∂-subgroup U_A of $\widetilde{A}$ vanishes, and that $a = 0 \in L\widetilde{A}(K)$. Theorem 2.6 then implies that any $y \in \widetilde{A}^{\partial}(K^{\mathrm{diff}})$ is actually defined over K, so $K^{\sharp}_{\widetilde{A}} = K$. As in [Bertrand 2009; 2011], this property of K-largeness of $\widetilde{A}$ (when $U_A = 0$) is in fact one of the main ingredients in the proof of Theorem 2.6. As explained in [Marker and Pillay 1997] it is based on the strong minimality of $\widetilde{A}^{\partial}$ (see [Hrushovski and Sokolović 1994]) in the context above. But it has recently been noted in [Benoist et al. 2014] that this K-largeness property can be seen rather more directly, using only the simplicity of A.

Our last Galois-theoretic result requires the *semiconstant* notions introduced in [Bertrand and Pillay 2010], although our notation will be a slight modification of the notation there. First, a connected algebraic group G over K is said to be constant if G is isomorphic (as an algebraic group) to an algebraic group defined over $\mathbb{C}$ (equivalently, G arises via base change from an algebraic group $G_{\mathbb{C}}$ over $\mathbb{C}$). For G an algebraic group over K, G_0 will denote the largest (connected) constant algebraic subgroup of G. We will concentrate on the case $G = \widetilde{B}$ for a semiabelian variety B over K, with $0 \to T \to B \to A \to 0$ the canonical exact sequence, where T is the maximal linear algebraic subgroup of B (which is an algebraic torus) and A is an abelian variety. So now A_0, B_0 denote the constant parts of A, B, respectively. The inverse image of A_0 in B will be called the semiconstant part of B and will now be denoted by B_{sc}. We call B semiconstant if $B = B_{\mathrm{sc}}$, which is equivalent to requiring that $A = A_0$, and moreover allows the possibility that $B = B_0$ is constant. (Of course, when B is constant, $\widetilde{B}$, which is also constant, obviously satisfies Conjecture 2.3, in view of Fact 2.1.)

Theorem 2.7. *Suppose that $K = \mathbb{C}(t)^{\mathrm{alg}}$ and that $B = B_{\mathrm{sc}}$ is a semiconstant semiabelian variety over K with toric part of dimension ≤ 1. Then Conjecture 2.3 holds for $G = \widetilde{B}$.*

2C. ***Lindemann–Weierstrass via Galois theory.*** We are now ready to describe the impact of the previous Galois-theoretic results on Ax–Lindemann problems, where $a = \partial_{LG}(\tilde{x}) \in \partial_{LG}(LG(K))$.

Firstly, from Theorem 2.6 we will deduce directly the main result of [Bertrand and Pillay 2010, Theorem 1.4], when B is semiabelian with toric part at most $\mathbb{G}_m$, but now with transcendence degree computed over $K^{\sharp}_{\widetilde{B}}$.

Corollary 2.8. *Let B be a semiabelian variety over $K = \mathbb{C}(t)^{\mathrm{alg}}$ such that the toric part of B is of dimension ≤ 1 and $B_{\mathrm{sc}} = B_0$ (i.e., the semiconstant part B_{sc} of B is constant). Let $x \in LB(K)$, and lift x to $\tilde{x} \in L\widetilde{B}(K)$. Assume that*

(∗) *for no proper algebraic subgroup H of $\widetilde{B}$ defined over K is $\tilde{x} \in LH(K) + (L\widetilde{B})^{\partial}(K)$,*

which under the current assumptions is equivalent to demanding that for no proper semiabelian subvariety H of B is $x \in LH(K) + LB_0(\mathbb{C})$. Then

(i) *any solution $\tilde{y} \in B(\mathcal{U})$ of $\partial\ell n_{\widetilde{B}}(-) = \partial_{L\widetilde{B}}(\tilde{x})$ satisfies*

$$\mathrm{tr.deg}\big(K^{\sharp}_{\widetilde{B}}(\tilde{y})/K^{\sharp}_{\widetilde{B}}\big) = \dim(\widetilde{B});$$

(ii) *in particular, $y := \exp_B(x)$ satisfies $\mathrm{tr.deg}\big(K^{\sharp}_{\widetilde{B}}(y)/K^{\sharp}_{\widetilde{B}}\big) = \dim(B)$, i.e., is a generic point over $K^{\sharp}_{\widetilde{B}}$ of B.*

See [Bertrand and Pillay 2010] for the analytic description of $\exp_B(x)$ in (ii) above. In particular $\exp_B(x)$ can be viewed as a point of $B(\mathcal{U})$. We recall briefly the argument. Consider B as the generic fiber of a family $\mathbf{B} \to S$ of complex semiabelian varieties over a complex curve S, and x as a rational section $x : S \to L\mathbf{B}$ of the corresponding family of Lie algebras. Fix a small disc U in S such that $x : U \to L\mathbf{B}$ is holomorphic, and let $\exp(x) = y : U \to \mathbf{B}$ be the holomorphic section obtained by composing with the exponential map in the fibers. So y lives in the differential field of meromorphic functions on U, which contains K, and can thus be embedded over K in the universal differentially closed field $\mathcal{U}$. So talking about $\mathrm{tr.deg}\big(K^{\sharp}_{\widetilde{B}}(y)/K^{\sharp}_{\widetilde{B}}\big)$ makes sense.

Let us comment on the methods. In [Bertrand and Pillay 2010] an essential use was made of the so-called "socle theorem" (see Section 4.1 of [Bertrand and Pillay 2010] for a discussion of this expression) in order to prove Theorem 1.4 there. As recalled in the introduction, a differential Galois-theoretic approach was also mentioned [Bertrand and Pillay 2010, Section 6], but could be worked out only when $\widetilde{B}$ is K-large. In the current paper, we dispose of this hypothesis, and obtain a stronger result, namely over $K^{\sharp}_{\widetilde{B}}$, but for the time being at the expense of restricting the toric part of B.

When $B = A$ is an abelian variety, one obtains a stronger statement than Corollary 2.8. This is Theorem 4.4 of [Bertrand 2011], which for the sake of completeness we restate, and will deduce from Proposition 2.5 in Section 4A.

Corollary 2.9. *Let A be an abelian variety over $K = \mathbb{C}(t)^{\mathrm{alg}}$. Let $x \in LA(K)$, and let B be the smallest abelian subvariety of A such that $x \in LB(K) + LA_0(\mathbb{C})$. Let*

$\tilde{x} \in L\widetilde{A}(K)$ *be a lift of* x *and let* $\tilde{y} \in \widetilde{A}(\mathcal{U})$ *be such that* $\partial\ell n_{\widetilde{A}}(\tilde{y}) = \partial_{L\widetilde{A}}(\tilde{x})$. *Then* $\widetilde{B}^{\partial}$ *is the Galois group of* $K^{\sharp}_{\widetilde{A}}(\tilde{y})$ *over* $K^{\sharp}_{\widetilde{A}}$, *so*

(i) $\mathrm{tr.deg}\big(K^{\sharp}_{\widetilde{A}}(\tilde{y})/K^{\sharp}_{\widetilde{A}}\big) = \dim(\widetilde{B}) = 2\dim(B)$, *and in particular*,

(ii) $y := \exp_A(x)$ *satisfies* $\mathrm{tr.deg}\big(K^{\sharp}_{\widetilde{A}}(y)/K^{\sharp}_{\widetilde{A}}\big) = \dim(B)$.

We now return to the semiabelian context. Corollary 2.8 is not true without the assumption that the semiconstant part of B is constant. The simplest possible counterexample is given in Section 5.3 of [Bertrand and Pillay 2010]: B is a nonconstant extension of a constant elliptic curve E_0 by $\mathbb{G}_m$, with judicious choices of x and $\tilde{x}$. Moreover $\tilde{x}$ will satisfy assumption $(*)$ in Corollary 2.8, but $\mathrm{tr.deg}(K(\tilde{y})/K) \leq 1$, which is strictly smaller than $\dim(\widetilde{B}) = 3$. We will use Theorems 2.6 and 2.7 as well as material from [Bertrand et al. 2013] to give a full account of this situation (now over $K^{\sharp}_{\widetilde{B}}$, of course), and more generally, for all semiabelian surfaces B/K, as follows:

Corollary 2.10. *Let* B *be an extension over* $K = \mathbb{C}(t)^{\mathrm{alg}}$ *of an elliptic curve* E/K *by* $\mathbb{G}_m$. *Let* $x \in LB(K)$ *satisfy*

$(*)$ *for any proper algebraic subgroup* H *of* B, $x \notin LH + LB_0(\mathbb{C})$.

Let $\tilde{x} \in L\widetilde{B}(K)$ *be a lift of* x, *let* $\bar{x}$ *be its projection to* $LE(K)$, *and let* $\tilde{y} \in \widetilde{B}(\mathcal{U})$ *be such that* $\partial\ell n_{\widetilde{B}}(\tilde{y}) = \tilde{x}$. *Then* $\mathrm{tr.deg}\big(K^{\sharp}_{\widetilde{B}}(\tilde{y})/K^{\sharp}_{\widetilde{B}}\big) = 3$, *unless* $\bar{x} \in LE_0(\mathbb{C})$, *in which case* $\mathrm{tr.deg}\big(K^{\sharp}_{\widetilde{B}}(\tilde{y})/K^{\sharp}_{\widetilde{B}}\big)$ *is precisely* 1.

Here, E_0 is the constant part of E. Notice that in view of $(*)$, E must descend to $\mathbb{C}$ and B must be nonconstant (hence not isotrivial) if x projects to $LE_0(\mathbb{C})$.

2D. ***Manin maps.*** We finally discuss the results on the Manin maps attached to abelian varieties. The expression "Manin map" covers at least two maps. The original one was introduced by Manin [1963] (see also [Coleman 1990]), and is discussed at the end of this section. Here we are mainly concerned with the model-theoretic or *differential algebraic Manin map* (see [Buium and Cassidy 1999, Section 2.5; Pillay 1997]). We identify our algebraic, differential algebraic groups with their sets of points in a universal differential field $\mathcal{U}$ (or alternatively, points in a differential closure of whatever differential field of definition we work over). So for now let K be a differential field, and A an abelian variety over K. A has a smallest Zariski-dense differential algebraic (definable in $\mathcal{U}$) subgroup $A^{\sharp}$, which can also be described as the smallest definable subgroup of A containing the torsion. The definable group $A/A^{\sharp}$ embeds definably in a commutative unipotent algebraic group (i.e., a vector group) by results of Buium, and results of Cassidy on differential algebraic vector groups yield a (noncanonical) differential algebraic isomorphism between $A/A^{\sharp}$ and $\mathbb{G}_a^n$ where $n = \dim(A)$. This differential algebraic isomorphism is defined over K, and we call it the Manin homomorphism.

There is a somewhat more intrinsic account of this isomorphism. Let $\widetilde{A}$ be the universal vectorial extension of A as discussed above, equipped with its unique algebraic ∂-group structure, and let W_A be the unipotent part of $\widetilde{A}$. We have the surjective differential algebraic homomorphism $\partial\ell n_{\widetilde{A}} : \widetilde{A} \to L\widetilde{A}$. Note that if $\tilde{y} \in \widetilde{A}$ lifts $y \in A$, then the image of $\tilde{y}$ under $\partial\ell n_{\widetilde{A}}$ modulo the subgroup $\partial\ell n_{\widetilde{A}}(W_A)$ depends only on y. This gives a surjective differential algebraic homomorphism from A to $L\widetilde{A}/\partial\ell n(W_A)$, which is defined over K, and which we call μ_A.

Remark 2.11. *Any abelian variety A/K satisfies* $\ker(\mu_A) = A^\sharp$.

Proof. Let U_A be the maximal algebraic subgroup of W_A which is a ∂-subgroup of $\widetilde{A}$. Then $\widetilde{A}/U_A$ has the structure of an algebraic ∂-group, and as explained in [Bertrand and Pillay 2010], the canonical map $\pi : \widetilde{A} \to A$ induces an isomorphism between $(\widetilde{A}/U_A)^\partial$ and $A^\sharp$. As (by functoriality) $(\widetilde{A})^\partial$ maps onto $(\widetilde{A}/U_A)^\partial$, the map $\pi : \widetilde{A} \to A$ also induces a surjective map $(\widetilde{A})^\partial \to A^\sharp$. Now, as the image of μ_A is torsion-free, $\ker(\mu_A)$ contains $A^\sharp$. On the other hand, if $y \in \ker(\mu_A)$ and $\tilde{y} \in \widetilde{A}$ lifts y, then there is $z \in W_A$ such that $\partial\ell n_{\widetilde{A}}(\tilde{y}) = \partial\ell n_{\widetilde{A}}(z)$. So $\partial\ell n_{\widetilde{A}}(\tilde{y} - z) = 0$ and $\pi(\tilde{y} - z) = y$, hence $y \in A^\sharp$. $\square$

Hence we call μ_A the (differential algebraic) Manin map. The target space embeds in an algebraic vector group and thus has the structure of a $\mathscr{C}$-vector space which is unique (any definable isomorphism between two commutative unipotent differential algebraic groups is an isomorphism of $\mathscr{C}$-vector spaces).

Now assume that $K = \mathbb{C}(t)^{\mathrm{alg}}$ and that A is an abelian variety over K with $\mathbb{C}$-trace $A_0 = 0$. Then the "model-theoretic/differential algebraic theorem of the kernel" is (see Corollary K.3 of [Bertrand and Pillay 2010]):

Fact 2.12 ($K = \mathbb{C}(t)^{\mathrm{alg}}$, A/K traceless). *The kernel* $\ker(\mu_A) \cap A(K)$ *is precisely the subgroup* $\mathrm{Tor}(A)$ *of torsion points of A.*

In Section 5 we generalize Fact 2.12 by proving:

Theorem 2.13 ($K = \mathbb{C}(t)^{\mathrm{alg}}$, A/K traceless). *Let $y_1, \dots, y_n \in A(K)$. Suppose that $a_1, \dots, a_n \in \mathbb{C}$ are not all 0, and that $a_1\mu_A(y_1) + \cdots + a_n\mu_A(y_n) = 0$ in $L\widetilde{A}(K)/\partial\ell n_{\widetilde{A}}(W_A)$. Then $y_1, \dots, y_n$ are linearly dependent over* $\mathrm{End}(A)$.

Note that on reducing to a simple abelian variety, Fact 2.12 is the special case of Theorem 2.13 when $n = 1$. Hrushovski asked whether the conclusion of the theorem can be strengthened to the linear dependence of $y_1, \dots, y_n$ over $\mathbb{Z}$. Namely, is the extension $\mu_A \otimes 1$ of μ_A to $A(K) \otimes_{\mathbb{Z}} \mathbb{C}$ injective? An example of André (see [Bertrand and Pillay 2010, p. 504; Lange and Birkenhake 1992, Chapter 9 §6]) of a traceless abelian variety A with $U_A \neq W_A$ yields a counterexample:

Proposition 2.14. *There exist*

- *a simple traceless* 4*-dimensional abelian variety A over $K = \mathbb{C}(t)^{\mathrm{alg}}$, such that* $\mathrm{End}(A)$ *is an order in a CM number field F of degree* 4 *over* $\mathbb{Q}$;
- *four points $y_1, \dots, y_4$ in $A(K)$ which are linearly dependent over* $\mathrm{End}(A)$, *but linearly independent over* $\mathbb{Z}$; *and*
- *four complex numbers $a_1, \dots, a_4$, not all zero*;

such that $a_1\mu_A(y_1) + \cdots + a_4\mu_A(y_4) = 0$.

In fact, for $i = 1, \dots, 4$, we will construct lifts $\tilde{y}_i \in \widetilde{A}(K)$ of the points y_i, and solutions $\tilde{x}_i \in L\widetilde{A}(K^{\mathrm{diff}})$ to the equations $\nabla(\tilde{x}_i) = \partial\ell n_{\widetilde{A}}\tilde{y}_i$ (where we have set $\nabla := \nabla_{L\widetilde{A}} = \partial_{L\widetilde{A}}$, with $\nabla|_{LW_A} = \partial\ell n_{\widetilde{A}}|_{W_A}$ in the identification $W_A = LW_A$), and will find a nontrivial relation

$$(\mathfrak{R}) \qquad a_1\tilde{x}_1 + \cdots + a_4\tilde{x}_4 := u \in U_A(K^{\mathrm{diff}}).$$

Since U_A is a ∇-submodule of $L\widetilde{A}$, this implies that $a_1\partial\ell n_{\widetilde{A}}\tilde{y}_1 + \cdots + a_4\partial\ell n_{\widetilde{A}}\tilde{y}_4$ lies in U_A. And since $U_A \subseteq W_A$, this in turn shows that

$$a_1\mu_A(y_1) + \cdots + a_4\mu_A(y_4) = 0 \quad \text{in } L\widetilde{A}/\partial\ell n_{\widetilde{A}}(W_A),$$

contradicting the injectivity of $\mu_A \otimes 1$.

We conclude with a remark on the more classical *differential arithmetic* Manin map $M_{K,A}$, where the stronger version *is* true. Again A is an abelian variety over $K = \mathbb{C}(t)^{\mathrm{alg}}$ with $\mathbb{C}$-trace 0. As above, we let ∇ denote $\partial_{L\widetilde{A}} : L\widetilde{A} \to L\widetilde{A}$. The map $M_{K,A}$ is then the homomorphism from $A(K)$ to $L\widetilde{A}(K)/\nabla(L\widetilde{A}(K))$, which attaches to a point $y \in A(K)$ the class $M_{K,A}(y)$ of $\partial\ell n_{\widetilde{A}}(\tilde{y})$ in $L\widetilde{A}(K)/\nabla(L\widetilde{A}(K))$, for any lift $\tilde{y}$ of y to $\widetilde{A}(K)$. This class is independent of the lift, since $\partial\ell n_{\widetilde{A}}$ and $\partial_{L\widetilde{A}}$ coincide on $W_A = LW_A$. Again $L\widetilde{A}(K)/\nabla(L\widetilde{A}(K))$ is a $\mathbb{C}$-vector space. The initial theorem of Manin (see [Coleman 1990]) says that $\ker(M_{K,A}) = \mathrm{Tor}(A) + A_0(\mathbb{C})$, so in the traceless case the kernel is precisely $\mathrm{Tor}(A)$.

Proposition 2.15 ($K = \mathbb{C}(t)^{\mathrm{alg}}$, A/K traceless). *The $\mathbb{C}$-linear extension*

$$M_{K,A} \otimes 1 : A(K) \otimes_Z \mathbb{C} \to L\widetilde{A}(K)/\nabla(L\widetilde{A}(K))$$

is injective.

3. Computation of Galois groups

Here we prove the Galois-theoretic statements Proposition 2.5 and Theorems 2.6 and 2.7 stated in Section 2B. We assume throughout that $K = \mathbb{C}(t)^{\mathrm{alg}}$.

3A. ***The abelian case.*** Let us first set up the notation. Let A be an abelian variety over K, and let A_0 be its $\mathbb{C}$-trace, which we view as a subgroup of A defined over $\mathbb{C}$. Let $\widetilde{A}$ be the universal vectorial extension of A. We have the short exact sequence $0 \to W_A \to \widetilde{A} \to A \to 0$. Let U_A denote the (unique) maximal ∂-subgroup of $\widetilde{A}$ contained in W_A. By Remarque 7.2 of [Bertrand 2009], we have:

Fact 3.1. $$\widetilde{A}^{\partial}(K^{\text{diff}}) = \widetilde{A}_0(\mathbb{C}) + \operatorname{Tor}(\widetilde{A}) + U_A^{\partial}(K^{\text{diff}}).$$

Let us briefly remark that the ingredients behind Fact 3.1 include Chai's theorem (see [Chai 1991] and Appendix K of [Bertrand and Pillay 2010]), as well as the strong minimality of $A^{\sharp}$ when A is simple and traceless from [Hrushovski and Sokolović 1994]. As already pointed out in connection with K-largeness, the reference to [Hrushovski and Sokolović 1994] can be replaced by the easier arguments from [Benoist et al. 2014]. Let $K^{\sharp}_{\widetilde{A}}$ be the (automatically differential) field generated over K by $\widetilde{A}^{\partial}(K^{\text{diff}})$, and likewise with $K^{\sharp}_{U_A}$ for $(U_A)^{\partial}(K^{\text{diff}})$. So by Fact 3.1, $K^{\sharp}_{\widetilde{A}} = K^{\sharp}_{U_A}$. Also, as recalled at the beginning of Section 8 of [Bertrand 2009], we have:

Remark 3.2. *$K^{\sharp}_{U_A}$ is a Picard–Vessiot extension of K whose Galois group (a linear algebraic group over $\mathbb{C}$) is semisimple.*

Proof of Proposition 2.5. Here, G is an *almost abelian* ∂-group over K. We first treat the case where $G = \widetilde{A}$.

Let $y \in G(K^{\text{diff}})$ be such that $a = \partial\ell n_G(y)$ lies in $LG(K)$. Note that in the setup of Conjecture 2.3, y could very well be an element of U_A, for instance when $a \in LU_A \simeq U_A$, so in a sense we are moving outside the almost abelian context. In any case, let H be a minimal ∂-subgroup of G defined over K such that $y \in H + G(K) + G^{\partial}(K^{\text{diff}})$. Since $G(K)$ contains all the torsion points, H is automatically connected. We will prove that $H^{\partial}(K^{\text{diff}})$ is the differential Galois group of $K^{\sharp}(y)$ over $K^{\sharp}$ where $K^{\sharp} = K^{\sharp}_G$. We recall from the remark after Fact 2.2 on the commutative case that we can and do assume that this Galois group is connected. Also, these statements imply that H is actually the smallest ∂-subgroup of G over K such that $y \in H + G(K) + G^{\partial}(K^{\text{diff}})$, as required.

Let H_1^{∂} be the Galois group of $K^{\sharp}(y)$ over $K^{\sharp}$ with H_1 a ∂-subgroup of G which on the face of it is defined over $K^{\sharp}$. So, H_1 is a connected ∂-subgroup of H, and we aim to show that $H = H_1$.

Claim. H_1 is defined over K as an algebraic group.

Proof. It is enough to show that H_1^{∂} is defined over K as a differential algebraic group. This is a very basic model-theoretic argument, but may be a bit surprising at the algebraic-geometric level, as $K^{\sharp}(y)$ need not be a "differential Galois extension" of K in any of the usual meanings. We use the fact that any definable (with parameters) set in the differentially closed field K^{diff} which is $\operatorname{Aut}(K^{\text{diff}}/K)$-invariant, is definable over K. This follows from model-theoretic homogeneity of K^{diff}

over K as well as elimination of imaginaries in DCF_0. Now $H_1^{\partial}(K^{\mathrm{diff}})$ is the set of $g \in G^{\partial}(K^{\mathrm{diff}})$ such that $y_1 g$ and y_1 have the same type over $K^{\sharp}$ for some (any) $y_1 \in G(K^{\mathrm{diff}})$ such that $\partial\ell n_G(y_1) = a$. As $a \in LG(K)$ and $K^{\sharp}$ is setwise invariant under $\mathrm{Aut}(K^{\mathrm{diff}}/K)$, it follows that $H_1^{\partial}(K^{\mathrm{diff}})$ is also $\mathrm{Aut}(K^{\mathrm{diff}}/K)$-invariant, and so defined over K. This proves the claim.

Note that since one of its translates by $G(K)$ lies in H, we may assume that $y \in H$, whereby $\partial\ell n_G(y) = a \in LH(K)$.

Let B be the image of H in A, and B_1 the image of H_1 in A. So $B_1 \le B$ are abelian subvarieties of A. Let V be the maximal unipotent ∂-subgroup of H, and V_1 the maximal unipotent subgroup of H_1. So $V_1 \le V$, and using the assumptions and the claim, everything is defined over K. Note also that the surjective homomorphism $H \to B$ induces an isomorphism between H/V and $\widetilde{B}/U_B$ (where as above U_B denotes the maximal unipotent ∂-subgroup of $\widetilde{B}$), and likewise for H_1/V_1 and the quotient of $\widetilde{B_1}$ by its maximal unipotent ∂-subgroup.

Case I. $B = B_1$.

Then by the previous paragraph, we have a canonical isomorphism ι (of ∂-groups) between H/H_1 and V/V_1, defined over K, so there is no harm in identifying them, although we need to remember where they came from. Let us denote V/V_1 by $\overline{V}$, a unipotent ∂-group. This isomorphism respects the logarithmic derivatives in the obvious sense. Let $\bar{y}$ denote the image of y in H/H_1. So $\partial\ell n_{H/H_1}(\bar{y}) = \bar{a}$ where $\bar{a}$ is the image of a in $L(H/H_1)(K)$. Via ι we identify $\bar{y}$ with a point in $\overline{V}(K^{\sharp})$ and $\bar{a}$ with $\partial\ell n_{\overline{V}}(\bar{y}) \in L(\overline{V})(K)$.

By Remark 3.2 we identify $\mathrm{Aut}(K^{\sharp}/K)$ with a group $J(\mathbb{C})$ where J is a semisimple algebraic group. We have a natural action of $J(\mathbb{C})$ on $\overline{V}^{\partial}(K^{\mathrm{diff}}) = \overline{V}^{\partial}(K^{\sharp})$. Now the latter is a $\mathbb{C}$-vector space, and this action can be checked to be a (rational) representation of $J(\mathbb{C})$. On the other hand, for $\sigma \in J(\mathbb{C})$, $\sigma(\bar{y})$ (which is well-defined since $\bar{y}$ is $K^{\sharp}$-rational) is also a solution of $\partial\ell n_{\overline{V}}(-) = \bar{a}$, hence $\sigma(\bar{y}) - \bar{y} \in \overline{V}^{\partial}(K^{\mathrm{diff}})$. The map taking σ to $\sigma(\bar{y}) - \bar{y}$ is then a cocycle c from $J(\mathbb{C})$ to $V^{\partial}(K^{\mathrm{diff}})$ which is a morphism of algebraic varieties. Now the corresponding $H^1(J(\mathbb{C}), \overline{V}^{\partial}(K^{\mathrm{diff}}))$ is trivial as it equals $\mathrm{Ext}_{J(\mathbb{C})}(1, \overline{V}^{\partial}(K^{\mathrm{diff}}))$, the group of isomorphism classes of extensions of the trivial representation of $J(\mathbb{C})$ by $\overline{V}^{\partial}(K^{\mathrm{diff}})$. But $J(\mathbb{C})$ is semisimple, so reductive, whereby every rational representation is completely reducible (see pp. 26 and 27 of [Mumford and Fogarty 1982], and [Bertrand 2001] for Picard–Vessiot applications, which actually cover the case when a lies in LU_A). Putting everything together, the original cocycle is trivial. Therefore there is $\bar{z} \in \overline{V}^{\partial}(K^{\sharp})$ such that $\sigma(\bar{y}) - \bar{y} = \sigma(z) - z$ for all $\sigma \in J(\mathbb{C})$. So $\sigma(\bar{y} - \bar{z}) = \bar{y} - \bar{z}$ for all σ. Hence $\bar{y} - \bar{z} \in (H/H_1)(K)$. Lift $\bar{z}$ to a point $z \in H^{\partial}(K^{\mathrm{diff}})$. So $\overline{y - z} \in \overline{V}(K)$. As K is algebraically closed, there is $d \in H(K)$ such that $y - z + d \in H_1$. This contradicts the minimal choice of H, unless $H = H_1$. So the proof is complete in Case I.

Case II. B_1 is a proper subgroup of B.

Consider the group $H_1 \cdot V$ a ∂-subgroup of H, defined over K, which also projects onto B_1. It is now easy to extend $H_1 \cdot V$ to a ∂-subgroup H_2 of H over K such that H/H_2 is canonically isomorphic to $\overline{B_2}$, where B_2 is a simple abelian variety, and $\overline{B_2}$ denotes the quotient of $\widetilde{B_2}$ by its maximal unipotent subgroup. Now let $\bar{y}$ denote $y/H_2 \in H/H_2$. Hence $\partial\ell n_{\overline{B_2}}(\bar{y}) = \bar{a} \in L(\overline{B_2})(K)$. As $H_1 \subseteq H_2$, $\bar{y} \in \overline{B_2}(K^\sharp)$. Now we have two cases. If B_2 descends to $\mathbb{C}$, then $\bar{y}$ generates a strongly normal extension of K whose Galois group is a connected algebraic subgroup of $B_2(\mathbb{C})$. As this Galois group will be a homomorphic image of the linear (in fact semisimple) complex algebraic group $\mathrm{Aut}(K^\sharp/K)$, we have a contradiction unless $\bar{y}$ is K-rational. On the other hand, if B_2 does not descend to $\mathbb{C}$, then by Fact 2.2(ii) $\bar{y}$ generates over K a (generalized) differential Galois extension of K with Galois group contained in $\overline{B_2}^\partial(K^{\mathrm{diff}})$, which again will be a homomorphic image of a complex semisimple linear algebraic group (cf. [Bertrand 2009, 8.2(i)]). We get a contradiction by various possible means (for example as in Remarque 8.2 of [Bertrand 2009]) unless $\bar{y}$ is K-rational. So either way we are forced into $\bar{y} \in (H/H_2)(K)$. But then, as K is algebraically closed, $y - d \in H_2$ for some $d \in H(K)$, again a contradiction. So Case II is impossible. This concludes the proof of Proposition 2.5 when $G = \widetilde{A}$.

Finally, consider a general almost abelian ∂-group G, given as a quotient of $\widetilde{A}$ by a unipotent ∂-subgroup $U \subset U_A$ defined over K. Taking the quotient by $U^\partial(K^{\mathrm{diff}})$ of the decomposition of $\widetilde{A}^\partial(K^{\mathrm{diff}})$ given by Fact 3.1, we obtain a similar decomposition for $G^\partial(K^{\mathrm{diff}})$. Therefore $K^\sharp_G = K((U_A/U)^\partial)$ is also a Picard–Vessiot extension of K, and we deduce from Remark 3.2 that its Galois group is again semisimple. The various cases of the previous proof therefore also apply to the quotient $G = \widetilde{A}/U$, and Proposition 2.5 holds for any almost abelian ∂-group. □

3B. ***The semiabelian case.*** We now aim towards proofs of Theorems 2.6 and 2.7. Here, $G = \widetilde{B}$ for B a semiabelian variety over K, equipped with its unique algebraic ∂-group structure.

We have:

- $0 \to T \to B \to A \to 0$, where T is an algebraic torus and A an abelian variety, all over K,
- $G = \widetilde{B} = B \times_A \widetilde{A}$, where $\widetilde{A}$ is the universal vectorial extension of A, and
- $0 \to T \to G \to \widetilde{A} \to 0$.

We use the same notation for A as at the beginning of this section, namely

$$0 \longrightarrow W_A \longrightarrow \widetilde{A} \longrightarrow A \longrightarrow 0.$$

We denote by A_0 the $\mathbb{C}$-trace of A (so up to isogeny we can write A as a product $A_0 \times A_1$, all defined over K, where A_1 has $\mathbb{C}$-trace 0), and by U_A the maximal

∂-subgroup of $\widetilde{A}$ contained in W_A. So U_A is a unipotent subgroup of G, though not necessarily one of its ∂-subgroups. Finally, we have the exact sequence

$$0 \longrightarrow T^{\partial} \longrightarrow G^{\partial} \xrightarrow{\pi} \widetilde{A}^{\partial} \longrightarrow 0.$$

Note that $T^{\partial} = T(\mathbb{C})$. Let $K_G^{\sharp}$ be the (differential) field generated over K by $G^{\partial}(K^{\text{diff}})$. We have already noted above that $K_{\widetilde{A}}^{\sharp}$ equals $K_{U_A}^{\sharp}$. So $K_{U_A}^{\sharp} < K_G^{\sharp}$, and we deduce from the last exact sequence above the following:

Remark 3.3. *$G^{\partial}(K^{\text{diff}})$ is the union of the $\pi^{-1}(b)$ for $b \in \widetilde{A}^{\partial}$, each $\pi^{-1}(b)$ being a coset of $T(\mathbb{C})$ defined over $K_{U_A}^{\sharp}$. Hence $K_G^{\sharp}$ is (generated by) a union of Picard–Vessiot extensions over $K_{U_A}^{\sharp}$, each with Galois group contained in $T(\mathbb{C})$.*

Proof of Theorem 2.6. Bearing in mind Proposition 2.5 we may assume that $T = \mathbb{G}_m$. We have $a \in LG(K)$ and $y \in G(K^{\text{diff}})$ such that $\partial \ell n_G(y) = a$ and $y \notin H + G(K)$ for any proper ∂-subgroup H of G. The latter is a little weaker than the condition that $a \notin LH(K) + \partial \ell n_G(G(K))$ for any proper H, but (thanks to Fact 3.1) will suffice for the special case we are dealing with.

Fix a solution y of $\partial \ell n_G(-) = a$ in $G(K^{\text{diff}})$ and let $H^{\partial}(K^{\text{diff}})$ be the differential Galois group of $K_G^{\sharp}(y)$ over $K_G^{\sharp}$. As said after Fact 2.2, there is no harm in assuming that H is connected. So H is a connected ∂-subgroup of G, defined over $K_G^{\sharp}$.

As in the proof of the claim in the proof of Proposition 2.5, we have:

Claim 1. H (equivalently H^{∂}) is defined over K.

We assume for a contradiction that $H \neq G$.

Case I. H maps onto a proper (∂-)subgroup of $\widetilde{A}$.

This is similar to Case II in the proof of Proposition 2.5 above. Some additional complications come from the structure of $K_G^{\sharp}$. We will make use of Remark 3.3 all the time.

As $\widetilde{A}$ is an essential extension of A by W_A, it follows that we can find a connected ∂-subgroup H_1 of G containing H and defined over K such that the surjection $G \to \widetilde{A}$ induces an isomorphism between G/H_1 and $\overline{A_2}$, where A_2 is a simple abelian subvariety of A (over K of course) and $\overline{A_2}$ is the quotient of $\widetilde{A_2}$ by its maximal unipotent ∂-subgroup. Let η and α be such that the quotient map taking G to $\overline{A_2}$ takes y to η and also induces a surjection $LG \to L(\overline{A_2})$ which takes a to α.

As $\eta = y/H_1$ and $H \subseteq H_1$, we see that η is fixed by $\operatorname{Aut}\big(K_G^{\sharp}(y)/K_G^{\sharp}\big)$, establishing the following:

Claim 2. We have $\eta \in \overline{A_2}(K_G^{\sharp})$.

On the other hand, η is a solution of the logarithmic differential equation $\partial \ell n_{\overline{A_2}}(-) = \alpha$ over K. By K-largeness of $\overline{A_2}$, we have $K_{\overline{A_2}}^{\sharp} = K$, hence $K(\eta)$ is a differential Galois extension of K whose Galois group is either trivial (in which case $\eta \in \overline{A_2}(K)$), or equal to $\overline{A_2}^{\partial}(K^{\text{diff}})$, in view of the strong minimality of $\overline{A_2}^{\partial}$.

Claim 3. We have $\eta \in \overline{A_2}(K)$.

Proof. Suppose not. We first claim that η is independent from $K^{\sharp}_{U_A}$ over K (in the sense of differential fields). Indeed, the Galois theory would otherwise give us some proper definable subgroup in the product of $\overline{A_2}^{\partial}(K^{\mathrm{diff}})$ by the Galois group of $K^{\sharp}_{U_A}$ over K (or equivalently, these two groups would share a nontrivial definable quotient). As the latter is a complex semisimple algebraic group (Remark 3.2), we get a contradiction. Alternatively, we could proceed as in Remarque 8.2 of [Bertrand 2009].

So the Galois group of $K^{\sharp}_{U_A}(\eta)$ over $K^{\sharp}_{U_A}$ is $\overline{A_2}^{\partial}(K^{\mathrm{diff}})$. As there are no nontrivial definable subgroups of $\overline{A_2}(K^{\mathrm{diff}}) \times \mathbb{G}_m(\mathbb{C})^n$, we see that η is independent of $K^{\sharp}_G$ over $K^{\sharp}_{U_A}$, contradicting Claim 2.

By Claim 3, the coset of y modulo H_1 is defined over K (differential algebraically), so as in the proof of Fact 2.1, as K is algebraically closed there is $y_1 \in G(K)$ in the same coset of H_1 as y. So $y \in H_1 + G(K)$, contradicting the assumptions. Thus Case I is complete.

Case II. H projects onto $\widetilde{A}$.

Our assumption that H is a proper subgroup of G and that the toric part is $\mathbb{G}_m$ implies that (up to isogeny) G splits as $T \times H = T \times \widetilde{A}$. This case is essentially dealt with in [Bertrand 2009], but nevertheless we continue with the proof. We identify G/H with T. So $y/H = d \in T$ and the image a_0 of a under the projection $G \to T$ is in $LT(K)$. As $H^{\partial}(K^{\mathrm{diff}})$ is the Galois group of $K^{\sharp}_G(y)$ over $K^{\sharp}_G$, we see that $y \in T(K^{\sharp}_G)$. Now $K(d)$ is a Picard–Vessiot extension of K with Galois group a subgroup of $\mathbb{G}_m(\mathbb{C})$. Moreover, since G splits as $T \times \widetilde{A}$, we have $G^{\partial} = T^{\partial} \times \widetilde{A}^{\partial}$. Hence by Fact 3.1, $K^{\sharp}_G = K^{\sharp}_{\widetilde{A}}$, and by Remark 3.2, it is a Picard–Vessiot extension of K whose Galois group is a semisimple algebraic group in the constants. We deduce from the Galois theory that d is independent from $K^{\sharp}_G$ over K, and hence $d \in T(K)$. So the coset of y modulo H has a representative $y_1 \in G(K)$ and $y \in H + G(K)$, contradicting our assumption. This concludes Case II and the proof of Theorem 2.6. □

Proof of Theorem 2.7. $G = \widetilde{B}$ for $B = B_{\mathrm{sc}}$ a semiconstant semiabelian variety over K and we may assume it has toric part $\mathbb{G}_m$. So although the toric part is still $\mathbb{G}_m$, both the hypothesis and conclusion of Theorem 2.7 are stronger than in Theorem 2.6.

We have $0 \to \mathbb{G}_m \to B \to A$ where $A = A_0$ is over $\mathbb{C}$. Hence $\widetilde{A}$ is also over $\mathbb{C}$ and we have $0 \to \mathbb{G}_m \to \widetilde{B} \to \widetilde{A} \to 0$, and $G = \widetilde{B}$. As $\widetilde{A}^{\partial} = \widetilde{A}(\mathbb{C}) \subseteq \widetilde{A}(K)$, we see:

Fact 3.4. *$G^{\partial}(K^{\mathrm{diff}})$ is a union of cosets of $\mathbb{G}_m(\mathbb{C})$, each defined over K.*

We are given a logarithmic differential equation $\partial \ell n_G(-) = a \in LG(K)$ and solution $y \in G(K^{\mathrm{diff}})$. We let H be a minimal connected ∂-subgroup of G, defined over K, such that $a \in LH + \partial\ell n_G(G(K))$, or equivalently, $y \in H + G(K) + G^\partial(K^{\mathrm{diff}})$. We want to prove that $H^\partial(K^{\mathrm{diff}})$ is the Galois group of $K_G^\sharp(y)$ over $K_G^\sharp$.

By Theorem 2.6, we may assume that $H \neq G$. Note that after translating y by an element of $G(K)$ plus an element of $G^\partial(K^{\mathrm{diff}})$, we can assume that $y \in H$. If H is trivial then everything is clear.

We go through the cases.

Case I. $H = \mathbb{G}_m$.

Then by Fact 2.1, $K(y)$ is a Picard–Vessiot extension of K, with Galois group $\mathbb{G}_m(\mathbb{C})$, and all that remains to be proved is that y is algebraically independent from $K^\sharp$ over K. Let $z_1, \dots, z_n \in G^\partial(K^{\mathrm{diff}})$, and we want to show that y is independent from $z_1, \dots, z_n$ over K (in the sense of DCF_0). By Fact 3.4, $K(z_1, \dots, z_n)$ is a Picard–Vessiot extension of K and we can assume the Galois group is $\mathbb{G}_m^n(\mathbb{C})$. Suppose towards a contradiction that $\mathrm{tr.deg}(K(y, z_1, \dots, z_n)/K) < n+1$, and so must equal n. Hence the differential Galois group of $K(y, z_1, \dots, z_n)/K$ is of the form $L(\mathbb{C})$ where L is the algebraic subgroup of $\mathbb{G}_m^{n+1}$ defined by $x^k x_1^{k_1} \cdots x_n^{k_n} = 1$ for k, k_i integers such that $k \neq 0$ and not all $k_i = 0$. It easily follows that in additive notation, $ky + k_1 z_1 + \cdots + k_n z_n \in G(K)$. So ky is of the form $z + g$ for $z \in G^\partial(K^{\mathrm{diff}})$ and $g \in G(K)$. Let $z' \in G^\partial$ and $g' \in G(K)$ be such that $kz' = z$ and $kg' = g$. Then $k(y - (z' + g')) = 0$, so $y - (z' + g)$ is a torsion point of G and hence also in G^∂. We conclude that $y \in G^\partial(K^{\mathrm{diff}}) + G(K)$, contradicting our assumptions on y. This concludes the proof in Case I.

Case II. H projects onto $\widetilde{A}$.

So our assumption that $G \neq H$ implies that up to isogeny G is $T \times \widetilde{A}$, and so defined over $\mathbb{C}$. Now everything follows from Fact 2.1.

Case III. Otherwise.

This is more or less a combination of the previous cases. To begin, suppose H is disjoint from T (up to a finite set). So $H \leq \widetilde{A}$ is a constant group, and by Fact 2.1, $H^\partial(K^{\mathrm{diff}}) = H(\mathbb{C})$ is the Galois group of $K(y)$ over K. By Fact 3.4 the Galois theory tells us that y is independent from $K_G^\sharp$ over K, so $H(\mathbb{C})$ is the Galois group of $K^\sharp(y)$ over $K^\sharp$ as required.

So we may assume that $T \leq H$. Let $H_1 \leq H$ be the differential Galois group of $K_G^\sharp(y)$ over $K_G^\sharp$, and we suppose for a contradiction that $H_1 \neq H$. As in the proof of Proposition 2.5, H_1 is defined over K. By the remark after Fact 2.2, we can assume that H_1 is connected.

Case III(a). H_1 is a complement of T in H (in the usual sense that $H_1 \times T \to H$ is an isogeny).

So $y/H_1 \in T(K_G^\sharp)$. Let $y_1 = y/H_1$. If $y_1 \notin T(K)$, then $K(y_1)$ is a Picard–Vessiot extension of K with Galois group $\mathbb{G}_m(\mathbb{C})$. The proof in Case I above shows that

$y_1 \in G^\partial(K^{\mathrm{diff}}) + G(K)$, whereby $y \in H_1 + G^\partial(K^{\mathrm{diff}}) + G(K)$, contradicting the minimality assumptions on H.

Case III(b). $H_1 + T$ is a proper subgroup of H.

Note that since we are assuming $H_1 \neq H$, the negation of Case III(a) forces Case III(b) to hold. Let $H_2 = H_1 + T$, so H/H_2 is a constant group, say H_3, which is a vectorial extension of an abelian variety. Then $y_2 = y/H_2 \in H_3(K^\sharp_G)$, and $K(y_2)$ is a Picard–Vessiot extension of K with Galois group a subgroup of $H_3(\mathbb{C})$. Fact 3.4 and the Galois theory imply that $y_2 \in H_3(K)$. Hence $y \in H_2 + G(K)$, contradicting the minimality of H again.

This completes the proof of Theorem 2.7. □

3C. ***Discussion on nongeneric cases.*** We complete this section with a discussion of some complications arising when one would like to drop either the genericity assumption in Theorem 2.6, or the restriction on the toric part in both Theorems 2.6 and 2.7.

Let us first give an example which will have to be considered if we drop the genericity assumption in Theorem 2.6, and give some positive information as well as identify some technical complications. Let A be a simple abelian variety over K which has $\mathbb{C}$-trace 0 and such that $U_A \neq 0$. (Note that such an example appears below in Section 5B connected with Manin map issues.) Let B be a nonsplit extension of A by $\mathbb{G}_m$, and let $G = \widetilde{B}$. We have $\pi : G \to \widetilde{A}$ with kernel $\mathbb{G}_m$, and let H be $\pi^{-1}(U_A)$, a ∂-subgroup of G. Let $a \in LH(K)$ and $y \in H(K^{\mathrm{diff}})$ with $d\ell n_H(y) = a$. We have to compute $\mathrm{tr.deg}\big(K^\sharp_G(y)/K^\sharp_G\big)$. Conjecture 2.3 predicts that it is the dimension of the smallest algebraic ∂-subgroup H_1 of H such that $y \in H_1 + G(K) + G^\partial(K^{\mathrm{diff}})$.

Lemma 3.5. *With the above notation, suppose $y \notin H_1 + G(K) + G^\partial(K^{\mathrm{diff}})$ for any proper algebraic ∂-subgroup H_1 of H over K. Then* $\mathrm{tr.deg}\big(K^\sharp_G(y)/K^\sharp_G\big) = \dim(H)$ (*and H is the Galois group*).

Proof. Let z and α be the images of y and a, respectively, under the maps $H \to U_A$ and $LH \to L(U_A) = U_A$ induced by $\pi : G \to \widetilde{A}$. So $\partial\ell n_{\widetilde{A}}(z) = \alpha$ with $\alpha \in L\widetilde{A}(K)$.

Claim. We have $z \notin U + \widetilde{A}(K) + \widetilde{A}^\partial(K^{\mathrm{diff}})$ for any proper algebraic ∂-subgroup U of U_A over K.

Proof of claim. Suppose otherwise. Then lifting suitable $z_2 \in \widetilde{A}(K)$ and $z_3 \in \widetilde{A}(K^{\mathrm{diff}})$ to $y_2 \in G(K)$ and $y_3 \in G^\partial(K^{\mathrm{diff}})$, respectively, we see that $y - (y_2 + y_3) \in \pi^{-1}(U)$, a proper algebraic ∂-subgroup of H, a contradiction.

As in Case I in the proof of Proposition 2.5 above, we may now conclude that $\mathrm{tr.deg}\big(K^\sharp_{\widetilde{A}}(z)/K^\sharp_{\widetilde{A}}\big) = \dim(U_A)$, and U_A is the Galois group. Now $K^\sharp_G$ is a union of Picard–Vessiot extensions of $K^\sharp_{\widetilde{A}} = K^\sharp_{U_A}$, each with Galois group $\mathbb{G}_m$ (by

Remark 3.3), so the Galois theory tells us that z is independent from $K_G^\sharp$ over $K_{\widetilde{A}}^\sharp$. Hence the differential Galois group of $K_G^\sharp(z)$ over $K_G^\sharp$ is U_A^∂. But then the Galois group of $K_G^\sharp(y)$ over $K_G^\sharp$ will be the group of ∂-points of a ∂-subgroup of H which projects onto U_A. The only possibility is H itself, because otherwise H splits as $\mathbb{G}_m \times U_A$ as a ∂-group, which contradicts (v) of Section 2 of [Bertrand 2009]. This completes the proof. □

Essentially the same argument applies if we replace H by the preimage under π of some nontrivial ∂-subgroup of U_A. So this shows that the scenario described right before Lemma 3.5 reduces to the case where $a \in LT$ where T is the toric part $\mathbb{G}_m$ (of both G and H), and we may assume $y \in T(K^{\mathrm{diff}})$. We would like to show (in analogy with Lemma 3.5) that if $y \notin G(K)+G^\partial(K^{\mathrm{diff}})$ then $\mathrm{tr.deg}\big(K_G^\sharp(y)/K_G^\sharp\big)=1$. Of course already $K(y)$ is a Picard–Vessiot extension of K with Galois group $T(\mathbb{C})$, and we have to prove that y is independent from $K_G^\sharp$ over K. One deduces from the Galois theory that y is independent from $K_{U_A}^\sharp$ over K. It remains to show that for any $z_1, \ldots, z_n \in G^\partial(K^{\mathrm{diff}})$, y is independent from $z_1, \ldots, z_n$ over $K_{U_A}^\sharp$. If not, the discussion in Case I of the proof of Theorem 2.7 gives that $y = z + g$ for some $z \in G^\partial(K^{\mathrm{diff}})$ and $g \in G(K_{U_A}^\sharp)$, but an additional argument seems necessary to yield a contradiction.

Similar and other issues arise when we want to drop the restriction on the toric part. For example in Case II in the proof of Theorem 2.6, we can no longer deduce the splitting of G as $T \times \widetilde{A}$. And in the proof of Theorem 2.7, both the analogues of Case I ($H = T$) and Case II (H projects on to $\widetilde{A}$) present technical difficulties.

4. Lindemann–Weierstrass

We here prove Corollaries 2.8, 2.9, and 2.10.

4A. *General results.*

Proof of Corollary 2.8. We first prove (i). Write G for $\widetilde{B}$. Let $\tilde{x} \in LG(K)$ be a lift of x and $\tilde{y} \in G(\mathcal{U})$ a solution of $\partial\ell n_G(-) = \tilde{x}$. We refer to Section 1.2 and Lemma 4.2 of [Bertrand and Pillay 2010] for a discussion of the equivalence of the hypotheses

$$x \notin LH(K) + LB_0(\mathbb{C}) \text{ for any proper semiabelian subvariety } H \text{ of } B,$$

and

$$(*)\quad \tilde{x} \notin LH(K) + (LG)^\partial(K) \text{ for any proper algebraic subgroup } H \text{ of } G \text{ over } K.$$

Let $a = \partial_{LG}(\tilde{x})$. So $\tilde{y}$ is a solution of the logarithmic differential equation (over K) $\partial\ell n_G(-) = a$. We want to show that $\mathrm{tr.deg}\big(K_G^\sharp(\tilde{y})/K_G^\sharp\big) = \dim(G)$. If not, we may assume that $\tilde{y} \in G(K^{\mathrm{diff}})$, and so by Theorem 2.6, $\tilde{y} \in H + G(K)$ for some proper connected algebraic ∂-subgroup H of G defined over K. Extend H to a maximal proper connected ∂-subgroup H_1 of G defined over K. Then G/H_1 is either

(a) $\mathbb{G}_m$, or

(b) a simple abelian variety A_0 over $\mathbb{C}$, or

(c) the quotient of $\widetilde{A}_1$ by a maximal unipotent ∂-subgroup, where A_1 is a simple abelian variety over K with $\mathbb{C}$-trace 0.

Let x', y' be the images of $\tilde{x}$, $\tilde{y}$ under the map $G \to G/H_1$ and induced map $LG \to L(G/H_1)$. So both x' and y' are K-rational. Moreover the hypothesis (∗) is preserved in G/H_1 (by our assumptions on G and Lemma 4.2(ii) of [Bertrand and Pillay 2010]). As $\partial\ell n_{G/H_1}(y') = \partial_{L(G/H_1)}(x')$, we have a contradiction in each of the cases (a), (b), and (c) listed above, by virtue of the truth of Ax–Lindemann in the constant case, as well as Manin–Chai (Proposition 4.4 in [Bertrand and Pillay 2010]).

(ii) Immediate as in [Bertrand and Pillay 2010]: choosing $\tilde{y} = \exp_G(\tilde{x})$, then $\exp_B(y)$ is the projection of $\tilde{y}$ on B. □

Proof of Corollary 2.9. This is like the proof of Corollary 2.8. So $x \in LA(K)$. Let $\tilde{x} \in L\widetilde{A}(K)$ lift x and let $\tilde{y} \in \widetilde{A}(K^{\text{diff}})$ be such that $\partial\ell n_{\widetilde{A}}(\tilde{y}) = \partial_{L\widetilde{A}}(\tilde{x}) = a$, say. Let B be a minimal abelian subvariety of A such that $x \in LB(K) + LA_0(\mathbb{C})$, and we want to prove that $\text{tr.deg}\big(K^\sharp_{\widetilde{A}}(\tilde{y})/K^\sharp_{\widetilde{A}}\big) = \dim(\widetilde{B})$.

Claim. We may assume that $x \in LB(K)$, $\tilde{x} \in L\widetilde{B}(K)$, and $\tilde{y} \in \widetilde{B}(K^{\text{diff}})$.

Proof of claim. Let $x = x_1 + c$ for $x_1 \in LB$ and $c \in LA_0(\mathbb{C})$. Let $\tilde{x}_1 \in L\widetilde{B}(K)$ be a lift of x_1 and $\tilde{c} \in L\widetilde{A}_0(\mathbb{C})$ be a lift of c. Finally let $\tilde{y}_1 \in \widetilde{B}(K^{\text{diff}})$ be such that $\partial\ell n_{\widetilde{A}}(\tilde{y}_1) = \partial_{L\widetilde{A}}(\tilde{x}_1) = a_1$, say. As $\tilde{x}_1 + \tilde{c}$ projects onto x, it differs from $\tilde{x}$ by an element $z \in LW(K)$. Now $\partial_{L\widetilde{A}}(z) = \partial\ell n_{\widetilde{A}}(z)$. So

$$a = \partial_{L\widetilde{A}}(\tilde{x}) = \partial_{L\widetilde{A}}(\tilde{x}_1 + \tilde{c} + z) = \partial_{L\widetilde{A}}(\tilde{x}_1) + \partial\ell n_{\widetilde{A}}(z) = a_1 + \partial\ell n_{\widetilde{A}}(z).$$

Hence $\partial\ell n(\tilde{y}_1 + z) = a$, and so $\tilde{y}_1 + z$ differs from $\tilde{y}$ by an element of $\widetilde{A}^\partial$. Hence $\text{tr.deg}\big(K^\sharp_{\widetilde{A}}(\tilde{y}_1)/K^\sharp_{\widetilde{A}}\big) = \text{tr.deg}\big(K^\sharp_{\widetilde{A}}(\tilde{y}_1)/K^\sharp_{\widetilde{A}}\big)$. Moreover the same hypothesis remains true of x_1 (namely B is minimal such that $x_1 \in LB + LA_0(\mathbb{C})$). So we can replace $x, \tilde{x}, \tilde{y}$ by $x_1, \tilde{x}_1, \tilde{y}_1$.

As recalled in the proof of Corollary 2.8 (see Corollary H.5 of [Bertrand and Pillay 2010]), the condition that $x \notin B_1(K) + LA_0(\mathbb{C})$ for any proper abelian subvariety B_1 of B is equivalent to

(∗) $\tilde{x} \notin LH(K) + (L\widetilde{A})^\partial(K)$ for any proper algebraic subgroup H of $\widetilde{B}$ defined over K.

Now we can use the Galois-theoretic result Proposition 2.5, namely the truth of Conjecture 2.3 for $\widetilde{A}$, as above. That is, if to obtain a contradiction we suppose $\text{tr.deg}\big(K^\sharp_{\widetilde{A}}(\tilde{y})/K^\sharp_{\widetilde{A}}\big) < \dim(\widetilde{B})$, then $\tilde{y} \in H + \widetilde{A}(K) + (\widetilde{A})^\partial(K^{\text{diff}})$ for some proper connected algebraic ∂-subgroup of $\widetilde{B}$, defined over K, and moreover H^∂ is the differential Galois group of $K^\sharp_{\widetilde{A}}(\tilde{y})/K^\sharp_{\widetilde{A}}$. As at the end of the proof of Corollary 2.8 above, we get a contradiction by choosing H_1 to be a maximal proper connected algebraic ∂-subgroup of $\widetilde{A}$ containing H and defined over K. This concludes the proof of Corollary 2.9. □

4B. ***Semiabelian surfaces.*** We first recall the counterexample from Section 5.3 of [Bertrand and Pillay 2010]. This example shows that in Corollary 2.8, we cannot drop the assumption that the semiconstant part is constant. We go through it again briefly. Let B over K be a nonconstant extension of a constant elliptic curve $E = E_0$ by $\mathbb{G}_m$, and let $G = \widetilde{B}$. Let $\tilde{x} \in LG(K)$ map onto a point $\check{x}$ in $L\widetilde{E}(\mathbb{C})$ which itself maps onto a nonzero point $\bar{x}$ of $LE(\mathbb{C})$. As pointed out in [Bertrand and Pillay 2010], we have $(LG)^\partial(K) = (L\mathbb{G}_m)(\mathbb{C})$, whereby $\tilde{x}$ satisfies the hypothesis $(*)$ from Corollary 2.8: $\tilde{x} \notin LH(K) + (LG)^\partial(K)$ for any proper algebraic subgroup H of G. Let $a = \partial_{LG}(\tilde{x}) \in LG(K)$, and $\tilde{y} \in G(K^{\mathrm{diff}})$ such that $\partial\ell n_G(\tilde{y}) = a$. Then as the image of a in $L\widetilde{E}$ is 0, $\tilde{y}$ projects onto a point of $\tilde{E}(\mathbb{C})$, and hence $\tilde{y}$ is in a coset of $\mathbb{G}_m$ defined over K, whereby $\mathrm{tr.deg}(K(\tilde{y})/K) \le 1$, so a fortiori the same is true with $K^\sharp_G$ in place of K. A consequence of Corollary 2.10, in fact the main part of its proof, is that with the above choice of $\tilde{x}$, we have $\mathrm{tr.deg}\big(K^\sharp_G(\tilde{y})/K^\sharp_G\big) = 1$ (as announced in [Bertrand et al. 2013, Footnote 5]).

Proof of Corollary 2.10. Let us fix notation: B is a semiabelian variety over K with toric part $\mathbb{G}_m$ and abelian quotient a not necessarily constant elliptic curve E/K, with constant part E_0; G denotes the universal vectorial extension $\widetilde{B}$ of B and $\widetilde{E}$ the universal vectorial extension of E. For $x \in LB(K)$, $\tilde{x}$ denotes a lift of x to a point of $LG(K)$, $\bar{x}$ denotes the projection of x to $LE(K)$, and $\check{x}$ denotes the projection of $\tilde{x}$ to $L\widetilde{E}(K)$.

Recall the hypothesis $(*)$ in Corollary 2.10: $x \notin LH + LB_0(\mathbb{C})$ for any proper algebraic subgroup H of B. As pointed out after the statement of Corollary 2.10, under this hypothesis, the condition $\bar{x} \in LE_0(\mathbb{C})$ can occur only if B is semiconstant and not constant. Indeed, if B were not semiconstant then $E_0 = 0$, so $x \in L\mathbb{G}_m$, contradicting the hypothesis on x. And if B were constant then $B = B_0$ and $\bar{x}$ would have a lift in $LB_0(\mathbb{C})$, whereby $x \in L\mathbb{G}_m + LB_0(\mathbb{C})$, contradicting the hypothesis.

Now if the semiconstant part of B is constant, then we can simply quote Corollary 2.8, bearing in mind the paragraph above which rules out the possibility that $\bar{x} \in LE_0(\mathbb{C})$. So we will assume that $B_{\mathrm{sc}} \ne B_0$, namely $E = E_0$ and $B_0 = \mathbb{G}_m$.

Case I. We have $\bar{x} \in LE(\mathbb{C})$ $(= LE_0(\mathbb{C})$ as $E = E_0)$.

This is where the bulk of the work goes. We first check that we are essentially in the situation of the "counterexample" mentioned above. The argument is a bit like in the proof of the claim in Corollary 2.9. Note that $\bar{x} \ne 0$ by hypothesis $(*)$. Let $\check{x}'$ be a lift of $\bar{x}$ to a point in $L\widetilde{E}(\mathbb{C})$ (noting that $\widetilde{E}$ is also over $\mathbb{C}$). Then $\check{x}' = \check{x} - \beta$ for some

$\beta \in L\mathbb{G}_a(K)$. Let $\tilde{x}' = \tilde{x} - \beta$. Let $a' = \partial_{LG}(\tilde{x}')$. Then (as $\partial_{LG}(\beta) = \partial\ell n_G(\beta)$, under the usual identifications) $a' = a + \partial\ell n_G(\beta)$, and if $\tilde{y}' \in G$ is such that $\partial\ell n_G(\tilde{y}') = a'$ then $\partial\ell n_G(\tilde{y}' - \beta) = a$. As $\beta \in G(K)$, $\mathrm{tr.deg}\big(K_G^\sharp(\tilde{y}')/K_G^\sharp\big) = \mathrm{tr.deg}\big(K_G^\sharp(\tilde{y})/K_G^\sharp\big)$.

The end result is that we can assume that $\tilde{x} \in LG(K)$ maps onto $\check{x}' \in L\widetilde{E}(\mathbb{C})$ which in turn maps on to our nonzero $\bar{x} \in LE(\mathbb{C})$, precisely the situation in the example above from Section 5.1 of [Bertrand and Pillay 2010]. So to deal with Case I, we need to prove:

Claim 1. We have $\mathrm{tr.deg}\big(K_G^\sharp(\tilde{y})/K_G^\sharp\big) = 1$.

Proof of Claim 1. Remember that a denotes $\partial_{LG}(\tilde{x})$. Now by Theorem 2.7, it suffices to prove that $a \notin \partial\ell n_G(G(K))$.

We assume for a contradiction that there is $\tilde{s} \in G(K)$ such that

$$a = \partial_{LG}(\tilde{x}) = \partial\ell n_G(\tilde{s}). \tag{†}$$

This is the semiabelian analogue of a Manin kernel statement, which can probably be studied directly, but we will deduce the contradiction from [Bertrand et al. 2013]. Let $\tilde{x}_1 = \log_G(\tilde{s})$ be a solution given by complex analysis to the linear inhomogeneous equation $\partial_{LG}(-) = \partial\ell n_G(\tilde{s})$. Namely, with notations as in the appendix to [Bertrand and Pillay 2010] (generalizing those given after Corollary 2.8 above), a local analytic section of $L\mathbf{G}^{\mathrm{an}}/S^{\mathrm{an}}$ such that $\exp_{\mathbf{G}}(\tilde{x}_1) = \tilde{s}$. Let $\xi \in (LG)^\partial$ be $\tilde{x} - \tilde{x}_1$. Then ξ lives in a differential field (of meromorphic functions on some disc in S) which extends K and has the same constants as K, namely $\mathbb{C}$. As ξ is the solution of a linear homogeneous differential equation over K, it follows that ξ lives in $(LG)^\partial(K^{\mathrm{diff}})$. Hence, as $\tilde{x} \in LG(K)$, this implies that $\tilde{x}_1 \in LG(K_{LG}^\sharp)$ where $K_{LG}^\sharp$ is the differential field generated over K by $(LG)^\partial(K^{\mathrm{diff}})$.

Now from Section 5.1 of [Bertrand et al. 2013], $K_{LG}^\sharp$ coincides with the "field of periods" F_q attached to the point $q \in \hat{E}(K)$ which parametrizes the extension B of E by $\mathbb{G}_m$. Hence from (†) we conclude that $F_q(\log_G(\tilde{s})) = F_q$.

Let $s \in B(K)$ be the projection of $\tilde{s}$, and $p \in E(K)$ the projection of s. By the discussion in Section 5.1 of [Bertrand et al. 2013], $F_{pq}(\log_B(s)) = F_q(\log_G(\tilde{s}))$. Therefore, $F_q = F_{pq} = F_{pq}(\log_B(s))$.

Now as $\tilde{x} \in LG(K)$ maps onto the constant point $\check{x} \in L\widetilde{E}(\mathbb{C})$, so also $\tilde{s}$ maps onto a constant point $\check{p} \in \widetilde{E}(\mathbb{C})$ and hence $p \in E(\mathbb{C})$. So we are in Case (SC2) of the proof of the Main Lemma of [Bertrand et al. 2013, Section 6], namely p constant while q nonconstant. The conclusion of (SC2) is that $\log_B(s)$ is transcendental over F_{pq} if p is nontorsion. So the previous equality forces $p \in E(\mathbb{C})$ to be torsion.

Let $\tilde{s}_{\mathrm{tor}} \in G(K)$ be a torsion point lifting p, hence $\tilde{s} - \tilde{s}_{\mathrm{tor}}$ is a K-point of the kernel of the surjection $G \to E$. Thus $\tilde{s} = \tilde{s}_{\mathrm{tor}} + \delta + \beta$ where $\beta \in \mathbb{G}_a(K)$ and $\delta \in \mathbb{G}_m(K)$. Taking logs, putting again $\xi = \tilde{x} - \tilde{x}_1$, and using that $\log_G(-)$ restricted to $\mathbb{G}_a(K)$ is the identity, we see that $\tilde{x} = \xi + \log_G(\tilde{s}_{\mathrm{tor}}) + \log_G(\delta) + \beta = \xi' + \log_{\mathbb{G}_m}(\delta) + \beta$

where $\xi' \in (LG)^\partial$. It follows that $\ell = \log_{\mathbb{G}_m}(\delta) \in K_G^\sharp = F_q$. But by Lemma 1 of [Bertrand et al. 2013] (proof of Main Lemma in isotrivial case, but reversing roles of p and q), such ℓ is transcendental over F_q unless δ is constant.

Hence $\delta \in \mathbb{G}_m(\mathbb{C})$, whereby $\log_{\mathbb{G}_m}(\delta) \in L\mathbb{G}_m(\mathbb{C})$ so is in $(LG)^\partial(K^{\mathrm{diff}})$, and we conclude that $\tilde{x} - \beta \in (LG)^\partial(K^{\mathrm{diff}})$. As also $\tilde{x} - \beta \in LG(K)$, from Claim III in Section 5.3 of [Bertrand and Pillay 2010] (alternatively, using the fact that $K_{LG}^\sharp = F_q$ has transcendence degree 2 over K), we conclude that $\tilde{x} - \beta \in L\mathbb{G}_m(\mathbb{C})$ whereby $\tilde{x} \in L\mathbb{G}_a(K) + L\mathbb{G}_m(\mathbb{C})$, contradicting that x projects onto a nonzero element of LE. This contradiction completes the proof of Claim 1 and hence of Case I of Corollary 2.10.

Case II. The point $\bar{x} \in LE(K) \setminus LE(\mathbb{C})$ is a nonconstant point of $LE(K) = LE_0(K)$.

Let $\tilde{y} \in G(K^{\mathrm{diff}})$ be such that $\partial\ell n_G(\tilde{y}) = a = \partial_{LG}(\tilde{x})$. Let $\check{y}$ be the projection of $\tilde{y}$ to $\widetilde{E}$. Hence $\partial\ell n_{\widetilde{E}}(\check{y}) = \partial_{L\widetilde{A}}(\check{x})$ (remembering that $\check{x}$ is the projection of $\tilde{x}$ to $L\widetilde{E}$). Noting that $\check{x}$ lifts $\bar{x} \in LE(K)$, and using our case hypothesis, we can apply Corollary 2.9 to E to conclude that $\mathrm{tr.deg}(K(\check{y})/K) = 2$ with Galois group $\widetilde{E}^\partial(K^{\mathrm{diff}}) = \widetilde{E}(\mathbb{C})$. (In fact as E is constant this is already part of the Ax–Kolchin framework and appears in [Bertrand 2008].)

Claim 2. We have $\mathrm{tr.deg}\big(K_G^\sharp(\check{y})/K_G^\sharp\big) = 2$.

Proof of Claim 2. Fact 3.4 applies to the current situation, showing that $K_G^\sharp$ is a directed union of Picard–Vessiot extensions of K each with Galois group some product of $\mathbb{G}_m^n(\mathbb{C})$'s. As there are no proper algebraic subgroups of $\widetilde{E}(\mathbb{C}) \times \mathbb{G}_m^n(\mathbb{C})$ projecting onto each factor, it follows from the Galois theory that $\check{y}$ is independent from $K_G^\sharp$ over K, yielding Claim 2.

Now $K_G^\sharp(\tilde{y})/K_G^\sharp$ is a differential Galois extension with Galois group of the form $H^\partial(K^{\mathrm{diff}})$ where H is a connected algebraic ∂-subgroup of G. So H^∂ projects onto the (differential) Galois group of $K_G^\sharp(\check{y})$ over $K_G^\sharp$, which by Claim 2 is $\widetilde{E}^\partial(K^{\mathrm{diff}})$. In particular, H projects onto $\widetilde{E}$. If H is a proper subgroup of G, then projecting H and $\tilde{E}$ to B and E, respectively, shows that B splits (up to isogeny), so $B = B_0$ is constant, contradicting the current assumptions. Hence the (differential) Galois group of $K_G^\sharp(\tilde{y})$ over $K_G^\sharp$ is $G^\partial(K^{\mathrm{diff}})$, whereby $\mathrm{tr.deg}\big(K_G^\sharp(\tilde{y})/K_G^\sharp\big)$ is 3. This concludes the proof of Corollary 2.10. □

4C. ***An Ax–Schanuel conjecture.*** As a conclusion to the first two themes of the paper, we may say that both at the Galois-theoretic level and for Lindemann–Weierstrass, we have obtained rather definitive results for families of abelian varieties, and working over a suitable base $K^\sharp$. There remain open questions for families of semiabelian varieties, such as Conjecture 2.3, as well as dropping the restriction on the toric part in Theorems 2.6 and 2.7 and Corollaries 2.8 and 2.10. It also remains to formulate a qualitative description of $\mathrm{tr.deg}\big(K^\sharp(\exp_B(x))/K^\sharp\big)$

where B is a semiabelian variety over K of dimension > 2, and $x \in LB(K)$, under the nondegeneracy hypothesis that $x \notin LH + LB_0(\mathbb{C})$ for any proper semiabelian subvariety H of B.

Before turning to our third theme, it seems fitting to propose a more general *Ax–Schanuel* conjecture for families of abelian varieties:

Conjecture 4.1. *Let A be an abelian variety over $K = \mathbb{C}(S)$ for a curve $S/\mathbb{C}$, and let F be the field of meromorphic functions on some disc in S. Let $K^\sharp$ now denote $K^\sharp_{L\widetilde{A}}$ (which contains $K^\sharp_{\widetilde{A}}$). Let $\tilde{x}$, $\tilde{y}$ be F-rational points of $L\widetilde{A}$, $\widetilde{A}$, respectively, such that $\exp_{\widetilde{A}}(\tilde{x}) = \tilde{y}$, and let y be the projection of $\tilde{y}$ on A. Assume that $y \notin H + A_0(\mathbb{C})$ for any proper algebraic subgroup H of A. Then* $\operatorname{tr.deg}(K^\sharp(\tilde{x}, \tilde{y})/K^\sharp) \geq \dim(\widetilde{A})$.

We point out that the assumption concerns y, and not the projection x of $\tilde{x}$ to LA. Indeed, the conclusion would in general not hold true under the weaker hypothesis that $x \notin LH + LA_0(\mathbb{C})$ for any proper abelian subvariety H of A. As a counterexample, take for A a simple nonconstant abelian variety over K, and for $\tilde{x}$ a nonzero period of $L\widetilde{A}$. Then $x \neq 0$ satisfies the hypothesis above and $\tilde{x}$ is defined over $K^\sharp = K^\sharp_{L\widetilde{A}}$, but $\tilde{y} = \exp_{\widetilde{A}}(\tilde{x}) = 0$, so $\operatorname{tr.deg}(K^\sharp(\tilde{x}, \tilde{y})/K^\sharp) = 0$.

Finally, here is a concrete corollary of the conjecture. Let $\boldsymbol{E} : y^2 = x(x-1)(x-t)$ be the universal Legendre elliptic curve over $S = \mathbb{C} \setminus \{0, 1\}$, and let $\omega_1(t)$, $\omega_2(t)$ be a basis of the group of periods of $\boldsymbol{E}$ over some disk, so $K^\sharp = K^\sharp_{L\widetilde{E}}$ is the field generated over $K = \mathbb{C}(t)$ by ω_1, ω_2 and their first derivatives. Let $\wp = \wp_t(z)$, $\zeta = \zeta_t(z)$ be the standard Weierstrass functions attached to $\{\omega_1(t), \omega_2(t)\}$. For $g \geq 1$, consider $2g$ algebraic functions $\alpha_1^{(i)}(t), \alpha_2^{(i)}(t) \in K^{\mathrm{alg}}$, $i = 1, \ldots, g$, and assume that the vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \alpha_1^{(1)} \\ \alpha_2^{(1)} \end{pmatrix}, \ldots, \begin{pmatrix} \alpha_1^{(g)} \\ \alpha_2^{(g)} \end{pmatrix}$$

are linearly independent over $\mathbb{Z}$. Then the $2g$ functions

$$\wp\big(\alpha_1^{(i)}\omega_1 + \alpha_2^{(i)}\omega_2\big), \zeta\big(\alpha_1^{(i)}\omega_1 + \alpha_2^{(i)}\omega_2\big), \quad i = 1, \ldots, g,$$

of the variable t are algebraically independent over $K^\sharp$. In the language of [Bertrand et al. 2013, Section 3.3], this says in particular that a g-tuple of $\mathbb{Z}$-linearly independent local analytic sections of $\boldsymbol{E}/S$ with algebraic *Betti* coordinates forms a generic point of $\boldsymbol{E}^g/S$. Such a statement is not covered by our Lindemann–Weierstrass results, which concern analytic sections with algebraic logarithms.

5. Manin maps

5A. ***Injectivity.*** We here prove Theorem 2.13 and Proposition 2.15. Both statements will follow fairly quickly from Fact 5.1 below, which is Theorem 4.3 of [Bertrand 2011] and relies on the strongest version of "Manin–Chai", namely formula (2*) from Section 4.1 of [Bertrand 2011]. We should mention that a more direct proof of Proposition 2.15 can be extracted from the proof of Proposition J.2 (Manin–Coleman) in [Bertrand and Pillay 2010]. But we will stick with the current proof below, as it provides a good introduction to the counterexample in Section 5B.

We set up some notation: K is $\mathbb{C}(t)^{\mathrm{alg}}$ as usual, A is an abelian variety over K, and A_0 is the $\mathbb{C}$-trace of A. For $y \in \widetilde{A}(K)$, we let $\bar{y}$ be its image in $A(K)$. Let $b = \partial\ell n_{\widetilde{A}}(y)$. We consider the differential system in unknown x:

$$\nabla_{L\widetilde{A}}(x) = b,$$

where we write $\nabla_{L\widetilde{A}}$ for $\partial_{L\widetilde{A}}$. Let $K^{\sharp}_{L\widetilde{A}}$ be the differential field generated, over K, by $(L\widetilde{A})^{\partial}(K^{\mathrm{diff}})$. So for x a solution in $L\widetilde{A}(K^{\mathrm{diff}})$, the differential Galois group of $K^{\sharp}_{L\widetilde{A}}(x)$ over $K^{\sharp}_{L\widetilde{A}}$ pertains to Picard–Vessiot theory, and is well-defined as a $\mathbb{C}$-subspace of the $\mathbb{C}$-vector space $(L\widetilde{A})^{\partial}(K^{\mathrm{diff}})$.

Fact 5.1 (A = any abelian variety over $K = \mathbb{C}(t)^{\mathrm{alg}}$). *Let $y \in \widetilde{A}(K)$. Let B be the smallest abelian subvariety of A such that a multiple of $\bar{y}$ by a nonzero integer is in $B + A_0(\mathbb{C})$. Let x be a solution of $\nabla_{L\widetilde{A}}(-) = b$ in $L\widetilde{A}(K^{\mathrm{diff}})$. Then the differential Galois group of $K^{\sharp}_{L\widetilde{A}}(x)$ over $K^{\sharp}_{L\widetilde{A}}$ is $(L\widetilde{B})^{\partial}(K^{\mathrm{diff}})$. In particular,* $\mathrm{tr.deg}\big(K^{\sharp}_{L\widetilde{A}}(x)/K^{\sharp}_{L\widetilde{A}}\big) = \dim \widetilde{B} = 2\dim B$.

Proof of Theorem 2.13. Here, the abelian variety A has $\mathbb{C}$-trace 0. By assumption we have $y_1, \dots, y_n \in A(K)$ and $a_1, \dots, a_n \in \mathbb{C}$ not all 0 such that

$$a_1\mu_A(y_1) + \cdots + a_n\mu_A(y_n) = 0$$

in $L\widetilde{A}(K)/\partial\ell n_{\widetilde{A}}(W_A)$. Lifting y_i to $\tilde{y}_i \in \widetilde{A}(K)$, we derive that

$$a_1\partial\ell n_{\widetilde{A}}(\tilde{y}_1) + \cdots + a_n\partial\ell n_{\widetilde{A}}(\tilde{y}_n) = \partial\ell n_{\widetilde{A}}(z)$$

for some $z \in W_A$. Via our identification of W_A with LW_A we write the right hand side as $\nabla_{L\widetilde{A}}z$ with $z \in LW_A \subset L\widetilde{A}$. Let $\tilde{x}_i \in L\widetilde{A}$ be such that $\nabla_{L\widetilde{A}}(\tilde{x}_i) = \partial\ell n_{\widetilde{A}}(\tilde{y}_i)$. Hence $a_1\tilde{x}_1 + \cdots + a_n\tilde{x}_n - z \in (L\widetilde{A})^{\partial}$, and there exists $d \in (L\widetilde{A})^{\partial}$ such that

$$a_1\tilde{x}_1 + \cdots + a_n\tilde{x}_n - d = z \in LW_A.$$

Suppose for a contradiction that $y_1, \dots, y_n$ are linearly independent with respect to $\mathrm{End}(A)$. Then no multiple of $y = (y_1, \dots, y_n)$ by a nonzero integer lies in any proper abelian subvariety B of the traceless abelian variety $A^n = A \times \cdots \times A$. By Fact 5.1, we have $\mathrm{tr.deg}(K^{\sharp}(\tilde{x}_1, \dots, \tilde{x}_n)/K^{\sharp}) = \dim(\widetilde{A^n})$, where we have set $K^{\sharp} := K^{\sharp}_{L\widetilde{A^n}} = K^{\sharp}_{L\widetilde{A}}$. So, the points $\tilde{x}_1, \dots, \tilde{x}_n$ of $L\widetilde{A}$ are generic and independent over $K^{\sharp}$. Hence, because $a_1, \dots, a_n$ are in $\mathbb{C}$ and therefore $K^{\sharp}$, it follows that $a_1\tilde{x}_1 + \cdots + a_n\tilde{x}_n$ is a generic point of $L\widetilde{A}$ over $K^{\sharp}$. And as d is a $K^{\sharp}$-rational point of $(L\widetilde{A})^{\partial}$, also $a_1\tilde{x}_1 + \cdots + a_n\tilde{x}_n - d = z$ is a generic point of $L\widetilde{A}$ over $K^{\sharp}$, so cannot lie in its strict subspace LW_A. This contradiction concludes the proof of Theorem 2.13. □

Proof of Proposition 2.15. We use the same notation as at the end of Section 2D, and recall that A is traceless. Furthermore, the functoriality of $M_{K,A}$ in A allows us to assume that A is a simple abelian variety.

Step I. We show, as in the proof of Theorem 2.13, that if $M_{K,A}(y_1), \dots, M_{K,A}(y_n)$ are $\mathbb{C}$-linearly dependent, then $y_1, \dots, y_n$ are $\mathrm{End}(A)$-linearly dependent. Indeed, assume that $a_i \in \mathbb{C}$ are not all 0 and that $a_1 M_{K,A}(y_1) + \cdots + a_n M_{K,A}(y_n) = 0$ in the target space $L\widetilde{A}(K)/\nabla(L\widetilde{A}(K))$. Lifting y_i to $\tilde{y}_i \in \widetilde{A}(K)$, we derive that

$$a_1 \partial\ell n_{\widetilde{A}}(\tilde{y}_1) + \cdots + a_n \partial\ell n_{\widetilde{A}}(\tilde{y}_n) \in \nabla(L\widetilde{A}(K)).$$

Letting $\tilde{x}_i \in L\widetilde{A}(K^{\mathrm{diff}})$ be such that $\nabla\tilde{x}_i = \partial\ell n_{\widetilde{A}}(\tilde{y}_i)$, we obtain a K-rational point $z \in L\widetilde{A}(K)$ such that

$$a_1\tilde{x}_1 + \cdots + a_n\tilde{x}_n - z := d \in (L\widetilde{A})^{\partial}(K^{\mathrm{diff}}).$$

Taking $K^{\sharp} := K^{\sharp}_{L\widetilde{A}}$ as in the proof of Theorem 2.13, we get

$$\mathrm{tr.deg}(K^{\sharp}(\tilde{x}_1, \dots, \tilde{x}_n)/K^{\sharp}) < \dim(\widetilde{A}^n).$$

Hence by Fact 5.1, some integral multiple of $(y_1, \dots, y_n)$ lies in a proper abelian subvariety of A^n, whereby $y_1, \dots, y_n$ are $\mathrm{End}(A)$-linearly dependent.

Step II. Assuming that $y_1, \dots, y_n$ are $\mathrm{End}(A)$-linearly dependent, given by Step I, as well as the relation on the point d above with not all $a_i = 0$, we will show that the points y_i are $\mathbb{Z}$-linearly dependent. Equivalently we will show that if a similar relation holds with the a_i linearly independent over $\mathbb{Z}$, then $y = (y_1, \dots, y_n)$ is a torsion point of A^n. Let $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)$. Let B be the connected component of the Zariski closure of the group $\mathbb{Z}\cdot y$ of multiples of y in A^n. By Fact 5.1, the differential Galois group of $K^{\sharp}(\tilde{x})$ over $K^{\sharp} := K^{\sharp}_{L\widetilde{A}}$ is $(L\widetilde{B})^{\partial}$. More precisely, the set of $\sigma(\tilde{x}) - \tilde{x}$ as σ varies in $\mathrm{Aut}_{\partial}(K^{\sharp}(\tilde{x})/K^{\sharp})$ is precisely $(L\widetilde{B})^{\partial} \subseteq (L\widetilde{A}^n)^{\partial}$. Since z and d are defined over $K^{\sharp}$, the relation on d implies that

$$\forall(\tilde{c}_1, \dots, \tilde{c}_n) \in (L\widetilde{B})^{\partial},\ a_1\tilde{c}_1 + \cdots + a_n\tilde{c}_n = 0.$$

Let now

$$\mathcal{B} = \left\{\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathrm{End}(A))^n = \mathrm{Hom}(A, A^n) : \alpha(A) \subseteq B \subset A^n\right\}.$$

Claim. Assume that $a_1, \dots, a_n$ are linearly independent over $\mathbb{Z}$. Then any $\alpha \in \mathcal{B}$ is identically 0.

It follows from the claim that $B = 0$ and hence some multiple of y by a nonzero integer vanishes, namely y is a torsion point of A^n. This completes the proof of Step II, hence of Proposition 2.15, and we are now reduced to proving the claim.

Proof of claim. Since A is simple, $\mathrm{End}(A)$ is an order in a simple algebra D over $\mathbb{Q}$. For $i = 1, \dots, n$, denote by $\rho(\alpha_i)$ the $\mathbb{C}$-linear map induced on $(L\widetilde{A})^\partial$ by the endomorphism α_i of A. So we view $(L\widetilde{A})^\partial$ as a complex representation, of degree $2\dim A$, of the $\mathbb{Z}$-algebra $\mathrm{End}(A)$, or more generally, of D. Let f^2 be the dimension of D over its center F, let e be the degree of F over $\mathbb{Q}$ and let R be a reduced representation of D, viewed as a complex representation of degree ef. As the representation ρ is defined over $\mathbb{Q}$ (since it preserves the Betti homology), ρ is equivalent to the direct sum $R^{\oplus r}$ of $r = 2\dim A/ef$ copies of R (cf. [Shimura and Taniyama 1961, Section 5.1]). Furthermore,

$$R : D \to \mathrm{Mat}_f(F \otimes \mathbb{C}) \simeq (\mathrm{Mat}_f(\mathbb{C}))^e \subset \mathrm{Mat}_{ef}(\mathbb{C})$$

extends by $\mathbb{C}$-linearity to an injection $R \otimes 1 : D \otimes \mathbb{C} \simeq (\mathrm{Mat}_f(\mathbb{C}))^e \subset \mathrm{Mat}_{ef}(\mathbb{C})$.

Recall now that $a_1\tilde{c}_1 + \cdots + a_n\tilde{c}_n = 0$ for any $(\tilde{c}_1, \dots, \tilde{c}_n)$ in $(L\widetilde{B})^\partial$. Applied to the image under $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathcal{B}$ of the generic element of $(L\widetilde{A})^\partial$, this relation implies that

$$a_1\rho(\alpha_1) + \cdots + a_n\rho(\alpha_n) = 0 \in \mathrm{End}_{\mathbb{C}}((L\widetilde{A})^\partial).$$

So $a_1R(\alpha_1) + \cdots + a_nR(\alpha_n) = 0$ in $(\mathrm{Mat}_f(\mathbb{C}))^e$. From the injectivity of $R \otimes 1$ on $D \otimes \mathbb{C}$ and the $\mathbb{Z}$-linear independence of the a_i, we derive that each $\alpha_i \in D$ vanishes, hence $\alpha = 0$, proving the claim. $\square$

5B. *A counterexample.* We conclude with the promised counterexample to the injectivity of $\mu_A \otimes 1$, namely Proposition 2.14.

Construction of A. We will use Yves André's example of a simple traceless abelian variety A over $\mathbb{C}(t)^{\mathrm{alg}}$ with $0 \neq U_A \subsetneq W_A$ (cf. [Bertrand and Pillay 2010], just before Remark 3.10). Since $U_A \neq W_A$, this A is not constant, but we will derive this property and the simplicity of A from another argument, borrowed from [Lange and Birkenhake 1992, Chapter 9 §6].

We start with a CM field F of degree $2k$ over $\mathbb{Q}$, over a totally real number field F_0 of degree $k \geq 2$, and denote by $\{\sigma_1, \bar{\sigma}_1, \dots, \sigma_k, \bar{\sigma}_k\}$ the complex embeddings of F. We further fix the CM type $S := \{\sigma_1, \bar{\sigma}_1, 2\sigma_2, \dots, 2\sigma_k\}$. By [Lange and Birkenhake 1992, Chapter 9 §6], we can attach to S and to any $\tau \in \mathcal{H}$ (the Poincaré half-plane, or equivalently, the open unit disk) an abelian variety $A = A_\tau$ of dimension $g = 2k$ and an embedding of F into $\mathrm{End}(A) \otimes \mathbb{Q}$ such that the representation r of F on W_A is given by the type S. The representation ρ of F on $L\widetilde{A}$ is then $r \oplus \bar{r}$, equivalent to twice the regular representation. (The notation used by [Lange and Birkenhake 1992] here read: $e_0 = k, d = 1, m = 2, r_1 = s_1 = 1, r_2 = \cdots = r_{e_0} = 2, s_2 = \cdots = s_{e_0} = 0$, so, the product of the $\mathcal{H}_{r_i,s_i}$ of [loc. cit.] is just $\mathcal{H}$. Also, [loc. cit.] considers the more standard "analytic" representation of F on the Lie algebra $LA = L\widetilde{A}/W_A$, which is $\bar{r}$ in our notation.)

From the bottom of [Lange and Birkenhake 1992, p. 271], one infers that the moduli space of such abelian varieties A_τ is an analytic curve $\mathscr{H}/\Gamma$. But Shimura has shown that it can be compactified to an algebraic curve $\mathscr{X}$ (cf. [Lange and Birkenhake 1992, p. 247]). So, we can view the universal abelian variety $A_\tau = A$ of this moduli space as an abelian variety over $\mathbb{C}(\mathscr{X})$, hence as an abelian variety A over $K = \mathbb{C}(t)^{\mathrm{alg}}$. This will be our A; it is by construction not constant — and it is a fourfold if we take $k = 2$, as we will in what follows.

Finally, since A is the general element over $\mathscr{H}/\Gamma$, Theorem 9.1 of [Lange and Birkenhake 1992] and the hypothesis $k \geq 2$ imply that $\mathrm{End}(A) \otimes \mathbb{Q}$ is *equal* to F. Therefore, A is a simple abelian variety, necessarily traceless since it is not constant. We denote by $\mathbb{O}$ the order $\mathrm{End}(A)$ of F.

Action of F and of ∇ on $L\widetilde{A}$. For simplicity, we will now restrict to the case $k = 2$, but the general case (requiring $2k$ points) would work in exactly the same way. So, F is a totally imaginary quadratic extension of a real quadratic field F_0, and $L\widetilde{A}$ is 8-dimensional. As said in [Bertrand and Pillay 2010], and by definition of the CM-type S, the action ρ of F splits $L\widetilde{A}$ into eigenspaces for its irreducible representations σ's, as follows:

- $W_A = D_{\sigma_1} \oplus D_{\bar\sigma_1} \oplus P_{\sigma_2}$, where the D's are lines and P_{σ_2} is a plane;
- LA lifts to $L\widetilde{A}$ into $D'_{\sigma_1} \oplus D'_{\bar\sigma_1} \oplus P_{\bar\sigma_2}$, with the same notation.

Since $\nabla := \nabla_{L\widetilde{A}} = \partial_{L\widetilde{A}}$ commutes with the action ρ of F and since A is not constant, we infer that the maximal ∂-submodule of W_A is

$$U_A = P_{\sigma_2},$$

while $W_A + \nabla(W_A) = \Pi_{\sigma_1} \oplus U_A \oplus \Pi_{\bar\sigma_1}$, with the planes

$$\Pi_{\sigma_1} = D_{\sigma_1} \oplus D'_{\sigma_1},$$
$$\Pi_{\bar\sigma_1} = D_{\bar\sigma_1} \oplus D'_{\bar\sigma_1},$$

each stable under ∇ (just as is $P_{\bar\sigma_2}$, of course). In fact, for our proof, we only need to know that $P_{\sigma_2} \subset U_A$.

Now let $\tilde{y} \in \widetilde{A}(K)$ be a lift of a point $y \in A(K)$. Going into a complex analytic setting, we choose a logarithm $\tilde{x} \in L\widetilde{A}(K^{\mathrm{diff}})$ of $\tilde{y}$, locally analytic on a small disk in $\mathscr{X}(\mathbb{C})$. Let further $\alpha \in \mathbb{O}$, which canonically lifts to $\mathrm{End}(\widetilde{A})$. Then $\rho(\alpha)\tilde{x}$ is a logarithm of $\alpha \cdot \tilde{y} \in \widetilde{A}(K)$, and therefore satisfies

$$\nabla(\rho(\alpha)\tilde{x}) = \partial\ell n_{\widetilde{A}}(\alpha \cdot \tilde{y}).$$

In fact, this appeal to analysis is not necessary; the formula just says that $\partial\ell n_{\widetilde{A}}$ (and ∇) commutes with the actions of $\mathbb{O}$. But once one $\tilde{y}$ and one $\tilde{x}$ are chosen, it will be crucial, for the desired relation $(\mathfrak{R})$ following Proposition 2.14, that we take these $\rho(\alpha)\tilde{x}$ as solutions to the equations on the $\mathbb{O}$-orbit of $\tilde{y}$.

Concretely, if

$$\tilde{x} = x_{\sigma_2} \oplus x_{\sigma_1} \oplus x_{\bar{\sigma}_1} \oplus x_{\bar{\sigma}_2}$$

is the decomposition of $\tilde{x}$ in

$$L\widetilde{A} = P_{\sigma_2} \oplus \Pi_{\sigma_1} \oplus \Pi_{\bar{\sigma}_1} \oplus P_{\bar{\sigma}_2},$$

then for any $\alpha \in \mathbb{O}$, we have

$$\rho(\alpha)(\tilde{x}) = \sigma_2(\alpha)x_{\sigma_2} \oplus \sigma_1(\alpha)x_{\sigma_1} \oplus \bar{\sigma}_1(\alpha)x_{\bar{\sigma}_1} \oplus \bar{\sigma}_2(\alpha)x_{\bar{\sigma}_2}.$$

Conclusion. Let $y \in A(K)$ be a nontorsion point of the simple abelian variety A, for which we choose at will a lift $\tilde{y}$ to $\widetilde{A}(K)$ and a logarithm $\tilde{x} \in L\widetilde{A}(K^{\text{diff}})$. Let $\{\alpha_1, \dots, \alpha_4\}$ be an integral basis of F over $\mathbb{Q}$. We will consider the 4 points $y_i = \alpha_i \cdot y$ of $A(K)$, $i = 1, \dots, 4$. Since the action of $\mathbb{O}$ on A is faithful, they are linearly independent over $\mathbb{Z}$. For each $i = 1, \dots, 4$, we consider the lift $\tilde{y}_i = \alpha_i \tilde{y}$ of y_i to $L\widetilde{A}(K)$, and set as above $\tilde{x}_i = \rho(\alpha_i)\tilde{x}$, which satisfies $\nabla(\tilde{x}_i) = \partial\ell n_{\widetilde{A}}\tilde{y}_i$.

We claim that there exist complex numbers $a_1, \dots, a_4$, not all zero, such that

$$u := a_1\tilde{x}_1 + \dots + a_4\tilde{x}_4 = \big(a_1\rho(\alpha_1) + \dots + a_4\rho(\alpha_4)\big)(\tilde{x}) \in U_A(K^{\text{diff}}),$$

i.e., such that in the decomposition above of $L\widetilde{A} = P_{\sigma_2} \oplus \Pi_{\sigma_1} \oplus \Pi_{\bar{\sigma}_1} \oplus P_{\bar{\sigma}_2}$, the components of $u = u_{\sigma_2} \oplus u_{\sigma_1} \oplus u_{\bar{\sigma}_1} \oplus u_{\bar{\sigma}_2}$ on the last three planes vanish.

The whole point is that the complex representation $\hat{\sigma}^{\oplus 2}$ of F which ρ induces on $\Pi_{\sigma_1} \oplus \Pi_{\bar{\sigma}_1} \oplus P_{\bar{\sigma}_2}$ is twice the representation $\hat{\sigma} := \sigma_1 \oplus \bar{\sigma}_1 \oplus \bar{\sigma}_2$ of F on $\mathbb{C}^3$, and so, does not contain the full regular representation of F. More concretely, the 4 vectors $\hat{\sigma}(\alpha_1), \dots, \hat{\sigma}(\alpha_4)$ of $\mathbb{C}^3$ are of necessity linearly dependent over $\mathbb{C}$, so, there exists a nontrivial linear relation

$$a_1\hat{\sigma}(\alpha_1) + \dots + a_4\hat{\sigma}(\alpha_4) = 0 \quad \text{in } \mathbb{C}^3$$

(where the complex numbers a_i lie in the normal closure of F). Therefore, *any* element $\tilde{x}_{\hat{\sigma}} = (x_{\sigma_1}, x_{\bar{\sigma}_1}, x_{\bar{\sigma}_2})$ of $\Pi_{\sigma_1} \oplus \Pi_{\bar{\sigma}_1} \oplus P_{\bar{\sigma}_2}$ satisfies

$$\big(a_1\hat{\sigma}^{\oplus 2}(\alpha_1) + \dots + a_4\hat{\sigma}^{\oplus 2}(\alpha_4)\big)\tilde{x}_{\hat{\sigma}} = 0 \quad \text{in } \Pi_{\sigma_1} \oplus \Pi_{\bar{\sigma}_1} \oplus P_{\bar{\sigma}_2}$$

(viewing each $\hat{\sigma}^{\oplus 2}(\alpha_i)$ as a (6×6) diagonal matrix inside the (8×8) diagonal matrix $\rho(\alpha_i)$), i.e., the 3 plane-components $u_{\sigma_1}, u_{\bar{\sigma}_1}, u_{\bar{\sigma}_2}$ of u all vanish, and u indeed lies in P_{σ_2}, and so in U_A.

The existence of such a point $u = a_1\tilde{x}_1 + \dots + a_4\tilde{x}_4$ in $U_A(K^{\text{diff}})$ establishes relation $(\mathfrak{R})$ of Section 2D, and concludes the proof of Proposition 2.14.

References

[Benoist et al. 2014] F. Benoist, E. Bouscaren, and A. Pillay, "On function field Mordell–Lang and Manin–Mumford", preprint, 2014. arXiv 1404.6710

[Bertrand 2001] D. Bertrand, "Unipotent radicals of differential Galois group and integrals of solutions of inhomogeneous equations", *Math. Ann.* **321**:3 (2001), 645–666. MR 2002k:12014 Zbl 1016.12007

[Bertrand 2008] D. Bertrand, "Schanuel's conjecture for non-isoconstant elliptic curves over function fields", pp. 41–62 in *Model theory with applications to algebra and analysis*, vol. 1, edited by Z. Chatzidakis et al., London Math. Soc. Lecture Note Ser. **349**, Cambridge University Press, 2008. MR 2010c:14046 Zbl 1215.14050

[Bertrand 2009] D. Bertrand, "Théories de Galois différentielles et transcendance", *Ann. Inst. Fourier (Grenoble)* **59**:7 (2009), 2773–2803. MR 2011f:12004 Zbl 1226.12002

[Bertrand 2011] D. Bertrand, "Galois descent in Galois theories", pp. 1–24 in *Arithmetic and Galois theories of differential equations*, edited by L. Di Vizio and T. Rivoal, Sémin. Congr. **23**, Société Mathématique de France, Paris, 2011. MR 3076077 Zbl 1316.12004

[Bertrand and Pillay 2010] D. Bertrand and A. Pillay, "A Lindemann–Weierstrass theorem for semi-abelian varieties over function fields", *J. Amer. Math. Soc.* **23**:2 (2010), 491–533. MR 2011f:12005 Zbl 1276.12003

[Bertrand et al. 2013] D. Bertrand, D. Masser, A. Pillay, and U. Zannier, "Relative Manin–Mumford for semi-abelian surfaces", preprint, 2013. arXiv 1307.1008

[Buium and Cassidy 1999] A. Buium and P. J. Cassidy, "Differential algebraic geometry and differential algebraic groups: from algebraic differential equation to Diophantine geometry", pp. 567–636 in *Selected works of Ellis Kolchin with commentary*, edited by H. Bass et al., American Mathematical Society, Providence, RI, 1999.

[Chai 1991] C.-L. Chai, "A note on Manin's theorem of the kernel", *Amer. J. Math.* **113**:3 (1991), 387–389. MR 93b:14036 Zbl 0759.14017

[Coleman 1990] R. F. Coleman, "Manin's proof of the Mordell conjecture over function fields", *Enseign. Math.* (2) **36**:3-4 (1990), 393–427. MR 92e:11069 Zbl 0729.14018

[Hrushovski and Sokolović 1994] E. Hrushovski and Ž. Sokolović, "Strongly minimal sets in differentially closed fields", 1994.

[Lange and Birkenhake 1992] H. Lange and C. Birkenhake, *Complex abelian varieties*, Grundlehren der Mathematischen Wissenschaften **302**, Springer, Berlin, 1992. MR 94j:14001 Zbl 0779.14012

[Manin 1963] J. I. Manin, "Рациональные точки алгебраических крибых над функциональными полями", *Izv. Akad. Nauk SSSR Ser. Mat.* **27**:6 (1963), 1395–1440. Translated as "Rational points of algebraic curves over function fields", pp. 189–234 in *Fifteen papers on algebra*, Transl. Amer. Math. Soc. (2) **50**, Amer. Math. Soc., Providence, RI, 1966. MR 28 #1199 Zbl 0166.16901

[Marker and Pillay 1997] D. Marker and A. Pillay, "Differential Galois theory, III: Some inverse problems", *Illinois J. Math.* **41**:3 (1997), 453–461. MR 99m:12011 Zbl 0927.03065

[Mumford and Fogarty 1982] D. Mumford and J. Fogarty, *Geometric invariant theory*, 2nd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete **34**, Springer, Berlin, 1982. MR 86a:14006 Zbl 0504.14008

[Pillay 1997] A. Pillay, "Differential Galois theory II", *Ann. Pure Appl. Logic* **88**:2-3 (1997), 181–191. MR 99m:12010 Zbl 0927.03064

[Pillay 1998] A. Pillay, "Differential Galois theory I", *Illinois J. Math.* **42**:4 (1998), 678–699. MR 99m:12009 Zbl 0916.03028

[Pillay 2004] A. Pillay, "Algebraic *D*-groups and differential Galois theory", *Pacific J. Math.* **216**:2 (2004), 343–360. MR 2005k:12007 Zbl 1093.12004

[Shimura and Taniyama 1961] G. Shimura and Y. Taniyama, *Complex multiplication of abelian varieties and its applications to number theory*, Publications of the Mathematical Society of Japan **6**, The Mathematical Society of Japan, Tokyo, 1961. MR 23 #A2419 Zbl 0112.03502

Received February 15, 2015. Revised July 9, 2015.

DANIEL BERTRAND
INSTITUT DE MATHÉMATIQUES DE JUSSIEU
UNIVERSITÉ PIERRE ET MARIE CURIE
CASE 247
4 PLACE JUSSIEU
75252 PARIS CEDEX 05
FRANCE
daniel.bertrand@imj-prg.fr

ANAND PILLAY
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NOTRE DAME
281 HURLEY HALL
NOTRE DAME, IN 46556
UNITED STATES
Anand.Pillay.3@nd.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.83

MORSE AREA AND SCHARLEMANN–THOMPSON WIDTH FOR HYPERBOLIC 3-MANIFOLDS

DIANE HOFFOSS AND JOSEPH MAHER

Scharlemann and Thompson define a numerical complexity for a 3-manifold using handle decompositions of the manifold. We show that for compact hyperbolic 3-manifolds, this is linearly related to a definition of metric complexity in terms of the areas of level sets of Morse functions.

1. Introduction

Let M be a closed Riemannian 3-manifold, and let $f : M \to \mathbb{R}$ be a Morse function; i.e., f is a smooth function, all of whose critical points are nondegenerate, and for which distinct critical points have distinct images in $\mathbb{R}$. We define the *area* of f to be the maximum area of any level set $F_t = f^{-1}(t)$ over all points $x \in \mathbb{R}$. We define the *Morse area* of M to be the infimum of the area of all Morse functions $f : M \to \mathbb{R}$.

For hyperbolic 3-manifolds, the hyperbolic metric is a topological invariant by Mostow rigidity, and the critical points of a Morse function determine a handle decomposition of the manifold, so one might hope that Morse area is related to a topological measure of complexity defined in terms of handle decompositions of the manifold. We show that Morse area is linearly related to a definition of topological complexity we call *Scharlemann–Thompson width* or *linear width*, and which we now describe.

For a closed (possibly disconnected) surface S, we define the complexity, or genus, of S to be the sums of the genera of each connected component. For a compact (possibly disconnected) surface with boundary, we define the genus of S

MSC2010: primary 57M50; secondary 57N16.
Keywords: hyperbolic 3-manifold, Heegaard splitting, Morse function, Scharlemann–Thompson width.

to be the genus of the surface obtained by capping off all boundary curves with discs. We shall write $|\partial S|$ for the number of boundary components of S.

A *handlebody* is a compact 3-manifold with boundary, homeomorphic to the regular neighborhood of a graph in $\mathbb{R}^3$. Up to homeomorphism, a handlebody is determined by the genus g of its boundary surface. Every 3-manifold M has a *Heegaard splitting*, which is a decomposition of the manifold into two handlebodies. This immediately gives a notion of complexity for a 3-manifold, called the *Heegaard genus*, which is the smallest genus of any Heegaard splitting of the 3-manifold.

There is a refinement of this, due to Scharlemann and Thompson [1994], which we now describe. Let S be a closed surface, which need not be connected. A *compression body* C is a compact 3-manifold with boundary, constructed by attaching some number of 2-handles to one side $S \times \{0\}$ of $S \times I$. We do not require compression bodies to be connected. We shall refer to $S \times \{1\}$ as the *top boundary* $\partial_+ C$ of the compression body, and the other boundary components of C as the *lower boundary* $\partial_- C$. The lower boundary may be disconnected, even if C is connected, and any 2-sphere components are capped off with 3-balls. In particular, if a maximal number of nonparallel 2-handles are attached, then the resulting compression body is a handlebody, so a handlebody is a special case of a compression body. A *generalized Heegaard splitting*, which we shall call a *linear splitting*, is a decomposition of a closed 3-manifold M into a linearly ordered sequence of compression bodies $C_1, \dots C_{2n}$, which need not be connected, such that the upper boundary of an odd numbered compression body C_{2i+1} is equal to the top boundary of the compression body C_{2i+2}, and the lower boundary of C_{2i+1} is equal to the lower boundary of the previous compression body C_{2i}. For the even numbered compression bodies C_{2i}, the top boundary is equal to the upper boundary of C_{2i-1}, and the lower boundary is equal to the lower boundary of C_{2i+1}. In the case of the first and last compression bodies C_1 and C_{2n}, the lower boundaries are empty. Let H_i be the sequence of surfaces consisting of the upper boundaries of the compression bodies C_{2i-1} and C_{2i}; these are often referred to as the *odd surfaces*, and the surfaces corresponding to the lower boundaries as the *even surfaces*. A linear splitting has a natural height function, i.e., a Morse function onto $\mathbb{R}$, in which each odd or even surface, which may not be connected, is the pre-image of a single point, and the compression bodies are the pre-images of the closed intervals determined by these points.

The complexity $c(H_i)$ of the surface H_i is the genus of H_i, i.e., the sum of the genera of each connected component, and the *width* of the linear splitting is the maximum value of $c(H_i)$ over all upper boundaries. The *Scharlemann–Thompson width*, which we shall also refer to as the *linear width*, of a 3-manifold M is the minimum width over all possible linear splittings. As a Heegaard splitting is a special case of a linear splitting, the Heegaard genus of M is an upper bound for the linear width of M.

There is a refinement of linear width known as *thin position*, which we discuss when we use it in Section 3.

Results. In order to bound Morse area in terms of linear width, we shall assume the following result announced by Pitts and Rubinstein [1986] (see also [Rubinstein 2005]).

Theorem 1.1 [Pitts and Rubinstein 1986; Rubinstein 2005]. *Let M be a Riemannian 3-manifold with a strongly irreducible Heegaard splitting. Then the Heegaard surface is isotopic to a minimal surface, or to the boundary of a regular neighborhood of a nonorientable minimal surface with a small tube attached vertically in the I-bundle structure.*

A full proof of this result has not yet appeared in the literature, though recent progress has been made by Colding and De Lellis [2003], De Lellis and Pellandrini [2010], and Ketover [2013].

We shall show this:

Theorem 1.2. *There is a constant $K > 0$ such that for any closed hyperbolic 3-manifold,*

$$K(\text{linear width}(M)) \leqslant \text{Morse area}(M) \leqslant 4\pi(\text{linear width}(M)), \tag{1}$$

where the right-hand bound holds assuming Theorem 1.1.

Our methods are effective, and the constant K may be estimated using a bound on the Margulis constant for $\mathbb{H}^3$, though we omit the details of this calculation, as our methods seem unlikely to give an optimal constant.

Outline. In Section 2, we show how to bound linear width in terms of Morse area. A bound on the Morse area of M gives a Morse function $f\colon M \to \mathbb{R}$ with bounded area level sets, but with no a priori bound on the topological complexity of the level sets.

We use a Voronoi decomposition of M to give a polyhedral approximation of the Morse function, which we now describe in a simple case. Let $\mathcal{V}$ be a Voronoi decomposition of M in which every Voronoi cell V_i is a topological ball, and has size bounded above and below; i.e., there is an $\epsilon > 0$ such that $B(x_i, \epsilon/2) \subset V_i \subset B(x_i, \epsilon)$, where x_i is the center of the Voronoi cell. Let M_t be the sublevel set of the Morse function, i.e., $M_t = f^{-1}((-\infty, t])$. The sublevel sets are a monotonically increasing collection of subsets of M, which start off empty, and eventually contain all of M; so, in particular, for each Voronoi cell V_i, there is a t_i such that the volume of $M_t \cap B(x_i, \epsilon/2)$ is exactly half the volume of $B(x_i, \epsilon/2)$, and we shall call t_i the *cell splitter* for the Voronoi cell V_i. Furthermore, we may assume that the t_i are distinct for distinct Voronoi cells. This gives a linear order to the Voronoi cells, and we wish to show that constructing the manifold by adding the Voronoi cells in this order gives a bounded linear-width handle decomposition for M. Let P_t be the union of the Voronoi cells whose cell splitters t_i are at most t. Each Voronoi cell is a ball, with a bounded number of faces, so adding a Voronoi cell corresponds to adding a bounded number of handles. It remains to show that the boundary of each P_t has genus bounded in terms of the area of the level set $F_t = f^{-1}(t)$. Let V_i and V_j be two adjacent Voronoi cells, with V_i contained in P_t and V_j outside P_t, so their common face is a subset of ∂P_t. Consider the sequence of balls $B(x, \epsilon/2)$, as x runs along the geodesic from x_i to x_j. At least half the volume of $B(x_i, \epsilon/2)$ is contained in M_t, and at most half the volume of $B(x_j, \epsilon/2)$ is contained in M_t; so there is an x such that exactly half the volume of $B(x, \epsilon/2)$ is contained in M_t, and so there is a lower bound on the area of $F_t \cap B(x, \epsilon/2)$. Therefore, a bound on the area of F_t gives a bound on the number of faces of ∂P_t. As each face has a bounded number of edges, this gives a bound on the genus of ∂P_t, and hence a bound on the linear width of M, though this bound depends on ϵ.

In order to produce a bound which works for any compact hyperbolic manifold M, we use the Margulis lemma and the thick-thin decomposition for hyperbolic manifolds. There is constant μ, called a Margulis constant, such that any compact hyperbolic manifold may be decomposed into a thick part X_μ, where each point has injectivity radius greater than μ, and a thin part, where each point has injectivity radius at most μ, and which is a disjoint union of solid tori. If we choose ϵ sufficiently small, then we may choose a Voronoi decomposition of the thick part in which each Voronoi cell has size bounded above and below, and run the argument in the previous paragraph to control the genus of ∂P_t inside the thick part. We do not control the complexity of ∂P_t in the thin part, but as each component of the thin part is a solid torus, we may cap off $\partial P_t \cap X_\mu$ with surfaces parallel to $P_t \cap \partial X_\mu$, while still obtaining bounds on the genus. In order to bound the number of handles corresponding to adding a Voronoi cell, we use a result of Kobayashi and Rieck [2011] which gives bounds on the topological complexity of the intersection of a Voronoi cell with the thin part.

The key problem for the upper bound is that the techniques of Pitts and Rubinstein use sweepouts, so although their minimax construction produces a sweepout of bounded area, we do not know how to directly replace a bounded area sweepout with a bounded area foliation. However, the upper bound is obtained in recent work of Colding and Gabai [2015], using work of Colding and Minicozzi [2015] on the mean curvature flow, and we describe their results in Section 3.

2. Morse area bounds Scharlemann–Thompson width

In this section we show that we can bound the Scharlemann–Thompson width of a hyperbolic manifold in terms of its Morse area. We will approximate level sets by surfaces which are unions of faces of Voronoi cells, and we start by describing the properties of the Voronoi decompositions that we will use.

Voronoi cells. We will approximate the level sets of f by surfaces consisting of faces of Voronoi cells. We now describe in detail the Voronoi cell decompositions we shall use, and their properties.

A *polygon* in $\mathbb{H}^3$ is a compact convex subset of a hyperbolic plane whose boundary consists of a finite number of geodesic segments. A *polyhedron* in $\mathbb{H}^3$ is a convex topological 3-ball in $\mathbb{H}^3$ whose boundary consists of a finite collection of polygons. A *polyhedral cell decomposition* of $\mathbb{H}^3$ is a cell decomposition in which every 3-cell is a polyhedron, each 2-cell is a polygon, and the edges are all geodesic segments. We say a cell decomposition of a hyperbolic manifold M is *polyhedral* if its preimage in the universal cover gives a polyhedral cell decomposition of $\mathbb{H}^3$.

Let $X = \{x_i\}$ be a discrete collection of points in 3-dimensional hyperbolic space $\mathbb{H}^3$. The Voronoi cell V_i determined by $x_i \in X$ consists of all points of M which are closer to x_i than any other $x_j \in X$, i.e.,

$$V_i = \{x \in \mathbb{H}^3 \mid d(x, x_i) \leqslant d(x, x_j) \text{ for all } x_j \in X\}.$$

We shall call x_i the *center* of the Voronoi cell V_i, and we shall write $\mathcal{V} = \{V_i\}$ for the collection of Voronoi cells determined by X. Voronoi cells are convex sets in $\mathbb{H}^3$, and hence topological balls. The set of points equidistant from both x_i and x_j is a totally geodesic hyperbolic plane in $\mathbb{H}^3$. A *face* F of the Voronoi decomposition consists of all points which lie in two distinct Voronoi cells V_i and V_j, so F is contained in a geodesic plane. An *edge* e of the Voronoi decomposition consists of all points which lie in three distinct Voronoi cells V_i, V_j and V_k, which is a geodesic segment, and a *vertex* v is a point lying in four distinct Voronoi cells V_i, V_j, V_k and V_l. By general position, we may assume that all edges of the Voronoi decomposition are contained in exactly three distinct faces, the collection of vertices is a discrete set, and there are no points which lie in more than four distinct Voronoi cells. We shall call such a Voronoi decomposition a *regular* Voronoi decomposition, and it is a polyhedral decomposition of $\mathbb{H}^3$ if every cell is compact. As each edge is 3-valent, and each vertex is 4-valent, this implies that the dual cell structure is a simplicial triangulation of $\mathbb{H}^3$, which we shall refer to as the *dual triangulation*. The dual triangulation may be realised in $\mathbb{H}^3$ by choosing the vertices to be the centers x_i of the Voronoi cells and the edges to be geodesic segments connecting the vertices, and we shall always assume that we have done this. In this case, the triangles and tetrahedra are geodesic triangles and geodesic tetrahedra in $\mathbb{H}^3$.

Given a collection of points $X = \{x_i\}$ in a hyperbolic 3-manifold M, let $\widetilde{X}$ be the pre-image of X in the universal cover of M, which is isometric to $\mathbb{H}^3$. We say a subset of $\mathbb{H}^3$ is *equivariant* if it is preserved by the covering translations determined by the quotient M. As $\widetilde{X}$ is equivariant, the k-skeleton of the corresponding Voronoi cell decomposition $\mathcal{V}$ of $\mathbb{H}^3$ is also equivariant for $0 \leqslant k \leqslant 3$, as are the k-skeletons of the dual triangulation.

We now show that the interior of each Voronoi cell V is mapped down homeomorphically by the covering projection. Suppose y is a point in the interior of a Voronoi cell V with center x, so $d(x, y) < d(x', y)$ for any other $x' \in X$. Let g be a covering translation, which is an isometry, so $d(x, y) = d(gx, gy)$. As covering translations act freely, this implies that gy lies in the interior of the Voronoi cell corresponding to $gx \neq x$. Therefore interior(V) has disjoint translates under the group of covering translations, and so is mapped down homeomorphically into M, though the covering projection may identify distinct faces of a Voronoi cell under projection into M.

By abuse of notation, we shall refer to the resulting polyhedral decomposition of M as the Voronoi decomposition $\mathcal{V}$ of M. By general position, we may assume that $\mathcal{V}$ is regular. The dual triangulation also projects down to a triangulation of M, which we will also refer to as the dual triangulation, though this triangulation may no longer be simplicial.

We say a collection $X = \{x_i\}$ of points in M is *ϵ-separated* if the distance between any pair of points is at least ϵ, i.e., $d(x_i, x_j) \geqslant \epsilon$ for all $i \neq j$.

Definition 2.1. Let M be a compact hyperbolic 3-manifold. We say a Voronoi decomposition $\mathcal{V}$ is ϵ-regular if it is regular and it arises from a maximal collection of ϵ-separated points.

We shall write $B(x, r)$ for the closed metric ball of radius r about x in M,

$$B(x, r) = \{y \in M \mid d(x, y) \leqslant r\},$$

which need not be a topological ball. As the cells of an ϵ-regular Voronoi decomposition are determined by a maximal collection of ϵ-separated points in M, each Voronoi cell is contained in a metric ball of radius ϵ about its center. Furthermore, as the points x_i are distance at least ϵ apart, each Voronoi cell contains a metric ball of radius $\epsilon/2$ about its center, i.e.,

$$B(x_i, \epsilon/2) \subset V_i \subset B(x_i, \epsilon).$$

One useful property of ϵ-regular Voronoi decompositions is that the boundary of any union of Voronoi cells is an embedded surface, in fact an embedded normal surface in the dual triangulation, as we now describe.

A *simple arc* in the boundary of a tetrahedron is a properly embedded arc in a face of the tetrahedron with endpoints in distinct edges. A *triangle* in a tetrahedron is a properly embedded disc whose boundary is a union of three simple arcs, and a *quadrilateral* is a properly embedded disc whose boundary is the union of four simple arcs. A *normal surface* in a triangulated 3-manifold is a surface that intersects each tetrahedron in a union of normal triangles and quadrilaterals.

Proposition 2.2. *Let M be a compact hyperbolic manifold, and let $\mathcal{V}$ be an ϵ-regular Voronoi decomposition. Let P be a union of Voronoi cells in $\mathcal{V}$, and let S be the boundary of P. Then S is an embedded surface in M.*

Proof. The collection of Voronoi cells P intersects a tetrahedron T in the dual triangulation in a regular neighborhood of the vertices of T. If a tetrahedron T has one or three vertices corresponding to Voronoi cells in P, then S intersects T in a single normal triangle. If T has exactly two vertices corresponding to Voronoi cells in P, then S intersects T in a single normal quadrilateral. Therefore S consists of at most one triangle or quadrilateral in each tetrahedron, and so is an embedded normal surface. □

We shall write $\mathrm{inj}_M(x)$ for the injectivity radius of M at x, i.e., the radius of the largest embedded ball in M centered at x. We shall write $\mathrm{inj}(M)$ for the injectivity radius of M, which is defined to be

$$\mathrm{inj}(M) = \inf_{x \in M} \mathrm{inj}_M(x).$$

We shall say a Voronoi cell V_i with center x_i is a *deep* Voronoi cell if the injectivity radius at x_i is at least 4ϵ, i.e., $\mathrm{inj}_M(x_i) \geqslant 4\epsilon$, and, in particular, this implies that the metric ball $B(x_i, 3\epsilon)$ is a topological ball. We shall also call centers, faces, edges and vertices of deep Voronoi cells deep. We shall write $\mathcal{W}$ for the subset of $\mathcal{V}$ consisting of deep Voronoi cells. The fact that a deep Voronoi cell V_i has injectivity radius at least 4ϵ at its center x_i guarantees that every adjacent Voronoi cell is also a topological ball.

We now show that there are bounds, which only depend on ϵ, on the number of faces of a deep Voronoi cell, and the number of edges and faces of a deep Voronoi cell.

Proposition 2.3. *Let M be a compact hyperbolic* 3*-manifold with an ϵ-regular Voronoi decomposition $\mathcal{V}$, and let $\mathcal{W}$ be the collection of deep Voronoi cells. Then there is a number J, which only depends on ϵ, such that each deep Voronoi cell $W_i \in \mathcal{W}$ has at most J faces, edges and vertices.*

Proof. Let W be a deep Voronoi cell with center x, and with faces $F_1, \dots, F_n$. Let x_i be the center of the Voronoi cell W_i adjacent to the face F_i. As W is deep, the Voronoi cell W_i is also a topological ball.

If two Voronoi cells share a common face, then the distance between their centers is at most 2ϵ. Therefore all of the centers of the Voronoi cells corresponding to the faces of W are contained in the metric ball $B(x, 2\epsilon)$. This implies that the balls of radius $\epsilon/2$ around the x_i are contained in the metric ball $B(x, 5\epsilon/2)$. As the $B(x_i, \epsilon/2)$ are all disjoint, this implies that the number of faces is at most

$$J_1 = \frac{\mathrm{vol}_{\mathbb{H}^3}(B(x, 5\epsilon/2))}{\mathrm{vol}_{\mathbb{H}^3}(B(x, \epsilon/2))}.$$

Note that J_1 is also an upper bound for the maximum number of edges in any face of a Voronoi cell because every edge of that face is contained in another face in that cell. So the total number of edges is at most J_1^2, and by the formula for Euler characteristic, the number of vertices is at most $J_1^2 + C$. Therefore we may choose J to be $J_1^2 + 2$. □

A similar volume bound argument to the one above proves the following:

Proposition 2.4. *Let M be a compact hyperbolic 3-manifold with an ϵ-regular Voronoi decomposition $\mathcal{V}$. Then there is a number L, which depends only on ϵ, such that for any deep Voronoi center x_i, the number of Voronoi centers contained in $B(x_i, 3\epsilon)$ is at most L.*

Polyhedral surfaces. We may choose a Morse function $f: M \to \mathbb{R}$ such that the complexity of f is within some small $\delta > 0$ of the infimum, i.e.,

$$\mathrm{area}(F_t) \leqslant \text{Morse area}(M) + \delta$$

for all $t \in \mathbb{R}$. We now describe how to use the Morse function f to give a linear ordering to the Voronoi cells in $\mathcal{V}$.

Definition 2.5. Let M be a compact hyperbolic 3-manifold, and let $f: M \to \mathbb{R}$ be a Morse function. Given $t \in \mathbb{R}$, define the *sublevel set of M at t*, which we shall denote M_t, to be the subset of M consisting of the union of all level sets F_t with $t \in (-\infty, t]$, i.e.,

$$M_t = f^{-1}((-\infty, t]).$$

For t sufficiently small, M_t is the empty set, and for t sufficiently large, M_t is equal to all of M. The region M_t varies continuously in t and is monotonically increasing in t.

Definition 2.6. Let M be a compact hyperbolic 3-manifold with an ϵ-regular Voronoi decomposition $\mathcal{V}$. Let $f: M \to \mathbb{R}$ be a Morse function. For each Voronoi cell V_i with center x_i, there is a unique $t_i \in \mathbb{R}$ such that the surface F_{t_i} divides the metric ball $B(x_i, \epsilon/2)$ exactly in half by volume, i.e.,

$$\mathrm{vol}(M_t \cap B(x_i, \epsilon/2)) = \tfrac{1}{2}\,\mathrm{vol}(B(x_i, \epsilon/2)).$$

We call this t_i the *cell splitter* of V_i.

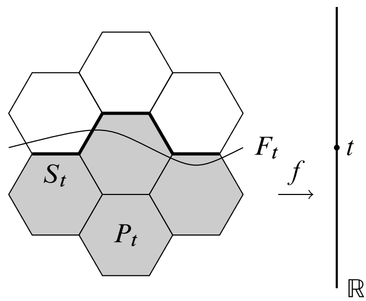


Figure 1. A polyhedral surface S_t determined by a level set F_t.

Definition 2.7. We say that a Morse function $f: M \to \mathbb{R}$ is *generic* with respect to a Voronoi decomposition $\mathcal{V}$ if the cell splitters for distinct Voronoi cells V_i correspond to distinct points $t_i \in \mathbb{R}$, and no cell splitter is also a critical point for the Morse function. We say a point $t \in \mathbb{R}$ is *generic* if it is not a critical point for the Morse function, and is not a cell splitter.

We may assume that f is generic by an arbitrarily small perturbation of f, and we shall always assume that f is generic from now on.

Definition 2.8. Let M be a compact hyperbolic 3-manifold with an ϵ-regular Voronoi decomposition $\mathcal{V}$, and let $f: M \to \mathbb{R}$ be a generic Morse function. Let $\mathcal{V}$ be the Voronoi decomposition ordered by the order inherited from the cell splitters t_i. Given $t \in \mathbb{R}$, Let M_t be the sublevel set of M at t. We define P_t, the *polyhedral approximation to* M_t, to be the union of the Voronoi cells V_i with $t_i \leqslant t$, and call $S_t = \partial P_t$ the *polyhedral surface* determined by $t \in \mathbb{R}$.

The polyhedral surface S_t is a union of faces of the Voronoi cells, and so is a normal surface in the dual triangulation. We shall write $\|S_t\|$ for the number of Voronoi faces the polyhedral surface S_t contains. We shall write $\|S_t \cap \mathcal{W}\|$ for the number of faces in the polyhedral surface $S_t \cap \mathcal{W}$, which may have boundary. A schematic picture of a polyhedral surface is given in Figure 1.

In this section, we will show the following bound on the complexity of the polyhedral surface in the deep part $\mathcal{W}$.

Proposition 2.9. *Let M be a compact hyperbolic* 3*-manifold, with an ϵ-regular Voronoi decomposition $\mathcal{V}$, deep part $\mathcal{W}$, and a generic Morse function $f: M \to \mathbb{R}$. For $t \in \mathbb{R}$, let S_t be the polyhedral surface associated to t. Then there is a constant K, which only depends on ϵ, such that*

$$|\partial(S_t \cap \mathcal{W})| \leqslant K \operatorname{area}(F_t),$$
$$\operatorname{genus}(S_t \cap \mathcal{W}) \leqslant K \operatorname{area}(F_t).$$

In particular, this bounds the genus of $S_t \cap \mathcal{W}$ as a constant times the Morse width of M, where the constant depends only on ϵ. We start by showing that the area of the level sets bounds the number of faces of the polyhedral surface in the deep part $\mathcal{W}$.

Proposition 2.10. *Let M be a compact hyperbolic 3-manifold, with an ϵ-regular Voronoi decomposition $\mathcal{V}$, deep part $\mathcal{W}$, and a generic Morse function $f\colon M \to \mathbb{R}$. For $t \in \mathbb{R}$, let S_t be the polyhedral surface associated to t. Then there is a constant K, which only depends on ϵ, such that*

$$\|S_t \cap \mathcal{W}\| \leqslant K \operatorname{area}(F_t).$$

Proof. Let P_t be the polyhedral approximation to M_t. Let C be a face of $S_t \cap \mathcal{W}$, and let W_i and W_j be the two adjacent Voronoi cells in $\mathcal{V}$. Up to relabeling, we may assume that W_i is contained in P_t, and W_j is not. Let γ be a geodesic connecting x_i to x_j, and consider $B(s, \epsilon/2)$ for $s \in \gamma$. As W_i and W_j are deep, the metric balls $B(x_i, \epsilon/2)$, $B(x_j, \epsilon/2)$ and $B(s, \epsilon/2)$ are all topological balls, isometric to the ball $B(x, \epsilon/2)$ in $\mathbb{H}^3$. At least half of the volume of $B(x_i, \epsilon/2)$ is contained in P_t, and strictly less than half of the volume of $B(x_j, \epsilon/2)$ is contained in P_t, so there is some $s \in \gamma$ such that exactly half the volume of $B(s, \epsilon/2)$ is contained in P_t. There is a constant A, depending only on ϵ, such that any surface dividing a ball in hyperbolic space into regions of equal volume has area at least A. In fact, we may take A to be the area of the equatorial disc, which is $2\pi(\cosh(\epsilon/2) - 1)$; see, for example, [Bachman et al. 2004].

Recall that the Voronoi decomposition has a dual triangulation in which each edge is a geodesic segment, and we shall write Γ for the geodesic graph in M formed by the 1-skeleton of the dual triangulation. We shall write Γ_d for the subset of Γ consisting of vertices corresponding to deep Voronoi cells, and edges connecting two deep Voronoi cells, and we shall refer to this as the *deep graph*. Each geodesic edge between two deep Voronoi cells has length strictly less than 2ϵ. Therefore the choice of geodesic is unique for the Voronoi cells in $\mathcal{W}$, as its length is smaller than the injectivity radius at each deep Voronoi cell center x_i. By Proposition 2.3, the geodesic dual graph Γ_d has valence at most J.

Claim 2.11. Consider a collection of points $\{s_i\}$ such that each point s_i lies in a distinct edge γ_i of the deep graph Γ_d. Then any ball $B(s_i, \epsilon/2)$ intersects at most L other balls $B(s_j, \epsilon/2)$, where L is the constant from Proposition 2.4.

Proof of Claim 2.11. If two balls $B(s_i, \epsilon/2)$ and $B(s_j, \epsilon/2)$ intersect, then the distance between their corresponding edges γ_i and γ_j is at most ϵ, and so there is a pair of vertices, $x_k \in \gamma_i$ and $x_l \in \gamma_j$ with $d(x_k, x_l) \leqslant 3\epsilon$. By Proposition 2.4, there are at most L other vertices within distance 3ϵ of a given vertex. Therefore the total number of balls intersecting $B(s_i, \epsilon/2)$ is at most L, which only depends on ϵ. □

If there are N faces in S_t then there are at least N/L disjoint balls $B(s_i, \epsilon/2)$, each containing a part of F_t of area at least A. Therefore, the total number of faces is at most

$$\|S_t \cap \mathcal{W}\| \leqslant \frac{L}{A}\,\mathrm{area}(F_t), \tag{2}$$

where the constants only depend on ϵ, as required. □

We now show that the bound on the number of faces of S_t in the deep part $\mathcal{W}$ gives a bound on the genus of $S_t \cap \mathcal{W}$.

Proposition 2.12. *Let M be a compact hyperbolic 3-manifold, with an ϵ-regular Voronoi decomposition $\mathcal{V}$, deep part $\mathcal{W}$, and a generic Morse function $f\colon M \to \mathbb{R}$. For $t \in \mathbb{R}$, let S_t be the polyhedral surface associated to t. Then there is a constant J, which only depends on ϵ, such that*

$$|\partial(S_t \cap \mathcal{W})| \leqslant J\|S_t \cap \mathcal{W}\|,$$
$$\mathrm{genus}(S_t \cap \mathcal{W}) \leqslant J\|S_t \cap \mathcal{W}\|,$$

where J is the constant from Proposition 2.3.

Proof. We shall write S for S_t to simplify notation. The first bound follows as each boundary component must contain at least one edge, so the number of boundary components is at most the number of edges in $S \cap \mathcal{W}$, which is at most $J\|S \cap \mathcal{W}\|$ by Proposition 2.3.

We shall write $\widehat{S}$ for the surface $S \cap \mathcal{W}$ with all boundary curves capped off with discs. Recall that the genus of a disconnected surface is the sum of the genera of each component, and this in turn is equal to the number of connected components minus half the Euler characteristic, i.e.,

$$\mathrm{genus}(\widehat{S}) = |\widehat{S}| - \tfrac{1}{2}\chi(\widehat{S}),$$

where $|\widehat{S}|$ is the number of connected components of $\widehat{S}$.

As capping off with discs does not change the number of connected components, this is at most the number of connected components of $S \cap \mathcal{W}$, which is at most the number of faces $\|S \cap \mathcal{W}\|$. Furthermore, capping off boundary components with discs may only increase the Euler characteristic, so

$$\mathrm{genus}(\widehat{S}) \leqslant \|S \cap \mathcal{W}\| - \tfrac{1}{2}\chi(S \cap \mathcal{W}).$$

Therefore

$$\mathrm{genus}(\widehat{S}) \leqslant \|S \cap \mathcal{W}\| - \tfrac{1}{2}(V - E + F),$$

where V, E and F are the numbers of vertices, edges and faces of $S \cap \mathcal{W}$. As each face of a deep Voronoi cell has at most J edges, this implies

$$\mathrm{genus}(\widehat{S}) \leqslant (1 + J/2)\|S \cap \mathcal{W}\|.$$

As we may assume that J is at least 2, this gives the second inequality. □

Proposition 2.9 now follows immediately from Propositions 2.10 and 2.12.

Capped surfaces. We have constructed surfaces with bounded complexity in the deep part. The complement of the deep part is contained in a union of solid tori by the Margulis lemma, and we now explain how to cap off the surfaces in the deep part with surfaces in the solid tori to produce bounded genus surfaces.

We will use the Margulis lemma and the *thick-thin* decomposition for finite volume hyperbolic 3-manifolds, which we now review. Given a number $\mu > 0$, let $X_\mu = M_{[\mu,\infty)}$ be the *thick part* of M, i.e., the union of all points x of M with $\mathrm{inj}_M(x) \geqslant \mu$. We shall refer to the closure of the complement of the thick part as the *thin part* and denote it by $T_\mu = \overline{M \setminus X_\mu}$.

The Margulis lemma states that there is a constant $\mu_0 > 0$, such that for any compact hyperbolic 3-manifold, the thin part is a disjoint union of solid tori, and each of these solid tori is a regular metric neighborhood of an embedded closed geodesic of length less than μ_0. We shall call a number μ_0 for which this result holds a *Margulis constant* for $\mathbb{H}^3$. If μ_0 is a Margulis constant for $\mathbb{H}^3$, then so is μ for any $0 < \mu < \mu_0$, and furthermore, given μ and μ_0, there is a number $\delta > 0$ such that $N_\delta(T_\mu) \subseteq T_{\mu_0}$. For the remainder of this section, we shall fix a pair of numbers (μ, ϵ) such that there are Margulis constants $0 < \mu_1 < \mu < \mu_2$, a number δ such that $N_\delta(T_\mu) \subseteq T_{\mu_2} \setminus T_{\mu_1}$, and $\epsilon = \frac{1}{4}\min\{\mu_1, \delta\}$. We shall call (μ, ϵ) a choice of MV-constants for $\mathbb{H}^3$. This choice of constants ensures that the deep part $\mathcal{W}$ is nonempty.

Let (μ, ϵ) be a choice of MV-constants, and consider an ϵ-regular Voronoi decomposition of M. The fact that $N_\delta(T_\mu) \subseteq T_{\mu_2} \setminus T_{\mu_1}$ means that we adjust the boundary of T_μ by an arbitrarily small isotopy so that it is transverse to the Voronoi cells, and we will assume that we have done this for the remainder of this section. Our choice of ϵ implies that the thick part X_μ is contained in the Voronoi cells in the deep part, i.e., $X_\mu \subset \bigcup_{W_i \in \mathcal{W}} W_i$, so, in particular, $\partial X_\mu = \partial T_\mu$ is contained in the deep part. Furthermore, as $\epsilon < \delta$, each deep Voronoi cell hits at most one component of T_μ.

Each boundary component of the surface $S_t \cap X_\mu$ is contained in T_μ, so $S_t \cap X_\mu$ is a properly embedded surface in X_μ. We now bound the number of boundary components of $S_t \cap X_\mu$ in terms of the number of polyhedral faces in the deep part, $\|S_t \cap \mathcal{W}\|$.

Proposition 2.13. *Let (μ, ϵ) be MV-constants, and let M be a compact hyperbolic 3-manifold with thin part T_μ, an ϵ-regular Voronoi decomposition $\mathcal{V}$ with deep part $\mathcal{W}$, and a generic Morse function f. Let S_t be a polyhedral surface in M. Then there is a constant J, depending only on ϵ, such that*

$$\mathrm{genus}(S_t \cap X_\mu) \leqslant J\|S_t \cap \mathcal{W}\|,$$
$$|\partial(S_t \cap X_\mu)| \leqslant 2J\|S_t \cap \mathcal{W}\|.$$

Proof. The properly embedded surface $S_t \cap X_\mu$ is obtained from $S_t \cap \mathcal{W}$ by cutting $S_t \cap \mathcal{W}$ along simple closed curves and discarding some connected components. This does not increase the genus, which gives the first bound, using Proposition 2.12.

Each face C of a Voronoi cell is a totally geodesic convex polygon, and a component of T_μ lifts to a convex set in the universal cover $\mathbb{H}^3$, so $C \cap T_\mu$ is a convex subset of C. Therefore $C \cap \partial T_\mu$ consists either of a simple closed curve, or a collection of properly embedded arcs which have at most two endpoints in each edge of C, so there are at most as many arcs as the number of edges of C. Therefore, the number of components of $C \cap \partial T_\mu$ has at most (number of faces of $S_t \cap \mathcal{W}$) plus (number of edges of $S_t \cap \mathcal{W}$) components, and this gives the second bound, again using Proposition 2.12. □

We now wish to cap off the properly embedded surfaces $S_t \cap X_\mu$ with properly embedded surfaces in T_μ to form closed surfaces. For each torus T_i in ∂T_μ, let U_i be the subsurface consisting of $\partial T_i \cap M_t$. Let $S_t^+ = (S \cap X_\mu) \cup \bigcup_i U_i$, and we shall call the resulting closed surface the *T-capped surface* S_t^+. We now bound the genus of the resulting T-capped surfaces.

Proposition 2.14. *Let (μ, ϵ) be MV-constants, and let M be a compact hyperbolic 3-manifold with thin part T_μ, an ϵ-regular Voronoi decomposition $\mathcal{V}$ with deep part $\mathcal{W}$, and a generic Morse function f. Let S_t be a polyhedral surface in M, and let S_t^+ be the corresponding T-capped surface. Then there is a constant K, depending only on ϵ, such that*

$$\text{genus}(S_t^+) \leqslant K \,\text{area}(F_t).$$

Furthermore, for any finite collection of generic points $\{u_i\}$ in $\mathbb{R}$, the corresponding T-capped surfaces $\{S_{u_i}^+\}$ may be isotoped to be disjoint.

Proof. By Proposition 2.10, it suffices to bound the genus of the T-capped surface in terms of the number of polyhedral faces of the surface in the deep part. We will show

$$\text{genus}(S_t^+) \leqslant (5J+1)\|S_t \cap \mathcal{W}\|,$$

where J is the constant from Proposition 2.3, which only depends on ϵ.

Each surface U_i is a subsurface of a torus, and so consists of a union of planar surfaces, together with at most one surface which is a torus with (possibly many) boundary components.

Capping off components of $S_t \cap X_\mu$ with planar surfaces cannot increase the genus by more than twice the number of boundary components, and capping off with punctured tori increases the genus by at most the number of boundary components, plus the number of punctured tori. As each Voronoi cell hits at most one component of T_μ, there are at most $\|S_t \cap \mathcal{W}\|$ components of the U_i surfaces which may be punctured tori. This implies

$$\text{genus}(S_t^+) \leqslant \text{genus}(S_t \cap X_\mu) + 2|\partial(S_t \cap X_\mu)| + \|S_t \cap \mathcal{W}\|.$$

Using the bounds from Proposition 2.13, we obtain

$$\text{genus}(S_t^+) \leqslant (5J+1)\|S_t \cap \mathcal{W}\|,$$

as required.

Finally, we show that for any finite collection of generic points $\{u_i\}$ in $\mathbb{R}$, we may isotope the corresponding T-capped surfaces to be disjoint. To simplify notation, given a generic point $u_i \in \mathbb{R}$, we will write M_i and S_i for the corresponding polyhedral approximation and polyhedral surface determined by u_i.

For any two distinct points $u_i < u_j$ in $\mathbb{R}$, the polyhedral approximation M_i is a strict subset of M_j, so S_i and S_j are disjoint normal surfaces. Let T be a single solid torus component of T_μ. Take a small product neighborhood $\partial T \times [0,1]$, and choose the parameterization such that $\partial T \times \{0\}$ is equal to ∂T, and the product neighborhood is contained in T. Let U_i be the subsurface of ∂T given by $\partial T \cap M_i$. Let U_i^+ be the properly embedded surface in the product $\partial T \times [0,1]$ given by placing U_i at depth i/n, together with a product neighborhood of the boundary ∂U_i connecting U_i to the boundary of S_i, i.e.,

$$U_i^+ = (U_i \times \{i/n\}) \cup (\partial U_i \times [0, i/n]).$$

As the submanifolds M_i are strictly nested, the subsurfaces U_i are also strictly nested, i.e., $U_i \subset U_j$ for $i < j$, and so the resulting surfaces $S_i \cup U_i^+$ are disjoint. □

Bounded handles. We now bound the number of handles between a pair of T-capped surfaces S_i^+ and S_j^+ whose corresponding points in u_i and u_j in $\mathbb{R}$ bound an interval containing a single cell splitter.

Proposition 2.15. *Let (μ, ϵ) be MV-constants and M be a hyperbolic 3-manifold with an ϵ-regular Voronoi decomposition $\mathcal{V}$ and a generic Morse function $f : M \to \mathbb{R}$. Let $u_1 < u_2$ be a pair of points in $\mathbb{R}$, which bound an interval containing a single cell splitter t. Let S_1^+ and S_2^+ be T-capped surfaces corresponding to the level sets for u_1 and u_2, bounding regions P_1 and P_2, with $P_1 \subset P_2$. Then P_2 is homeomorphic to a manifold obtained from P_1 by adding at most $60J^2 \max\{\|S_i \cap \mathcal{W}\|\}$ handles, where J is the constant from Proposition 2.3, which depend only on ϵ.*

We start by observing that attaching a compression body P to a 3-manifold Q by a subsurface S of the upper boundary component of P, requires a number of handles which is bounded in terms of the Heegaard genus of P, and the number of boundary components of the attaching surface.

Proposition 2.16. *Let Q be a compact 3-manifold with boundary and let $R = Q \cup P$, where P is a compression body of genus g, attached to Q by a homeomorphism* *along a (possibly disconnected) subsurface S contained in the upper boundary component of P of genus g. Then R is homeomorphic to a 3-manifold obtained from Q by the addition of at most* $(4\,\mathrm{genus}(P) + 2|\partial S|)$ 1-*and* 2-*handles, where* $|\partial S|$ *is the number of boundary components of S.*

Proof. Recall that the genus of a disconnected surface with boundary is the sum of the genus of each closed component obtained by capping off all boundary components with discs. Therefore, the genus of S is at most the genus of P. For a connected surface of genus g with b boundary components, cutting along a nonseparating arc with endpoints in the same boundary component produces a surface of genus $g-1$ with $b+1$ boundary components. A planar surface with b boundary components may be cut into at most b discs by $b-1$ nonseparating arcs. Therefore we may choose at most $2g+b$ arcs which cut the surface S into at most $g+b$ discs. We can add a 1-handle to Q for each arc, and then a 2-handle for each disc, to produce a manifold Q_+ which is homeomorphic to Q union a regular neighborhood of ∂P. We may then form R by adding at most g 2-handles. The total number of 1- and 2-handles required is at most $4g+2b$. □

Proof of Proposition 2.15. Let V be the Voronoi cell corresponding to the single cell splitter t contained in the interval $[u_1, u_2]$. The surfaces S_1^+ and S_2^+ are parallel everywhere, except in a regular neighborhood of V. If the Voronoi cell V is disjoint from T_μ, then it is a ball, and is attached to P_1 along a subsurface consisting of a union of faces of V. Therefore the number of boundary components of the attaching surface is at most J, where J is the constant from Proposition 2.3, so by Proposition 2.16, P_2 is obtained from P_1 by attaching at most $2J$ handles.

If the Voronoi cell V intersects T_μ, then P_2 is obtained from P_1 by adding regions of $V \setminus T_\mu$, which we shall refer to as the *complementary regions*, together with regions of $T_\mu \cap (P_2 \setminus P_1)$. The complementary regions may not be topological balls, but Kobayashi and Rieck [2011] show that they are handlebodies of bounded genus.

Proposition 2.17 [Kobayashi and Rieck 2011]. *Let μ be a Margulis constant for $\mathbb{H}^3$ and M be a finite volume hyperbolic* 3-*manifold; let $0 < \epsilon < \mu$, and let $\mathcal{V}$ be a regular Voronoi decomposition of M arising from a maximal collection of ϵ-separated points. Then there is a number G, depending only on μ and ϵ, such that for any Voronoi cell V_i, there are at most G connected components of $V_i \cap X_\mu$, each of which is a handlebody of genus at most G, attached to T_μ by a surface with at most G boundary components.*

We state a simplified version of their result which suffices for our purposes. Their stated result involves extra parameters d and R, but if d is chosen close to 0, then R is close to μ, and we obtain the result above. Their proof involves showing that in the universal cover, for any point p in $T_\mu \cap V_i$, projection to ∂T_μ along geodesic rays based at p gives a topological product structure to $V_i \cap X_\mu$ as $(V_i \cap \partial T_\mu) \times I$. An examination of their proof shows that we may choose $G = 3J$, where J is the constant from Proposition 2.3. Then Proposition 2.16 implies that adding the complementary regions of a Voronoi cell which intersects ∂T_μ requires at most $6G^2 = 54J^2$ handles.

However, if the Voronoi cell intersects a solid torus component T of T_μ, then the surfaces S_1^+ and S_2^+ need not be parallel inside T, and so we now bound the number of handles needed to add the region corresponding to $(P_2 \setminus P_1) \cap T$. If U_2 is equal to all of ∂T, then the additional region is a solid torus attached along $\partial T \setminus U_1$, so adding this region requires at most $4 + 2|\partial U_1|$ handles, by Proposition 2.15. If U_2 is not equal to all of ∂T, then this region is homeomorphic to $(U_2 \times [0,1]) \setminus \left(U_1 \times \left[0, \frac{1}{2}\right]\right)$, and so is homeomorphic to $U_2 \times I$, which is a handlebody of genus at most $|\partial U_2|$. The region is attached along U_2, so adding this region requires at most $4|\partial U_2| + 2|\partial U_1|$ handles, and so in either case, at most $4J\|S_2 \cap \mathcal{W}\| + 2J\|S_1 \cap \mathcal{W}\|$ are required.

Therefore P_2 may be constructed from P_1 by adding at most

$$54J^2 + 4J\|S_2 \cap \mathcal{W}\| + 2J\|S_1 \cap \mathcal{W}\| \leqslant 60J^2 \max\{\|S_i \cap \mathcal{W}\|\}$$

handles, as required. □

The manifold M may be constructed by adding the Voronoi cells in the order arising from the cell splitters t_i in $\mathbb{R}$. Choose a finite collection of generic points $\{u_i\}$, so that each pair of adjacent cell splitters is separated by one of the u_i, and let $\{S_i^+\}$ be the corresponding collection of T-capped surfaces. The linear width is at most the largest genus of any surface in the collection $\{S_i^+\}$, plus the maximum number of handles added by attaching a single Voronoi cell. Therefore the bounds from Propositions 2.14 and 2.15 imply

$$\text{linear width}(M) \leqslant (5J+1)K(\text{Morse area}(M)) + 60J^2K(\text{Morse area}(M)).$$

As J is at least 1, this gives

$$\text{linear width}(M) \leqslant (66J^2K)\,\text{Morse width}(M).$$

The constants J and K only depend on the choice of MV-constants, which may be chosen independently of the hyperbolic 3-manifold M, and so this completes the proof of the left-hand bound of Theorem 1.2.

3. Scharlemann–Thompson width bounds Morse area

We will show linear bounds for Morse area in terms of Scharlemann–Thompson width, assuming the Pitts and Rubinstein result, Theorem 1.1; i.e., we will show the right-hand bound of Theorem 1.2. This result is due to Gabai and Colding [2015, Appendix A], using recent work of Colding and Minicozzi [2015], but we give a brief description for the convenience of the reader, as they do not state this result explicitly.

We will use properties of a refinement of linear width, known as *thin position*, which we now describe. Let $\{H_i\}$ be the collection of upper boundaries of compression bodies in the linear splitting, and let $c(H_i)$ be the complexity of the surface H_i, i.e., the sum of the genera of its connected components. We say that the complexity of the linear splitting is the collection of integers $\{c(H_i)\}$, arranged in decreasing order. A linear splitting which gives the minimum complexity of all possible linear splittings in the lexicographic ordering on sets of integers is called a *thin position* linear splitting. Scharlemann and Thompson [1994] showed that thin position linear splittings have the following property.

Theorem 3.1 [Scharlemann and Thompson 1994]. *Let H be a linear splitting that is in thin position. Then every even surface is incompressible in M and the odd surfaces form strongly irreducible Heegaard surfaces for the components of M cut along the even surfaces.*

If follows from [Freedman et al. 1983] that the incompressible surfaces may be chosen to be disjoint least area minimal surfaces, and in fact the odd surfaces may also be chosen to be disjoint minimal surfaces, possibly up to compression; see, for example, [Lackenby 2006] or [Renard 2014] for a detailed statement of the result in this case. In a hyperbolic manifold, the intrinsic curvature of a minimal surface is at most -1, so the Gauss–Bonnet formula gives an upper bound for the area of the minimal surface. Therefore the area of a minimal surface of genus g is at most $-2\pi\chi(S) \leqslant 4\pi g$.

We say that a hyperbolic 3-manifold M has *least area boundary* if its boundary components are (possibly empty) least area minimal surfaces, and we say that a Heegaard splitting H for M is *minimal* if it is isotopic to an unstable minimal surface. The right-hand bound of Theorem 1.2 is a consequence of the following result of Colding and Gabai [2015], which constructs bounded area foliations for a pair of compression bodies with least area lower boundaries, sharing a common minimal Heegaard splitting surface.

Theorem 3.2 [Colding and Gabai 2015]. *Let M be a hyperbolic manifold, with* (*possibly empty*) *least area boundary, with a minimal Heegaard splitting H of genus g. Then, assuming Theorem 1.1, the manifold M has a* (*possibly singular*) *foliation by compact leaves, containing the boundary surfaces as leaves, such that each leaf has area at most $4\pi g$.*

As they do not state this explicitly in their paper, we give a brief outline for the convenience of the reader.

Definition 3.3. A mean convex foliation on a Riemannian 3-manifold with boundary is a smooth codimension-1 foliation, possibly with singularities of standard type, such that each leaf is mean convex.

In a 3-manifold, a foliation with singularities of "standard type" means that almost all leaves are completely smooth (i.e., without any singularities). In particular, any connected subset of the singular set is completely contained in a leaf. Furthermore, the entire singular set is contained in finitely many (compact) embedded Lipschitz curves with cylinder singularities together with a countable set of spherical singularities.

The following result is shown by Colding and Gabai [2015].

Theorem 3.4 [Colding and Gabai 2015, Appendix A]. *Let Σ be an unstable minimal surface in a hyperbolic manifold M. Then there is a regular neighborhood of Σ with a smooth mean convex product foliation Σ_t, $t \in [-\epsilon, \epsilon]$, with nonminimal boundary leaves $\Sigma_{-\epsilon}$ and Σ_ϵ.*

In particular, each leaf in the foliation has area at most $4\pi g$. As the boundary leaves $\Sigma_{-\epsilon}$ and Σ_ϵ are nonminimal mean convex surfaces, we may apply the mean curvature flow results of Colding and Minicozzi [2015], which show that the mean curvature flow gives rise to a mean convex foliation with standard singularities. As the mean curvature flow gives a foliation by surfaces of decreasing area, the only possible singularities which may arise are disc compressions, 2-spheres collapsing to a point or tori collapsing to circles. In particular, each nonsingular leaf bounds a compression body in the interior of the compression body it is contained in.

If all leaves eventually collapse, then the compression body has empty lower boundary, i.e., it is a handlebody, and this gives a mean convex foliation, and hence area-decreasing foliation, of the handlebody. Otherwise, the mean curvature flow limits to a stable minimal surface Γ whose components bound compression bodies together with the lower boundary of the original compression body.

If the stable minimal surface Γ is not equal to the stable boundary of the compression body, then it bounds a subcompression body with stable boundary, whose standard Heegaard splitting is strongly irreducible, so we may apply the argument again. Anderson [1985] and White [1987] showed that there are only finitely many minimal surfaces of bounded genus in a compact Riemannian manifold, and so this process may occur only finitely many times, resulting in a foliation of the entire compression body. This completes the proof of Theorem 3.2.

Finally we deduce the right-hand bound of Theorem 1.2 from Theorem 3.2.

Proof of right-hand bound of Theorem 1.2. By Theorem 1.1, the irreducible Heegaard surface for a hyperbolic 3-manifold M with stable boundary is either isotopic to an unstable minimal surface Σ, to which we may apply Theorem 3.2 directly, or isotopic to a regular neighborhood of a one-sided stable minimal surface union a small tube parallel to one of the normal fibers. In the latter case, the Heegaard surface bounds a handlebody on at least one side, and cutting along the stable one-sided surface leaves a compression body homeomorphic to the Heegaard surface cut along the disc corresponding to the tube, where all boundary components are stable minimal surfaces. As the standard Heegaard splitting of a compression body is strongly irreducible, we may now apply Theorem 3.2 in this case as well. □

Acknowledgements

The authors would like to thank Dick Canary, David Futer, David Gabai, Joel Hass, Daniel Ketover, Sadayoshi Kojima, Yair Minksy, Yo'av Rieck and Dylan Thurston for helpful conversations, and the Tokyo Institute of Technology for its hospitality and support. Maher was supported by the Simons Foundation and PSC-CUNY.

References

[Anderson 1985] M. T. Anderson, "Curvature estimates for minimal surfaces in 3-manifolds", *Ann. Sci. École Norm. Sup.* (4) **18**:1 (1985), 89–105. MR 87e:53098 Zbl 0578.49027

[Bachman et al. 2004] D. Bachman, D. Cooper, and M. E. White, "Large embedded balls and Heegaard genus in negative curvature", *Algebr. Geom. Topol.* **4** (2004), 31–47. MR 2004m:57029 Zbl 1056.57014

[Colding and De Lellis 2003] T. H. Colding and C. De Lellis, "The min-max construction of minimal surfaces", pp. 75–107 in *Lectures on geometry and topology held in honor of Calabi, Lawson, Siu, and Uhlenbeck* (Boston, 2002), edited by S.-T. Yau, Surveys in Differential Geometry **8**, International Press, Somerville, MA, 2003. MR 2005a:53008 Zbl 1051.53052

[Colding and Gabai 2015] T. H. Colding and D. Gabai, "Effective finiteness of irreducible Heegaard splittings of non Haken 3-manifolds", preprint, 2015. arXiv 1411.2509

[Colding and Minicozzi 2015] T. H. Colding and W. P. Minicozzi, II, "The singular set of mean curvature flow with generic singularities", preprint, 2015. arXiv 1405.5187

[De Lellis and Pellandini 2010] C. De Lellis and F. Pellandini, "Genus bounds for minimal surfaces arising from min-max constructions", *J. Reine Angew. Math.* **644** (2010), 47–99. MR 2011g:53012 Zbl 1201.53009

[Freedman et al. 1983] M. Freedman, J. Hass, and P. Scott, "Least area incompressible surfaces in 3-manifolds", *Invent. Math.* **71**:3 (1983), 609–642. MR 85e:57012 Zbl 0482.53045

[Ketover 2013] D. Ketover, "Degeneration of min-max sequences in 3-manifolds", preprint, 2013. arXiv 1312.2666

[Kobayashi and Rieck 2011] T. Kobayashi and Y. Rieck, "A linear bound on the tetrahedral number of manifolds of bounded volume (after Jørgensen and Thurston)", pp. 27–42 in *Topology and geometry in dimension three*, edited by W. Li et al., Contemp. Math. **560**, Amer. Math. Soc., Providence, RI, 2011. MR 2866921

[Lackenby 2006] M. Lackenby, "Heegaard splittings, the virtually Haken conjecture and property (τ)", *Invent. Math.* **164**:2 (2006), 317–359. MR 2007c:57030 Zbl 1110.57015

[Pitts and Rubinstein 1986] J. T. Pitts and J. H. Rubinstein, "Existence of minimal surfaces of bounded topological type in three-manifolds", pp. 163–176 in *Miniconference on geometry and partial differential equations* (Canberra, 1985), edited by L. Simon and N. S. Trudinger, Proc. Centre Math. Anal. Austral. Nat. Univ. **10**, Austral. Nat. Univ., Canberra, 1986. MR 87j:49074 Zbl 0602.49028

[Renard 2014] C. Renard, "Detecting surface bundles in finite covers of hyperbolic closed 3-manifolds", *Trans. Amer. Math. Soc.* **366**:2 (2014), 979–1027. MR 3130323 Zbl 1286.57008

[Rubinstein 2005] J. H. Rubinstein, "Minimal surfaces in geometric 3-manifolds", pp. 725–746 in *Global theory of minimal surfaces*, edited by D. Hoffman, Clay Math. Proc. **2**, Amer. Math. Soc., Providence, RI, 2005. MR 2006g:57038 Zbl 1119.53042

[Scharlemann and Thompson 1994] M. Scharlemann and A. Thompson, "Thin position for 3-manifolds", pp. 231–238 in *Geometric topology* (Haifa, 1992), edited by C. Gordon et al., Contemp. Math. **164**, Amer. Math. Soc., Providence, RI, 1994. MR 95e:57032 Zbl 0818.57013

[White 1987] B. White, "Curvature estimates and compactness theorems in 3-manifolds for surfaces that are stationary for parametric elliptic functionals", *Invent. Math.* **88**:2 (1987), 243–256. MR 88g:58037 Zbl 0615.53044

Received March 30, 2015. Revised July 23, 2015.

DIANE HOFFOSS
DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF SAN DIEGO
5998 ALCALA PARK
SAN DIEGO, CA 92110-2492
UNITED STATES

dhoffoss@sandiego.edu

JOSEPH MAHER
DEPARTMENT OF MATHEMATICS
CUNY COLLEGE OF STATEN ISLAND AND CUNY GRADUATE CENTER
2800 VICTORY BOULEVARD
STATEN ISLAND, NY 10314
UNITED STATES

joseph.maher@csi.cuny.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.103

RICCI TENSOR OF REAL HYPERSURFACES

MAYUKO KON

Let M be a real hypersurface of a complex space form $M^n(c)$, $c \neq 0$, and suppose that the structure vector field ξ is an eigen vector field of the Ricci tensor S, which satisfies $S\xi = \beta\xi$ where β is a function. We show that if $(\nabla_X S)Y$ is proportional to ξ for any vector fields X and Y orthogonal to ξ, then M is a Hopf hypersurface, and if it is perpendicular to ξ, then M is a ruled real hypersurface.

1. Introduction

Takagi [1973] gave a classification of the homogeneous real hypersurface (see also [Takagi 1975a; 1975b]). As a consequence of this result, the structure vector ξ of any homogeneous real hypersurface in $\mathbb{C}P^n$ is principal. If ξ satisfies this property, then M is said to be a *Hopf hypersurface*. When the ambient manifold is a complex hyperbolic space, Lohnherr [1998] (see also [Lohnherr and Reckziegel 1999]) discovered a homogeneous ruled real hypersurface in $\mathbb{C}H^n$ that is not a Hopf hypersurface, and further examples were given (see [Berndt and Brück 2001]). The classification theorem for homogeneous real hypersurfaces in $\mathbb{C}H^n$, $n \geq 2$, was given by Berndt and Tamaru [2007].

When a real hypersurface is Hopf, fundamental formulas are simple. So many classification theorems are given under that assumption (see, for example, [Niebergall and Ryan 1997]). Kimura [1986] has given a classification of Hopf hypersurfaces of $\mathbb{C}P^n$, $n \geq 2$, with constant principal curvatures. He showed that a real hypersurface in $\mathbb{C}P^n$ with constant principal curvatures is a Hopf hypersurface if and only if it is an open part of a homogeneous real hypersurface. A classification theorem for Hopf hypersurfaces with constant principal curvatures in $\mathbb{C}H^n$, $n \geq 2$, was given by Berndt [1989].

On the other hand, the Ricci tensor of the real hypersurfaces is an interesting subject. It is well known that any real hypersurface of $M^n(c)$, $c \neq 0$, is not Einstein. If the Ricci tensor S is of the form $S(X, Y) = ag(X, Y) + b\eta(X)\eta(Y)$, then the real hypersurface is said to be pseudo-Einstein. The classification theorems for pseudo-Einstein real hypersurfaces in a complex space form $M^n(c)$ have been

MSC2010: primary 53C40; secondary 53C55, 53C25.
Keywords: real hypersurface, Ricci tensor.

completed [Cecil and Ryan 1982; Kim and Ryan 2008; Kon 1979; Montiel 1985]. Ki [1989] showed that there are no real hypersurfaces with parallel Ricci tensor, $\nabla S = 0$, in $M^n(c)$, $n \geq 3$. Several conditions that weaken the condition $\nabla S = 0$ have been studied (see [Ki et al. 1990; Suh 1990]).

We focus on the Ricci tensor S and consider a condition $S\xi = \beta\xi$, where β is a function. We note that this condition contains not only Hopf hypersurfaces, $A\xi = \alpha\xi$, but also some non-Hopf hypersurfaces. For example, ruled hypersurfaces, which are an important example of non-Hopf hypersurfaces, also satisfy $S\xi = \beta\xi$. Under this assumption, we study some Hopf hypersurfaces and ruled real hypersurfaces according to the direction of a covariant differentiation of S.

Our main result is the following theorem:

Theorem 1.1. *Let M be a connected real hypersurface of $M^n(c)$, $c \neq 0$, and suppose that the Ricci tensor S of M satisfies $S\xi = \beta\xi$ for some function β.*

(1) *If $(\nabla_X S)Y$ is proportional to the structure vector field ξ for any vector fields X and Y orthogonal to ξ, then M is a Hopf hypersurface.*

(2) *If $(\nabla_X S)Y$ is perpendicular to the structure vector field ξ for any vector fields X and Y orthogonal to the structure vector field ξ, then M is a ruled real hypersurface.*

When $n = 2$, the author gave a corresponding result in [Kon 2014].

2. Preliminaries

Let $M^n(c)$ denote the complex space form of complex dimension n (real dimension $2n$) with constant holomorphic sectional curvature $4c$. We denote by J the almost complex structure of $M^n(c)$. The Hermitian metric of $M^n(c)$ is denoted by G.

Let M be a real $(2n-1)$-dimensional hypersurface immersed in $M^n(c)$. Throughout this paper, we suppose that M is connected. We denote by g the Riemannian metric induced on M from G. We take the unit normal vector field N of M in $M^n(c)$. For any vector field X tangent to M, we define ϕ, η and ξ by

$$JX = \phi X + \eta(X)N, \quad JN = -\xi,$$

where ϕX is the tangential part of JX, ϕ is a tensor field of type (1,1), η is a 1-form, and ξ is the unit vector field on M. We call ξ the *structure vector field*. Then

$$\phi^2 X = -X + \eta(X)\xi, \quad \phi\xi = 0, \quad \eta(\phi X) = 0$$

for any vector field X tangent to M. Moreover, we have

$$g(\phi X, Y) + g(X, \phi Y) = 0, \quad \eta(X) = g(X, \xi), \quad g(\phi X, \phi Y) = g(X, Y) - \eta(X)\eta(Y).$$

Thus (ϕ, ξ, η, g) defines an almost contact metric structure on M.

We denote by $\tilde{\nabla}$ the operator of covariant differentiation in $M^n(c)$, and by ∇ the operator of covariant differentiation in M determined by the induced metric. Then the *Gauss and Weingarten formulas* are given respectively by

$$\tilde{\nabla}_X Y = \nabla_X Y + g(AX, Y)N, \quad \tilde{\nabla}_X N = -AX,$$

for any vector fields X and Y tangent to M.

For the contact metric structure on M, we have

$$\nabla_X \xi = \phi AX, \quad (\nabla_X \phi)Y = \eta(Y)AX - g(AX, Y)\xi.$$

We call A the *shape operator* of M. If the shape operator A of M satisfies $A\xi = \alpha\xi$ for some function α, then M is called a *Hopf hypersurface*. By the Codazzi equation, we have the following result (see [Maeda 1976]).

Proposition A. *Let M be a Hopf hypersurface in $M^n(c)$, $n \geq 2$. If $X \perp \xi$ and $AX = \lambda X$, then $\alpha = g(A\xi, \xi)$ is constant and*

$$(2\lambda - \alpha)A\phi X = (\lambda\alpha + 2c)\phi X.$$

We offer an important example of a non-Hopf hypersurface. Take a regular curve γ in $M^n(c)$ with tangent vector field X. At each point of γ there is a unique complex projective or hyperbolic hyperplane cutting γ so as to be orthogonal to X and JX. The union of these hyperplanes is called a *ruled real hypersurface* (see [Kimura and Maeda 1989; Lohnherr and Reckziegel 1999; Niebergall and Ryan 1997]).

We remark that the shape operator A is η-parallel if it satisfies $g((\nabla_X A)Y, Z) = 0$ for any X, Y and Z orthogonal to ξ.

We denote by R the Riemannian curvature tensor field of M. Then the *equation of Gauss* is given by

$$\begin{aligned} R(X, Y)Z \\ &= c\{g(Y, Z)X - g(X, Z)Y + g(\phi Y, Z)\phi X - g(\phi X, Z)\phi Y - 2g(\phi X, Y)\phi Z\} \\ &\quad + g(AY, Z)AX - g(AX, Z)AY, \end{aligned}$$

and the *equation of Codazzi* by

$$(\nabla_X A)Y - (\nabla_Y A)X = c\{\eta(X)\phi Y - \eta(Y)\phi X - 2g(\phi X, Y)\xi\}.$$

From the equation of Gauss, the Ricci tensor S of M is given by

$$g(SX, Y) = (2n+1)cg(X, Y) - 3c\eta(X)\eta(Y) + \operatorname{tr} A g(AX, Y) - g(AX, AY), \tag{1}$$

where $\operatorname{tr} A$ is the trace of A. Taking a covariant differentiation, we have

$$\begin{aligned} g((\nabla_X S)Y, Z) = &-3cg(Y, \phi AX)\eta(Z) - 3cg(\phi AX, Z)\eta(Y) + (X \operatorname{tr} A)g(AY, Z) \\ &+ \operatorname{tr} A g((\nabla_X A)Y, Z) - g((\nabla_X A)AY, Z) - g((\nabla_X A)Y, AZ). \end{aligned} \tag{2}$$

Now we develop some lemmas needed to prove our main theorem. Suppose $n \geq 3$.

Lemma 2.1. *Let M be a real hypersurface in a complex space form $M^n(c)$, $n \geq 3$, $c \neq 0$. If there exists an orthonormal frame $\{\xi, e_1, \ldots, e_{2n-2}\}$ on a sufficiently small neighborhood $\mathcal{N}$ of $x \in M$ such that the shape operator A can be represented as*

$$A = \begin{pmatrix} \alpha & h_1 & 0 & \cdots & 0 \\ h_1 & a_1 & & & \\ 0 & & a_2 & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & & \cdots & 0 & a_{2n-2} \end{pmatrix},$$

then we have

$$(a_j - a_k)g(\nabla_{e_i} e_j, e_k) - (a_i - a_k)g(\nabla_{e_j} e_i, e_k) = 0, \tag{3}$$

$$(a_j - a_1)g(\nabla_{e_i} e_j, e_1) - (a_i - a_1)g(\nabla_{e_j} e_i, e_1) = h_1(a_i + a_j)g(e_i, \phi e_j), \tag{4}$$

$$h_1 g(\nabla_{e_i} e_j, e_1) - h_1 g(\nabla_{e_j} e_i, e_1) = \{2c - 2a_i a_j + \alpha(a_i + a_j)\}g(\phi e_i, e_j), \tag{5}$$

$$(e_j a_i) = (a_j - a_i)g(\nabla_{e_i} e_j, e_i), \tag{6}$$

$$(e_1 a_i) = (a_1 - a_i)g(\nabla_{e_i} e_1, e_i), \tag{7}$$

$$(a_1 - a_j)g(\nabla_{e_i} e_1, e_j) + (a_j - a_i)g(\nabla_{e_1} e_i, e_j) = a_i h_1 g(e_i, \phi e_j), \tag{8}$$

$$(e_i h_1) = \{2c - 2a_1 a_i + \alpha(a_i + a_1)\}g(e_i, \phi e_1) - h_1 g(\nabla_{e_1} e_i, e_1), \tag{9}$$

$$(e_i a_1) = h_1(2a_i + a_1)g(e_i, \phi e_1) + (a_i - a_1)g(\nabla_{e_1} e_i, e_1), \tag{10}$$

$$(\xi a_i) = h_1 g(\nabla_{e_i} e_1, e_i), \tag{11}$$

$$h_1 g(\nabla_{e_i} e_1, e_j) + (a_j - a_i)g(\nabla_\xi e_i, e_j) = (c + a_i\alpha - a_i a_j)g(e_i, \phi e_j), \tag{12}$$

$$(e_i h_1) = (c + a_i\alpha - a_1 a_i + h_1^2)g(e_i, \phi e_1) + (a_i - a_1)g(\nabla_\xi e_i, e_1), \tag{13}$$

$$(e_i \alpha) = h_1(\alpha - 3a_i)g(e_i, \phi e_1) - h_1 g(\nabla_\xi e_i, e_1), \tag{14}$$

$$(e_1 h_1) = (\xi a_1), \tag{15}$$

$$(e_1 \alpha) = (\xi h_1), \tag{16}$$

$$(a_1 - a_i)g(\nabla_\xi e_1, e_i) - h_1 g(\nabla_{e_1} e_1, e_i) = (c + a_1\alpha - a_1 a_i - h_1^2)g(e_i, \phi e_1), \tag{17}$$

for any $i, j \geq 2$, $i \neq j$.

Proof. By the equation of Codazzi, we have

$$g((\nabla_{e_i} A)e_1 - (\nabla_{e_1} A)e_i, e_j) = 0,$$

where $i, j = 2, \ldots, 2n-2$. On the other hand, we have

$$\begin{aligned} &g((\nabla_{e_i} A)e_1 - (\nabla_{e_1} A)e_i, e_j) \\ &\qquad = g(\nabla_{e_i}(Ae_1) - A\nabla_{e_i} e_1 - \nabla_{e_1}(Ae_i) + A\nabla_{e_1} e_i, e_j) \\ &\qquad = (a_1 - a_j)g(\nabla_{e_i} e_1, e_j) + (a_j - a_i)g(\nabla_{e_1} e_i, e_j) + a_i h_1 g(\phi e_i, e_j). \end{aligned}$$

Thus we obtain (8). We obtain the other results through similar computations. □

We remark that these equations hold in the case that M is a Hopf hypersurface, i.e., $h_1 = 0$. When $n = 2$, we showed the corresponding result in [Kon 2014].

We define the subspace $L_x \subset T_x(M)$ as the smallest subspace that contains ξ and is invariant under the shape operator A. Then M is Hopf if and only if L_x is one-dimensional at each point x.

Lemma 2.2. *Let M be a real hypersurface of $M^n(c)$. If the Ricci tensor S of M satisfies $S\xi = \beta\xi$ for some function β, then* $\dim L_x \leq 2$ *at each point x of $M^n(c)$.*

Proof. By (1), we have

$$0 = g(S\xi, Y) = -g(A^2\xi, Y)$$

for any Y orthogonal to ξ and $A\xi$. So $A^2\xi$ is spanned by ξ and $A\xi$. Thus we see that $\dim L_x \leq 2$. □

Suppose that M is not a Hopf hypersurface and that $S\xi = \beta\xi$. By Lemma 2.2, we can take an orthonormal frame $\{\xi, e_1, \dots, e_{2n-2}\}$, locally, such that A is of the form

$$A = \begin{pmatrix} \alpha & h_1 & & & 0 \\ h_1 & a_1 & & & \\ & & a_2 & & \\ & & & \ddots & \\ 0 & & & & a_{2n-2} \end{pmatrix},$$

where $h_1 = g(Ae_1, \xi)$, $a_i = g(Ae_i, e_i)$ for $i = 1, \dots, 2n-2$, $g(Ae_i, e_j) = 0$ for $i \neq j$ and $\alpha = g(A\xi, \xi)$. By (1), we obtain

$$\begin{aligned} S\xi &= (2n-2)c\xi + (\operatorname{tr} A)(h_1e_1 + \alpha\xi) - A(h_1e_1 + \alpha\xi) \\ &= (\operatorname{tr} A - \alpha - a_1)h_1e_1 + \{(2n-2)c + (\operatorname{tr} A)\alpha - h_1^2 - \alpha^2\}\xi = \beta\xi. \end{aligned}$$

So we see that

$$\operatorname{tr} A = \alpha + a_1, \quad a_2 + \cdots + a_{2n-2} = 0.$$

Moreover, (1) implies that the Ricci tensor S can be represented as

$$S = \begin{pmatrix} \beta & & & 0 \\ & \lambda_1 & & \\ & & \ddots & \\ 0 & & & \lambda_{2n-2} \end{pmatrix},$$

where β and λ_i satisfy

$$\begin{aligned} \beta &= (2n-2)c + (\alpha a_1 - h_1^2), \quad \lambda_1 = (2n+1)c + (\alpha a_1 - h_1^2), \\ \lambda_j &= (2n+1)c + \operatorname{tr} A \cdot a_j - a_j^2, \quad j = 2, \dots, 2n-2. \end{aligned}$$

3. Real hypersurfaces with η-parallel Ricci tensor

In this section, we consider the additional condition that the Ricci operator S is η-parallel, that is,

$$g((\nabla_X S)Y, Z) = 0$$

for any vector fields X, Y and Z orthogonal to ξ. This is equivalent to the condition that $(\nabla_X S)Y$ is proportional to ξ [Suh 1990].

Theorem 3.1. *Let M be a real hypersurface of $M^n(c)$, $c \neq 0$, with η-parallel Ricci tensor. If the Ricci tensor S of M satisfies $S\xi = \beta\xi$ for some function β, then M is a Hopf hypersurface.*

Before proving Theorem 3.1, we need the following lemma.

Lemma 3.2. *Let M be a real hypersurface of $M^n(c)$, $c \neq 0$, with η-parallel Ricci tensor. If the Ricci tensor S of M satisfies $S\xi = \beta\xi$ for some function β, then we have*

$$\begin{aligned} g((R(W, X)S)Y, Z) = &-g(S\phi AX, Z)g(\phi AW, Y) - g(S\phi AX, Y)g(\phi AW, Z) \\ &+ g(S\phi AW, Z)g(\phi AX, Y) + g(S\phi AW, Y)g(\phi AX, Z) \\ &- g((\nabla_\xi S)Y, Z)g((\phi A + A\phi)X, W) \end{aligned}$$

for any X, Y, Z and W orthogonal to ξ.

Proof. Since S is η-parallel, we have

$$\begin{aligned} &g((R(W, X)S)Y, Z) \\ &\quad = g(R(W, X)SY, Z) - g(R(W, X)Y, SZ) \\ &\quad = g(\nabla_W\nabla_X SY - \nabla_X\nabla_W SY - \nabla_{[W,X]}SY, Z) \\ &\qquad - g(\nabla_W\nabla_X Y - \nabla_X\nabla_W Y - \nabla_{[W,X]}Y, SZ) \\ &\quad = -g((\nabla_X S)Y, \nabla_W Z) + g(\nabla_W(S\nabla_X Y), Z) + g((\nabla_W S)Y, \nabla_X Z) \\ &\qquad - g(\nabla_X(S\nabla_W Y), Z) - g((\nabla_{[W,X]}S)Y, Z) - g(\nabla_W\nabla_Y, SZ) \\ &\qquad + g(\nabla_X\nabla_W Y, SZ) \\ &\quad = -g((\nabla_X S)Y, \xi)g(\xi, \nabla_W Z) + g((\nabla_W S)\nabla_X Y, Z) \\ &\qquad + g((\nabla_W S)Y, \xi)g(\xi, \nabla_X Z) - g((\nabla_X S)\nabla_W Y, Z) \\ &\qquad - g((\nabla_\xi S)Y, Z)g(\xi, [W, X]) \\ &\quad = -g(S\phi AX, Y)g(\phi AW, Z) + g(S\phi AW, Z)g(\phi AX, Y) \\ &\qquad + g(S\phi AW, Y)g(\phi AX, Z) - g(S\phi AX, Z)g(\phi AW, Y) \\ &\qquad - g((\nabla_\xi S)Y, Z)g((\phi A + A\phi)X, W). \qquad \square \end{aligned}$$

From Lemma 3.2 we obtain the following:

Lemma 3.3. *Let M be a real hypersurface of $M^n(c)$, $c \neq 0$, with η-parallel Ricci tensor. Suppose that the Ricci tensor S of M satisfies $S\xi = \beta\xi$ for some function β. If $SY = \lambda Y$ and if Y is orthogonal to ξ, then we have*

$$g((\nabla_\xi S)Y, Y)g((\phi A + A\phi)X, W) = 0$$

for any X, Y and W orthogonal to ξ.

Proof of Theorem 3.1.

In the following, we suppose that M is not a Hopf hypersurface. We work in an open set where $h_1 \neq 0$.

Case (I): First we consider the case $g((\nabla_\xi S)Y, Y) = 0$.

Lemma 3.4. $\beta, \lambda_1, \dots, \lambda_{2n-2}$ *are constant.*

Proof. Since the Ricci tensor S is η-parallel and since $g((\nabla_\xi S)Y, Y) = 0$, we have

$$0 = g((\nabla_Z S)Y, Y) = g(\nabla_Z SY, Y) - g(S\nabla_Z Y, Y) = Z\lambda$$

for any tangent vector field Z. So we see that $\lambda_1, \dots, \lambda_{2n-2}$ are constant. On the other hand, since $\beta = \lambda_1 - 3c$, we see that β is also constant. □

Lemma 3.5. *If $\lambda_i \neq \lambda_j$, $i, j = 1, \dots, 2n-2$, then we have $g(\nabla_X e_i, e_j) = 0$ for any X orthogonal to ξ.*

Proof. Since we have $Se_i = \lambda_i e_i$ and $Se_j = \lambda_j e_j$ and since S is η-parallel, we obtain

$$0 = g((\nabla_X S)e_i, e_j) = (\lambda_i - \lambda_j)g(\nabla_X e_i, e_j). \qquad \square$$

If $\lambda_1 = \dots = \lambda_{2n-2} = \lambda$, then M is pseudo-Einstein, i.e., $SX = \lambda X + (\beta - \lambda)\eta(X)\xi$, and so it is a Hopf hypersurface (see [Kon 1979]).

Suppose that M is non-Hopf and that there exist λ_t and λ_j, $t, j \geq 2$, satisfying $\lambda_1 \neq \lambda_t$ and $\lambda_t \neq \lambda_j$. By Lemma 3.5,

$$\begin{aligned} g(\nabla_j \nabla_t e_t, e_j) &= -g(\nabla_{e_t} e_t, \nabla_{e_j} e_j) \\ &= -g(\nabla_{e_t} e_t, \xi)(\xi, \nabla_{e_j} e_j) - \sum_k g(\nabla_{e_t} e_t, e_k)g(e_k, \nabla_{e_j} e_j) \\ &= -g(e_t, \phi A e_t)g(\phi A e_j, e_j) = 0, \end{aligned}$$

$$\begin{aligned} g(\nabla_t \nabla_j e_t, e_j) &= -g(\nabla_{e_j} e_t, \nabla_{e_t} e_j) = -g(\nabla_{e_j} e_t, \xi)g(\xi, \nabla_{e_t} e_g) \\ &= -g(e_t, \phi A e_j)g(\phi A e_t, e_j) = -a_j a_t g(e_t, \phi e_j)g(\phi e_t, e_j). \end{aligned}$$

On the other hand, from (8),

$$(a_1 - a_t)g(\nabla_{e_j} e_1, e_t) + (a_t - a_j)g(\nabla_{e_1} e_j, e_t) + a_j h_1 g(\phi e_j, e_t) = 0.$$

From Lemma 3.5, we have $g(\nabla_{e_j} e_1, e_t) = 0$, $g(\nabla_{e_1} e_j, e_t) = 0$. Since $h_1 \neq 0$,

$$a_j g(\phi e_j, e_t) = 0,$$

from which we obtain

$$g(\nabla_{e_t}\nabla_{e_j}e_t, e_j) = 0.$$

Moreover, we have

$$\begin{aligned} g(\nabla_{[e_j,e_t]}e_t, e_j) &= g(\nabla_\xi e_t, e_j)g(\xi, [e_j, e_t]) \\ &= g(\nabla_\xi e_t, e_j)(-g(\phi Ae_j, e_t) + g(\phi Ae_t, e_j)) \\ &= g(\nabla_\xi e_t, e_j)(a_t - a_j)g(\phi e_t, e_j) \\ &= g(\nabla_\xi e_t, e_j)a_t g(\phi e_t, e_j). \end{aligned}$$

Using (12), we see that

$$(c + a_j\alpha - a_j a_t)g(\phi e_j, e_t) + h_1 g(\nabla_{e_j}e_1, e_t) + (a_t - a_j)g(\nabla_\xi e_j, e_t) = 0.$$

From these equations, we obtain

$$cg(\phi e_j, e_t)^2 + a_t g(\phi e_j, e_t)g(\nabla_\xi e_j, e_t) = 0.$$

Hence we have

$$g(\nabla_{[e_j,e_t]}e_t, e_j) = -cg(\phi e_j, e_t)^2.$$

Therefore,

$$g(R(e_j, e_t)e_t, e_j) = cg(\phi e_j, e_t)^2.$$

On the other hand, the equation of Gauss implies

$$g(R(e_j, e_t)e_t, e_j) = c + 3cg(\phi e_j, e_t)^2 + a_t a_j.$$

From these equations, we have

$$c(1 + 2g(\phi e_j, e_t)^2) + a_t a_j = 0.$$

Sine $c \neq 0$, we see that $a_t \neq 0$ and $a_j \neq 0$. Thus $g(\phi e_j, e_t) = 0$ and $c + a_t a_j = 0$. So we can represent A as

$$A = \begin{pmatrix} \alpha & h_1 & & & & & & \\ h_1 & a_1 & & & & & & \\ & & a & & & & & \\ & & & \ddots & & & & \\ & & & & a & & & \\ & & & & & b & & \\ & & & & & & \ddots & \\ & & & & & & & b \end{pmatrix}$$

by setting $a = a_j,\ b = a_t$ and taking a suitable permutation of $\{e_2, \ldots, e_{2n-2}\}$.

Suppose there exist j and t such that $g(\phi e_j, e_1) \neq 0$ and $g(\phi e_t, e_1) \neq 0$. Then ϕe_j and ϕe_t satisfy

$$\phi e_j = \sum_k g(\phi e_j, e_k)e_k + g(\phi e_j, e_1)e_1, \quad Ae_k = ae_k,$$
$$\phi e_t = \sum_l g(\phi e_t, e_l)e_l + g(\phi e_t, e_1)e_1, \quad Ae_l = be_l.$$

So we have

$$0 = g(\phi e_j, \phi e_t) = g(\phi e_j, e_1)g(\phi e_t, e_1),$$

from which we see that $g(\phi e_j, e_1) = 0$ or $g(\phi e_t, e_1) = 0$, and hence $A\phi e_1 = a\phi e_1$ or $A\phi e_1 = b\phi e_1$.

When $A\phi e_1 = a\phi e_1$, we have $A\phi e_t = b\phi e_t$. By (4),

$$(b - a_1)g(\nabla_{e_t}\phi e_t, e_1) - (b - a_1)g(\nabla_{\phi e_t} e_t, e_1) + 2h_1 bg(\phi e_t, \phi e_t) = 0.$$

Thus we obtain $b = 0$, which contradicts $c + ab = 0$ and $c \neq 0$. By a similar computation, the case $A\phi e_1 = b\phi e_1$ does not occur.

Next we consider the case $\lambda_2 = \cdots = \lambda_{2n-2} \neq \lambda_1$. We set $\lambda = \lambda_j$, $j = 2, \dots, 2n-2$. From Lemma 3.5, we have $g(\nabla_X e_1, e_i) = 0$, $i \geq 2$, for any X orthogonal to ξ.

By (4) and (5),

$$h_1(a_i + a_j)g(\phi e_i, e_j) = 0, \quad (2c - 2a_i a_j + \alpha(a_i + a_j))g(\phi e_i, e_j) = 0.$$

Since a_j satisfies

$$\lambda = (2n+1)c + \operatorname{tr} A \cdot a_j - a_j^2,$$

we can represent A as

$$A = \begin{pmatrix} \alpha & h_1 & & & & & & \\ h_1 & a_1 & & & & & & \\ & & a & & & & & \\ & & & \ddots & & & & \\ & & & & a & & & \\ & & & & & b & & \\ & & & & & & \ddots & \\ & & & & & & & b \end{pmatrix}$$

by taking a suitable permutation of $\{e_2, \dots, e_{2n-2}\}$.

There exist i and j satisfying $g(\phi e_i, e_j) \neq 0$. Therefore, using $h_1 \neq 0$,

$$a_i + a_j = 0, \quad 2c - 2a_i a_j + \alpha(a_i + a_j) = 0.$$

We notice that $\operatorname{tr} A = a_1 + \alpha$ and $\sum_{j=2}^{2n-2} a_j = ka + lb = 0$, where k and l are the multiplicities of a and b, respectively.

When $a_i = a_j = a$, then we have $a_i + a_j = 2a = 0$. Combining this with the above equations, we obtain $b = 0$ and $c = 0$. This is a contradiction. Similarly, the case $a_i = a_j = b$ does not occur.

Next, when $a_i = a,\ a_j = b$ and $a = b$, we have $a = b = 0$ and $c = 0$. This is a contradiction.

Finally we consider the case $a_i = a,\ a_j = b$ and $a \neq b$. Then we have $a = -b \neq 0$. Since $ka + lb = 0$, we obtain $k = l$. This contradicts the fact that M is an odd-dimensional real hypersurface.

Case (II): Next we consider the case

$$g((\phi A + A\phi)X, W) = 0 \tag{18}$$

for any X and W orthogonal to ξ.

Since $\{\xi, \phi e_1, \ldots, \phi e_{2n-2}\}$ is an orthonormal basis of the tangent space, we have

$$\begin{aligned} \mathrm{tr}A &= g(A\xi, \xi) + \sum_{i=1}^{2n-2} g(A\phi e_i, \phi e_i) \\ &= \alpha - \sum_{i=1}^{2n-2} g(\phi A e_i, \phi e_i) = \alpha - \sum_{i=1}^{2n-2} g(Ae_i, e_i). \end{aligned}$$

Since $\mathrm{tr}\, A = \alpha + \sum_{i=1}^{2n-2} g(Ae_i, e_i)$, we obtain $\sum_{i=1}^{2n-2} g(Ae_i, e_i) = 0$ and $\mathrm{tr}\, A = \alpha$. On the other hand, from $\mathrm{tr}A = a_1 + \alpha$, we have $a_1 = 0$. Substituting $X = e_1$ in (18), we see that $g(A\phi e_1, W) = 0$ for any W orthogonal to ξ. Since

$$g(A\phi e_1, \xi) = g(\phi e_1, A\xi) = 0,$$

we have $A\phi e_1 = 0$. Without loss of generality, we can set $\phi e_1 = e_2$. From (13) and (17), we obtain

$$(e_2 h_1) = c + h_1^2, \tag{19}$$

$$(c - h_1^2) + h_1 g(\nabla_{e_1} e_2, e_1) = 0. \tag{20}$$

On the other hand, since S is η-parallel, putting $X = Y = e_1$ and $Z = e_2$ into (2), we have

$$0 = \mathrm{tr}A g((\nabla_{e_1} A)e_1, e_2) - g((\nabla_{e_1} A)Ae_1, e_2) = h_1^2 g(e_1, \nabla_{e_1} e_2).$$

Since $h_1 \neq 0$, we have $g(\nabla_{e_1} e_2, e_1) = 0$. Combining this with (20), we see that $h_1^2 = c$. This contradicts (19), finishing the proof. □

We remark that Suh [1990] and Maeda [2013] classified Hopf hypersurfaces of nonflat complex space forms with η-parallel Ricci tensor.

4. Ruled real hypersurfaces

In the previous sections, under the condition that the Ricci tensor S of M satisfies $S\xi = \beta\xi$, we gave sufficient conditions for M to be a Hopf hypersurface with respect to the covariant derivative of the Ricci tensor of S. The purpose of this section is to give a condition on the Ricci tensor for M to be a ruled real hypersurface.

Theorem 4.1. *Let M be a real hypersurface of $M^n(c)$, $c \neq 0$. If the Ricci tensor S of M satisfies $S\xi = \beta\xi$ for some function β and if $g((\nabla_X S)Y, \xi) = 0$ for any vector fields X and Y orthogonal to ξ, then M is a ruled real hypersurface.*

Proof. To prove Theorem 4.1, we need the following proposition:

Proposition 4.2. *Let M be a real hypersurface of $M^n(c)$, $c \neq 0$. If the Ricci tensor S of M satisfies $S\xi = \beta\xi$ for some function β and if $g((\nabla_X S)Y, \xi) = 0$ for any vector fields X and Y orthogonal to ξ, then M is not Hopf.*

Proof. Suppose that M is a Hopf hypersurface. Then we have $A\xi = \alpha\xi$, and hence $S\xi = \beta\xi$. We note that α is constant. Therefore, we have

$$\begin{aligned} g((\nabla_X S)Y, \xi) &= g((\nabla_X S)\xi, Y) \\ &= g(\nabla_X S\xi, Y) - g(S\phi AX, Y) \\ &= \beta g(\phi AX, Y) - g(\phi AX, SY) \end{aligned}$$

for any X and Y orthogonal to ξ. We take an orthonormal basis $\{\xi, e_1, \ldots, e_{2n-2}\}$ that satisfies $e_{2i} = \phi e_{2i-1}$, $i = 1, \ldots, n-1$, and set $Ae_t = a_t e_t$, $t = 1, \ldots, 2n-2$. Then we have $A\phi e_t = \bar{a}_t \phi e_t$ since M is Hopf. Then the Ricci operator S satisfies $S\xi = \beta\xi$ and $Se_t = \lambda_t e_t$, $t = 1, \ldots, 2n-2$, where

$$\beta = (2n-2)c + \operatorname{tr} A \cdot \alpha - \alpha^2, \quad \lambda_t = (2n+1)c + \operatorname{tr} A \cdot a_t - a_t^2.$$

Thus we obtain

$$0 = (\beta - \lambda_t) g(\phi AX, e_t) = -(\beta - \lambda_t) g(X, A\phi e_t)$$

for any X orthogonal to ξ. Since $A\xi = \alpha\xi$, we have $g(A\phi e_t, \xi) = 0$. From these equations, we have:

Lemma 4.3. *If $\beta \neq \lambda_t$, then $A\phi e_t = 0$, that is, $\bar{a}_t = 0$.*

We suppose $\beta \neq \lambda_t$. Then, from (1), we have

$$\overline{\lambda_t} = g(S\phi e_t, \phi e_t) = (2n+1)c.$$

Using Proposition A and $c \neq 0$, we have $\alpha \neq 0$ and

$$a_t = -\frac{2c}{\alpha}.$$

If $\beta \neq \lambda_t$ and $\beta \neq \overline{\lambda_t} = g(S\phi e_t, \phi e_t)$, then we have $a_t = \bar{a_t} = 0$. This is a contradiction. Thus we obtain:

Lemma 4.4. *If $\beta \neq \lambda_t$, then $\beta = \overline{\lambda_t} = (2n+1)c$.*

Since M is not Einstein, there exists a t such that $\beta \neq \lambda_t$. So we see that λ_t satisfies $\beta = \lambda_t = \overline{\lambda_t}$ or $\beta = \overline{\lambda_t} \neq \lambda_t$.

When $\beta = \lambda_t = \bar{\lambda}_t$, since $\beta = (2n+1)c$, we have

$$0 = a_t(\operatorname{tr} A - a_t).$$

So we obtain $a_t = 0$ or $a_t = \operatorname{tr} A$. If $a_t = 0$, then $\bar{a}_t = -2c/\alpha$. There exists an s that satisfies $\lambda_s \neq \beta$, and hence $a_s = -2c/\alpha$. Thus we have

$$\beta \neq \lambda_s = (2n+1)c + \operatorname{tr} A\Big(\frac{-2c}{\alpha}\Big) - \Big(-\frac{2c}{\alpha}\Big)^2.$$

Thus $\bar{\lambda}_t = \lambda_s \neq \beta$. This is a contradiction. So we see that $a_t = \operatorname{tr} A \neq 0$. In the following, we set $a = a_t = \operatorname{tr} A$. Since $a_t = \bar{a}_t = \operatorname{tr} A$, we have

$$(2a-\alpha)a = (\alpha a + 2c).$$

Thus a satisfies $a^2 - \alpha a - c = 0$, and hence a turns to be constant. In the following, we set $a_1 = -2c/\alpha$ and $\bar{a}_1 = a_2 = 0$.

Next we compute $g(R(e_1, e_2)e_2, e_1)$. By the equation of Gauss,

$$g(R(e_1, e_2)e_2, e_1) = g(R(e_1, \phi e_1)\phi e_1, e_1) = 4c.$$

Using (7), $a_1 g(\nabla_{e_2} e_1, e_2) = 0$. Since $a_1 \neq 0$, we have $g(\nabla_{e_2} e_2, e_1) = 0$. Moreover,

$$g(\nabla_{e_2} e_2, e_2) = 0, \quad g(\nabla_{e_2} e_2, \xi) = -g(e_2, \phi A e_2) = 0.$$

When $k \geq 3$, by (6),

$$a_k g(\nabla_{e_2} e_2, e_k) = 0.$$

When $a_k \neq 0$, we have $g(\nabla_{e_2} e_2, e_k) = 0$. By (10), $g(\nabla_{e_1} e_1, e_2) = 0$. Moreover,

$$g(\nabla_{e_1} e_1, e_1) = 0, \quad g(\nabla_{e_1} e_1, \xi) = 0.$$

Since $k \geq 3$, by (10) and the fact that a_1 is constant,

$$(a_1 - a_k) g(\nabla_{e_1} e_k, e_1) = 0.$$

By $a_1 \neq 0$, if $a_k = 0$, then $g(\nabla_{e_1} e_1, e_k) = 0$. Thus we have

$$\sum_{k=1}^{2n-2} g(\nabla_{e_1} e_1, e_k) g(e_k, \nabla_{e_2} e_2) = 0.$$

So we have

$$\begin{aligned} g(\nabla_{e_1}\nabla_{e_2}e_2, e_1) &= e_1 g(\nabla_{e_2}e_2, e_1) - g(\nabla_{e_2}e_2, \nabla_{e_1}e_1) \\ &= -\sum_k g(\nabla_{e_2}e_2, e_k)g(e_k, \nabla_{e_1}e_1) = 0, \end{aligned}$$

$$\begin{aligned} g(\nabla_{e_2}\nabla_{e_1}e_2, e_1) &= e_2 g(\nabla_{e_1}e_2, e_1) - g(\nabla_{e_1}e_2, \nabla_{e_2}e_1) = -g(\nabla_{e_1}\phi e_1, \nabla_{e_2}e_1) \\ &= g(\nabla_{e_1}e_1, \phi\nabla_{e_2}e_1) = g(\nabla_{e_1}e_1, \nabla_{e_2}e_2) = 0, \end{aligned}$$

and

$$\begin{aligned} &g(\nabla_{[e_1,e_2]}e_2, e_1) \\ &\quad = g(\nabla_\xi e_2, e_1)g(\xi, [e_1, e_2]) + \sum_{k\geq 3} g(\nabla_k e_2, e_1)g(e_k, [e_1, e_2]) \\ &\quad = -a_1 g(\nabla_\xi e_2, e_1) + \sum_{k\geq 3} g(\nabla_{e_k}e_2, e_1)g(e_k, \nabla_{e_1}e_2) - \sum_{k\geq 3} g(\nabla_{e_k}e_2, e_1)g(e_k, \nabla_{e_2}e_1). \end{aligned}$$

By (13),

$$a_1 g(\nabla_\xi e_2, e_1) = c.$$

Using (4), we have

$$g(\nabla_{e_k}e_2, e_1) = \frac{a_k - a_1}{a_1} g(\nabla_{e_2}e_1, e_k).$$

On the other hand, by (8),

$$g(\nabla_{e_k}e_2, e_1) = \frac{a_k}{a_1} g(\nabla_{e_1}e_2, e_k).$$

So we obtain

$$\begin{aligned} &\sum_{k\geq 3} g(\nabla_{e_k}e_2, e_1)(e_k, \nabla_{e_1}e_2) - \sum_{k\geq 3} g(\nabla_{e_k}e_2, e_1)g(e_k, \nabla_{e_2}e_1) \\ &\qquad = \sum \frac{(a_k - a_1)}{a_1} g(\nabla_{e_2}e_1, e_k)g(e_k, \nabla_{e_1}e_2) - \sum \frac{a_k}{a_1} g(\nabla_{e_1}e_2, e_k)(e_k, \nabla_{e_2}e_1) \\ &\qquad = -\sum g(\nabla_{e_2}e_1, e_k)g(e_k, \nabla_{e_1}e_2) \\ &\qquad = -\sum g(\nabla_{e_2}e_1, \phi e_k)g(\phi e_k, \nabla_{e_1}e_2) \\ &\qquad = \sum g(\nabla_{e_2}e_2, e_k)g(e_k, \nabla_{e_1}e_1) = 0. \end{aligned}$$

Thus we have

$$g(R(e_1, e_2)e_2, e_1) = c,$$

from which we obtain $c = 0$. This is a contradiction. Hence we see that M is not Hopf. Thus we have proven Proposition 4.2. □

From Proposition 4.2, if $g((\nabla_X S)Y, \xi) = 0$ for $X, Y \in H$, then M is not Hopf. In the following, we suppose that M is not Hopf, that is, $h_1 \neq 0$. Then, by Lemma 2.2, we can take an orthonormal basis $\{\xi, e_1, \dots, e_{2n-2}\}$ such that

$$(21)\quad A\xi = \alpha\xi + h_1 e_1, \quad Ae_1 = a_1 e_1 + h_1\xi, \quad Ae_j = a_j e_j, \quad j = 2, \dots, 2n-2,$$
$$\operatorname{tr} A = \alpha + a_1, \quad a_2 + \cdots + a_{2n-2} = 0.$$

Then we have

$$\beta = g(S\xi, \xi) = (2n-2)c + (a_1\alpha - h_1^2),$$
$$\lambda_1 = g(Se_1, e_1) = (2n+1)c + (a_1\alpha - h_1^2),$$
$$\lambda_j = g(Se_j, e_j) = (2n+1)c + \operatorname{tr} A \cdot a_j - a_j^2, \quad j \geq 2.$$

By the assumption, for any X and Y orthogonal to ξ,

$$0 = g((\nabla_X S)\xi, Y) = g(\nabla_X S\xi, Y) - g(S\phi AX, Y).$$

We set $SY = \lambda Y$. Then we have

$$0 = (\beta - \lambda) g(\phi AX, Y).$$

Since $\beta \neq \lambda_1$, we see that

$$g(\phi AX, e_1) = -g(AX, \phi e_1) = -g(X, A\phi e_1) = 0$$

for any $X \in H$. We also have $g(\xi, A\phi e_1) = 0$. Thus we have $A\phi e_1 = 0$. In the following, we set $\phi e_1 = e_2$. Then we have

$$0 = (\beta - \lambda_2) g(\phi Ae_1, e_2) = (-3c + a_1\alpha - h_1^2)a_1.$$

Lemma 4.5. *If $h_1 \neq 0$, then $a_2 = 0$. Moreover, $a_1 = 0$ or $a_1\alpha - h_1^2 = 3c$.*

Case (I): Suppose $a_1 = 0$.

Since $a_1 = a_2 = 0$, (13) implies

$$(e_2 h_1) = c + h_1^2.$$

If $\beta = (2n+1)c = \lambda_2$, then $h_1^2 = -3c$ and $e_2 h_1 = 0$. Then we have $h_1^2 = -c$ and $c = 0$. This is a contradiction. So we have:

Lemma 4.6. *If $a_1 = 0$, then $\beta \neq (2n+1)c = \lambda_2$.*

For any $X \in H$, we see that

$$(\beta - \lambda_k) g(\phi AX, e_k) = 0, \quad k \geq 3.$$

If $\beta \neq \lambda_k$, then $g(A\phi e_k, X) = 0$, and moreover $g(A\phi e_k, \xi) = 0$. This shows that $A\phi e_k = 0$ and that ϕe_k is a principal vector of A. We set

$$\bar{\lambda}_k = g(S\phi e_k, \phi e_k).$$

Since $a_1\alpha - h_1^2 \neq 3c$, we have $\bar{\lambda}_k = (2n+1)c \neq \beta$. Then, from

$$(\beta - \bar{\lambda}_k)g(\phi AX, \phi e_k) = 0,$$

we have $g(Ae_k, X) = 0$. We also have $g(Ae_k, \xi) = 0$ since $k \geq 3$. Hence we obtain $Ae_k = 0$ for e_k satisfying $\beta \neq \lambda_k$.

We next consider the case $\beta = \lambda_j$ for some $j \geq 3$. If $\beta = \lambda_j = \lambda_i$, then

$$\beta = (2n+1)c + \operatorname{tr} A \cdot a_j - a_j^2 = (2n+1)c + \operatorname{tr} A \cdot a_i - a_i^2.$$

Therefore, at most two a_j are different. By this equation, we have

$$0 = (a_j - a_i)(\operatorname{tr} A - (a_j + a_i)).$$

If $a_j = a_i = a$ for all j and i, then (21) implies $\sum a_j = 0$. Thus we have all $a_j = 0$, $j = 2, \ldots, 2n-2$. Since $a_1 = 0$, M is a ruled real hypersurface.

Let us suppose that two a_j are different. We set

$$T_a = \{X \mid AX = aX,\ X \in H_x\}, \quad T_b = \{X \mid AX = bX,\ X \in H_x\},$$

where $\beta = \lambda_a = \lambda_b,\ a \neq b$. We notice $\operatorname{tr} A = a + b$. If $a = 0$ or $b = 0$, then, by (21), $a = b = 0$. This contradicts the assumption that $a \neq b$. So we obtain $a \neq 0$ and $b \neq 0$. We notice that $\dim T_a + \dim T_b$ is even number.

Let $e_i, e_j \in T_a$. By (8) and (12),

$$-ag(\nabla_{e_i} e_1, e_j) + ah_1 g(\phi e_i, e_j) = 0,$$
$$(c + a\alpha - a^2)g(\phi e_i, e_j) + h_1 g(\nabla_{e_i} e_1, e_j) = 0.$$

From these, we obtain

$$(c + a\alpha - a^2 + h_1^2)g(\phi e_i, e_j) = 0.$$

If there exist e_i and e_j such that $g(\phi e_i, e_j) \neq 0$, then

$$c + a\alpha - a^2 + h_1^2 = 0.$$

On the other hand, we have

$$\beta = (2n-2)c - h_1^2 = (2n+1)c + \operatorname{tr} A \cdot a - a^2.$$

Since $\operatorname{tr} A = \alpha + a_1 = \alpha$, we have

$$3c + \alpha a - a^2 + h_1^2 = 0.$$

Therefore, we have $2c = 0$. This contradicts $c \neq 0$. Hence $g(\phi e_i, e_j) = 0$ for all e_i and e_j of T_a. So we have $\phi T_a \subset T_b$. Similarly, we also have $\phi T_b \subset T_a$. Consequently, we see that

$$\phi T_a = T_b, \quad \phi T_b = T_a.$$

If $\dim T_a = \dim T_b = 1$, then $\phi T_a = T_b$. We see that if $Ae_j = ae_j$, then $A\phi e_j = b\phi e_j$ and $a + b = \operatorname{tr} A$. From (21), we have $a + b = 0$ and $\operatorname{tr} A = 0$. Therefore, we obtain $\operatorname{tr} A = \alpha = 0$.

We will prove that there is no real hypersurface that satisfies

$$a + b = 0, \quad \alpha = 0, \quad a_1 = 0, \quad a_2 = 0, \quad \operatorname{tr} A = 0,$$

and also

$$a^2 - h_1^2 = 3c.$$

By (5),

$$(2c + 2a^2)g(\phi e_i, \phi e_i) - h_1 g(\nabla_{e_i}\phi e_i, e_1) + h_1 g(\nabla_{\phi e_i} e_i, e_1) = 0. \tag{22}$$

On the other hand, we have

$$g(\nabla_{e_i}\phi e_i, e_1) = g(\phi\nabla_{e_i} e_i, e_1) = -g(\nabla_{e_i} e_i, e_2).$$

By (6),

$$(a_2 - a_i)g(\nabla_{e_i} e_2, e_i) - (e_2 a_i) = 0.$$

Using $a_2 = 0$ and $a_i = a$, we obtain

$$ag(\nabla_{e_i} e_i, e_2) = (e_2 a).$$

From this equation and $a \neq 0$, we have

$$g(\nabla_{e_i} e_i, e_2) = \frac{(e_2 a)}{a}.$$

On the other hand,

$$g(\nabla_{\phi e_i} e_i, e_1) = g(\phi\nabla_{\phi e_i} e_i, \phi e_1) = g(\nabla_{\phi e_i}\phi e_i, e_2).$$

By (6), we obtain

$$(a_2 + a)g(\nabla_{\phi e_i} e_2, \phi e_i) + (e_2 a) = 0,$$

and hence

$$g(\nabla_{\phi e_i}\phi e_i, e_2) = \frac{(e_2 a)}{a}.$$

Substituting these equations into (22), we get

$$2(c + a^2) + h_1\frac{(e_2 a)}{a} + h_1\frac{(e_2 a)}{a} = 0.$$

Thus we have

$$(c + a^2)a = -h_1(e_2 a). \tag{23}$$

On the other hand, since $a^2 - h_1^2 = 3c$,

$$a(e_2 a) = h_1(e_2 h_1).$$

Since $a_1 = a_2 = 0$, by (13), we have

$$e_2 h_1 = c + h_1^2,$$

from which we obtain

$$e_2 a = \frac{h_1}{a}(c + h_1^2).$$

Substituting this into (23), we get

$$(c + a^2)a = -\frac{h_1^2}{a}(c + h_1^2) = -\frac{1}{a}(a^2 - 3c)(a^2 - 2c).$$

Thus we obtain

$$(a^2 - c)^2 + 2c^2 = 0.$$

So we have $c = 0$. This is a contradiction. Consequently, if $a_1 = 0$, then M is a ruled real hypersurface.

Case (II): Suppose $a_1 \neq 0$.

We notice that $a_2 = 0$ and $\alpha a_1 h_1^2 = 3c$ by Lemma 4.5. So we have

$$(Xa_1)\alpha + a_1(X\alpha) - 2h_1(Xh_1) = 0 \tag{24}$$

for any tangent vector field X.

Lemma 4.7. *$\nabla_{e_1} e_1$ and $\nabla_{e_2} e_2$ are perpendicular to ξ, e_1 and e_2.*

Proof. By (14),

$$(e_2\alpha) = \alpha h_1 + h_1 g(\nabla_\xi e_1, e_2).$$

By (10),

$$(e_2 a_1) = a_1 h_1 + a_1 g(\nabla_{e_1} e_1, e_2).$$

Substituting these into (24), we get

$$2a_1\alpha h_1 + \alpha a_1 g(\nabla_{e_1} e_1, e_2) + a_1 h_1 g(\nabla_\xi e_1, e_2) - 2h_1(e_2 h_1) = 0.$$

By (9) and (13),

$$\begin{aligned}
(e_2 h_1) &= (2c + \alpha a_1) + h_1 g(\nabla_{e_1} e_1, e_2) = (5c + h_1^2) + h_1 g(\nabla_{e_1} e_1, e_2), \\
(e_2 h_1) &= (c + h_1^2) + a_1 g(\nabla_\xi e_1, e_2).
\end{aligned}$$

From these equations and (24), we have

$$2h_1(a_1\alpha - h_1^2 - 3c) + (a_1\alpha - h_1^2) g(\nabla_{e_1} e_1, e_2) = 0.$$

Since $a_1\alpha - h_1^2 = 3c$, we have

$$g(\nabla_{e_1}e_1, e_2) = 0.$$

By (7), $a_1 \neq 0$ and $a_2 = 0$,

$$g(\nabla_{e_2}e_2, e_1) = 0.$$

Moreover, we have

$$g(\nabla_{e_2}e_2, \xi) = -g(e_2, \phi Ae_2) = 0, \quad g(\nabla_{e_1}e_1, \xi) = -g(e_1, \phi Ae_1) = 0.$$

These equations prove our lemma. □

Lemma 4.8. *Suppose $j \geq 3$. If $a_j = 0$, then $g(\nabla_{e_1}e_1, e_j) = 0$. If $a_j \neq 0$, then $g(\nabla_{e_2}e_2, e_j) = 0$.*

Proof. By (6), we have

$$a_j g(\nabla_{e_2}e_2, e_j) = 0, \quad j \geq 3.$$

If $a_j \neq 0$, then $g(\nabla_{e_2}e_2, e_j) = 0$ for $j \geq 3$. Suppose $a_j = 0,\ j \geq 3$. Then, by (10), (14), (9) and (13),

$$\begin{aligned} (e_j a_1) &= a_1 g(\nabla_{e_1}e_1, e_j), \quad & (e_j\alpha) &= h_1 g(\nabla_\xi e_1, e_j), \\ (e_j h_1) &= h_1 g(\nabla_{e_1}e_1, e_j), \quad & (e_j h_1) &= a_1 g(\nabla_\xi e_1, e_j). \end{aligned}$$

Substituting these into (24), we get

$$\begin{aligned} 0 &= (e_j a_1)\alpha + a_1(e_j\alpha) - 2h_1(e_j h_1) \\ &= \alpha a_1 g(\nabla_{e_1}e_1, e_j) + a_1 h_1 g(\nabla_\xi e_1, e_j) - h_1^2 g(\nabla_{e_1}e_1, e_j) - h_1 a_1 g(\nabla_\xi e_1, e_j) \\ &= (\alpha a_1 - h_1^2) g(\nabla_{e_1}e_1, e_j). \end{aligned}$$

Since $a_1\alpha - h_1^2 = 3c$, we have our lemma. □

Using these lemmas, we compute $g(R(e_1, e_2)e_2, e_1)$. We note that $e_2 = \phi e_1$ and $a_2 = 0$. First, we have

$$\begin{aligned} g(\nabla_{e_1}\nabla_{e_2}e_2, e_1) &= e_1 g(\nabla_{e_2}e_2, e_1) - g(\nabla_{e_2}e_2, \nabla_{e_1}e_1) \\ &= -g(\nabla_{e_2}e_2, \xi)g(\xi, \nabla_{e_1}e_1) - g(\nabla_{e_2}e_2, e_1)g(e_1, \nabla_{e_1}e_1) \\ &\quad - g(\nabla_{e_2}e_2, e_2)g(e_2, \nabla_{e_1}e_1) - \sum_{k\geq 3} g(\nabla_{e_2}e_2, e_j)g(e_j, \nabla_{e_1}e_1) = 0. \end{aligned}$$

Next, we have

$$\begin{aligned} &g(\nabla_{e_2}\nabla_{e_1}e_2, e_1)\\ &\qquad = e_2 g(\nabla_{e_1}e_2, e_1) - g(\nabla_{e_1}e_2, \nabla_{e_2}e_1)\\ &\qquad = -g(\nabla_{e_1}e_2, \xi)g(\xi, \nabla_{e_2}e_1) - g(\nabla_{e_1}e_2, e_1)g(e_1, \nabla_{e_2}e_1)\\ &\qquad\qquad - g(\nabla_{e_1}e_2, \xi)g(\xi, \nabla_{e_2}e_1) - \sum_{k\geq 3} g(\nabla_{e_1}e_2, e_k)g(e_k, \nabla_{e_2}e_1)\\ &\qquad = -\sum_{k\geq 3} g(\nabla_{e_1}e_2, e_k)g(e_k, \nabla_{e_2}e_1) = -\sum_{k\geq 3} g(\nabla_{e_1}\phi e_1, e_k)g(\phi e_k, \phi\nabla_{e_2}e_1)\\ &\qquad = \sum_{k\geq 3} g(\nabla_{e_1}e_1, \phi e_k)g(\phi e_k, \nabla_{e_2}e_2) = \sum_{l\geq 3} g(\nabla_{e_1}e_1, e_l)g(e_l, \nabla_{e_2}e_2) = 0. \end{aligned}$$

Moreover, we obtain

$$\begin{aligned} g(\nabla_{[e_1,e_2]}e_2, e_1) &= g(\nabla_{\xi}e_2, e_1)g(\xi, [e_1, e_2]) + g(\nabla_{e_1}e_2, e_1)g(e_1, [e_1, e_2])\\ &\qquad + g(\nabla_{e_2}e_2, e_1)g(e_2.[e_1, e_2]) + \sum_{k\geq 3} g(\nabla_{e_k}e_2, e_1)g(e_k, [e_1, e_2])\\ &= g(\nabla_{\xi}e_2, e_1)g(\xi, \nabla_{e_1}e_2)\\ &\qquad + \sum_{k\geq 3}(g(\nabla_{e_k}e_2, e_1)g(e_k, \nabla_{e_1}e_2) - g(\nabla_{e_k}e_2, e_1)g(e_k, \nabla_{e_2}e_1)). \end{aligned}$$

On the other hand, by (8), when $j \geq 3$,

$$a_1 g(\nabla_{e_j}e_2, e_1) - a_j g(\nabla_{e_1}e_2, e_j) = 0,$$
$$(a_1 - a_j)g(\nabla_{e_2}e_1, e_j) + a_j g(\nabla_{e_1}e_2, e_j) = 0.$$

Thus, if $a_1 = a_j$, then we see that $a_j \neq 0$ and hence $g(\nabla_{e_1}e_2, e_j) = 0$ since $a_1 \neq 0$. Next, when $a_1 \neq a_j$ we have

$$g(\nabla_{e_2}e_1, e_j) = -\frac{a_j}{(a_1 - a_j)} g(\nabla_{e_1}e_2, e_j).$$

On the other hand,

$$g(\nabla_{e_j}e_2, e_1) = \frac{a_j}{a_1} g(\nabla_{e_1}e_2, e_j) = -\frac{(a_1 - a_j)}{a_1} g(\nabla_{e_2}e_1, e_j).$$

So we have

$$\begin{aligned} &\sum_{k\geq 3}(g(\nabla_{e_k}e_2, e_1)g(e_k, \nabla_{e_1}e_2) - g(\nabla_{e_k}e_2, e_1)g(e_k, \nabla_{e_2}e_1)\\ &\qquad = -\sum_{k\geq 3} g(\nabla_{e_2}e_1, e_k)g(e_k, \nabla_{e_1}e_2) = -\sum_{k\geq 3} g(\phi\nabla_{e_2}e_1, e_k)g(\phi e_k, \nabla_{e_1}e_2)\\ &\qquad = \sum_{l\geq 3} g(\nabla_{e_1}e_1, e_l)g(e_l, \nabla_{e_2}e_2) = 0. \end{aligned}$$

Thus we obtain

$$\begin{aligned} g(\nabla_{[e_1,e_2]}e_2, e_1) &= g(\nabla_\xi e_2, e_1)g(\xi, \nabla_{e_1}e_2) \\ &= -g(\nabla_\xi e_2, e_1)g(\phi A e_1, e_2) = -a_1 g(\nabla_\xi e_2, e_1), \end{aligned}$$

and so

$$g(R(e_1, e_2)e_2, e_1) = a_1 g(\nabla_\xi e_2, e_1).$$

On the other hand, by (9),

$$-(2c + \alpha a_1) + h_1 g(\nabla_{e_1} e_2, e_1) + (e_2 h_1) = 0.$$

Using Lemma 4.7 and $a_1\alpha - h_1^2 = 3c$, we have

$$(e_2 h_1) = 2c + \alpha a_1 = 5c + h_1^2.$$

By (13),

$$-(c + h_1^2) + a_1 g(\nabla_\xi e_2, e_1) + e_2 h_1 = 0,$$

from which we obtain

$$a_1 g(\nabla_\xi e_2, e_1) = -4c,$$

and so

$$g(R(e_1, e_2)e_2, e_1) = -4c.$$

On the other hand, the equation of Gauss implies

$$g(R(e_1, e_2)e_2, e_1) = 4c,$$

and hence $c = 0$. This is a contradiction.

Consequently, M is a ruled real hypersurface.

From (2), any ruled real hypersurface satisfies $g((\nabla_X S)Y, \xi) = 0$ for any X and Y orthogonal to ξ, and $S\xi = \beta\xi$ for some function β. □

From Theorems 3.1 and 4.1, we have Theorem 1.1.

References

[Berndt 1989] J. Berndt, "Real hypersurfaces with constant principal curvatures in complex hyperbolic space", *J. Reine Angew. Math.* **395** (1989), 132–141. MR 90d:53062 Zbl 0655.53046

[Berndt and Brück 2001] J. Berndt and M. Brück, "Cohomogeneity one actions on hyperbolic spaces", *J. Reine Angew. Math.* **541** (2001), 209–235. MR 2002j:53059 Zbl 1014.53042

[Berndt and Tamaru 2007] J. Berndt and H. Tamaru, "Cohomogeneity one actions on noncompact symmetric spaces of rank one", *Trans. Amer. Math. Soc.* **359**:7 (2007), 3425–3438. MR 2008d:53063 Zbl 1117.53041

[Cecil and Ryan 1982] T. E. Cecil and P. J. Ryan, "Focal sets and real hypersurfaces in complex projective space", *Trans. Amer. Math. Soc.* **269**:2 (1982), 481–499. MR 83b:53049 Zbl 0492.53039

[Ki 1989] U.-H. Ki, "Real hypersurfaces with parallel Ricci tensor of a complex space form", *Tsukuba J. Math.* **13**:1 (1989), 73–81. MR 90c:53135 Zbl 0678.53046

[Ki et al. 1990] U.-H. Ki, H. Nakagawa, and Y. J. Suh, "Real hypersurfaces with harmonic Weyl tensor of a complex space form", *Hiroshima Math. J.* **20**:1 (1990), 93–102. MR 91c:53051 Zbl 0716.53026

[Kim and Ryan 2008] H. S. Kim and P. J. Ryan, "A classification of pseudo-Einstein hypersurfaces in $\mathbb{C}P^2$", *Differential Geom. Appl.* **26**:1 (2008), 106–112. MR 2008m:53135 Zbl 1143.53050

[Kimura 1986] M. Kimura, "Real hypersurfaces and complex submanifolds in complex projective space", *Trans. Amer. Math. Soc.* **296**:1 (1986), 137–149. MR 87k:53133 Zbl 0597.53021

[Kimura and Maeda 1989] M. Kimura and S. Maeda, "On real hypersurfaces of a complex projective space", *Math. Z.* **202**:3 (1989), 299–311. MR 90h:53067 Zbl 0661.53015

[Kon 1979] M. Kon, "Pseudo-Einstein real hypersurfaces in complex space forms", *J. Differential Geom.* **14**:3 (1979), 339–354. MR 81k:53050 Zbl 0461.53031

[Kon 2014] M. Kon, "3-dimensional real hypersurfaces and Ricci operator", *Differential Geom. Dyn. Syst.* **16** (2014), 189–202. MR 3226614

[Lohnherr 1998] M. Lohnherr, *On ruled real hypersurfaces of complex space forms*, Ph.D. thesis, University of Cologne, 1998.

[Lohnherr and Reckziegel 1999] M. Lohnherr and H. Reckziegel, "On ruled real hypersurfaces in complex space forms", *Geom. Dedicata* **74**:3 (1999), 267–286. MR 99m:53120 Zbl 0932.53018

[Maeda 1976] Y. Maeda, "On real hypersurfaces of a complex projective space", *J. Math. Soc. Japan* **28**:3 (1976), 529–540. MR 53 #11543 Zbl 0324.53039

[Maeda 2013] S. Maeda, "Hopf hypersurfaces with η-parallel Ricci tensors in a nonflat complex space form", *Sci. Math. Jpn.* **76**:3 (2013), 449–456. MR 3310013

[Montiel 1985] S. Montiel, "Real hypersurfaces of a complex hyperbolic space", *J. Math. Soc. Japan* **37**:3 (1985), 515–535. MR 86i:53027 Zbl 0554.53021

[Niebergall and Ryan 1997] R. Niebergall and P. J. Ryan, "Real hypersurfaces in complex space forms", pp. 233–305 in *Tight and taut submanifolds* (Berkeley, CA, 1994), edited by T. E. Cecil and S.-S. Chern, Math. Sci. Res. Inst. Publ. **32**, Cambridge University Press, 1997. MR 98j:53066 Zbl 0904.53005

[Suh 1990] Y. J. Suh, "On real hypersurfaces of a complex space form with η-parallel Ricci tensor", *Tsukuba J. Math.* **14**:1 (1990), 27–37. MR 91h:53047 Zbl 0721.53029

[Takagi 1973] R. Takagi, "On homogeneous real hypersurfaces in a complex projective space", *Osaka J. Math.* **10** (1973), 495–506. MR 49 #1433 Zbl 0274.53062

[Takagi 1975a] R. Takagi, "Real hypersurfaces in a complex projective space with constant principal curvatures", *J. Math. Soc. Japan* **27** (1975), 43–53. MR 50 #8380 Zbl 0292.53042

[Takagi 1975b] R. Takagi, "Real hypersurfaces in a complex projective space with constant principal curvatures II", *J. Math. Soc. Japan* **27**:4 (1975), 507–516. MR 53 #3955 Zbl 0311.53064

Received May 7, 2015. Revised July 21, 2015.

MAYUKO KON
FACULTY OF EDUCATION
SHINSHU UNIVERSITY
6-RO, NISHINAGANO
NAGANO CITY 380-8544
JAPAN
mayuko_k@shinshu-u.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.125

MONOTONICITY FORMULAE AND VANISHING THEOREMS

JINTANG LI

We study Cartan–Hadamard manifolds with pinching conditions. Using the stress-energy tensor, we establish some monotonicity formulae for vector bundle-valued *p*-forms and pluriharmonic maps between Kähler manifolds. Some vanishing theorems follow immediately from the monotonicity formulae under suitable growth conditions on the energy of *p*-forms and pluriharmonic maps.

1. Introduction

Harmonic maps between Riemannian manifolds are defined as the critical points of energy functionals. They are important in both classical and modern differential geometry. As is well-known, any harmonic map $\phi : R^n \to S^m$ with finite energy must be constant [Garber et al. 1979]. This result has been generalized by Sealey [1982] to harmonic maps from a space form of nonpositive sectional curvature to any Riemannian manifold with finite energy. In 1980, Baird and Eells [1981] introduced and studied the stress-energy tensor for maps between Riemannian manifolds. Sealey [≥ 2016] introduced the stress-energy tensor for vector bundle-valued p-forms and established some vanishing theorems for L^2 harmonic p-forms. The stress-energy tensors have become a useful tool for investigating the energy behavior of vector bundle-valued p-forms in various problems. Dong and Lin [2014] introduced the notion of J-invariant p-forms on Kähler manifolds. They established a monotonicity formula by means of the stress-energy tensor. Using this monotonicity formula they proved the following vanishing theorem for vector bundle-valued J-invariant p-forms satisyfing the conservation law:

Theorem A. *Let M be a complex n-dimensional* $(n \geq 3)$ *complete Kähler manifold with radial curvature K_r satisfying $-a^2 \leq K_r \leq -b^2 < 0$ with $a \geq b > 0$ and* $(2n-1)b - 2pa \geq 0$. *Let $\xi : E \to M$ be a smooth Riemannian vector bundle*

This research is supported by the Natural Science Foundation of Fujian Province of China (No. 2014J01022) and the Fundamental Research Funds for the Central Universities (No. 20720150008).
MSC2010: primary 53C43; secondary 53C55.
Keywords: monotonicity formulae, Kähler manifolds, p-forms, pluriharmonic maps.

over (M, g). *If* $\omega \in A^p(\xi)$ *is J-invariant and satisfies the conservation law*, *that is*, $\operatorname{div} S_\omega = 0$, *then*

$$\frac{1}{r_1^C}\int_{B_{r_1}(x_0)} |\omega|^2\,dv \le \frac{1}{r_2^C}\int_{B_{r_2}(x_0)} |\omega|^2\,dv$$

for any $0 < r_1 < r_2$, *where* $C = 2n - 2pa/b$ *and* $B_r(x_0) \subseteq M$ *is a geodesic ball of radius* r *centered at* x_0 *in* M. *In particular*, *if*

$$\frac{1}{r^C}\int_{B_r(x_0)} |\omega|^2\,dv \to 0 \quad \textit{as } r \to +\infty,$$

then $\omega = 0$.

We shall establish a monotonicity formula for vector bundle-valued J-invariant p-forms satisfying the conservation law by means of the stress-energy tensor too. Using this monotonicity formula we can improve Theorem A as follows:

Theorem 1. *Let* M *be a complex* n*-dimensional* $(n \ge 3)$ *complete Kähler manifold with radial curvature* K_r *satisfying* $-a^2 \le K_r \le -b^2 < 0$ *with* $a \ge b > 0$ *and* $(n-1)b - (p-1)a \ge 0$. *Let* $\xi : E \to M$ *be a smooth Riemannian vector bundle over* (M, g). *If* $\omega \in A^p(\xi)$ *is J-invariant and satisfies the conservation law*, *that is*, $\operatorname{div} S_\omega = 0$, *then*

$$\frac{1}{\sinh^C(ar_1)}\int_{B_{r_1}(x_0)} \cosh(ar)\big[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\big]\,dv \\ \le \frac{1}{\sinh^C(ar_2)}\int_{B_{r_2}(x_0)} \cosh(ar)\big[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\big]\,dv$$

for any $0 < r_1 < r_2$, *where* $C = [2(n-2)b - 2(p-2)a]/a$. *In particular*, *if*

$$\frac{\int_{B_r(x_0)} |\omega|^2\,dv}{e^{a(C-1)r}} \to 0 \quad \textit{as } r \to +\infty,$$

then $\omega = 0$. (See Section 2 for the definition of $i_{\partial/\partial r}\omega$.)

For the case of Cartan–Hadamard manifolds with some pinching conditions, Xin [1986] established a monotonicity formula for vector bundle-valued p-forms satisfying the conservation law by means of the stress-energy tensor. Using this monotonicity formula, Xin proved the following vanishing theorem:

Theorem B. *Let* M *be an* n*-dimensional* $(n \ge 3)$ *complete Riemannian manifold with radial curvature* K_r *satisfying* $-a^2 \le K_r \le -b^2 < 0$ *with* $a \ge b > 0$ *and* $(n-1)b - 2pa \ge 0$. *Let* $\xi : E \to M$ *be a smooth Riemannian vector bundle over* (M, g). *If* $\omega \in A^p(\xi)$ *satisfies the conservation law*, *that is*, $\operatorname{div} S_\omega = 0$, *and*

$$\frac{1}{r^C}\int_{B_r(x_0)} |\omega|^2\,dv \to 0 \quad \textit{as } r \to +\infty,$$

where $C = n - 2pa/b$, *then* $\omega = 0$.

We shall establish a monotonicity formula for vector bundle-valued p-forms satisfying the conservation law by means of the stress-energy tensor. Using this monotonicity formula we can improve Theorem B as follows:

Theorem 2. *Let M be an n-dimensional $(n \geq 3)$ complete Riemannian manifold with radial curvature K_r satisfying $-a^2 \leq K_r \leq -b^2 < 0$ with $a \geq b > 0$ and $(n-1)b-(2p-1)a \geq 0$. Let $\xi : E \to M$ be a smooth Riemannian vector bundle over (M, g). If $\omega \in A^p(\xi)$ satisfies the conservation law, that is,* $\operatorname{div} S_\omega = 0$, *then*

$$\frac{1}{\sinh^C(ar_1)} \int_{B_{r_1}(x_0)} \cosh(ar)\big[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\big]\,dv \leq \frac{1}{\sinh^C(ar_2)} \int_{B_{r_2}(x_0)} \cosh(ar)\big[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\big]\,dv$$

for any $0 < r_1 < r_2$, where $C = [(n-2)b-(2p-2)a]/a$ and $B_r(x_0) \subseteq M$ is a geodesic ball of radius r centered at x_0 in M. In particular, if

$$\frac{\int_{B_r(x_0)} |\omega|^2\,dv}{e^{a(C-1)r}} \to 0 \quad \text{as } r \to +\infty,$$

then $\omega = 0$.

Siu [1980] introduced and studied pluriharmonic maps from a compact Kähler manifold to a Kähler manifold. When the domain of such a map is complete, Dong [2013] proved the following:

Theorem C. *Let M be a complex n-dimensional $(n \geq 2)$ complete Kähler manifold with radial curvature K_r satisfying $K_r \leq -b^2 < 0$ with $b > 0$. Suppose $\phi : M \to N$ is either a pluriharmonic map between Kähler manifolds or a harmonic map into a Kähler manifold with strongly seminegative curvature. Then*

$$\frac{\int_{B_{r_1}(x_0)} |\bar\partial\phi|^2\,dv}{r_1^{2(n-1)}} \leq \frac{\int_{B_{r_2}(x_0)} |\bar\partial\phi|^2\,dv}{r_2^{2(n-1)}} \quad \text{and} \quad \frac{\int_{B_{r_1}(x_0)} |\partial\phi|^2\,dv}{r_1^{2(n-1)}} \leq \frac{\int_{B_{r_2}(x_0)} |\partial\phi|^2\,dv}{r_2^{2(n-1)}}$$

for any $0 < r_1 < r_2$. In particular, if

$$\frac{\int_{B_r(x_0)} |d\phi|^2\,dv}{r^{(2n-2)}} \to 0 \quad \text{as } r \to +\infty,$$

then ϕ is constant.

We can also improve Theorem C as follows:

Theorem 3. *Let M be a complex n-dimensional $(n \geq 2)$ complete Kähler manifold with radial curvature K_r satisfying $K_r \leq -b^2 < 0$ with $b > 0$. Suppose $\phi : M \to N$ is either a pluriharmonic map between Kähler manifolds or a harmonic map into a*

Kähler manifold with strongly seminegative curvature. Then

$$\frac{\int_{B_{r_1}(x_0)} \cosh(br)|\bar\partial\phi|^2\,dv}{\sinh^{2(n-1)}(br_1)} \le \frac{\int_{B_{r_2}(x_0)} \cosh(br)|\bar\partial\phi|^2\,dv}{\sinh^{2(n-1)}(br_2)}$$

and

$$\frac{\int_{B_{r_1}(x_0)} \cosh(br)|\partial\phi|^2\,dv}{\sinh^{2(n-1)}(br_1)} \le \frac{\int_{B_{r_2}(x_0)} \cosh(br)|\partial\phi|^2\,dv}{\sinh^{2(n-1)}(br_2)}$$

for any $0 < r_1 < r_2$. *In particular, if*

$$\frac{\int_{B_r(x_0)} |d\phi|^2\,dv}{e^{(2n-3)br}} \to 0 \qquad \text{as } r \to +\infty,$$

then ϕ *is constant.*

2. Preliminaries

Let (M, g) be an n-dimensional complete Riemannian manifold. Let $\xi : E \to M$ be a smooth Riemannian vector bundle over (M, g). Let $A^p(\xi) = \Gamma(\Lambda^p T^*M \otimes E)$ be the space of smooth p-forms on M with values in the vector bundle $\xi : E \to M$. For $\omega \in A^p(\xi)$, we define the energy functional of ω by

$$E(\omega) = \int_M \tfrac{1}{2}|\omega|^2\,dv_g.$$

The stress-energy tensor associated with $E(\omega)$ is defined by

$$S_\omega(X, Y) = \tfrac{1}{2}|\omega|^2 g(X, Y) - (\omega \odot \omega)(X, Y), \tag{2-1}$$

where $\omega \odot \omega$ denotes the 2-tensor

$$(\omega \odot \omega)(X, Y) = \langle i_X\omega, i_Y\omega\rangle.$$

Here $\langle\cdot,\cdot\rangle$ is the induced inner product on $A^{p-1}(\xi)$ and $i_X\omega$ is the interior multiplication by the vector field X given by

$$(i_X\omega)(Y_1, \dots, Y_{p-1}) = \omega(X, Y_1, \dots, Y_{p-1})$$

for $\omega \in A^p(\xi)$ and any vector fields $Y_1, \dots, Y_{p-1}$ on M.

Let D be any bounded domain of M with C^1 boundary. We have the integral formula [Dong 2013]

$$\int_{\partial D} S_\omega(X, \nu)\,dv = \int_D \{\langle S_\omega, \nabla\theta_X\rangle + (\operatorname{div} S_\omega)(X)\}\,dv, \tag{2-2}$$

where ν is the unit normal vector field along ∂D in D, and θ_X is the dual 1-form of X and $\nabla\theta_X$ is given by

$$(\nabla\theta_X)(Y, Z) = g(\nabla_Y X, Z). \tag{2-3}$$

Proposition 2.1 [Greene and Wu 1979]. *Let (M, g) be a complete Riemannian manifold with a pole x_0 and let r be the distance function relative to x_0. Denote by K_r the radial curvature of M. If $-a^2 \le K_r \le -b^2 < 0$, where $a \ge b > 0$, then*

$$b \coth(br)[g - dr \otimes dr] \le \operatorname{Hess}(r) \le a \coth(ar)[g - dr \otimes dr],$$

where $\operatorname{Hess}(r)$ *is the Hessian of the distance function r.*

3. Monotonicity formulae for Kähler manifolds

A Hermitian metric on a complex manifold M is a Riemannian metric g such that $g(JX, JY) = g(X, Y)$ for any $X, Y \in TM$, where J denotes the complex structure of M. We say that (M, g) is a Kähler manifold if $\nabla J = 0$, where ∇ is the Levi-Civita connection of g. A p-form $\omega \in A^p(\xi)$ is called J-invariant if $(\omega \odot \omega)(JX, JY) = (\omega \odot \omega)(X, Y)$. Now we consider J-invariant p-forms on Kähler manifolds and can prove the following:

Theorem 3.1. *Let M be a complex n-dimensional $(n \ge 3)$ complete Kähler manifold with radial curvature K_r satisfying $-a^2 \le K_r \le -b^2 < 0$ with $a \ge b > 0$ and $(n-1)b \ge (p-1)a$. Let $\xi : E \to M$ be a smooth Riemannian vector bundle over (M, g). If $\omega \in A^p(\xi)$ is J-invariant and satisfies the conservation law, that is,* $\operatorname{div} S_\omega = 0$, *then*

$$\frac{1}{\sinh^C(ar_1)} \int_{B_{r_1}(x_0)} \cosh(ar)\left[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\right] dv \le \frac{1}{\sinh^C(ar_2)} \int_{B_{r_2}(x_0)} \cosh(ar)\left[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\right] dv$$

for any $0 < r_1 < r_2$, where $C = [2(n-2)b - 2(p-2)a]/a$ and $B_r(x_0) \subseteq M$ is a geodesic ball of radius r centered at x_0 in M.

Proof. If $X = \operatorname{grad}(\psi)$ is the gradient of a smooth function ψ on M, then $\theta_X = d\psi$ and $\nabla\theta_X = \operatorname{Hess}(\psi)$. Let $\psi = \cosh(ar)$. It is easy to see that

$$\operatorname{Hess}(\cosh(ar)) = a^2 \cosh(ar)\, dr \otimes dr + a \sinh(ar) \operatorname{Hess}(r). \tag{3-1}$$

Let $\{e_i, Je_i\}$ with $e_n = \partial/\partial r$ be an orthonormal frame field around $x_0 \in M$. Then, for $\omega \in A^p(\xi)$, we have

$$\begin{aligned} |\omega|^2 &= \frac{1}{p}\Bigg[(\omega \odot \omega)\Big(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}\Big) + (\omega \odot \omega)\Big(J\frac{\partial}{\partial r}, J\frac{\partial}{\partial r}\Big) \\ &\qquad + \sum_{\lambda=1}^{n-1}(\omega \odot \omega)(e_\lambda, e_\lambda) + \sum_{\lambda=1}^{n-1}(\omega \odot \omega)(Je_\lambda, Je_\lambda)\Bigg] \\ &= \frac{2}{p}\Bigg\{|i_{\partial/\partial r}\omega|^2 + \sum_{\lambda=1}^{n-1}(\omega \odot \omega)(e_\lambda, e_\lambda)\Bigg\}. \end{aligned} \tag{3-2}$$

It follows from (3-1), (3-2) and Proposition 2.1 that

$$
\begin{aligned}
(3\text{-}3)\quad &\langle S_\omega, \nabla\theta_X\rangle \\
&= \tfrac{1}{2}|\omega|^2\langle g, \mathrm{Hess}(\mathrm{ch}(ar))\rangle - \langle(\omega\odot\omega), \mathrm{Hess}(\mathrm{ch}(ar))\rangle \\
&= \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\Big\{\sum_\lambda a\sinh(ar)\,\mathrm{Hess}(r)(e_\lambda, e_\lambda) \\
&\qquad + \sum_\lambda a\sinh(ar)\,\mathrm{Hess}(r)(Je_\lambda, Je_\lambda) \\
&\qquad - (p-1)a^2\cosh(ar) - pa\sinh(ar)\,\mathrm{Hess}(r)\Big(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}\Big) \\
&\qquad - (p-1)a\sinh(ar)\,\mathrm{Hess}(r)\Big(J\frac{\partial}{\partial r}, J\frac{\partial}{\partial r}\Big)\Big\} \\
&\quad + \sum_\lambda \tfrac{1}{p}(\omega\odot\omega)(e_\lambda, e_\lambda)\Big\{a^2\cosh(ar) + a\sinh(ar)\,\mathrm{Hess}(r)\Big(J\frac{\partial}{\partial r}, J\frac{\partial}{\partial r}\Big) \\
&\qquad + \sum_\mu a\sinh(ar)\,\mathrm{Hess}(r)(e_\mu, e_\mu) + \sum_\mu a\sinh(ar)\,\mathrm{Hess}(r)(Je_\mu, Je_\mu) \\
&\qquad - p\,\mathrm{Hess}(\cosh(ar))(e_\lambda, e_\lambda) - p\,\mathrm{Hess}(\cosh(ar))(Je_\lambda, Je_\lambda)\Big\} \\
&\ge \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\{2(n-1)ab\sinh(ar)\coth(br) - 2(p-1)a^2\cosh(ar)\} \\
&\quad + \sum_\lambda \tfrac{1}{p}(\omega\odot\omega)(e_\lambda, e_\lambda)\{2(n-2)ab\sinh(ar)\coth(br) \\
&\qquad - 2(p-2)a^2\cosh(ar)\} \\
&\ge \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2 a\cosh(ar)\{2(n-1)b - 2(p-1)a\} \\
&\quad + \sum_\lambda \tfrac{1}{p}(\omega\odot\omega)(e_\lambda, e_\lambda)a\cosh(ar)\{2(n-2)b - 2(p-2)a\} \\
&\ge \sum_\lambda \tfrac{1}{p}(\omega\odot\omega)(e_\lambda, e_\lambda)a\cosh(ar)\{2(n-2)b - 2(p-2)a\}.
\end{aligned}
$$

On the other hand, we have

$$
\begin{aligned}
(3\text{-}4)\qquad S_\omega\Big(X, \frac{\partial}{\partial r}\Big) &= \tfrac{1}{2}|\omega|^2 a\sinh(ar) - a\sinh(ar)(\omega\odot\omega)\Big(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}\Big) \\
&\le \big[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\big]a\sinh(ar).
\end{aligned}
$$

Substituting (3-3) and (3-4) into (2-2), we obtain

$$
\begin{aligned}
(3\text{-}5)\quad &\int_{\partial B_r(x_0)} \big[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\big]a\sinh(ar)\,ds \\
&\qquad \ge \int_{B_r(x_0)} \big[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\big][2(n-2)b - 2(p-2)a]a\cosh(ar)\,dv.
\end{aligned}
$$

It can be seen from (3-5) that

$$\frac{\cosh(ar)\int_{\partial B_r(x_0)}\left[\frac{1}{2}|\omega|^2-\frac{1}{p}|i_{\partial/\partial r}\omega|^2\right]ds}{\int_{B_r(x_0)}\left[\frac{1}{2}|\omega|^2-\frac{1}{p}|i_{\partial/\partial r}\omega|^2\right]\cosh(ar)\,dv}\geq\frac{aC\cosh(ar)}{\sinh(ar)},\tag{3-6}$$

where $C=[2(n-2)b-2(p-2)a]/a$.

Thus we obtain from (3-6)

$$\frac{d}{dr}\ln\left\{\int_{B_r(x_0)}\left[\tfrac{1}{2}|\omega|^2-\tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\right]\cosh(ar)\,dv\right\}\geq\frac{d}{dr}\{C\ln[\sinh(ar)]\}.\tag{3-7}$$

Integrating (3-7) over $[r_1,r_2]$, we have

$$\begin{aligned}\ln\int_{B_{r_2}(x_0)}&\left[\tfrac{1}{2}|\omega|^2-\tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\right]\cosh(ar)\,dv\\&-\ln\int_{B_{r_1}(x_0)}\left[\tfrac{1}{2}|\omega|^2-\tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\right]\cosh(ar)\,dv\\&\geq C\ln[\sinh(ar_2)]-C\ln[\sinh(ar_1)].\end{aligned}\tag{3-8}$$ □

Now we can deduce the following vanishing theorem from the above monotonicity formula.

Theorem 3.2. *Let M be a complex n-dimensional ($n\geq 3$) complete Kähler manifold with radial curvature K_r satisfying $-a^2\leq K_r\leq -b^2<0$ with $a\geq b>0$ and $(n-1)b\geq(p-1)a$. Let $\xi:E\to M$ be a smooth Riemannian vector bundle over (M,g). If the J-invariant p-form $\omega\in A^p(\xi)$ satisfies the conservation law, that is,* div $S_\omega=0$, *and*

$$\frac{\int_{B_r(x_0)}|\omega|^2\,dv}{e^{a(C-1)r}}\to 0\quad as\ r\to+\infty,$$

where $C=[2(n-2)b-2(p-2)a]/a$, then $\omega\equiv 0$.

Proof. **Case 1.** If $1\geq(n-1)b-(p-1)a\geq 0$, i.e., $C\leq 1$, it is obvious that $\omega\equiv 0$.

Case 2. If $(n-1)b-(p-1)a>1$, i.e., $C>1$, using the fact $\coth(ar)\to 1$ as $r\to+\infty$ and our condition, we have

$$\begin{aligned}\frac{1}{\sinh^C(ar_2)}\int_{B_{r_2}(x_0)}&\cosh(ar)\left[\tfrac{1}{2}|\omega|^2-\tfrac{1}{p}|i_{\partial/\partial r}\omega|^2\right]dv\\&\leq\frac{\cosh(ar_2)\int_{B_{r_2}(x_0)}\frac{1}{2}|\omega|^2\,dv}{\sinh^C(ar_2)}\\&=\frac{\int_{B_{r_2}(x_0)}\frac{1}{2}|\omega|^2\,dv}{e^{a(C-1)r_2}}\left[\frac{e^{ar_2}}{\sinh(ar_2)}\right]^{C-1}\coth(ar_2)\\&\to 0\quad\text{as } r_2\to+\infty.\end{aligned}\tag{3-9}$$

It follows from (3-9) and Theorem 3.1 that

$$\tfrac{1}{2}|\omega|^2 - \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2 = 0, \quad \text{i.e., } (\omega \odot \omega)(e_\lambda, e_\lambda) = 0. \tag{3-10}$$

Set $X = r\partial/\partial r$. It is easy to see from (3-10), (3-4), (3-3), (3-2) and (2-2) that

$$\begin{aligned}
&\int_{\partial B_r(x_0)} -\frac{p-1}{p} r|i_{\partial/\partial r}\omega|^2\, ds \\
&= \int_{\partial B_r(x_0)} \bigl[\tfrac{r}{2}|\omega|^2 - r|i_{\partial/\partial r}\omega|^2\bigr]\, ds \\
&= \int_{\partial B_r(x_0)} S_\omega\Bigl(X, \frac{\partial}{\partial r}\Bigr)\, ds \\
&= \tfrac{1}{p}|i_{\partial/\partial r}\omega|^2 \Bigl\{ \sum_\lambda r\,\mathrm{Hess}(r)(e_\lambda, e_\lambda) + \sum_\lambda r\,\mathrm{Hess}(r)\Bigl(Je_\lambda, Je_\lambda\Bigr) \\
&\qquad - (p-1) - pr\,\mathrm{Hess}(r)\Bigl(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}\Bigr) - (p-1)r\,\mathrm{Hess}(r)\Bigl(J\frac{\partial}{\partial r}, J\frac{\partial}{\partial r}\Bigr)\Bigr\} \\
&\geq \int_{B_r(x_0)} \tfrac{1}{p}[2(n-1)br\coth(br) - p + 1 - (p-1)ar\coth(ar)]|i_{\partial/\partial r}\omega|^2\, dv \\
&\geq \int_{B_r(x_0)} \tfrac{1}{p}\bigl\{(n-1)br\coth(br) - p + 1 \\
&\qquad + [(n-1)br - (p-1)ar]\coth(br)\bigr\}|i_{\partial/\partial r}\omega|^2\, dv \\
&\geq \int_{B_r(x_0)} \tfrac{1}{p}[(n-1)br\coth(br) - p + 1]|i_{\partial/\partial r}\omega|^2\, dv \\
&\geq \int_{B_r(x_0)} \tfrac{1}{p}[n-p]|i_{\partial/\partial r}\omega|^2\, dv.
\end{aligned} \tag{3-11}$$

Using our condition $(n-1)b - (p-1)a \geq 0$, we get $n - p \geq 0$, which, together with (3-11) and $x \coth x > 1$ for $x > 0$, yields $|i_{\partial/\partial r}\omega|^2 = 0$. □

4. Monotonicity formulae for Riemannian manifolds

Theorem 4.1. *Let M be an n-dimensional $(n \geq 3)$ complete Riemannian manifold with radial curvature K_r satisfying $-a^2 \leq K_r \leq -b^2 < 0$ with $a \geq b > 0$ and $(n-1)b - (2p-1)a \geq 0$. Let $\xi : E \to M$ be a smooth Riemannian vector bundle over (M, g). If $\omega \in A^p(\xi)$ satisfies the conservation law, that is,* $\operatorname{div} S_\omega = 0$, *then*

$$\begin{aligned}
&\frac{1}{\sinh^C(ar_1)} \int_{B_{r_1}(x_0)} \cosh(ar)\bigl[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\bigr]\, dv \\
&\qquad\qquad \leq \frac{1}{\sinh^C(ar_2)} \int_{B_{r_2}(x_0)} \cosh(ar)\bigl[\tfrac{1}{2}|\omega|^2 - \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\bigr]\, dv
\end{aligned}$$

for any $0 < r_1 < r_2$, *where* $C = [(n-2)b - (2p-2)a]/a$ *and* $B_r(x_0) \subseteq M$ *is a geodesic ball of radius* r *centered at* x_0 *in* M.

Proof. Set $X = \sinh(ar)\partial/\partial r$, where $\partial/\partial r$ denotes the unit radial vector. Obviously, the unit normal vector to $\partial B_r(x_0)$ is $\partial/\partial r$. Let $\{e_\lambda, \partial/\partial r\}$ be an orthonormal frame field on $B_r(x_0)$, where $\lambda = 1, \dots, n-1$. Then we have that

$$\nabla_{\partial/\partial r} X = a\cosh(ar)\frac{\partial}{\partial r} \quad \text{and} \quad \nabla_{e_\lambda} X = \sinh(ar)\sum_\mu h_{\lambda\mu} e_\mu, \tag{4-1}$$

where the $-h_{\lambda\mu}$ are the components of the second fundamental form of $\partial B_r(x_0)$ in $B_r(x_0)$.

On the other hand, we have

$$\mathrm{Hess}(r)(e_\lambda, e_\mu) = \langle e_\lambda, \nabla_{\partial/\partial r} e_\mu\rangle = \langle e_\lambda, h_{\mu\nu} e_\nu\rangle = h_{\lambda\mu}. \tag{4-2}$$

We can choose an orthonormal frame field $\{e_\lambda\}$ on $\partial B_r(x_0)$ such that $h_{\lambda\mu} = \delta_{\lambda\mu} h_{\lambda\lambda}$. It follows from (4-1), (4-2), (2-1) and (2-3) that

$$\begin{aligned}
\text{(4-3)}\ \langle S_\omega, \nabla\theta_X\rangle
&= \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\Big\{a\cosh(ar) + \sinh(ar)\sum_\lambda h_{\lambda\lambda} - 2pa\cosh(ar)\Big\}\\
&\quad + \sum_{\lambda=1}^{n-1}\tfrac{1}{2p}(\omega\odot\omega)(e_\lambda, e_\lambda)\Big\{a\cosh(ar) + \sinh(ar)\sum_\mu h_{\nu\nu} - 2p\sinh(ar)h_{\lambda\lambda}\Big\}\\
&\geq \big[\tfrac12|\omega|^2 - \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\big]aC\cosh(ar).
\end{aligned}$$

On the other hand, we have

$$S_\omega\Big(X, \frac{\partial}{\partial r}\Big) \leq \big[\tfrac12|\omega|^2 - \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\big]\sinh(ar). \tag{4-4}$$

Substituting (4-3) and (4-4) into (2-2), we obtain

$$\begin{aligned}
\text{(4-5)}\quad \int_{\partial B_r(x_0)} \big[\tfrac12|\omega|^2 - \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\big]\sinh(ar)\,ds \\
\geq \int_{B_r(x_0)} \big[\tfrac12|\omega|^2 - \tfrac{1}{2p}|i_{\partial/\partial r}\omega|^2\big]aC\cosh(ar)\,dv.
\end{aligned}$$

The proof is completed using (4-5) along with the same arguments used in the proof of Theorem 3.1. □

Similarly, using Theorem 4.1 along with the same arguments used in the proof of Theorem 3.2, we get the following vanishing theorem:

Theorem 4.2. *Let* M *be an* n*-dimensional* $(n \geq 3)$ *complete Riemannian manifold with radial curvature* K_r *satisfying* $-a^2 \leq K_r \leq -b^2 < 0$ *with* $a \geq b > 0$ *and*

$(n-1)b-(2p-1)a \geq 0$. *Let* $\xi : E \to M$ *be a smooth Riemannian vector bundle over* (M, g). *If* $\omega \in A^p(\xi)$ *satisfies the conservation law*, *that is*, $\operatorname{div} S_\omega = 0$, *and*

$$\frac{\int_{B_r(x_0)} |\omega|^2 \, dv}{e^{a(C-1)r}} \to 0 \quad \text{as } r \to +\infty,$$

where $C = [(n-2)b-(2p-2)a]/a$, *then* $\omega \equiv 0$.

5. Monotonicity formulae for pluriharmonic maps

Let M be a complex n-dimensional ($n \geq 3$) Kähler manifold. The complex structure of M gives a decomposition of TM^C into tangent vectors of types (1,0) and (0,1), i.e,

$$TM^C = T^{1,0}M \oplus T^{0,1}M.$$

Let $\phi : M \to N$ be a smooth map between Kähler manifolds. Then we have the following bundle maps:

$$\begin{aligned} \partial\phi &: T^{1,0}M \to T^{1,0}N, & \bar{\partial}\phi &: T^{0,1}M \to T^{1,0}N, \\ \partial\bar{\phi} &: T^{1,0}M \to T^{0,1}N, & \overline{\partial\phi} &: T^{0,1}M \to T^{0,1}N. \end{aligned}$$

A direct computation gives

$$|\bar{\partial}\phi|^2 = \frac{1}{4}\sum_{i=1}^{n}\bigl\{\langle d\phi(e_i), d\phi(e_i)\rangle + \langle d\phi(Je_i), d\phi(Je_i)\rangle - 2\langle d\phi(Je_i), J'd\phi(e_i)\rangle\bigr\} \tag{5-1}$$

and

$$|\partial\phi|^2 = \frac{1}{4}\sum_{i=1}^{n}\bigl\{\langle d\phi(e_i), d\phi(e_i)\rangle + \langle d\phi(Je_i), d\phi(Je_i)\rangle + 2\langle d\phi(Je_i), J'd\phi(e_i)\rangle\bigr\}, \tag{5-2}$$

where $\{e_i, Je_i\}$ is an orthonormal frame field on M, and J and J' are the complex structures of M and N, respectively.

We introduce two 1-forms $\sigma, \tau \in A^1(\phi^{-1}TN)$ given by

$$\sigma(X) = \frac{d\phi(X) + J'd\phi(JX)}{2} \quad \text{and} \quad \tau(X) = \frac{d\phi(X) - J'd\phi(JX)}{2}$$

for any $X \in TM$.

Lemma 5.1 [Dong 2013]. σ, τ *are J-invariant*, *and* $|\sigma|^2 = 2|\bar{\partial}\phi|^2$, $|\tau|^2 = 2|\partial\phi|^2$.

Siu [1980] introduced pluriharmonic maps. A smooth map $\phi : M \to N$ between Kähler manifolds is called pluriharmonic if $(\nabla d\phi)(X, \overline{Y}) = 0$, for all $X, Y \in T^{1,0}M$.

Lemma 5.2 [Dong 2013]. *If a map* $\phi : M \to N$ *between Kähler manifolds is pluriharmonic*, *then we have* $\operatorname{div} S_\sigma = \operatorname{div} S_\tau = 0$, *where* $S_\sigma = \frac{1}{2}|\sigma|^2 g - \sigma \odot \sigma$ *and* $S_\tau = \frac{1}{2}|\tau|^2 g - \tau \odot \tau$.

In this section, we will establish monotonicity formulae for pluriharmonic maps and harmonic maps.

Theorem 5.3. *Let M be a complex n-dimensional $(n \geq 2)$ complete Kähler manifold with radial curvature K_r satisfying $K_r \leq -b^2 < 0$ with $b > 0$. Suppose $\phi : M \to N$ is either a pluriharmonic map between Kähler manifolds or a harmonic map into a Kähler manifold with strongly seminegative curvature. Then*

$$\frac{\int_{B_{r_1}(x_0)} \cosh(br)|\bar{\partial}\phi|^2\,dv}{\sinh^{2(n-1)}(br_1)} \leq \frac{\int_{B_{r_2}(x_0)} \cosh(br)|\bar{\partial}\phi|^2\,dv}{\sinh^{2(n-1)}(br_2)}$$

and

$$\frac{\int_{B_{r_1}(x_0)} \cosh(br)|\partial\phi|^2\,dv}{\sinh^{2(n-1)}(br_1)} \leq \frac{\int_{B_{r_2}(x_0)} \cosh(br)|\partial\phi|^2\,dv}{\sinh^{2(n-1)}(br_2)}$$

for any $0 < r_1 < r_2$.

Proof. When $\phi : M \to N$ is a pluriharmonic map between Kähler manifolds, it follows from Lemmas 5.1 and 5.2 and (3-4), in which $p = 1$ and $\omega = \sigma$, that

$$\begin{aligned}
(5\text{-}3)\quad \langle S_\sigma, \nabla\theta_X\rangle &\geq |i_{\partial/\partial r}\sigma|^2 2(n-1)b^2\cosh(br) + (\sigma \odot \sigma)(e_\lambda, e_\lambda)2(n-1)b^2\cosh(br)\\
&= (n-1)b^2\cosh(br)|\sigma|^2 = 2(n-1)b^2\cosh(br)|\bar{\partial}\phi|^2.
\end{aligned}$$

On the other hand, we have

$$S_\sigma(X, v) \leq b\sinh(br)|\bar{\partial}\phi|^2. \tag{5-4}$$

Substituting (5-3) and (5-4) into (2-2) yields

$$\int_{\partial B_r} b\sinh(br)|\bar{\partial}\phi|^2\,ds \geq \int_{B_r} 2(n-1)b^2\cosh(br)|\bar{\partial}\phi|^2\,dv. \tag{5-5}$$

When $\phi : M \to N$ is a harmonic map into a Kähler manifold with strongly seminegative curvature, we have $\int_{B_r}(\operatorname{div} S_\sigma)(X)\,dv = \int_{B_r}(\operatorname{div} S_\tau)(X)\,dv \geq 0$ [Dong 2013]. Then ϕ also satisfies the integral formula (5-5).

The proof is completed using (5-5) and the same arguments used in the proof of Theorem 3.1. □

Similarly, using Theorem 5.3 along with the same arguments used in the proof of Theorem 3.2, we get the following theorem:

Theorem 5.4. *Let M be a complex n-dimensional $(n \geq 2)$ complete Kähler manifold with radial curvature K_r satisfying $K_r \leq -b^2 < 0$ with $b > 0$. Suppose $\phi : M \to N$ is either a pluriharmonic map between Kähler manifolds or a harmonic map into a*

Kähler manifold with strongly seminegative curvature. If

$$\frac{\int_{B_r(x_0)} |\bar{\partial}\phi|^2\,dv}{e^{(2n-3)br}} \to 0 \quad \left(\textit{resp.}\ \frac{\int_{B_r(x_0)} |\partial\phi|^2\,dv}{e^{(2n-3)br}} \to 0\right) \quad \textit{as } r \to +\infty$$

then ϕ is holomorphic (*resp. antiholomorphic*). *In particular*, *if*

$$\frac{\int_{B_r(x_0)} |d\phi|^2\,dv}{e^{(2n-3)br}} \to 0 \quad \textit{as } r \to +\infty,$$

then ϕ is constant.

Acknowledgement

The author thanks the referee for valuable comments and suggestions.

References

[Baird and Eells 1981] P. Baird and J. Eells, "A conservation law for harmonic maps", pp. 1–25 in *Geometry Symposium* (Utrecht, 1980), edited by E. Looijenga et al., Lecture Notes in Math. **894**, Springer, Berlin, 1981. MR 83i:58031 Zbl 0485.58008

[Dong 2013] Y. X. Dong, "Monotonicity formulae and holomorphicity of harmonic maps between Kähler manifolds", *Proc. Lond. Math. Soc.* (3) **107**:6 (2013), 1221–1260. MR 3149846 Zbl 1295.53070

[Dong and Lin 2014] Y. X. Dong and H. Z. Lin, "Monotonicity formulae, vanishing theorems and some geometric applications", *Q. J. Math.* **65**:2 (2014), 365–397. MR 3230367 Zbl 1303.53046

[Garber et al. 1979] W.-D. Garber, S. N. M. Ruijsenaars, E. Seiler, and D. Burns, "On finite action solutions of the nonlinear σ-model", *Ann. Physics* **119**:2 (1979), 305–325. MR 80e:81059 Zbl 0412.35089

[Greene and Wu 1979] R. E. Greene and H. Wu, *Function theory on manifolds which possess a pole*, Lecture Notes in Math. **699**, Springer, Berlin, 1979. MR 81a:53002 Zbl 0414.53043

[Sealey 1982] H. C. J. Sealey, "Some conditions ensuring the vanishing of harmonic differential forms with applications to harmonic maps and Yang–Mills theory", *Math. Proc. Cambridge Philos. Soc.* **91**:3 (1982), 441–452. MR 83i:58038 Zbl 0494.58002

[Sealey $\geq$ 2016] H. C. J. Sealey, "The stress-energy tensor and vanishing of L^2 harmonic forms", preprint.

[Siu 1980] Y. T. Siu, "The complex-analyticity of harmonic maps and the strong rigidity of compact Kähler manifolds", *Ann. of Math.* (2) **112**:1 (1980), 73–111. MR 81j:53061 Zbl 0517.53058

[Xin 1986] Y. L. Xin, "Differential forms, conservation law and monotonicity formula", *Sci. Sinica Ser. A* **29**:1 (1986), 40–50. MR 87m:58046 Zbl 0616.58001

Received January 27, 2015. Revised May 25, 2015.

JINTANG LI
SCHOOL OF MATHEMATICAL SCIENCES
XIAMEN UNIVERSITY
XIAMEN, FUJIAN 361005
CHINA
dli66@xmu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.137

JET SCHEMES OF THE CLOSURE OF NILPOTENT ORBITS

ANNE MOREAU AND RUPERT WEI TZE YU

We study in this paper the jet schemes of the closure of nilpotent orbits in a finite-dimensional complex reductive Lie algebra. For the nilpotent cone, which is the closure of the regular nilpotent orbit, all the jet schemes are irreducible. This was first observed by Eisenbud and Frenkel, and follows from a strong result of Mustață (2001). Using induction and restriction of "little" nilpotent orbits in reductive Lie algebras, we show that for a large number of nilpotent orbits, the jet schemes of their closures are reducible. As a consequence, we obtain certain geometric properties of these nilpotent orbit closures.

1. Introduction

Throughout this paper, the ground field will be the field $\mathbb{C}$ of complex numbers. We shall work with the Zariski topology, and by *variety* we mean a reduced, irreducible, and separated scheme of finite type over $\mathbb{C}$.

For X a scheme of finite type over $\mathbb{C}$ and $m \in \mathbb{N}$, we denote by $\mathscr{J}_m(X)$ the *m-th jet scheme of X*. It is a scheme of finite type over $\mathbb{C}$ whose $\mathbb{C}$-valued points are naturally in bijection with the $\mathbb{C}[t]/(t^{m+1})$-valued points of X; see, e.g., [Mustaţă 2001; Ein and Mustaţă 2009; Ishii 2011]. We have $\mathscr{J}_0(X) \simeq X$ and $\mathscr{J}_1(X) \simeq \mathrm{T}X$,

MSC2010: 14L30, 17B20, 17B08.
Keywords: nilpotent orbits, jet schemes, induction.

where TX is the total tangent bundle of X; see Section 2 for more details about generalities on jet schemes. From Nash [1995], it is known that the geometry of the jet schemes is deeply related to the singularities of X. As an illustration of that phenomenon, we have the following result, first conjectured by Eisenbud and Frenkel [Mustaţă 2001, Introduction], which will be important for us.

Theorem 1 [Mustaţă 2001, Theorem 1]. *Let X be an irreducible scheme of finite type over $\mathbb{C}$. If X is locally a complete intersection, then $\mathscr{J}_m(X)$ is irreducible for every $m \in \mathbb{N}$ if and only if X has rational singularities.*

According to Kolchin [1973], in contrast to the above theorem, the arc space $\mathscr{J}_\infty(X) = \varprojlim \mathscr{J}_m(X)$ of X is always irreducible when X is irreducible. In this paper, we shall be interested in the irreducibility of the jet schemes for the closure of nilpotent orbits in a complex reductive Lie algebra.

Let G be a complex connected reductive algebraic group, $\mathfrak{g}$ its Lie algebra, and $\mathcal{N}(\mathfrak{g})$ the nilpotent cone of $\mathfrak{g}$. It is the subscheme of $\mathfrak{g}$ associated to the augmentation ideal of $\mathbb{C}[\mathfrak{g}]^G$. It is a finite union of nilpotent G-orbits, and there is a unique nilpotent orbit of $\mathfrak{g}$, called the regular nilpotent orbit and denoted by $\mathcal{O}_{\mathrm{reg}}$, such that $\mathcal{N}(\mathfrak{g}) = \overline{\mathcal{O}_{\mathrm{reg}}}$.

According to Kostant [1963], the nilpotent cone is a complete intersection which is irreducible, reduced, and normal. Furthermore, by [Hesselink 1976], it has rational singularities. Hence by Theorem 1, the jet scheme $\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$ is irreducible for every $m \geqslant 1$. In fact, by [Mustaţă 2001, Propositions 1.4 and 1.5], $\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$ is also a complete intersection which is reduced for every $m \geqslant 1$.

In [op. cit., Appendix], Eisenbud and Frenkel used these results to extend certain results of Kostant [1963] in the setting of jet schemes. In particular, they proved that $\mathbb{C}[\mathscr{J}_m(\mathfrak{g})]$ is free over the ring $\mathbb{C}[\mathscr{J}_m(\mathfrak{g})]^{\mathscr{J}_m(G)}$ of $\mathscr{J}_m(G)$-invariants of $\mathbb{C}[\mathscr{J}_m(\mathfrak{g})]$.

Other nilpotent orbit closures do not share these geometric properties in general. Indeed, according to a recent result of Namikawa [2013], for a nonzero and nonregular nilpotent orbit $\mathcal{O}$, $\bar{\mathcal{O}}$ is not a complete intersection. In addition, $\bar{\mathcal{O}}$ does not always have rational singularities since it is not always normal; see, e.g., [Levasseur and Smith 1988; Kraft and Procesi 1982; Kraft 1989; Broer 1998; Sommers 2003].

Thus, it is quite natural to ask the following question.

Question 1. Let $\mathcal{O}$ be a nilpotent orbit of $\mathfrak{g}$, and $m \in \mathbb{N}^*$. Is $\mathscr{J}_m(\bar{\mathcal{O}})$ irreducible?

Answering Question 1 is the main purpose of this paper. For the zero orbit and the regular orbit, the answer is positive for every $m \in \mathbb{N}$. Outside these extreme cases, we will see that these jet schemes are rarely irreducible.

Motivations. Since $\bar{\mathcal{O}}$ is not a complete intersection for $\mathcal{O}$ nonzero and nonregular, Theorem 1 cannot be applied directly to answer Question 1. Very recently, Brion and Fu [2015] gave another proof of Namikawa's result, which is more uniform and slightly shorter. An interesting question, posed by Michel Brion to the first author, is whether jet schemes can be used to provide another proof of Namikawa's result.

Let us explain how we can tackle this problem using jet schemes. Let $\mathcal{O}$ be a nilpotent orbit of $\mathfrak{g}$. The singular locus of $\overline{\mathcal{O}}$ is exactly $\overline{\mathcal{O}} \setminus \mathcal{O}$. This follows from [Kaledin 2006, Lemma 1.4; Panyushev 1991]; see also [Henderson 2014, Section 2] for a recent review. Moreover, we have

$$\operatorname{codim}_{\overline{\mathcal{O}}}(\overline{\mathcal{O}} \setminus \mathcal{O}) \geqslant 2.$$

For the nilpotent cone, we have precisely $\operatorname{codim}_{\mathcal{N}(\mathfrak{g})}(\mathcal{N}(\mathfrak{g}) \setminus \mathcal{O}_{\mathrm{reg}}) = 2$, and the equality $\mathcal{N}(\mathfrak{g})_{\mathrm{reg}} = \mathcal{O}_{\mathrm{reg}}$ is a consequence of [Kostant 1963, Theorem 9] (thus the notation $\mathcal{O}_{\mathrm{reg}}$ does not bear any confusion).

So, if we assume that $\overline{\mathcal{O}}$ is a complete intersection, then $\overline{\mathcal{O}}$ is normal and so it has rational singularities by [Hinich 1991] or [Panyushev 1991]. Hence, in that event, Mustaţă's Theorem implies that $\mathscr{J}_m(\overline{\mathcal{O}})$ is irreducible for every $m \geqslant 1$. So if we can show that $\mathscr{J}_m(\overline{\mathcal{O}})$ is reducible for some $m \geqslant 1$, then we would obtain a contradiction.[1] The above was our original motivation to look into Question 1.

It may happen that a variety X is not a complete intersection, that X has rational singularities, and that nonetheless $\mathscr{J}_m(X)$ is irreducible for every $m \geqslant 1$. The cone over the Segre embedding

$$\mathbb{P}^1 \times \mathbb{P}^{n-1} \hookrightarrow \mathbb{P}^{2n-1}, \quad n \geqslant 2,$$

shows that this situation is possible; see [Mustaţă 2001, Example 4.7]. We do not know so far whether this situation may happen in the context of nilpotent orbit closures.

More generally, following Nash's philosophy, it would be interesting to understand what kind of properties on the singularities of $\overline{\mathcal{O}}$ we can deduce from the study of $\mathscr{J}_m(\overline{\mathcal{O}})$, $m \geqslant 1$. The fact that $\overline{\mathcal{O}}$ is not a complete intersection (with $\mathcal{O}$ nonzero and nonregular) whenever $\mathscr{J}_m(\overline{\mathcal{O}})$ is reducible for some $m \geq 1$ is one illustration of such a phenomenon.

Nilpotent orbit closures also form an interesting family of varieties, providing examples and counterexamples in the context of jet schemes. For instance, Examples 7.6 and 7.7 illustrate that the locally complete intersection hypothesis cannot be removed from Lemma 2.7(3), and Theorem 2.8(3). Another example is that the normality is not conserved when we pass to jet schemes. By Kostant, the nilpotent cone $\mathcal{N}(\mathfrak{g})$ is normal, and we show in Proposition 7.3 that $\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$, $m \geqslant 1$, is not normal for a simple Lie algebra $\mathfrak{g}$.

[1] There are other approaches that use jet schemes to show that $\overline{\mathcal{O}}$ is not a complete intersection; see Example 7.2.

Main results. Let us describe the main techniques used to study Question 1 and summarize the main results of the paper. To avoid technical details, we shall assume here that $\mathfrak{g}$ is simple.

Let X be an irreducible variety, and $m \in \mathbb{N}$. Then $\mathscr{J}_m(X)$ is irreducible if and only if

$$\pi_{X,m}^{-1}(X_{\mathrm{sing}}) \subset \overline{\pi_{X,m}^{-1}(X_{\mathrm{reg}})},$$

where $\pi_{X,m} : \mathscr{J}_m(X) \to X$ is the canonical projection from $\mathscr{J}_m(X)$ onto X (see Section 2), X_{reg} is the smooth part of X, and X_{sing} is its complement (see Lemma 2.7). This is our starting point.

For $\mathcal{O}$ a nilpotent orbit of $\mathfrak{g}$, the singular locus of $\overline{\mathcal{O}}$ is $\overline{\mathcal{O}} \setminus \mathcal{O}$ (see Section 3). The above criterion leads us to the following two conditions which will be central in our paper (see Definition 3.3).

Definition 1. Let $\mathcal{O}$ be a nilpotent orbit of $\mathfrak{g}$.

(1) We say that $\mathcal{O}$ *verifies* RC_1 if $\pi_{\overline{\mathcal{O}},1}^{-1}(0)$ is not contained in the closure of $\pi_{\overline{\mathcal{O}},1}^{-1}(\mathcal{O})$.

(2) Let $m \in \mathbb{N}^*$. We say that $\mathcal{O}$ *verifies* $\mathrm{RC}_2(m)$ if for some nilpotent orbit $\mathcal{O}'$ contained in $\overline{\mathcal{O}} \setminus \mathcal{O}$, we have $\dim \pi_{\overline{\mathcal{O}},m}^{-1}(\mathcal{O}') \geqslant \dim \pi_{\overline{\mathcal{O}},m}^{-1}(\mathcal{O})$.

Here the letters RC stand for "reducibility condition".

It follows readily (see Lemma 3.4) that if a nilpotent orbit $\mathcal{O}$ of $\mathfrak{g}$ verifies RC_1, then $\mathscr{J}_1(\overline{\mathcal{O}})$ is reducible. Similarly, if a nilpotent orbit $\mathcal{O}$ of $\mathfrak{g}$ verifies $\mathrm{RC}_2(m)$ for some $m \in \mathbb{N}^*$, then $\mathscr{J}_m(\overline{\mathcal{O}})$ is reducible.

We have a characterization for the condition RC_1 (see Proposition 3.6) which allows us, for example, to show that the nilpotent orbits of $\mathfrak{sl}_{2p}(\mathbb{C})$, with $p \geqslant 2$, associated with partitions of the form (2^p) verify RC_1 (see Example 3.7). Note that these orbits do not verify $\mathrm{RC}_2(1)$ (see again Example 3.7).

A nilpotent orbit $\mathcal{O}$ is called *little* if $0 < 2\dim\mathcal{O} \leqslant \dim\mathfrak{g}$ (see Definition 4.1). For example, the minimal nilpotent orbit of $\mathfrak{g}$ is little (see Corollary 4.3), and the nilpotent orbits of $\mathfrak{sl}_n(\mathbb{C})$ associated with partitions of the form $(2^p, 1^q)$, with $p, q \in \mathbb{N}^*$, are little (see Example 4.4). There are many other examples (see Section 4). Little nilpotent orbits verify both RC_1 and $\mathrm{RC}_2(m)$ for every $m \in \mathbb{N}^*$ (see Proposition 4.2), and they turn out to be useful to study the reducibility of jet schemes of many other orbits via "restriction" or "induction" of orbits.

Firstly, by "restriction" to some Levi subalgebras of $\mathfrak{g}$ (see Proposition 4.6), we can obtain from nilpotent orbits $\mathcal{O}$ which verify $0 < 2\dim\mathcal{O} < \dim\mathfrak{g}$ examples of nilpotent orbits which verify RC_1 (and that are not necessarily little); see Table 1. More precisely, as we shall see (in a slightly more general context) in Proposition 4.6, we have:

Proposition 1. *Let* $\mathfrak{l}$ *be a Levi subalgebra of* $\mathfrak{g}$ *with a center of dimension one, and such that* $\mathfrak{a} := [\mathfrak{l}, \mathfrak{l}]$ *is simple. Denote by* A *the connected subgroup of* G *whose*

Lie algebra is $\mathfrak{a}$. *Let* e *be a nilpotent element of* $\mathfrak{a}$ *and suppose that the following conditions are satisfied*:

(i) $\mathfrak{a}$ *contains a regular semisimple element of* $\mathfrak{g}$,

(ii) $2\dim G\cdot e < \dim\mathfrak{g}$.

Then $A\cdot e$ *verifies* RC_1.

Secondly, by "induction", we can reach from nilpotent orbits of reductive Lie subalgebras of $\mathfrak{g}$ many nilpotent orbits of $\mathfrak{g}$. Here, we consider induction in the sense of Lusztig and Spaltenstein [1979]. We refer the reader to Section 5 for the precise definition of a nilpotent orbit of $\mathfrak{g}$ induced from another one in some proper Levi subalgebra $\mathfrak{l}$ of $\mathfrak{g}$. Our next statement says that condition $\mathrm{RC}_2(m)$, for $m\in\mathbb{N}^*$, passes through induction.

Theorem 2. *Let* $\mathfrak{l}$ *be a Levi subalgebra of* $\mathfrak{g}$, $\mathcal{O}_{\mathfrak{l}}$ *a nilpotent orbit of* $\mathfrak{l}$ *and* $\mathcal{O}_{\mathfrak{g}}$ *the induced nilpotent orbit of* $\mathfrak{g}$ *from* $\mathcal{O}_{\mathfrak{l}}$. *If* $\mathcal{O}_{\mathfrak{l}}$ *verifies* $\mathrm{RC}_2(m)$ *for some* $m\in\mathbb{N}^*$, *then* $\mathcal{O}_{\mathfrak{g}}$ *also verifies* $\mathrm{RC}_2(m)$.

From this result, we are able to deal with a large number of nilpotent orbits. First of all, any nilpotent orbit induced from a nilpotent orbit that has a little factor verifies $\mathrm{RC}_2(m)$ for every $m\in\mathbb{N}^*$ (see Theorem 6.1). In particular, if $\mathfrak{g}$ is not of type A_1, $\mathrm{B}_2=\mathrm{C}_2$, or G_2, then the subregular nilpotent orbit $\mathcal{O}_{\mathrm{subreg}}$ of $\mathfrak{g}$ verifies $\mathrm{RC}_2(m)$ for every $m\in\mathbb{N}^*$ (see Corollary 6.2), and so $\mathscr{J}_m(\overline{\mathcal{O}_{\mathrm{subreg}}})$ is reducible for every $m\in\mathbb{N}^*$.

It turns out that many nilpotent orbits can be induced from a nilpotent orbit that has a little factor. This allows us to obtain the following result when $\mathfrak{g}$ is of type A (see Theorem 6.5).

Theorem 3. *Any nilpotent orbit of* $\mathfrak{sl}_n(\mathbb{C})$ *associated with a nonrectangular partition of* n *verifies* $\mathrm{RC}_2(m)$ *for every* $m\in\mathbb{N}^*$.

For the other simple Lie algebras of classical types, we have the following (see Theorem 6.7).

Theorem 4. *Let* $n\in\mathbb{N}^*$, $\boldsymbol{\lambda}=(\lambda_1,\dots,\lambda_t)$ *be a partition of* n, *and* $\lambda_{t+1}=0$. *Suppose that there exist* $1\leqslant k<\ell\leqslant t$ *such that* $\lambda_k\geqslant\lambda_{k+1}+2$ *and* $\lambda_\ell\geqslant\lambda_{\ell+1}+2$.

(1) *If* $\mathcal{O}$ *is a nilpotent orbit of* $\mathfrak{so}_n(\mathbb{C})$ *whose associated partition is* $\boldsymbol{\lambda}$, *then* $\mathcal{O}$ *verifies* $\mathrm{RC}_2(m)$ *for every* $m\in\mathbb{N}^*$.

(2) *If* n *is even and* $\mathcal{O}$ *is a nilpotent orbit of* $\mathfrak{sp}_n(\mathbb{C})$ *whose associated partition is* $\boldsymbol{\lambda}$, *then* $\mathcal{O}$ *verifies* $\mathrm{RC}_2(m)$ *for every* $m\in\mathbb{N}^*$.

While our result in the special linear case is exhaustive relative to induction, in the orthogonal and symplectic cases, other nilpotent orbits can be obtained by induction from a little orbit (see Theorem 6.7 and Remark 6.8). For a simple Lie algebra of exceptional type, we have a list of nilpotent orbits which can be induced from a little one (see Appendix C).

Organization of the paper. In Section 2, we state some basic properties on jet schemes with some proofs for the convenience of the reader.

In Section 3, we recall some standard properties of nilpotent orbit closures, and of their jet schemes. We introduce here the two sufficient conditions RC_1 and $\mathrm{RC}_2(m)$, $m \geqslant 1$, to study the reducibility of these jet schemes, and we state some first properties of these conditions.

Section 4 is devoted to little nilpotent orbits. We show that little nilpotent orbits verify both RC_1 and $\mathrm{RC}_2(m)$ for every $m \geqslant 1$, and we show how they can be used to prove condition RC_1 via the "restriction" of orbits (see Proposition 4.6).

In Section 5, we study the induction of nilpotent orbits the sense of Luzstig and Spaltenstein [1979]. The main result is that condition $\mathrm{RC}_2(m)$, for $m \geqslant 1$, passes through induction (see Theorem 5.6). We describe in Section 6 how to use Theorem 5.6 to obtain the reducibility of nilpotent orbit closures in simple Lie algebras according to their Dynkin type. The details of some of the conclusions are presented in Appendices B and C.

We present in Section 7 some applications of our results to geometric properties of nilpotent orbit closures. We also discuss in this section some open problems.

The standard notations relative to nilpotent orbits in classical simple Lie algebras are gathered together in Appendix A. Appendix B contains some numerical data for classical simple Lie algebras, and Appendix C summarizes our conclusions for simple Lie algebras of exceptional type.

2. Generalities on jet schemes

In this section, we present some general facts on jet schemes. Our main references on the topic are [Mustaţă 2001; Ein and Mustaţă 2009; Ishii 2011], and [de Fernex et al. 2013, Chapter 8].

Let X be a scheme of finite type over $\mathbb{C}$, and $m \in \mathbb{N}$.

Definition 2.1. An *m-jet of X* is a morphism

$$\operatorname{Spec} \mathbb{C}[t]/(t^{m+1}) \longrightarrow X.$$

The set of all m-jets of X carries the structure of a scheme $\mathscr{J}_m(X)$, called the *m-th jet scheme of X*. It is a scheme of finite type over $\mathbb{C}$ characterized by the following functorial property: for every scheme Z over $\mathbb{C}$, we have

$$\mathrm{Hom}(Z, \mathscr{J}_m(X)) = \mathrm{Hom}(Z \times_{\mathrm{Spec}\,\mathbb{C}} \mathrm{Spec}\,\mathbb{C}[t]/(t^{m+1}), X).$$

The $\mathbb{C}$-points of $\mathscr{J}_m(X)$ are thus the $\mathbb{C}[t]/(t^{m+1})$-points of X. From Definition 2.1, we have for example that $\mathscr{J}_0(X) \simeq X$ and that $\mathscr{J}_1(X) \simeq \mathrm{T}X$ where $\mathrm{T}X$ denotes the total tangent bundle of X.

For $p \in \{0, \dots, m\}$, the canonical projection $\mathbb{C}[t]/(t^{m+1}) \to \mathbb{C}[t]/(t^{p+1})$ induces a *truncation morphism*,

$$\pi_{X,m,p} : \mathscr{J}_m(X) \to \mathscr{J}_p(X).$$

We shall simply denote by $\pi_{X,m}$ the morphism $\pi_{X,m,0}$,

$$\pi_{X,m} : \mathscr{J}_m(X) \to \mathscr{J}_0(X) \simeq X.$$

Also, the canonical injection $\mathbb{C} \hookrightarrow \mathbb{C}[t]/(t^{m+1})$ induces a morphism $\iota_{X,m} : X \to \mathscr{J}_m(X)$, and we have $\pi_{X,m} \circ \iota_{X,m} = \mathrm{Id}_X$. Hence $\iota_{X,m}$ is injective and $\pi_{X,m}$ is surjective. We shall always view X as a subscheme of $\mathscr{J}_m(X)$.

If $f : X \to Y$ is a morphism of schemes, then we naturally obtain a morphism $f_m : \mathscr{J}_m(X) \to \mathscr{J}_m(Y)$ making the following diagram commutative:

$$\begin{tikzcd} \mathscr{J}_m(X) \arrow[r, "f_m"] \arrow[d, "\pi_{X,m}"'] & \mathscr{J}_m(Y) \arrow[d, "\pi_{Y,m}"] \\ X \arrow[r, "f"'] & Y \end{tikzcd}$$

Remark 2.2. In the case where X is affine, we have the following explicit description of $\mathscr{J}_m(X)$.

Let $n \in \mathbb{N}^*$ and $X \subset \mathbb{C}^n$ be the affine subscheme defined by an ideal $I = (f_1, \dots, f_r)$ of $\mathbb{C}[x_1, \dots, x_n]$. Thus

$$X = \mathrm{Spec}\,\mathbb{C}[x_1, \dots, x_n]/I.$$

For $k \in \{1, \dots, r\}$, we extend f_k as a map from $(\mathbb{C}[t]/(t^{m+1}))^n$ to $\mathbb{C}[t]/(t^{m+1})$ via base extension. Then giving a morphism $\gamma : \mathrm{Spec}\,\mathbb{C}[t]/(t^{m+1}) \to X$ is equivalent to giving a morphism $\gamma^* : \mathbb{C}[x_1, \dots, x_n]/I \to \mathbb{C}[t]/(t^{m+1})$, or to giving

$$\gamma^*(x_i) = \sum_{j=0}^{m} \gamma_i^{(j)} t^j, \quad 1 \leqslant i \leqslant n$$

such that for any $k \in \{1, \dots, r\}$,

$$f_k(\gamma^*(x_1), \dots, \gamma^*(x_n)) = 0 \quad \text{in } \mathbb{C}[t]/(t^{m+1}).$$

For $k \in \{1, \ldots, r\}$, there exist functions $f_k^{(0)}, \ldots, f_k^{(m)}$, which depend only on f, in the variables $\boldsymbol{\gamma} = (\gamma_i^{(j)})$, for $1 \leqslant i \leqslant n$ and $0 \leqslant j \leqslant m$, such that

$$f_k(\gamma^*(x_1), \ldots, \gamma^*(x_n)) = \sum_{j=0}^{m} f_k^{(j)}(\boldsymbol{\gamma})\, t^j. \tag{1}$$

The jet scheme $\mathscr{J}_m(X)$ is then the closed subscheme in $\mathbb{C}^{(m+1)n}$ defined by the ideal generated by the polynomials $f_k^{(j)}$, where $k \in \{1, \ldots, r\}$ and $j \in \{0, \ldots, m\}$. More precisely,

$$\mathscr{J}_m(X) \simeq \operatorname{Spec} \mathbb{C}[x_1^{(j)}, \ldots, x_n^{(j)} : j = 0, \ldots, m] \big/ (f_k^{(j)} : k = 1, \ldots, r;\ j = 0, \ldots, m).$$

In particular, if X is an n-dimensional vector space, then $\mathscr{J}_m(X) \simeq \mathbb{C}^{(m+1)n}$ and for $p \in \{0, \ldots, m\}$, the projection $\mathscr{J}_m(X) \to \mathscr{J}_p(X)$ corresponds to the projection onto the first $(p+1)n$ coordinates.

Example 2.3. Let us consider a concrete example. Let

$$X = \operatorname{Spec} \mathbb{C}[x, y, z]/(x^2 + yz) \subset \mathbb{C}^3,$$

and let us compute $\mathscr{J}_1(X)$ and $\mathscr{J}_2(X)$. We have

$$\begin{aligned}(x_0 + x_1 t + x_2 t^2)^2 &+ (y_0 + y_1 t + y_2 t^2)(z_0 + z_1 t + z_2 t^2)\\ &= x_0^2 + y_0 z_0 + (2x_0 x_1 + y_0 z_1 + y_1 z_0)t\\ &\quad + (2x_0 x_2 + x_1^2 + y_0 z_2 + y_2 z_0 + y_1 z_1)t^2 \mod t^3.\end{aligned}$$

Hence $\mathscr{J}_1(X)$ is the subscheme of

$$\mathscr{J}_1(\mathbb{C}^3) \simeq \mathbb{C}[x_0, y_0, z_0, x_1, y_1, z_1]$$

defined by the ideal

$$(x_0^2 + y_0 z_0,\ 2x_0 x_1 + y_0 z_1 + y_1 z_0),$$

and $\mathscr{J}_2(X)$ is the subscheme of

$$\mathscr{J}_2(\mathbb{C}^3) \simeq \mathbb{C}[x_0, y_0, z_0, x_1, y_1, z_1, x_2, y_2, z_3]$$

defined by the ideal

$$(x_0^2 + y_0 z_0,\ 2x_0 x_1 + y_0 z_1 + y_1 z_0,\ 2x_0 x_2 + x_1^2 + y_0 z_2 + y_1 z_1 + y_2 z_0).$$

We now list some basic properties that we need in the sequel. Their proofs can found in [Ein and Mustaţă 2009, Lemma 2.3, Remarks 2.8 and 2.10].

Lemma 2.4. (1) *For every open subset U of X, we have* $\mathscr{J}_m(U) = \pi_{X,m}^{-1}(U)$.

(2) *For every scheme Y, we have a canonical isomorphism*

$$\mathscr{J}_m(X \times Y) \simeq \mathscr{J}_m(X) \times \mathscr{J}_m(Y).$$

(3) *If G is a group scheme over $\mathbb{C}$, then $\mathscr{J}_m(G)$ is also a group scheme over $\mathbb{C}$. Moreover, if G acts on X, then $\mathscr{J}_m(G)$ acts on $\mathscr{J}_m(X)$.*

(4) *If $f : X \to Y$ is a smooth surjective morphism between schemes, then f_m is also smooth and surjective for every $m \in \mathbb{N}^*$.*

Geometric properties. It is known that the geometry of the jet schemes $\mathscr{J}_m(X)$, $m \geqslant 1$, is closely linked to that of X. More precisely, we can transport some geometric properties from $\mathscr{J}_m(X)$ to X.

The following proposition gives examples of such phenomena.

Proposition 2.5 [Mumford et al. 1994; Ishii 2011, Theorem 3.5]. *Let $m \in \mathbb{N}^*$. If $\mathscr{J}_m(X)$ is smooth (respectively, irreducible, reduced, normal, locally a complete intersection) for some m, then so is X.*

For smoothness, the converse is true, even with "every m" instead of "for some m". In fact, for smooth varieties, we have the following more precise statement.

Proposition 2.6 [Ein and Mustaţă 2009, Corollary 2.12]. *If X is a smooth variety of dimension n, then the truncation morphism $\pi_{m,p}$, for $p \in \{0, \ldots, m\}$, is a locally trivial projection with fiber isomorphic to $\mathbb{C}^{(m-p)n}$. In particular, $\mathscr{J}_m(X)$ is a smooth variety of dimension $(m+1)n$.*

For the other properties stated in Proposition 2.5, the converse is not true in general. We refer to [Ishii 2011, §3] for counterexamples. We shall encounter other counterexamples in this paper in the setting of nilpotent orbit closures. In this setting, our main purpose is to study the irreducibility of jet schemes. The following lemma gives a necessary and sufficient condition for the converse of Proposition 2.5 to hold for irreducibility.

We denote by X_{reg} the smooth part of X, and by X_{sing} its complement.

Lemma 2.7. *Assume that X is an irreducible reduced scheme of finite type over $\mathbb{C}$, and let $m \in \mathbb{N}^*$.*

(1) *$\overline{\pi_{X,m}^{-1}(X_{\text{reg}})}$ is an irreducible component of $\mathscr{J}_m(X)$.*

(2) *$\mathscr{J}_m(X)$ is irreducible if and only if $\pi_{X,m}^{-1}(X_{\text{sing}})$ is contained in $\overline{\pi_{X,m}^{-1}(X_{\text{reg}})}$.*

(3) *If X is a complete intersection, then $\mathscr{J}_m(X)$ is irreducible if and only if $\dim \pi_{X,m}^{-1}(X_{\text{sing}}) < \dim \pi_{X,m}^{-1}(X_{\text{reg}})$.*

In particular, if $\dim \pi_{X,m}^{-1}(X_{\text{sing}}) \geqslant \dim \overline{\pi_{X,m}^{-1}(X_{\text{reg}})}$, then $\mathscr{J}_m(X)$ is reducible.

Proof. Part (3) is proved in [Mustaţă 2001, Proposition 1.4], and parts (1) and (2) follow from its proof. More precisely, since X_{reg} is smooth and irreducible, $\overline{\pi_{X,m}^{-1}(X_{\text{reg}})}$ is an irreducible closed subset of $\mathscr{J}_m(X)$ of dimension $(m+1)\dim X$;

see Proposition 2.6. Then parts (1) and (2) follow easily from the fact that we have the decomposition

$$\mathscr{J}_m(X) = \pi_{X,m}^{-1}(X_{\mathrm{sing}}) \cup \overline{\pi_{X,m}^{-1}(X_{\mathrm{reg}})}$$

of closed subsets, and that $\pi_{X,m}^{-1}(X_{\mathrm{sing}}) \not\supset \overline{\pi_{X,m}^{-1}(X_{\mathrm{reg}})}$. □

There are also subtle connections between the geometry of $\mathscr{J}_m(X)$, $m \geqslant 1$, and the singularities of X which are important for us. In particular, according to [Mustaţă 2001, Theorem 0.1, Propositions 1.5 and 4.12], we have the theorem:

Theorem 2.8 (Mustață). *Let X be an irreducible variety over $\mathbb{C}$.*

(1) *If X is locally a complete intersection, then $\mathscr{J}_m(X)$ is irreducible for every $m \geqslant 1$ if and only if X has rational singularities.*

(2) *If X is locally a complete intersection and if $\mathscr{J}_m(X)$ is irreducible for some $m \geqslant 1$, then $\mathscr{J}_m(X)$ is also reduced.*

(3) *If X is locally a complete intersection, then $(\mathscr{J}_1(X))_{\mathrm{reg}} = \pi_{X,1}^{-1}(X_{\mathrm{reg}})$.*

Let us give an easy counterexample to the converse implication of Proposition 2.5 for normality. This example turns out to be a particular case of a more general situation that will be studied in Proposition 7.3.

Example 2.9. Let X be as in Example 2.3. Then X is a complete intersection and it is normal since the singular locus is reduced to $\{0\}$ which has codimension 2 in X. Next, it is not difficult to verify that $\mathscr{J}_1(X)$ is irreducible, reduced, and that it is a complete intersection. But $\mathscr{J}_1(X)$ is not normal. Indeed, by Theorem 2.8(3),

$$(\mathscr{J}_1(X))_{\mathrm{sing}} = \pi_{X,1}^{-1}(\{0\}) \simeq \{0\} \times \mathbb{C}^3.$$

Hence, the singular locus of $\mathscr{J}_1(X)$ has codimension 1 in $\mathscr{J}_1(X)$ since

$$\dim \mathscr{J}_1(X) = 2 \dim X = 4.$$

Group actions. Let G be a connected algebraic group, acting on a variety X, and $m \in \mathbb{N}$. Denote by

$$\rho : G \times X \to X, \quad (g, x) \mapsto g \cdot x$$

the corresponding action. As stated in Lemma 2.4, the morphism

$$\rho_m : \mathscr{J}_m(G \times X) \simeq \mathscr{J}_m(G) \times \mathscr{J}_m(X) \to \mathscr{J}_m(X)$$

defines an action of $\mathscr{J}_m(G)$ on $\mathscr{J}_m(X)$.

Recall that we embed X into $\mathscr{J}_m(X)$ through $\iota_{X,m}$. For $x \in X$, let us denote by G^x the stabilizer of x in G, and for $m \in \mathbb{N}$, we denote by $\mathscr{J}_m(G)^x$ its stabilizer in $\mathscr{J}_m(G)$. The following results are probably standard. Since we have not found any reference, we shall include their proofs.

Lemma 2.10. *Let $x \in X$. Then,*

$$\mathscr{J}_m(G)\cdot x = \mathscr{J}_m(G\cdot x), \quad \mathscr{J}_m(G^x) = \mathscr{J}_m(G)^x, \quad \pi^{-1}_{\overline{G\cdot x},m}(G\cdot x) = \mathscr{J}_m(G\cdot x).$$

Proof. The morphism $G\times\{x\}\to G\cdot x$, $(g,x)\mapsto g\cdot x$ is a submersion at all points of $G\times\{x\}$. Hence, according to [Hartshorne 1977, Chapter III, Proposition 10.4], it is a smooth morphism onto $G\cdot x$. So, by Lemma 2.4(4), the induced morphism $\mathscr{J}_m(G)\times\{x\}\to\mathscr{J}_m(G\cdot x)$ is also smooth and surjective. Consequently, we have the first equality $\mathscr{J}_m(G)\cdot x = \mathscr{J}_m(G\cdot x)$.

By applying the first equality to the algebraic group G^x, we get $\mathscr{J}_m(G^x)\cdot x = \mathscr{J}_m(G^x\cdot x)$, whence the inclusion $\mathscr{J}_m(G^x)\subset\mathscr{J}_m(G)^x$.

Conversely, let $\gamma : \operatorname{Spec}\mathbb{C}[t]/(t^{m+1})\to G$ be an element of $\mathscr{J}_m(G)^x$. Then $\rho_m(\gamma,x)=x$; hence, viewing x as a morphism $x:\operatorname{Spec}\mathbb{C}[t]/(t^{m+1})\to X$,

$$\rho(\gamma(\tau),x(\tau)) = x(\tau),$$

where τ is the unique element of $\operatorname{Spec}\mathbb{C}[t]/(t^{m+1})$. Thus $\gamma(\tau)\in G^x$ and $x(\tau)=x$. So we have $\gamma\in\mathscr{J}_m(G^x)$, and the second equality follows.

The third equality is a direct consequence of Lemma 2.4(1) since $G\cdot x$ is open in its closure. □

Let $\mathfrak{g}$ be the Lie algebra of G. We consider now the adjoint action of G on $\mathfrak{g}$. For the results we present here, we refer the reader to [Mustaţă 2001, Appendix]. Denote by

$$\mathfrak{g}_m := \mathfrak{g}\otimes_{\mathbb{C}}\mathbb{C}[t]/(t^{m+1})$$

the *generalized Takiff Lie algebra* whose Lie bracket is given by

$$[u\otimes x(t), v\otimes y(t)] = [u,v]\otimes x(t)y(t), \quad u,v\in\mathfrak{g},\ x(t),y(t)\in\mathbb{C}[t]/(t^{m+1}).$$

As Lie algebras, we have

$$\mathscr{J}_m(\mathfrak{g})\simeq\mathfrak{g}_m\simeq\operatorname{Lie}(\mathscr{J}_m(G)).$$

In the sequel, when there is no confusion, we shall use the notations $\mathfrak{g}_m$ and G_m for $\mathscr{J}_m(\mathfrak{g})$ and $\mathscr{J}_m(G)$ respectively. If $\mathfrak{a}$ is a Lie subalgebra of $\mathfrak{g}$, then $\mathscr{J}_m(\mathfrak{a})\simeq\mathfrak{a}_m$ is a Lie subalgebra of $\mathfrak{g}_m$. In particular, for $x\in\mathfrak{g}$, we have $(\mathfrak{g}_m)^x = (\mathfrak{g}^x)_m$, where for any subalgebra $\mathfrak{m}$ of $\mathfrak{g}_k$, with $k\geqslant 0$, $\mathfrak{m}^x$ stands for the centralizer of x in $\mathfrak{m}$.

We can identify $\mathfrak{g}_m$ with $\mathfrak{g}^{m+1}\simeq\mathscr{J}_m(\mathfrak{g})$ as a variety through the map

$$\mathfrak{g}^{m+1}\to\mathfrak{g}_m, \quad (x_0,x_1,\dots,x_m)\mapsto x_0+x_1\otimes t+\cdots+x_m\otimes t^m.$$

Let G_m be a connected algebraic group whose Lie algebra is $\mathfrak{g}_m$. Let $\mathbb{C}[\mathfrak{g}_m]$ be the coordinate ring of $\mathfrak{g}_m$, and let $\mathbb{C}[\mathfrak{g}_m]^{G_m}$ be the subring of G_m-invariants. We conclude in this section with the following result.

Lemma 2.11. *For $f \in \mathbb{C}[\mathfrak{g}]^G$, the polynomials $f^{(0)}, \dots, f^{(m)}$, defined in Remark 2.2, are elements of $\mathbb{C}[\mathfrak{g}_m]^{G_m}$.*

Proof. This is straightforward from the explicit description of the polynomials $f^{(0)}, \dots, f^{(m)}$ given in Remark 2.2. $\square$

3. Nilpotent orbit closures

From now on, we let G to be a connected reductive algebraic group over $\mathbb{C}$, $\mathfrak{g}$ its Lie algebra, and $\mathcal{N}(\mathfrak{g})$ the nilpotent cone of $\mathfrak{g}$. Recall that $\mathcal{N}(\mathfrak{g})$ is the subscheme of $\mathfrak{g}$ defined by the augmentation ideal of $\mathbb{C}[\mathfrak{g}]^G$, and that $\mathcal{N}(\mathfrak{g}) = \overline{\mathcal{O}_{\mathrm{reg}}}$ where $\mathcal{O}_{\mathrm{reg}}$ is the regular nilpotent orbit of $\mathfrak{g}$ (see the introduction). As mentioned there, we are interested in this paper in the irreducibility of jet schemes of the closure of nilpotent orbits.

Recall that for an arbitrary nilpotent orbit $\mathcal{O}$ of $\mathfrak{g}$, the singular locus of $\overline{\mathcal{O}}$ is $\overline{\mathcal{O}} \setminus \mathcal{O}$ and that $\operatorname{codim}_{\overline{\mathcal{O}}}(\overline{\mathcal{O}} \setminus \mathcal{O}) \geqslant 2$ (see Section 1).

Definition 3.1. Let $\mathcal{O}$ be a nonzero nilpotent orbit of $\mathfrak{g}$. Define $\mathfrak{g}_{\mathcal{O}}$ to be the smallest semisimple ideal of $\mathfrak{g}$ containing $\mathcal{O}$.

More precisely, if $\mathfrak{g} \simeq \mathfrak{z}(\mathfrak{g}) \times \mathfrak{s}_1 \times \cdots \times \mathfrak{s}_m$, with $\mathfrak{z}(\mathfrak{g})$ the center of $\mathfrak{g}$ and $\mathfrak{s}_1, \dots, \mathfrak{s}_m$ the simple factors of $\mathfrak{g}$, then $\mathcal{O} = \mathcal{O}_1 \times \cdots \times \mathcal{O}_m$, with $\mathcal{O}_i$ a nilpotent orbit of $\mathfrak{s}_i$ for $i = 1, \dots, m$, and

$$\mathfrak{g}_{\mathcal{O}} = \mathfrak{s}_{i_1} \times \cdots \times \mathfrak{s}_{i_k},$$

where $\{i_1, \dots, i_k\}$ is the set of integers $j \in \{1, \dots, m\}$ such that $\mathcal{O}_j$ is nonzero. In particular, if $\mathcal{O}$ is zero, then $\mathfrak{g}_{\mathcal{O}} = 0$, and if $\mathcal{O}$ is nonzero and $\mathfrak{g}$ is simple, then $\mathfrak{g}_{\mathcal{O}} = \mathfrak{g}$.

For $\mathcal{O}$ a nilpotent orbit of $\mathfrak{g}$, we denote by $\mathcal{I}_{\overline{\mathcal{O}}}$ the defining ideal of $\overline{\mathcal{O}}$ in $\mathfrak{g}_{\mathcal{O}}$. Thus,

$$\overline{\mathcal{O}} = \operatorname{Spec} \mathbb{C}[\mathfrak{g}_{\mathcal{O}}]/\mathcal{I}_{\overline{\mathcal{O}}}.$$

Recall that $\overline{\mathcal{O}}$ is conical, so $\mathcal{I}_{\overline{\mathcal{O}}}$ is a homogeneous ideal.

Lemma 3.2. *Let $\mathcal{O}$ be a nonzero nilpotent orbit of $\mathfrak{g}$. If $f_1, \dots, f_s$ are homogeneous generators of $\mathcal{I}_{\overline{\mathcal{O}}}$, then the minimum degree of the f_i is exactly 2.*

Proof. By the above discussion, $\mathcal{O}$ is a product of nilpotent orbits. We may therefore assume that $\mathfrak{g} = \mathfrak{g}_{\mathcal{O}}$ is simple.

Assume that for some $i \in \{1, \dots, s\}$, $\deg f_i = 1$. A contradiction is expected. Let $\mathcal{V}$ be the intersection of all the hyperplanes $\mathcal{H}_g$, $g \in G$, defined by the linear form

$$g \cdot f_i : \mathfrak{g} \to \mathbb{C}, \quad x \mapsto f_i(g^{-1}(x)).$$

Since $\overline{\mathcal{O}}$ is G-invariant and is contained in the zero locus of f_i, $\overline{\mathcal{O}}$ is contained in $\mathcal{V}$. Thus $\mathcal{V}$ is a nonzero G-invariant subspace of $\mathfrak{g}$ which is different from $\mathfrak{g}$ (because $\mathcal{V}$ is contained in the hyperplane $\mathcal{H}_{1_G}$), whence the contradiction since $\mathfrak{g}$ is simple.

The Casimir element, $x \mapsto \langle x, x \rangle$ with $\langle \cdot, \cdot \rangle$ the Killing form of $\mathfrak{g}$, vanishes on the nilpotent cone of $\mathfrak{g}$. Hence it is contained in $\mathcal{I}_{\overline{\mathcal{O}}}$. Since it has degree 2, the minimal degree of the f_i is exactly 2. □

To determine the reducibility of $\mathscr{J}_m(\overline{\mathcal{O}})$ for $\mathcal{O}$ a (nonzero) nilpotent orbit of $\mathfrak{g}$, we introduce the two sufficient conditions below.

Definition 3.3. Let $\mathcal{O}$ be a nilpotent orbit of $\mathfrak{g}$.

(1) We say that $\mathcal{O}$ *verifies* RC_1 if $\pi_{\overline{\mathcal{O}},1}^{-1}(0)$ is not contained in the closure of $\pi_{\overline{\mathcal{O}},1}^{-1}(\mathcal{O})$.

(2) Let $m \in \mathbb{N}^*$. We say that $\mathcal{O}$ *verifies* $\mathrm{RC}_2(m)$ if for some nilpotent orbit $\mathcal{O}'$ contained in $\overline{\mathcal{O}} \setminus \mathcal{O}$, we have $\dim \pi_{\overline{\mathcal{O}},m}^{-1}(\mathcal{O}') \geqslant \dim \pi_{\overline{\mathcal{O}},m}^{-1}(\mathcal{O}) = (m+1) \dim \mathcal{O}$.

The following Lemma directly results from Lemma 2.7(2).

Lemma 3.4. *Let $\mathcal{O}$ be a nilpotent orbit of* $\mathfrak{g}$.

(1) *If $\mathcal{O}$ verifies* RC_1, *then $\mathscr{J}_1(\overline{\mathcal{O}})$ is reducible.*

(2) *If $\mathcal{O}$ verifies* $\mathrm{RC}_2(m)$ *for some $m \in \mathbb{N}^*$, then $\mathscr{J}_m(\overline{\mathcal{O}})$ is reducible.*

The zero nilpotent orbit verifies neither RC_1 nor $\mathrm{RC}_2(m)$ for $m \in \mathbb{N}^*$. Since $\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$ is irreducible for every $m \in \mathbb{N}^*$ (see Section 1), the same goes for the regular nilpotent orbit according to Lemma 3.4.

In view of the conditions above, let us study the zero fiber of $\pi_{\overline{\mathcal{O}},1} : \mathscr{J}_1(\overline{\mathcal{O}}) \to \overline{\mathcal{O}}$. As in Section 2, we identify $(\mathfrak{g}_{\mathcal{O}})_m$ with $(\mathfrak{g}_{\mathcal{O}})^{m+1} = \underbrace{\mathfrak{g}_{\mathcal{O}} \times \cdots \times \mathfrak{g}_{\mathcal{O}}}_{(m+1)\text{ factors}}$.

Lemma 3.5. *Let $\mathcal{O}$ be a nonzero nilpotent orbit of* $\mathfrak{g}$, *and $m \in \mathbb{N}^*$.*

(1) *We have* $\pi_{\overline{\mathcal{O}},1}^{-1}(0) \simeq \{0\} \times \mathfrak{g}_{\mathcal{O}}$. *In particular*, $\dim \pi_{\overline{\mathcal{O}},1}^{-1}(0) = \dim \mathfrak{g}_{\mathcal{O}}$.

(2) *If* $m \geqslant 2$, *then* $\dim \pi_{\overline{\mathcal{O}},m}^{-1}(0) \geqslant \dim \mathscr{J}_{m-2}(\overline{\mathcal{O}}) + \dim \mathfrak{g}_{\mathcal{O}} \geqslant m \dim \mathcal{O} + \mathrm{codim}_{\mathfrak{g}_{\mathcal{O}}}(\mathcal{O})$.

Part (2) of Lemma 3.5 remains valid for an affine variety in $\mathbb{C}^n$ defined by homogeneous polynomials of degree at least 2. The special case where all the generators have the same degree is treated in [Yuen 2007, Proposition 5.2].

Proof. Clearly we may assume that $\mathfrak{g}_{\mathcal{O}} = \mathfrak{g}$. Let $f_1, \ldots, f_r$ be homogeneous generators of $\mathcal{I}_{\overline{\mathcal{O}}}$ that we order so that $2 = d_1 \leqslant \ldots \leqslant d_r$, with $d_i = \deg f_i$ for any $i = 1, \ldots, r$ (see Lemma 3.2).

(1) Through our identification, we can write

$$\pi_{\overline{\mathcal{O}},1}^{-1}(0) \simeq \{0\} \times \{x \in \mathfrak{g} \mid f_i(tx) = 0 \bmod t^2 \text{ for any } i = 1, \ldots, r\},$$

whence the statement since for any $x \in \mathfrak{g}$ and $i \in \{1, \ldots, r\}$, we have $f_i(tx) = t^{d_i} f_i(x)$ and $d_i \geqslant 2$.

(2) Assume that $m \geqslant 2$. Let $(x_1, x_2, \dots, x_{m-1})$ be an element of $\mathscr{J}_{m-2}(\overline{\mathcal{O}})$, and let $x_m \in \mathfrak{g}$. Then for any $i \in \{1, \dots, r\}$, we get

$$f_i(tx_1 + t^2 x_2 + \cdots + t^m x_m) = f_i(tx_1 + t^2 x_2 + \cdots + t^{m-1} x_{m-1}) \mod t^{m+1}$$

since f_i is homogeneous of degree at least 2. Hence,

$$f_i(tx_1 + t^2 x_2 + \cdots + t^m x_m) = t^{d_i} f_i(x_1 + t x_2 + \cdots + t^{m-2} x_{m-1}) \mod t^{m+1}.$$

But $f_i(x_1 + tx_2 + \cdots + t^{m-2} x_{m-1}) = 0 \bmod t^{m-1}$ because $(x_1, x_2, \dots, x_{m-1}) \in \mathscr{J}_{m-2}(\overline{\mathcal{O}})$. So,

$$t^{d_i} f_i(x_1 + t x_2 + \cdots + t^{m-2} x_{m-1}) = 0 \mod t^{m+1}$$

since $d_i \geqslant 2$. In other words, $(0, x_1, x_2, \dots, x_m)$ is an element of $\pi^{-1}_{\overline{\mathcal{O}},m}(0)$.

Thus we obtain an embedding from $\mathscr{J}_{m-2}(\overline{\mathcal{O}}) \times \mathfrak{g}$ into $\pi^{-1}_{\overline{\mathcal{O}},m}(0)$ given by

$$\mathscr{J}_{m-2}(\overline{\mathcal{O}}) \times \mathfrak{g} \to \pi^{-1}_{\overline{\mathcal{O}},m}(0), \quad ((x_1, x_2, \dots, x_{m-1}), x_m) \mapsto (0, x_1, x_2, \dots, x_{m-1}, x_m).$$

The assertions follows. $\square$

Let $\mathcal{O}$ be a nonzero nilpotent orbit of $\mathfrak{g}$, and fix $e \in \mathcal{O}$. The tangent space at e to $\overline{\mathcal{O}}$ is the space $[e, \mathfrak{g}]$. Consider the morphism

$$\eta_{\mathfrak{g},e} : G \times [e, \mathfrak{g}] \to \mathfrak{g}, \quad (g, x) \mapsto g(x).$$

Proposition 3.6. *The nonzero nilpotent orbit $\mathcal{O}$ verifies* RC_1 *if and only if the closure of the image of $\eta_{\mathfrak{g},e}$ is strictly contained in $\mathfrak{g}_{\mathcal{O}}$.*

Proof. Since $[e, \mathfrak{g}] = [e, \mathfrak{g}_{\mathcal{O}}]$, we may assume that $\mathfrak{g} = \mathfrak{g}_{\mathcal{O}}$. Thus, by the definition of condition RC_1, we have to show that $\pi^{-1}_{\overline{\mathcal{O}},1}(0)$ is contained in

$$\overline{\pi^{-1}_{\overline{\mathcal{O}},1}(\mathcal{O})}$$

if and only if $\eta_{\mathfrak{g},e}$ is dominant, i.e., $\overline{G \cdot [e, \mathfrak{g}]} = \mathfrak{g}$.

By Lemma 3.5(1), we have $\pi^{-1}_{\overline{\mathcal{O}},1}(0) \simeq \{0\} \times \mathfrak{g}$. On the other hand,

$$\pi^{-1}_{\overline{\mathcal{O}},1}(\mathcal{O}) = G \cdot (\{e\} \times [e, \mathfrak{g}]).$$

So, if $\pi^{-1}_{\overline{\mathcal{O}},1}(0) \subset \overline{\pi^{-1}_{\overline{\mathcal{O}},1}(\mathcal{O})}$, then

$$\{0\} \times \mathfrak{g} \subset \overline{G \cdot (\{e\} \times [e, \mathfrak{g}])} \subset \overline{G \cdot e} \times \overline{G \cdot [e, \mathfrak{g}]},$$

whence the inclusion $\mathfrak{g} \subset \overline{G \cdot [e, \mathfrak{g}]}$, and $\eta_{\mathfrak{g},e}$ is dominant.

For the other direction, observe that

$$\overline{\pi^{-1}_{\overline{\mathcal{O}},1}(\mathcal{O})}$$

is a closed bicone of $\mathfrak{g} \times \mathfrak{g}$ since $\mathcal{O}$ and $\overline{\mathcal{O}}$ are both subcones of $\mathfrak{g}$. Here, by *bicone*, we mean a subset of $\mathfrak{g} \times \mathfrak{g}$ stable under the natural $(\mathbb{C}^* \times \mathbb{C}^*)$-action on $\mathfrak{g} \times \mathfrak{g}$. Therefore, if $\overline{G \cdot [e, \mathfrak{g}]} = \mathfrak{g}$, then

$$\overline{G \cdot (\{e\} \times [e, \mathfrak{g}])} = \overline{G \cdot (\mathbb{C}^* e \times [e, \mathfrak{g}])} \supset \{0\} \times \overline{G \cdot [e, \mathfrak{g}]} = \{0\} \times \mathfrak{g},$$

whence $\pi_{\overline{\mathcal{O}},1}^{-1}(0) \subset \overline{\pi_{\overline{\mathcal{O}},1}^{-1}(\mathcal{O})}$. □

Example 3.7. Let $p \in \mathbb{N}^*$ with $p \geqslant 2$, and $\mathfrak{g} = \mathfrak{sl}_{2p}(\mathbb{C})$. In the notation of Appendix A, we claim that the nilpotent orbit $\mathcal{O}_{(2^p)}$ of $\mathfrak{g}$ associated with the partition (2^p) verifies RC_1. According to Proposition 3.6, it suffices to prove that for the element

$$e := \begin{pmatrix} 0 & I_p \\ 0 & 0 \end{pmatrix} \in \mathcal{O}_{(2^p)},$$

the morphism $\eta_{\mathfrak{g},e}$ is not dominant. We readily verify that $[e, \mathfrak{g}]$ consists of matrices of the form

$$\begin{pmatrix} A & C \\ 0 & -A \end{pmatrix}$$

with A and C of size p. In particular, $[e, \mathfrak{g}]$ is contained in the closed subset $\mathcal{Z}$ of $\mathfrak{g}$ consisting of the matrices whose characteristic polynomial is even. Since $\overline{G([e, \mathfrak{g}])}$ and $\mathcal{Z}$ are both closed G-stable subsets of $\mathfrak{g}$, we get

$$\overline{G([e, \mathfrak{g}])} \subset \mathcal{Z}.$$

The diagonal matrix $\mathrm{diag}(1, \ldots, 1, -2p+1)$ is in $\mathfrak{g}$ but does not lie in $\mathcal{Z}$ for $p \geqslant 2$. Hence, $\mathcal{Z}$ is strictly contained in $\mathfrak{g}$, and $\eta_{\mathfrak{g},e}$ is not dominant. Thus $\mathcal{O}_{(2^p)}$ verifies RC_1.

According to Lemma 3.4(1), $\mathscr{J}_1(\overline{\mathcal{O}_{(2^p)}})$ is reducible. In fact, we can be more precise. By [Weyman 2002, Theorem 1] (see also [Weyman 1989] or [Weyman 2003, Proposition 8.2.15]), the defining ideal of $\overline{\mathcal{O}_{(2^p)}}$ is generated by the entries of the matrix X^2 as functions of $X \in \mathfrak{sl}_{2p}(\mathbb{C})$. It follows that $\mathscr{J}_1(\overline{\mathcal{O}_{(2^p)}})$ can be identified with the scheme of pairs $(X_0, X_1) \in \mathfrak{sl}_{2p}(\mathbb{C}) \times \mathfrak{sl}_{2p}(\mathbb{C})$ defined by the equations $X_0^2 = 0$ and $X_0X_1 + X_1X_0 = 0$. Using this identification, we obtain from direct computations that

- $\mathscr{J}_1(\overline{\mathcal{O}_{(2^p)}})$ has exactly one irreducible component of dimension $4p^2 = 2\dim \mathcal{O}_{(2^p)}$,
- all the other irreducible components have dimension $4p^2 - 1$, and $\pi_{\overline{\mathcal{O}_{(2^p)}},1}^{-1}(0)$ is one of them.

Remark 3.8. Assume that $\mathfrak{g} = \mathfrak{g}_{\mathcal{O}}$. A nilpotent element e is distinguished if its centralizer is contained in the nilpotent cone. In particular, if e is distinguished, then the centralizer of an $\mathfrak{sl}_2$-triple (e, h, f) in $\mathfrak{g}$ is zero, and the theory of representations of $\mathfrak{sl}_2$ shows that $[e, \mathfrak{g}]$ contains $\mathfrak{g}^h$, and hence contains a Cartan subalgebra of $\mathfrak{g}$. Consequently, $G \cdot e$ does not verify RC_1.

Remark 3.9. Assume that $\mathfrak{g} = \mathfrak{g}_{\mathcal{O}}$. Since $G \times [e, \mathfrak{g}]$ and $\mathfrak{g}$ are irreducible varieties, $\eta_{\mathfrak{g},e}$ is dominant if and only if there is a nonempty open set U consisting of points $a \in G \times [e, \mathfrak{g}]$ such that $(\mathrm{d}\eta_{\mathfrak{g},e})_a$ is surjective. The differential of $\eta_{\mathfrak{g},e}$ at $a = (g, [e, x])$, with $(g, x) \in G \times \mathfrak{g}$ is given by

$$\mathfrak{g} \times [e, \mathfrak{g}] \to \mathfrak{g}, \quad (v, [e, w]) \mapsto [v, [e, x]] + g([e, w]).$$

Let us endow $G \times [e, \mathfrak{g}]$ with the action of G by left multiplication on the first factor. Since $\eta_{\mathfrak{g},e}$ is G-equivariant, we may assume that a is of the form $a = (1_G, [e, x])$ with $x \in \mathfrak{g}$. Then $(\mathrm{d}\eta_{\mathfrak{g},e})_a$ is surjective if and only if $[\mathfrak{g}, [e, x]] + [e, \mathfrak{g}] = \mathfrak{g}$.

Consequently, $\eta_{\mathfrak{g},e}$ is dominant if and only if there exists $x \in \mathfrak{g}$ such that $[\mathfrak{g}, [e, x]] + [e, \mathfrak{g}] = \mathfrak{g}$. This allows us to affirm in some cases that $\eta_{\mathfrak{g},e}$ is dominant. For example, for e in the nondistinguished nilpotent orbit $\mathcal{O}_{(3^2)}$ of $\mathfrak{sl}_6(\mathbb{C})$, the map $\eta_{\mathfrak{g},e}$ is dominant.

4. Little nilpotent orbits

We introduce in this section a family of nonzero nilpotent orbits which verify both RC_1 and $\mathrm{RC}_2(m)$ for every $m \in \mathbb{N}^*$. This family turns out to be useful to study the reducibility of jet schemes of many other orbits.

Lemma 3.5 leads us to the following definition.

Definition 4.1. Let $\mathcal{O}$ be a nilpotent orbit of $\mathfrak{g}$ and let $\mathfrak{g}_{\mathcal{O}}$ be as in Definition 3.1. We say that $\mathcal{O}$ is *little* if $0 < 2 \dim \mathcal{O} \leqslant \dim \mathfrak{g}_{\mathcal{O}}$.

In particular, neither the zero orbit nor the regular nilpotent orbit is little.

Proposition 4.2. *If $\mathcal{O}$ is a little nilpotent orbit of $\mathfrak{g}$, then $\mathcal{O}$ verifies RC_1 and $\mathrm{RC}_2(m)$ for every $m \in \mathbb{N}^*$.*

Proof. Let $\mathcal{O}$ be a little nilpotent orbit of $\mathfrak{g}$. As in the preceding proofs, we may assume that $\mathfrak{g} = \mathfrak{g}_{\mathcal{O}}$. According to Lemma 3.5(1), $\dim \pi^{-1}_{\overline{\mathcal{O}},1}(0) = \dim \mathfrak{g}$, and since $\pi^{-1}_{\overline{\mathcal{O}},1}(\mathcal{O})$ has dimension $2 \dim \mathcal{O} \leqslant \dim \mathfrak{g}$, it follows that $\mathcal{O}$ verifies $\mathrm{RC}_2(1)$ and RC_1. Now let $m \geqslant 2$. According to Lemma 3.5(2), we have

$$\dim \pi^{-1}_{\overline{\mathcal{O}},m}(0) \geqslant m \dim \mathcal{O} + \operatorname{codim}_{\mathfrak{g}}(\mathcal{O}) \geqslant (m+1) \dim \mathcal{O},$$

since $\operatorname{codim}_{\mathfrak{g}}(\mathcal{O}) \geqslant \dim \mathcal{O}$ because $\mathcal{O}$ is little. Hence $\mathcal{O}$ verifies $\mathrm{RC}_2(m)$. $\square$

When $\mathfrak{g}$ is simple, there is a unique nonzero nilpotent orbit $\mathcal{O}_{\min}$, called the *minimal* nilpotent orbit of $\mathfrak{g}$, of minimal dimension and it is contained in the closure of all nonzero nilpotent orbits.

Corollary 4.3. *Assume that $\mathfrak{g}$ is simple and not of type A_1. Then $\mathcal{O}_{\min}$ is little. In particular, $\mathscr{J}_m(\overline{\mathcal{O}_{\min}})$ is reducible for every $m \in \mathbb{N}^*$.*

Proof. Let $e \in \mathcal{O}_{\min}$ that we embed into an $\mathfrak{sl}_2$-triple (e, h, f) of $\mathfrak{g}$, and consider the corresponding Dynkin grading,

$$\mathfrak{g} = \bigoplus_{i \in \mathbb{Z}} \mathfrak{g}(i) \quad \text{with } \mathfrak{g}(i) := \{x \in \mathfrak{g} \mid [h, x] = ix\}.$$

By [Collingwood and McGovern 1993, Lemma 4.1.3],

$$\dim \mathcal{O} = \dim \mathfrak{g} - \dim \mathfrak{g}(0) - \dim \mathfrak{g}(1).$$

In addition, since $e \in \mathcal{O}_{\min}$, we have $\dim \mathfrak{g}(2) = 1$ and $\mathfrak{g} = \sum_{-2 \leqslant i \leqslant 2} \mathfrak{g}(i)$ [Tauvel and Yu 2005, Proposition 34.4.1]. As a result,

$$\dim \mathfrak{g} - 2 \dim \mathcal{O} = \dim \mathfrak{g}(0) - 2.$$

The Levi subalgebra $\mathfrak{g}(0)$ contains a Cartan subalgebra which has dimension at least two by our hypothesis. Hence, $\dim \mathfrak{g} - 2 \dim \mathcal{O} \geqslant 0$, and so $\mathcal{O}_{\min}$ is little. □

For classical simple Lie algebras, there are explicit formulas (see Appendix A) for the dimension of nilpotent orbits. This allows us to readily obtain examples of little nilpotent orbits.

Example 4.4. Let $n \in \mathbb{N}^*$ and $p, q \in \mathbb{N}$.

(i) A nilpotent orbit of $\mathfrak{sl}_n(\mathbb{C})$ corresponding to a rectangular partition is never little.

(ii) The nilpotent orbit $\mathcal{O}_{(2^p, 1^q)}$ of $\mathfrak{sl}_{2p+q}(\mathbb{C})$ is little if and only if $p, q \in \mathbb{N}^*$.

(iii) The nilpotent orbit $\mathcal{O}_{(p, 1^q)}$ of $\mathfrak{sl}_{p+q}(\mathbb{C})$ is little for $q \gg p$.

Explicit computations suggest that it is unlikely that there is a nice description of little nilpotent orbits in terms of partitions.

For the notation $\mathscr{P}_\varepsilon(n)$, $\varepsilon \in \{0, 1\}$, and $\mathcal{O}_{\boldsymbol{\lambda}}$ with $\boldsymbol{\lambda} \in \mathscr{P}_\varepsilon(n)$, $n \in \mathbb{N}^*$, refer to Appendix A.

Example 4.5. Let $\boldsymbol{\lambda} = (2^p, 1^q)$, with $p \in \mathbb{N}^*$ and $q \in \mathbb{N}$.

(i) If p is even, then $\boldsymbol{\lambda} \in \mathscr{P}_1(n)$, and the nilpotent orbit $\mathcal{O}_{\boldsymbol{\lambda}}$ of $\mathfrak{so}_{2p+q}(\mathbb{C})$ is little.

(ii) If q is even, then $\boldsymbol{\lambda} \in \mathscr{P}_{-1}(n)$, and the nilpotent orbit $\mathcal{O}_{\boldsymbol{\lambda}}$ of $\mathfrak{sp}_{2p+q}(\mathbb{C})$ is little if and only if $p \leqslant q(q+1)/2$.

The next proposition will allow us to produce new examples of nilpotent orbits which verify RC_1 by the "restriction" of certain little nilpotent orbits to Levi subalgebras.

Recall that for $\mathcal{O}$ a nilpotent orbit of some reductive Lie algebra $\mathfrak{a}$, the semisimple Lie algebra $\mathfrak{a}_{\mathcal{O}}$ was defined in Definition 3.1.

Proposition 4.6. *Assume that* $\mathfrak{g}$ *is simple. Let* $\mathfrak{l}$ *be a Levi subalgebra of* $\mathfrak{g}$ *with center* $\mathfrak{z}(\mathfrak{l})$, *and denote by* A *the connected subgroup of* G *whose Lie algebra is* $\mathfrak{a} := [\mathfrak{l}, \mathfrak{l}]$. *Let* e *be a nilpotent element of* $\mathfrak{a}$ *and suppose that the following conditions are satisfied*:

(i) $\mathfrak{a}$ *contains a regular semisimple element of* $\mathfrak{g}$,

(ii) $\mathfrak{a}_{A \cdot e} = \mathfrak{a}$,

(iii) $2 \dim G \cdot e \leqslant \dim \mathfrak{g} - \dim \mathfrak{z}(\mathfrak{l})$.

Then $A \cdot e$ *verifies* RC_1.

Proof. Define the following maps

$$\theta : G \times \mathfrak{a} \to \mathfrak{g},\ (g, x) \mapsto g(x), \quad \eta = \eta_{\mathfrak{g},e} : G \times [e, \mathfrak{g}] \to \mathfrak{g},\ (g, x) \mapsto g(x).$$

Observe that the image of each of these maps is irreducible. Moreover, for any $x \in \mathfrak{g}$, the map $g \mapsto (g^{-1}, g(x))$ defines a bijection between $G_\theta(x) := \{g \in G \mid g(x) \in \mathfrak{a}\}$ and $\theta^{-1}(\{x\})$. Similarly, we have a bijection between $G_\eta(x) := \{g \in G \mid g(x) \in [e, \mathfrak{g}]\}$ and $\eta^{-1}(\{x\})$. These bijections are isomorphisms of varieties.

Step 1. We shall first compute the dimension of the image of θ.

Let L be the connected subgroup of G whose Lie algebra is $\mathfrak{l}$. By condition (i), $\mathfrak{a}$ contains regular semisimple elements of $\mathfrak{g}$. If s is such an element, then $\mathfrak{g}^s$ is a Cartan subalgebra of $\mathfrak{l}$. Let $g \in G_\theta(s)$. Then $g(s) \in \mathfrak{a}$ and $\mathfrak{g}^{g(s)} = g(\mathfrak{g}^s)$ is another Cartan subalgebra of $\mathfrak{l}$. It follows that there exists $\tau \in L$ such that $\tau g \in N_G(\mathfrak{g}^s)$, with $N_G(\mathfrak{g}^s)$ the normalizer of $\mathfrak{g}^s$ in G. Hence, $g \in L N_G(\mathfrak{g}^s)$. Thus, we have obtained the inclusion $G_\theta(s) \subset L N_G(\mathfrak{g}^s)$. On the other hand, since L normalizes $\mathfrak{a}$, we get $L \subset G_\theta(s)$ and therefore $\dim L \leqslant \dim G_\theta(s)$.

Let $C_G(\mathfrak{g}^s)$ and $C_L(\mathfrak{g}^s)$ be the centralizers of $\mathfrak{g}^s$ in G and L, respectively. Since $\mathfrak{g}^s$ is a Cartan subalgebra, $C_G(\mathfrak{g}^s)$ is connected, so $C_G(\mathfrak{g}^s) = C_L(\mathfrak{g}^s)$ is contained in L. It follows that $L N_G(\mathfrak{g}^s)$ is a finite union of right L-cosets. We deduce that

$$\dim \theta^{-1}(\{s\}) = \dim G_\theta(s) = \dim L = \dim \mathfrak{a} + \mathfrak{z}(\mathfrak{l}).$$

Since the set of regular semisimple elements in $\mathfrak{g}$ is open and dense, we obtain that for s as above,

$$\dim \overline{\operatorname{im} \theta} = \dim \mathfrak{g} + \dim \mathfrak{a} - \dim \theta^{-1}(\{s\}) = \dim \mathfrak{g} - \dim \mathfrak{z}(\mathfrak{l}).$$

Step 2. We now consider the image of η.

Let (e, h, f) be an $\mathfrak{sl}_2$-triple of $\mathfrak{g}$. We easily check that $\mathfrak{c} := \mathbb{C}h \oplus \mathfrak{g}^e$ is a Lie subalgebra, and that $\mathfrak{c}$ stabilizes $[e, \mathfrak{g}]$. Let C be the connected subgroup of G whose Lie algebra is $\mathfrak{c}$. Then C is contained in $G_\eta(x)$ for any $x \in [e, \mathfrak{g}]$. In particular, $\dim G_\eta(x) \geqslant \dim C = 1 + \dim \mathfrak{g}^e$ for $x \in [e, \mathfrak{g}]$, and so

$$\dim \overline{\operatorname{im} \eta} \leqslant \dim \mathfrak{g} + \dim[e, \mathfrak{g}] - 1 - \dim \mathfrak{g}^e = 2 \dim G \cdot e - 1.$$

Step 3. By condition (iii) and Steps 1 and 2, we deduce that $\dim \overline{\operatorname{im}\theta} > \dim \overline{\operatorname{im}\eta}$. Thus $\overline{\operatorname{im}\theta} \not\subset \overline{\operatorname{im}\eta}$. We claim that this implies that $A \cdot e$ is RC_1. Let us suppose on the contrary that $A \cdot e$ is not RC_1. By condition (ii) and Lemma 3.5(1), $\pi^{-1}_{\overline{A\cdot e},1}(0) = \{0\} \times \mathfrak{a}$. So, $\pi^{-1}_{\overline{A\cdot e},1}(0)$ is contained in

$$\overline{\pi^{-1}_{\overline{A\cdot e},1}(A \cdot e)}.$$

Recall from the end of Section 2 the notation G_1 and A_1 for $\mathscr{J}_1(G)$ and $\mathscr{J}_1(A)$, respectively. It follows that

$$\{0\} \times G \cdot \mathfrak{a} \subset G_1 \cdot (\{0\} \times \mathfrak{a}) \subset G_1 \overline{A_1 \cdot e} \subset \overline{G_1 \cdot e},$$

whence

$$\{0\} \times \overline{G \cdot \mathfrak{a}} \subset \overline{G_1 \cdot e}.$$

Since $\overline{\pi^{-1}_{\overline{G\cdot e},1}(G \cdot e)} = \overline{G_1 \cdot e}$ (see Lemma 2.10), it follows from the proof of Proposition 3.6 that

$$\overline{G_1 \cdot e} \cap (\{0\} \times \mathfrak{g}) = \overline{\pi^{-1}_{\overline{G\cdot e},1}(G \cdot e)} \cap (\{0\} \times \mathfrak{g}) = \{0\} \times \overline{G \cdot [e, \mathfrak{g}]}.$$

Hence we get $\overline{\operatorname{im}\theta} \subset \overline{\operatorname{im}\eta}$ and the contradiction. □

Suppose that $\mathfrak{g}$ is simple. Let us fix a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$. Denote by Δ the root system relative to $(\mathfrak{g}, \mathfrak{h})$ and let us fix a system of simple roots Π. Given $S \subset \Pi$, we denote $\Delta_S = \mathbb{Z}S \cap \Delta$ the subroot system generated by S, and

$$\mathfrak{l}_S = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta_S} \mathfrak{g}_\alpha$$

where $\mathfrak{g}_\alpha$ denotes the root subspace relative to α. Then $\mathfrak{l}_S$ is a Levi subalgebra of $\mathfrak{g}$ and any Levi subalgebra of $\mathfrak{g}$ is conjugate to one in this form.

Given $S \subset \Pi$, set $\mathfrak{t} = [\mathfrak{l}_S, \mathfrak{l}_S] \cap \mathfrak{h}$. Then, $\mathfrak{l}_S$ verifies condition (i) if and only if $\mathfrak{t} \not\subset \bigcup_{\alpha \in \Delta} \ker \alpha$. To check the latter condition, it is enough to verify that for every $\alpha \in \Delta$, there is $\beta \in S$ such that $\langle \beta^\vee, \alpha \rangle \neq 0$.

Thus not all Levi subalgebras of $\mathfrak{g}$ verify condition (i) of Proposition 4.6. For example, if $\mathfrak{g}$ is simple of type B_ℓ, then a (maximal) Levi subalgebra whose semisimple part is simple of type $\mathrm{B}_{\ell-1}$ does not verify the condition. The same goes for a Levi subalgebra in type C_ℓ whose semisimple part is simple of type $\mathrm{C}_{\ell-1}$.

However, if $\mathfrak{g}$ is simple of type D_ℓ and if $\mathfrak{l}$ is a Levi subalgebra whose semisimple part is simple of type $\mathrm{D}_{\ell-1}$, then $\mathfrak{l}$ verifies the condition (i). Likewise, if $\mathfrak{g}$ is simple of type E_7 and if $\mathfrak{l}$ is a Levi subalgebra whose semisimple part is simple of type E_6, then $\mathfrak{l}$ verifies the condition (i). Applying Proposition 4.6, we obtain examples of nilpotent orbits in types D or E_6 which verify RC_1 that are not little.

We list in Table 1 some nilpotent orbits that we obtain in this way. In all the examples presented in the table, the center of the Levi subalgebra is 1-dimensional, and $\mathfrak{a}$ is simple. The first and second columns give the type of the simple Lie

$\mathfrak{g}$	$\mathfrak{a}$	$G \cdot e$	$A \cdot e$
D_6	D_5	$(3, 2^2, 1^5)$	$(3, 2^2, 1^3)$
D_7	D_6	$(3^2, 1^8)$	$(3^2, 1^6)$
D_9	D_8	$(3^2, 2^2, 1^8)$	$(3^2, 2^2, 1^6)$
D_{10}	D_9	$(3^3, 1^{11})$	$(3^3, 1^9)$
D_{10}	D_9	$(4^2, 1^{12})$	$(4^2, 1^{10})$
D_{10}	D_9	$(5, 2^2, 1^{11})$	$(5, 2^2, 1^9)$
D_{10}	D_9	$(5, 3, 1^{12})$	$(5, 3, 1^{10})$
E_7	E_6	$(3A_1)'$	$3A_1$
E_7	E_6	A_2	A_2

Table 1. Examples of nonlittle nilpotent orbits satisfying RC_1 obtained by restriction.

algebras $\mathfrak{g}$ and $\mathfrak{a}$. Condition (ii) is verified in view of the discussion above. We describe the nilpotent orbits $G \cdot e$ and $A \cdot e$ in the third and fourth columns, respectively. The description for an orbit in $\mathfrak{g}$ of type D is given in terms of partitions (see Appendix A), while for an orbit in $\mathfrak{g}$ of type E_6 or E_7, it is given by its Bala–Carter label.

Remark 4.7. (1) The first and last lines of Table 1 provide examples of a *rigid*[2] nilpotent orbit which verifies RC_1 and which is not little.

(2) Propositions 3.6, 4.2, and 4.6, together with Remark 3.9, allow us to classify all nilpotent orbits verifying RC_1 in simple Lie algebras of exceptional type. They are listed in Appendix C.

5. Induced nilpotent orbits

Let $\mathfrak{l}$ be a proper Levi subalgebra of $\mathfrak{g}$, and let $\mathfrak{p}$ be a parabolic subalgebra of $\mathfrak{g}$ with Levi decomposition $\mathfrak{p} = \mathfrak{l} \oplus \mathfrak{u}$ so that $\mathfrak{u}$ is the nilpotent radical of $\mathfrak{p}$. Let P, L, and U be the connected closed subgroups of G whose Lie algebra are $\mathfrak{p}$, $\mathfrak{l}$, and $\mathfrak{u}$, respectively. Then $P = LU$.

The following definitions and results on induced nilpotent orbits are mostly extracted from [Richardson 1974; Lusztig and Spaltenstein 1979]. We refer to [Collingwood and McGovern 1993, Chapter 7] for a recent survey.

Theorem 5.1. *Let $\mathcal{O}_{\mathfrak{l}}$ be a nilpotent orbit of $\mathfrak{l}$. There exists a unique nilpotent orbit $\mathcal{O}_{\mathfrak{g}}$ in $\mathfrak{g}$ whose intersection with $\mathcal{O}_{\mathfrak{l}} + \mathfrak{u}$ is a dense open subset of $\mathcal{O}_{\mathfrak{l}} + \mathfrak{u}$. Moreover, the intersection of $\mathcal{O}_{\mathfrak{g}}$ with $\mathcal{O}_{\mathfrak{l}} + \mathfrak{u}$ consists of a single P-orbit and* $\operatorname{codim}_{\mathfrak{g}}(\mathcal{O}_{\mathfrak{g}}) = \operatorname{codim}_{\mathfrak{l}}(\mathcal{O}_{\mathfrak{l}})$.

[2] See Section 5 for the notion of rigid nilpotent orbit, and Appendices A and C for the description of rigid nilpotent orbits in simple Lie algebras.

The nilpotent orbit $\mathcal{O}_{\mathfrak{g}}$ only depends on $\mathfrak{l}$, and not on the choice of a parabolic subalgebra $\mathfrak{p}$ containing it. The nilpotent orbit $\mathcal{O}_{\mathfrak{g}}$ is called the *induced nilpotent orbit of* $\mathfrak{g}$ *from* $\mathcal{O}_{\mathfrak{l}}$, and it is denoted by $\mathrm{Ind}_{\mathfrak{l}}^{\mathfrak{g}}(\mathcal{O}_{\mathfrak{l}})$. A nilpotent orbit which is not induced in a proper way from another one is called *rigid*. In type A, only the zero orbit is rigid.

Remark 5.2. (1) Let $\mathfrak{s}_1, \dots, \mathfrak{s}_n$ be the simple factors of $[\mathfrak{g}, \mathfrak{g}]$ and denote by $\mathfrak{z}(\mathfrak{g})$ the center of $\mathfrak{g}$. Then there are Levi subalgebras $\mathfrak{r}_1, \dots, \mathfrak{r}_n$ of $\mathfrak{s}_1, \dots, \mathfrak{s}_n$, respectively, such that

$$\mathfrak{l} = \mathfrak{z}(\mathfrak{g}) \times \mathfrak{r}_1 \times \cdots \times \mathfrak{r}_n.$$

If $\mathcal{O}_{\mathfrak{l}}$ is a nilpotent orbit of $\mathfrak{l}$, then $\mathcal{O}_{\mathfrak{l}} = \mathcal{O}_{\mathfrak{r}_1} \times \cdots \times \mathcal{O}_{\mathfrak{r}_n}$, where $\mathcal{O}_{\mathfrak{r}_1}, \dots, \mathcal{O}_{\mathfrak{r}_n}$ are nilpotent orbits in the semisimple parts of $\mathfrak{r}_1, \dots, \mathfrak{r}_n$, respectively. Then

$$\mathrm{Ind}_{\mathfrak{l}}^{\mathfrak{g}}(\mathcal{O}_{\mathfrak{l}}) = \mathrm{Ind}_{\mathfrak{r}_1}^{\mathfrak{s}_1}(\mathcal{O}_{\mathfrak{r}_1}) \times \cdots \times \mathrm{Ind}_{\mathfrak{r}_n}^{\mathfrak{s}_n}(\mathcal{O}_{\mathfrak{r}_n}) = \mathrm{Ind}_{[\mathfrak{g},\mathfrak{g}]\cap\mathfrak{l}}^{[\mathfrak{g},\mathfrak{g}]}(\mathcal{O}_{\mathfrak{l}}).$$

(2) The induction property is transitive in the following sense [Collingwood and McGovern 1993, Proposition 7.1.4]: if $\mathfrak{l}_1$ and $\mathfrak{l}_2$ are two Levi subalgebras of $\mathfrak{g}$ with $\mathfrak{l}_1 \subset \mathfrak{l}_2$, then

$$\mathrm{Ind}_{\mathfrak{l}_2}^{\mathfrak{g}}(\mathrm{Ind}_{\mathfrak{l}_1}^{\mathfrak{l}_2}(\mathcal{O}_{\mathfrak{l}_1})) = \mathrm{Ind}_{\mathfrak{l}_1}^{\mathfrak{g}}(\mathcal{O}_{\mathfrak{l}_1}).$$

(3) If $\Omega_{\mathfrak{l}}$ is an L-orbit in $\overline{\mathcal{O}_{\mathfrak{l}}} \setminus \mathcal{O}_{\mathfrak{l}}$, then the induced nilpotent orbit of $\mathfrak{g}$ from $\Omega_{\mathfrak{l}}$ is contained in $\overline{\mathcal{O}_{\mathfrak{g}}} \setminus \mathcal{O}_{\mathfrak{g}}$.

Let $\mathcal{O}_{\mathfrak{l}}$ be a nilpotent orbit of $\mathfrak{l}$ and denote by $\mathcal{O}_{\mathfrak{g}}$ the induced nilpotent orbit of $\mathfrak{g}$ from $\mathcal{O}_{\mathfrak{l}}$. According to Theorem 5.1, $\mathcal{O}_{\mathfrak{g}} \cap (\mathcal{O}_{\mathfrak{l}} + \mathfrak{u})$ is a single P-orbit that we shall denote by $\mathcal{O}_{\mathfrak{p}}$; that is,

$$\mathcal{O}_{\mathfrak{p}} := \mathcal{O}_{\mathfrak{g}} \cap (\mathcal{O}_{\mathfrak{l}} + \mathfrak{u}).$$

Lemma 5.3. *The orbits satisfy*

$$\overline{\mathcal{O}_{\mathfrak{p}}} = \overline{\mathcal{O}_{\mathfrak{l}}} + \mathfrak{u}, \quad \overline{\mathcal{O}_{\mathfrak{p}}} \cap \mathcal{O}_{\mathfrak{g}} = \mathcal{O}_{\mathfrak{p}}, \quad \overline{\mathcal{O}_{\mathfrak{g}}} = G \cdot (\overline{\mathcal{O}_{\mathfrak{l}}} + \mathfrak{u}).$$

Proof. The first equality is obvious since $\mathcal{O}_{\mathfrak{p}}$ is dense in $\mathcal{O}_{\mathfrak{l}} + \mathfrak{u}$ by definition.

Next, the inclusion $\mathcal{O}_{\mathfrak{p}} \subset \overline{\mathcal{O}_{\mathfrak{p}}} \cap \mathcal{O}_{\mathfrak{g}}$ is clear. To show the other inclusion, assume that there is $x \in \overline{\mathcal{O}_{\mathfrak{p}}} \cap \mathcal{O}_{\mathfrak{g}}$, with $x \notin \mathcal{O}_{\mathfrak{p}}$. A contradiction is expected. Since $x \in \overline{\mathcal{O}_{\mathfrak{p}}} \setminus \mathcal{O}_{\mathfrak{p}}$, $\dim P \cdot x < \dim P \cdot e$. Hence,

$$\dim \mathfrak{g}^x \geqslant \dim \mathfrak{p}^x > \dim \mathfrak{p}^e = \dim \mathfrak{g}^e.$$

As a consequence, x is not in $\mathcal{O}_{\mathfrak{g}}$, whence the contradiction.

A proof of the last equality can be found in [op. cit., Theorem 7.1.3]. □

We have the following generalization.

Lemma 5.4. *The jet schemes satisfy*

(1) $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})} = \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{l})} + \mathfrak{u}_m$,
(2) $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})} \cap \mathscr{J}_m(\mathcal{O}_\mathfrak{g}) = \mathscr{J}_m(\mathcal{O}_\mathfrak{p}) = (\mathscr{J}_m(\overline{\mathcal{O}}_\mathfrak{l}) + \mathfrak{u}_m) \cap \mathscr{J}_m(\mathcal{O}_\mathfrak{g})$,
(3) $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{g})}$ *is the closure of* $G_m \cdot \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})}$.

Proof. (1) Since $\mathcal{O}_\mathfrak{p} \subset \mathcal{O}_\mathfrak{l} + \mathfrak{u}$, we get $\mathscr{J}_m(\mathcal{O}_\mathfrak{p}) \subset \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{l})} + \mathfrak{u}_m$ because $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{l})} + \mathfrak{u}_m$ is closed. Let $e' \in \mathcal{O}_\mathfrak{l}$ and $x \in \mathfrak{u}$ be such that $e := e' + x$ is in $\mathcal{O}_\mathfrak{p}$. From the above inclusion, we deduce that

$$\dim \mathfrak{p} - \dim \mathfrak{p}^e \leqslant \dim \mathfrak{l} - \dim \mathfrak{l}^{e'} + \dim \mathfrak{u} = \dim \mathfrak{p} - \dim \mathfrak{g}^e,$$

because $\dim \mathfrak{l}^{e'} = \dim \mathfrak{g}^e$ by Theorem 5.1. Since $\dim \mathfrak{p}^e \leqslant \dim \mathfrak{g}^e$, we get $\mathfrak{p}^e = \mathfrak{g}^e$, whence $\dim \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})} = \dim(\mathscr{J}_m(\mathcal{O}_\mathfrak{l}) + \mathfrak{u}_m)$ by Lemma 2.10 and Proposition 2.6. So $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})}$ and $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{l})} + \mathfrak{u}_m$ are irreducible varieties of the same dimension, and the equality follows.

(2) Taking into account Lemma 2.10 and Proposition 2.6, the result follows from the same arguments as in the proof of the second equality of Lemma 5.3.

(3) By Lemma 2.10,

$$\mathscr{J}_m(\mathcal{O}_\mathfrak{g}) = G_m \cdot \mathscr{J}_m(\mathcal{O}_\mathfrak{p}) \subset G_m \cdot \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})}.$$

As a result, the jet scheme $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{g})}$ is in the closure of $G_m \cdot \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})}$. On the other hand, since $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{g})}$ is G_m-stable, we get

$$G_m \cdot \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})} \subset \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{g})}.$$

So the closure of $G_m \cdot \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})}$ is contained in $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{g})}$, whence the expected equality. □

Question 5.5. For $m = 0$, $G_m \cdot \overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{p})}$ is closed (see Lemma 5.3) essentially because G/P is compact. For $m \geqslant 1$, G_m/P_m is a trivial fibration over G/P with m-dimensional affine fiber. Can we show nevertheless that $G_m \cdot (\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{l})} + \mathfrak{u}_m)$ is closed, in other words that $\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{g})} = G_m \cdot (\overline{\mathscr{J}_m(\mathcal{O}_\mathfrak{l})} + \mathfrak{u}_m)$?

Theorem 5.6. *Let* $\mathfrak{l}$ *be a Levi subalgebra of* $\mathfrak{g}$, $\mathcal{O}_\mathfrak{l}$ *a nilpotent orbit of* $\mathfrak{l}$, *and* $\mathcal{O}_\mathfrak{g}$ *the induced nilpotent orbit of* $\mathfrak{g}$ *from* $\mathcal{O}_\mathfrak{l}$. *If* $\mathcal{O}_\mathfrak{l}$ *verifies* $\mathrm{RC}_2(m)$ *for some* $m \in \mathbb{N}^*$, *then* $\mathcal{O}_\mathfrak{g}$ *also verifies* $\mathrm{RC}_2(m)$.

The rest of the section will be devoted to the proof of Theorem 5.6.

Definition 5.7. Let $\mathfrak{l}$ be a Levi subalgebra of $\mathfrak{g}$. We say that $\mathfrak{l}$ is a *maximal Levi subalgebra of* $\mathfrak{g}$ if the center of $[\mathfrak{g}, \mathfrak{g}] \cap \mathfrak{l}$ has dimension one.

Let us first assume that $\mathfrak{g}$ is simple and that $\mathfrak{l}$ is a maximal Levi subalgebra of $\mathfrak{g}$. Thus, the center $\mathfrak{z}(\mathfrak{l})$ of $\mathfrak{l}$ has dimension one. Let us fix a Cartan subalgebra $\mathfrak{h}$ in $\mathfrak{l}$ and Δ the root system relative to $(\mathfrak{g}, \mathfrak{h})$. There exists a simple root system Π and a subset $\Pi' \subseteq \Pi$ verifying $\operatorname{card}(\Pi \setminus \Pi') = 1$ such that $\mathfrak{l}$ is the sum of $\mathfrak{h}$ and all the α-root spaces for α in the root subsystem generated by Π'. Define z to be the element in $\mathfrak{h}$ such that

$$\alpha(z) = \begin{cases} 0 & \text{if } \alpha \in \Pi', \\ 1 & \text{if } \alpha \in \Pi \setminus \Pi'. \end{cases}$$

Then z is a generator of $\mathfrak{z}(\mathfrak{l})$ and all the eigenvalues of $\operatorname{ad} z$ are integers.

Let $m \in \mathbb{N}$. Then $\operatorname{ad} z$ induces a $\mathbb{Z}$-grading on $\mathfrak{g}_m$,

$$\mathfrak{g}_m = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_m(k) \quad \text{with} \quad \mathfrak{g}_m(k) := \{y \in \mathfrak{g}_m \mid [z, y] = ky\}.$$

Set

$$\mathfrak{p} = \bigoplus_{k \geqslant 0} \mathfrak{g}_0(k) \quad \text{and} \quad \mathfrak{u} = \bigoplus_{k > 0} \mathfrak{g}_0(k).$$

Then $\mathfrak{p}$ is a parabolic subalgebra of $\mathfrak{g}$, where $\mathfrak{l} = \mathfrak{g}_0(0)$ is a Levi factor and whose nilpotent radical is $\mathfrak{u}$. Denote by P, L, and U the connected closed subgroups of G whose Lie algebras are $\mathfrak{p}$, $\mathfrak{l}$, and $\mathfrak{u}$, respectively.

Observe that

$$\mathfrak{l}_m = \mathfrak{z}(\mathfrak{l})_m \oplus [\mathfrak{l}_m, \mathfrak{l}_m] = \mathfrak{g}_m(0), \quad \mathfrak{p}_m = \bigoplus_{k \geqslant 0} \mathfrak{g}_m(k), \quad \mathfrak{u}_m = \bigoplus_{k > 0} \mathfrak{g}_m(k).$$

Remark 5.8. Clearly, for any nonzero integer k, we have $[z, \mathfrak{g}_m(k)] = \mathfrak{g}_m(k)$. In particular, $\mathfrak{g}_m(0) = (\mathfrak{g}_m)^z = \mathfrak{n}_{\mathfrak{g}_m}(\mathbb{C}z)$ where $\mathfrak{n}_{\mathfrak{g}_m}(\mathbb{C}z)$ is the normalizer of z in $\mathfrak{g}_m$. Also, if $x \in \mathfrak{g}_m(k)$, with $k \in \mathbb{N}^*$, then x is ad-nilpotent, and $e^{\operatorname{ad} x} z = z + [x, z] = z - kx$.

Lemma 5.9. *Let $\lambda \in \mathbb{C}^*$, $x \in \mathfrak{g}_m(0)$, and $y \in \mathfrak{u}_m$. If x is* ad*-nilpotent in $\mathfrak{g}_m$ then there exists $\tau \in U_m$ such that $\tau(\lambda z + x + y) = \lambda z + x$.*

Proof. For some $p > 0$, $y = y_p + t$ with $y_p \in \mathfrak{g}_m(p)$ and $t \in \sum_{k \geqslant p+1} \mathfrak{g}_m(k)$. Since x is ad-nilpotent, the sequence $((\operatorname{ad} x)^n \mathfrak{g}_m(p))_{n \in \mathbb{N}}$ is decreasing and $(\operatorname{ad} x)^n \mathfrak{g}_m(p) = \{0\}$ for $n \geqslant \dim \mathfrak{g}_m(p)$. Let $q \in \mathbb{N}$ be such that $y_p \in (\operatorname{ad} x)^q \mathfrak{g}_m(p)$. Then

$$\begin{aligned} e^{(1/p\lambda) \operatorname{ad} y_p}(\lambda z + x + y) &= \lambda z + e^{(1/p\lambda) \operatorname{ad} y_p} x + e^{(1/p\lambda) \operatorname{ad} y_p} t \\ &= \lambda z + x + (1/p\lambda)[y_p, x] + t' = \lambda z + x + y' \end{aligned}$$

with $t' \in \sum_{k \geqslant p+1} \mathfrak{g}_m(k)$, $y' := (1/p\lambda)[y_p, x] + t'$, and

$$(1/p\lambda)[y_p, x] \in (\operatorname{ad} x)^{q+1} \mathfrak{g}_m(p).$$

Therefore we may start again with y'. After a finite number of steps, we come to an element in $\sum_{k \geqslant p+1} \mathfrak{g}_m(k)$. Then we start again with $p + 1$ instead of p and, after a finite number of steps, we come to an element of the expected form, $\lambda z + x$. $\square$

Lemma 5.10. *Let Ω be an L-orbit contained in $\overline{\mathcal{O}}_{\mathfrak{l}}$ and let X be an irreducible component of $\pi^{-1}_{\overline{\mathcal{O}}_{\mathfrak{l}},m}(\overline{\Omega})$. Then*

$$\dim G_m \cdot (\mathfrak{z}(\mathfrak{l}) + X + \mathfrak{u}_m) = \dim X + 2\dim \mathfrak{u}_m + 1.$$

Proof. Set

$$C := \mathfrak{z}(\mathfrak{l}) + X + \mathfrak{u}_m.$$

Since Ω and $\overline{\Omega}$ are L-stable, $\pi^{-1}_{\overline{\mathcal{O}}_{\mathfrak{l}},m}(\overline{\Omega})$ is L_m-stable and so is X. In addition, $\mathfrak{z}(\mathfrak{l})$ is L_m-stable too. Hence, C is P_m-stable because

$$P_m \cdot C = L_m U_m \cdot (\mathfrak{z}(\mathfrak{l}) + X + \mathfrak{u}_m) = L_m \cdot (\mathfrak{z}(\mathfrak{l}) + X + \mathfrak{u}_m) \subset C.$$

Observe also that the elements of X are all ad-nilpotent.

Consider the action of P_m on $G_m \times C$ given by $\rho \cdot (\sigma, c) = (\sigma\rho^{-1}, \rho(c))$. Denote by $\overline{(\sigma, c)}$ the P_m-orbit of $(\sigma, c) \in G_m \times C$ with respect to this action, and denote by $G_m \times_{P_m} C$ the corresponding quotient space. The natural morphism

$$G_m \times C \to \mathfrak{g}, \quad (\sigma, c) \mapsto \sigma(c)$$

factors through the quotient and we obtain a morphism

$$\psi : G_m \times_{P_m} C \to \mathfrak{g}$$

whose image is $G_m \cdot C$. Since X and $\mathfrak{u}_m$ are both closed cones, $z = 1_{G_m}(z)$ lies in the image of ψ and

$$\psi^{-1}(z) = \{\overline{(\sigma, c)} \in G_m \times_{P_m} C \mid \sigma(c) = z\}.$$

Let $\overline{(\sigma, c)} \in \psi^{-1}(z)$. Because z is ad-semisimple, c is also ad-semisimple. Since all elements of X are ad-nilpotent, we deduce that c does not belong to $X + \mathfrak{u}_m$. Also, since $U_m \subset P_m$, we may assume by Lemma 5.9 that c is of the form $\lambda z + x$ with $\lambda \in \mathbb{C}^*$ and $x \in X$. Since $x \in \mathfrak{g}_m(0) = (\mathfrak{g}_m)^z$, we deduce from the uniqueness of the Jordan decomposition that $c = \lambda z$. In particular, σ is in the normalizer $N_G(\mathbb{C}z)$ of z in G, and $c = \sigma^{-1}(z)$.

According to Remark 5.8, the identity component of the centralizer $C_{G_m}(z)$ of z in G_m is contained in P_m and has finite index in $N_{G_m}(\mathbb{C}z)$. Consequently, $\psi^{-1}(z)$ is a finite set. Thus, we get that $\dim G_m \cdot C = \dim G_m \times_{P_m} C$ because they are both irreducible subsets. To conclude, it suffices to observe that $\dim G_m - \dim P_m = \dim \mathfrak{u}_m$ and $\dim C = 1 + \dim X + \dim \mathfrak{u}_m$ since $\mathfrak{z}(\mathfrak{l}) = \mathbb{C}z$. □

Since $\mathfrak{g}$ is simple, its Killing form $\langle \cdot\,, \cdot \rangle$ is nondegenerate. Let us denote by ϕ the element of $\mathbb{C}[\mathfrak{g}]^G$ defined for all $x \in \mathfrak{g}$ by

$$\phi(x) = \langle x, x \rangle.$$

By our choice of z, $\phi(z)$ is a nonzero positive integer. Set

$$\mathscr{C} := \mathfrak{z}(\mathfrak{l}) + \overline{\mathcal{O}}_{\mathfrak{l}} + \mathfrak{u}$$

Lemma 5.11. *The nullvariety in $\mathscr{C}$ of ϕ is $\overline{\mathcal{O}}_{\mathfrak{l}} + \mathfrak{u}$.*

Proof. First of all, $\overline{\mathcal{O}}_{\mathfrak{l}} + \mathfrak{u}$ is contained in the nullvariety in $\mathscr{C}$ of ϕ. For the other inclusion, let $u = \lambda z + x + y$ be in $\mathscr{C}$, with $\lambda \in \mathbb{C}$, $x \in \overline{\mathcal{O}}_{\mathfrak{l}}$, and $y \in \mathfrak{u}$, such that $\phi(u) = 0$. We have

$$0 = \phi(u) = \langle \lambda z + x + y, \lambda z + x + y\rangle = \lambda^2 \langle z, z\rangle + \langle x, x\rangle = \lambda^2 \langle z, z\rangle$$

since $\mathfrak{u}$ is orthogonal to $\mathfrak{p}$, $\mathfrak{z}(\mathfrak{l})$ is orthogonal to $[\mathfrak{l}, \mathfrak{l}] \oplus \mathfrak{u}$, and $\langle x, x\rangle = \phi(x) = 0$. Hence $\lambda = 0$ since $\phi(z) \neq 0$. So, u lies in $\overline{\mathcal{O}}_{\mathfrak{l}} + \mathfrak{u}$, whence the other inclusion. □

Let $\phi^{(0)}, \dots, \phi^{(m)} \in \mathbb{C}[\mathfrak{g}_m]$ be the polynomials as defined in Remark 2.2 relative to ϕ. According to Lemma 2.11, they are G_m-invariant. In particular, $\phi^{(0)}$ is G_m-invariant.

Lemma 5.12. *Let $\Omega_{\mathfrak{l}}$ be an L-orbit contained in $\overline{\mathcal{O}}_{\mathfrak{l}}$ and set $\Omega_{\mathfrak{g}} := \mathrm{Ind}_{\mathfrak{l}}^{\mathfrak{g}}(\Omega_{\mathfrak{l}})$. Then*:

(1) *the nullvariety in $G_m \cdot (\mathfrak{z}(\mathfrak{l}) + \pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}^{-1}(\overline{\Omega_{\mathfrak{l}}}) + \mathfrak{u}_m)$ of $\phi^{(0)}$ is contained in $\pi_{\overline{\mathcal{O}}_{\mathfrak{g}},m}^{-1}(\overline{\Omega_{\mathfrak{g}}})$*,

(2) $\dim \pi_{\overline{\mathcal{O}}_{\mathfrak{g}},m}^{-1}(\overline{\Omega_{\mathfrak{g}}}) \geqslant \dim \pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}^{-1}(\overline{\Omega_{\mathfrak{l}}}) + 2 \dim \mathfrak{u}_m$.

Proof. Let us denote by Y the nullvariety in $G_m \cdot (\mathfrak{z}(\mathfrak{l}) + \pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}^{-1}(\overline{\Omega_{\mathfrak{l}}}) + \mathfrak{u}_m)$ of $\phi^{(0)}$. First of all, observe that Y contains 0 because each of the spaces $\mathfrak{z}(\mathfrak{l})$, $\pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}^{-1}(\overline{\Omega_{\mathfrak{l}}})$, and $\mathfrak{u}_m$ is a closed cone. In particular, Y is nonempty.

(1) Let $u = g \cdot (\lambda z + x + y)$ be in Y, with $g \in G_m$, $\lambda \in \mathbb{C}$, $x \in \pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}^{-1}(\overline{\Omega_{\mathfrak{l}}})$, and $y \in \mathfrak{u}_m$, such that $\phi^{(0)}(u) = 0$. Since $\phi^{(0)}$ is G_m-invariant, setting $x_0 := \pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}(x)$ and $y_0 := \pi_{\mathfrak{u},m}(y)$, we get

$$0 = \phi^{(0)}(u) = \phi^{(0)}(\lambda z + x + y) = \phi(\lambda z + x_0 + y_0) = \lambda^2 \phi(z)$$

by the computations of the proof of Lemma 5.11. Hence $\lambda = 0$ since $\phi(z) \neq 0$. So u lies in $G_m \cdot (\pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}^{-1}(\overline{\Omega_{\mathfrak{l}}}) + \mathfrak{u}_m)$. But

$$G_m \cdot (\pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}^{-1}(\overline{\Omega_{\mathfrak{l}}}) + \mathfrak{u}_m) \subset G_m \cdot (\mathscr{J}_m(\overline{\mathcal{O}}_{\mathfrak{l}}) + \mathfrak{u}_m) \subset G_m \cdot \mathscr{J}_m(\overline{\mathcal{O}}_{\mathfrak{g}}) = \mathscr{J}_m(\overline{\mathcal{O}}_{\mathfrak{g}})$$

because $\mathscr{J}_m(\overline{\mathcal{O}}_{\mathfrak{g}})$ is G_m-invariant. Thus Y is contained in $\mathscr{J}_m(\overline{\mathcal{O}}_{\mathfrak{g}})$. Then it remains to observe that for $u \in Y$,

$$\pi_{\overline{\mathcal{O}}_{\mathfrak{g}},m}(u) \in G \cdot (\overline{\Omega_l} + \mathfrak{u}) = \overline{\Omega_{\mathfrak{g}}}$$

by Lemma 5.3. In conclusion, Y is contained in $\pi_{\overline{\mathcal{O}}_{\mathfrak{g}},m}^{-1}(\overline{\Omega_{\mathfrak{g}}})$.

(2) Let X be an irreducible component of $\pi_{\overline{\mathcal{O}}_{\mathfrak{l}},m}^{-1}(\overline{\Omega})$ of maximal dimension, and let Y' be the nullvariety in $G_m \cdot (\mathfrak{z}(\mathfrak{l}) + X + \mathfrak{u}_m)$ of $\phi^{(0)}$. The function $\phi^{(0)}$ is not identically zero on $G_m \cdot (\mathfrak{z}(\mathfrak{l}) + X + \mathfrak{u}_m)$ since $z \in G_m \cdot (\mathfrak{z}(\mathfrak{l}) + X + \mathfrak{u}_m)$ and

$\phi^{(0)}(z) = \phi(z) \neq 0$. Since Y' is irreducible, we deduce by Lemma 5.10 and our choice of X that

$$\begin{aligned}\dim Y' &= \dim G_m \cdot (\mathfrak{z}(\mathfrak{l}) + X + \mathfrak{u}_m) - 1\\ &= \dim X + 2\dim \mathfrak{u}_m = \dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{l}},m}(\overline{\Omega_\mathfrak{l}}) + 2\dim \mathfrak{u}_m,\end{aligned}$$

whence the statement, by (1). $\square$

Proposition 5.13. *If for some L-orbit $\Omega_\mathfrak{l}$ in $\overline{\mathcal{O}_\mathfrak{l}}$, we have*

$$\dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{l}},m}(\overline{\Omega_\mathfrak{l}}) \geqslant \dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{l}},m}(\mathcal{O}_\mathfrak{l}),$$

then

$$\dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{g}},m}(\overline{\Omega_\mathfrak{g}}) \geqslant \dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{g}},m}(\mathcal{O}_\mathfrak{g}),$$

where $\Omega_\mathfrak{g}$ is the induced nilpotent orbit of $\mathfrak{g}$ from $\Omega_\mathfrak{l}$.

Proof. Assume that for some L-orbit $\Omega_\mathfrak{l}$ in $\overline{\mathcal{O}_\mathfrak{l}}$, we have

$$\dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{l}},m}(\overline{\Omega_\mathfrak{l}}) \geqslant \dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{l}},m}(\mathcal{O}_\mathfrak{l}).$$

Then by Lemma 5.12,

$$\begin{aligned}\dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{g}},m}(\overline{\Omega_\mathfrak{g}}) &\geqslant \dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{l}},m}(\overline{\Omega_\mathfrak{l}}) + 2\dim \mathfrak{u}_m\\ &\geqslant \dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{l}},m}(\mathcal{O}_\mathfrak{l}) + 2\dim \mathfrak{u}_m = (m+1)\dim \mathcal{O}_\mathfrak{l} + 2(m+1)\dim \mathfrak{u}.\end{aligned}$$

To conclude, it remains to observe that $\pi^{-1}_{\overline{\mathcal{O}_\mathfrak{g}},m}(\mathcal{O}_\mathfrak{g})$ has dimension

$$(m+1)\dim \mathcal{O}_\mathfrak{l} + 2(m+1)\dim \mathfrak{u}$$

because $\dim \mathcal{O}_\mathfrak{g} = 2\dim \mathfrak{u} + \dim \mathcal{O}_\mathfrak{l}$ from Theorem 5.1. $\square$

Remark 5.14. The above proof actually shows that $\pi^{-1}_{\overline{\mathcal{O}_\mathfrak{g}},m}(\overline{\Omega_\mathfrak{g}})$ has dimension at least $2(m+1)\dim \mathfrak{u} + \dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{l}},m}(\overline{\Omega_\mathfrak{l}})$ even if $\Omega_\mathfrak{l}$ does not verify the hypothesis of the proposition. This can be used in practice to give an estimate of $\dim \pi^{-1}_{\overline{\mathcal{O}_\mathfrak{g}},m}(\overline{\mathcal{O}_\mathfrak{g}} \setminus \mathcal{O}_\mathfrak{g})$.

We are now in a position to prove the main result of the section.

Proof of Theorem 5.6. Let $\mathfrak{l}$ be a Levi subalgebra of $\mathfrak{g}$. Then there is a finite sequence of Levi subalgebras

$$\mathfrak{l} = \mathfrak{l}_0 \subset \mathfrak{l}_1 \subset \mathfrak{l}_1 \subset \cdots \subset \mathfrak{l}_k = \mathfrak{g}$$

such that $\mathfrak{l}_{i-1}$ is a maximal Levi subalgebra of $\mathfrak{l}_i$ for every $i \in \{1, \ldots, k\}$.

Let $\mathcal{O}_\mathfrak{l}$ be a nilpotent orbit of $\mathfrak{l} = \mathfrak{l}_0$ verifying $\mathrm{RC}_2(m)$ for some $m \in \mathbb{N}$, and set for $i \in \{1, \ldots, k\}$,

$$\mathcal{O}_{\mathfrak{l}_i} = \mathrm{Ind}^{\mathfrak{l}_i}_{\mathfrak{l}_{i-1}}(\mathcal{O}_{\mathfrak{l}_{i-1}}).$$

Since induction is transitive (see Remark 5.2(2)),

$$\mathcal{O}_\mathfrak{g} := \mathrm{Ind}^{\mathfrak{g}}_{\mathfrak{l}}(\mathcal{O}_\mathfrak{l}) = \mathrm{Ind}^{\mathfrak{l}_k}_{\mathfrak{l}_{k-1}}(\mathrm{Ind}^{\mathfrak{l}_{k-1}}_{\mathfrak{l}_{k-2}}(\cdots(\mathrm{Ind}^{\mathfrak{l}_1}_{\mathfrak{l}_0}(\mathcal{O}_{\mathfrak{l}_0})))).$$

So, in order to proof Theorem 5.6, we may assume that $\mathfrak{l}$ is maximal in $\mathfrak{g}$. Let us write $\mathcal{O}_{\mathfrak{l}}$ as a product $\mathcal{O}_{\mathfrak{l}} = \mathcal{O}_{\mathfrak{r}_1} \times \cdots \times \mathcal{O}_{\mathfrak{r}_n}$, with the $\mathfrak{r}_j$ as in Remark 5.2(1). Since $\mathcal{O}_{\mathfrak{l}}$ verifies $\mathrm{RC}_2(m)$, $\mathcal{O}_{\mathfrak{r}_j}$ verifies $\mathrm{RC}_2(m)$ for some $j \in \{1, \dots, n\}$. Since $\mathfrak{l}$ is maximal in $\mathfrak{g}$, either $\mathfrak{r}_j = \mathfrak{s}_j$ and $\mathrm{Ind}_{\mathfrak{r}_j}^{\mathfrak{s}_j}(\mathcal{O}_{\mathfrak{r}_j})$ obviously verifies $\mathrm{RC}_2(m)$ too, or $\mathfrak{r}_j$ is maximal in $\mathfrak{s}_j$ and by Proposition 5.13, $\mathrm{Ind}_{\mathfrak{r}_j}^{\mathfrak{s}_j}(\mathcal{O}_{\mathfrak{r}_j})$ verifies $\mathrm{RC}_2(m)$ as well. Indeed, since $\mathcal{O}_{\mathfrak{r}_j}$ verifies $\mathrm{RC}_2(m)$, for some $\Omega_{\mathfrak{r}_j}$ in $\overline{\mathcal{O}_{\mathfrak{r}_j}} \setminus \mathcal{O}_{\mathfrak{r}_j}$,

$$\dim \pi^{-1}_{\overline{\mathcal{O}_{\mathfrak{r}_j}},m}(\overline{\Omega_{\mathfrak{r}_j}}) \geqslant \dim \pi^{-1}_{\overline{\mathcal{O}_{\mathfrak{r}_j}},m}(\mathcal{O}_{\mathfrak{r}_j})$$

and Proposition 5.13 applies. In both cases, by Remark 5.2(3), we conclude that $\mathcal{O}_{\mathfrak{g}} := \mathrm{Ind}_{\mathfrak{l}}^{\mathfrak{g}}(\mathcal{O}_{\mathfrak{l}})$ verifies $\mathrm{RC}_2(m)$. □

6. Consequence of Theorem 5.6

Theorem 5.6 allows us to answer the reducibility problem for many nilpotent orbits.

Recall from the beginning of Section 3 that if $\mathcal{O}$ is a nilpotent orbit of a reductive Lie algebra $\mathfrak{g}$ with simple factors $\mathfrak{s}_1, \dots, \mathfrak{s}_m$, then $\mathcal{O} = \mathcal{O}_1 \times \cdots \times \mathcal{O}_m$ where $\mathcal{O}_i$ is a nilpotent orbit of $\mathfrak{s}_i$. We shall say that $\mathcal{O}$ *has a little factor* if there exists i such that $\mathcal{O}_i$ is a little nilpotent orbit of $\mathfrak{s}_i$.

The following result is a direct consequence of Theorem 5.6 and Proposition 4.2.

Theorem 6.1. *Any nilpotent orbit induced from a nilpotent orbit that has a little factor verifies* $\mathrm{RC}_2(m)$ *for every* $m \in \mathbb{N}^*$.

When $\mathfrak{g}$ is simple, there is a unique nilpotent orbit $\mathcal{O}_{\mathrm{subreg}}$ of $\mathfrak{g}$, called the subregular nilpotent orbit, such that $\mathcal{N}(\mathfrak{g}) \setminus \mathcal{O}_{\mathrm{reg}} = \overline{\mathcal{O}_{\mathrm{subreg}}}$. It has codimension $\mathrm{rk}\,\mathfrak{g} + 2$ in $\mathfrak{g}$.

Corollary 6.2. *Assume that* $\mathfrak{g}$ *simple and not of type* A_1, $\mathrm{B}_2 = \mathrm{C}_2$, *or* G_2. *Then the subregular nilpotent orbit* $\mathcal{O}_{\mathrm{subreg}}$ *of* $\mathfrak{g}$ *verifies* $\mathrm{RC}_2(m)$ *for every* $m \in \mathbb{N}^*$. *In particular,* $\mathscr{J}_m(\overline{\mathcal{O}_{\mathrm{subreg}}})$ *is reducible for every* $m \in \mathbb{N}^*$.

Proof. Assume first that $\mathfrak{g}$ has type A_2. Then $\mathfrak{g} = \mathfrak{sl}_3(\mathbb{C})$ and $\mathcal{O}_{\mathrm{subreg}} = \mathcal{O}_{\mathrm{min}} = \mathcal{O}_{(2,1)}$. Hence, $\mathcal{O}_{\mathrm{subreg}}$ is little and verifies $\mathrm{RC}_2(m)$ for every $m \in \mathbb{N}^*$ according to Corollary 4.3.

Assume now that $\mathfrak{g}$ is simple with rank $\geqslant 3$. Then there exists a Levi subalgebra $\mathfrak{l}$ of $\mathfrak{g}$ such that $[\mathfrak{l}, \mathfrak{l}]$ is simple of type A_2, and the subregular nilpotent orbit of $\mathfrak{g}$ is induced from that of $[\mathfrak{l}, \mathfrak{l}]$ for dimension reasons (see Theorem 5.1). Therefore, the theorem follows from the case $\mathfrak{sl}_3(\mathbb{C})$ and Theorem 6.1. □

Remark 6.3. Outside types A and B, the subregular nilpotent orbit of a simple Lie algebra is distinguished. Thus Corollary 6.2 provides examples of distinguished nilpotent orbits which verify $\mathrm{RC}_2(m)$ for every $m \in \mathbb{N}^*$. In particular, according to Remark 3.8, these nilpotent orbits verify $\mathrm{RC}_2(1)$ but not RC_1.

Remark 6.4. For $\mathfrak{g}=\mathfrak{sp}_4(\mathbb{C})\simeq\mathfrak{so}_5(\mathbb{C})$, we can show that $\mathscr{J}_1(\overline{\mathcal{O}_{\text{subreg}}})$ is irreducible.

Let us detail this example where the computations are explicit. Let $\mathfrak{g}=\mathfrak{sp}_4(\mathbb{C})$. The subregular nilpotent orbit is $\mathcal{O}_{(2^2)}$. By Appendix A, it has dimension 6, and its singular locus is the union of two nilpotent orbits, $\mathcal{O}_{(2,1^2)}=\mathcal{O}_{\min}$ and the zero orbit.

Using [Weyman 2002, Theorem 1] — see also [Weyman 1989] or [Weyman 2003, Proposition 8.2.15] — and the realization of $\mathfrak{sp}_4(\mathbb{C})$ as the set of anti-self-adjoint matrices for the symplectic form, we can show that the defining ideal of $\overline{\mathcal{O}_{(2^2)}}$ is generated by the entries of the matrix X^2 as functions of $X\in\mathfrak{sp}_4(\mathbb{C})$.[3] It follows that $\mathscr{J}_1(\overline{\mathcal{O}_{(2^2)}})$ can be identified with the scheme of pairs $(X_0, X_1)\in\mathfrak{sp}_4(\mathbb{C})\times\mathfrak{sp}_4(\mathbb{C})$ defined by the equations $X_0^2=0$ and $X_0X_1+X_1X_0=0$.

Using this identification, we obtain from direct computations that

$$\dim\pi^{-1}_{\overline{\mathcal{O}_{(2^2)}},1}(\mathcal{O}_{(2,1^2)})=11 \quad\text{and}\quad \dim\pi^{-1}_{\overline{\mathcal{O}_{(2^2)}},1}(0)=10.$$

Furthermore, there are no smooth points of $\mathscr{J}_1(\overline{\mathcal{O}_{(2^2)}})$ in

$$\pi^{-1}_{\overline{\mathcal{O}_{(2^2)}},1}(\mathcal{O}_{(2,1^2)})\cup\pi^{-1}_{\overline{\mathcal{O}_{(2^2)}},1}(0).$$

To see this, we have computed the dimension of the tangent space to $\mathscr{J}_1(\overline{\mathcal{O}_{(2^2)}})$ at generic points in $\pi^{-1}_{\overline{\mathcal{O}_{(2^2)}},1}(\mathcal{O}_{(2,1^2)})$ and $\pi^{-1}_{\overline{\mathcal{O}_{(2^2)}},1}(0)$, and the smallest dimensions turn out to be 13 and 14, respectively.

Now, if $\mathscr{J}_1(\overline{\mathcal{O}_{(2^2)}})$ were reducible, it would have an irreducible component of dimension 10 or 11 by the above equalities. This is not possible according to the computations of the tangent space dimensions. Hence, $\mathscr{J}_1(\overline{\mathcal{O}_{(2^2)}})$ is irreducible.

Classical types. We now summarize our conclusions for the case where $\mathfrak{g}$ is simple of classical type. We refer to Appendix A for the notation relative to the induction of nilpotent orbits in the classical cases.

Theorem 6.5 (Type A). *Let $n\in\mathbb{N}^*$, $n\geqslant 2$, and let $\boldsymbol{\lambda}\in\mathscr{P}(n)$ be nonrectangular. Then the nilpotent orbit $\mathcal{O}_{\boldsymbol{\lambda}}$ of $\mathfrak{sl}_n(\mathbb{C})$ verifies* $\mathrm{RC}_2(m)$ *for every $m\in\mathbb{N}^*$. In particular, $\mathscr{J}_m(\overline{\mathcal{O}_{\boldsymbol{\lambda}}})$ is reducible for every $m\in\mathbb{N}^*$.*

Proof. Suppose that $\boldsymbol{\lambda}=(\lambda_1,\ldots,\lambda_r)\in\mathscr{P}(n)$ is nonrectangular, with $1<r<n$. Then there exists $1\leqslant p<r$ such that $\lambda_p>\lambda_{p+1}$. It follows that

$$\boldsymbol{\lambda}=\operatorname{Ind}^n_{(n-p-r,\,p+r)}\boldsymbol{\Lambda},$$

where

$$\boldsymbol{\Lambda}=\big((\lambda_1-2,\ldots,\lambda_p-2,\lambda_{p+1}-1,\ldots,\lambda_r-1),\,(2^p,1^{r-p})\big).$$

[3]Here, we have used the computer program `Macaulay2` to check that these equations indeed generate a reduced ideal.

Thus any nonrectangular partition of n can be induced from a partition of the form $(2^p, 1^q)$ with $p, q \in \mathbb{N}^*$. According to Example 4.4, $\mathcal{O}_{(2^p,1^q)}$ is little for $p, q \in \mathbb{N}^*$. Hence the theorem follows from Theorem 6.1. □

Remark 6.6. It is not difficult to see that rectangular partitions can only be induced from a rectangular one. So they cannot be induced from a nilpotent orbit that has a little factor (see Example 4.4).

In fact, for the rectangular case, the theorem is not true. First of all, it is obviously not true for $\boldsymbol{\lambda} = (n)$ and $\boldsymbol{\lambda} = (1^n)$. Let us look at some special cases.

(1) Let $\boldsymbol{\lambda} = (2^p)$ with $2p = n$. Then we saw in Example 3.7 that $\mathcal{O}_{\boldsymbol{\lambda}}$ is RC_1, and that all the irreducible components of $\mathscr{J}_1(\overline{\mathcal{O}_{\boldsymbol{\lambda}}})$ different from $\pi^{-1}_{\overline{\mathcal{O}_\lambda},1}(\mathcal{O}_{\boldsymbol{\lambda}})$ have codimension one. In particular, it is not $\mathrm{RC}_2(1)$.

(2) Let $\boldsymbol{\lambda} = (3^2)$. By [Weyman 2002] — see also [Weyman 1989] or [Weyman 2003, Proposition 8.2.15] — the defining ideal of $\overline{\mathcal{O}_{\boldsymbol{\lambda}}}$ is generated by $\mathrm{tr}(X^2)$ and the entries of the matrix X^3 as functions of $X \in \mathfrak{sl}_6(\mathbb{C})$. By Appendix A, the singular locus of $\overline{\mathcal{O}_{\boldsymbol{\lambda}}}$ is the finite union of the nilpotent orbits $\mathcal{O}_{\boldsymbol{\mu}}$ with

$$\mu \in \{(3, 2, 1), (3, 1^3), (2^3), (2^2, 1^2), (2, 1^4), (1^6)\} \subset \mathscr{P}(6),$$

and the respective dimensions of $\pi^{-1}_{\overline{\mathcal{O}_\lambda},1}(\mathcal{O}_{\boldsymbol{\mu}})$ are 47, 44, 44, 47, 44, 35. Note that $\mathscr{J}_1(\overline{\mathcal{O}_{\boldsymbol{\lambda}}})$ has dimension 48. Next, we obtain that the respective dimensions of the tangent space to $\mathscr{J}_1(\overline{\mathcal{O}_{\boldsymbol{\lambda}}})$ at generic points in $\pi^{-1}_{\overline{\mathcal{O}_\lambda},1}(\mathcal{O}_{\boldsymbol{\mu}})$, with $\boldsymbol{\mu}$ running through the above set, are 49, 51, 51, 48, 52, 69. Arguing as in Remark 6.4, we conclude that $\mathscr{J}_1(\overline{\mathcal{O}})$ is irreducible.

Therefore, from Remark 6.6(1) and (2), we have complete answers for the reducibility of $\mathscr{J}_1(\overline{\mathcal{O}})$ for any nilpotent orbit $\mathcal{O}$ in $\mathfrak{sl}_n(\mathbb{C})$, for $n \leqslant 7$, and for any nilpotent orbit $\mathcal{O}$ in $\mathfrak{sl}_p(\mathbb{C})$, with p a prime number.

In the other classical simple Lie algebras, we have the following result.

Theorem 6.7 (Types B, C, D). *Let $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_t) \in \mathscr{P}_\varepsilon(n)$ with $\varepsilon \in \{+1, -1\}$, and set $\lambda_{t+1} = 0$.*

(1) *Suppose that $\varepsilon = +1$ and there exist $1 \leqslant k < \ell \leqslant t$ such that $\lambda_k \geqslant \lambda_{k+1} + 2$ and $\lambda_\ell \geqslant \lambda_{\ell+1} + 2$, then the nilpotent orbit $\mathcal{O}_{\boldsymbol{\lambda}}$ of $\mathfrak{so}_n(\mathbb{C})$ verifies $\mathrm{RC}_2(m)$ for every $m \in \mathbb{N}^*$.*

(2) *Suppose that $\varepsilon = -1$ and there exist $1 \leqslant k < \ell \leqslant t$ such that $\lambda_k \geqslant \lambda_{k+1} + 2$ and $\lambda_\ell \geqslant \lambda_{\ell+1} + 2$, then the nilpotent orbit $\mathcal{O}_{\boldsymbol{\lambda}}$ of $\mathfrak{sp}_n(\mathbb{C})$ verifies $\mathrm{RC}_2(m)$ for every $m \in \mathbb{N}^*$.*

(3) *Suppose that $\varepsilon = +1$ and that $\boldsymbol{\lambda}$ is very even. Then both $\mathcal{O}^{\mathrm{I}}_{\boldsymbol{\lambda}}$ and $\mathcal{O}^{\mathrm{II}}_{\boldsymbol{\lambda}}$ verify $\mathrm{RC}_2(m)$ for every $m \in \mathbb{N}^*$. (See Appendix A for the definition of "very even".)*

In particular, $\mathscr{J}_m(\overline{\mathcal{O}_{\boldsymbol{\lambda}}})$ is reducible for every $m \in \mathbb{N}^*$.

Proof. Let $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_t) \in \mathscr{P}_\varepsilon(n)$, set $\lambda_{t+1} = 0$, and suppose that there exist $1 \leqslant k < \ell \leqslant t$ such that $\lambda_k \geqslant \lambda_{k+1} + 2$ and $\lambda_\ell \geqslant \lambda_{\ell+1} + 2$ as in the theorem. Then

$$\boldsymbol{\lambda} = \mathrm{Ind}^{n,\varepsilon}_{(\ell+k,n-2(\ell+k))} \boldsymbol{\Gamma}$$

where

$$\boldsymbol{\Gamma} := \big((2^k, 1^{\ell-k});\ (\lambda_1 - 4, \dots, \lambda_k - 4, \lambda_{k+1} - 2, \dots, \lambda_\ell - 2, \lambda_{\ell+1}, \dots, \lambda_t)\big).$$

So $\boldsymbol{\lambda}$ is induced from a partition in $\mathscr{P}(n)$ of the form $(2^p, 1^q)$, with $p, q \in \mathbb{N}^*$. By Example 4.4, the partition $(2^p, 1^q)$ is little. This concludes the proof of parts (1) and (2) according to Theorem 6.1.

Finally, if $\boldsymbol{\lambda} \in \mathscr{P}_1(n)$ is very even, then $\mathcal{O}_{\boldsymbol{\lambda}}$ is induced from the nilpotent orbit $\mathcal{O}_{(2^t)}$ of $\mathfrak{so}_{2t}(\mathbb{C})$ which is little by Example 4.5. Again, we conclude, thanks to Theorem 6.1. □

Remark 6.8. Unlike the type A case, in types B, C, D, orbits other than the ones considered in Theorem 6.7 can be induced from little ones. For example, for $\lambda, p, q \in \mathbb{N}^*$ with p even, we have $\boldsymbol{\lambda} = ((2\lambda)^p, (2\lambda-1)^q) \in \mathscr{P}_1(2\lambda(p+q)-q)$ and $\boldsymbol{\lambda}$ does not verify the conditions of Theorem 6.7. However, we have

$$\boldsymbol{\lambda} = ((2\lambda)^p, (2\lambda-1)^q) = \mathrm{Ind}^{2\lambda(p+q)-q,1}_{((\lambda-1)(p+q),2p+q)}\big((\lambda-1)^{p+q}, (2^p, 1^q)\big)$$

Since the nilpotent orbit of $\mathfrak{so}_{2p+q}(\mathbb{C})$ corresponding to the partition $(2^p, 1^q)$ is little (see Example 4.5), $\mathcal{O}_{\boldsymbol{\lambda}}$ verifies $\mathrm{RC}_2(m)$ for all $m \in \mathbb{N}^*$.

Unfortunately, in types B, C, D, we have not found a nice exhaustive description of nilpotent orbits that can be reached by induction from a little nilpotent orbit. Computations using GAP4 show that a big proportion of partitions can be induced from little ones. See Appendix B for some numerical data.

Exceptional types. Our conclusions for the exceptional types are summarized in Appendix C. More precisely, we can find in Appendix C the list of nilpotent orbits in a simple Lie algebra of exceptional type which can be induced from a little one.

7. Applications, remarks and comments

We give in this section applications to geometric properties of nilpotent orbit closures.

Nilpotent orbits closures and complete intersections. Let $\mathcal{O}$ be a nilpotent orbit of the reductive Lie algebra $\mathfrak{g}$.

Theorem 7.1. *If $\mathcal{O}$ verifies RC_1 or $\mathrm{RC}_2(m)$ for some $m \geqslant 1$, then $\bar{\mathcal{O}}$ is not a complete intersection.*

Proof. Since the singular locus of $\bar{\mathcal{O}}$ is $\bar{\mathcal{O}} \setminus \mathcal{O}$ (see Section 1), it has codimension at least two in $\bar{\mathcal{O}}$. Hence, $\bar{\mathcal{O}}$ is normal if it is a complete intersection. If so, by [Hinich 1991] or [Panyushev 1991], it has rational singularities. The theorem is then of direct consequence of Theorem 2.8. □

In [Namikawa 2013; Brion and Fu 2015], the authors use symplectic resolutions of singularities of nilpotent orbit closures to prove the above corollary for arbitrary nilpotent orbits in $\mathfrak{g}$. The foregoing provides an alternative method to obtain that result through jet schemes in a large number of cases (see Section 6). There are other approaches in the jet scheme setting to show that $\bar{\mathcal{O}}$ is not a complete intersection. Let us give an example.

Example 7.2. The computations described in Remark 6.6(2), show that for generic

$$x \in \pi^{-1}_{\overline{\mathcal{O}_{(3^2)}},1}(\mathcal{O}_{(2^2,1^2)}),$$

the tangent space at x of $\mathscr{J}_1(\overline{\mathcal{O}_{(3^2)}})$ has dimension $48 = \dim \mathscr{J}_1(\overline{\mathcal{O}_{(3^2)}})$. Hence, such an x is a smooth point of $\mathscr{J}_1(\overline{\mathcal{O}_{(3^2)}})$, because $\mathscr{J}_1(\overline{\mathcal{O}_{(3^2)}})$ is irreducible, which does not belong to $\pi^{-1}_{\overline{\mathcal{O}_{(3^2)}},1}(\mathcal{O}_{(3^2)})$. So,

$$(\mathscr{J}_1(\overline{\mathcal{O}_{(3^2)}}))_{\mathrm{reg}} \neq \pi^{-1}_{\overline{\mathcal{O}_{(3^2)}},1}(\mathcal{O}_{(3^2)})$$

and by Theorem 2.8(3), $\overline{\mathcal{O}_{(3^2)}}$ is not a complete intersection.

Unfortunately, theses arguments cannot be used for the nilpotent orbit $\mathcal{O}_{(2^2)}$ of $\mathfrak{sp}_4(\mathbb{C})$ because, in this case, the computations of Remark 6.4 show that we exactly have $(\mathscr{J}_1(\overline{\mathcal{O}_{(2^2)}}))_{\mathrm{reg}} = \pi^{-1}_{\overline{\mathcal{O}_{(2^2)}},1}(\mathcal{O}_{(2^2)})$.

Examples and counterexamples. Our results provide many examples showing that the converse of Proposition 2.5 for irreducibility is not true. Since the nilpotent cone $\mathcal{N}(\mathfrak{g})$ is normal, the following result illustrates that the converse of Proposition 2.5 for normality is also not true.

Proposition 7.3. *Assume that $\mathfrak{g}$ simple, and let $m \in \mathbb{N}$. Then $\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$ is normal if and only if $m = 0$.*

Proof. Since $\mathscr{J}_0(\mathcal{N}(\mathfrak{g})) \simeq \mathcal{N}(\mathfrak{g})$ is normal, we have to show that for any $m \in \mathbb{N}^*$, $\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$ is not normal.

Fix $m \in \mathbb{N}^*$. Let ℓ be the rank of $\mathfrak{g}$, and let $p_1, \dots, p_\ell$ be homogeneous generators of $\mathbb{C}[\mathfrak{g}]^G$ so that

$$\mathcal{N}(\mathfrak{g}) = \operatorname{Spec} \mathbb{C}[\mathfrak{g}]/(p_1, \dots, p_\ell).$$

By Remark 2.2, we get

$$\mathscr{J}_m(\mathcal{N}(\mathfrak{g})) \simeq \operatorname{Spec} \mathbb{C}[\mathfrak{g}_m]/(p_i^{(j)} \mid i = 1, \dots, \ell,\ j = 0, \dots, m).$$

Since $\mathcal{N}(\mathfrak{g})$ is a complete intersection with rational singularities, $\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$ is irreducible and reduced by Theorem 2.8. So, it is generically reduced. Furthermore,

$\big(\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))\big)_{\mathrm{reg}}$ consists of the set of $x = x_0 + x_1 t + \ldots x_m t^m \in \mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$ such that, for $i = 1, \ldots, \ell$ and $j = 0, \ldots, m$,

$$\mathrm{d}p_i^{(j)}(x_0, x_1, \ldots, x_m) \text{ are linearly independent.} \tag{2}$$

Let $x_0+x_1t+\cdots+x_mt^m \in \mathfrak{g}_m$. By [Raïs and Tauvel 1992, Lemma 3.3(i)], the vectors $\mathrm{d}p_i^{(j)}(x_0, x_1, \ldots, x_m)$ for $i \in \{1, \ldots, \ell\}$ and $j \in \{0, \ldots, m\}$ are linearly independent if and only if the vectors $\mathrm{d}p_1(x_0), \ldots, \mathrm{d}p_\ell(x_0)$ are linearly independent. But by [Kostant 1963], the later condition is satisfied if and only if x_0 is a regular element of $\mathfrak{g}$. Therefore by (2),

$$\big(\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))\big)_{\mathrm{reg}} = \pi^{-1}_{\mathcal{N}(\mathfrak{g}),m}(\mathcal{O}_{\mathrm{reg}}) \quad \text{and} \quad \big(\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))\big)_{\mathrm{sing}} = \pi^{-1}_{\mathcal{N}(\mathfrak{g}),m}(\overline{\mathcal{O}_{\mathrm{subreg}}}) \tag{3}$$

since $\mathcal{N}(\mathfrak{g}) \setminus \mathcal{O}_{\mathrm{reg}} = \overline{\mathcal{O}_{\mathrm{subreg}}}$. Then by Serre's criterion, it is enough to show that $\pi^{-1}_{\mathcal{N}(\mathfrak{g}),m}(\overline{\mathcal{O}_{\mathrm{subreg}}})$ has codimension one in $\mathscr{J}_m(\mathcal{N}(\mathfrak{g}))$, or else that

$$\dim \pi^{-1}_{\mathcal{N}(\mathfrak{g}),m}(\overline{\mathcal{O}_{\mathrm{subreg}}}) \geqslant \dim \mathscr{J}_m(\mathcal{N}(\mathfrak{g})) - 1. \tag{4}$$

The zero orbit of $\mathfrak{sl}_2(\mathbb{C})$ has codimension 2 in $\mathcal{N}(\mathfrak{sl}_2(\mathbb{C}))$. Hence, for dimension reasons, $\mathcal{O}_{\mathrm{subreg}}$ is the induced nilpotent orbit from 0 in any Levi subalgebra $\mathfrak{l}$ of $\mathfrak{g}$ with semisimple part $[\mathfrak{l}, \mathfrak{l}]$ isomorphic to $\mathfrak{sl}_2(\mathbb{C})$. So by Remark 5.14, in order to prove (4), it suffices to show the statement for $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$.

If $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$, then $\overline{\mathcal{O}_{\mathrm{subreg}}} = 0$, but by Lemma 3.5(2),

$$\begin{aligned}\dim \pi^{-1}_{\mathcal{N}(\mathfrak{sl}_2(\mathbb{C})),m}(0) &\geqslant \dim \mathscr{J}_{m-2}(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C}))) + \dim \mathfrak{sl}_2(\mathbb{C}) \\ &= 2(m-1) + 3 = 2m+1,\end{aligned}$$

whence the expected result since $\dim \mathscr{J}_m(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C}))) = 2(m+1) = 2m+2$. ☐

Remark 7.4. For $m = 1$, (3) is also a consequence of Theorem 2.8(3).

We now give an example illustrating the fact that the converse of Proposition 2.5 is also not true for reducedness.

Example 7.5. The scheme $\mathscr{J}_1(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C})))$ is irreducible and reduced. We readily obtain from the description of $\mathscr{J}_1(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C})))$ given in Example 2.3 that $\mathscr{J}_1(\mathscr{J}_1(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C}))))$ is defined by the ideal $\mathcal{J}$ of

$$\mathbb{C}[x_0, y_0, z_0, x_1, y_1, z_1, x_0', y_0', z_0', x_1', y_1', z_1']$$

generated by the polynomials

$$x_0^2 + y_0z_0, \quad 2x_0x_1 + y_0z_1 + y_1z_0,$$
$$2x_0x_0' + y_0z_0' + z_0y_0', \quad 2x_0x_1' + 2x_1x_0' + y_0z_1' + y_1z_0' + z_1y_0' + z_0y_1'.$$

A computation made with the program `Macaulay2` shows that $\mathcal{J}$ is not radical, and that the radical of $\mathcal{J}$ is the intersection of two prime ideals. So, $\mathscr{J}_1(\mathscr{J}_1(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C}))))$ is neither reduced nor irreducible.

Example 7.5 gives another piece of evidence that $\mathscr{J}_1(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C})))$ does not have rational singularities (see Proposition 7.3). Indeed, if it did, then by Theorem 2.8, $\mathscr{J}_1(\mathscr{J}_1(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C}))))$ would be irreducible (and reduced) because $\mathscr{J}_1(\mathcal{N}(\mathfrak{sl}_2(\mathbb{C})))$ is a complete intersection.

We now turn to other interesting phenomena.

Example 7.6. As has been observed in Example 3.7, for the nilpotent orbit $\mathcal{O}_{(2^p)}$ of $\mathfrak{sl}_{2p}(\mathbb{C})$, with $p \geqslant 2$, $\mathscr{J}_1(\overline{\mathcal{O}_{(2^p)}})$ is reducible and

$$\dim \pi^{-1}_{\overline{\mathcal{O}_{(2^p)}},1}((\overline{\mathcal{O}_{(2^p)}})_{\mathrm{sing}}) < \dim \pi^{-1}_{\overline{\mathcal{O}_{(2^p)}},1}(\mathcal{O}_{(2^p)}).$$

This shows that Lemma 2.7(3), does not hold in general if X is not a complete intersection.

Example 7.7. As has been observed in Remark 6.6(2), for the nilpotent orbit $\mathcal{O}_{(3^2)}$ of $\mathfrak{sl}_6(\mathbb{C})$, $\mathscr{J}_1(\overline{\mathcal{O}_{(3^2)}})$ is irreducible and

$$(\mathscr{J}_1(\overline{\mathcal{O}_{(3^2)}}))_{\mathrm{reg}} \neq \pi^{-1}_{\overline{\mathcal{O}_{(3^2)}},1}(\mathcal{O}_{(3^2)}).$$

This shows that Theorem 2.8(3) is not true for varieties that are not locally complete intersection.

Example 7.8. For the nilpotent orbit $\mathcal{O}_{(2^2)}$ of $\mathfrak{sp}_4(\mathbb{C})$, we have observed (see Remark 6.4) that

$$(\mathscr{J}_1(\overline{\mathcal{O}_{(2^2)}}))_{\mathrm{reg}} = \pi^{-1}_{\overline{\mathcal{O}_{(2^2)}},1}(\mathcal{O}_{(2^2)}).$$

This shows that the equality of Theorem 2.8(3) may hold even if X is not locally a complete intersection.

Questions and remarks. Although we have determined the reducibility of the closure of many nilpotent orbits, we would like to complete the cases where our methods do not apply. Here are some open questions.

Question 7.9. We have seen that jet schemes of nilpotent orbits in $\mathfrak{sl}_n(\mathbb{C})$ corresponding to rectangular partitions can be irreducible or reducible. Is there an explicit characterization?

Question 7.10. In all our examples of nilpotent orbits $\mathcal{O}$ with $\mathscr{J}_1(\overline{\mathcal{O}})$ reducible, the orbit $\mathcal{O}$ verifies RC_1 or $\mathrm{RC}_2(1)$. Are these conditions necessary or are there examples of $\mathcal{O}$ for which $\mathscr{J}_1(\overline{\mathcal{O}})$ is reducible and that verify neither RC_1 nor $\mathrm{RC}_2(1)$?

We have used the reducibility of jet schemes to study the property of complete intersection for nilpotent orbit closures. It is very likely that other geometric properties of nilpotent orbit closures can be studied using jet schemes.

Appendix A: Nilpotent orbits in classical simple Lie algebras

We fix in this appendix some notation and basic results, relative to nilpotent orbits in simple Lie algebras of classical type. Our main references are [Collingwood and McGovern 1993; Kempken 1983]. The results concerning the induction of nilpotent orbits are mostly taken from [loc. cit.].

Let $n \in \mathbb{N}^*$, and denote by $\mathscr{P}(n)$ the set of partitions of n. As a rule, unless otherwise specified, we write an element $\boldsymbol{\lambda}$ of $\mathscr{P}(n)$ as a decreasing sequence $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_r)$ omitting the zeroes. Thus,

$$\lambda_1 \geqslant \cdots \geqslant \lambda_r \geqslant 1 \quad \text{and} \quad \lambda_1 + \cdots + \lambda_r = n.$$

We shall denote the dual partition of a partition $\boldsymbol{\lambda} \in \mathscr{P}(n)$ by ${}^t\boldsymbol{\lambda}$. The *concatenation* of two partitions $\boldsymbol{\lambda}$ and $\boldsymbol{\lambda}'$ will be the rearrangement of the parts in decreasing order, and shall be denoted by $\boldsymbol{\lambda} \smile \boldsymbol{\lambda}'$.

Let us denote by $\geqslant$ the partial order on $\mathscr{P}(n)$ relative to dominance. More precisely, given $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_r)$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_s) \in \mathscr{P}(n)$, we have $\boldsymbol{\lambda} \geqslant \boldsymbol{\mu}$ if

$$\sum_{i=1}^{k} \lambda_i \geqslant \sum_{i=1}^{k} \mu_i$$

for $1 \leqslant k \leqslant \min(r, s)$.

***Case* $\mathfrak{sl}_n(\mathbb{C})$.** According to [Collingwood and McGovern 1993, Theorem 5.1.1], nilpotent orbits of $\mathfrak{sl}_n(\mathbb{C})$ are parametrized by $\mathscr{P}(n)$. For $\boldsymbol{\lambda} \in \mathscr{P}(n)$, we shall denote by $\mathcal{O}_{\boldsymbol{\lambda}}$ the corresponding nilpotent orbit of $\mathfrak{sl}_n(\mathbb{C})$, and if we write ${}^t\boldsymbol{\lambda} = (d_1, \dots, d_s)$, then

$$\dim \mathcal{O}_{\boldsymbol{\lambda}} = n^2 - \sum_{i=1}^{s} d_i^2.$$

Also, if $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathscr{P}(n)$, then $\mathcal{O}_{\boldsymbol{\mu}} \subset \overline{\mathcal{O}_{\boldsymbol{\lambda}}}$ if and only if $\boldsymbol{\mu} \leqslant \boldsymbol{\lambda}$.

The Levi subalgebras of $\mathfrak{sl}_n(\mathbb{C})$ are parametrized by compositions of n. Let $\boldsymbol{m} = (m_1, \dots, m_r)$ be a composition of n and

$$\boldsymbol{\lambda} = (\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(r)}) \in \mathscr{P}(m_1) \times \cdots \times \mathscr{P}(m_r).$$

It corresponds to a nilpotent orbit in the Levi subalgebra associated to the composition $\boldsymbol{m}$. Set

$$\boldsymbol{\mu} := {}^t\boldsymbol{\lambda}^{(1)} \smile \cdots \smile {}^t\boldsymbol{\lambda}^{(r)} \quad \text{and} \quad \boldsymbol{\nu} = {}^t\boldsymbol{\mu}.$$

Then the partition associated to the induced nilpotent orbit from $\mathcal{O}_{(\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(r)})}$ is $\boldsymbol{\nu}$. Note that we have $\nu_i = \lambda_i^{(1)} + \cdots + \lambda_i^{(k)}$ which is much simpler to compute in practice. We shall denote $\boldsymbol{\nu}$ by $\mathrm{Ind}_{\boldsymbol{m}}^{n}(\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(r)})$ and we shall say that $\boldsymbol{\nu}$ is *induced from* $(\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(r)})$.

Case $\mathfrak{so}_n(\mathbb{C})$. For $n \in \mathbb{N}^*$, set

$$\mathscr{P}_1(n) := \{\boldsymbol{\lambda} \in \mathscr{P}(n) \mid \text{number of parts of each even number in } \boldsymbol{\lambda} \text{ is even}\}.$$

According to [Collingwood and McGovern 1993, Theorems 5.1.2 and 5.1.4], nilpotent orbits of $\mathfrak{so}_n(\mathbb{C})$ are parametrized by $\mathscr{P}(n)$, with the exception that each *very even* partition $\boldsymbol{\lambda} \in \mathscr{P}_1(n)$ (i.e., $\boldsymbol{\lambda}$ has only even parts) corresponds to two nilpotent orbits. For $\boldsymbol{\lambda} \in \mathscr{P}_1(n)$, not very even, we shall denote by $\mathcal{O}_{\boldsymbol{\lambda}}$ the corresponding nilpotent orbit of $\mathfrak{so}_n(\mathbb{C})$. For very even $\boldsymbol{\lambda} \in \mathscr{P}_1(n)$, we shall denote by $\mathcal{O}^{\mathrm{I}}_{\boldsymbol{\lambda}}$ and $\mathcal{O}^{\mathrm{II}}_{\boldsymbol{\lambda}}$ the two corresponding nilpotent orbits of $\mathfrak{so}_n(\mathbb{C})$. In fact, their union form a single $\mathrm{O}_n(\mathbb{C})$-orbit.

Let $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_r) \in \mathscr{P}_1(n)$ and ${}^t\boldsymbol{\lambda} = (d_1, \dots, d_s)$. Then

$$\dim \mathcal{O}^{\bullet}_{\boldsymbol{\lambda}} = \frac{n(n-1)}{2} - \frac{1}{2}\left(\sum_{i=1}^{s} d_i^2 - \#\{i \mid \lambda_i \text{ odd}\}\right)$$

where $\mathcal{O}^{\bullet}_{\boldsymbol{\lambda}}$ is either $\mathcal{O}_{\boldsymbol{\lambda}}$, $\mathcal{O}^{\mathrm{I}}_{\boldsymbol{\lambda}}$, or $\mathcal{O}^{\mathrm{II}}_{\boldsymbol{\lambda}}$ according to whether $\boldsymbol{\lambda}$ is very even or not. Using the same notation, if $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathscr{P}_1(n)$, then $\overline{\mathcal{O}^{\bullet}_{\boldsymbol{\mu}}} \subsetneq \overline{\mathcal{O}^{\bullet}_{\boldsymbol{\lambda}}}$ if and only if $\boldsymbol{\mu} < \boldsymbol{\lambda}$.

Given $\boldsymbol{\lambda} \in \mathscr{P}(n)$, there exists a unique $\boldsymbol{\lambda}^+ \in \mathscr{P}_1(n)$ such that $\boldsymbol{\lambda}^+ \leqslant \boldsymbol{\lambda}$, and if $\boldsymbol{\mu} \in \mathscr{P}_1(n)$ satisfies $\boldsymbol{\mu} \leqslant \boldsymbol{\lambda}$, then $\boldsymbol{\mu} \leqslant \boldsymbol{\lambda}^+$. More precisely, let $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$ (adding zeroes if necessary). If $\boldsymbol{\lambda} \in \mathscr{P}_1(n)$, then $\boldsymbol{\lambda}^+ = \boldsymbol{\lambda}$, and if $\boldsymbol{\lambda} \notin \mathscr{P}_1(n)$, set

$$\boldsymbol{\lambda}' = (\lambda_1, \dots, \lambda_r, \lambda_{r+1} - 1, \lambda_{r+2}, \dots, \lambda_{s-1}, \lambda_s + 1, \lambda_{s+1}, \dots, \lambda_n)$$

where r is maximum such that $(\lambda_1, \dots, \lambda_r) \in \mathscr{P}_1(\lambda_1 + \dots + \lambda_r)$, and s is the index of the first even part in $(\lambda_{r+2}, \dots, \lambda_n)$. Note that $r = 0$ if such a maximum does not exist, while s is always defined. If $\boldsymbol{\lambda}'$ is not in $\mathscr{P}_1(n)$, then we repeat the process until we obtain an element of $\mathscr{P}_1(n)$ which will be our $\boldsymbol{\lambda}^+$.

The Levi subalgebras in $\mathfrak{so}_n(\mathbb{C})$ are parametrized by

$$\mathcal{L}(n) := \left\{(p_1, \dots, p_k; r) \;\middle|\; 2\sum_{i=1}^{k} p_i + r = n\right\}.$$

Let $(p_1, \dots, p_k; r) \in \mathcal{L}(n)$, $(\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(k)}) \in \mathscr{P}(p_1) \times \dots \times \mathscr{P}(p_k)$ and $\boldsymbol{\mu} \in \mathscr{P}_1(r)$, and set

$$\boldsymbol{\nu} := \mathrm{Ind}^n_{(p_1,\dots,p_k,r,p_k,\dots,p_1)}(\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(k)}, \boldsymbol{\mu}, \boldsymbol{\lambda}^{(k)}, \dots, \boldsymbol{\lambda}^{(1)})$$

in the notation of the $\mathfrak{sl}_n(\mathbb{C})$ case. Thus $\boldsymbol{\nu}$ is the partition associated to the nilpotent orbit in $\mathfrak{sl}_n(\mathbb{C})$ induced from the nilpotent orbit in the Levi subalgebra of $\mathfrak{sl}_n(\mathbb{C})$ associated to the composition $(p_1, \dots, p_k, r, p_k, \dots, p_1)$ and the multipartition $(\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(k)}, \boldsymbol{\mu}, \boldsymbol{\lambda}^{(k)}, \dots, \boldsymbol{\lambda}^{(1)})$. The partition associated to the nilpotent orbit induced from $(\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(k)}; \boldsymbol{\mu})$ is $\boldsymbol{\nu}^+$. We shall denote $\boldsymbol{\nu}^+$ by

$$\mathrm{Ind}^{n,+}_{(p_1,\dots,p_k;r)}(\boldsymbol{\lambda}^{(1)}, \dots, \boldsymbol{\lambda}^{(k)}; \boldsymbol{\mu}).$$

The partition $\boldsymbol{\lambda} \in \mathscr{P}_1(n)$ corresponds to a rigid orbit if and only if

(i) $\lambda_i - \lambda_{i+1} \leqslant 1$ for all i, so the last part of $\boldsymbol{\lambda}$ is 1;

(ii) no odd number occurs exactly twice in $\boldsymbol{\lambda}$.

Note that in the case that $\boldsymbol{\lambda}$ is a very even partition, $\boldsymbol{\nu}^+$ is also very even, and we obtain both nilpotent orbits corresponding to $\boldsymbol{\nu}^+$ via induction of the nilpotent orbits corresponding to $\boldsymbol{\lambda}$; see [Collingwood and McGovern 1993, Theorem 7.3.3(iii)].

Case $\mathfrak{sp}_{2n}(\mathbb{C})$**.** For $n \in \mathbb{N}^*$, set

$$\mathscr{P}_{-1}(2n) := \{\boldsymbol{\lambda} \in \mathscr{P}(2n) \mid \text{ number of parts of each odd number is even}\}.$$

According to [op. cit., Theorem 5.1.3], nilpotent orbits of $\mathfrak{sp}_{2n}(\mathbb{C})$ are parametrized by $\mathscr{P}_{-1}(2n)$. For $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_r) \in \mathscr{P}_{-1}(2n)$, we shall denote by $\mathscr{O}_{\boldsymbol{\lambda}}$ the corresponding nilpotent orbit of $\mathfrak{sp}_{2n}(\mathbb{C})$, and if we write ${}^t\boldsymbol{\lambda} = (d_1, \ldots, d_s)$, then

$$\dim \mathscr{O}_{\boldsymbol{\lambda}} = n(2n+1) - \frac{1}{2}\left(\sum_{i=1}^{s} d_i^2 + \#\{i \mid \lambda_i \text{ odd}\}\right).$$

As in the case of $\mathfrak{sl}_n(\mathbb{C})$, if $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathscr{P}_{-1}(2n)$, then $\mathscr{O}_{\boldsymbol{\mu}} \subset \overline{\mathscr{O}_{\boldsymbol{\lambda}}}$ if and only if $\boldsymbol{\mu} \leqslant \boldsymbol{\lambda}$.

Given $\boldsymbol{\lambda} \in \mathscr{P}(2n)$, there exists a unique $\boldsymbol{\lambda}^- \in \mathscr{P}_{-1}(2n)$ such that $\boldsymbol{\lambda}^- \leqslant \boldsymbol{\lambda}$, and if $\boldsymbol{\mu} \in \mathscr{P}_{-1}(2n)$ satisfies $\boldsymbol{\mu} \leq \boldsymbol{\lambda}$, then $\boldsymbol{\mu} \leq \boldsymbol{\lambda}^-$. The construction of $\boldsymbol{\lambda}^-$ is the same as in the orthogonal case except that s is the index of the first odd part in $(\lambda_{r+2}, \ldots, \lambda_{2n})$.

As in the orthogonal case, Levi subalgebras are parametrized by $\mathcal{L}(2n)$. Let us conserve the same notations as in the orthogonal case. The partition associated to the nilpotent orbit induced from $(\boldsymbol{\lambda}^{(1)}, \ldots, \boldsymbol{\lambda}^{(k)}; \boldsymbol{\mu})$ is $\boldsymbol{\nu}^-$. We shall denote $\boldsymbol{\nu}^-$ by

$$\operatorname{Ind}^{2n,-}_{(p_1,\ldots,p_k;r)}(\boldsymbol{\lambda}^{(1)}, \ldots, \boldsymbol{\lambda}^{(k)}; \boldsymbol{\mu}).$$

The partition $\boldsymbol{\lambda} \in \mathscr{P}_{-1}(2n)$ corresponds to a rigid orbit if and only if

(i) $\lambda_i - \lambda_{i+1} \leqslant 1$ for all i, so the last part of $\boldsymbol{\lambda}$ is 1;

(ii) no even number occurs exactly twice in $\boldsymbol{\lambda}$.

Appendix B: Statistics in types B, C, and D

As mentioned in Remark 6.8, many nilpotent orbits in $\mathfrak{so}_n(\mathbb{C})$ and $\mathfrak{sp}_{2n}(\mathbb{C})$ can be obtained by induction from little nilpotent orbits. In particular, these induced orbits verify $\mathrm{RC}_2(m)$ for all $m \in \mathbb{N}^*$. Computations using GAP4 gave us the following numerical data supporting our claim.

For $\varepsilon \in \{-1, +1\}$ and $n \in \mathbb{N}^*$, we denote by $\mathscr{P}^{\ell}_{\varepsilon}(n)$ the set of partitions in $\mathscr{P}_{\varepsilon}(n)$ induced from little ones.

Case $\mathfrak{so}_n(\mathbb{C})$.

n	$\#\mathscr{P}_1^\ell(n)$	$\#\mathscr{P}_1(n)$	n	$\#\mathscr{P}_1^\ell(n)$	$\#\mathscr{P}_1(n)$	n	$\#\mathscr{P}_1^\ell(n)$	$\#\mathscr{P}_1(n)$
2	0	1	19	111	130	36	2370	2741
3	0	2	20	130	161	37	2821	3206
4	1	3	21	170	196	38	3265	3740
5	1	4	22	195	236	39	3852	4368
6	2	5	23	250	287	40	4460	5096
7	4	7	24	291	350	41	5242	5922
8	6	10	25	367	420	42	6064	6868
9	9	13	26	423	501	43	7086	7967
10	10	16	27	527	602	44	8182	9233
11	16	21	28	609	722	45	9536	10670
12	20	28	29	751	858	46	10986	12306
13	27	35	30	869	1016	47	12748	14193
14	32	43	31	1055	1206	48	14667	16357
15	45	55	32	1223	1431	49	16974	18803
16	52	70	33	1474	1687	50	19485	21581
17	73	86	34	1710	1981	51	22464	24766
18	83	105	35	2039	2331			

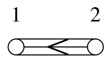
Case $\mathfrak{sp}_{2n}(\mathbb{C})$.

n	$\#\mathscr{P}_{-1}^\ell(2n)$	$\#\mathscr{P}_{-1}(2n)$	n	$\#\mathscr{P}_{-1}^\ell(2n)$	$\#\mathscr{P}_{-1}(2n)$
1	0	2	13	594	728
2	1	4	14	857	1040
3	3	8	15	1223	1472
4	9	14	16	1726	2062
5	15	24	17	2421	2864
6	28	40	18	3378	3948
7	45	64	19	4652	5400
8	77	100	20	6374	7336
9	119	154	21	8677	9904
10	182	232	22	11728	13288
11	273	344	23	15755	17728
12	409	504	24	21061	23528

Appendix C: Tables for exceptional types

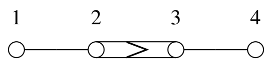
We list on the next few pages nilpotent orbits in a simple Lie algebra of exceptional type specifying when possible whether they are RC_1 or $\mathrm{RC}_2(m)$. Condition RC_1 is

***Type* G₂.**



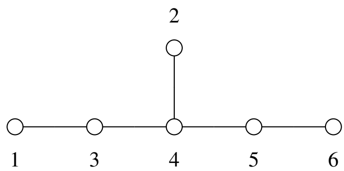
$\mathcal{O}$		$\dim \mathcal{O}$	RC_1	RC_2	rigid
A_1	[0, 1]	6	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
$\tilde{A}_1$	[1, 0]	8	$\times$	?	$\surd$
$G_2(a_1)$	[2, 0]	10	$\times$	?	$\times$

***Type* F₄.**



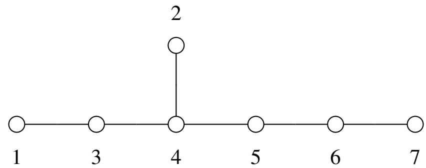
$\mathcal{O}$		$\dim \mathcal{O}$	RC_1	RC_2	rigid
A_1	[1, 0, 0, 0]	16	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
$\tilde{A}_1$	[0, 0, 0, 1]	22	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
$A_1 + \tilde{A}_1$	[0, 1, 0, 0]	28	$\times$	?	$\surd$
A_2	[2, 0, 0, 0]	30	$\times$	?	$\times$
$\tilde{A}_2$	[0, 0, 0, 2]	30	$\times$	?	$\times$
$A_2 + \tilde{A}_1$	[0, 0, 1, 0]	34	$\times$	?	$\surd$
B_2	[2, 0, 0, 1]	36	$\times$	$\surd \leftarrow \mathcal{O}_{\min}^{\{2,3,4\}}$	$\times$
$\tilde{A}_2 + A_1$	[0, 1, 0, 1]	36	$\times$	?	$\surd$
$C_3(a_1)$	[1, 0, 1, 0]	38	$\times$	$\surd \leftarrow \mathcal{O}_{\min}^{\{1,2,3\}}$	$\times$
$F_4(a_3)$	[0, 2, 0, 0]	40	$\times$	$\surd \leftarrow \mathcal{O}_{[0,1,0]}^{\{2,3,4\}}$	$\times$
B_3	[2, 2, 0, 0]	42	$\times$	?	$\times$
C_3	[1, 0, 1, 2]	42	$\times$	?	$\times$
$F_4(a_2)$	[0, 2, 0, 2]	44	$\times$	$\surd \leftarrow \mathcal{O}_{(2,1),(1^2)}^{\{1,2,4\}}$	$\times$
$F_4(a_1)$	[2, 2, 0, 2]	46	$\times$	$\surd \leftarrow \mathcal{O}_{(2,1),(2)}^{\{1,2,4\}}$	$\times$

***Type* $\mathbf{E_6}$.** The notation $\mathrm{Res}_{E_6}^{E_7}\mathcal{O}$ means that the orbit is obtained by restriction from the little nilpotent orbit $\mathcal{O}$ in E_7 as explained in Table 1.



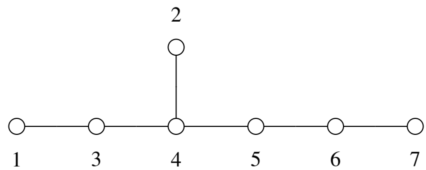
$\mathcal{O}$		$\dim\mathcal{O}$	RC_1	RC_2	rigid
A_1	$[0,1,0,0,0,0]$	22	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
$2A_1$	$[1,0,0,0,0,1]$	32	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\times$
$3A_1$	$[0,0,0,1,0,0]$	40	$\surd \leftarrow \mathrm{Res}_{E_6}^{E_7}\mathcal{O}_{(3A_1)'}$	?	$\surd$
A_2	$[0,2,0,0,0,0]$	42	$\surd \leftarrow \mathrm{Res}_{E_6}^{E_7}\mathcal{O}_{A_2}$	?	$\times$
A_2+A_1	$[1,1,0,0,0,1]$	46	$\times$	$\surd \leftarrow \mathcal{O}_{\min}^{\{1,2,3,4,5\}}$	$\times$
$2A_2$	$[2,0,0,0,0,2]$	48	$\times$	$\surd \leftarrow \mathcal{O}_{(3,1^7)}^{\{1,2,3,4,5\}}$	$\times$
A_2+2A_1	$[0,0,1,0,1,0]$	50	$\times$	?	$\times$
A_3	$[1,2,0,0,0,1]$	52	$\times$	$\surd \leftarrow \mathcal{O}_{(2,1^4)}^{\{1,3,4,5,6\}}$	$\times$
$2A_2+A_1$	$[1,0,0,1,0,1]$	54	$\times$	?	$\surd$
A_3+A_1	$[0,1,1,0,1,0]$	56	$\times$	?	$\times$
$D_4(a_1)$	$[0,0,0,2,0,0]$	58	$\times$	$\surd \leftarrow \mathcal{O}_{(2^2,1^2)}^{\{1,3,4,5,6\}}$	$\times$
A_4	$[2,2,0,0,0,2]$	60	$\times$	$\surd \leftarrow \mathcal{O}_{(3,1^3)}^{\{1,3,4,5,6\}}$	$\times$
D_4	$[0,2,0,2,0,0]$	60	$\times$	?	$\times$
A_4+A_1	$[1,1,1,0,1,1]$	62	$\times$	$\surd \leftarrow \mathcal{O}_{((2,1^2);(1^2))}^{\{1,2,3,4,6\}}$	$\times$
A_5	$[2,1,1,0,1,2]$	64	$\times$	?	$\times$
$D_5(a_1)$	$[1,2,1,0,1,1]$	64	$\times$	$\surd \leftarrow \mathcal{O}_{(3,2,1)}^{\{1,3,4,5,6\}}$	$\times$
$E_6(a_3)$	$[2,0,0,2,0,2]$	66	$\times$	$\surd \leftarrow \mathcal{O}_{(4,1^2)}^{\{1,3,4,5,6\}}$	$\times$
D_5	$[2,2,0,2,0,2]$	68	$\times$	$\surd \leftarrow \mathcal{O}_{(4,2)}^{\{1,3,4,5,6\}}$	$\times$
$E_6(a_1)$	$[2,2,2,0,2,2]$	70	$\times$	$\surd \leftarrow \mathcal{O}_{(5,1)}^{\{1,3,4,5,6\}}$	$\times$

***Type* E$_7$.** Note that the characteristics [0, 0, 0, 0, 2, 0] and [0, 0, 0, 0, 0, 2] of nilpotent orbits in D$_6$ correspond to the very even partition (2^6).



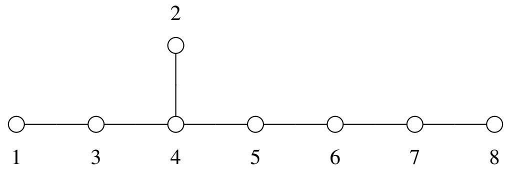
$\mathcal{O}$		$\dim\mathcal{O}$	RC_1	RC_2	rigid
A_1	$[1,0,0,0,0,0,0]$	34	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
$2A_1$	$[0,0,0,0,0,1,0]$	52	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
$(3A_1)''$	$[0,0,0,0,0,0,2]$	54	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\times$
$(3A_1)'$	$[0,0,1,0,0,0,0]$	64	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
A_2	$[2,0,0,0,0,0,0]$	66	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\times$
$4A_1$	$[0,1,0,0,0,0,1]$	70	$\times$	?	$\surd$
A_2+A_1	$[1,0,0,0,0,1,0]$	76	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,2,3,4,5,6\}}_{[0,1,0,0,0,0]}$	$\times$
A_2+2A_1	$[0,0,0,1,0,0,0]$	82	$\times$	?	$\surd$
A_3	$[2,0,0,0,0,1,0]$	84	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7\}}_{(2^2,1^8)}$	$\times$
$2A_2$	$[0,0,0,0,0,2,0]$	84	$\times$	?	$\times$
A_2+3A_1	$[0,2,0,0,0,0,0]$	84	$\times$	?	$\times$
$(A_3+A_1)''$	$[2,0,0,0,0,0,2]$	86	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7\}}_{(3,1^9)}$	$\times$
$2A_2+A_1$	$[0,0,1,0,0,1,0]$	90	$\times$	?	$\surd$
$(A_3+A_1)'$	$[1,0,0,1,0,0,0]$	92	$\times$	?	$\surd$
$D_4(a_1)$	$[0,0,2,0,0,0,0]$	94	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7\}}_{(2^4,1^4)}$	$\times$
A_3+2A_1	$[1,0,0,0,1,0,1]$	94	$\times$	?	$\times$
D_4	$[2,0,2,0,0,0,0]$	96	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7\}}_{[0,0,0,0,2,0]}$	$\times$
$D_4(a_1)+A_1$	$[0,1,1,0,0,0,1]$	96	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7\}}_{[0,0,0,0,0,2]}$	$\times$
A_3+A_2	$[0,0,0,1,0,1,0]$	98	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7\}}_{(3,2^2,1^5)}$	$\times$
A_4	$[2,0,0,0,0,2,0]$	100	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7\}}_{(3^2,1^6)}$	$\times$
$A_3+A_2+A_1$	$[0,0,0,0,2,0,0]$	100	$\times$	?	$\times$
$(A_5)''$	$[2,0,0,0,0,2,2]$	102	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7\}}_{(5,1^7)}$	$\times$

Type* E_7 *(continued). The characteristics [0, 2, 0, 0, 2, 0] and [0, 2, 0, 0, 0, 2] of nilpotent orbits in D_6 correspond to the very even partition $(4^2, 2^2)$, while the characteristics [0, 2, 0, 2, 2, 0] and [0, 2, 0, 2, 0, 2] correspond to (6^2).



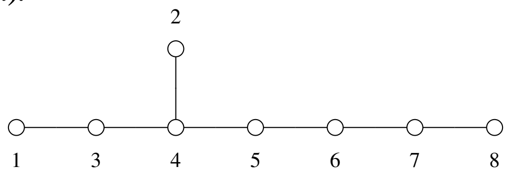
$\mathcal{O}$		$\dim\mathcal{O}$	RC_1	RC_2	rigid
D_4+A_1	$[2,1,1,0,0,0,1]$	102	$\times$	?	$\times$
A_4+A_1	$[1,0,0,1,0,1,0]$	104	$\times$	$\surd\leftarrow\mathcal{O}^{\{1,3,4,5,6,7\}}_{(2^2,1^3)}$	$\times$
$D_5(a_1)$	$[2,0,0,1,0,1,0]$	106	$\times$	$\surd\leftarrow\mathcal{O}^{\{2,3,4,5,6,7\}}_{(3^2,2^2,1^2)}$	$\times$
A_4+A_2	$[0,0,0,2,0,0,0]$	106	$\times$	?	$\times$
$(A_5)'$	$[1,0,0,1,0,2,0]$	108	$\times$	?	$\times$
A_5+A_1	$[1,0,0,1,0,1,2]$	108	$\times$	?	$\times$
$D_5(a_1)+A_1$	$[2,0,0,0,2,0,0]$	108	$\times$	$\surd\leftarrow\mathcal{O}^{\{1,2,3,4,6,7\}}_{(2,1^3),(1^3)}$	$\times$
$D_5(a_2)$	$[0,1,1,0,1,0,2]$	110	$\times$	?	$\times$
$E_6(a_3)$	$[0,0,2,0,0,2,0]$	110	$\times$	$\surd\leftarrow\mathcal{O}^{\{1,2,3,5,6,7\}}_{(2,1),(1^2),(1^4)}$	$\times$
D_5	$[2,0,2,0,0,2,0]$	112	$\times$	$\surd\leftarrow\mathcal{O}^{\{2,3,4,5,6,7\}}_{[0,2,0,0,2,0]}$	$\times$
$E_7(a_5)$	$[0,0,0,2,0,0,2]$	112	$\times$	$\surd\leftarrow\mathcal{O}^{\{2,3,4,5,6,7\}}_{[0,2,0,0,0,2]}$	$\times$
A_6	$[0,0,0,2,0,2,0]$	114	$\times$	?	$\times$
D_5+A_1	$[2,1,1,0,1,1,0]$	114	$\times$	$\surd\leftarrow\mathcal{O}^{\{1,2,3,4,6,7\}}_{(3,1^2),(1^3)}$	$\times$
$D_6(a_1)$	$[2,1,1,0,1,0,2]$	114	$\times$	$\surd\leftarrow\mathcal{O}^{\{1,2,4,5,6,7\}}_{(2),(3,1^3)}$	$\times$
$E_7(a_4)$	$[2,0,0,2,0,0,2]$	116	$\times$	$\surd\leftarrow\mathcal{O}^{\{1,3,4,5,6,7\}}_{(3^2,1)}$	$\times$
D_6	$[2,1,1,0,1,2,2]$	118	$\times$	?	$\times$
$E_6(a_1)$	$[2,0,0,2,0,2,0]$	118	$\times$	$\surd\leftarrow\mathcal{O}^{\{2,3,4,5,6,7\}}_{(5^2,1^2)}$	$\times$
E_6	$[2,0,2,2,0,2,0]$	120	$\times$	$\surd\leftarrow\mathcal{O}^{\{2,3,4,5,6,7\}}_{[0,2,0,2,2,0]}$	$\times$
$E_7(a_3)$	$[2,0,0,2,0,2,2]$	120	$\times$	$\surd\leftarrow\mathcal{O}^{\{2,3,4,5,6,7\}}_{[0,2,0,2,0,2]}$	$\times$
$E_7(a_2)$	$[2,2,2,0,2,0,2]$	122	$\times$	$\surd\leftarrow\mathcal{O}^{\{1,3,4,5,6,7\}}_{(5,2)}$	$\times$
$E_7(a_1)$	$[2,2,2,0,2,2,2]$	124	$\times$	$\surd\leftarrow\mathcal{O}^{\{1,3,4,5,6,7\}}_{(6,1)}$	$\times$

***Type* $\mathbf{E_8}$.**



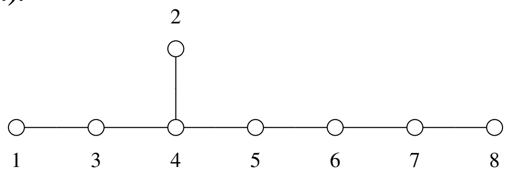
$\mathcal{O}$		$\dim\mathcal{O}$	RC_1	RC_2	rigid
A_1	$[0,0,0,0,0,0,0,1]$	58	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
$2A_1$	$[1,0,0,0,0,0,0,0]$	92	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
$3A_1$	$[0,0,0,0,0,0,1,0]$	112	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\surd$
A_2	$[0,0,0,0,0,0,0,2]$	114	$\surd \leftarrow$ little	$\surd \leftarrow$ little	$\times$
$4A_1$	$[0,1,0,0,0,0,0,0]$	128	$\times$	?	$\surd$
A_2+A_1	$[1,0,0,0,0,0,0,1]$	136	$\times$	?	$\surd$
A_2+2A_1	$[0,0,0,0,0,1,0,0]$	146	$\times$	?	$\surd$
A_3	$[1,0,0,0,0,0,0,2]$	148	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,2,3,4,5,6,7\}}_{[1,0,0,0,0,0,0]}$	$\times$
A_2+3A_1	$[0,0,1,0,0,0,0,0]$	154	$\times$	?	$\surd$
$2A_2$	$[2,0,0,0,0,0,0,0]$	156	$\times$	?	$\times$
$2A_2+A_1$	$[1,0,0,0,0,0,1,0]$	162	$\times$	?	$\surd$
A_3+A_1	$[0,0,0,0,0,1,0,1]$	164	$\times$	?	$\surd$
$D_4(a_1)$	$[0,0,0,0,0,0,2,0]$	166	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,2,3,4,5,6,7\}}_{[0,0,0,0,0,1,0]}$	$\times$
D_4	$[0,0,0,0,0,0,2,2]$	168	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,2,3,4,5,6,7\}}_{[0,0,0,0,0,0,2]}$	$\times$
$2A_2+2A_1$	$[0,0,0,0,1,0,0,0]$	168	$\times$	?	$\surd$
A_3+2A_1	$[0,0,1,0,0,0,0,1]$	172	$\times$	?	$\surd$
$D_4(a_1)+A_1$	$[0,1,0,0,0,0,1,0]$	176	$\times$	?	$\surd$
A_3+A_2	$[1,0,0,0,0,1,0,0]$	178	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(2^2,1^{10})}$	$\times$
A_4	$[2,0,0,0,0,0,0,2]$	180	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(3,1^{11})}$	$\times$
$A_3+A_2+A_1$	$[0,0,0,1,0,0,0,0]$	182	$\times$	?	$\surd$
D_4+A_1	$[0,1,0,0,0,0,1,2]$	184	$\times$	?	$\times$
$D_4(a_1)+A_2$	$[0,2,0,0,0,0,0,0]$	184	$\times$	?	$\times$
A_4+A_1	$[1,0,0,0,0,1,0,1]$	188	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,2,3,4,5,6,8\}}_{[0,1,0,0,0,0],(1^2)}$	$\times$

Type* $\mathbf{E_8}$ *(continued).



	$\mathcal{O}$	$\dim\mathcal{O}$	RC_1	RC_2	rigid
$2A_3$	$[1,0,0,0,1,0,0,0]$	188	$\times$	?	$\surd$
$D_5(a_1)$	$[1,0,0,0,0,1,0,2]$	190	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(2^4,1^6)}$	$\times$
A_4+2A_1	$[0,0,0,1,0,0,0,1]$	192	$\times$	?	$\times$
A_4+A_2	$[0,0,0,0,0,2,0,0]$	194	$\times$	?	$\times$
A_5	$[2,0,0,0,0,1,0,1]$	196	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(3,2^2,1^7)}$	$\times$
$D_5(a_1)+A_1$	$[0,0,0,1,0,0,0,2]$	196	$\times$	?	$\times$
$A_4+A_2+A_1$	$[0,0,1,0,0,1,0,0]$	196	$\times$	?	$\times$
D_4+A_2	$[0,2,0,0,0,0,0,2]$	198	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(2^6,1^2)}$	$\times$
$E_6(a_3)$	$[2,0,0,0,0,0,2,0]$	198	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(3^2,1^8)}$	$\times$
D_5	$[2,0,0,0,0,0,2,2]$	200	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(5,1^9)}$	$\times$
A_4+A_3	$[0,0,0,1,0,0,1,0]$	200	$\times$	?	$\surd$
A_5+A_1	$[1,0,0,1,0,0,0,1]$	202	$\times$	?	$\surd$
$D_5(a_1)+A_2$	$[0,0,1,0,0,1,0,1]$	202	$\times$	?	$\surd$
$D_6(a_2)$	$[0,1,1,0,0,0,1,0]$	204	$\times$	?	$\times$
$E_6(a_3)+A_1$	$[1,0,0,0,1,0,1,0]$	204	$\times$	?	$\times$
$E_7(a_5)$	$[0,0,0,1,0,1,0,0]$	206	$\times$	?	$\times$
D_5+A_1	$[1,0,0,0,1,0,1,2]$	208	$\times$	?	$\times$
$E_8(a_7)$	$[0,0,0,0,2,0,0,0]$	208	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(3^2,2^2,1^4)}$	$\times$
A_6	$[2,0,0,0,0,2,0,0]$	210	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,2,3,4,5,7,8\}}_{(3,1^7),(1^3)}$	$\times$
$D_6(a_1)$	$[0,1,1,0,0,0,1,2]$	210	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(3^2,2^4)}$	$\times$
A_6+A_1	$[1,0,0,1,0,1,0,0]$	212	$\times$	?	$\times$
$E_7(a_4)$	$[0,0,0,1,0,1,0,2]$	212	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,2,3,4,5,6,7\}}_{[0,0,0,1,0,1,0]}$	$\times$
$E_6(a_1)$	$[2,0,0,0,2,0,2]$	214	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(5,3,1^6)}$	$\times$

Type* $\mathbf{E}_8$ *(continued).



$\mathcal{O}$		$\dim \mathcal{O}$	RC_1	RC_2	rigid
$D_5 + A_2$	[0, 0, 0, 0, 2, 0, 0, 2]	214	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,3,4,5,6,7,8\}}_{(2^3,1^2)}$	$\times$
D_6	[2, 1, 1, 0, 0, 0, 1, 2]	216	$\times$	?	$\times$
E_6	[2, 0, 0, 0, 0, 2, 2, 2]	216	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(7,1^7)}$	$\times$
$D_7(a_2)$	[1, 0, 0, 1, 0, 1, 0, 1]	216	$\times$	$\surd \leftarrow \mathcal{O}^{\{1,2,4,5,6,7,8\}}_{(1^2),(2^2,1^3)}$	$\times$
A_7	[1, 0, 0, 1, 0, 1, 1, 0]	218	$\times$	?	$\times$
$E_6(a_1) + A_1$	[1, 0, 0, 1, 0, 1, 0, 2]	218	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(4^2,2^2,1^2)}$	$\times$
$E_7(a_3)$	[2, 0, 0, 1, 0, 1, 0, 2]	220	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(5,3,2^2,1)}$	$\times$
$E_8(b_6)$	[0, 0, 0, 2, 0, 0, 0, 2]	220	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(4^3,3,1^3)}$	$\times$
$D_7(a_1)$	[2, 0, 0, 0, 2, 0, 0, 2]	222	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(4^2,3^2)}$	$\times$
$E_6 + A_1$	[1, 0, 0, 1, 0, 1, 2, 2]	222	$\times$	?	$\times$
$E_7(a_2)$	[0, 1, 1, 0, 1, 0, 2, 2]	224	$\times$	?	$\times$
$E_8(a_6)$	[0, 0, 0, 2, 0, 0, 2, 0]	224	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(5,3^3)}$	$\times$
D_7	[2, 1, 1, 0, 1, 1, 0, 1]	226	$\times$	?	$\times$
$E_8(b_5)$	[0, 0, 0, 2, 0, 0, 2, 2]	226	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(5^2,2^2)}$	$\times$
$E_7(a_1)$	[2, 1, 1, 0, 1, 0, 2, 2]	228	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(7,3,2^2)}$	$\times$
$E_8(a_5)$	[2, 0, 0, 2, 0, 0, 2, 0]	228	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(5^2,3,1)}$	$\times$
$E_8(b_4)$	[2, 0, 0, 2, 0, 0, 2, 2]	230	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(6^2,1^2)}$	$\times$
E_7	[2, 1, 1, 0, 1, 2, 2, 2]	232	$\times$	?	$\times$
$E_8(a_4)$	[2, 0, 0, 2, 0, 2, 0, 2]	232	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(7,5,1^2)}$	$\times$
$E_8(a_3)$	[2, 0, 0, 2, 0, 2, 2, 2]	234	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(7^2)}$	$\times$
$E_8(a_2)$	[2, 2, 2, 0, 2, 0, 2, 2]	236	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(9,5)}$	$\times$
$E_8(a_1)$	[2, 2, 2, 0, 2, 2, 2, 2]	238	$\times$	$\surd \leftarrow \mathcal{O}^{\{2,3,4,5,6,7,8\}}_{(11,3)}$	$\times$

checked using Propositions 3.6, 4.2, 4.6, and Remark 3.9. When condition RC_1 is obtained via Proposition 4.6, we give an example of the bigger simple Lie algebra and the little nilpotent orbit satisfying condition (iii) of Proposition 4.6 from which it is obtained.

As for determining whether condition $\mathrm{RC}_2(m)$ is verified, our main method is to list the orbits induced by nilpotent orbits that have a little factor (Theorem 6.1). Thus they are $\mathrm{RC}_2(m)$ for all $m \in \mathbb{N}^*$. Since induction is transitive, we can proceed by induction on the rank of the Lie algebra, where at each step, we only need to consider induction from orbits in maximal Levi subalgebras which are themselves induced from nilpotent orbits with a little factor. For an orbit verifying condition $\mathrm{RC}_2(m)$, we give an example of a maximal Levi subalgebra $\mathfrak{l}$ and an orbit in $\mathfrak{l}$ induced from a nilpotent orbit with a little factor.

In both cases, if the orbit is little, then we just label it little. The subscript of an orbit indicates: its characteristics, the associated partition, or its Bala-Carter label. If a superscript of an orbit is present, it indicates the corresponding maximal Levi subalgebra.

We have omitted the zero orbit and the regular orbit because they are neither RC_1 nor $\mathrm{RC}_2(m)$.

All the computations are done using the package `sla` of GAP4.

Acknowledgments

We thank Michel Brion for suggesting us this topic, and bringing to our attention Namikawa's result. We are also indebted to *Mathematisches Forschungsinstitut Oberwolfach* (MFO) for its hospitality during our stay as "Research in pairs" in October, 2014.

We are grateful to Michaël Bulois for pointing out an error in a previous version, and to Jean-Yves Charbonnel and Michel Brion for their comments.

We are also thankful to the referee for careful reading and thoughtful suggestions.

This research is supported by the ANR Project GERCHER Grant number ANR-2010-BLAN-110-02.

References

[Brion and Fu 2015] M. Brion and B. Fu, "Symplectic resolutions for conical symplectic varieties", *Int. Math. Res. Not.* **2015**:12 (2015), 4335–4343. MR 3356756 Zbl 06471154

[Broer 1998] A. Broer, "Normal nilpotent varieties in F_4", *J. Algebra* **207**:2 (1998), 427–448. MR 99e:14056 Zbl 0918.17003

[Collingwood and McGovern 1993] D. H. Collingwood and W. M. McGovern, *Nilpotent orbits in semisimple Lie algebras*, Van Nostrand Reinhold, New York, 1993. MR 94j:17001 Zbl 0972.17008

[Ein and Mustaţă 2009] L. Ein and M. Mustaţă, "Jet schemes and singularities", pp. 505–546 in *Algebraic geometry, part 2* (Seattle, 2005), edited by D. Abramovich et al., Proceedings of

Symposia in Pure Mathematics **80**:2, Amer. Math. Soc., Providence, RI, 2009. MR 2010h:14004 Zbl 1181.14019

[de Fernex et al. 2013] T. de Fernex, L. Ein, and M. Mustaţă, "Vanishing theorems and singularities in birational geometry", book in preparation, 2013, http://www.math.utah.edu/∼defernex/book.pdf.

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Henderson 2014] A. Henderson, "Singularities of nilpotent orbit closures", preprint, 2014. arXiv 1408.3888v1

[Hesselink 1976] W. H. Hesselink, "Cohomology and the resolution of the nilpotent variety", *Math. Ann.* **223**:3 (1976), 249–252. MR 54 #5253 Zbl 0318.14007

[Hinich 1991] V. Hinich, "On the singularities of nilpotent orbits", *Israel J. Math.* **73**:3 (1991), 297–308. MR 92m:14005 Zbl 0766.14038

[Ishii 2011] S. Ishii, "Geometric properties of jet schemes", *Comm. Algebra* **39**:5 (2011), 1872–1882. MR 2012g:14017 Zbl 1230.14052

[Kaledin 2006] D. Kaledin, "Symplectic singularities from the Poisson point of view", *J. Reine Angew. Math.* **600** (2006), 135–156. MR 2007j:32030 Zbl 1121.53056

[Kempken 1983] G. Kempken, "Induced conjugacy classes in classical Lie algebras", *Abh. Math. Sem. Univ. Hamburg* **53** (1983), 53–83. MR 85i:17009 Zbl 0495.17003

[Kolchin 1973] E. R. Kolchin, *Differential algebra and algebraic groups*, Pure and Applied Mathematics **54**, Academic Press, New York, 1973. MR 58 #27929 Zbl 0264.12102

[Kostant 1963] B. Kostant, "Lie group representations on polynomial rings", *Amer. J. Math.* **85** (1963), 327–404. MR 28 #1252 Zbl 0124.26802

[Kraft 1989] H. Kraft, "Closures of conjugacy classes in G_2", *J. Algebra* **126**:2 (1989), 454–465. MR 91a:17020 Zbl 0693.17004

[Kraft and Procesi 1982] H. Kraft and C. Procesi, "On the geometry of conjugacy classes in classical groups", *Comment. Math. Helv.* **57**:4 (1982), 539–602. MR 85b:14065 Zbl 0511.14023

[Levasseur and Smith 1988] T. Levasseur and S. P. Smith, "Primitive ideals and nilpotent orbits in type G_2", *J. Algebra* **114**:1 (1988), 81–105. MR 89f:17013 Zbl 0644.17005

[Lusztig and Spaltenstein 1979] G. Lusztig and N. Spaltenstein, "Induced unipotent classes", *J. London Math. Soc.* (2) **19**:1 (1979), 41–52. MR 82g:20070 Zbl 0407.20035

[Mumford et al. 1994] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, 3rd ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (2) **34**, Springer, Berlin, 1994. MR 95m:14012 Zbl 0797.14004

[Mustaţă 2001] M. Mustaţă, "Jet schemes of locally complete intersection canonical singularities", *Invent. Math.* **145**:3 (2001), 397–424. MR 2002f:14005 Zbl 1091.14004

[Namikawa 2013] Y. Namikawa, "On the structure of homogeneous symplectic varieties of complete intersection", *Invent. Math.* **193**:1 (2013), 159–185. MR 3069115 Zbl 1285.14051

[Nash 1995] J. F. Nash, Jr., "Arc structure of singularities", *Duke Math. J.* **81**:1 (1995), 31–38. MR 98f:14011 Zbl 0880.14010

[Panyushev 1991] D. I. Panyushev, "Рациональность особенностей и Горенштейновость нильпотентных орбит", *Funktsional. Anal. i Prilozhen.* **25**:3 (1991), 76–78. Translated as "Rationality of singularities and the Gorenstein property of nilpotent orbits" in *Funct. Anal. Appl.* **25**:3 (1991), 225–226. MR 92i:14047 Zbl 0749.14030

[Raïs and Tauvel 1992] M. Raïs and P. Tauvel, "Indice et polynômes invariants pour certaines algèbres de Lie", *J. Reine Angew. Math.* **425** (1992), 123–140. MR 93g:17012 Zbl 0743.17008

[Richardson 1974] R. W. Richardson, Jr., "Conjugacy classes in parabolic subgroups of semisimple algebraic groups", *Bull. London Math. Soc.* **6** (1974), 21–24. MR 48 #8648 Zbl 0287.20036

[Sommers 2003] E. Sommers, "Normality of nilpotent varieties in E_6", *J. Algebra* **270**:1 (2003), 288–306. MR 2004i:20085 Zbl 1041.22009

[Tauvel and Yu 2005] P. Tauvel and R. W. T. Yu, *Lie algebras and algebraic groups*, Springer, Berlin, 2005. MR 2006c:17001 Zbl 1068.17001

[Weyman 1989] J. Weyman, "The equations of conjugacy classes of nilpotent matrices", *Invent. Math.* **98**:2 (1989), 229–245. MR 91g:20070 Zbl 0717.20033

[Weyman 2002] J. Weyman, "Two results on equations of nilpotent orbits", *J. Algebraic Geom.* **11**:4 (2002), 791–800. MR 2003k:20071 Zbl 1009.20052

[Weyman 2003] J. Weyman, *Cohomology of vector bundles and syzygies*, Cambridge Tracts in Mathematics **149**, Cambridge University Press, 2003. MR 2004d:13020 Zbl 1075.13007

[Yuen 2007] C. Yuen, "Jet schemes of determinantal varieties", pp. 261–270 in *Algebra, geometry and their interactions*, edited by A. Corso et al., Contemporary Mathematics **448**, Amer. Math. Soc., Providence, RI, 2007. MR 2009b:14093 Zbl 1142.14032

Received April 21, 2015. Revised August 3, 2015.

ANNE MOREAU
LABORATOIRE DE MATHÉMATIQUES ET APPLICATIONS
UNIVERSITE DE POITIERS
TEÉLÉPORT 2 - BP 30179
BOULEVARD MARIE ET PIERRE CURIE
86962 FUTUROSCOPE (POITIERS) CHASSENEUIL CEDEX
FRANCE

anne.moreau@math.univ-poitiers.fr

RUPERT WEI TZE YU
LABORATOIRE DE MATHÉMATIQUES DE REIMS (LMR) - EA 4535
UNIVERSITÉ DE REIMS CHAMPAGNE ARDENNE
U.F.R. SCIENCES EXACTES ET NATURELLES
MOULIN DE LA HOUSSE - BP 1039
51687 REIMS CEDEX 2
FRANCE

rupert.yu@univ-reims.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.185

COMPONENTS OF SPACES OF CURVES WITH CONSTRAINED CURVATURE ON FLAT SURFACES

NICOLAU C. SALDANHA AND PEDRO ZÜHLKE

Let S be a complete flat surface, such as the Euclidean plane. We obtain direct characterizations of the connected components of the space of all curves on S which start and end at given points in given directions, and whose curvatures are constrained to lie in a given interval, in terms of all parameters involved. Many topological properties of these spaces are investigated. Some conjectures of L. E. Dubins are proved.

0. Introduction

To abbreviate the notation, we shall identify $\mathbb{R}^2$ with $\mathbb{C}$ throughout. A curve $\gamma : [0, 1] \to \mathbb{C}$ is called *regular* if its derivative is continuous and never vanishes. Its *unit tangent* is then defined as

$$\boldsymbol{t}_\gamma : [0, 1] \to \mathbb{S}^1, \quad \boldsymbol{t}_\gamma(t) = \frac{\dot\gamma(t)}{|\dot\gamma(t)|}.$$

Lifting γ to the unit tangent bundle $UT\mathbb{C} \equiv \mathbb{C} \times \mathbb{S}^1$, we obtain its *frame*

$$\Phi_\gamma : [0, 1] \to \mathbb{C} \times \mathbb{S}^1, \quad \Phi_\gamma(t) = (\gamma(t), \boldsymbol{t}_\gamma(t)). \tag{1}$$

MSC2010: primary 53A04; secondary 53C42, 57N20.
Keywords: curve, curvature, Dubins path, flat surface, topology of infinite-dimensional manifolds.

Let $P=(p,w)$, $Q=(q,z)\in\mathbb{C}\times\mathbb{S}^1$ and consider the spaces of curves

$$\text{(2)}\qquad \begin{aligned}\mathcal{S}(P,Q)&=\{\gamma:[0,1]\xrightarrow{C^r}\mathbb{C}:\ \gamma\text{ is regular, }\Phi_\gamma(0)=P\text{ and }\Phi_\gamma(1)=Q\},\\ \Omega UT\mathbb{C}(P,Q)&=\{\omega:[0,1]\xrightarrow{C^{r-1}}UT\mathbb{C}:\ \omega(0)=P\text{ and }\omega(1)=Q\}\end{aligned}$$

endowed with the C^r and C^{r-1} topologies, respectively ($1\le r\in\mathbb{N}$). In 1956, S. Smale proved that the map

$$\Phi:\mathcal{S}(P,Q)\to\Omega UT\mathbb{C}(P,Q),\quad \gamma\mapsto\Phi_\gamma,$$

is a weak homotopy equivalence (that is, it induces isomorphisms on homotopy groups). Actually, Smale's theorem [1958, Theorem C] is much more general in that it holds for any manifold, not just $\mathbb{C}$. Using standard results on Banach manifolds which were discovered later, one can conclude that the spaces in (2) are in fact homeomorphic, and that the value of r is unimportant.

Given a regular plane curve γ, an *argument* of $\boldsymbol{t}_\gamma$ is a continuous function $\theta_\gamma:[0,1]\to\mathbb{R}$ such that $\boldsymbol{t}_\gamma=e^{i\theta_\gamma}$. The *total turning* of γ is defined to be $\theta_\gamma(1)-\theta_\gamma(0)$; note that this is independent of the choice of $\theta_\gamma(0)$. It is easy to see that $\Omega UT\mathbb{C}(P,Q)$ is homotopy equivalent to $\Omega\mathbb{S}^1(w,z)$. The latter possesses infinitely many connected components, one for each θ_1 satisfying $e^{i\theta_1}=z\bar{w}=zw^{-1}$, all of which are contractible. Therefore, the components of $\mathcal{S}(P,Q)$ are all contractible as well, and two curves in $\mathcal{S}(P,Q)$ lie in the same component if and only if they have the same total turning. This generalizes the Whitney–Graustein theorem [Whitney 1937, Theorem 1] to nonclosed curves.

The main purpose of this work is to investigate the topology of subspaces of $\mathcal{S}(P,Q)$ obtained by imposing constraints on the curvature of the curves.

(0.1) Definition. Suppose $-\infty\le\kappa_1<\kappa_2\le+\infty$ and $r\in\{2,3,\dots,\infty\}$. For $P=(p,w)$, $Q=(q,z)\in\mathbb{C}\times\mathbb{S}^1$, define $\mathcal{C}_{\kappa_1}^{\kappa_2}(P,Q)$ to be the set of all C^r regular curves $\gamma:[0,1]\to\mathbb{C}$ such that

(i) $\Phi_\gamma(0)=P$ and $\Phi_\gamma(1)=Q$;

(ii) the curvature κ_γ of γ satisfies $\kappa_1<\kappa_\gamma(t)<\kappa_2$ for each $t\in[0,1]$.

Let this set be furnished with the C^r topology.

Condition (i) means that γ starts at p in the direction of w and ends at q in the direction of z. In this notation, $\mathcal{S}(P,Q)$ becomes $\mathcal{C}_{-\infty}^{+\infty}(P,Q)$. The connected components of $\mathcal{C}_{-\kappa_0}^{+\kappa_0}(P,Q)$ ($\kappa_0>0$) were first studied by L. E. Dubins [1961]. His main result (Theorem 5.3, slightly rephrased) implies that there may exist curves with the same total turning which are not homotopic within this space.

Theorem (Dubins). *Suppose $x>0$ and $Q_x=(x,1)$, $O=(0,1)\in\mathbb{C}\times\mathbb{S}^1$. Let $\eta\in\mathcal{C}_{-1}^{+1}(O,Q_x)$ be the line segment parametrized by $\eta(t)=xt$. Then the con-*

catenation of η with a figure eight curve lies in the same connected component of $\mathcal{C}_{-1}^{+1}(O, Q_x)$ as η if and only if $x > 4$.

Here a "figure eight" curve means any closed curve of total turning 0 whose curvature takes values in $(-1, 1)$, such as the one depicted in Figure 7(d).

Naturally, we always have the following decomposition of $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$ into closed-open subspaces:

$$\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q) = \bigsqcup_{\theta_1} \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1),$$

where $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ consists of those curves in $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$ which have total turning equal to θ_1 and the union is over all $\theta_1 \in \mathbb{R}$ satisfying $e^{i\theta_1} = z\bar{w}$.

If $\kappa_1\kappa_2 \geq 0$, it will be shown that each $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ is either empty or a contractible connected component of $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$.[1] If $\kappa_1\kappa_2 < 0$, then $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ is never empty, and it is a contractible connected component provided that $|\theta_1| \geq \pi$. However, the remaining subspace $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ with $|\theta_1| < \pi$ may not be contractible, nor even connected, as implied by Dubins' theorem. It turns out that one can obtain simple and explicit characterizations of its components in terms of κ_1, κ_2, P and Q by using a homeomorphism with a space of the form $\mathcal{C}_{-1}^{+1}(P_0, Q_0; \theta_1)$ and an elementary geometric construction (see Figure 1).

This paper is close in spirit to Dubins' [1961], and some of his conjectures will be settled; it is *not* assumed, however, that the reader is familiar with his work. In the sequel to this article [Saldanha and Zühlke 2015], we determine the homotopy type of $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$. Granted the results described above, the only remaining task is the determination of the homotopy type of the exceptional subspace $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1) \subset \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$ with $|\theta_1| < \pi$ $(\kappa_1\kappa_2 < 0)$ containing the curves in the latter of least total turning. It is proved in [Saldanha and Zühlke 2015] that this subspace may be homotopy equivalent to an n-sphere for any $n \in \{0, 1, \ldots, \infty\}$ (recall that $\mathbb{S}^\infty$ is contractible). The value of n can be determined in terms of all parameters by first reducing to the case where $\kappa_1 = -1$, $\kappa_2 = +1$ through the homeomorphism mentioned above, and then using a construction extending the one depicted in Figure 1 (which only tells whether $n = 0$ or not).

Outline of the sections. Many useful constructions, such as the concatenation of elements of $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$ and $\mathcal{C}_{\kappa_1}^{\kappa_2}(Q, R)$, yield curves which need not be of class C^2. To avoid having to smoothen curves all the time, we work with curves which have a continuously varying unit tangent at all points, but whose curvatures are defined only almost everywhere. The resulting spaces, denoted by $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$, are defined in Section 1, where it will also be seen that the set inclusion $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q) \to \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ is a homotopy equivalence with dense image and that these spaces are homeomorphic.

[1] In determining the sign of $\kappa_1\kappa_2$, we adopt the convention that $0(\pm\infty) = 0$.

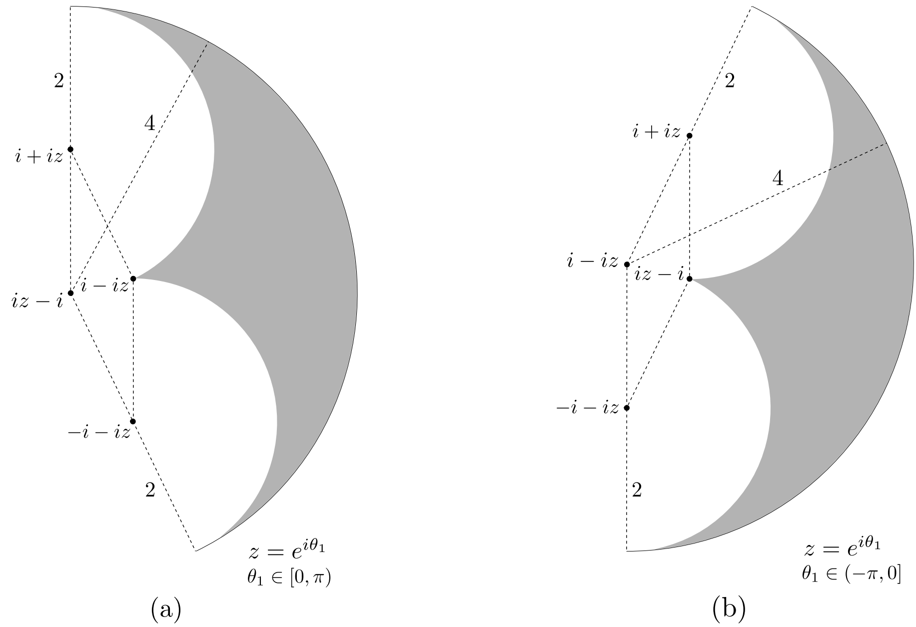


Figure 1. Let $\theta_1 \in \mathbb{R}$ be fixed, $z = e^{i\theta_1}$ and $Q = (q, z)$. Then $\mathcal{C}_{-1}^{+1}(Q; \theta_1)$ is disconnected if and only if $|\theta_1| < \pi$ and q lies in the gray region. The region contains the arc of circle of radius 4, but not the arcs of circle of radius 2. Figure (a) depicts the case $\theta_1 \in [0, \pi)$, and (b) the case $\theta_1 \in (-\pi, 0]$ (here $\theta_1 \approx \pm 26°$). The theorem of Dubins stated above corresponds to the case where $\theta_1 = 0$ and $q \in \mathbb{R}$.

Let $O = (0, 1) \in \mathbb{C} \times \mathbb{S}^1$ denote the canonical element of $UT\mathbb{C}$, and let us denote $\mathcal{C}_{\kappa_1}^{\kappa_2}(O, Q)$ simply by $\mathcal{C}_{\kappa_1}^{\kappa_2}(Q)$. Using Euclidean motions, dilatations and a construction called *normal translation* (see Figure 2 on p. 200), we obtain in (2.4) an explicit homeomorphism between any space $\mathcal{C}_{\kappa_1}^{\kappa_2}(P_0, Q_0)$ and a space of one of the following types: $\mathcal{C}_0^{+\infty}(Q)$, $\mathcal{C}_1^{+\infty}(Q)$ or $\mathcal{C}_{-1}^{+1}(Q)$, according to whether $\kappa_1\kappa_2 = 0$, $\kappa_1\kappa_2 > 0$ or $\kappa_1\kappa_2 < 0$, respectively. Moreover, this homeomorphism preserves the total turning of curves up to sign. Among these three, $\mathcal{C}_{-1}^{+1}(Q)$ has the most interesting topological properties.

We call a regular curve $\gamma : [0, 1] \to \mathbb{C}$ *condensed*, *critical* or *diffuse*, according to whether its *amplitude*

$$\omega = \sup_{t\in[0,1]} \theta_\gamma(t) - \inf_{t\in[0,1]} \theta_\gamma(t)$$

satisfies $\omega < \pi$, $\omega = \pi$ or $\omega > \pi$, respectively. Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and θ_1 be such that $e^{i\theta_1} = z$. Let $\mathcal{U}_c \subset \mathcal{C}_{-1}^{+1}(Q; \theta_1)$ (resp. $\mathcal{U}_d \subset \mathcal{C}_{-1}^{+1}(Q; \theta_1)$) denote the subspace consisting of all condensed (resp. diffuse) curves. Both are open and

$\mathcal{U}_d \neq \varnothing$, since we may always concatenate a curve in $\mathcal{C}_{-1}^{+1}(Q;\theta_1)$ with a curve of total turning 0 (an eight curve, as in Figure 7(e) on p. 218). Clearly, $\mathcal{U}_c$ must be empty if $|\theta_1| \geq \pi$, but it may also be empty otherwise, depending on Q. We determine exactly when this occurs in Section 3.

A condensed curve may be viewed as the graph of a function with respect to some axis. This leads to a direct, albeit involved, proof that $\mathcal{U}_c$ is contractible when nonempty. In fact, if the curvatures are allowed to be discontinuous and to take values in the closed interval $[-1, 1]$, then one can exhibit a contraction of the subspace of condensed curves to the unique curve of minimal length (*Dubins path*) in the corresponding space. This is also done in Section 3.

In Section 4, an indirect proof that $\mathcal{U}_d$ is contractible is obtained. If γ is diffuse, then we can "graft" straight line segments onto γ, as illustrated in Figure 8, p. 223. Such a segment can be deformed so that in the end an eight curve of large radius traversed n times has been attached to it. These eights are then spread along the curve, as in Figure 7(f). If $n \in \mathbb{N}$ is large enough, then the whole process can be carried out within $\mathcal{C}_{-1}^{+1}(Q)$. The result is a curve whose curvature is uniformly small, and hence easily deformable.

In Section 5 we determine when the set $\mathcal{T}$ of all critical curves in $\mathcal{C}_{-1}^{+1}(Q;\theta_1)$ is empty. The main result in this section is that $\mathcal{T} = \partial\mathcal{U}_c = \partial\mathcal{U}_d$. When $\mathcal{T} \neq \varnothing$, a finer analysis of how $\partial\mathcal{U}_c$ and $\partial\mathcal{U}_d$ fit together is required to determine the homeomorphism class of $\mathcal{C}_{-1}^{+1}(Q;\theta_1)$. This problem will be treated in [Saldanha and Zühlke 2015].

In (6.1) we obtain various characterizations of the connected components of $\mathcal{C}_{-1}^{+1}(Q;\theta_1)$. Perhaps the simplest one is the following: this space is disconnected if and only if $|\theta_1| < \pi$ and q lies in the region illustrated in Figure 1, or, equivalently, its subset $\mathcal{T}$ is empty, but $\mathcal{U}_c$ is not. In this case, it has exactly two components, $\mathcal{U}_c$ and $\mathcal{U}_d$, which are contractible. As mentioned previously, this is sufficient to determine explicitly the components of any space $\mathcal{C}_{\kappa_1}^{\kappa_2}(P_0, Q_0)$ with $\kappa_1\kappa_2 < 0$.

In Section 7, it is established that when $\kappa_1\kappa_2 \geq 0$, the space $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$ has one connected component for each realizable total turning, and they are all contractible. The set of possible total turnings can be described in terms of all parameters using normal translation and elementary geometry. The detailed solution to this problem is not carried out to shorten the paper, but it can be found in the earlier unpublished version [Saldanha and Zühlke 2014].

In Section 8 these results are extended to spaces of curves with constrained curvature on any complete flat surface S (orientable or not) using the fact that if S is connected then it must be the quotient of $\mathbb{C}$ by a group of isometries.

Even though we have imposed that the curvatures should lie in an open interval, the main results obtained here have analogues for spaces (defined in Section 1) where the curvature is constrained to lie in $[\kappa_1, \kappa_2]$. For $\kappa_1 = -\kappa_2$, this is the class with which Dubins actually worked [1957; 1961]. The necessary modifications in the statements and proofs are sketched in Section 9, where we also prove some conjectures appearing in [Dubins 1961] and discuss a few additional conjectures on curves of minimal length.

Related work. The problem treated here and in [Saldanha and Zühlke 2015] for flat surfaces can be generalized to any smooth (or even C^2) surface S equipped with a Riemannian metric: if u, v are elements of its unit tangent bundle UTS, then one can study the space $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ of curves on S whose lift to UTS joins u to v and whose geodesic curvature takes values in (κ_1, κ_2). When S is nonorientable, only the unsigned curvature makes sense, so in this case we require that $\kappa_2 = -\kappa_1 > 0$ (cf. Section 8 below). This topic is largely unexplored, and even the problem of determining when $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v) \neq \varnothing$ is open (and probably difficult). The topology of these spaces is very closely related to the geometry of S.

A special case which has been more intensively studied is that of the space of *nondegenerate* curves on S, i.e., curves of nonvanishing curvature. In our notation, this corresponds to $\mathcal{C}S_0^{+\infty}(u, v) \sqcup \mathcal{C}S_{-\infty}^0(u, v)$. There is also an obvious generalization to higher-dimensional manifolds, obtained by replacing the (geodesic) curvature by the generalized curvature of a curve $\gamma : [0, 1] \to M^n$. To say that the latter does not vanish is equivalent to requiring that the first n (covariant) derivatives of γ at $\gamma(t)$ span the tangent space to M at this point for each $t \in [0, 1]$. Some papers treating this problem, especially for spaces of closed curves on the simplest manifolds, such as $\mathbb{R}^n$, $\mathbb{S}^n$ or $\mathbb{RP}^n$, include [Anisov 1998; Feldman 1968; 1971, Khesin and Shapiro 1992; 1999, Little 1970; Mostovoy and Sadykov 2012; Saldanha 2015; Saldanha and Shapiro 2012; Shapiro and Shapiro 1991, Shapiro 1993]. Most of these are concerned with obtaining characterizations of the connected components of the corresponding spaces.

In [Saldanha 2015], the homotopy type of spaces of (not necessarily closed) nondegenerate curves on $\mathbb{S}^2$ is determined, and in [Saldanha and Zühlke 2013], the connected components of spaces of closed curves on $\mathbb{S}^2$ with curvature in an arbitrary interval (κ_1, κ_2) are characterized. In the sequel [Saldanha and Zühlke 2015], we determine the homotopy type of $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ for any flat surface S in terms of κ_1, κ_2 and $u, v \in UTS$. Many of the ideas appearing in the present paper (normal translation, diffuse vs. condensed, grafting, curvature spreading, etc.) appear in [Saldanha 2015] or [Saldanha and Zühlke 2013] in some form as well, although sometimes the connection is only heuristical.

1. Spaces of plane curves

Basic terminology. Let $\gamma : [a, b] \to \mathbb{C}$ be a regular curve. The *unit normal* $\boldsymbol{n} = \boldsymbol{n}_\gamma : [a, b] \to \mathbb{S}^1$ is given by $\boldsymbol{n} = i\boldsymbol{t}$, where $i \in \mathbb{C}$ denotes the imaginary unit and

$\boldsymbol{t} = \boldsymbol{t}_\gamma$ is the unit tangent to γ. The *arc-length parameter* s of γ is defined by

$$s(t) = \int_a^t |\dot{\gamma}(\tau)|\, d\tau \quad (t \in [a, b]),$$

and $L = \int_a^b |\dot{\gamma}(\tau)|\, d\tau$ is the *length* of γ. Assuming γ is twice differentiable, its *curvature* $\kappa = \kappa_\gamma$ is given by

$$\kappa(s) = \langle \boldsymbol{t}'(s), \boldsymbol{n}(s)\rangle \quad (s \in [0, L]). \tag{3}$$

In terms of a general parameter,

$$\kappa = \frac{1}{|\dot{\gamma}|}\langle \dot{\boldsymbol{t}}, \boldsymbol{n}\rangle = \frac{1}{|\dot{\gamma}|^2}\langle \ddot{\gamma}, \boldsymbol{n}\rangle = \frac{\det(\dot{\gamma}, \ddot{\gamma})}{|\dot{\gamma}|^3}. \tag{4}$$

(We denote derivatives with respect to arc-length by a $'$ (prime) and derivatives with respect to other parameters by a $\dot{}$ (dot).) Notice that the curvature at each point is not altered by an orientation-preserving reparametrization of the curve, while its sign changes if the reparametrization is orientation-reversing. It follows from (3) that if $\theta_\gamma : [0, L] \to \mathbb{R}$ is an argument of $\boldsymbol{t}$, then

$$\kappa(s) = \theta_\gamma'(s). \tag{5}$$

The following example illustrates one reason why it is more convenient to require that curvatures lie in an open interval, as in (0.1).

(1.1) Example. Consider the space of all C^2 regular curves $\gamma : [0, 1] \to \mathbb{C}$ whose curvatures are restricted to lie in $[-1, 1]$ and which satisfy $\Phi_\gamma(0) = (1, i)$, $\Phi_\gamma(1) = (i, -1)$, where we have identified $UT\mathbb{C}$ with $\mathbb{C} \times \mathbb{S}^1$. The arc α of the unit circle given by $t \mapsto \exp(\pi i t/2)$ $(t \in [0, 1])$ is a curve in this space. In fact, it is not hard to see that α is an isolated point; i.e., its connected component does not contain any other curve.

In contrast, the spaces $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)^r$ are Banach manifolds (for $r \neq \infty$). Still, some useful constructions, such as the concatenation of curves, lead out of this class of spaces. To avoid having to smoothen curves all the time, we shall work with another class of spaces, which possess the additional advantage of being Hilbert manifolds.

The group structure of $UT\mathbb{C}$. The group of all orientation-preserving isometries of $\mathbb{C}$ (i.e., proper Euclidean motions) acts simply transitively on $UT\mathbb{C}$. An element of this group is thus uniquely determined by where it maps $(0, 1) \in UT\mathbb{C}$, and may be identified with this image. Therefore, $UT\mathbb{C}$ carries a natural Lie group structure as a semidirect product $\mathbb{C} \rtimes \mathbb{S}^1$, wherein the operation is

$$(p, w) \cdot (q, z) = (p + wq, wz) \quad (p, q \in \mathbb{C},\ w, z \in \mathbb{S}^1).$$

Accordingly, viewed as a one-parameter family of Euclidean motions, the frame Φ_γ of a regular curve $\gamma : [0,1] \to \mathbb{C}$ operates on $\mathbb{C}$ through

$$\Phi_\gamma(t)a = \gamma(t) + \boldsymbol{t}(t)a \quad (a \in \mathbb{C},\ t \in [0,1]). \tag{6}$$

If we identify the Lie algebra of $UT\mathbb{C}$ with $\mathbb{C} \times \mathbb{R}$, then the bracket operation is given by

$$[(a,\theta),(b,\varphi)] = (i(\theta b - \varphi a), 0) \quad (a, b \in \mathbb{C}, \theta,\ \varphi \in \mathbb{R}). \tag{7}$$

We can also realize $UT\mathbb{C}$ as a matrix group if we identify

$$P = (p, w) \text{ with } \begin{pmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{where } p = x + iy,\ w = e^{i\theta}.$$

Then Φ_γ corresponds to the map

$$\Phi_\gamma : [0,1] \to \boldsymbol{GL}_3, \quad \Phi_\gamma(t) = \begin{pmatrix} \cos\theta_\gamma(t) & -\sin\theta_\gamma(t) & \gamma_1(t) \\ \sin\theta_\gamma(t) & \cos\theta_\gamma(t) & \gamma_2(t) \\ 0 & 0 & 1 \end{pmatrix}, \tag{8}$$

where $\theta_\gamma : [0,1] \to \mathbb{R}$ is an argument of $\boldsymbol{t}_\gamma$ and $\gamma(t) = \gamma_1(t) + i\gamma_2(t)$.[2] Moreover, under this identification the Lie algebra $\mathfrak{a}$ of $UT\mathbb{C}$ becomes a subalgebra of $\mathfrak{gl}_3$ generated by

$$A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9}$$

The expression for the bracket in (7) can be easily derived from this.

Spaces of admissible curves. Suppose now that $\gamma : [0,1] \to \mathbb{C}$ is not only regular, but also smooth. Let κ denote its curvature and $\sigma = |\dot\gamma|$ its speed. Using (5), we deduce that

$$\dot\Phi_\gamma = |\dot\gamma| \begin{pmatrix} -\kappa\sin\theta_\gamma & -\kappa\cos\theta_\gamma & \cos\theta_\gamma \\ \kappa\cos\theta_\gamma & -\kappa\sin\theta_\gamma & \sin\theta_\gamma \\ 0 & 0 & 0 \end{pmatrix} = \Phi_\gamma\Lambda_\gamma, \quad \text{where } \Lambda_\gamma = \sigma \begin{pmatrix} 0 & -\kappa & 1 \\ \kappa & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let $\mathfrak{h} \subset \mathfrak{a}$ denote the half-plane

$$\mathfrak{h} = \{aA + bB : a \in \mathbb{R}, b > 0\}. \tag{10}$$

The map $\Lambda_\gamma : [0,1] \to \mathfrak{h}$ is called the *logarithmic derivative* of γ. The crucial observation for us is that Φ_γ (and hence γ) is uniquely determined as the solution

[2]Notice that the first column of Φ_γ gives the coordinates of $\boldsymbol{t}_\gamma$, the second the coordinates of $\boldsymbol{n}_\gamma$ and the third the coordinates of γ. This justifies our terminology "frame" for Φ_γ.

of an initial value problem

$$(11)\quad \Phi(0)=P\in UT\mathbb{C},\quad \dot{\Phi}=\Phi\Lambda,\quad \text{where } \Lambda:[0,1]\to\mathfrak{h},\ \Lambda=\sigma\begin{pmatrix}0&-\kappa&1\\ \kappa&0&0\\ 0&0&0\end{pmatrix}.$$

Equivalently, γ is uniquely determined by $P=\Phi_\gamma(0)$ and the pair of functions $\kappa:[0,1]\to\mathbb{R}$ and $\sigma:[0,1]\to\mathbb{R}^+$. Our preferred class of spaces is obtained by relaxing the requirements that σ and κ be smooth.

Let $h=h_{0,+\infty}:(0,+\infty)\to\mathbb{R}$ be the smooth diffeomorphism

$$h(t)=t-t^{-1}.$$

More generally, for each pair $\kappa_1<\kappa_2\in\mathbb{R}$, let $h_{\kappa_1,\kappa_2}:(\kappa_1,\kappa_2)\to\mathbb{R}$ be the smooth diffeomorphism

$$h_{\kappa_1,\kappa_2}(t)=(\kappa_1-t)^{-1}+(\kappa_2-t)^{-1},$$

and, similarly, set

$$\begin{aligned}
&h_{-\infty,+\infty}:\mathbb{R}\to\mathbb{R}, && h_{-\infty,+\infty}(t)=t,\\
&h_{-\infty,\kappa_2}:(-\infty,\kappa_2)\to\mathbb{R}, && h_{-\infty,\kappa_2}(t)=t+(\kappa_2-t)^{-1},\\
&h_{\kappa_1,+\infty}:(\kappa_1,+\infty)\to\mathbb{R}, && h_{\kappa_1,+\infty}(t)=t+(\kappa_1-t)^{-1}.
\end{aligned}$$

(1.2) Remark. All of these functions are monotone increasing; hence so are their inverse functions. Also, if $\hat{\kappa}\in L^2[0,1]$, then $\kappa=h^{-1}_{\kappa_1,\kappa_2}\circ\hat{\kappa}\in L^2[0,1]$ as well. This is obvious if (κ_1,κ_2) is bounded, and if one of κ_1,κ_2 is infinite then it is a consequence of the fact that $h^{-1}_{\kappa_1,\kappa_2}(t)$ diverges linearly to $\pm\infty$ with respect to t.

In all that follows, $\boldsymbol{E}$ denotes the separable Hilbert space $L^2[0,1]\times L^2[0,1]$. The (i,j)-entry of a matrix A will be denoted by $A^{(i,j)}$.

(1.3) Definition. Let $-\infty\le\kappa_1<\kappa_2\le+\infty$ and $P\in UT\mathbb{C}$. A curve $\gamma:[0,1]\to\mathbb{C}$, $\gamma=\gamma_1+i\gamma_2$, will be called (κ_1,κ_2)-*admissible* if $\gamma_1=\Phi^{(1,3)}$, $\gamma_2=\Phi^{(2,3)}$ for $\Phi:[0,1]\to UT\mathbb{C}$ satisfying (11), with

$$(12)\qquad \sigma=h^{-1}\circ\hat{\sigma},\quad \kappa=h^{-1}_{\kappa_1,\kappa_2}\circ\hat{\kappa},\quad (\hat{\sigma},\hat{\kappa})\in\boldsymbol{E}.$$

When it is not important to keep track of the bounds κ_1,κ_2, we will simply say that γ is *admissible*.

The differential equation (11) has a unique solution Φ for any $(\hat{\sigma},\hat{\kappa})\in\boldsymbol{E}$ and $P\in UT\mathbb{C}$. This follows from [Younes 2010, Theorem C.3] using the fact that $\sigma,\kappa\in L^2[0,1]\subset L^1[0,1]$. Moreover, Φ is absolutely continuous [Younes 2010, p. 385] and defined over all of $[0,1]$ (since $\mathbb{C}$ is complete). The resulting maps $\boldsymbol{t}:[0,1]\to\mathbb{S}^1$, $\boldsymbol{n}:[0,1]\to\mathbb{S}^1$ and $\gamma:[0,1]\to\mathbb{C}$, obtained from the first, second and third columns of Φ, respectively, are thus absolutely continuous. It follows from (11) that

$$\dot{\gamma} = \sigma \boldsymbol{t}, \quad \dot{\boldsymbol{t}} = \sigma\kappa \boldsymbol{n} \quad \text{and} \quad \dot{\boldsymbol{n}} = -\sigma\kappa \boldsymbol{t}. \tag{13}$$

Furthermore, if Ψ denotes the 2×2 matrix obtained from Φ by discarding its third column and line, then $\Psi : [0, 1] \to \boldsymbol{SO}_2$, as one sees by differentiating $\Psi\Psi^T$, using (11) and noting that $\Psi(0) \in \boldsymbol{SO}_2$. Hence, $\boldsymbol{n} = i\boldsymbol{t}$. Differentiation of $|\boldsymbol{t}|^2$ yields that

$$|\boldsymbol{n}(t)| = |\boldsymbol{t}(t)| = |\boldsymbol{t}(0)| = 1 \quad \text{for all } t \in [0, 1].$$

Comparing with (13), it is thus natural to define $\boldsymbol{t}_\gamma = \boldsymbol{t}$, $\boldsymbol{n}_\gamma = \boldsymbol{n}$, $\Phi_\gamma = \Phi$, and to call σ and κ the *speed* and *curvature* of γ, respectively, even though $\sigma, \kappa \in L^2[0, 1]$. With this definition, $\boldsymbol{t}_\gamma$, $\boldsymbol{n}_\gamma$, Φ_γ and any argument $\theta_\gamma = \arg \circ \boldsymbol{t}_\gamma$ are absolutely continuous functions, as remarked above. Although $\dot{\gamma} = \sigma \boldsymbol{t}_\gamma$ is defined only almost everywhere on $[0, 1]$, if we reparametrize γ by arc-length then it becomes a regular curve, since $\gamma' = \boldsymbol{t}_\gamma$. Instead of thinking of γ as corresponding to a pair of L^2 functions, it is more helpful to regard γ as a regular curve whose curvature is defined only a.e. In fact, all of the concrete examples of admissible curves considered below are piecewise C^2 curves.

(1.4) Definition. Let $-\infty \le \kappa_1 < \kappa_2 \le +\infty$. For $P \in UT\mathbb{C}$, define $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, \cdot)$ to be the set of all (κ_1, κ_2)-admissible curves $\gamma : [0, 1] \to \mathbb{C}$ with $\Phi_\gamma(0) = P$. This set is identified with $\boldsymbol{E}$ via the correspondence $\gamma \leftrightarrow (\hat{\sigma}, \hat{\kappa})$, thus furnishing $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, \cdot)$ with a trivial Hilbert manifold structure.

The "$\mathcal{L}$" is intended to remind one of L^2 functions.

(1.5) Lemma. *Let $-\infty \le \kappa_1 < \kappa_2 \le +\infty$ and $P \in UT\mathbb{C}$. Then*

$$F : \mathcal{L}_{\kappa_1}^{\kappa_2}(P, \cdot) \to UT\mathbb{C}, \quad \gamma \mapsto \Phi_\gamma(1),$$

is a submersion. Consequently, it is an open map.

Proof. Let $\delta > 0$, $r \in (-\delta, \delta)$ and $\hat{\sigma}(r), \hat{\kappa}(r) \in L^2[0, 1]$ be one-parameter families of functions; set $\sigma(r) = h^{-1} \circ \hat{\sigma}(r)$, $\kappa(r) = h_{\kappa_1,\kappa_2}^{-1} \circ \hat{\kappa}(r)$. Define a corresponding family of curves $\Lambda(r) : [0, 1] \to \mathfrak{h}$ by

$$\Lambda(r) = \sigma(r) \begin{pmatrix} 0 & -\kappa(r) & 1 \\ \kappa(r) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Denoting derivatives with respect to t (resp. r) by a $\dot{}$ (resp. $'$), let the map $\Phi(r) : [0, 1] \to UT\mathbb{C}$, $t \mapsto \Phi(r)(t)$, be the solution of $\dot{\Phi}(r) = \Phi(r)\Lambda(r)$. A straightforward computation shows that

$$\Phi'(r)(t)\big(\Phi(r)(t)\big)^{-1} = \int_0^t \Phi(r)(\tau)\,\Lambda'(r)(\tau)\,\big(\Phi(r)(\tau)\big)^{-1}\,d\tau \quad (r \in (-\delta, \delta),\ t \in [0, 1]).$$

Let $\Lambda'(0)$ consist of three smooth narrow bumps at times $t = t_0$, $t = t_1$ and $t = t_2$, with each $t_i \in (0, 1)$ close to 1. Let $\Psi = \Phi(0)$; setting $r = 0$ in the previous expression, we deduce that

$$\Psi(1)^{-1}\Phi'(0)(1) \approx \sum_{i=1}^{3}\big(\Psi(t_i)^{-1}\Psi(1)\big)^{-1}\,\Lambda'(0)(t_i)\,\big(\Psi(t_i)^{-1}\Psi(1)\big).$$

Since each $\Lambda(r)$ is a curve in the open convex cone

$$\{aA + bB : a \in \mathbb{R}, b > 0 \text{ and } \kappa_1 b < a < \kappa_2 b\},$$

we can make $\Lambda'(0)(t_i)$ assume any value in the vector subspace $\mathfrak{v}$ generated by A and B (with A, B as in (9)). Another computation using the fact that $\sigma(0) > 0$ a.e. shows that the planes $\mathfrak{v}$ and $\big(\Psi(t_i)^{-1}\Psi(1)\big)^{-1}\mathfrak{v}\big(\Psi(t_i)^{-1}\Psi(1)\big)$ are transversal for small $1 - t_i$, with the angle between them proportional to $(1 - t_i) + o(1 - t_i)$. Hence, any vector in $\mathfrak{a}$ can be written in the form $\Psi(1)^{-1}\Phi'(0)(1)$ for a suitable choice of $\Lambda'(0)$, which shows that F is a submersion. □

(1.6) Definition. Let $-\infty \le \kappa_1 < \kappa_2 \le +\infty$ and $P, Q \in UT\mathbb{C}$. Define $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ to be the subspace of $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, \cdot)$ consisting of all $\gamma \in \mathcal{L}_{\kappa_1}^{\kappa_2}(P, \cdot)$ such that $\Phi_\gamma(1) = Q$.

It follows from (1.5) that $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ is a closed submanifold of codimension 3 in $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, \cdot) \equiv \boldsymbol{E}$; the proof that $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ is always nonempty is postponed until Section 4.

The following lemmas contain all the results on infinite-dimensional manifolds that we shall use.

(1.7) Lemma. *Let $\mathcal{M}$, $\mathcal{N}$ be (infinite-dimensional) separable Banach manifolds. Then:*

(a) *$\mathcal{M}$ is locally path-connected. In particular, its connected and path components coincide.*

(b) *If $F : \mathcal{M} \to \mathcal{N}$ is a weak homotopy equivalence, then F is homotopic to a homeomorphism.*

(c) *Let $\boldsymbol{E}$ and $\boldsymbol{F}$ be separable Banach spaces. Suppose $i : \boldsymbol{F} \to \boldsymbol{E}$ is a bounded, injective linear map with dense image and $\mathcal{M} \subset \boldsymbol{E}$ is a smooth closed submanifold of finite codimension. Then $\mathcal{N} = i^{-1}(\mathcal{M})$ is a smooth closed submanifold of $\boldsymbol{F}$ and $i : \mathcal{N} \to \mathcal{M}$ is a homotopy equivalence.*

Proof. Part (a) is obvious. Part (b) follows from [Palais 1966, Theorem 15], [Burghelea and Kuiper 1969, Theorem 9] and [Henderson 1969, Corollary 3]. Part (c) is [Burghelea et al. 2003, Theorem 2]. □

(1.8) Lemma. *Let $\boldsymbol{E}$ be a separable Hilbert space, $D \subset \boldsymbol{E}$ a dense vector subspace, $L \subset \boldsymbol{E}$ a submanifold of finite codimension and U an open subset of L. If K is a finite simplicial complex and $f : |K| \to U$ a continuous map, then f is homotopic within U to a map $|K| \to D \cap U$.*

Proof. See [Saldanha and Zühlke 2013, Lemma 1.10]. □

(1.9) Corollary. *Let $\kappa_1 < \kappa_2$ and $P, Q \in UT\mathbb{C}$. Then the subset of all smooth curves in $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ is dense in the latter.*

Proof. Take $\boldsymbol{E} = L^2[0, 1] \times L^2[0, 1]$, $D = C^\infty[0, 1] \times C^\infty[0, 1]$ and U an open subset of $L = \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$. Then it is a trivial consequence of (1.8) that $D \cap U \neq \varnothing$ if $U \neq \varnothing$. □

(1.10) Lemma. *Let $(\kappa_1, \kappa_2) \subset (\bar{\kappa}_1, \bar{\kappa}_2)$ and $P, Q \in UT\mathbb{C}$. Then*

$$j : \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)^r \to \mathcal{L}_{\bar{\kappa}_1}^{\bar{\kappa}_2}(P, Q), \quad \gamma \mapsto (\hat{\sigma}, \hat{\kappa}), \tag{14}$$

where $\hat{\sigma} = h \circ |\dot{\gamma}|$ and $\hat{\kappa} = h_{\bar{\kappa}_1, \bar{\kappa}_2} \circ \kappa_\gamma$, is a continuous injection for all $r \geq 2$. Furthermore, the actual curve in $\mathbb{C}$ corresponding to $j(\gamma) \in \mathcal{L}_{\bar{\kappa}_1}^{\bar{\kappa}_2}(P, Q)$ is γ itself.

Proof. The curve corresponding to the right side of (14) in $\mathcal{L}_{\bar{\kappa}_1}^{\bar{\kappa}_2}(P, \cdot)$ is the solution of (11) with

$$\sigma = h^{-1} \circ \hat{\sigma} = |\dot{\gamma}| \quad \text{and} \quad \kappa = h_{\bar{\kappa}_1, \bar{\kappa}_2}^{-1} \circ \hat{\kappa} = \kappa_\gamma.$$

By uniqueness, this solution must equal γ. In particular, j is injective and its image is indeed contained in $\mathcal{L}_{\bar{\kappa}_1}^{\bar{\kappa}_2}(P, Q)$. Continuity of j is clear: if η is C^r-close to γ, then σ_η (resp. κ_η) is C^1-close (resp. C^0-close) to σ_γ (resp. κ_γ); hence $j(\eta)$ is close to $j(\gamma)$ in the L^2-norm. □

(1.11) Corollary. *Let $\kappa_1 < \kappa_2$, $P, Q \in UT\mathbb{C}$ and $\mathcal{U} \subset \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ be open. Let K be a finite simplicial complex and $f : |K| \to \mathcal{U}$ be a continuous map. Then there exists a continuous $g : |K| \to \mathcal{U}$ such that*

(i) *$f \simeq g$ within $\mathcal{U}$;*

(ii) *$g(a)$ is a smooth curve for all $a \in K$;*

(iii) *all derivatives of $g(a)$ with respect to t depend continuously on $a \in K$.*

In particular, the set inclusion $j : \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q) \hookrightarrow \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ in (14) induces surjections $\pi_k(j^{-1}(\mathcal{U})) \to \pi_k(\mathcal{U})$ for all $k \in \mathbb{N}$.

Proof. Parts (i), (ii) follow immediately from (1.8) by setting $\boldsymbol{E} = L^2[0, 1] \times L^2[0, 1]$, $D = C^\infty[0, 1] \times C^\infty[0, 1]$, $L = \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ and $U = \mathcal{U}$. The image of the function $g = H_2 : |K| \to \mathcal{U}$ constructed in the proof of (1.8) [Saldanha and Zühlke 2013, Lemma 1.10] is contained in a finite-dimensional vector subspace of D, namely, the one generated by all $\tilde{v}_{ij}$, so (iii) also holds. $\square$

(1.12) Lemma. *Let $\kappa_1 < \kappa_2$ and $P, Q \in UT\mathbb{C}$. Then the inclusion*

$$j : \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)^r \to \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$$

of (14) is a homotopy equivalence for any $r \in \mathbb{N}$, $r \geq 2$. Consequently, $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)^r$ is homeomorphic to $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ for any $r \in \mathbb{N}$, $r \geq 2$.

Proof. Let $\boldsymbol{E} = L^2[0, 1] \times L^2[0, 1]$, let $\boldsymbol{F} = C^{r-1}[0, 1] \times C^{r-2}[0, 1]$ (where $C^k[0, 1]$ denotes the set of all C^k functions $[0, 1] \to \mathbb{R}$ with the C^k norm) and let $i : \boldsymbol{F} \to \boldsymbol{E}$ be set inclusion. Setting $\mathcal{M} = \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$, we conclude from (1.7)(c) that $i : \mathcal{N} = i^{-1}(\mathcal{M}) \hookrightarrow \mathcal{M}$ is a homotopy equivalence. We claim that $\mathcal{N}$ is homeomorphic to $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)^r$, where the homeomorphism G is obtained by associating a pair $(\hat{\sigma}, \hat{\kappa}) \in \mathcal{N}$ to the curve γ obtained by solving (11), with σ and κ as in (12). The lemma will follow from this and the easily verified commutativity of

$$\begin{tikzcd} \mathcal{N} \arrow[r, "G"] \arrow[dr, "i"'] & \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)^r \arrow[d, "j"] \\ & \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \end{tikzcd}$$

Suppose first that $\gamma \in \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)^r$. Then $|\dot{\gamma}|$ (resp. κ) is a function $[0, 1] \to \mathbb{R}$ of class C^{r-1} (resp. C^{r-2}). Hence, so are $\hat{\sigma} = h \circ |\dot{\gamma}|$ and $\hat{\kappa} = h_{\kappa_1}^{\kappa_2} \circ \kappa$, since h and $h_{\kappa_1}^{\kappa_2}$ are smooth. Moreover, if $\gamma, \eta \in \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)^r$ are close in the C^r topology, then $\hat{\kappa}_\gamma$ is C^{r-2}-close to $\hat{\kappa}_\eta$ and $\hat{\sigma}_\gamma$ is C^{r-1}-close to $\hat{\sigma}_\eta$.

Conversely, if $(\hat{\sigma}, \hat{\kappa}) \in \mathcal{N}$, then $\sigma = h^{-1} \circ \hat{\sigma}$ is of class C^{r-1} and $\kappa = (h_{\kappa_1}^{\kappa_2})^{-1} \circ \hat{\kappa}$ is of class C^{r-2}. Since all functions on the right side of (11) are of class (at least) C^{r-2}, the solution $\boldsymbol{t} = \boldsymbol{t}_\gamma$ to this initial value problem is of class C^{r-1}. Moreover, $\dot{\gamma} = \sigma \boldsymbol{t}$; hence the velocity vector of γ is seen to be of class C^{r-1}. We conclude that γ is a curve of class C^r. Further, continuous dependence on the parameters of a differential equation shows that the correspondence $(\hat{\sigma}, \hat{\kappa}) \mapsto \boldsymbol{t}_\gamma$ is continuous. Since γ is obtained by integrating $\sigma \boldsymbol{t}_\gamma$, we deduce that the map $(\hat{\sigma}, \hat{\kappa}) \mapsto \gamma$ is likewise continuous.

The last assertion of the lemma follows from (1.7)(b). $\square$

(1.13) Definition. Let $P = (p, w)$, $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$. Given $\theta_1 \in \mathbb{R}$ satisfying $e^{i\theta_1} = z\bar{w}$, we denote by $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ the subspace of $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ consisting of all curves which have total turning equal to θ_1. When $P = (0, 1)$, the space $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ (resp. $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$) will be denoted simply by $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q)$ (resp. $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q; \theta_1)$).

Notice that $(0, 1) \in \mathbb{C} \times \mathbb{S}^1$ corresponds to the identity element in the group structure of $UT\mathbb{C}$. It will be proved in Section 4 that $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ is never empty if $\kappa_1\kappa_2 < 0$, but may be empty if $\kappa_1\kappa_2 \geq 0$, depending on the value of θ_1.

The next two results let us reparametrize a family of curves to better suit our needs.

(1.14) Lemma. *Let $\mathcal{M} = \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ or $\mathcal{M} = \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$. Let A be a topological space and $A \to \mathcal{M}$, $a \mapsto \gamma_a$, be a continuous map. Then there exists a homotopy $\gamma_a^r : [0, 1] \to \mathcal{M}$, $r \in [0, 1]$, such that for any $a \in A$,*

(i) *$\gamma_a^0 = \gamma_a$ and γ_a^1 is parametrized so that $|\dot{\gamma}_a^1(t)|$ is independent of t;*

(ii) *γ_a^r is an orientation-preserving reparametrization of γ_a for all $r \in [0, 1]$.*

Proof. Let $s_a(t) = \int_0^t |\dot{\gamma}_a(\tau)|\, d\tau$ be the arc-length parameter of γ_a, L_a its length and $\tau_a : [0, L_a] \to [0, 1]$ the inverse function of s_a. Define $\gamma_a^r : [0, 1] \to M$ by

$$\gamma_a^r(t) = \gamma_a\big((1-r)t + r\tau_a(L_a t)\big) \quad (r, t \in [0, 1],\ a \in A).$$

Then γ_a^r is the desired homotopy. □

(1.15) Corollary. *Let $\mathcal{M} = \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ or $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$. Let A be a topological space and $f : \mathbb{S}^0 \times A \to \mathcal{M}$ a continuous map such that for all $a \in A$, $f(1, a)$ is an orientation-preserving reparametrization of $f(-1, a)$. Then f admits a continuous extension $F : A \times [-1, 1] \to \mathcal{M}$ with the property that $f(r, a)$ is an orientation-preserving reparametrization of $f(-1, a)$ for all $r \in [-1, 1]$, $a \in A$.* □

It is assumed in (1.14) and (1.15) that the parametrizations of all curves have domain $[0, 1]$, but it is clearly possible to have these intervals depend (continuously) on a. A typical application is to reparametrize all curves in a family by arc-length, not just proportionally to arc-length as in (1.14).

Spaces of curves with curvature in a closed interval.

(1.16) Definition. Let $P, Q \in UT\mathbb{C}$ and $-\infty < \kappa_1 < \kappa_2 < +\infty$. Define $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$ to be the set of all C^1 regular plane curves $\gamma : [0, 1] \to \mathbb{C}$ satisfying

(i) $\Phi_\gamma(0) = P$ and $\Phi_\gamma(1) = Q$;

(ii)
$$\kappa_1 \leq \frac{\theta(s_1) - \theta(s_2)}{s_1 - s_2} \leq \kappa_2$$

for any $s_1 \neq s_2 \in [0, L]$. (Here the parameter is the arc-length of γ, L is its length and $\theta : [0, L] \to \mathbb{R}$ is an argument of $\boldsymbol{t}_\gamma$.)

Condition (ii) implies that θ is a Lipschitz function. In particular, it is absolutely continuous, and its derivative κ_γ lies in L^2, since it is bounded. We give this set the topology induced by the following distance function d: given $\gamma, \eta \in \hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$, set

$$d(\gamma, \eta) = \|\gamma - \eta\|_2 + \|\dot{\gamma} - \dot{\eta}\|_2 + \|\kappa_\gamma - \kappa_\eta\|_2.$$

For $P = (0, 1) \in \mathbb{C} \times \mathbb{S}^1$, we will denote $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$ simply by $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(Q)$.

Remark. This definition is essentially due to L. E. Dubins, who studied paths of minimal length, now called *Dubins paths*, in $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(P, Q)$ ($\kappa_0 > 0$). Such shortest paths always exist, but may not be unique in some special cases (see Proposition 1 and the corollary to Theorem I of [Dubins 1957]). His main result states that any Dubins path is the concatenation of at most three pieces, each of which is either a line segment or an arc of circle of radius $1/\kappa_0$ (see [loc. cit., Theorem I] for the precise statement). Dubins paths and variations thereof have many applications in engineering and are the subject of a vast literature. The space $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$ will play a minor role in our investigations. Its topology has been chosen to ensure that the following result holds.

(1.17) Lemma. *Let* $(\kappa_1, \kappa_2) \subset [\bar{\kappa}_1, \bar{\kappa}_2] \subset (\bar{\bar{\kappa}}_1, \bar{\bar{\kappa}}_2)$ *and* $P, Q \in UT\mathbb{C}$. *Then the set inclusions*

$$\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q) \to \hat{\mathcal{L}}_{\bar{\kappa}_1}^{\bar{\kappa}_2}(P, Q) \quad \text{and} \quad \hat{\mathcal{L}}_{\bar{\kappa}_1}^{\bar{\kappa}_2}(P, Q) \to \mathcal{L}_{\bar{\bar{\kappa}}_1}^{\bar{\bar{\kappa}}_2}(P, Q)$$

are continuous injections.

Proof. The proof is a straightforward verification, which will be left to the reader. □

2. Normal translation

The *radius of curvature* ρ of an admissible curve γ is given by $\rho = 1/\kappa$; when $\kappa(t) = 0$, it is to be understood that $\rho(t) = \infty$ (unsigned infinity). An analogue of the following construction has already appeared in [Saldanha and Zühlke 2013]. It can be used to uniformly shift the radii of curvature of a family of curves.

(2.1) Definition. Let $\gamma : [0, 1] \to \mathbb{C}$ be admissible and $u \in \mathbb{R}$. The *normal translation* γ_u of γ by u is the curve given by

$$\gamma_u(t) = \gamma(t) + u\boldsymbol{n}(t) \quad (t \in [0, 1]).$$

Observe that the normal translation α_u of a circle α of radius of curvature $\rho \in \mathbb{R} \setminus \{0\}$ is a circle of radius of curvature $\rho - u$ for any u in the component of $\mathbb{R} \setminus \{\rho\}$ containing 0 (see Figure 2). The following lemma generalizes this to arbitrary curves.

(2.2) Lemma. *Let* $\gamma \in \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ *be parametrized proportionally to arc-length and let* $\boldsymbol{t}, \kappa, \rho$ *denote its unit tangent, curvature and radius of curvature, respectively. Suppose* $u \in \mathbb{R}$ *satisfies* $1 - uk > 0$ *for all* $k \in (\kappa_1, \kappa_2)$ *and set*

$$\bar{\kappa}_i = \frac{\kappa_i}{1 - u\kappa_i} \quad (i = 1, 2). \tag{15}$$

Then the normal translation γ_u *of* γ *by* u *has the following properties*:

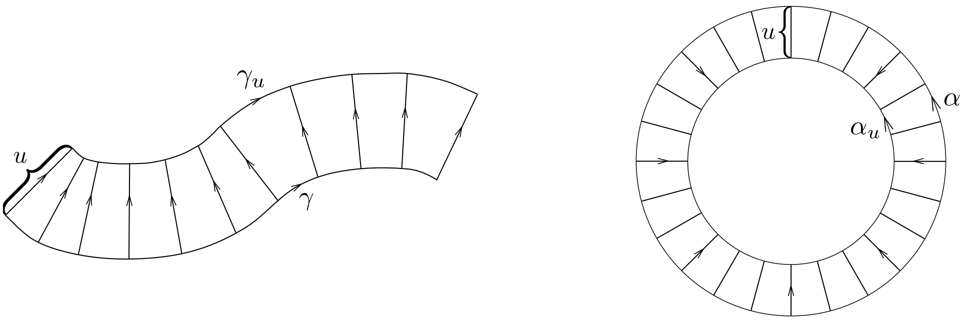


Figure 2. The normal translation of a general curve γ and of a circle α.

(a) *$\gamma_u \in \mathcal{L}_{\bar\kappa_1}^{\bar\kappa_2}(\bar P, \bar Q)$ for $\bar P = (p + iuw, w)$, $\bar Q = (q + iuz, z)$ and its unit tangent $\bar{\boldsymbol{t}}$ satisfies $\bar{\boldsymbol{t}}(t) = \boldsymbol{t}(t)$ for each $t \in [0, 1]$. In particular, γ and γ_u have the same total turning.*

(b) $(\gamma_u)_{-u} = \gamma$.

(c) *If η is a reparametrization of γ, then η_u is a reparametrization of γ_u.*

(d) *For almost every $t \in [0, 1]$, the curvature $\bar\kappa$ of γ_u is given by*

$$\bar\kappa(t) = \frac{\kappa(t)}{1 - u\kappa(t)}$$

and its radius of curvature $\bar\rho$ by

$$\bar\rho(t) = \rho(t) - u.$$

In (15) above, it should be understood that $\bar\kappa_i = -1/u$ if κ_i is infinite and that $\bar\kappa_i = \pm\infty$ has the same sign as κ_i if $1 - u\kappa_i = 0$.

Proof. Let $\theta_\gamma : [0, 1] \to \mathbb{R}$ be an argument of $\boldsymbol{t} = \boldsymbol{t}_\gamma$ and define $\Psi : [0, 1] \to \boldsymbol{GL}_3$ by

$$\Psi = \begin{pmatrix} \cos\theta_\gamma & -\sin\theta_\gamma & \gamma_1 - u\sin\theta_\gamma \\ \sin\theta_\gamma & \cos\theta_\gamma & \gamma_2 + u\cos\theta_\gamma \\ 0 & 0 & 1 \end{pmatrix}. \tag{16}$$

Let L be the length of γ. Since γ is parametrized proportionally to arc-length, a straightforward calculation shows that Ψ satisfies $\dot\Psi = \Psi\Lambda$ for

$$\Lambda : [0, 1] \to \mathfrak{a} \subset \mathfrak{gl}_3, \quad \Lambda = L \begin{pmatrix} 0 & -\kappa & 1 - u\kappa \\ \kappa & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{17}$$

By hypothesis, the image of Λ is contained in the half-plane $\mathfrak{h}$ of (10). Comparing the third column of (16) with the definition of γ_u, we deduce that Ψ is the frame of γ_u. Further, looking at the first and second columns, we deduce that $\bar{\boldsymbol{t}} = \boldsymbol{t}$ and

$\bar{\boldsymbol{n}} = \boldsymbol{n}$. That $\Phi_{\gamma_u}(0) = \bar{P}$ and $\Phi_{\gamma_u}(1) = \bar{Q}$ then follows immediately from the definition. This establishes (a) except for the fact that γ_u is $(\bar{\kappa}_1, \bar{\kappa}_2)$-admissible, which will be proved below.

Part (b) is an easy verification:

$$(\gamma_u)_{-u} = \gamma_u - u\bar{\boldsymbol{n}} = (\gamma + u\boldsymbol{n}) - u\boldsymbol{n} = \gamma.$$

Part (c) is obvious.

We know that the curvature $\bar{\kappa}$ of γ_u is given by the quotient of $\Lambda^{(2,1)}$ by $\Lambda^{(1,3)}$, that is,

$$\bar{\kappa} = \frac{\kappa}{1 - u\kappa} = \frac{1}{\rho - u} = \frac{1}{\bar{\rho}}.$$

This proves (d).

It is straightforward to check that $u \in \mathbb{R}$ satisfies $1 - uk > 0$ for all $k \in (\kappa_1, \kappa_2)$ if and only if u lies in the maximal closed interval J containing 0 and not containing any number of the form $1/k$ for $k \in (\kappa_1, \kappa_2)$. More explicitly:

(i) If $0 \le \kappa_1 < \kappa_2$ then $J = (-\infty, \rho_2]$;

(ii) If $\kappa_1 < 0 < \kappa_2$ then $J = [\rho_1, \rho_2]$;

(iii) If $\kappa_1 < \kappa_2 \le 0$ then $J = [\rho_1, +\infty)$.

By (17), $|\dot{\gamma}_u| = L(1 - u\kappa)$. To establish that γ_u is $(\bar{\kappa}_1, \bar{\kappa}_2)$-admissible, it suffices to prove that

$$h_{0,+\infty} \circ (1 - u\kappa) = (1 - u\kappa) - (1 - u\kappa)^{-1} \in L^2[0, 1], \tag{18}$$

$$h_{\bar{\kappa}_1,\bar{\kappa}_2} \circ \bar{\kappa} \in L^2[0, 1]. \tag{19}$$

By (1.2), $\kappa \in L^2[0, 1]$; hence so is $(1 - u\kappa)$. Moreover, $(1 - u\kappa)^{-1}$ is bounded unless u is one of the endpoints of J, but we claim that even in this case $(1 - u\kappa)^{-1} \in L^2[0, 1]$. Suppose for concreteness that $u = \rho_2 \in \partial J$. If $\kappa_2 = +\infty$ ($\rho_2 = 0$) then there is nothing to prove, and otherwise

$$(1 - u\kappa)^{-1} = (1 - \rho_2\kappa)^{-1} = \kappa_2(\kappa_2 - \kappa)^{-1}. \tag{20}$$

Now by hypothesis, $\gamma \in \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$; therefore

$$h_{\kappa_1,\kappa_2} \circ \kappa = (\kappa_1 - \kappa)^{-1} + (\kappa_2 - \kappa)^{-1} \in L^2[0, 1].$$

This implies that both

$$(\kappa_1 - \kappa)^{-1} \in L^2[0, 1] \quad \text{and} \quad (\kappa_2 - \kappa)^{-1} \in L^2[0, 1], \tag{21}$$

since as one of them increases in absolute value, the other one decreases. Consequently, (20) lies in $L^2[0, 1]$ and (18) follows from Minkowski's inequality.

The proof of (19) involves the tedious consideration of several cases, because it depends on which of the four h functions defined on p. 193 is used, both for (κ_1, κ_2)

and $(\bar\kappa_1, \bar\kappa_2)$. Assume first that $\kappa_1 < \kappa_2$ are both finite. If $u \notin \partial J$, then $\bar\kappa_1, \bar\kappa_2$ are also finite, so

$$h_{\bar\kappa_1,\bar\kappa_2} \circ \bar\kappa = (1-u\kappa)(1-u\kappa_1)(\kappa_1-\kappa)^{-1} + (1-u\kappa)(1-u\kappa_2)(\kappa_2-\kappa)^{-1}.$$

Since $\kappa \in (\kappa_1, \kappa_2)$ is bounded, this is a sum of two functions in $L^2[0, 1]$ by (21); hence it lies in $L^2[0, 1]$. If u is an endpoint ρ_i of J then $\bar\kappa_i$ is infinite. For instance, if $u = \rho_2$ then

$$h_{\bar\kappa_1,\bar\kappa_2} \circ \bar\kappa = h_{\bar\kappa_1,+\infty} \circ \bar\kappa = (1-\rho_2\kappa)(1-\rho_2\kappa_1)(\kappa_1-\kappa)^{-1} + \kappa_2\kappa(\kappa_2-\kappa)^{-1}.$$

Because κ is bounded, we conclude from (21) that (19) holds in this case also.

If one of the κ_i, say κ_2, is infinite, then the hypothesis that γ is admissible implies that

$$h_{\kappa_1,+\infty} \circ \kappa = (\kappa_1-\kappa)^{-1} + \kappa \in L^2[0, 1].$$

As above, it follows that each of the summands lies in $L^2[0, 1]$. If $u \neq \rho_1$ then $\bar\kappa_1 = \kappa_1/(1-u\kappa_1)$, $\bar\kappa_2 = -1/u$ are both finite, and

$$h_{\bar\kappa_1,\bar\kappa_2} \circ \bar\kappa = (1-u\kappa_1)(1-u\kappa)(\kappa_1-\kappa)^{-1} - u(1-u\kappa).$$

Observe that $(1-u\kappa)(\kappa_1-\kappa)^{-1} \in L^2[0, 1]$ because as κ increases to $+\infty$, $(\kappa_1-\kappa)^{-1}$ remains bounded, while as $\kappa \to \kappa_1$, obviously $(1-u\kappa)$ remains bounded. Thus, (19) holds. We leave the similar verification in the remaining cases to the reader. □

(2.3) Remark. The necessity of reparametrizing an admissible curve by arc-length before applying normal translation stems from the fact that the product of two L^2 functions need not be of class L^2: for a general parametrization, the speed of γ_u is given by $\sigma(1-u\kappa)$, where σ, κ are the speed and curvature of γ; hence, γ_u need not be admissible. This has no serious consequences because of (1.14).

The next result greatly simplifies the study of the spaces $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$. In all that follows, the notation $X \approx Y$ means that X is homeomorphic to Y.

(2.4) Theorem. *Let $P = (p, w)$, $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$, $-\infty \le \kappa_1 < \kappa_2 \le +\infty$ and $\rho_i = 1/\kappa_i$.*

(a) *Suppose $\kappa_1 < 0 < \kappa_2$. If at least one of κ_1, κ_2 is finite, then $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \approx \mathcal{L}_{-1}^{+1}(Q_1)$ for*

$$Q_1 = \Big(\frac{2}{\rho_2-\rho_1}\bar w\Big((q-p) + \frac{i}{2}(\rho_1+\rho_2)(z-w)\Big), z\bar w\Big).$$

(b) *Suppose $0 < \kappa_1 < \kappa_2$. Then $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \approx \mathcal{L}_1^{+\infty}(Q_2)$ for*

$$Q_2 = \Big(\frac{\bar w}{\rho_1-\rho_2}\big((q-p) + i\rho_2(z-w)\big), z\bar w\Big).$$

(c) *Suppose* $0 = \kappa_1 < \kappa_2$. *Then* $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \approx \mathcal{L}_0^{+\infty}(Q_3)$ *for*

$$Q_3 = \big(\bar{w}\big((q-p) + i\rho_2(z-w)\big), z\bar{w}\big).$$

(d) *Suppose* $\kappa_1 < \kappa_2 < 0$. *Then* $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \approx \mathcal{L}_1^{+\infty}(Q_4)$ *for*

$$Q_4 = \Big(\frac{\bar{z}}{\rho_1 - \rho_2}\big((q-p) + i\rho_1(z-w)\big), w\bar{z}\Big).$$

(e) *Suppose* $\kappa_1 < \kappa_2 = 0$. *Then* $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \approx \mathcal{L}_0^{+\infty}(Q_5)$ *for*

$$Q_5 = \big(\bar{z}\big((q-p) + i\rho_1(z-w)\big), w\bar{z}\big).$$

In cases (a)–(c) (*resp.* (d)–(e)), *the total turning of the image of a curve under the homeomorphism is equal* (*resp. opposite*) *to that of the original curve.*

Proof. Suppose first that $\kappa_1 < 0 < \kappa_2$ and let $k \in (\kappa_1, \kappa_2)$ be arbitrary. If $\rho_1 + \rho_2 \le 0$, then

$$1 - \left(\frac{\rho_1 + \rho_2}{2}\right)k > 1 - \left(\frac{\rho_1 + \rho_2}{2}\right)\kappa_1 = \tfrac{1}{2}(1 - \rho_2\kappa_1) \ge \tfrac{1}{2} > 0,$$

and if $\rho_1 + \rho_2 \ge 0$, then

$$1 - \left(\frac{\rho_1 + \rho_2}{2}\right)k > 1 - \left(\frac{\rho_1 + \rho_2}{2}\right)\kappa_2 = \tfrac{1}{2}(1 - \rho_1\kappa_2) \ge \tfrac{1}{2} > 0.$$

Consequently, $u = (\rho_1 + \rho_2)/2$ satisfies the hypothesis of (2.2). Let

$$\kappa_0 = \frac{2}{\rho_2 - \rho_1}.$$

Note that $0 < \kappa_0 < +\infty$; in the notation of (2.2), $-\kappa_0 = \bar{\kappa}_1$ and $\kappa_0 = \bar{\kappa}_2$. Define a map $F : \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \to \mathcal{L}_{-\kappa_0}^{+\kappa_0}(\bar{P}, \bar{Q})$ by letting $F(\gamma)$ be the translation by u of its reparametrization (still with domain $[0, 1]$) by a multiple of arc-length. This is continuous by (1.14). In fact, it is a homotopy equivalence: there is a similarly defined map $G : \mathcal{L}_{-\kappa_0}^{+\kappa_0}(\bar{P}, \bar{Q}) \to \mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ using translation by $-u$, and $GF(\gamma)$ is just a reparametrization of γ by (2.2)(b) and (c).

Let $T : \mathbb{C} \to \mathbb{C}$ be the dilatation $x \mapsto \kappa_0 x$. If $\gamma \in \mathcal{L}_{-\kappa_0}^{+\kappa_0}(\bar{P}, \bar{Q})$, then $T \circ \gamma$ lies in $\mathcal{L}_{-1}^{+1}(\tilde{P}, \tilde{Q})$, where

$$\tilde{P} = \left(\kappa_0\left(p + \frac{\rho_1 + \rho_2}{2} iw\right), w\right), \quad \tilde{Q} = \left(\kappa_0\left(q + \frac{\rho_1 + \rho_2}{2} iz\right), z\right),$$

and the correspondence $\gamma \mapsto T \circ \gamma$ yields a homeomorphism between these two spaces. Write $\tilde{P} = (\tilde{p}, w) \in \mathbb{C} \times \mathbb{S}^1$ and let $E : \mathbb{C} \to \mathbb{C}$ be the Euclidean motion given by $E(x) = \bar{w}(x - \tilde{p})$. Then the map $\gamma \mapsto E \circ \gamma$ is a homeomorphism from $\mathcal{L}_{-1}^{+1}(\tilde{P}, \tilde{Q})$ onto $\mathcal{L}_{-1}^{+1}(Q_1)$, with Q_1 as in the statement. The composition of all of these maps yields a homotopy equivalence $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \to \mathcal{L}_{-1}^{+1}(Q_1)$, which is homotopic to a homeomorphism by (1.7)(b).

The proofs of parts (b) and (c) are analogous, so only a brief outline will be provided. In part (b), we first use normal translation by ρ_2, and then compose with the dilatation $x \mapsto x/(\rho_1 - \rho_2)$ and an Euclidean motion; in part (c) the dilatation is not necessary. Parts (d) and (e) follow from (b) and (c), respectively, by reversing the orientation of all curves in the corresponding space.

By (2.2)(a), the normal translations used in establishing (a)–(c) preserve the total turning of a curve. Clearly, so do dilatations and Euclidean motions, while a reversal of orientation changes the sign of the total turning. This proves the last assertion of the theorem. $\square$

(2.5) Remark. Normal translations, and hence also the homotopy equivalences constructed in (2.4), do not generally respect inequalities between lengths. This is clear from Figure 2: two circles of the same radius $r > 0$ but different orientations are mapped to circles of radii equal to $r \pm u$ under normal translation by $u \in (0, r)$. See also the remarks at the end of Section 9.

A more concise version of (2.4) is the following; recall that $0(\pm\infty) = 0$ by convention.

(2.6) Corollary. *Let $P, Q \in UT\mathbb{C}$. Then $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ is homeomorphic to a space of type $\mathcal{L}_{-1}^{+1}(Q_0)$, $\mathcal{L}_{0}^{+\infty}(Q_0)$ or $\mathcal{L}_{1}^{+\infty}(Q_0)$, according to whether $\kappa_1\kappa_2 < 0$, $\kappa_1\kappa_2 = 0$ or $\kappa_1\kappa_2 > 0$, respectively.* $\square$

Out of the three possibilities, the spaces of type $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ with $\kappa_1\kappa_2 < 0$ are the ones with the most interesting topological properties. We deal with the two remaining cases in Section 7.

(2.7) Remark. We may replace $\mathcal{L}$ with $\mathcal{C}$ throughout in the statement of (2.4). In fact, the difficulty indicated in (2.3) disappears in this case, so the proof is simpler because it is not necessary to reparametrize the curves by arc-length before applying normal translation. This yields explicit homeomorphisms of the corresponding spaces, without relying on (1.7)(b). Because the curves in a space of type $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}$ are C^1 regular by definition, this simpler proof also works for this class (except that here $\kappa_1 < \kappa_2$ must be finite); see (9.1) for the precise statement.

3. Topology of $\mathcal{U}_c$

(3.1) Definition. Let $\gamma : [0, 1] \to \mathbb{C}$ be a regular curve and $\theta : [0, 1] \to \mathbb{R}$ be an argument of $\boldsymbol{t}_\gamma$. The *amplitude* of γ is given by

$$\omega = \sup_{t \in [0,1]} \theta(t) - \inf_{t \in [0,1]} \theta(t).$$

We call γ *condensed*, *critical* or *diffuse* according to whether $\omega < \pi$, $\omega = \pi$ or $\omega > \pi$.

Our main objective now is to understand the topology of $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ when $\kappa_1\kappa_2 < 0$. By (2.6), no generality is lost in assuming that $\kappa_1 = -1$, $\kappa_2 = +1$ and $P = (0, 1) \in \mathbb{C} \times \mathbb{S}^1$

(3.2) Definition. Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and $\theta_1 \in \mathbb{R}$ satisfy $e^{i\theta_1} = z$. We denote by $\mathcal{U}_c$, $\mathcal{U}_d$ and $\mathcal{T}$ the subspaces of $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ consisting of all condensed, diffuse and critical curves, respectively.

(3.3) Theorem. *The subspace $\mathcal{U}_c \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ consisting of all condensed curves is either empty or homeomorphic to $\boldsymbol{E}$, and hence contractible.*

Recall that $\boldsymbol{E}$ denotes the separable Hilbert space. In what follows, a function ϕ of a real variable will be called *increasing* (resp. *decreasing*) if $x < y$ (resp. $x > y$) implies that $\phi(x) \leq \phi(y)$. The previous theorem will be derived as a corollary of the following result.

(3.4) Proposition. *Let $\kappa_0 > 0$ and $\hat{\mathcal{U}}_c \subset \hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q; \theta_1)$ be the subspace consisting of all condensed curves. If $\hat{\mathcal{U}}_c \neq \emptyset$, then there exists a continuous $H : [0, 1] \times \hat{\mathcal{U}}_c \to \hat{\mathcal{U}}_c$ such that for all $\gamma \in \hat{\mathcal{U}}_c$,*

(i) *$H(1, \gamma) = \gamma$ and $H(0, \gamma) = \gamma_0$ (where γ_0 is independent of γ);*

(ii) *the amplitude of $\gamma_s = H(s, \gamma)$ is an increasing function of $s \in [0, 1]$;*

(iii) *the length of $\gamma_s = H(s, \gamma)$ is an increasing function of $s \in [0, 1]$.*

In particular, $\hat{\mathcal{U}}_c$ is contractible. Moreover, γ_0 is the unique curve of minimal length in $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$.

We believe that this proposition and its proof may be useful for other purposes which are not pursued here, e.g., for calculating the minimal length of curves in $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$. We shall first describe the effect of H on a single curve $\gamma \in \hat{\mathcal{U}}_c$ and then derive its main properties separately as lemmas. First we record two results which will be used to show that $H(0, \gamma)$ is independent of γ.

(3.5) Lemma. *Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$, $\gamma \in \hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$ and L be the length of γ. Suppose that q lies on the line through i/κ_0 having direction $-ie^{i\alpha}$ for some $\alpha \in [0, \pi)$. Then $L \geq \alpha/\kappa_0$ and equality holds if and only if γ is a reparametrization of the arc of the circle centered at i/κ_0 joining 0 to $1/\kappa_0(i - ie^{i\alpha})$.*

Proof. We lose no generality in assuming that $\kappa_0 = 1$. If $\alpha = 0$, there is nothing to prove, so suppose $\alpha \in (0, \pi)$. Let $\gamma : [0, L] \to \mathbb{C}$ be parametrized by arc-length, and let $\eta : [0, \alpha] \to \mathbb{C}$ be given by

$$\eta(s) = \int_0^s e^{i\sigma}\, d\sigma = i - ie^{is} \quad (s \in [0, \alpha]),$$

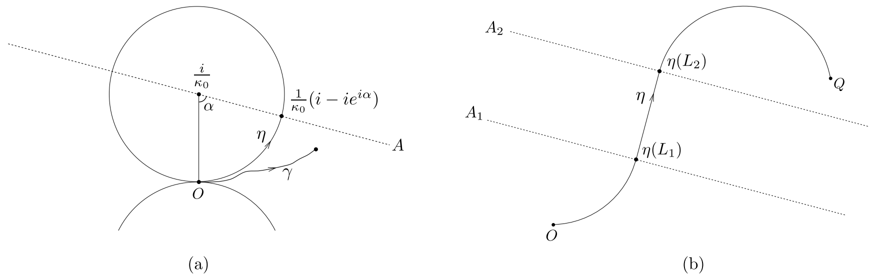


Figure 3. An illustration of (3.5) and (3.6).

so that η is the parametrization by arc-length of the arc of circle described in (3.5); see Figure 3(a). Set

$$f : [0, L] \to \mathbb{R}, \quad f(s) = \langle \gamma(s) - i, e^{i\alpha} \rangle,$$
$$g : [0, \alpha] \to \mathbb{R}, \quad g(s) = \langle \eta(s) - i, e^{i\alpha} \rangle.$$

Let A denote the line in the statement. Note that $f(s) = 0$ if and only if $\gamma(s) \in A$. We need to prove that $f(s) < 0$ for all $s \in [0, \alpha) \cap [0, L]$. Let θ_γ be the argument of $\boldsymbol{t}_\gamma$ satisfying $\theta_\gamma(0) = 0$. Then

$$(22)\quad f'(s) = \langle e^{i\theta_\gamma(s)}, e^{i\alpha} \rangle = \cos(\alpha - \theta_\gamma(s)) \quad \text{and} \quad g'(s) = \langle e^{is}, e^{i\alpha} \rangle = \cos(\alpha - s).$$

We have $f(0) = g(0)$. Since $g(s) < 0$ for all $s \in [0, \alpha)$, it suffices to establish that $f'(s) \leq g'(s)$ for all $s \in [0, \alpha] \cap [0, L]$. By the definition of $\hat{\mathcal{L}}_{-1}^{+1}(Q)$, θ_γ is 1-Lipschitz. Hence, $|\theta_\gamma(s)| \leq s$ for all $s \in [0, L]$. Consequently,

$$\alpha - s \leq \alpha - \theta_\gamma(s) \leq \alpha + s \quad \text{for all } s \in [0, L].$$

In particular, $\alpha - \theta_\gamma(s) \in [0, 2\pi]$ for all $s \in [0, \alpha] \cap [0, L]$. Since the cosine is decreasing over $[0, \pi]$, it follows immediately from (22) that if $\alpha - \theta_\gamma(s) \leq \pi$, then $f'(s) \leq g'(s)$. On the other hand, if $\alpha - \theta_\gamma(s) \in [\pi, 2\pi]$, then from $\alpha - \theta_\gamma(s) \leq \alpha + s$, we obtain that

$$\cos(\alpha - \theta_\gamma(s)) \leq \cos(\alpha + s) \leq \cos(\alpha - s),$$

the latter inequality coming from $\alpha \in (0, \pi)$ and $s \in [0, \alpha]$. Thus, $f'(s) \leq g'(s)$ in this case also. We conclude that $f(s) \leq g(s) < 0$ for all $s \in [0, \alpha) \cap [0, L]$. In particular, $L \geq \alpha$, as $\gamma(L) \in A$.

If $f(\alpha) = g(\alpha) = 0$, then we must have $f' = g'$, that is, $\theta_\gamma(s) = s$ for all $s \in [0, \alpha]$. Thus, in this case, $\gamma|_{[0,\alpha]}$ is a reparametrization of $\eta|_{[0,\alpha]}$. □

(3.6) Corollary. *Suppose that $\eta \in \hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$ is a concatenation of an arc of circle of curvature $\pm\kappa_0$, a line segment, and another arc of circle of curvature $\pm\kappa_0$, where some of these may be degenerate and both arcs have length less than π/κ_0. Then η is the unique curve in $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$ of minimal length.*

This result should be compared to [Dubins 1957, Proposition 9]. Their proofs are essentially the same.

Proof. Let $\eta : [0, L] \to \mathbb{C}$ be parametrized by arc-length, with $\eta|_{[0,L_1]}$, $\eta|_{[L_1,L_2]}$ and $\eta|_{[L_2,L]}$ corresponding to the first arc, line segment and second arc, respectively (see Figure 3(b)). Let A_i be the line perpendicular to $\eta'(L_i)$ passing through $\eta(L_i)$, $i = 1, 2$. Notice that A_1 and A_2 are parallel (or equal). Suppose that $\gamma : [0, M] \to \mathbb{C}$ is another curve in $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$, parametrized by arc-length. Let

$$M_1 = \inf\{s \in [0, M] : \gamma(s) \in A_1\}, \quad M_2 = \sup\{s \in [0, M] : \gamma(s) \in A_2\}.$$

By (3.5), we have $M_1 \geq L_1$ and $M - M_2 \geq L - L_2$. It is clear that $M_2 - M_1 \geq L_2 - L_1$ since any path joining a point of A_1 to a point of A_2 must have length greater than or equal to the distance between these lines. Hence, $M \geq L$. Furthermore, if equality holds, then $M_1 = L_1$, $M - M_2 = L - L_2$ and $M_2 - M_1 = L_2 - L_1$. By (3.5), the two former equalities imply that $\gamma|_{[0,M_1]} = \eta|_{[0,L_1]}$ and $\gamma|_{[M_2,M]} = \eta|_{[L_2,L]}$. The condition $M_2 - M_1 = L_2 - L_1$ then implies that $\gamma|_{[M_1,M_2]}$ must coincide with the line segment $\eta|_{[L_1,L_2]}$. □

(3.7) Remark. Notice that a condensed curve must be an embedding of [0, 1]. In fact, its image is the graph of a function of x, after a suitable choice of the x-axis.

(3.8) Construction. Let $\gamma \in \hat{\mathfrak{U}}_c$, $\theta : [0, 1] \to \mathbb{R}$ be the argument of $\boldsymbol{t}_\gamma$ satisfying $\theta(0) = 0$. A number $\varphi \in (-\pi/2, \pi/2)$ will be called an *axis* of γ if $\langle \boldsymbol{t}_\gamma(t), e^{i\varphi}\rangle > 0$ for all $t \in [0, 1]$. Since γ is condensed, the set of all axes of γ is an open interval. The most natural axis, and the center of this interval, is

$$\bar{\varphi}_\gamma = \tfrac{1}{2}\Big(\sup_{t\in[0,1]} \theta(t) + \inf_{t\in[0,1]} \theta(t)\Big). \tag{23}$$

Let φ be any axis of γ. Rotating around the origin through φ and writing $\gamma(t) = (x(t), y(t))$ in terms of the new x- and y-axes, the hypothesis that $\langle \boldsymbol{t}_\gamma, e^{i\varphi}\rangle > 0$ becomes equivalent to the fact that $\dot{x}$ is bounded and positive over [0, 1]. Let

$$\gamma(x) = (x, y(x)) \quad (x \in [0, b])$$

be the reparametrization of γ by x and define

$$f : [0, b] \to \mathbb{R} \quad \text{by} \quad f(x) = \dot{y}(x).$$

Let $f_s : [0, b] \to \mathbb{R}$ ($s \in [0, 1]$) be a family of absolutely continuous functions and set

$$\gamma_s(x) = \left(x, \int_0^x f_s(u)\,du\right) \quad (x \in [0, b]).$$

A straightforward computation shows that the curvature of γ_s is given by

$$\kappa_{\gamma_s}(x) = \frac{\dot{f}_s(x)}{(1 + f_s(x)^2)^{3/2}} \quad (x \in [0, b]).$$

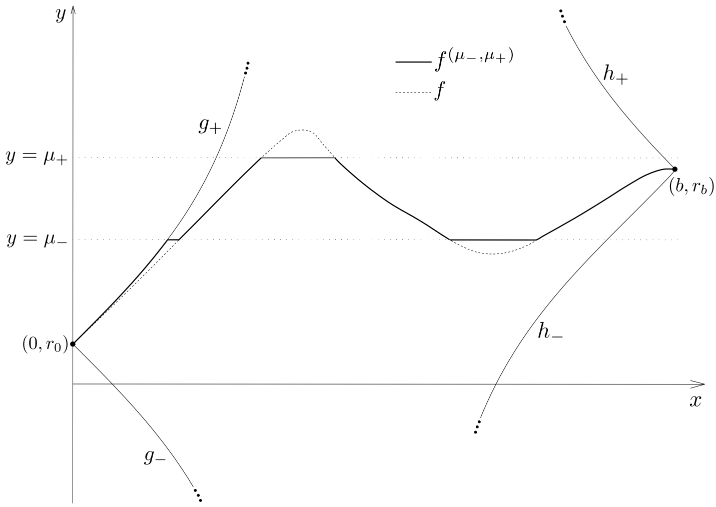


Figure 4. An illustration of (3.8).

Therefore, γ_s lies in $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q;\theta_1)$ if and only if f_s satisfies

(i) $|\dot{f}_s(x)|\le\kappa_0(1+f_s(x)^2)^{3/2}$ for almost every $x\in[0,b]$ (i.e., $\kappa_{\gamma_s}\in[-\kappa_0,+\kappa_0]$ a.e.);

(ii) $f_s(0)=r_0:=\dot{y}(0)$ and $f_s(b)=r_b:=\dot{y}(b)$ (i.e., $\boldsymbol{t}_{\gamma_s}(0)=\boldsymbol{t}_\gamma(0)$ and $\boldsymbol{t}_{\gamma_s}(b)=\boldsymbol{t}_\gamma(b)$);

(iii) $\int_0^b f_s(x)\,dx=A_1:=y(b)-y(0)$ (i.e., $\gamma_s(b)=\gamma(b)$).

We will now produce a homotopy of $f=f_1$ through absolutely continuous functions satisfying (i)–(iii).

Define

(24) $$\alpha_\pm=\mp\frac{r_0}{\sqrt{1+r_0^2}},\quad g_\pm(x)=\pm\frac{\kappa_0x-\alpha_\pm}{\sqrt{1-(\kappa_0x-\alpha_\pm)^2}}\quad\text{for } x\in\left(\frac{\alpha_\pm-1}{\kappa_0},\frac{\alpha_\pm+1}{\kappa_0}\right)$$

(see Figure 4) and, similarly,

(25)
$$\beta_\pm=\kappa_0b\pm\frac{r_b}{\sqrt{1+r_b^2}},\quad h_\pm(x)=\mp\frac{\kappa_0x-\beta_\pm}{\sqrt{1-(\kappa_0x-\beta_\pm)^2}}\quad\text{for } x\in\left(\frac{\beta_\pm-1}{\kappa_0},\frac{\beta_\pm+1}{\kappa_0}\right).$$

The functions $g_\pm$ are the solutions of the differential equations $\dot{g}=\pm\kappa_0(1+g^2)^{3/2}$ with $g(0)=r_0$. Similarly, $h_\pm$ are the solutions of the differential equations $\dot{h}=\mp\kappa_0(1+h^2)^{3/2}$ with $h(b)=r_b$. Extend their domains to all of $\mathbb{R}$ by setting

$$g_\pm(x)=\pm\infty\ \text{ if }\ x\ge\frac{\alpha_\pm+1}{\kappa_0}\quad\text{and}\quad g_\pm(x)=\mp\infty\ \text{ if }\ x\le\frac{\alpha_\pm-1}{\kappa_0},$$

and do similarly for $h_\pm$. Since the curvature of $\gamma = \gamma_1$ takes values in $[-\kappa_0, +\kappa_0]$, condition (i) applied to $f = f_1$ gives

$$g_-(x), h_-(x) \le f(x) \le g_+(x), h_+(x) \quad \text{for all } x \in [0, b]. \tag{26}$$

Let

$$\begin{gathered} m_- = \inf_{x\in[0,b]} f(x), \quad m_+ = \sup_{x\in[0,b]} f(x), \\ \Delta = \{(\mu_-, \mu_+) \in [m_-, m_+] : \mu_- \le \mu_+\}. \end{gathered} \tag{27}$$

For $(\mu_-, \mu_+) \in \Delta$, let $f^{(\mu_-,\mu_+)} : [0, b] \to \mathbb{R}$ be given by

$$f^{(\mu_-,\mu_+)}(x) = \operatorname{median}\big(h_-(x), g_-(x), \mu_-, f(x), \mu_+, g_+(x), h_+(x)\big) \tag{28}$$

(see Figure 4). The functions $f^{(\mu_-,\mu_+)}$ automatically satisfy conditions (i) and (ii). Define $A : \Delta \to \mathbb{R}$ to be the area under the graph of $f^{(\mu_-,\mu_+)}$:

$$A(\mu_-, \mu_+) = \int_0^b f^{(\mu_-,\mu_+)}(x)\, dx.$$

It is immediate from (28) that

(A) A is increasing as a function of either μ_- or μ_+;

(B) A is a Lipschitz function of (μ_-, μ_+). In fact,

$$\big|A(\mu_- + u, \mu_+ + v) - A(\mu_-, \mu_+)\big| \le b(|u| + |v|).$$

By (A), for each $s \in [0, 1]$, the set

$$\big\{(\mu_-, \mu_+) \in \Delta : A(\mu_-, \mu_+) = A_1 \text{ and } \mu_+ - \mu_- = (m_+ - m_-)s\big\}$$

is an interval of the latter line in the (μ_-, μ_+)-plane. Let $(\mu_-(s), \mu_+(s))$ be the coordinates of the center of this interval. By (B), $\mu_-(s)$ and $\mu_+(s)$ are continuous (even Lipschitz), and (A) implies that μ_- is a decreasing, while μ_+ is an increasing function of $s \in [0, 1]$. The functions

$$f_s : [0, b] \to \mathbb{R}, \quad f_s = f^{(\mu_-(s),\mu_+(s))},$$

satisfy all of conditions (i)–(iii) by construction. We repeat their definition for convenience:

$$\begin{aligned} f_s(x) &= \operatorname{median}\big(h_-(x), g_-(x), \mu_-(s), f(x), \mu_+(s), g_+(x), h_+(x)\big), \\ \gamma_s(x) &= \left(x, \int_0^x f_s(u)\, du\right) \quad (x \in [0, b]). \end{aligned} \tag{29}$$

We will denote $\mu_+(0) = \mu_-(0)$ by μ_0. The monotonicity of μ_-, μ_+ implies that

$$\mu_-(s) \le \mu_0 \le \mu_+(s) \quad \text{for all } s \in [0, 1]. \tag{30}$$

□

(3.9) Remark. We deduce from (26) and (29) that

$$f_0 = \operatorname{median}(h_-, g_-, \mu_0, g_+, h_+).$$

The graph of f_0 is composed of at most three parts: a piece of the graph of g_- or g_+, a piece of the graph of the constant function $y = \mu_0$, and a piece of the graph of h_- or h_+. The corresponding curve γ_0 is thus the concatenation of an arc of circle of curvature $\pm\kappa_0$, a line segment and another arc of circle of curvature $\pm\kappa_0$, though some of these may degenerate to a point. It is an immediate consequence of (3.6) that γ_0 (and hence f_0) is independent of γ and of the chosen axis φ.

(3.10) Lemma. *Let φ be an axis of $\gamma \in \hat{\mathfrak{U}}_c$ and $s \mapsto \gamma_s$ $(s \in [0, 1])$ be the deformation described in (3.8). Then $\gamma_0 \in \hat{\mathfrak{U}}_c$ is the unique curve of minimal length in $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$.* □

Remark. Notice that this proves Dubins' Theorem I [1957] in the case where $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$ contains condensed curves. Furthermore, given Q and κ_0, we can use (3.8) to describe γ_0 explicitly.

(3.11) Lemma. *Let $S_+ = \{x \in [0, b] : f(x) \geq \mu_0\}$ and $S_- = \{x \in [0, b] : f(x) \leq \mu_0\}$. Then $f_s(x)$ is an increasing (resp. decreasing) function of $s \in [0, 1]$ if $x \in S_+$ (resp. S_-). Moreover, for all $s \in [0, 1]$, $f_s(x) \geq \mu_0$ if $x \in S_+$ and $f_s(x) \leq \mu_0$ if $x \in S_-$.*

Proof. Suppose that $x \in S_+$. From (26) and (30), we deduce that

$$g_-(x), h_-(x), \mu_-(s) \leq f(x) \leq g_+(x), h_+(x).$$

Hence, $f_s(x) = \min\{\mu_+(s), f(x)\} \geq \mu_0$ and $f_s(x)$ increases with s since $\mu_+(s)$ does. The proof for $x \in S_-$ is analogous. □

(3.12) Corollary. *Let $m_-(s) = \inf_{x\in[0,b]} f_s(x)$ and $m_+(s) = \sup_{x\in[0,b]} f_s(x)$. Then $m_+(s)$ is an increasing and $m_-(s)$ a decreasing function of $s \in [0, 1]$.* □

(3.13) Lemma. *Let $\gamma \in \hat{\mathfrak{U}}_c$ and $s \mapsto \gamma_s \in \hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q; \theta_1)$ be the homotopy described in (3.8). Let ω_s denote the amplitude of γ_s. Then ω_s is an increasing function of s; in particular, γ_s is condensed (i.e., $\gamma_s \in \hat{\mathfrak{U}}_c$) for all $s \in [0, 1]$.*

Proof. Let φ be the axis of γ chosen for the construction. Recall that, by definition,

$$\omega_s = \sup_{x\in[0,b]} \theta_s(x) - \inf_{x\in[0,b]} \theta_s(x) \quad (s \in [0, 1]), \tag{31}$$

where θ_s is the argument of $\boldsymbol{t}_{\gamma_s}$ such that $\theta_s(0) = 0$. By (29),

$$f_s(x) = \tan(\theta_s(x) - \varphi). \tag{32}$$

Because the tangent is an increasing function, (3.12) immediately implies that ω_s is increasing. □

Remark. Although γ_0 has minimal amplitude in $\hat{\mathfrak{U}}_c$ by the previous lemma, there may be other curves in $\hat{\mathfrak{U}}_c$ with the same amplitude. This is the case, for instance, for the curves γ_0 and γ corresponding to the functions f and f_0 of Figure 4.

(3.14) Lemma. *Let $\gamma \in \hat{\mathfrak{U}}_c$ and $s \mapsto \gamma_s$ be the deformation described in (3.8). Then the length of γ_s is an increasing function of $s \in [0, 1]$.*

Proof. Let $\lambda : \mathbb{R} \to \mathbb{R}$ be given by $\lambda(u) = (1+u^2)^{1/2}$. A straightforward computation shows that

$$\lambda''(u) = (1+u^2)^{-3/2} > 0 \quad \text{for all } u \in \mathbb{R}. \tag{33}$$

Moreover, by the definition (29), the length L_s of γ_s is given by

$$L_s = \int_0^b (\lambda \circ f_s)(x)\, dx.$$

Let $s_1 \leq s_2 \in [0, 1]$, S_+, S_- be as in (3.11) and

$$\begin{aligned} T_+ =& \{(x, y) \in [0, b] \times \mathbb{R} : f_{s_1}(x) \leq y \leq f_{s_2}(x)\}, \\ T_- =& \{(x, y) \in [0, b] \times \mathbb{R} : f_{s_2}(x) \leq y \leq f_{s_1}(x)\}. \end{aligned}$$

Using (3.11), we deduce that

$$\begin{aligned} L_{s_2} - L_{s_1} &= \int_0^b (\lambda \circ f_{s_2})(x) - (\lambda \circ f_{s_1})(x)\, dx \\ &= \Big(\int_{S_+} + \int_{S_-}\Big)(\lambda \circ f_{s_2})(x) - (\lambda \circ f_{s_1})(x)\, dx \\ &= \Big(\int_{T_+} - \int_{T_-}\Big)\lambda'(y)\, dy\, dx \\ &\geq \Big(\int_{T_+} - \int_{T_-}\Big)\lambda'(\mu_0)\, dy\, dx \quad \text{(by (33))} \\ &= \lambda'(\mu_0)\Big(\int_0^b f_{s_2} - \int_0^b f_{s_1}\Big) = 0 \quad \text{(by the definition of } f_s). \end{aligned}$$

Therefore, L_s is an increasing function of $s \in [0, 1]$. □

We are finally ready to prove (3.4) and (3.3).

Proof of (3.4). For each $\gamma \in \hat{\mathfrak{U}}_c$, let

$$\bar{\varphi}_\gamma = \tfrac{1}{2}\big(\sup_{t \in [0,1]} \theta_\gamma(t) + \inf_{t \in [0,1]} \theta_\gamma(t)\big), \tag{34}$$

where $\theta_\gamma : [0, 1] \to \mathbb{R}$ is the argument of $\boldsymbol{t}_\gamma$ satisfying $\theta_\gamma(0) = 0$. It is clear that $\bar{\varphi}_\gamma$ depends continuously on $\gamma \in \hat{\mathfrak{U}}_c$. Define $H : [0, 1] \times \hat{\mathfrak{U}}_c \to \hat{\mathfrak{U}}_c$ by $H(s, \gamma) = \gamma_s$, where γ_s is the curve (29) constructed in (3.8) with chosen axis $\bar{\varphi}_\gamma$. Then part (ii) of (3.4) follows from (3.13), and part (iii) from (3.14). The last assertion of (3.4) and part (i) were established in (3.9). □

Proof of (3.3). Assume that $\mathcal{U}_c$ is nonempty. It is certainly open in $\mathcal{L}_{-1}^{+1}(Q;\theta_1)$. Hence, by (1.7), it suffices to prove that $\mathcal{U}_c$ is weakly contractible. Let K be a compact manifold and $g: K \to \mathcal{U}_c$, $a \mapsto \gamma^a$, be a continuous map. Using (1.11), we may assume that the image of g is contained in (the image under set inclusion of) $\mathcal{C}_{-\kappa_0}^{+\kappa_0}(Q;\theta_1)$ for some $\kappa_0 \in (0,1)$. By (1.17), we have continuous injections

$$\mathcal{C}_{-\kappa_0}^{+\kappa_0}(Q;\theta_1) \to \hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q;\theta_1) \to \mathcal{L}_{-1}^{+1}(Q;\theta_1).$$

Let $G:[0,1]\times K \to \mathcal{L}_{-1}^{+1}(Q;\theta_1)$ map (s,a) to (the image under set inclusion of) $H(s,\gamma^a)$, with H as in (3.4). Then G is a null-homotopy of g in $\mathcal{U}_c$. ☐

The next couple of lemmas will only be needed in later sections.

(3.15) Lemma. *Suppose that there exists $\hat{\omega}\in(0,\pi)$ such that if $\gamma\in\hat{\mathcal{U}}_c\subset\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q;\theta_1)$ then its amplitude ω_γ satisfies $\omega_\gamma \le \hat{\omega}$. Let $L(\eta)$ denote the length of η. Then* $\sup_{\gamma\in\hat{\mathcal{U}}_c} L(\gamma)$ *is finite. In particular, the images of $\gamma\in\hat{\mathcal{U}}_c$ are all contained in some bounded subset of $\mathbb{C}$.*

Proof. Let $\gamma\in\hat{\mathcal{U}}_c$ and $\bar{\varphi}_\gamma$ be as in (34). By hypothesis, the image of $\theta_\gamma:[0,1]\to\mathbb{R}$ is contained in $\left[\bar{\varphi}_\gamma-\hat{\omega}/2,\, \bar{\varphi}_\gamma+\hat{\omega}/2\right]$. Let $f:[0,b]\to\mathbb{R}$ be the function corresponding to γ and the axis $\bar{\varphi}_\gamma$, in the notation of (3.8). Note that $b=\langle e^{i\bar{\varphi}_\gamma}, q\rangle \le |q|$, where q is the $\mathbb{C}$-coordinate of Q. By (32),

$$|f(x)| \le \tan\Big(\frac{\hat{\omega}}{2}\Big) \quad \text{for all } x\in[0,b].$$

Therefore, the length $L(\gamma)$ of γ satisfies

$$L(\gamma) = \int_0^b \sqrt{1+f(x)^2}\,dx \le b\sec\Big(\frac{\hat{\omega}}{2}\Big) \le |q|\sec\Big(\frac{\hat{\omega}}{2}\Big).$$ ☐

(3.16) Lemma. *Let $\hat{\mathcal{U}}_c\subset\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q;\theta_1)$ and $H:[0,1]\times\hat{\mathcal{U}}_c\to\hat{\mathcal{U}}_c$ be the deformation described in (3.4) and (3.8). Suppose that $\theta_1=0$. Then $\omega_0<\omega_1$ unless $\gamma_1=\gamma_0$.*

Proof. It is obvious that $\omega_1=\omega_0$ if $\gamma_1=\gamma_0$. The condition $\theta_1=0$ is equivalent to $r_0=r_b$, in the notation of (3.8). Suppose without loss of generality that $\mu_0\ge r_0$, so that $m_+(0)=\mu_0$.

If $m_+(1)\le\mu_0$, then $S_-=[0,b]$. Hence, by (3.11), $f_1(x)\le f_0(x)$ for all $x\in[0,b]$. Since f_1 and f_0 have the same area, we conclude that $f_1=f_0$, that is, $\gamma_1=\gamma_0$.

By (3.12), $m_-(1)\le m_-(0)$. Hence, if $m_+(1)>\mu_0=m_+(0)$, then $\omega_0<\omega_1$ by (31) and (32). ☐

Existence of condensed curves. The question of whether $\mathcal{U}_c\ne\varnothing$ is settled by means of an elementary geometric construction. In all that follows, $C_r(a)$ denotes the circle of radius $r>0$ centered at $a\in\mathbb{C}$.

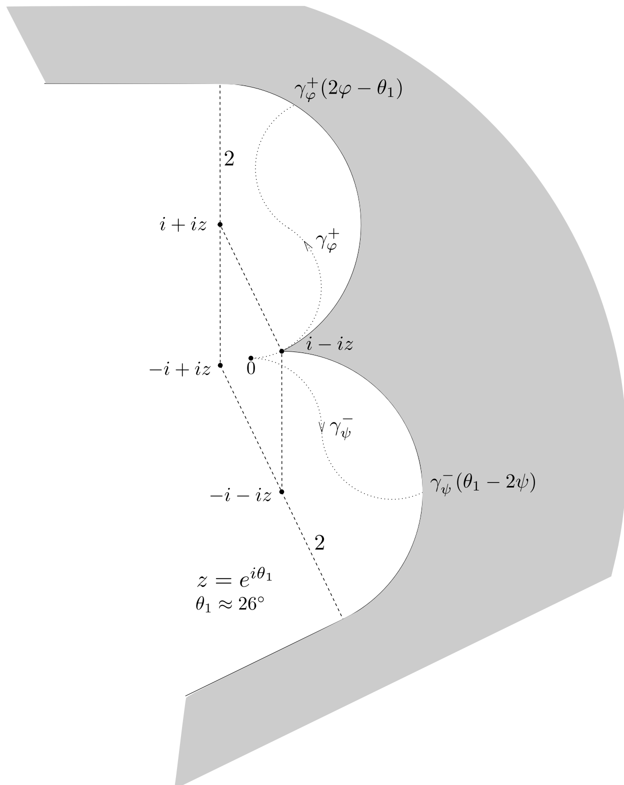


Figure 5. Let $\theta_1 \in [0, \pi)$ be fixed and $Q = (q, z)$, where $z = e^{i\theta_1}$. There exist condensed curves in $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ if and only if q belongs to the open gray region.

(3.17) Proposition. *Let $\theta_1 \in [0, \pi)$ be fixed, $z = e^{i\theta_1}$ and $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$. Let $R_{\mathcal{U}_c}$ be the open region of the plane which does* not *contain $-i + iz$ and which is bounded by the shortest arcs of the circles $C_2(\pm(i + iz))$ joining $i - iz$ to $i - iz \pm 2(i + iz)$ and their tangent lines at the latter points. Then $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ contains condensed curves if and only if $q \in R_{\mathcal{U}_c}$.* (*See Figure 5.*)

It is clear from the definition of condensed curve that $\mathcal{U}_c \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ is empty if $|\theta_1| \geq \pi$. In other words, all condensed curves in $\mathcal{L}_{-1}^{+1}(Q)$ must be contained in the subspace $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ with θ_1 the unique number in $(-\pi, \pi)$ satisfying $e^{i\theta_1} = z$ (for $z \neq -1$). We have assumed that $\theta_1 \in [0, \pi)$ just to simplify the statement. If $\theta_1 \in (-\pi, 0]$, the only difference is that $i - iz$ should be interchanged with $-i + iz$. The proof is analogous to the one given below. Alternatively, it can be deduced from the proposition by applying a reflection across the x-axis. When $\theta_1 = 0$, the statement becomes ambiguous; in this case the arcs of circles which bound $R_{\mathfrak{U}_c}$ are centered at $\pm 2i$, bounded by 0 and $\pm 4i$, and pass through the points $2 \pm 2i$, respectively.

Proof of (3.17). Let $\eta : [0, 1] \to \mathbb{C}$ be condensed and let $\theta_\eta : [0, 1] \to \mathbb{R}$ be the argument of $\boldsymbol{t}_\eta$ satisfying $\theta_\eta(0) = 0$. Observe that

$$\inf\{\theta_\eta(t) : t \in [0, 1]\} \in [\theta_1 - \pi, 0] \quad \text{and} \quad \sup\{\theta_\eta(t) : t \in [0, 1]\} \in [\theta_1, \pi].$$

The proof relies on the study of the following curves. For each $\varphi \in [\theta_1, \pi]$, define $\gamma_\varphi^+ : [0, 2\varphi - \theta_1] \to \mathbb{C}$ to be the unique curve parametrized by arc-length satisfying

$$\gamma_\varphi^+(0) = 0 \quad \text{and} \quad \boldsymbol{t}_{\gamma_\varphi^+}(s) = \begin{cases} e^{is} & \text{if } s \in [0, \varphi], \\ e^{i(2\varphi - s)} & \text{if } s \in [\varphi, 2\varphi - \theta_1]. \end{cases}$$

Then γ_φ^+ is the concatenation of two arcs of circles of radius 1,

$$\inf_{t \in [0,1]} \theta_{\gamma_\varphi^+}(t) = 0, \qquad \sup_{t \in [0,1]} \theta_{\gamma_\varphi^+}(t) = \varphi, \qquad \boldsymbol{t}_{\gamma_\varphi^+}(2\varphi - \theta_1) = z,$$

$$\gamma_\varphi^+(2\varphi - \theta_1) = \int_0^\varphi e^{is}\, ds + \int_\varphi^{2\varphi - \theta_1} e^{i(2\varphi - s)}\, ds = (i + iz) - 2ie^{i\varphi}.$$

Thus, as φ increases from θ_1 to π, the endpoints of the γ_φ^+ trace out the arc of $C_2(i + iz)$ bounded by $i - iz$ and $3i + iz$. Further, the tangent line to $C_2(i + iz)$ at $\gamma_\varphi^+(2\varphi - \theta_1)$ is parallel to $e^{i\varphi}$, for it must be orthogonal to $-2ie^{i\varphi}$.

Similarly, for each $\psi \in [\theta_1 - \pi, 0]$, let $\gamma_\psi^- : [0, \theta_1 - 2\psi] \to \mathbb{C}$ be the curve, parametrized by arc-length, which satisfies

$$\gamma_\psi^-(0) = 0 \quad \text{and} \quad \boldsymbol{t}_{\gamma_\psi^-}(s) = \begin{cases} e^{-is} & \text{for } s \in [0, -\psi], \\ e^{i(2\psi + s)} & \text{for } s \in [-\psi, \theta_1 - 2\psi]. \end{cases}$$

Then γ_ψ^- is the concatenation of two arcs of circles of radius 1, $\boldsymbol{t}_{\gamma_\psi^-}(\theta_1 - 2\psi) = z$ for all $\psi \in [\theta_1 - \pi, 0]$, and as ψ decreases from 0 to $\theta_1 - \pi$, the endpoints of the γ_ψ^- traverse the arc of $C_2(-i - iz)$ bounded by $i - iz$ and $-i - 3iz$. Moreover, the tangent line to this circle at $\gamma_\psi^-(\theta_1 - 2\psi)$ is parallel to $e^{i\psi}$.

Any $q \in \bar{R}_{\mathfrak{U}_c}$ is the endpoint of a curve of one of the following three types:

(i) The concatenation of a γ_φ^+ or a γ_ψ^- with a line segment of direction z.

(ii) The concatenation of $\gamma_\pi^+|_{[0,\pi]}$, a line segment of length $\ell \geq 0$ having direction -1, the arc $-\ell + \gamma_\pi^+|_{[\pi, 2\pi - \theta_1]}$, and a line segment of direction z.

(iii) The concatenation of $\gamma_{\theta_1 - \pi}^-|_{[0, \pi - \theta_1]}$, a line segment of length $\ell_1 \geq 0$ of direction $-z$, the arc $-\ell_1 z + \gamma_{\theta_1 - \pi}^-|_{[\pi - \theta_1, 3\pi/2 - \theta_1]}$, a line segment of length $\ell_2 \geq 0$ and direction $-iz$, the arc $-\ell_1 z - \ell_2 iz + \gamma_{\theta_1 - \pi}^-|_{[3\pi/2 - \theta_1, 2\pi - \theta_1]}$, and a line segment of direction z.

The curves which we have described have curvature equal to ± 1 over intervals of positive measure and, additionally, may be critical curves. Nevertheless, for any $q \in R_{\mathfrak{U}_c}$, we can find a condensed $\gamma \in \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ by composing one of these curves with a dilatation through a factor $c > 1$, with c close to 1 if q lies close to $\partial R_{\mathcal{U}_c}$, and by avoiding the argument π (for a curve of type (i)) or $\theta_1 - \pi$ (for a curve of type (iii)).

Conversely, suppose that $\mathcal{L}_{-1}^{+1}(Q)$ contains condensed curves. Let $\eta : [0, L] \to \mathbb{C}$ be such a curve, parametrized by arc-length, and let $\varphi = \sup \theta$, where $\theta : [0, L] \to \mathbb{R}$ is an argument of $\boldsymbol{t}_\eta$ satisfying $\theta(0) = 0$. Define

$$g : [0, L] \to \mathbb{R} \quad \text{by} \quad g(s) = \big\langle \eta(s) - \gamma_\varphi^+(2\varphi - \theta_1), i e^{i\varphi} \big\rangle.$$

Note that $g(s) > 0$ if and only if $\eta(s)$ lies to the left of the line through $\gamma_\varphi^+(2\varphi - \theta_1) \in C_2(i + iz)$ having direction $e^{i\varphi}$; we have already seen that this line is tangent to this circle at this point. We claim that $g(L) < 0$. Since η is admissible, $\theta = \arg \circ \boldsymbol{t}_\eta$ is an absolutely continuous function, and $|\theta'| = |\kappa_\eta| < 1$ almost everywhere by (5). Moreover, $\theta(s) \in [\varphi - \pi, \varphi]$ for all s because η is condensed. Hence,

$$g'(s) = \langle e^{i\theta(s)}, i e^{i\varphi} \rangle = \cos\Big(\theta(s) - \varphi - \frac{\pi}{2}\Big) \leq 0 \quad \text{for all } s \in [0, L]. \tag{35}$$

Let $J_i = (a_i, b_i) \subset (0, L)$ $(i = 1, 2, 3)$ be disjoint intervals such that

(I) $\theta(a_1) = 0$ and $\theta(b_1) = \theta_1$;

(II) $\theta(a_2) = \theta_1$ and $\theta(b_2) = \varphi$;

(III) $\theta(a_3) = \varphi$ and $\theta(b_3) = \theta_1$.

Such intervals exist because θ is a continuous function satisfying $\theta(0) = 0$, $\theta_1 \leq \varphi = \sup \theta$ and $\theta(L) = \theta_1$. Let λ denote the Lebesgue measure on $\mathbb{R}$. Fix i and let $[\alpha, \beta]$ be any nondegenerate subinterval of $\theta((a_i, b_i))$. Since θ is strictly 1-Lipschitz, if $S = \{s \in (a_i, b_i) : \alpha \leq \theta(s) \leq \beta\}$, then $\lambda(S) > \beta - \alpha$. Combining this with (35), we deduce that

$$\begin{aligned} g(L) - g(0) &\leq \bigg(\int_{a_1}^{b_1} + \int_{a_2}^{b_2} + \int_{a_3}^{b_3}\bigg) g'(s)\, ds \\ &< \int_0^{\theta_1} \langle e^{it}, e^{i\varphi} \rangle\, dt + 2\int_{\theta_1}^{\varphi} \langle e^{it}, e^{i\varphi} \rangle\, dt = \langle \gamma_\varphi^+(2\varphi - \theta_1), i e^{i\varphi} \rangle. \end{aligned}$$

Therefore, $g(L) < 0$ as claimed. Similarly, if $\psi = \inf \theta$, then $\eta(L)$ lies on the side of the tangent to $C_2(-i - iz)$ at $\gamma_\psi^-(\theta_1 - 2\psi)$ which does not contain $-i + iz$. It follows that $q = \eta(L) \in R_{\mathcal{U}_c}$. □

4. Topology of $\mathcal{U}_d$

Throughout this section, let K denote a compact manifold, possibly with boundary. Also, let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ be fixed (but otherwise arbitrary) and let $\mathcal{U}_d \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ denote the subset consisting of all diffuse curves in $\mathcal{L}_{-1}^{+1}(Q)$ having total turning θ_1, for some fixed $\theta_1 \in \mathbb{R}$ satisfying $e^{i\theta_1} = z$. Finally, let $O = (0, 1) \in \mathbb{C} \times \mathbb{S}^1$, the identity element of the group $UT\mathbb{C}$.

Our next objective is to prove that $\mathcal{U}_d$ is contractible. The idea behind the proof is quite simple. If γ is diffuse, then we can "graft" a straight line segment of length greater than 4 onto γ, as illustrated in Figure 8. By the theorem of Dubins stated in the introduction, this segment can be deformed so that in the end an eight curve of large radius traversed n times has been attached to it, as in Figure 7(e). These eights are then spread along the curve, as in Figure 7(f). If $n \in \mathbb{N}$ is large enough, the spreading can be carried out within $\mathcal{L}_{-1}^{+1}(Q)$. The result is a curve which is so loose that the constraints on the curvature may be safely forgotten, allowing us to use the following fact.

(4.1) Theorem (Smale). *Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$. Then $\mathcal{C}_{-\infty}^{+\infty}(Q)$ and $\mathcal{L}_{-\infty}^{+\infty}(Q)$ have $\aleph_0$ connected components, one for each $\theta_1 \in \mathbb{R}$ satisfying $e^{i\theta_1} = z$, all of which are contractible.*

Proof. For the space $\mathcal{C}_{-\infty}^{+\infty}(Q)$, the proof was discussed in the introduction. We may replace $\mathcal{C}_{-\infty}^{+\infty}(Q)$ by $\mathcal{L}_{-\infty}^{+\infty}(Q)$ using (1.12). □

(4.2) Lemma. *Let $P \in UT\mathbb{C}$. Then $\mathcal{C}_{-1}^{+1}(P, P)$ and $\mathcal{L}_{-1}^{+1}(P, P)$ have $\aleph_0$ connected components, one for each turning number $n \in \mathbb{Z}$, all of which are contractible.*

Proof. By (1.12), it suffices to prove the result for $\mathcal{C}_{-1}^{+1}(P, P)$. Let $\mathcal{C}_n \subset \mathcal{C}_{-1}^{+1}(P, P)$ denote the subset of all curves which have turning number n. Then each $\mathcal{C}_n$ is closed and open. Hence, to establish that $\mathcal{C}_n$ is a contractible component, it suffices, by (1.7)(b), to prove that it is weakly contractible.

Recall that $\mathcal{C}_{-1}^{+1}(P, P) \approx \mathcal{C}_{-1}^{+1}(O)$, the homeomorphism coming from composing all curves with a suitable Euclidean motion. We may thus assume that $P = O$. Let K be a compact manifold and $f : K \to \mathcal{C}_n$ a continuous map. By (4.1), there exists a continuous $F : [0, 1] \times K \to \mathcal{C}_{-\infty}^{+\infty}(O)$ such that $F_0 = f$ and F_1 is a constant map. Let

$$M = 2 \sup\left\{ |\kappa_{F(s,a)}(t)| : s, t \in [0, 1], a \in K \right\}.$$

Given a curve γ, let $M\gamma$ denote the dilated curve $t \mapsto M\gamma(t)$. It is easy to see that $\kappa_{M\gamma} = \kappa_\gamma / M$. Hence, MF is a homotopy between Mf and a constant map within $\mathcal{C}_{-1}^{+1}(O)$. But f and Mf are homotopic within $\mathcal{C}_{-1}^{+1}(O)$ through $u \mapsto uf$ ($u \in [1, M]$). Therefore, f is null-homotopic. □

Loops and eights. We shall now explain how to attach loops and eights to a curve, and how to spread eights along it (Figure 7).

(4.3) Definition. We denote by $\alpha : \mathbb{R} \to \mathbb{C}$ the *loop* of radius 2 and by $\beta : \mathbb{R} \to \mathbb{C}$ the *eight* curve of the same radius (see Figure 7(b) and (d)) given by

$$\alpha(t) = 2i\bigl(1 - \exp(2\pi i t)\bigr),$$

$$\beta(t) = \begin{cases} \alpha(2t) & \text{for } t \in \left[\frac{m}{2}, \frac{m+1}{2}\right], m \equiv 0 \pmod 2, \\ -\alpha(-2t) & \text{for } t \in \left[\frac{m}{2}, \frac{m+1}{2}\right], m \equiv 1 \pmod 2 \end{cases} \qquad (m \in \mathbb{Z}).$$

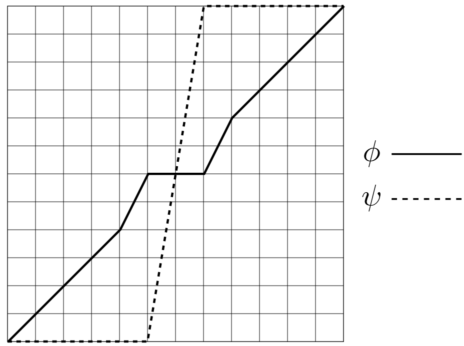


Figure 6. The graphs of ϕ and ψ given in (4.4).

We shall also denote by $\alpha_n : [0, 1] \to \mathbb{C}$ (resp. $\beta_n : [0, 1] \to \mathbb{C}$) a loop (resp. eight) traversed $n \geq 1$ times: $\alpha_n(t) = \alpha(nt)$ and $\beta_n(t) = \beta(nt)$ ($t \in [0, 1]$).

Note that $\alpha_n, \beta_n \in \mathcal{L}_{-1}^{+1}(O)$. The curvature of α_n is everywhere equal to $\frac{1}{2}$, and that of β_n equals $\pm\frac{1}{2}$ except at the $2n-1$ points where it is undefined. The turning number of α_n is n, and that of β_n is 0.

(4.4) Definition. Let $t_0 \in (0, 1)$, $0 < 2\varepsilon < \min\{1 - t_0, t_0\}$, $1 \leq n \in \mathbb{N}$ and γ be an admissible plane curve. Define piecewise linear functions $\phi, \psi : [0, 1] \to [0, 1]$ (whose graphs are depicted in Figure 6) by

$$\phi(t) = \begin{cases} t & \text{if } t \notin [t_0 - 2\varepsilon, t_0 + 2\varepsilon], \\ 2t - t_0 + 2\varepsilon & \text{if } t \in [t_0 - 2\varepsilon, t_0 - \varepsilon], \\ t_0 & \text{if } t \in [t_0 - \varepsilon, t_0 + \varepsilon], \\ 2t - t_0 - 2\varepsilon & \text{if } t \in [t_0 + \varepsilon, t_0 + 2\varepsilon], \end{cases} \tag{36}$$

$$\psi(t) = \begin{cases} 0 & \text{if } t \in [0, t_0 - \varepsilon], \\ (t - t_0 + \varepsilon)/2\varepsilon & \text{if } t \in [t_0 - \varepsilon, t_0 + \varepsilon], \\ 1 & \text{if } t \in [t_0 + \varepsilon, 1]. \end{cases}$$

Define curves A_{γ,n,t_0}, B_{γ,n,t_0} (attaching loops, eights) and $S_{\gamma,n} : [0, 1] \to \mathbb{C}$ (spreading eights) by (see Figure 7)

$$\begin{aligned} A_{\gamma,n,t_0}(t) &= \Phi_\gamma(\phi(t))\alpha_n(\psi(t)), \\ B_{\gamma,n,t_0}(t) &= \Phi_\gamma(\phi(t))\beta_n(\psi(t)), \qquad (t \in [0, 1]). \\ S_{\gamma,n}(t) &= \Phi_\gamma(t)\beta_n(t) \end{aligned}$$

Here $\Phi_\gamma : [0, 1] \to \mathbb{C} \times \mathbb{S}^1$ is the frame of γ (as in (1)), but viewed as a curve in the *group* $UT\mathbb{C}$: each $\Phi_\gamma(t)$ is an Euclidean motion, with $\Phi_\gamma(t)a = \gamma(t) + \boldsymbol{t}_\gamma(t)a$ for $a \in \mathbb{C}$. Different values of ε and t_0 yield curves which are homotopic in whichever space one is working with.

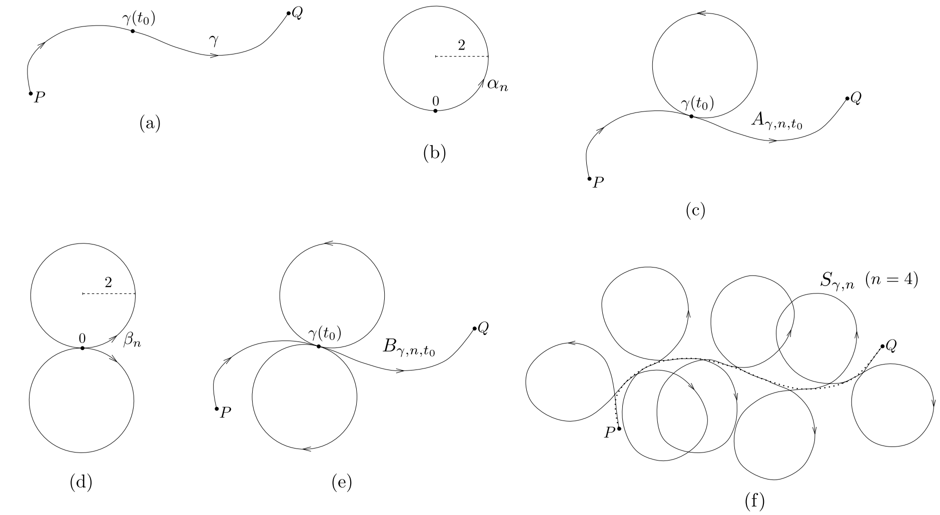


Figure 7. A depiction of how to attach loops and eights to a curve and how to spread eights along it.

(4.5) Lemma. *Let* $t_0 \in (0, 1)$, $1 \leq n \in \mathbb{N}$ *and* γ *be an admissible plane curve. Then*:

(a) A_{γ,n,t_0}, B_{γ,n,t_0} *and* $S_{\gamma,n}$ *have the same initial and final frames as* γ.

(b) B_{γ,n,t_0} *and* $S_{\gamma,n}$ *lie in the same connected component of* $\mathcal{L}_{-\infty}^{+\infty}(P, Q)$ $(P = \Phi_\gamma(0)$, $Q = \Phi_\gamma(1))$.

(c) *If* $\gamma \in \mathcal{L}_{-1}^{+1}(P, Q)$, *then* $A_{\gamma,n,t_0}, B_{\gamma,n,t_0} \in \mathcal{L}_{-1}^{+1}(P, Q)$ *also.*

(d) *Let* $O = (0, 1) \in \mathbb{C} \times \mathbb{S}^1$. *Then* α_1 *and* B_{α_1,n,t_0} *lie in the same connected component of* $\mathcal{L}_{-1}^{+1}(O)$ *for all* $n \geq 1$.

(e) *If* $f, g : K \to \mathcal{U}_d$ *are continuous and homotopic within* $\mathcal{U}_d$, *then so are* B_{f,n,t_0} *and* B_{g,n,t_0}.

(f) *If* γ *is a reparametrization of* α_1, *then* A_{γ,n,t_0} *is a reparametrization of* α_{n+1}.

Proof. It is clear that A_{γ,n,t_0}, B_{γ,n,t_0} have the same initial and final frames as γ, since they agree with γ in neighborhoods of the endpoints of $[0, 1]$. From the definition of $S_{\gamma,n}$, we find that

$$\dot{S}_{\gamma,n} = \dot{\gamma} + \dot{\boldsymbol{t}}_\gamma \beta_n + \boldsymbol{t}_\gamma \dot{\beta}_n.$$

Using that $\Phi_{\beta_n}(0) = \Phi_{\beta_n}(1) = (0, 1) \in \mathbb{C} \times \mathbb{S}^1$, we deduce that

$$S_{\gamma,n}(0) = \gamma(0) \quad \text{and} \quad \dot{S}_{\gamma,n}(0) = \big(|\dot{\gamma}(0)| + |\dot{\beta}_n(0)|\big)\boldsymbol{t}_\gamma(0).$$

Similarly, $S_{\gamma,n}(1) = \gamma(1)$ and $\dot{S}_{\gamma,n}(1)$ is a positive multiple of $\boldsymbol{t}_\gamma(1)$. This establishes (a).

Let $\phi, \psi : [0,1] \to [0,1]$ be as in (36), and set

$$\phi_s(t) = (1-s)\phi(t) + st \quad \text{and} \quad \psi_s(t) = (1-s)\psi(t) + st \quad (s, t \in [0,1]). \tag{37}$$

Then

$$(s,t) \mapsto \Phi_\gamma(\phi_s(t))\beta_n(\psi_s(t)) \quad (s, t \in [0,1])$$

defines a homotopy between B_{γ,n,t_0} and $S_{\gamma,n}$ in $\mathcal{L}_{-\infty}^{+\infty}(P, Q)$. This proves (b).

Part (c) follows from (a) and the fact that the curvatures of α_n, β_n equal $\pm\frac{1}{2}$ a.e.

Part (d) is a corollary of (4.2).

For part (e), let $H : [0,1] \times K \to \mathfrak{U}_d$ be a continuous map with $H_0 = f$ and $H_1 = g$. Set

$$\hat{H}(s,a)(t) = \Phi_{H(s,a)}(\phi(t))\beta_n(\psi(t)) \quad (s, t \in [0,1], a \in K).$$

Then $\hat{H}$ is a homotopy between $B_{f,n,t_0} = \hat{H}_0$ and $B_{g,n,t_0} = \hat{H}_1$ in $\mathfrak{U}_d$.

Part (f) is obvious. □

(4.6) Lemma. *Let $f : K \to \mathcal{C}_{-\infty}^{+\infty}(Q)$ be continuous. Then there exists $n_0 \in \mathbb{N}$ such that $S_{f(a),n} \in \mathcal{L}_{-1}^{+1}(Q)$ for all $a \in K$ whenever $n \geq n_0$ ($n \in \mathbb{N}$).*

Proof. For $a \in K$, let $\gamma_a = f(a)$ and $\boldsymbol{t}_a = \boldsymbol{t}_{\gamma_a}$. Let

$$T = \left\{\frac{1}{2n}, \frac{2}{2n}, \dots, \frac{2n-1}{2n}\right\}.$$

Then,

$$\begin{aligned}
S_{\gamma_a,n}(t) &= \Phi_{\gamma_a}(t)\beta_n(t) = \gamma_a(t) + \boldsymbol{t}_a(t)\beta(nt) && (t \in [0,1], a \in K),\\
\dot{S}_{\gamma_a,n}(t) &= \dot{\gamma}_a(t) + \dot{\boldsymbol{t}}_a(t)\beta(nt) + n\boldsymbol{t}_a(t)\dot{\beta}(nt) && (t \in [0,1], a \in K),\\
\ddot{S}_{\gamma_a,n}(t) &= \ddot{\gamma}_a(t) + \ddot{\boldsymbol{t}}_a(t)\beta(nt) + 2n\dot{\boldsymbol{t}}_a(t)\dot{\beta}(nt) + n^2\boldsymbol{t}_a(t)\ddot{\beta}(nt) && (t \in [0,1] \smallsetminus T, a \in K).
\end{aligned}$$

Since $f : K \to \mathcal{C}_{-\infty}^{+\infty}(Q)$ is continuous and K is compact, $|\gamma_a^{(k)}(t)|$ and $|\boldsymbol{t}_a^{(k)}(t)|$ ($k = 0, 1, 2$) are all bounded by some constant as (t, a) ranges over $[0,1] \times K$. Using the third expression for the curvature in (4) and the multilinearity of the determinant, we conclude that

$$\kappa_{S_{\gamma_a,n}}(t) = \frac{1}{2} + O\Big(\frac{1}{n}\Big) \quad (t \in [0,1] \smallsetminus T,\ a \in K),$$

where $O(1/n)$ is a function of (t, a) such that $n|O(1/n)|$ is uniformly bounded over $([0,1] \smallsetminus T) \times K$ as n ranges over $\mathbb{N}$. It follows that $S_{\gamma_a,n} \in \mathcal{L}_{-1}^{+1}(Q)$ for all sufficiently large n. □

(4.7) Lemma. *Let $f : K \to \mathcal{C}_{-1}^{+1}(Q)$ be continuous, $t_0 \in (0,1)$. Then for all sufficiently large $n \in \mathbb{N}$, there exists a continuous $H : [0,1] \times K \to \mathcal{L}_{-1}^{+1}(Q)$ with $H_0 = B_{f,n,t_0}$ and $H_1 = S_{f,n}$.*

Proof. Let H be given by

$$H(s,a)(t) = \Phi_{f(a)}(\phi_s(t))\beta_n(\psi_s(t)) \quad (s,t \in [0,1],\ a \in K),$$

where ϕ_s, ψ_s are as in (37). Then $H(0,a) = B_{f(a),n,t_0}$ and $H(1,a) = S_{f(a),n}$. A computation entirely similar to the one in the proof of (4.6) establishes that $H(s,a) \in \mathcal{L}_{-1}^{+1}(Q)$ for all $s \in [0,1]$, $a \in K$ if n is sufficiently large. The details will be left to the reader, but to make things easier, notice that ϕ_s, ψ_s are piecewise linear for all $s \in [0,1]$, so that $\ddot{\psi}_s = \ddot{\phi}_s = 0$ except at a finite set of points (which depends on s). □

The next result provides a sufficient condition, which does not involve g, for one to be able to write a compact family of curves f as $f = A_{g,n,t_0}$.

(4.8) Lemma. *Let X be a compact Hausdorff topological space and*

$$f : X \to \mathcal{U}_d \subset \mathcal{L}_{-1}^{+1}(Q;\theta_1), \quad t_0 : X \to (0,1)$$

be continuous maps. Then it is possible to reparametrize each $f(a)$ (continuously with a) and find a continuous $g : X \to \mathcal{L}_{-1}^{+1}(Q;\theta_1)$ so that $f(a) = A_{g(a),n,t_0(a)}$ for all $a \in X$ if and only if there exists a continuous function $\varepsilon : X \to (0,1)$ such that for all $a \in X$,

(i) $0 < t_0(a) - \varepsilon(a) < t_0(a) + \varepsilon(a) < 1$;

(ii) $f(a)|_{[t_0(a)-\varepsilon(a),\,t_0(a)+\varepsilon(a)]}$ *is some parametrization of* $\Phi_{f(a)}(t_0(a)-\varepsilon(a))\alpha_n$.

Proof. Suppose that such a function $\varepsilon : X \to (0,1)$ exists. Since X is compact, we may reparametrize all $f(a)$ so that ε becomes a constant function and, for all $a \in X$, satisfies

(I) $0 < t_0(a) - 2\varepsilon < t_0(a) + 2\varepsilon < 1$;

(II) $f(a)|_{[t_0(a)-\varepsilon,\,t_0(a)+\varepsilon]}$ is a parametrization of $\Phi_{f(a)}(t_0(a)-\varepsilon)\alpha_n$ by a multiple of arc-length.

Define $g : X \to \mathcal{L}_{-1}^{+1}(Q)$ by

$$g(a)(t) = \begin{cases} f(a)(t) & \text{if } t \notin [t_0(a)-2\varepsilon,\, t_0(a)+2\varepsilon], \\ f(a)\big(\tfrac{1}{2}(t+t_0(a)-2\varepsilon)\big) & \text{if } t \in [t_0(a)-2\varepsilon,\, t_0(a)], \\ f(a)\big(\tfrac{1}{2}(t+t_0(a)+2\varepsilon)\big) & \text{if } t \in [t_0(a),\, t_0(a)+2\varepsilon] \end{cases} \qquad (a \in X,\ t \in [0,1]).$$

Then g is continuous because f and t_0 are continuous, and $f(a) = A_{g(a),n,t_0(a)}$ for all $a \in X$. This proves the "if" part of the lemma. The converse is obvious. □

As a simple application of (4.6), we prove that this article is not a study of the empty set.

(4.9) Lemma. *Let $\kappa_1 < \kappa_2$, $P = (p, w)$, $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and let $\theta_1 \in \mathbb{R}$ satisfy $e^{i\theta_1} = z\bar{w}$. Then*:

(a) $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q) \neq \varnothing$.

(b) $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1) \neq \varnothing$ *if* $\kappa_1\kappa_2 < 0$.

(c) *If* $\kappa_1 < \kappa_2 \leq 0$, *then* $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1) = \varnothing$ *for all sufficiently large* θ_1. *If* $0 \leq \kappa_1 < \kappa_2$, *then* $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1) = \varnothing$ *for all sufficiently small* θ_1.

Proof. By (2.6), we need only consider spaces of the form $\mathcal{L}_{-1}^{+1}(Q)$, $\mathcal{L}_{0}^{+\infty}(Q)$ and $\mathcal{L}_{1}^{+\infty}(Q)$. It is clear that $\mathcal{C}_{-\infty}^{+\infty}(Q) \neq \varnothing$ for all $Q \in UT\mathbb{C}$. Let $\gamma \in \mathcal{C}_{-\infty}^{+\infty}(Q)$ be arbitrary.

By (4.6), if n is sufficiently large, then $S_{\gamma,n} \in \mathcal{L}_{-1}^{+1}(Q)$. Furthermore, attaching loops (possibly with reversed orientation) to $S_{\gamma,n}$, we can obtain a curve in $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ for any $\theta_1 \in \mathbb{R}$ satisfying $e^{i\theta_1} = z$. This proves (b), and also part (a) when $\kappa_1\kappa_2 < 0$.

Similarly, define a curve $\bar{S}_{\gamma,n}$ by $\bar{S}_{\gamma,n}(t) = \Phi_\gamma(t)\big(\tfrac{1}{4}\alpha_n(t)\big)$ ($t \in [0, 1]$). In words, $\bar{S}_{\gamma,n}$ is obtained from γ by spreading n loops of radius $\tfrac{1}{2}$, instead of n eights of radius 2. Using an argument analogous to the one which established (4.6), one sees that $\bar{S}_{\gamma,n} \in \mathcal{L}_{1}^{+\infty}(Q)$ for all sufficiently large n. This completes the proof of (a).

To see that $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ may be empty if $\kappa_1\kappa_2 \geq 0$, we use (5): if κ_1, κ_2 are both nonnegative, for example, then $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ can only contain curves having positive total turning. □

Remark. Invoking (1.11), we obtain a version of (4.9) with $\mathcal{C}$ in place of $\mathcal{L}$.

(4.10) Corollary. *Let $\mathcal{U}_d$ denote the subset of $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ consisting of all diffuse curves, where $Q = (q, z)$ and $e^{i\theta_1} = z$. Then $\mathcal{U}_d \neq \varnothing$.*

Proof. Lemma (4.5)(c) implies that $B_{\gamma,1,1/2} \in \mathcal{U}_d$ for any $\gamma \in \mathcal{L}_{-1}^{+1}(Q; \theta_1)$. Since the latter is nonempty by (4.9)(b), so is $\mathcal{U}_d$. □

(4.11) Theorem (Dubins). *Let $x > 0$, $Q = (x, 1)$ and $\eta \in \mathcal{L}_{-1}^{+1}(Q)$ be the line segment $\eta : t \mapsto xt$. Then η and $B_{\eta,1,1/2}$ lie in the same component of $\mathcal{L}_{-1}^{+1}(Q)$ if and only if $x > 4$.*

Proof. See [Dubins 1961, Theorem 5.3]. □

The next construction provides a homotopy of the straight line segment $[0, x]$ to the same segment with an eight attached which is continuous with respect to x.

(4.12) Construction. For $x > 0$, let $\eta_x : [0, x] \to \mathbb{C}$ be the line segment $t \mapsto t$. Take $t_0 = \frac{1}{2}$ in (36) and let $h : [0, 1] \times [0, 6] \to \mathbb{C}$ be a fixed homotopy between $h_0 = \eta_6$ and

$$h_1 = \Phi_{\eta_6}\Big(6\phi\Big(\frac{t}{6}\Big)\Big)\beta_1\Big(\psi\Big(\frac{t}{6}\Big)\Big) \quad (\eta_6 \text{ with an eight attached})$$

such that $t \mapsto h_s(6t)$ ($t \in [0, 1]$) is a curve in $\mathcal{L}_{-1}^{+1}(Q)$ for all $s \in [0, 1]$. The existence of h is guaranteed by (4.11). Let $\mu : [0, +\infty) \to [0, 1]$ be a smooth function such that $\mu(x) = 0$ if $x \in [0, 6]$ and $\mu(x) = 1$ if $x \geq 8$. Define a family of curves $\eta_x^u : [0, 1] \to \mathbb{C}$ by

$$(38)\quad \eta_x^u(t) = \begin{cases} \eta_x(t) & \text{if } t \geq 6 \text{ or } x \leq 6, \\ h(u\mu(x), t) & \text{if } t \leq 6 \text{ and } x \geq 6 \end{cases} \qquad (u \in [0, 1], t \in [0, x], x > 0).$$

Of course, $\eta_x^0 = \eta_x$ for all $x > 0$. If $x \geq 8$, then η_x^1 equals η_x with an eight attached; in particular, $\eta_x^1|_{[3-6\varepsilon,3]}$ is a loop.

Grafting. We now explain how to graft straight line segments onto a diffuse curve (see Figure 8).

(4.13) Definition. Let $\gamma \in \mathcal{L}_{-1}^{+1}(Q)$ be a curve of length L parametrized by arc-length, $\sigma_i \geq 0$ and $s_i \in [0, 1]$, $i = 1, \ldots, 2n$, where the s_i form a nondecreasing sequence. Suppose that there exists a bijection p of $\{1, \ldots, 2n\}$ onto itself such that for each i,

$$(*) \qquad \sigma_{p(i)} = \sigma_i \quad \text{and} \quad \boldsymbol{t}_\gamma(s_{p(i)}) = -\boldsymbol{t}_\gamma(s_i).$$

Then we define the *graft* $G_\gamma = G_{\gamma,(s_i),(\sigma_i)} : \left[0, L + \sum_{i=1}^{2n} \sigma_i\right] \to \mathbb{C}$ by

(39)

$$G_\gamma(s) = \begin{cases} \gamma(s) & \text{if } s \in [0, s_1], \\ \gamma(s_1) + (s - s_1)\boldsymbol{t}_\gamma(s_1) & \text{if } s \in [s_1, s_1 + \sigma_1], \\ \gamma(s - \sigma_1) + \sigma_1 \boldsymbol{t}_\gamma(s_1) & \text{if } s \in [s_1 + \sigma_1, s_2 + \sigma_1], \\ \gamma(s_2) + \sigma_1 \boldsymbol{t}_\gamma(s_1) + (s - s_2 - \sigma_1)\boldsymbol{t}_\gamma(s_2) & \text{if } s \in [s_2 + \sigma_1, s_2 + \sigma_1 + \sigma_2], \\ \qquad\vdots & \qquad\vdots \\ \gamma(s - \sum_{i=1}^{2n} \sigma_i) + \sum_{i=1}^{2n} \sigma_i \boldsymbol{t}_\gamma(s_i) & \text{if } s \in \left[s_{2n} + \sum_{i=1}^{2n} \sigma_i, L + \sum_{i=1}^{2n} \sigma_i\right]. \end{cases}$$

Although it simplifies the previous formula, the assumption that (s_i) is a nondecreasing sequence is not necessary for the construction to work, since we may always relabel the s_i.

(4.14) Lemma. *Let $\gamma \in \mathcal{L}_{-1}^{+1}(Q)$ be diffuse and G_γ be as in (4.13). Then G_γ is parametrized by arc-length and lies in the same connected component of $\mathcal{L}_{-1}^{+1}(Q)$ as γ.*

Proof. It is obvious from (39) that $\Phi_{G_\gamma}(0) = \Phi_\gamma(0)$. Looking at the last line of (39) and using $(*)$, we deduce that

$$G_\gamma(s) = \gamma\left(s - \sum_{i=1}^{2n} \sigma_i\right) \quad \text{for } s \in \left[s_{2n} + \sum_{i=1}^{2n} \sigma_i, L + \sum_{i=1}^{2n} \sigma_i\right].$$

Hence, $\Phi_{G_\gamma}(L + \sum_{i=1}^{2n} \sigma_i) = \Phi_\gamma(L)$. Since G_γ is made up of line segments and arcs of γ (composed with translations), $G_\gamma \in \mathcal{L}_{-1}^{+1}(Q)$. It is clear that G_γ is parametrized

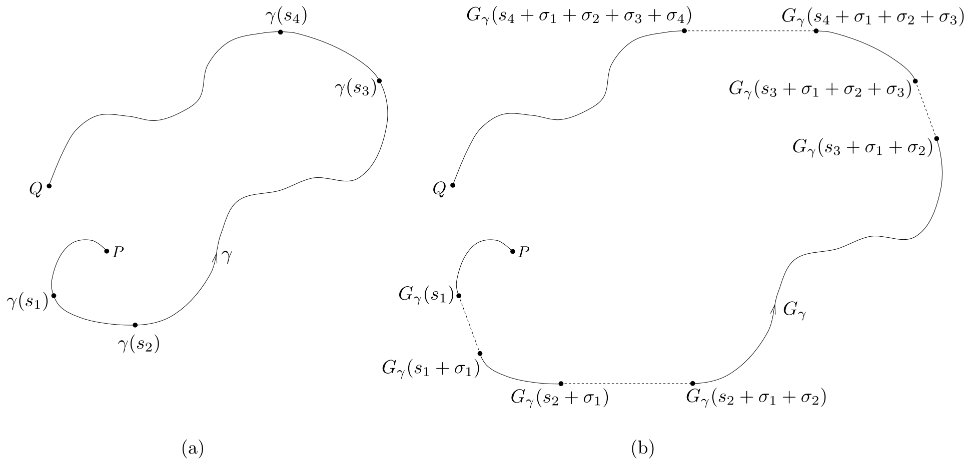


Figure 8. A diffuse curve γ and its graft $G_\gamma = G_{\gamma,(s_1,s_2,s_3,s_4),(\sigma_1,\sigma_2,\sigma_3,\sigma_4)}$.

by arc-length. Finally,

$$u \mapsto G_{\gamma,(s_i),(u\sigma_i)} \quad (u \in [0, 1])$$

defines a path in $\mathcal{L}_{-1}^{+1}(Q)$ joining γ to G_γ. $\square$

Contractibility of $\mathfrak{U}_d$. Recall that K denotes a compact manifold, possibly with boundary.

(4.15) Lemma. *Let $f : K \to \mathfrak{U}_d$ be continuous. Then there exist an open cover $(V_j)_{j=1}^m$ of K and continuous maps $\tau_j^\pm : K \to (0, 1)$, $f_1 : K \to \mathfrak{U}_d$ such that*

(i) *$f \simeq f_1$ within $\mathfrak{U}_d$ and f_1 satisfies conditions* (ii) *and* (iii) *of* (1.11).

(ii) *$\boldsymbol{t}_{f_1(a)}(\tau_j^+(a)) = -\boldsymbol{t}_{f_1(a)}(\tau_j^-(a))$ whenever $a \in V_j$.*

Proof. Apply (1.11) to f and $\mathfrak{U}_d$ to obtain f_1. The idea is to use the implicit function theorem to find $\tau_j^\pm$. However, some care must be taken since f_1 need not be differentiable with respect to a.

For each $a \in K$, let $\theta_a : [0, 1] \to \mathbb{R}$ be the argument of $\boldsymbol{t}_{f_1(a)}$ satisfying $\theta_a(0) = 0$, and set

$$\varphi_a = \tfrac{1}{2}\big(\sup_{t\in[0,1]} \theta_a(t) + \inf_{t\in[0,1]} \theta_a(t)\big).$$

Because each γ_a is diffuse and K is compact, we can find $\delta > 0$ such that

$$\theta_a([0, 1]) \supset \left(\varphi_a - \frac{\pi}{2} - \delta, \varphi_a + \frac{\pi}{2} + \delta\right) \quad \text{for all } a \in K.$$

Fix $a_0 \in K$. By Sard's theorem, we can find $\psi \in (\varphi_{a_0} + \pi/2, \varphi_{a_0} + \pi/2 + \delta)$ such that both ψ and $\psi - \pi$ are regular values of θ_{a_0}. Let $\tau^\pm(a_0) \in (0, 1)$ satisfy $\theta_{a_0}(\tau^+(a_0)) = \psi$ and $\theta_{a_0}(\tau^-(a_0)) = \psi - \pi$. No generality is lost in assuming that

$\dot\theta_{a_0}(\tau^+(a_0))>0$. From (5), $\dot\theta_a=|\dot\gamma_{f_1(a)}|\kappa_{f_1(a)}$. Thus, $\dot\theta_a$ depends continuously on a, so we can find $\mu,\varepsilon>0$ and a compact neighborhood $V\subset K$ of a_0 such that

$$\psi\in\theta_a\big((\tau^+(a_0)-\varepsilon,\tau^+(a_0)+\varepsilon)\big)\quad\text{and}\quad\dot\theta_a(t)>\mu$$

whenever $a\in V$, $|t-\tau^+(a_0)|<\varepsilon$. Hence, for each $a\in V$, there exists a *unique* $\tau^+(a)\in(\tau^+(a_0)-\varepsilon,\tau^+(a_0)+\varepsilon)$ with $\theta_a(\tau^+(a))=\psi$. We claim that the function $\tau^+:V\to(0,1)$ so defined is continuous. Consider the equation

$$\begin{aligned}\theta_a(\tau^+(b))-\theta_a(\tau^+(a))&\\ &=\big(\theta_b(\tau^+(b))-\theta_a(\tau^+(a))\big)+\big(\theta_a(\tau^+(b))-\theta_b(\tau^+(b))\big)\quad(a,b\in V).\end{aligned}$$

The first term on the right side equals 0 by the definition of τ^+, and the second converges to 0 as $b\to a$ since $\theta_b(t)$ is a uniformly continuous function of $(b,t)\in K\times[0,1]$. Hence, by the mean value theorem,

$$|\tau^+(b)-\tau^+(a)|<\frac{1}{\mu}\big|\theta_a(\tau^+(b))-\theta_a(\tau^+(a))\big|\to 0\ \text{ as }\ b\to a\quad(a,b\in V).$$

It follows that τ^+ is continuous. Similarly, reducing V if necessary, we can find a continuous function $\tau^-:V\to(0,1)$ with $\theta_a(\tau^-(a))=\psi-\pi$ for all $a\in V$. To finish the proof, cover K by finitely many such compact neighborhoods V_j, let $\tau_j^\pm:V_j\to(0,1)$ be the corresponding functions and extend each $\tau_j^\pm$ to K using the Tietze extension theorem. □

(4.16) Lemma. *Let $f:K\to\mathfrak{U}_d$ be continuous. Then there exist an open cover $(W_j)_{j=1}^m$ of K and continuous maps $t_j:K\to(0,1)$, $g_j:W_j\to\mathcal{L}_{-1}^{+1}(Q)$ and $f_2:K\to\mathfrak{U}_d$ such that*

(i) *$f\simeq f_2$ within $\mathfrak{U}_d$;*

(ii) *$f_2(a)=A_{g_j(a),1,t_j(a)}$ for all $a\in W_j$.*

Proof. Take f_1 as in (4.15). By (1.15), we may assume that each map $\gamma_a=f_1(a):[0,L_a]\to\mathbb{C}$ is parametrized by arc-length, so that now $\tau_j^\pm(a)\in(0,L_a)$ for each a. Let $(\lambda_j)_{j=1}^m$ be a partition of unity subordinate to $(V_j)_{j=1}^m$, with V_j as in (4.15). Set $\sigma_j=10m\lambda_j$ and $W_j=\{a\in K:\sigma_j(a)>8\}$. Then $\overline{W}_j\subset V_j$ and the W_j form an open cover of K. Define

$$\gamma_a^u=G_{\gamma_a,(\tau_1^-(a),\dots,\tau_m^-(a),\tau_1^+(a),\dots,\tau_m^+(a)),(u\sigma_1(a),\dots,u\sigma_m(a),u\sigma_1(a),\dots,u\sigma_m(a))}\quad(u\in[0,1],a\in K)$$

as in (4.13). Let us suppose that $\tau_1^-\le\cdots\le\tau_m^-(a)\le\tau_1^+(a)\le\cdots\le\tau_m^+(a)$ for each a to abbreviate the notation, and set

$$\xi_j^-(a)=\sum_{i<j}\sigma_i(a)\quad\text{and}\quad\xi_j^+(a)=10m+\sum_{i<j}\sigma_i(a)\quad(a\in K,\ j=1,\dots,m).$$

Then

$$\gamma_a^1\big([\tau_j^-(a)+\xi_j^-(a),\,\tau_j^-(a)+\xi_j^-(a)+\sigma_j(a)]\big)$$

is a line segment, corresponding to the graft at $\gamma_a(\tau_j^-(a))$. Its length $\sigma_j(a)$ is at least 8 if $a \in \overline{W}_j$. Of course, the same statements hold with $^+$ instead of $^-$. We obtain f_2 by deforming all of these segments to eights. More precisely, for $u \in [1, 2]$ and $a \in K$, let

$$\gamma_a^u(s) = \begin{cases} \Phi_{\gamma_a^1}(\tau_j^\pm(a) + \xi_j^\pm(a))\eta_{\sigma_j(a)}^{u-1}(s - \tau_j^\pm(a) - \xi_j^\pm(a)), \\ \gamma_a^1(s) \end{cases} \qquad (s \in [0, L_a+20m])$$

according to whether $s \in [\tau_j^\pm(a) + \xi_j^\pm(a), \tau_j^\pm(a) + \xi_j^\pm(a) + \sigma_j(a)]$ for some j or not, respectively. Here η_x^u is as in (4.12). Let $f_2 : K \to \mathcal{U}_d$ be given by $f_2(a) = \gamma_a^2$. Note that

$$\gamma_a^2\big([\tau_j^\pm(a) + \xi_j^\pm(a) + 3 - 6\varepsilon, \tau_j^\pm(a) + \xi_j^\pm(a) + 3]\big) \quad (j = 1, \dots, m)$$

is a loop whenever $a \in \overline{W}_j$. Thus (after reparametrizing the γ_a^2 so that their domains become $[0, 1]$), we may apply (4.8) to each family $f_2|_{\overline{W}_j}$ to find $g_j : \overline{W}_j \to \mathcal{L}_{-1}^{+1}(Q)$ and $t_j : \overline{W}_j \to (0, 1)$ such that

$$f_2(a) = A_{g_j(a),1,t_j(a)} \quad \text{for all } a \in W_j.$$

The functions t_j may be extended to all of K by the Tietze extension theorem. □

(4.17) Lemma. *Let $f : K \to \mathcal{U}_d$ be continuous. Suppose that there exist a covering of K by open sets W_j and continuous maps $t_j : K \to (0, 1)$, $g_j : W_j \to \mathcal{L}_{-1}^{+1}(Q)$ with $f(a) = A_{g_j(a),1,t_j(a)}$ whenever $a \in W_j$, $j = 1, \dots, m$. Then there exist continuous $g : K \to \mathcal{L}_{-1}^{+1}(Q)$ and $H : [0, 1] \times K \to \mathcal{U}_d$ with $H_0 = f$ and $H_1 = A_{g,1,1/2}$.*

Proof. The proof will be by induction on m. If $m = 1$ then $W_1 = K$, and H just slides the loop from t_1 to $\frac{1}{2}$:

$$H(s, a) = A_{g_1(a),1,(1-s)t_1(a)+s/2} \quad (s \in [0, 1], a \in K).$$

Suppose now that $m > 1$. Let W be an open set such that $\overline{W} \subset W_m$ and

$$W_1 \cup \cdots W_{m-1} \cup W = K.$$

Let $\lambda : K \to [0, 1]$ be a continuous function such that $\lambda(a) = 1$ if $a \in W$ and $\lambda(a) = 0$ if $a \notin W_m$. Define $\hat{H} : [0, 1] \times K \to \mathcal{U}_d$ by

$$\hat{H}(s, a) = \begin{cases} A_{g_m(a),1,(1-\lambda(a)s)t_m(a)+\lambda(a)st_{m-1}(a)} & \text{if } a \in W_m, \\ f(a) & \text{if } a \notin W_m \end{cases} \qquad (s \in [0, 1], a \in K).$$

Then the induction hypothesis applies to $\hat{H}_1 : K \to \mathcal{U}_d$, the open sets $\hat{W}_i = W_i$ ($i = 1, \dots, m-2$) and $\hat{W}_{m-1} = W_{m-1} \cup W$, and the same functions t_j as before, $j = 1, \dots, m-1$. The existence of $\hat{g}_{m-1}$ as in the statement is guaranteed by (4.8): using (4.5)(f), we deduce that there is at least one loop at $t_{m-1}(a)$ for $a \in \hat{W}_{m-1}$. □

(4.18) Proposition. *Let $f : K \to \mathcal{U}_d$ be continuous. Then $f \simeq B_{f,n,1/2}$ within $\mathcal{U}_d$ for all $n \geq 1$.*

Proof. Applying (4.16) and (4.17) to f, we obtain continuous maps $g : K \to \mathcal{L}_{-1}^{+1}(Q)$ and $h : K \to \mathcal{U}_d$ such that $f \simeq h$ in $\mathcal{U}_d$ and

$$h(a) = A_{g(a),1,1/2} \quad \text{for all } a \in K.$$

Using (4.5)(d), we may deform the loop at $t = \frac{1}{2}$ to attach n eights to h at $t = \frac{1}{2}$ (for arbitrary $n \geq 1$). Thus $h \simeq B_{h,n,1/2}$. Together with (4.5)(e), this implies that $f \simeq B_{f,n,1/2}$ within $\mathcal{U}_d$. □

(4.19) Theorem. *Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and $\theta_1 \in \mathbb{R}$ satisfy $e^{i\theta_1} = z$. Then the subspace $\mathcal{U}_d \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ consisting of all diffuse curves is homeomorphic to $\boldsymbol{E}$, and hence contractible.*

Proof. Because $\mathcal{U}_d$ is open, it suffices to prove that it is weakly contractible, by (1.7)(b). Let $k \in \mathbb{N}$, $f : \mathbb{S}^k \to \mathcal{U}_d$ be continuous and $g : \mathbb{S}^k \to \mathcal{U}_d$ be a map satisfying (i)–(iii) of (1.11) (with $\mathcal{U} = \mathcal{U}_d$). By (4.1), there exists $G : [0, 1] \times \mathbb{S}^k \to \mathcal{C}_{-\infty}^{+\infty}(Q)$ such that $G_0 = g$ and G_1 is a constant map. By (4.6), there exists $n_0 \in \mathbb{N}$ such that if $n \geq n_0$, then $S_{G(s,a),n} \in \mathcal{U}_d$ for all $s \in [0, 1]$, $a \in \mathbb{S}^k$. Applying (4.18) and (4.7) to g, we obtain $n_1 \geq n_0$ and a continuous $F : [-1, 0] \times \mathbb{S}^k \to \mathcal{U}_d$ with $F_{-1} = g$ and $F_0 = S_{g,n_1}$. Concatenating F and S_{G,n_1} we obtain a null-homotopy of g in $\mathcal{U}_d$. □

5. Critical curves

Fix $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and $\theta_1 \in \mathbb{R}$ satisfying $e^{i\theta_1} = z$. Let $\gamma \in \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ and $\theta : [0, 1] \to \mathbb{R}$ be the argument of $\boldsymbol{t}_\gamma$ satisfying $\theta(0) = 0$. Finally, let

$$\theta^+ = \sup_{t \in [0,1]} \theta(t) \quad \text{and} \quad \theta^- = \inf_{t \in [0,1]} \theta(t). \tag{40}$$

Recall that γ is called *critical* if $\theta^+ - \theta^- = \pi$. A curve $\eta \in \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ must be either condensed, diffuse or critical. It has already been shown that the subspace $\mathcal{U}_c$ (resp. $\mathcal{U}_d$) of $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ consisting of all condensed (resp. diffuse) curves is contractible. Let $\mathcal{T} \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ denote the subspace of all critical curves. Clearly, $\mathcal{T}$ is closed as the complement of $\mathcal{U}_c \cup \mathcal{U}_d$. Since the difference $\theta^+ - \theta^-$ depends continuously on γ, we deduce that $\partial\mathcal{U}_c \subset \mathcal{T}$ and $\partial\mathcal{U}_d \subset \mathcal{T}$, where $\partial\mathcal{U}_c$ denotes the topological boundary of $\mathcal{U}_c$ considered as a subspace of $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ and similarly for $\mathcal{U}_d$.

(5.1) Proposition. *Let $|\theta_1| < \pi$ and $\mathcal{U}_c, \mathcal{U}_d, \mathcal{T} \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ be as above. Then $\partial\mathcal{U}_c = \partial\mathcal{U}_d = \mathcal{T}$. Therefore, $\overline{\mathcal{U}}_c \cup \overline{\mathcal{U}}_d = \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ and $\overline{\mathcal{U}}_c \cap \overline{\mathcal{U}}_d = \mathcal{T}$.*

Observe that $\mathcal{T} = \varnothing$ if $|\theta_1| > \pi$ and $\mathcal{U}_c = \varnothing$ if $|\theta_1| \geq \pi$. However, in any case $\partial\mathcal{U}_d = \mathcal{T}$.

Proof. Let $\gamma \in \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ be a critical curve and $\mathcal{V} \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ be an open set containing γ. Let θ be the argument of $\boldsymbol{t}_\gamma$ satisfying $\theta(0) = 0$, and let θ^+, θ^- be as in (40).

We first prove that $\mathcal{V} \cap \mathcal{U}_c \neq \varnothing$. Our immediate objective is to replace γ with another curve in $\mathcal{V} \cap \mathcal{T}$ having smaller curvature. Choose $t_1 \in (0, 1)$ and $\delta > 0$ such that $\theta(t) \in (\theta^-, \theta^+)$ for all $t \in [t_1 - \delta, t_1]$. Let $Q_0 = \Phi_\gamma(t_1 - \delta)$, $Q_1 = \Phi_\gamma(t_1)$ and consider the map

$$F : \mathcal{L}_{-1}^{+1}(Q_0, \cdot\,) \to UT\mathbb{C}, \quad F(\eta) = \Phi_\eta(1).$$

(Recall that $\mathcal{L}_{-1}^{+1}(Q_0, \cdot\,)$ consists of all $(-1, 1)$-admissible curves having initial frame equal to Q_0 and arbitrary final frame.) By (1.5), F is an open map. It follows that for any $\tilde{Q}_1$ close enough to Q_1, we can find $\eta \in \mathcal{L}_{-1}^{+1}(Q_0, \tilde{Q}_1)$ such that

$$\theta_\eta([0, 1]) \subset (\theta^-, \theta^+). \tag{41}$$

Let $Q_1 = (q_1, z_1)$ and $Q = (q, z)$. Since γ is critical, the image of $\boldsymbol{t}_\gamma$ is contained in a semicircle. Consequently, $q \neq 0$. Choose $\kappa_0 \in (0, 1)$ close to 1. Replace the arc $\gamma|_{[t_1 - \delta, t_1]}$ by a curve η as above with $\tilde{Q}_1 = (q_1 + (\kappa_0 - 1)q, z_1)$, and the arc $\gamma|_{[t_1, 1]}$ by its translate $\gamma|_{[t_1,1]} + (\kappa_0 - 1)q$. Let γ_1 be the resulting curve; observe that γ_1 is critical, $\Phi_{\gamma_1}(0) = (0, 1)$ and $\Phi_{\gamma_1}(1) = (\kappa_0 q, z)$. Set $\gamma_2 = (1/\kappa_0)\gamma_1$ (that is, $\gamma_2(t)$ is obtained from $\gamma_1(t)$ by a dilatation through a factor of $1/\kappa_0$ for all $t \in [0, 1]$). Then

$$\Phi_{\gamma_2}(0) = (0, 1), \quad \Phi_{\gamma_2}(1) = (q, z),$$
$$\boldsymbol{t}_{\gamma_2}(t) = \boldsymbol{t}_{\gamma_1}(t), \quad \kappa_{\gamma_2}(t) = \kappa_0 \kappa_{\gamma_1}(t) \quad \text{for all } t \in [0, 1].$$

Thus, γ_2 is a critical curve in $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ whose curvature is constrained to $(-\kappa_0, \kappa_0)$. Moreover, if κ_0 is close enough to 1 and η is chosen appropriately, we can guarantee that $\gamma_2 \in \mathcal{V}$.

Having established the existence of γ_2 with these properties, let us return to the beginning, setting $\gamma = \gamma_2 \in \mathcal{V}$. Since $|\theta_1| < \pi$, either

$$\theta^{-1}(\{\theta^-\}) \cap \{0, 1\} = \varnothing \quad \text{or} \quad \theta^{-1}(\{\theta^+\}) \cap \{0, 1\} = \varnothing,$$

and we lose no generality in assuming the latter. Choose $\varepsilon > 0$ small enough to guarantee that

$$W = \theta^{-1}\big((\theta^+ - \varepsilon, \theta^+]\big) \subset (0, 1).$$

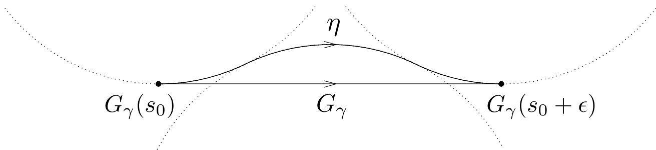
Cover $\theta^{-1}(\{\theta^+\})$ by the finite union of disjoint intervals $(a_i, b_i) \subset W$ with $\theta(a_i) = \theta(b_i) = \theta^+ - \varepsilon$, $i = 1, \dots, m$. Let $P_i = \Phi_\gamma(a_i)$, $Q_i = \Phi_\gamma(b_i)$. We can obtain a curve in $\mathcal{U}_c \cap \mathcal{V}$ by modifying γ in each of these intervals to avoid the argument θ^+ using (3.4): note that $P_i^{-1}\gamma|_{[a_i, b_i]}$ satisfies the hypotheses of (3.16) because it has curvature in the open interval $(-\kappa_0, +\kappa_0)$ and is not a line segment. Moreover, the inclusion $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(P_i^{-1}Q_i) \to \mathcal{L}_{-1}^{+1}(P_i^{-1}Q_i)$ is continuous by (1.17).

The proof that $\mathcal{V} \cap \mathcal{U}_d \neq \varnothing$ is easier. Let the critical curve $\gamma : [0, L] \to \mathbb{C}$ be parametrized by arc-length. Then we can find $s_0, s_1 \in [0, 1]$ with $\boldsymbol{t}_\gamma(s_0) = -\boldsymbol{t}_\gamma(s_1)$.

Choose $\varepsilon > 0$ and let

$$G_\gamma = G_{\gamma,(s_0,s_1),(\varepsilon,\varepsilon)}.$$

(See (4.13) and Figure 8.) Choose $\kappa_0 \in (0, 1)$ and construct a curve $\zeta \in \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ by replacing the line segment $G_\gamma|_{[s_0,s_0+\varepsilon]}$ by three small arcs of circles of radius $1/\kappa_0$ as indicated below:



If the bump is chosen to lie on the correct side, the curve ζ will be diffuse, and if $\varepsilon > 0$ is small enough, then $\zeta \in \mathcal{V}$. (Notice that this part of the proof works even if $|\theta_1| = \pi$.)

We have established that $\mathcal{T} \subset \partial\mathcal{U}_d \cap \partial\mathcal{U}_c$. As explained at the beginning of the section, $\partial\mathcal{U}_c \subset \mathcal{T}$ and $\partial\mathcal{U}_d \subset \mathcal{T}$. Thus, $\partial\mathcal{U}_c = \partial\mathcal{U}_d = \mathcal{T}$. □

Existence of critical curves. It is immediate from the definition of "critical curve" that if $|\theta_1| > \pi$, then the subspace $\mathcal{T} \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ must be empty. In this subsection we shall determine exactly when $\mathcal{T} = \varnothing$ for $|\theta_1| \le \pi$.

(5.2) Definition. A *sign string* σ is an alternating finite sequence of signs, such as $+-+$ or $-+-+$. As part of the definition we require that its *length* $|\sigma|$, the number of terms in the string, satisfy $|\sigma| \ge 2$. Let $\sigma(k)$ denote its k-th term ($1 \le k \le |\sigma|$). The *opposite* $-\sigma$ of σ is the unique sign string satisfying $|-\sigma| = |\sigma|$ and $(-\sigma)(k) = -\sigma(k)$.

A critical curve $\gamma : [0, 1] \to \mathbb{C}$ is *of type* σ if there exist $0 \le t_1 < t_2 < \cdots < t_{|\sigma|} \le 1$ with $\theta(t_k) = \theta^{\sigma(k)}$ (recall that $\theta^+ = \sup_{t\in[0,1]} \theta(t)$ and $\theta^- = \inf_{t\in[0,1]} \theta(t)$, where $e^{i\theta} = \boldsymbol{t}_\gamma$), but it is impossible to find $0 \le s_1 < \cdots < s_{|\sigma|+1} \le 1$ such that $\boldsymbol{t}_\gamma(s_k) = -\boldsymbol{t}_\gamma(s_{k+1})$ for each $k = 1, \ldots, |\sigma|$.

Given a sign string σ, one can determine whether there exist critical curves of type σ in $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ using an elementary geometric construction; see Figure 9.

(5.3) Proposition. *Let* $\theta_1 \in [0, \pi]$, $z = e^{i\theta_1}$ *and* $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$. *Let* σ *be a sign string,*

$$a = i\sigma(1)(1 + (-1)^{|\sigma|+1} z) \in \mathbb{C} \quad \textit{and} \quad r = 2|\sigma| \in \mathbb{N}.$$

Let R_σ *be the open region of the plane which does* not *contain* $-i + iz$ *and which is bounded by the shortest arc of* $C_r(a)$ *joining* $a + ri$ *to* $a - riz$ *and the tangent lines to* $C_r(a)$ *at these points. Then* $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ *contains critical curves of type* σ *if and only if* $q \in R_\sigma$.

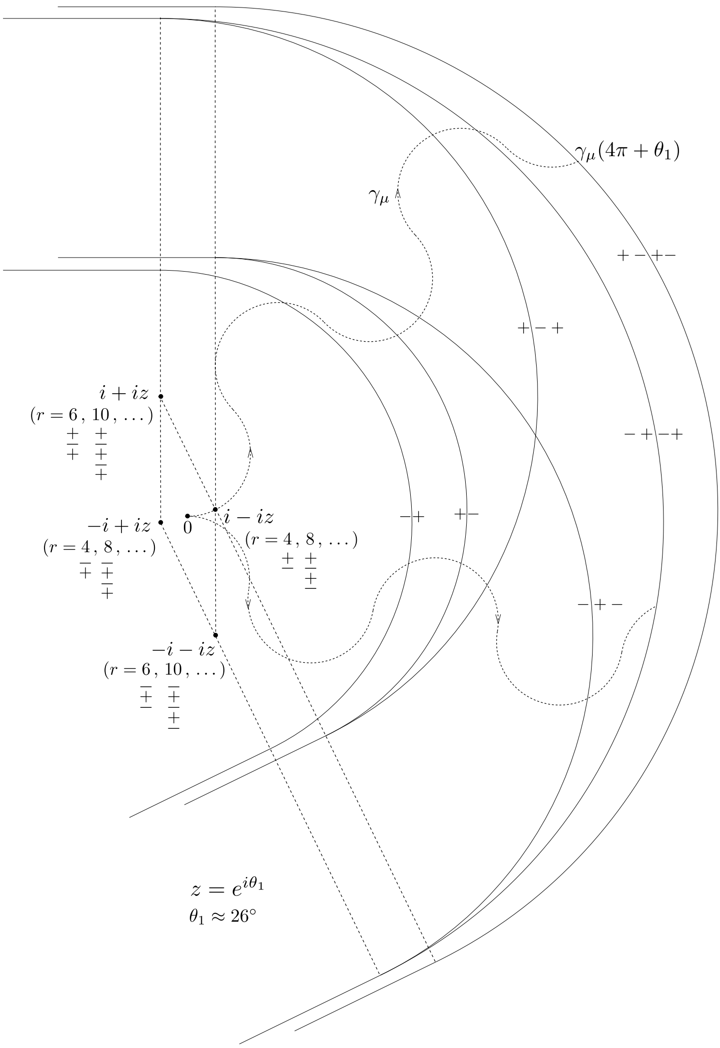


Figure 9. The regions R_σ of (5.3).

We have assumed that $\theta_1 \in [0, \pi]$ just to simplify the statement. If $\theta_1 \in [-\pi, 0]$, then the only differences are that the points bounding the arc of $C_r(a)$ are now $a-ri$ and $a+riz$ and the region R_σ is the one not containing $i-iz$. Indeed, reflection across the x-axis yields a homeomorphism between $\mathcal{L}_{-1}^{+1}(Q;\theta_1)$ and $\mathcal{L}_{-1}^{+1}(\bar{Q};-\theta_1)$, where $\bar{Q}=(\bar{q},\bar{z})$, which maps critical curves of type σ to critical curves of type $-\sigma$.

When $\theta_1 = 0$, the points $a+ri$ and $a-ri$ determine two shortest arcs of $C_r(a)$, not just one; the region R_σ is bounded by the one which goes through $a+r$. When $\theta_1 = \pm\pi$, the arc of circle degenerates to a single point. In this case, R_σ is the component of the complement of the horizontal line through $a + \operatorname{sign}(\theta_1) r i$ which does not contain the real axis.

Proof of (5.3). There are four essentially distinct types of sign strings to consider:

$$\underbrace{+-\cdots+-}_{2n},\quad \underbrace{-+\cdots-+}_{2n},\quad +\underbrace{-+\cdots-+}_{2n}\quad \text{and}\quad -\underbrace{+-\cdots+-}_{2n}\quad (n\in\mathbb{N},\, n\geq 1).$$

(Note that these are distinguished by the values of $\sigma(1)$ and $|\sigma|$ appearing in the expression for a.) We shall prove the theorem for a string of the first type; the proof in the remaining three cases is analogous. The argument given here is the same as the one which was used to prove (3.17), so some details will be omitted.

For each $\mu \in [\theta_1 - \pi, 0]$, let $\gamma_\mu : [0, 2n\pi + \theta_1] \to \mathbb{C}$ be the unique curve parametrized by arc-length satisfying

$$\gamma_\mu(0) = 0,$$
$$\boldsymbol{t}_{\gamma_\mu}(s) = \begin{cases} e^{is} & \text{if } s \in [0, \mu+\pi] \cup [\mu+2n\pi, \theta_1+2n\pi] \bigcup_k [\mu+k\pi, \mu+(k+1)\pi], \\ e^{i(2\mu-s)} & \text{if } s \in \bigcup_k [\mu+k\pi, \mu+(k+1)\pi], \end{cases}$$

where the first (resp. second) union is over all $k \equiv 0$ (resp. $k \equiv 1$) $(\mathrm{mod}\, 2)$, $1 \leq k \leq 2n-1$. Notice that γ_μ is the concatenation of arcs of circles of radius 1; see Figure 9. (Vaguely speaking, γ_μ is the "most efficient" critical curve γ of type σ with $\inf \theta_\gamma = \mu$ and $|\kappa_\gamma| \leq 1$.) We have

$$\Phi_{\gamma_\mu}(0) = (0, 1), \quad \boldsymbol{t}_{\gamma_\mu}(2n\pi + \theta_1) = z, \quad \inf \theta_{\gamma_\mu} = \mu, \quad \sup \theta_{\gamma_\mu} = \mu + \pi,$$
$$\gamma_\mu(2n\pi + \theta_1) = \left(\int_0^{\mu+\pi} + (2n-1)\int_\mu^{\mu+\pi} + \int_\mu^{\theta_1}\right) e^{is}\, ds = (i - iz) + 4nie^{i\mu}.$$

From the previous equation it follows that as μ increases from $\theta_1 - \pi$ to 0, the endpoint of γ_μ traces out the arc of $C_r(a)$ joining $a - riz$ to $a + ri$, where $a = i - iz$ and $r = 4n = 2|\sigma|$. Further, the tangent line to $C_r(a)$ at $\gamma_\mu(2n\pi + \theta_1)$ is parallel to $e^{i\mu}$, for it must be orthogonal to $4nie^{i\mu}$.

It is easy to see that any $q \in \bar{R}_\sigma$ is the endpoint of a curve of one of the following three types:

(i) The concatenation of γ_μ with a line segment of direction z for some $\mu \in [\theta_1 - \pi, 0]$.

(ii) The concatenation of $\gamma_0|_{[0,\pi]}$, a line segment of length $\ell \geq 0$ having direction -1, the arc $-\ell + \gamma_0|_{[\pi, 2n\pi+\theta_1]}$, and a line segment of direction z.

(iii) The concatenation of $\gamma_{\theta_1-\pi}|_{[0,\theta_1+\pi]}$, a line segment of length $\ell_1 \geq 0$ of direction $-z$, the arc $-\ell_1 z + \gamma_0|_{[\theta_1+\pi, \theta_1+3\pi/2]}$, a line segment of length $\ell_2 \geq 0$ and direction $-iz$, the arc $-\ell_1 z - \ell_2 iz + \gamma_0|_{[\theta_1+3\pi/2, 2n\pi+\theta_1]}$, and a line segment of direction z.

If $q \in R_\sigma$, then we can find a critical curve γ of type σ in $\mathcal{L}_{-1}^{+1}(Q;\theta_1)$ by a slight modification of one of these curves.

Conversely, suppose that $\mathcal{L}_{-1}^{+1}(Q;\theta_1)$ contains critical curves of type σ. Let $\eta\colon[0,L]\to\mathbb{C}$ be such a curve, parametrized by arc-length, and let $\mu=\inf\theta$, where $\theta\colon[0,L]\to\mathbb{R}$ is the argument of $\boldsymbol{t}_\eta$ satisfying $\theta(0)=0$. Define

$$g\colon[0,L]\to\mathbb{R}\quad\text{by}\quad g(s)=\big\langle\eta(s)-\gamma_\mu(2n\pi+\theta_1),\,ie^{i\mu}\big\rangle.$$

Note that $g(s)>0$ if and only if $\eta(s)$ lies to the left of the line through $\gamma_\mu(2n\pi+\theta_1)\in C_r(a)$ having direction $e^{i\mu}$; we have already seen that this line is tangent to $C_r(a)$ at this point. We claim that $g(L)>0$. Since η is critical, $\theta(s)\in[\mu,\mu+\pi]$ for all s. Hence,

$$g'(s)=\langle e^{i\theta(s)},ie^{i\mu}\rangle=\cos\big(\theta(s)-\big(\mu+\tfrac{\pi}{2}\big)\big)\geq 0\quad\text{for all } s\in[0,L]. \tag{42}$$

Let $J_i=(a_i,b_i)\subset(0,L)$, $i=0,\dots,2n=|\sigma|$, be disjoint intervals such that

(I) $\theta(a_0)=0$ and $\theta(b_1)=\mu+\pi$;

(II) $\theta(a_i)=\mu+\pi$ and $\theta(b_i)=\mu$ for $i=1,3,\dots,2n-1$;

(III) $\theta(a_i)=\mu$ and $\theta(b_i)=\mu+\pi$ for $i=2,4,\dots,2n-2$;

(IV) $\theta(a_{2n})=\mu+\pi$ and $\theta(b_{2n})=\theta_1$.

Such intervals exist because $\theta([0,L])\subset[\mu,\mu+\pi]$, $\theta(0)=0$, $\theta(L)=\theta_1$ and η is critical of type σ. It follows from (42) and the fact θ is strictly 1-Lipschitz (by (5) and the fact that $\theta=\arg\circ\boldsymbol{t}_\eta$ is absolutely continuous) that

$$\begin{aligned} g(L)-g(0) &\geq \Big(\sum_{i=0}^{2n}\int_{a_i}^{b_i}\Big)g'(s)\,ds\\ &> \Big(\int_0^{\mu+\pi}+(2n-1)\int_\mu^{\mu+\pi}+\int_\mu^{\theta_1}\Big)\langle e^{it},ie^{i\mu}\rangle\,dt=\big\langle\gamma_\mu(2n\pi+\theta_1),ie^{i\mu}\big\rangle. \end{aligned}$$

Therefore, $g(L)>0$ as claimed. We conclude that $q=\eta(L)$ lies on the side of a tangent to $C_r(a)$ which only contains points of R_σ. $\square$

(5.4) Corollary. *Let $Q=(q,z)\in\mathbb{C}\times\mathbb{S}^1$,*

$$a=\operatorname{sign}(\operatorname{Im}(z))i(z-1)\in\mathbb{C},$$

and let $R_{\mathfrak{I}}$ be the open region of the plane which does not *contain a and which is bounded by the shortest arc of $C_4(a)$ joining $a+\operatorname{sign}(\operatorname{Im}(z))4i$ to $a-\operatorname{sign}(\operatorname{Im}(z))4iz$ and the tangent lines to $C_4(a)$ at these points. Then $\mathcal{L}_{-1}^{+1}(Q)$ contains critical curves if and only if $q\in R_{\mathfrak{I}}$.*

Proof. Let $z=e^{i\theta_1}$, $|\theta_1|\leq\pi$. For $\theta_1\in[0,\pi]$ (resp. $\theta_1\in[-\pi,0]$), $R_{\mathfrak{I}}$ is the same as the region R_{-+} (resp. R_{+-}) appearing in (5.3). $\square$

If $z = \pm 1$, then $\operatorname{sign}(\operatorname{Im} z)$ is not defined. When $z = 1$, $R_{\mathfrak{T}}$ is bounded by the semicircle centered at 0 through 4 and $\pm 4i$ and the tangents to $C_4(0)$ at the latter two points. When $z = -1$, $R_{\mathfrak{T}}$ is bounded by the horizontal lines through $\pm 2i$.

(5.5) Corollary. *Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and $e^{i\theta_1} = z$. Then there exist condensed curves but not critical curves in $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ if and only if $|\theta_1| < \pi$ and q lies in the region illustrated in Figure 1.*

Proof. This is an immediate consequence of (3.17) and (5.4). □

(5.6) Lemma. *Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and $e^{i\theta_1} = z$, $|\theta_1| \le \pi$. Let $\omega \in [|\theta_1|, \pi]$ and $r(\omega) = 4\sin\big(\omega/2\big)$. Suppose that q lies inside of $C_{r(\omega)}\big(\operatorname{sign}(\theta_1) i (z-1)\big)$. Then there does not exist a curve in $\hat{\mathcal{L}}_{-1}^{+1}(Q; \theta_1)$ or $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ having amplitude ω.*

Proof. Assume that $\theta_1 \in [0, \pi]$; the proof for $\theta_1 \in [-\pi, 0]$ is analogous. Let $\omega \in [\theta_1, \pi]$, $\mu \in [\theta_1 - \omega, 0]$ and let $\gamma_\mu : [0, 2\omega - \theta_1] \to \mathbb{C}$ be the unique curve parametrized by arc-length satisfying

$$\gamma_\mu(0) = 0 \quad \text{and} \quad \boldsymbol{t}_{\gamma_\mu}(s) = \begin{cases} e^{-is} & \text{if } s \in [0, -\mu], \\ e^{i(s+2\mu)} & \text{if } s \in [-\mu, -\mu+\omega], \\ e^{-i(s-2\omega)} & \text{if } s \in [-\mu+\omega, 2\omega - \theta_1]. \end{cases}$$

Notice that $\boldsymbol{t}_{\gamma_\mu}(0) = 1$, $\boldsymbol{t}_{\gamma_\mu}(2\omega - \theta_1) = z$ and γ_μ is a concatenation of three arcs of circles of radius 1. Moreover, $\inf \theta_{\gamma_\mu} = \mu$ and $\sup \theta_{\gamma_\mu} = \mu + \omega$, where θ_{γ_μ} is the argument of $\boldsymbol{t}_{\gamma_\mu}$ satisfying $\theta_{\gamma_\mu}(0) = 0$. Consequently, γ_μ has amplitude ω. Further,

$$\gamma_\mu(2\omega - \theta_1) = \left(\int_\mu^0 + \int_\mu^{\mu+\omega} + \int_{\theta_1}^{\mu+\omega}\right) e^{is}\, ds = (-i + iz) + 4\sin\left(\frac{\omega}{2}\right) e^{i(\mu+\omega/2)}.$$

Thus, as μ increases from $\theta_1 - \omega$ to 0, the endpoint of γ_μ traverses an arc of $C_{r(\omega)}(-i + iz)$. Suppose that there exists $\eta \in \hat{\mathcal{L}}_{-1}^{+1}(Q; \theta_1)$ of amplitude ω, and let $\eta : [0, L] \to \mathbb{C}$ be parametrized by arc-length. Let θ_η be the argument of η satisfying $\theta_\eta(0) = 0$, take $\mu = \inf \theta_\eta$ and define

$$g : [0, L] \to \mathbb{R} \quad \text{by} \quad g(s) = \big\langle \eta(s) - \gamma_\mu(2\omega - \theta_1),\, e^{i(\mu+\omega/2)} \big\rangle.$$

Then the same reasoning used to establish (3.17) and (5.3) shows that $g(L) \ge 0$. This implies that $\eta(L) = q$ lies on or to the left of the line through $\gamma_\mu(2\omega - \theta_1)$ having direction $\exp\big(i(\mu + (\omega - \pi)/2)\big)$. This line is tangent to $C_{r(\omega)}(-i + iz)$ at this point; therefore q cannot lie inside of this circle. This proves the assertion about $\hat{\mathcal{L}}_{-1}^{+1}(Q; \theta_1)$. Since the latter contains $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ as a subset, the proof is complete. □

The next result will only be needed in [Saldanha and Zühlke 2015]. Recall the definition of $\bar{\varphi}_\gamma$ in (34).

(5.7) Corollary. *Let $|\theta_1| \leq \pi$, $e^{i\theta_1} = z$, $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and σ be a sign string. Then the set of all $\varphi \in \mathbb{R}$ such that there exists a critical curve $\gamma \in \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ of type σ for which $\bar{\varphi}_\gamma = \varphi$ is an open interval.*

Proof. No generality is lost in assuming that $\theta_1 \in [0, \pi)$. Let $\varphi \in \mathbb{R}$. It was established in the proof of (5.3) that a curve γ as in the statement exists if and only if $\varphi \in [\theta_1 - \pi/2, \pi/2]$ and q lies in the open external region E_φ determined by the tangent orthogonal to $e^{i\varphi}$ to a certain circle C (which depends only on θ_1 and σ). It is straightforward to check that $E_{\varphi_1} \cap E_{\varphi_2} \subset E_\varphi$ whenever $\varphi \in [\varphi_1, \varphi_2]$ with $\varphi_2 - \varphi_1 \leq \pi$. Moreover, if $q \in E_\varphi$, then $q \in E_{\tilde{\varphi}}$ for all $\tilde{\varphi}$ sufficiently close to φ. $\square$

6. Components of $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ for $\kappa_1\kappa_2 < 0$

Recall that $\boldsymbol{E}$ denotes the separable Hilbert space and $B_{\gamma,1,1/2}$ is obtained from γ by attaching a figure eight curve (at $t = 1/2$); see (4.4) and Figure 7(d).

(6.1) Theorem. *Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and $\theta_1 \in \mathbb{R}$ satisfy $e^{i\theta_1} = z$. Then the following assertions are equivalent*:

(i) *$\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ is disconnected.*

(ii) *$|\theta_1| < \pi$ and q lies in the region depicted in Figure 1.*

(iii) *$|\theta_1| < \pi$ and there exist condensed curves, but not critical curves, in $\mathcal{L}_{-1}^{+1}(Q)$.*

(iv) *$|\theta_1| < \pi$ and there exist condensed curves in $\mathcal{L}_{-1}^{+1}(Q)$, but no condensed curve is homotopic to a diffuse curve within $\mathcal{L}_{-1}^{+1}(Q)$.*

(v) *$|\theta_1| < \pi$ and there exists an embedding $\gamma \in \mathcal{L}_{-1}^{+1}(Q)$ which cannot be homotoped within this space to create self-intersections.*

(vi) *$|\theta_1| < \pi$ and there exists $\gamma \in \mathcal{L}_{-1}^{+1}(Q)$ which does not lie in the same component as $B_{\gamma,1,1/2}$.*

Furthermore, if $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ is disconnected, then it has exactly two components; one of them is $\mathcal{U}_c$ and the other is $\mathcal{U}_d \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$, and both are homeomorphic to $\boldsymbol{E}$, and hence contractible.

Proof. We know from (3.3) and (4.19) that each of $\mathcal{U}_c, \mathcal{U}_d \subset \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ is homeomorphic to $\boldsymbol{E}$, and hence connected. By (4.10), $\mathcal{U}_d \neq \varnothing$. By (5.1), $\overline{\mathcal{U}}_c \cup \overline{\mathcal{U}}_d = \mathcal{L}_{-1}^{+1}(Q; \theta_1)$, and $\overline{\mathcal{U}}_c \cap \overline{\mathcal{U}}_d$ consists of all the critical curves in $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$. Thus, the latter has at most two connected components. It has exactly two if and only if $\overline{\mathcal{U}}_c \neq \varnothing$ but $\overline{\mathcal{U}}_c \cap \overline{\mathcal{U}}_d = \varnothing$, that is, if and only if there exist condensed curves, but not critical curves. This proves the last assertion of the theorem and also the equivalence (i)$\Leftrightarrow$(iii). The equivalence (ii)$\Leftrightarrow$(iii) was proved in (5.5).

Suppose that $s \mapsto \gamma_s \in \mathcal{L}_{-1}^{+1}(Q; \theta_1)$ is a path joining a condensed curve to a diffuse curve. Let θ_s be the argument of $\boldsymbol{t}_{\gamma_s}$ satisfying $\theta_s(0) = 0$. By continuity, there must exist $s_0 \in (0,1)$ such that

$$\sup_{t\in[0,1]} \theta_{s_0}(t) - \inf_{t\in[0,1]} \theta_{s_0}(t) = \pi;$$

that is, there must exist s_0 such that γ_{s_0} is critical. Hence, (iii)⇒(iv).

Suppose that (iv) holds, and let $\gamma \in \mathcal{L}_{-1}^{+1}(Q)$ be smooth and condensed. Then γ is an embedding, but it cannot be deformed to have a self-intersection since any curve with double points must be diffuse. Thus, (iv)⇒(v).

Finally, it is obvious that (v)⇒(vi) and (vi)⇒(i). □

(6.2) Corollary. *Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and $\theta_1 \in \mathbb{R}$ satisfy $e^{i\theta_1} = z$. If $|\theta_1| \geq \pi$, then $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ is connected. If $|\theta_1| > \pi$, then $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ is homeomorphic to $\boldsymbol{E}$, and hence contractible.*

Proof. The first assertion is an immediate consequence of (6.1). If $|\theta_1| > \pi$ then $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ can only contain diffuse curves, and we know from (4.19) that $\mathcal{U}_d$ is homeomorphic to $\boldsymbol{E}$. □

Remark. The results of Section 4 go through to show that $\mathcal{L}_{-1}^{+1}(Q; \theta_1) = \mathcal{T} \cup \mathcal{U}_d$ is also contractible when $\theta_1 = \pm\pi$. Of course, if $|\theta_1| < \pi$ then $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ need not even be connected. We shall prove in the sequel [Saldanha and Zühlke 2015] that it may also be contractible, or connected but not contractible, depending on Q.

(6.3) Corollary. *Let $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$ and $\theta_1 \in \mathbb{R}$ satisfy $e^{i\theta_1} = z$. Then the subset $\mathcal{L}_{-1}^{+1}(Q; \theta_1)$ is either a connected component or the union of two contractible components of $\mathcal{L}_{-1}^{+1}(Q)$. The latter can occur only if $|\theta_1| < \pi$, that is, for at most one value of θ_1.* □

(6.4) Theorem. *Let $P = (p, w)$, $Q = (q, z) \in \mathbb{C} \times \mathbb{S}^1$, $\kappa_1 < 0 < \kappa_2$ and let θ_1 satisfy $e^{i\theta_1} = z\bar{w}$.*

(a) *If $|\theta_1| \geq \pi$, then the subspace $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ consisting of all curves having total turning θ_1 is a contractible connected component of $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$, homeomorphic to $\boldsymbol{E}$.*

(b) *If $|\theta_1| < \pi$, then $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ has at most two components. It is disconnected if and only if any of the conditions in (6.1) are satisfied for $\hat{Q} = (\hat{q}, z\bar{w})$, where*

$$\hat{q} = \frac{2}{\rho_2 - \rho_1}\bar{w}\Big((q-p) + \frac{i}{2}(\rho_1+\rho_2)(z-w)\Big) \quad \Big(\rho_i = \frac{1}{\kappa_i},\ i = 1, 2\Big).$$

In this case, one component consists of all condensed and the other of all diffuse curves in $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$, and both are homeomorphic to $\boldsymbol{E}$.

Proof. This is just a corollary of (2.4)(a), (6.1) and (6.2). □

We emphasize that the subspace of $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ which contains curves having least total turning, described in (b), does not have to be contractible even if it is connected. Observe also that we may replace $\mathcal{L}$ by $\mathcal{C}$ invoking (1.12).

7. Homeomorphism class of $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ for $\kappa_1\kappa_2 \geq 0$

An admissible plane curve γ is called locally convex if either $\kappa_\gamma > 0$ a.e. or $\kappa_\gamma < 0$ a.e. Notice that $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ consists of locally convex curves if and only if $\kappa_1\kappa_2 \geq 0$. This corresponds to parts (b)–(e) of (2.4). The topology of these spaces is very simple.

Suppose that $\gamma : [0, 1] \to \mathbb{C}$ is an admissible curve such that $\kappa_\gamma > 0$ a.e. and $\Phi_\gamma(0) = (0, 1)$. By (5), any argument $\theta : [0, 1] \to \mathbb{R}$ of $\boldsymbol{t}_\gamma$ must be strictly increasing; in particular, the total turning θ_1 of γ is positive. Thus, γ may be *parametrized by* its *argument* $\theta \in [0, \theta_1]$. By the chain rule,

$$\dot{\gamma}(\theta) = \rho(\theta)e^{i\theta} \quad (\theta \in [0, \theta_1]), \tag{43}$$

where $\rho : [0, \theta_1] \to (0, +\infty)$ is the radius of curvature of γ.[3]

(7.1) Theorem. *Let $P, Q \in UT\mathbb{C}$ and suppose that either $\kappa_1 \geq 0$ or $\kappa_2 \leq 0$. Then $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ has infinitely many connected components, one for each realizable total turning. All of these components are homeomorphic to $\boldsymbol{E}$, and hence contractible.*

Proof. Using an Euclidean motion if necessary, we may assume that $P = (0, 1)$. Further, by reversing the orientation of all curves, we pass from the case where $\kappa_2 \leq 0$ to the case where $\kappa_1 \geq 0$.

Let $Q = (q, z)$ and $e^{i\theta_1} = z$. The subspace $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q; \theta_1)$ is both open and closed in $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q)$ (but it may be empty; see (4.9). In particular, two curves which have different total turnings cannot lie in the same component of $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q)$. For any $k \in \mathbb{N}$, we may concatenate a curve in $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q)$ with a circle of curvature in (κ_1, κ_2) traversed k times. This shows that the number of components is infinite.

Suppose that $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q; \theta_1) \neq \varnothing$. Since $\kappa_1 \geq 0$ by hypothesis, we may reparametrize all curves in $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q; \theta_1)$ by the argument $\theta \in [0, \theta_1]$ of their unit tangent vectors using (1.15). Choose any $\gamma_0 \in \mathcal{L}_{\kappa_1}^{\kappa_2}(Q; \theta_1)$ and define a map H on $[0, 1] \times \mathcal{L}_{\kappa_1}^{\kappa_2}(Q; \theta_1)$ by

$$H(s, \gamma) = \gamma_s, \quad \gamma_s(\theta) = (1 - s)\gamma_0(\theta) + s\gamma(\theta) \quad (s \in [0, 1],\ \theta \in [0, \theta_1]).$$

Then $\gamma_s(0) = 0$, $\gamma_s(\theta_1) = q$ and the unit tangent vector $\boldsymbol{t}_{\gamma_s}$ to γ_s satisfies

$$\boldsymbol{t}_{\gamma_s}(\theta) = e^{i\theta} \quad \text{for all } \gamma \in \mathcal{L}_{\kappa_1}^{\kappa_2}(Q; \theta_1),\ s \in [0, 1] \text{ and } \theta \in [0, \theta_1].$$

Consequently, each γ_s has total turning θ_1, $\Phi_{\gamma_s}(0) = (0, 1)$ and $\Phi_{\gamma_s}(\theta_1) = Q$. Let $\rho_0, \rho : [0, 1] \to (0, +\infty)$ denote the radii of curvature of γ_0, γ, respectively. It follows from (43) that the radius of curvature ρ_s of γ_s is given by

$$\rho_s = (1 - s)\rho_0 + s\rho.$$

[3] The idea of parametrizing a locally convex curve by the argument of its unit tangent vector is not new. It appears in [Little 1970], where it is attributed to W. Pohl. We do not know whether it is older than that.

Therefore, the curvature $\kappa_s = 1/\rho_s$ of γ_s takes values in (κ_1, κ_2) and H is a contraction of $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q; \theta_1)$. We conclude that the latter is a connected component of $\mathcal{L}_{\kappa_1}^{\kappa_2}(Q)$ and, using (1.7)(b), that it is homeomorphic to $\boldsymbol{E}$. □

Possible total turnings of a curve in $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ when $\kappa_1\kappa_2 > 0$. Let T denote the set of all total turnings which are realized by some curve in $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$. If $P = (p, w)$, $Q = (q, z)$, then obviously

$$T \subset \{\theta_1 + 2k\pi : k \in \mathbb{Z}\}, \quad \text{where } e^{i\theta_1} = z\bar{w}.$$

If $\kappa_1\kappa_2 < 0$, this inclusion is an equality by (4.9)(b). However, it must be proper when $\kappa_1\kappa_2 \geq 0$. If κ_1, κ_2 are both positive, for instance, then, by (5) and the second paragraph of the above proof, T must have the form $\{\mu + 2k\pi : k \in \mathbb{N}\}$, where $\mu \in \mathbb{R}$ ($e^{i\mu} = z\bar{w}$) is the minimal attainable total turning in this space. It is possible to find the value of μ in terms of all parameters involved. Because this determination is of lesser interest and relatively technical, we shall not go into it here. However, interested readers can find the details, including the analogue for spaces of the form $\hat{\mathcal{L}}$, in [Saldanha and Zühlke 2014]. We mention only that (2.4) allows one to restrict attention to the two classes $\mathcal{L}_1^{+\infty}(Q)$ and $\mathcal{L}_0^{+\infty}(Q)$.

8. Components of spaces of curves on complete flat surfaces

By a *flat surface* we mean a connected Riemannian 2-manifold whose Gaussian curvature is identically zero; it will not be necessary to assume that S is a submanifold of some Euclidean space. The *unit tangent* $\boldsymbol{t} = \boldsymbol{t}_\gamma : [0, 1] \to UTS$ to a regular curve $\gamma : [0, 1] \to S$ is defined as before, $\boldsymbol{t} = \dot{\gamma}/|\dot{\gamma}|$. If S is orientable, the *unit normal* $\boldsymbol{n} = \boldsymbol{n}_\gamma : [0, 1] \to UTS$ to γ is defined by the condition that $(\boldsymbol{t}(t), \boldsymbol{n}(t))$ should be a positively oriented orthonormal basis of $TS_{\gamma(t)}$ for each $t \in [0, 1]$. For γ of class C^2, we can then define its *curvature* $\kappa_\gamma : [0, 1] \to \mathbb{R}$ by

$$\kappa_\gamma = \frac{1}{|\dot{\gamma}|}\left\langle \frac{D\boldsymbol{t}}{dt}, \boldsymbol{n} \right\rangle,$$

where D denotes covariant differentiation (along γ).

If S is nonorientable, we can still speak of the *unsigned curvature* $\kappa_\gamma : [0, 1] \to [0, +\infty)$ of a curve $\gamma : [0, 1] \to S$, given by

$$\kappa_\gamma = \frac{1}{|\dot{\gamma}|}\left|\left\langle \frac{D\boldsymbol{t}}{dt}, \boldsymbol{n} \right\rangle\right|,$$

where now $\boldsymbol{n}(t)$ denotes any of the two unit vectors in $TS_{\gamma(t)}$ orthogonal to $\boldsymbol{t}(t)$.

(8.1) Definition. Let S be an orientable flat surface, $u, v \in UTS$, $-\infty \leq \kappa_1 < \kappa_2 \leq +\infty$, and $2 \leq r \in \mathbb{N}$. Define $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ to be the set of all C^r regular curves $\gamma : [0, 1] \to S$ satisfying

(i) $\boldsymbol{t}_\gamma(0) = u$ and $\boldsymbol{t}_\gamma(1) = v$;

(ii) $\kappa_1 < \kappa_\gamma(t) < \kappa_2$ for each $t \in [0, 1]$.

In case S is nonorientable, define $\mathcal{C}S_{-\kappa_0}^{+\kappa_0}(u, v)$ ($\kappa_0 > 0$) as above, but replacing condition (ii) by

(ii′) $\kappa_\gamma(t) < \kappa_0$ for each $t \in [0, 1]$.

In both cases, let $\mathcal{C}_{\kappa_1}^{\kappa_2}(u, v)$ be furnished with the C^r topology.

Remark. A complete flat surface must be homeomorphic to one of the following five: $\mathbb{C}$ itself, a cylinder $\mathbb{S}^1 \times \mathbb{R}$, an open Möbius band, a torus or a Klein bottle. This is essentially a corollary of the following result; cf. [Hopf 1926, p. 319].

(8.2) Theorem (Killing–Hopf). *Any complete flat surface is isometric to the quotient of the Euclidean plane $\mathbb{C}$ by a group of isometries acting freely and properly discontinuously on $\mathbb{C}$.* □

Hence, if S is a complete flat surface, there exists a covering map $\mathbb{C} \to S$ which is a local isometry. Any curve on S may thus be lifted to a plane curve whose curvature is the same as that of the original curve, with the proviso that we ignore its sign if S is nonorientable. Let $\mathrm{pr} : UT\mathbb{C} \to UTS$ denote the natural projection induced by the covering map. In what follows, when referring to $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$, we adopt the convention that $\kappa_1 = -\kappa_2 < 0$ if S is nonorientable.

(8.3) Proposition. *Let S be a complete flat surface, $u, v \in UTS$, $-\infty \leq \kappa_1 < \kappa_2 \leq +\infty$ and $P \in UT\mathbb{C}$ be a fixed element of* $\mathrm{pr}^{-1}(u)$. *Then $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ is homeomorphic to $\coprod_{Q \in \mathrm{pr}^{-1}(v)} \mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$, where the homeomorphism maps a curve in the latter to its image under the quotient map $\mathbb{C} \to S$.* □

Here $\coprod$ denotes topological sum. Clearly, this decomposition is sufficient to determine the connected components of $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ explicitly, using (1.12) and (6.4) if $\kappa_1\kappa_2 < 0$ or (7.1) if $\kappa_1\kappa_2 \geq 0$.

(8.4) Corollary. *Let S be a complete flat surface, $\kappa_1 < \kappa_2$ and $u, v \in UTS$. Then $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ is nonempty and has an infinite number of connected components.*

Proof. By (4.9) and the remark which follows it, $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$ is always nonempty. The assertion is thus an immediate consequence of (8.3). □

Notice that it is irrelevant here whether S is compact. This should be compared to the case of $S = \mathbb{S}^2$, where, at least when $u = v$, the number of components of $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ is finite for any choice of $\kappa_1 < \kappa_2$ (see [Saldanha and Zühlke 2013, §7]). This is actually not surprising, since the fundamental group of $UT\mathbb{C}$ is isomorphic to $\mathbb{Z}$, but that of $UT\mathbb{S}^2 \approx \boldsymbol{SO}_3$ is isomorphic to $\mathbb{Z}_2$. We remark without proof that $\mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ may be empty for more general surfaces (for instance, $\mathcal{C}S_{-1}^{+1}(u, u) = \varnothing$ when S is the hyperbolic plane $\boldsymbol{H}^2$ for any $u \in UT\boldsymbol{H}^2$).

(8.5) Corollary. *Let S be a complete flat surface and $u, v \in UTS$. Let $\eta \in \mathcal{C}S_{-\infty}^{+\infty}(u, v)$ and suppose that $\kappa_1\kappa_2 < 0$. Then there exists $\gamma \in \mathcal{C}S_{\kappa_1}^{\kappa_2}(u, v)$ lying in the same component of $\mathcal{C}S_{-\infty}^{+\infty}(u, v)$ as η.*

In other words, given a regular curve η on S with $\boldsymbol{t}_\eta(0) = u$, $\boldsymbol{t}_\eta(1) = v$, we may deform η through regular curves, keeping $\boldsymbol{t}(0)$, $\boldsymbol{t}(1)$ fixed, to obtain a curve having curvature in (κ_1, κ_2) everywhere.

Proof. Take $P \in UT\mathbb{C}$ such that $\mathrm{pr}(P) = u$. Let $\tilde{\eta}$ be the lift of η to $\mathbb{C}$ with initial frame P; let Q be its final frame and θ_1 its total turning. By (4.9)(b), $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q; \theta_1)$ is nonempty. Let $\tilde{\gamma}$ be one of its elements. Then the projection γ of $\tilde{\gamma}$ on S satisfies the conclusion of the corollary because of (8.3) and the fact that $\tilde{\eta}$, $\tilde{\gamma}$ lie in the same component of $\mathcal{C}_{-\infty}^{+\infty}(P, Q)$. $\square$

Again, the analogue of this result does not hold for a general surface S, e.g., for $S = \mathbf{H}^2$. It is also false for a flat surface if $\kappa_1\kappa_2 \geq 0$. To see this, let $P, Q \in UT\mathbb{C}$ satisfy $\mathrm{pr}(P) = u$, $\mathrm{pr}(Q) = v$, choose $\tilde{\eta} \in \mathcal{C}_{-\infty}^{+\infty}(P, Q)$ to have a total turning which is unattainable for curves in $\mathcal{C}_{\kappa_1}^{\kappa_2}(P, Q)$ and let η be the projection of $\tilde{\eta}$ on S.

9. Final remarks

Spaces of curves with curvature in a closed interval. Dubins [1957; 1961] worked with the set $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(Q)$ of (1.16) (but with the C^1 topology), where the curvatures are restricted to lie in a *closed* interval. This choice is motivated by the fact that these spaces, unlike those of the form $\mathcal{L}_{-\kappa_0}^{+\kappa_0}(Q)$, always contain curves of minimal length (see [Dubins 1957, Proposition 1]). All of the main results in our paper concerning the topology of $\mathcal{L}_{\kappa_1}^{\kappa_2}(P, Q)$ have analogues for $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$. We shall now briefly indicate the modifications which are necessary.

Notice that $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$ is *not* a Banach manifold, and that the analogue of (1.5) is false for these spaces, as shown by (1.1). In contrast, (1.14) and (1.15) still hold when $\mathcal{M} = \hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$. The important corollary (2.6) has the following analogue, whose proof is essentially the same as that of (2.4); see (2.7).

(9.1) Proposition. *Let $P, Q \in UT\mathbb{C}$ and $\kappa_1 < \kappa_2$ be finite. Then there exists a homeomorphism between $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$ and a space of type $\hat{\mathcal{L}}_{-1}^{+1}(Q_0)$, $\hat{\mathcal{L}}_{0}^{1}(Q_0)$ or $\hat{\mathcal{L}}_{1}^{2}(Q_0)$, according to whether $\kappa_1\kappa_2 < 0$, $\kappa_1\kappa_2 = 0$ or $\kappa_1\kappa_2 > 0$, respectively. Moreover, this homeomorphism preserves the total turning of curves unless $\kappa_1 < \kappa_2 \leq 0$, in which case it reverses the sign.* $\square$

In case $\kappa_1\kappa_2 < 0$, we actually have $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q) \approx \hat{\mathcal{L}}_{-1}^{+1}(Q_1)$ with Q_1 as in the statement of (2.4). We leave the task of determining Q_0 in the other two cases to the interested reader.

Let us denote by $\hat{\mathcal{U}}_c$, $\hat{\mathcal{U}}_d$ and $\hat{\mathcal{T}} \subset \hat{\mathcal{L}}_{-1}^{+1}(Q; \theta_1)$ the subspaces consisting of all condensed, diffuse and critical curves, where $Q = (q, z)$, $e^{i\theta_1} = z$ and $\hat{\mathcal{L}}_{-1}^{+1}(Q; \theta_1)$

consists of those curves in $\hat{\mathcal{L}}_{-1}^{+1}(Q)$ which have total turning θ_1. The analogue of (3.3), stating that $\hat{\mathfrak{U}}_c$ is either empty or contractible, is, naturally, (3.4), which was used to prove (3.3). The results and proofs in Section 4 all need minimal or no modifications. In particular, $\hat{\mathfrak{U}}_d$ is always nonempty and weakly contractible.

The proof that $\partial\hat{\mathfrak{U}}_d = \hat{\mathfrak{T}}$ is the same as the one given in (5.1). The proof that $\partial\hat{\mathfrak{U}}_c = \hat{\mathfrak{T}}$, however, needs to be modified, since we have relied on (1.5). The idea is again to apply construction (3.8), but to all of γ, not just to some of its arcs as in the proof of (5.1). If γ is a critical curve, then the corresponding function f (see Figure 4) will attain the values $\pm\infty$, and at these points we need to assign weights, corresponding to the lengths of the line segments where θ_γ attains its maximum and minimum. Then we redefine $A(\mu_-, \mu_+)$ as the sum of the area under the graph of $f^{(\mu_-,\mu_+)}$ plus the weight at $+\infty$ minus the weight at $-\infty$. The process described in (3.8) will transform f into a bounded function of the same area; that is, it will decrease the amplitude of γ, making it a condensed curve.

The proofs of (3.17), (5.3) and (5.4), which deal with the existence of condensed and critical curves, go through unchanged; the only difference in the conclusions is that the corresponding regions $R_{\hat{\mathfrak{U}}_c}$, R_σ and $R_{\hat{\mathfrak{T}}}$ of the plane are now closed, instead of open. Thus, in the analogue of (6.1), the region of Figure 1 should contain the two circles of radius 2, but not the circle of radius 4, and we cannot assert that $\hat{\mathfrak{U}}_c$ and $\hat{\mathfrak{U}}_d$ are homeomorphic to $\boldsymbol{E}$, only that they are weakly contractible. The rest of the statement and the proof hold without modifications.

Similarly, the version of (7.1) for $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$ states that this space has one contractible connected component for each realizable total turning when $\kappa_1\kappa_2 \geq 0$. The proof is the same as that of (7.1) if $\kappa_1\kappa_2 > 0$. If $\kappa_1 = 0$, then we cannot really parametrize $\gamma \in \hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$ by argument. Nevertheless, the proof still works if we replace $\rho(\theta)d\theta$ by a measure $\mu(\theta)$ on the Borel subsets of $[0, \theta_1]$ which has an atom at θ if the curvature of γ vanishes at $\gamma(\theta)$; note that the convex combination of two measures is again a measure. The case where $\kappa_2 = 0$ can be deduced from this one by reversing orientations.

A few conjectures of Dubins. All of the results in the next proposition were conjectured by Dubins [1961, §6].

(9.2) Proposition. *Let $q \in \mathbb{C}$, $\theta_1 \in \mathbb{R}$, $z = e^{i\theta_1}$ and $Q = (q, z)$. Then*:

(a) *The set of all $(q, \theta_1) \in \mathbb{C} \times \mathbb{R}$ such that $\hat{\mathcal{L}}_{-1}^{+1}(Q; \theta_1)$ is disconnected is a bounded subset of $\mathbb{C} \times \mathbb{R}$, neither open nor closed.*[4]

(b) *$\hat{\mathcal{L}}_{-1}^{+1}(Q; \theta_1)$ has at most two components.*

[4] Actually, Dubins had guessed that this set would be bounded and open in $\mathbb{C} \times \mathbb{R}$.

(c) *If $\hat{\mathcal{L}}_{-1}^{+1}(Q;\theta_1)$ is disconnected, then one component ($\hat{\mathcal{U}}_d$) contains curves of arbitrarily large length, while the supremum of the lengths of curves in the other component ($\hat{\mathcal{U}}_c$) is finite.*

(d) *Every point of $\mathbb{C}$ lies in the image of some $\gamma \in \hat{\mathcal{U}}_d$, while the images of curves in $\hat{\mathcal{U}}_c$ are contained in a bounded subset of $\mathbb{C}$.*

Proof. Parts (a) and (b) are immediate from the analogue of (6.1) for $\hat{\mathcal{L}}$. As discussed above, $\hat{\mathcal{L}}_{-1}^{+1}(Q;\theta_1)$ is disconnected if and only if $|\theta_1| < \pi$ and q lies in the region in Figure 1 including the circles of radius 2 but not the circle of radius 4. Suppose that q does lie in this region. Choose $\hat{\omega} \in (\theta_1, \pi)$ such that

$$|q - \operatorname{sign}(\theta_1) i (z-1)| < 4\sin\Big(\frac{\hat{\omega}}{2}\Big).$$

Then (5.6) implies that there does not exist any curve in $\hat{\mathcal{L}}_{-1}^{+1}(Q;\theta_1)$ having amplitude in $[\hat{\omega}, \pi]$. The assertions about $\hat{\mathcal{U}}_c$ in (c) and (d) now follow from (3.15). The assertions about $\hat{\mathcal{U}}_d$ are obvious, because, by (the version for $\hat{\mathcal{L}}$ of) (4.18), this subspace always contains curves of amplitude $\geq 2\pi$, and onto such a curve we may graft line segments of any direction and arbitrary length. $\square$

As expected, there is a version of the foregoing proposition for $\mathcal{L}_{-1}^{+1}(Q;\theta_1)$. The corresponding assertions in (a) and (b) are immediate from (6.1), and the assertions about $\mathcal{U}_d$ are again obvious. A curve in $\mathcal{U}_c$ parametrized by arc-length can also be considered as an element of $\hat{\mathcal{U}}_c$, so the properties stated in (c) and (d) for $\mathcal{U}_c$ follow from those for $\hat{\mathcal{U}}_c$ unless q lies on the circle of radius 4 in Figure 1. In this case, $\mathcal{L}_{-1}^{+1}(Q;\theta_1)$ is disconnected, but $\hat{\mathcal{L}}_{-1}^{+1}(Q;\theta_1)$ is not. One can prove directly that the length of any $\gamma \in \mathcal{U}_c$ must be smaller than that of the "canonical" critical curves of type $+-$ or $-+$ that were constructed in the proof of (5.3).

Conjectures on minimal length. Let $L(\gamma)$ denote the length of γ and suppose that $\hat{\mathcal{L}}_{-1}^{+1}(Q;\theta_1)$ is disconnected. We believe that the results developed here may be used to prove that if $m = \sup_{\gamma\in\hat{\mathcal{U}}_c} L(\gamma)$ and $M = \inf_{\gamma\in\hat{\mathcal{U}}_d} L(\gamma)$, then $m < M$; this is another conjecture of Dubins. It would be interesting, and probably useful for applications, to find the values corresponding to m and M for the more general spaces $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(Q)$.

We observed in (2.5) that normal translations, and hence the homeomorphisms of (9.1) need not preserve inequalities between lengths. Since they do map circles to circles and lines to lines, it could still be expected that the image of a curve which minimizes length under these homeomorphisms is likewise of minimal length. Unfortunately, this is false. Suppose, for instance, that we apply the homeomorphism $\hat{\mathcal{L}}_{-1}^{+1}(Q) \to \hat{\mathcal{L}}_{-1}^{100}(Q_0)$ to the Dubins path in Figure 3(b). It should be clear that its image, which again consists of a line segment and two arcs of circles of opposite orientation with the same amplitude as before, does not minimize length in $\hat{\mathcal{L}}_{-1}^{100}(Q_0)$, since in the latter space it is generally much more efficient to curve to the left than to the right, even if this yields a path of greater total turning.

In spite of this difficulty, we conjecture that Dubins' theorem that any shortest path in $\hat{\mathcal{L}}_{-\kappa_0}^{+\kappa_0}(P, Q)$ must be a concatenation of three pieces, each of which is either an arc of circle or a line segment, still holds for the spaces $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(P, Q)$, $\kappa_1\kappa_2 < 0$. For $\kappa_1\kappa_2 > 0$, we conjecture that a curve of minimal length is a concatenation of n arcs of circles of curvature κ_1 and κ_2. However, for fixed $P, Q \in UT\mathbb{C}$, we must have $\lim_{\kappa_1,\kappa_2\to+\infty} n = \infty$. Indeed, a curve of this type has total turning at most $2n\pi$, and the minimal total turning of a curve in $\hat{\mathcal{L}}_{\kappa_1}^{\kappa_2}(Q)$ increases to infinity as $\kappa_1 > 0$ increases to infinity (for fixed $Q = (q, z)$ with $q \neq 0$).

Acknowledgements. Saldanha is partially supported by grants from CNPq, CAPES and FAPERJ (Brazil). Zühlke gratefully acknowledges the support of PUC-Rio (where this work was written while he was a post-doctoral fellow), and the financial support of CNPq and FAPESP. Both authors would like to thank the anonymous referee, whose comments have helped to make the paper clearer.

References

[Anisov 1998] S. S. Anisov, "Выпуклые кривые в $\mathbb{RP}^n$", *Tr. Mat. Inst. Steklova* **221** (1998), 9–47. Translated as "Convex curves in $\mathbb{RP}^n$" in *Proc. Steklov. Inst. Math.* **221** (1998), 3–39. MR 2000e:53014 Zbl 1017.52001

[Burghelea and Kuiper 1969] D. Burghelea and N. H. Kuiper, "Hilbert manifolds", *Ann. of Math.* (2) **90** (1969), 379–417. MR 40 #6589 Zbl 0195.53501

[Burghelea et al. 2003] D. Burghelea, N. C. Saldanha, and C. Tomei, "Results on infinite-dimensional topology and applications to the structure of the critical set of nonlinear Sturm–Liouville operators", *J. Differential Equations* **188**:2 (2003), 569–590. MR 2003m:47121 Zbl 1025.34023

[Dubins 1957] L. E. Dubins, "On curves of minimal length with a constraint on average curvature, and with prescribed initial and terminal positions and tangents", *Amer. J. Math.* **79**:3 (1957), 497–516. MR 19,678c Zbl 0098.35401

[Dubins 1961] L. E. Dubins, "On plane curves with curvature", *Pacific J. Math.* **11** (1961), 471–481. MR 24 #A2301 Zbl 0123.15503

[Feldman 1968] E. A. Feldman, "Deformations of closed space curves", *J. Differential Geometry* **2** (1968), 67–75. MR 38 #725 Zbl 0195.53401

[Feldman 1971] E. A. Feldman, "Nondegenerate curves on a Riemannian manifold", *J. Differential Geometry* **5** (1971), 187–210. MR 45 #1074 Zbl 0218.53010

[Henderson 1969] D. W. Henderson, "Infinite-dimensional manifolds are open subsets of Hilbert space", *Bull. Amer. Math. Soc.* **75**:4 (1969), 759–762. MR 40 #898 Zbl 0179.29101

[Hopf 1926] H. Hopf, "Zum Clifford–Kleinschen Raumproblem", *Math. Ann.* **95**:1 (1926), 313–339. MR 1512281 JFM 51.0439.05

[Khesin and Shapiro 1992] B. A. Khesin and B. Z. Shapiro, "Nondegenerate curves on S^2 and orbit classification of the Zamolodchikov algebra", *Comm. Math. Phys.* **145**:2 (1992), 357–362. MR 93f:58100 Zbl 0849.17026

[Khesin and Shapiro 1999] B. A. Khesin and B. Z. Shapiro, "Homotopy classification of nondegenerate quasiperiodic curves on the 2-sphere", *Publ. Inst. Math.* (*Beograd*) (*N.S.*) **66**:80 (1999), 127–156. MR 2001i:37116 Zbl 1274.53055

[Little 1970] J. A. Little, "Nondegenerate homotopies of curves on the unit 2-sphere", *J. Differential Geometry* **4** (1970), 339–348. MR 43 #1090 Zbl 0198.53603

[Mostovoy and Sadykov 2012] J. Mostovoy and R. Sadykov, "The space of non-degenerate closed curves in a Riemannian manifold", preprint, 2012. arXiv 1209.4109

[Palais 1966] R. S. Palais, "Homotopy theory of infinite dimensional manifolds", *Topology* **5** (1966), 1–16. MR 32 #6455 Zbl 0138.18302

[Saldanha 2015] N. C. Saldanha, "The homotopy type of spaces of locally convex curves in the sphere", *Geom. Topol.* **19**:3 (2015), 1155–1203. MR 3352233 Zbl 1318.53066

[Saldanha and Shapiro 2012] N. C. Saldanha and B. Z. Shapiro, "Spaces of locally convex curves in $\mathbb{S}^n$ and combinatorics of the group B^+_{n+1}", *J. Singul.* **4** (2012), 1–22. MR 2872212 Zbl 1292.58002

[Saldanha and Zühlke 2013] N. C. Saldanha and P. Zühlke, "On the components of spaces of curves on the 2-sphere with geodesic curvature in a prescribed interval", *Internat. J. Math.* **24**:14 (2013), Article ID #1350101. MR 3163615 Zbl 1288.53054

[Saldanha and Zühlke 2014] N. C. Saldanha and P. Zühlke, "Spaces of curves with constrained curvature on flat surfaces, I", preprint, 2014. arXiv 1312.1675v3

[Saldanha and Zühlke 2015] N. C. Saldanha and P. Zühlke, "Homotopy type of spaces of curves with constrained curvature on flat surfaces", preprint, 2015. arXiv 1410.8590

[Shapiro 1993] M. Z. Shapiro, "Топология пространства невырожденных кривых", *Izv. Ross. Akad. Nauk Ser. Mat.* **57**:5 (1993), 106–126. Translated as "Topology of the space of nondegenerate curves" in *Izv. Math.* **43**:2 (1994), 291–310. MR 94k:58019 Zbl 0821.57021

[Shapiro and Shapiro 1991] B. Z. Shapiro and M. Z. Shapiro, "On the number of connected components in the space of closed nondegenerate curves on S^n", *Bull. Amer. Math. Soc.* (*N.S.*) **25**:1 (1991), 75–79. MR 91j:58028 Zbl 0731.53003

[Smale 1958] S. Smale, "Regular curves on Riemannian manifolds", *Trans. Amer. Math. Soc.* **87** (1958), 492–512. MR 20 #1319 Zbl 0081.38103

[Whitney 1937] H. Whitney, "On regular closed curves in the plane", *Compositio Math.* **4** (1937), 276–284. MR 1556973 Zbl 0016.13804

[Younes 2010] L. Younes, *Shapes and diffeomorphisms*, Applied Mathematical Sciences **171**, Springer, Berlin, 2010. MR 2011h:53002 Zbl 1205.68355

Received November 5, 2014. Revised July 16, 2015.

NICOLAU C. SALDANHA
DEPARTAMENTO DE MATEMÁTICA
PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO
RUA MARQUÊS DE SAO VICENTE 225
RIO DE JANEIRO, RJ 22453-900
BRAZIL

nicolau@mat.puc-rio.br

PEDRO ZÜHLKE
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE DE SÃO PAULO
RUA DO MATÃO 1010
IME - CIDADE UNIVERSITÁRIA
SÃO PAULO, SP 05508-090
BRAZIL

pedroz@ime.usp.br

PACIFIC JOURNAL OF MATHEMATICS
Vol. 281, No. 1, 2016
dx.doi.org/10.2140/pjm.2016.281.243

A NOTE ON MINIMAL GRAPHS OVER CERTAIN UNBOUNDED DOMAINS OF HADAMARD MANIFOLDS

MIRIAM TELICHEVESKY

Given an unbounded domain Ω of a Hadamard manifold M, it makes sense to consider the problem of finding minimal graphs with prescribed continuous data on its cone topology boundary, i.e., on its ordinary boundary together with its asymptotic boundary. In this article it is proved that under the hypothesis that the sectional curvature of M is ≤ -1, this Dirichlet problem is solvable if Ω satisfies a certain convexity condition at infinity and if $\partial\Omega$ is mean convex. We also prove that mean convexity of $\partial\Omega$ is a necessary condition, extending to unbounded domains some results that are valid on bounded ones.

1. Introduction

The classical theorem of Jenkins and Serrin on minimal graphs theory, following the works of Bernstein [1910], Haar [1927], Radó [1930] and Finn [1965], states the following.

Theorem 1 [Jenkins and Serrin 1968, Theorem 1]. *Let $D \subset \mathbb{R}^n$ be a bounded domain whose boundary is of class C^2. Then the Dirichlet problem for the minimal surface equation in D is well posed for C^2 boundary data if and only if the mean curvature of ∂D is everywhere nonnegative.*

In the last four decades, several works considered problems related to Theorem 1 in distinct directions. Some of them are listed below together with some references.

- Unbounded domains of $\mathbb{R}^2$: [Hwang 1988; Collin and Krust 1991; Sá Earp and Rosenberg 1989; Ripoll and Tomi 2014; Krust 1989; Kuwert 1993; Kutev and Tomi 1998].
- Bounded domains of a Hadamard manifold M: [Folha and Rosenberg 2012; Mazet et al. 2011; Aiolfi et al. 2016].

MSC2010: primary 35-XX, 58-XX; secondary 58J05, 35J93.
Keywords: Dirichlet problem for minimal graphs, Hadamard manifolds.

- Asymptotic Dirichlet problems on Hadamard manifolds: [do Espírito Santo et al. 2010; Ripoll and Telichevesky 2015; Gálvez and Rosenberg 2010; Castéras et al. 2013].
- Replace the ambient space $\mathbb{R}^{n+1}$ by the hyperbolic spaces $\mathbb{H}^{n+1}$ [Barbosa and Sá Earp 1998; Guio and Sá Earp 2005; López 2001; Nitsche 2002] or other ambient spaces with a Killing field satisfying certain hypotheses [Alías and Dajczer 2007; Dajczer et al. 2008; 2013]. In this setting it is natural to consider CMC Killing graphs and there is an extensive bibliography on it.

The purpose of this article is to prove that similar existence and nonexistence results remain valid if in Theorem 1, $\mathbb{R}^n$ is replaced by a Hadamard manifold M with sectional curvature $K_M \leq -1$ and the domain D is unbounded and "strictly convex at infinity" (see Definition 4).

Classically, Dirichlet problems on unbounded domains are considered in $\mathbb{R}^n$ without prescribed values at infinity. In fact, sometimes the behavior at infinity of bounded solutions is determined by their boundary values. For instance, in $\mathbb{R}^2$ it is a consequence of Theorem 2 of [Collin and Krust 1991], which states that if u and v are distinct solutions of the Dirichlet problem in an unbounded domain $U \subset \mathbb{R}^2$ which coincide on ∂U, then $\sup |u - v|$ must have at least logarithmic growth. However, since the manifolds that we consider in this work have sectional curvature $K_M \leq -1$, it turns out that the asymptotic boundary of unbounded domains may be "good enough" to prescribe continuous data on them. It therefore makes sense to consider the generalized Dirichlet problem for the minimal hypersurface equation, Problem 2, described in the sequel. In order to state it, let us introduce some useful notations that are not standard.

Throughout this paper M denotes an m-dimensional Hadamard manifold, $m \geq 2$, with sectional curvature K_M satisfying $K_M \leq -1$. The asymptotic boundary $\partial_\infty M$ of M is defined by the set of equivalence classes of geodesic rays that stay at finite distance for all time, and it is possible to compactify M by adding $\partial_\infty M$ to it. $\overline{M} := M \cup \partial_\infty M$ carries the so-called cone topology (see [Eberlein and O'Neill 1973]), which makes it canonically homeomorphic to a closed ball. If $U \subset \overline{M}$ is any set, we denote by $\overline{U}^{\text{ct}} \subset \overline{M}$ and $\partial^{\text{ct}} U \subset \overline{M}$ its closure and boundary in terms of the cone topology; we also use the notation $\partial_\infty U := \partial^{\text{ct}} U \cap \partial_\infty M$.

Problem 2. Let $\Omega \subset M$ be a C^2 domain of M. Given $\varphi \in C(\partial^{\text{ct}}\Omega)$, find a minimal graph over Ω that attains φ on its boundary, or, equivalently, find a solution of the Dirichlet problem

$$\begin{cases} u \in C^2(\Omega) \cap C\overline{\Omega}^{\text{ct}}, \\ \mathcal{M}(u) := \operatorname{div}\left(\dfrac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right) = 0 \quad \text{in } \Omega, \\ u|_{\partial^{\text{ct}}\Omega} = \varphi. \end{cases}$$

Concerning the existence, perhaps the main difficulty dealing with unbounded domains is the nonexistence of natural barriers. In general, barriers are constructed using distance functions to a point or to the boundary of the domain, which cannot be adapted directly to points at infinity. Here the geometry of M at infinity plays an important role. For instance, the hyperbolic spaces $\mathbb{H}^n$ have "good geometry" at infinity by the existence of hyperspheres separating points at infinity and having their principal curvatures with the correct sign. The natural way to generalize this fact in order to use the Hessian comparison theorem and adapt barriers of $\mathbb{H}^n$ to other Hadamard manifolds is given by the *strict convexity condition* (*SC condition*) *at infinity*, introduced in [Ripoll and Telichevesky 2015]. In that work it is proved that Problem 2 is solvable for $\Omega = M$ (in this case, it is called the asymptotic Dirichlet problem) and any continuous boundary data if M satisfies the SC condition described below.

Definition 3. A Hadamard manifold M is said to satisfy the *strict convexity condition at infinity* if for all $x \in \partial_\infty M$ and all relatively open subsets $\Gamma \subset \partial_\infty M$ with $x \in \Gamma$, there exists an open set $V \subset M$ of class C^2 such that x is an interior point of $\partial_\infty V$ (with respect to the induced topology), $\partial_\infty V \subset \Gamma$ and $M \setminus V$ is convex.

At this point, it should be mentioned that under the hypothesis $K_M \leq -1$, the SC condition is always satisfied by 2-dimensional manifolds, by the rotationally symmetric ones and by those manifolds with controlled decay on sectional curvature (exponential decay) (see also [Ripoll and Telichevesky 2015]). We also should mention that under the same assumption on K_M, the SC condition is equivalent to the *convex conic neighborhood condition* presented by H. Choi [1984] in the study of the asymptotic Dirichlet problem with respect to Laplace's operator on Cartan–Hadamard manifolds; the equivalence is a consequence of a lemma of A. Borbély [1998b, Lemma 1]. In fact, both Dirichlet problems are closely related and may be studied together (see also [Ripoll and Telichevesky 2015]).

Contrasting with the existence results under the SC condition, we cite a counterexample constructed by I. Holopainen and J. Ripoll [2015]. In this work the authors present a Hadamard manifold with $K_M \leq -1$ that does not admit a solution to the asymptotic Dirichlet problem for the minimal hypersurface equation for any continuous $\varphi \in C(\partial_\infty M)$, although there are bounded nonconstant minimal graphs globally defined on M (see Theorem 1.1 of [Holopainen and Ripoll 2015]). This counterexample proves that the condition $K_M \leq -1$ is not sufficient to solve Problem 2 with any continuous boundary data.

Taking into account all these facts, the following definition is natural.

Definition 4. A domain $\Omega \subset \overline{M}$ is *strictly convex at infinity* if for any $x \in \partial_\infty \Omega$ and any relatively open neighborhood $\Gamma \subset \partial^{\mathrm{ct}}\Omega$ of x, there exists an open neighborhood $V = V(x, \Gamma, \Omega) \subset \Omega$ of x such that $\overline{V} \cap \partial^{\mathrm{ct}}\Omega \subset \Gamma$ and all the principal curvatures of $\partial V \cap \Omega$, oriented in the direction of $\Omega \setminus V$, are nonnegative.

Notice that when $\Omega = M$, this definition coincides with the SC condition. With Definition 4 it is now possible to state our main existence result.

Theorem 5. *Let $\Omega \subset M$ be a mean convex domain (with respect to the inward orientation) that is strictly convex at infinity. Then Problem 2 is solvable for any continuous boundary data.*

Returning our attention to Theorem 1, when $\Omega \subset \mathbb{R}^n$ is bounded, the mean convexity is a necessary condition to the solvability of Problem 2 for any continuous φ. The second part of this article is dedicated to proving that mean convexity is also necessary in M if we deal with unbounded domains and require boundedness of solutions. In Section 3 we present some necessary lemmata and the proof of the following nonexistence result.

Theorem 6. *Let $\Omega \subset \overline{M}$ be a domain and suppose that there exists $y \in \partial\Omega$ such that the mean curvature of $\partial\Omega$ at y (with respect to the inward orientation) satisfies $H(y) < 0$. Then there exists a continuous function $\varphi : \partial^{\mathrm{ct}}\Omega \to \mathbb{R}$ such that Problem 2 is not solvable.*

The construction of φ depends on two basic ingredients. First of all, on the local aspect concerning the negativity of the mean curvature $H(y)$, it is essential to guarantee that $\varphi(y)$ is bounded by values of the solution on a small sphere centered at y, say, on $S_r(y) \cap \Omega$. The second essential ingredient is the existence of a bounded barrier in $\Omega \setminus B_r(y)$ with some special properties. Similar results outside $\mathbb{R}^n$ were proved on bounded domains considering barriers dependent on the diameter of Ω, as in [Nitsche 2002]. Our main improvement is dropping the dependence on the size of the domain.

Combining the results of Theorems 5 and 6, we get:

Theorem 7. *Let $\Omega \subset M$ be a domain that is strictly convex at infinity. Then the Dirichlet problem (Problem 2) is solvable for any continuous boundary data if and only if Ω is mean convex.*

It remains an open question what happens if we assume that Ω is not strictly convex at infinity. We conjecture that in this case it is also possible to construct a continuous function on $\partial^{\mathrm{ct}}\Omega$ for which the Dirichlet problem is not solvable, and therefore strict convexity at infinity is also a necessary condition. Since it deals with nonexistence of solutions in arbitrarily large domains, Theorem 6 may have an important role in the study of this conjecture.

To finish, we should mention that there is a large gap between the behavior of K_M at infinity in the cases where Theorem 5 is true and in the ones where it is false. It also remains unknown if there exists some sharp condition on K_M that assures solvability of Problem 2 for any continuous boundary data.

2. Existence result

This section is dedicated to proving Theorem 5. We start with a very important tool, the comparison principle for unbounded domains. It plays an important role in both existence and uniqueness. For now, we just need to work with functions that extend continuously to the asymptotic boundary, however in Section 3 we treat a larger class of functions, as stated above.

Proposition 8 (comparison principle for unbounded domains). *Let $U \subset M$ be an unbounded domain of M. If $u, v \in C^2(U)$ are such that $\mathcal{M}(v) \le \mathcal{M}(u)$ on U with* $\limsup_{p\to x} u \le \liminf_{p\to x} v$ *for all $x \in \partial^{\mathrm{ct}} U$, then $u \le v$ in U.*

Proof. Choose $o \in M$. Let $\varepsilon > 0$ be an arbitrary constant. Using the basis of the cone topology of $\overline{M}$, we obtain that for all $x \in \partial_\infty U$, there is an open truncated cone $N_x := T_o(x, \alpha_x, R_x)$ (that is, the image of a truncated cone of opening angle α_x and radius R_x by the exponential map of a point o) such that $u < v + \varepsilon$ on N_x. Since $\partial_\infty U$ is compact, there exists uniform R such that $u < v + \varepsilon$ on $U \setminus B_R(o)$. In addition, notice that the hypothesis implies that $u \le v$ on ∂U. Therefore we have $u \le v + \varepsilon$ on $\partial(U \cap B_R(o))$, which implies, by the comparison principle on bounded domains, that $u \le v + \varepsilon$ on $U \cap B_R(o)$, and hence the last inequality holds on U. Since ε is arbitrary, the proof is complete. □

We now prove Theorem 5 using Perron's method.

A function $\Sigma \in C^0(\overline{\Omega}^{\mathrm{ct}})$ is called a *supersolution for $\mathcal{M}$* if, given a bounded subdomain $U \subset \Omega$, if $u \in C^2(U) \cap C^0(\overline{U})$ is a solution of $\mathcal{M} = 0$ in U, the condition $u|_{\partial U} \le \Sigma|_{\partial U}$ implies that $u \le \Sigma|_U$. A *subsolution* is defined by replacing $\le$ by $\ge$.

Let S_φ be defined by

$$S_\varphi := \{v \in C^0(\overline{\Omega}^{\mathrm{ct}}) \mid v \text{ is a subsolution for } \mathcal{M} \text{ with } v|_{\partial^{\mathrm{ct}}\Omega} \le \varphi\}.$$

By Proposition 8, $v_0 \equiv \min \varphi \in S_\varphi$, which implies that $S_\varphi \ne \varnothing$, and $w \equiv \max \varphi$ is such that $v \le w$ for all $v \in S_\varphi$. These facts imply that $u : \Omega \to \mathbb{R}$ given by

$$u(x) := \sup\{v(x) \mid v \in S_\varphi\} \tag{1}$$

is well-defined, and we shall prove that under the hypotheses of Theorem 5, we have $u \in C^\infty(\Omega) \cap C(\overline{\Omega}^{\mathrm{ct}})$, $\mathcal{M}(u) = 0$ and $u|_{\partial^{\mathrm{ct}}\Omega} = \varphi$.

We first prove that $u \in C^\infty(\Omega)$ and $\mathcal{M}(u) = 0$. Given $x \in \Omega$, let $r = r_x > 0$ be sufficiently small such that the open geodesic ball of center x and radius r satisfies $B_r(x) \subset \Omega$ and furthermore r satisfies the inequality

$$\frac{(n-1)^2}{n} \coth^2 r \ge - \inf_{B_r(x)} \mathrm{Ric}_M .$$

Such $r > 0$ exists because $\coth r \to +\infty$ as $r \to 0^+$ and Ric_M is of course bounded in bounded sets containing x.

The cylinder $\partial B_r(x) \times \mathbb{R} \subset M \times \mathbb{R}$ has mean curvature $\geq \frac{n-1}{n} \coth r$ (pointing inward) as a consequence of the Hessian comparison theorem, and therefore this choice of r implies, by Theorem 2 of [Dajczer et al. 2013], the existence of minimal graphs in $B_r(x)$ extending continuously to any prescribed continuous boundary data on $\partial B_r(x)$, and this is an essential fact when we use Perron's method.

Consider a sequence $(v_m)_m \subset S_\varphi$ such that $\lim_m v_m(x) = u(x)$. Theorem 2 of [Dajczer et al. 2013] again implies that for each $m \in \mathbb{N}$ there exists a solution $w_{m,x} \in C^\infty(B_r(x)) \cap C(\overline{B_r(x)})$ of $\mathcal{M} = 0$ such that $w_{m,x}|_{\partial B_r(x)} = v_m|_{\partial B_r(x)}$. The interior gradient estimate given by Theorem 1 of [Dajczer et al. 2013] implies that $(w_{m,x})_m$ contains a subsequence converging uniformly on compact subsets of $B_r(x)$ to a solution $w_x \in C^\infty(B_r(x))$ of $\mathcal{M} = 0$. As in [Gilbarg and Trudinger 1998, Section 2.8], one may prove that $w_x = u|_{B_r(x)}$, which implies that $u \in C^\infty(\Omega)$ and $\mathcal{M} = 0$. This is done by taking the limit of minimal lifts $u_m \in S_\varphi$ of each v_m defined by

$$u_m(y) := \begin{cases} v_m(y) & \text{if } y \in \Omega \setminus B_r(x), \\ w_{m,x}(y) & \text{if } y \in B_r(x). \end{cases}$$

We now need to prove that u extends continuously to the desired boundary data on $\partial^{\mathrm{ct}}\Omega$. Since $\partial\Omega$ is mean convex, standard arguments guarantee that the solution assumes the desired data on $\partial\Omega$. To conclude the proof it is necessary to guarantee that it also extends continuously to $\partial_\infty\Omega$, hence in the following we construct barriers at infinity.

Given $x \in \partial^{\mathrm{ct}}\Omega$ and an open subset V such that $x \in \partial^{\mathrm{ct}} V \cap \partial^{\mathrm{ct}}\Omega$, we call an *upper barrier for $\mathcal{M}$ relative to x and V with height C* a function $\Sigma \in C(\Omega)$ such that

(i) Σ is a supersolution for $\mathcal{M}$;

(ii) $\Sigma \geq 0$ and $\lim_{p \in \Omega,\, p \to x} \Sigma(p) = 0$, the limit with respect to the cone topology;

(iii) $\Sigma_{\Omega \setminus V} \geq C$.

Lower barriers are defined analogously.

A point $x \in \partial^{\mathrm{ct}}\Omega$ is said to be regular (with respect to the mean curvature operator $\mathcal{M}$) if it satisfies the following property: given $C > 0$ and a relatively open subset $\Gamma \subset \partial^{\mathrm{ct}}\Omega$ with $x \in \Gamma$, there exist an open set $V \subset \Omega$ such that x is an interior point of $\overline{V} \cap \partial^{\mathrm{ct}}\Omega$ (with respect to the topology induced on the boundary), with $\overline{V} \cap \partial^{\mathrm{ct}}\Omega \subset \Gamma$, and an upper barrier $\Sigma : \Omega \to \mathbb{R}$ relative to x and V with height C.

The following lemma is analogous to Theorem 2.7 of [Choi 1984], but we present the proof for the sake of completeness.

Lemma 9. *The function u given by (1) extends continuously to φ to each regular point $x \in \partial_\infty\Omega$.*

Proof. Given $x \in \partial_\infty\Omega$ and $\varepsilon > 0$, let $\Gamma \subset \partial^{\mathrm{ct}}\Omega$ be such that $|\varphi - \varphi(x)| < \varepsilon/2$ in Γ. Let $\Sigma : \Omega \to \mathbb{R}$ be an upper barrier relative to x and V with height $C = \max|\varphi|$, where V is given by the definition of regularity. It follows that $w := \Sigma + \varphi(x) + \varepsilon$ is a supersolution for $\mathcal{M}$. By the choice of Γ, it holds that $w > \varphi$ on Γ and since $w|_{\partial^{\mathrm{ct}}\Omega \setminus \Gamma} \geq \max|\varphi|$, it of course satisfies $w \geq \varphi$ on $\partial^{\mathrm{ct}}\Omega \setminus \Gamma$. Therefore $v \leq w$ for all $v \in S_\varphi$, which implies that

$$\lim_{p\in\Omega,\, p\to x} u(p) \leq \lim_{p\in\Omega,\, p\to x} w(p) = \varphi(x) + \varepsilon.$$

On the other hand, notice that $v_0 := \varphi(x) - \varepsilon - \Sigma$ belongs to S_φ and therefore $u \geq v_0$ in Ω, which implies that

$$\lim_{p\in\Omega,\, p\to x} u(p) \geq \lim_{p\in\Omega,\, p\to x} v_0(p) = \varphi(x) - \varepsilon.$$

Since ε is arbitrary, we have $\varphi(x) \leq \lim_{p\in\Omega,\, p\to x} u(p) \leq \varphi(x)$. □

To finish, we now prove regularity at the points of $\partial_\infty\Omega$.

Proposition 10. *Let $\Omega \subset M$ be a domain that is strictly convex at infinity. Then $\mathcal{M}$ is regular at each point of $\partial_\infty\Omega$.*

Proof. Let $x \in \partial_\infty\Omega$ and let $\Gamma \subset \partial^{\mathrm{ct}}\Omega$ be a relatively open neighborhood of x. By hypothesis, there exists an open neighborhood $V \subset \Omega$ of x such that $\overline{V} \cap \partial^{\mathrm{ct}}\Omega \subset \Gamma$ and $\partial V \cap \Omega$ has nonnegative principal curvatures.

Let $s : V \to \mathbb{R}$ be the distance function to $\partial V \cap \Omega$. Since $K_M \leq -1$ and all principal curvatures of $\partial V \cap \Omega$ are nonnegative, we have that the Laplacian of s satisfies

$$\Delta s \geq (n-1)\tanh s \tag{2}$$

(see, for instance, Theorem 4.3 of [Choi 1984]).

Define $g : (0, +\infty) \to \mathbb{R}$ by

$$g(s) := \int_s^{+\infty} \frac{dt}{\sqrt{\cosh^{2(n-1)} t - 1}}.$$

Notice that g is well-defined and $\lim_{s\to 0^+} g(s) = +\infty$, $\lim_{s\to+\infty} g(s) = 0$. Define now $w : V \to \mathbb{R}$ by $w(p) := g(s(p))$. A straightforward computation gives

$$\mathcal{M}(w) = (n-1)\cosh^{n-1} s \sinh s + (1-n)\cosh^{1-n} s\, \Delta s$$

and hence the estimate $\Delta s \geq \tanh s$ leads to $\mathcal{M}(w) \leq 0$.

We remark that w is a solution if $M = \mathbb{H}^n$ and V is a totally geodesic hypersphere.

To finish with the proof, define the supersolution $\Sigma \in C^0(\overline{\Omega})$ by

$$\Sigma(p) = \begin{cases} \min\{w(p), C\} & \text{if } p \in V, \\ C & \text{if } p \in \overline{\Omega} \setminus V, \end{cases}$$

which is of course an upper barrier relative to x and V with height C, and hence the proof is complete. □

3. Nonexistence result

We now prove that mean convexity of $\partial\Omega$ is a necessary condition, as stated in Theorem 6. We start with the next classical lemma, proved by Jenkins and Serrin [1968] in the case where the domain is bounded and contained on $\mathbb{R}^n$.

Lemma 11. *Let $U \subset M$ be an open domain and Γ a relatively C^1 open subset of ∂U. If $u \in C(\overline{U}) \cap C^2(U \cup \Gamma)$ and $w \in C(\overline{U}) \cap C^2(U)$ satisfy*

$$\mathcal{M}(w) < \mathcal{M}(u) \quad \text{in } U, \tag{3}$$

$$u \le w \quad \text{on } \partial U \setminus \Gamma, \text{ and} \tag{4}$$

$$\frac{\partial w}{\partial \nu} = -\infty \quad \text{in } \Gamma, \tag{5}$$

where ν is the inner normal vector to ∂U, then $u \le w$ in U.

Proof. If $u \le w$ on Γ, the result is a consequence of the comparison principle. Suppose, towards a contradiction, that there exists $y \in \operatorname{Int}\Gamma$ such that

$$d := \max_{\Gamma}(u - w) = u(y) - w(y) > 0.$$

Then $u \le w + d$ on ∂U, and hence by the comparison principle we have $u \le w + d$ in U. Therefore

$$\frac{\partial}{\partial \nu}(u - w)(y) \le 0 \Rightarrow \frac{\partial}{\partial \nu}(u)(y) \le -\infty,$$

contradicting the fact that $u \in C^2(U \cup \Gamma)$. □

Lemma 12. *Let $\Omega \subset M$ be an open C^2 domain (possibly unbounded) with mean curvature (with respect to the inner normal) $H : \partial\Omega \to \mathbb{R}$. Suppose that there exist $y \in \partial\Omega$ such that $H(y) < 0$. Then there exists $s > 0$ depending only on the local geometry of Ω near y and $C > 0$ depending only on $H(y)$ such that if $u \in C^2(\overline{\Omega}) \cap C(\overline{\Omega}^{\mathrm{ct}})$ satisfy $\mathcal{M}(u) = 0$ in Ω, then*

$$u(y) \le C + \sup_{\partial B_s(y) \cap \Omega} u.$$

Proof. Let $d : \widetilde{\Omega} \to \mathbb{R}$ be given by $d(x) = \operatorname{dist}(x, \partial\Omega)$, where $\widetilde{\Omega} \subset \Omega$ is the open subset where d is smooth. Since $H(y) < 0$, it holds that $\Delta d(y) = -H(y) > 0$. Since $\partial\Omega$ is C^2, there exists $s > 0$ such that $B_s(y) \cap \Omega \subset \widetilde{\Omega}$ and

$$\Delta d(x) > -\frac{H(y)}{2} =: \epsilon, \quad \forall x \in B_s(y) \cap \Omega.$$

This is the required s.

We claim that if $x \in B_s(y) \cap \Omega$, then $u(x) \le C + \sup_{\partial B_s(y) \cap \Omega} u$. To prove it, let Γ_x be the level set of d that contains x and Ω_x be the set enclosed by Γ_x and $\partial B_s(y)$, that is, $\Omega_x := \{p \in B_s(y) \mid d(p) > d(x)\}$.

Consider ψ given by

$$\psi(t) = \tfrac{\pi}{2} - \operatorname{arcsec}(t+1). \tag{6}$$

Then $\psi \geq 0$, $\psi(0) = \pi/2$ and $\lim_{t\to+\infty} \psi(t) = 0$. Its first and second derivatives are given below:

$$\begin{aligned}\psi'(t) &= -\frac{1}{(t+1)\sqrt{t^2+2t}},\\ \psi''(t) &= \frac{1}{(t+1)^2\sqrt{t^2+2t}} + \frac{1}{(t^2+2t)^{3/2}}.\end{aligned}$$

Define $w : B_s(y) \cap \Omega_x \to \mathbb{R}$ by

$$w(p) := A\psi(d(p)) + \sup_{\partial B_s(y)\cap\Omega} u,$$

where $A > 0$ is to be determined. After some computations we obtain

$$(1+|\nabla w|^2)^{3/2}\mathcal{M}(w)(p) = A\psi''(d(p)) + \big(A\psi'(d(p)) + A^3\psi'(d(p))^3\big)\Delta d(p).$$

Using then that $\Delta d(p) > \epsilon$ and $\psi' < 0$ in the domain we are considering, we obtain

$$\begin{aligned}(1+|\nabla w|^2)^{3/2}\mathcal{M}(w) &\leq A\big[\psi'' + \epsilon\psi' + \epsilon A^2\psi'^3\big]\\ &= A(t+1)^{-3}(t^2+2t)^{-3/2}\big[(t+1)(t^2+2t) + (t+1)^3\\ &\qquad - \epsilon(t+1)^2(t^2+2t) - \epsilon A^2\big].\end{aligned}$$

Notice that the term in the brackets is a polynomial of degree 4 with leading coefficient $-\epsilon < 0$ and constant term $1 - \epsilon A^2$. Then it is clear that there exists $A > 0$ large enough that this polynomial is negative for all $t \geq 0$; with this choice of A we obtain that $\mathcal{M}(w) < 0$ on Ω_x.

Furthermore, by definition of w we have $w \geq u$ on $\partial B_s(y)\cap\Omega_x$ and $\partial w/\partial\nu = -\infty$ on Γ_x, which is an open C^1 portion of $\partial\Omega_x$. We also notice that $u \in C^2(\Gamma_x)$. By Lemma 11, we obtain $w \geq u$ in Ω_x. Since x is arbitrary and u is continuous, it holds the desired inequality with $C = A\frac{\pi}{2}$, which concludes the proof. □

Proposition 13. *Let M be a Hadamard manifold with sectional curvature $K_M \leq -1$. There exists universal $C > 0$ such that if Ω is a C^1 domain of M and u satisfies $\mathcal{M}(u) = 0$ in Ω, then*

$$\sup_{\partial B_s(y)\cap\Omega} u \leq C + \sup_{\partial^{\mathrm{ct}}\Omega\setminus B_s(y)} u$$

for all $y \in \partial\Omega$ and $s > 0$ such that $\partial B_s(y) \cap \Omega$ is a nonempty connected set.

Proof. Consider $w : \Omega \setminus B_s(y) \to \mathbb{R}$ given by

$$w(x) = B\psi(r(x)) + \sup_{\partial^{\mathrm{ct}}\Omega\setminus B_s(y)} u,$$

where ψ is given by (6), $r(x) := \operatorname{dist}(x, \partial B_s(y))$ and B is an appropriate constant to be chosen latter. Since $K_M \leq -1$, we have by the Hessian comparison theorem that $\Delta r \geq n-1$. Hence, mimicking the computations of the previous lemma, we obtain the same polynomial, except that we have $n-1$ instead of ϵ and B instead of A. It is again clear that there exists B large enough that $\mathcal{M}(w) \leq 0$. We remark that such a constant does not depend on anything (except in the fact that $K_M \leq -1$) since we may choose the constant that is appropriate to the case $n = 2$ and it works on all dimensions.

We are again in the situation of the hypotheses of Lemma 11, with $U = \Omega \setminus B_s(y)$. Hence we obtain, for all $x \in \partial B_s(y) \cap \Omega$,

$$u(x) \leq \sup_{\partial^{\mathrm{ct}}\Omega \setminus B_s(y)} u + B\frac{\pi}{2}$$

and the proof is complete. □

Proof of Theorem 6. By combining the estimates obtained in Lemma 12 and Proposition 13, we obtain the existence of a continuous function $\varphi : \partial\Omega \to \mathbb{R}$ for which Problem 2 is not solvable: it suffices to put $\varphi(y) = \pi(A+B)$, where A and B are given by the previous results, and $\varphi = 0$ on $\partial\Omega \setminus B_s(y)$, where s is given in the proof of Lemma 12. □

4. Applications

Corollary 14. *Let Ω be a domain that has only finitely many points on $\partial_\infty\Omega$. Then Problem 2 is solvable for any continuous φ if and only if Ω is mean convex.*

Proof. Notice that since $\partial_\infty M$ is compact, $\partial_\infty\Omega$ is also compact and therefore "finitely many points on $\partial_\infty\Omega$" is equivalent to "isolated points on $\partial_\infty\Omega$". In order to apply Theorem 7, it suffices to prove that Ω is strictly convex at infinity.

Given $x \in \partial_\infty\Omega$, let $W \subset \overline{\Omega}^{\mathrm{ct}}$ be a relatively open neighborhood of x. We may suppose without loss of generality that x is the only point at infinity of W, otherwise we just work with any open subset of W where this property holds. Choosing $o \in M \setminus W$, we have that W is contained on some truncated cone centered at o with radius $R > 0$, and as a consequence we have $\partial W \subset M \setminus B_R(o)$. Set $V := \Omega \setminus \overline{B_R(o)}$, and it is clear that it satisfies the required conditions. □

Corollary 15. *If M satisfies the SC condition and Ω is a mean convex domain of M such that $\partial_\infty\Omega$ is composed only of open portions and isolated points, then Problem 2 is solvable in Ω. In particular, this is the case if either* $\dim M = 2$ *or M is rotationally symmetric, or*

$$\min\{K_M(\Pi) \mid \Pi \text{ is a 2-plane in } T_pM,\ p \in B_{R+1}(o)\} \geq -\frac{e^{2kR}}{R^{2+2\epsilon}}, \quad R \geq R^* \tag{7}$$

for some constants ϵ, $R^* > 0$.

Particular cases of Corollary 15 may be found in [Ripoll and Telichevesky 2015].

4.1. ***Application of the technique: Dirichlet problems for p-Laplacians.*** Consider now the following Dirichlet problem for the p-Laplacian operator, $p > 1$, for continuous u in the Sobolev space $W^{1,p}(\Omega)$:

$$\begin{cases} \Delta_p(u) := \operatorname{div}(|\nabla u|^{p-2}\nabla u) = 0 & \text{in } \Omega, \\ u|_{\partial\Omega} = \varphi. \end{cases} \tag{8}$$

Concerning the case $\Omega = M$, the counterexamples constructed by A. Ancona [1994] and by A. Borbély [1998a] show that some convexity at infinity is also needed to obtain existence of solutions of asymptotic Dirichlet problems related to the Laplacian operator Δ. I. Holopainen [2015] constructed a counterexample for the p-Laplacian operator Δ_p. The manifolds constructed by them contain a point in $\partial_\infty M$ with the property that any open neighborhood of it has the whole manifold as the convex hull, and hence M is not strictly convex at infinity.

On the other hand, in [Ripoll and Telichevesky 2015] the authors proved that the SC condition is sufficient for solvability of asymptotic Dirichlet problems with respect to Δ_p. We may extend this result to our case, proving that if Ω is strictly convex at infinity, then every $x \in \partial_\infty\Omega$ is regular with respect to the operator Δ_p.

The proof is *mutatis mutandis* the same as we have done above; it is sufficient to replace $\mathcal{M}$ by Δ_p and the function g constructed in Proposition 10 by

$$g(s) := c\int_s^{+\infty} \cosh^{(1-n)/(p-1)}(t)\,dt,$$

where c is a sufficiently large constant ($c = 2C(\cosh 1)^{(n-1)/(p-1)}$ works).

Together with the classical theory of existence of solutions over bounded domains that satisfy the exterior sphere condition, we obtain the following result.

Theorem 16. *Let M be a Hadamard manifold with sectional $K_M \le -1$. Let $\Omega \subset M$ be an unbounded domain that is strictly convex at infinity and that satisfies the exterior sphere condition on its finite part, namely, given $x \in \partial\Omega$, there exist a sphere contained in $M \setminus \Omega$ that is tangent to $\partial\Omega$ at x. Then (8) is solvable for any $\varphi \in C(\partial^{\mathrm{ct}}\Omega)$.*

Acknowledgment

We would like to thank the referee for valuable remarks and questions.

References

[Aiolfi et al. 2016] A. Aiolfi, J. B. Ripoll, and M. Soret, "The Dirichlet problem for the minimal hypersurface equation on arbitrary domains of a Riemannian manifold", *Manuscripta Math.* **149**:1–2 (2016), 71–81. MR 3447141

[Alías and Dajczer 2007] L. J. Alías and M. Dajczer, "Normal geodesic graphs of constant mean curvature", *J. Differential Geom.* **75**:3 (2007), 387–401. MR 2007m:53005 Zbl 1119.53040

[Ancona 1994] A. Ancona, "Convexity at infinity and Brownian motion on manifolds with unbounded negative curvature", *Rev. Mat. Iberoam.* **10**:1 (1994), 189–220. MR 95a:58132 Zbl 0804.58056

[Barbosa and Sá Earp 1998] J. L. M. Barbosa and R. Sá Earp, "Prescribed mean curvature hypersurfaces in $H^{n+1}(-1)$ with convex planar boundary, I", *Geom. Dedicata* **71**:1 (1998), 61–74. MR 99d:53064 Zbl 0922.53023

[Bernstein 1910] S. Bernstein, "Sur les surfaces définies au moyen de leur courbure moyenne ou totale", *Ann. Sci. École Norm. Sup.* (3) **27** (1910), 233–256. MR 1509123 JFM 41.0692.05

[Borbély 1998a] A. Borbély, "The nonsolvability of the Dirichlet problem on negatively curved manifolds", *Differential Geom. Appl.* **8**:3 (1998), 217–237. MR 99j:53043 Zbl 0947.53019

[Borbély 1998b] A. Borbély, "Some results on the convex hull of finitely many convex sets", *Proc. Amer. Math. Soc.* **126**:5 (1998), 1515–1525. MR 98j:53039 Zbl 0902.53030

[Castéras et al. 2013] J.-B. Castéras, I. Holopainen, and J. B. Ripoll, "On the asymptotic Dirichlet problem for the minimal hypersurface equation in a Hadamard manifold", preprint, 2013. arXiv 1311.5693v1

[Choi 1984] H. I. Choi, "Asymptotic Dirichlet problems for harmonic functions on Riemannian manifolds", *Trans. Amer. Math. Soc.* **281**:2 (1984), 691–716. MR 85b:53040 Zbl 0541.53035

[Collin and Krust 1991] P. Collin and R. Krust, "Le problème de Dirichlet pour l'équation des surfaces minimales sur des domaines non bornés", *Bull. Soc. Math. France* **119**:4 (1991), 443–462. MR 92m:53007 Zbl 0754.53013

[Dajczer et al. 2008] M. Dajczer, P. A. Hinojosa, and J. H. de Lira, "Killing graphs with prescribed mean curvature", *Calc. Var. Partial Differential Equations* **33**:2 (2008), 231–248. MR 2009m:53154 Zbl 1152.53046

[Dajczer et al. 2013] M. Dajczer, J. H. de Lira, and J. B. Ripoll, "An interior gradient estimate for the mean curvature equation of Killing graphs and applications", preprint, 2013. To appear in *J. Anal. Math.* arXiv 1206.2900

[Eberlein and O'Neill 1973] P. Eberlein and B. O'Neill, "Visibility manifolds", *Pacific J. Math.* **46** (1973), 45–109. MR 49 #1421 Zbl 0264.53026

[do Espírito Santo et al. 2010] N. do Espírito Santo, S. Fornari, and J. B. Ripoll, "The Dirichlet problem for the minimal hypersurface equation in $\mathbb{M}\times\mathbb{R}$ with prescribed asymptotic boundary", *J. Math. Pures Appl.* (9) **93**:2 (2010), 204–221. MR 2011e:35138 Zbl 1193.58012

[Finn 1965] R. Finn, "Remarks relevant to minimal surfaces, and to surfaces of prescribed mean curvature", *J. Analyse Math.* **14** (1965), 139–160. MR 32 #6337 Zbl 0163.34604

[Folha and Rosenberg 2012] A. Folha and H. Rosenberg, "The Dirichlet problem for constant mean curvature graphs in $\mathbb{M}\times\mathbb{R}$", *Geom. Topol.* **16**:2 (2012), 1171–1203. MR 2946806 Zbl 1281.53013

[Gálvez and Rosenberg 2010] J. A. Gálvez and H. Rosenberg, "Minimal surfaces and harmonic diffeomorphisms from the complex plane onto certain Hadamard surfaces", *Amer. J. Math.* **132**:5 (2010), 1249–1273. MR 2011j:53012 Zbl 1229.53064

[Gilbarg and Trudinger 1998] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, revised 2nd ed., Grundlehren der Mathematischen Wissenschaften **224**, Springer, Berlin, 1998. MR 2001k:35004 Zbl 1042.35002

[Guio and Sá Earp 2005] E. M. Guio and R. Sá Earp, "Existence and non-existence for a mean curvature equation in hyperbolic space", *Commun. Pure Appl. Anal.* **4**:3 (2005), 549–568. MR 2006i:35093 Zbl 1106.35014

[Haar 1927] A. Haar, "Über das Plateausche Problem", *Math. Ann.* **97**:1 (1927), 124–158. MR 1512358 JFM 52.0710.02

[Holopainen 2015] I. Holopainen, "Nonsolvability of the asymptotic Dirichlet problem for the p-Laplacian on Cartan–Hadamard manifolds", *Adv. Calc. Var.* (online publication February 2015).

[Holopainen and Ripoll 2015] I. Holopainen and J. B. Ripoll, "Nonsolvability of the asymptotic Dirichlet problem for some quasilinear elliptic PDEs on Hadamard manifolds", *Rev. Mat. Iberoam.* **31**:3 (2015), 1107–1129. MR 3420486 Zbl 06503648

[Hwang 1988] J.-F. Hwang, "Comparison principles and Liouville theorems for prescribed mean curvature equations in unbounded domains", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **15**:3 (1988), 341–355. MR 90m:35019 Zbl 0705.49022

[Jenkins and Serrin 1968] H. Jenkins and J. Serrin, "The Dirichlet problem for the minimal surface equation in higher dimensions", *J. Reine Angew. Math.* **229** (1968), 170–187. MR 36 #5519 Zbl 0159.40204

[Krust 1989] R. Krust, "Remarques sur le problème extérieur de Plateau", *Duke Math. J.* **59**:1 (1989), 161–173. MR 90i:49050 Zbl 0709.49022

[Kutev and Tomi 1998] N. Kutev and F. Tomi, "Existence and nonexistence for the exterior Dirichlet problem for the minimal surface equation in the plane", *Differential Integral Equations* **11**:6 (1998), 917–928. MR 99m:35062 Zbl 1074.35550

[Kuwert 1993] E. Kuwert, "On solutions of the exterior Dirichlet problem for the minimal surface equation", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **10**:4 (1993), 445–451. MR 94i:35034 Zbl 0820.35038

[López 2001] R. López, "Graphs of constant mean curvature in hyperbolic space", *Ann. Global Anal. Geom.* **20**:1 (2001), 59–75. MR 2002e:53009 Zbl 0996.53039

[Mazet et al. 2011] L. Mazet, M. M. Rodríguez, and H. Rosenberg, "The Dirichlet problem for the minimal surface equation, with possible infinite boundary data, over domains in a Riemannian surface", *Proc. Lond. Math. Soc.* (3) **102**:6 (2011), 985–1023. MR 2012f:53013 Zbl 1235.53007

[Nitsche 2002] P.-A. Nitsche, "Existence of prescribed mean curvature graphs in hyperbolic space", *Manuscripta Math.* **108**:3 (2002), 349–367. MR 2003f:53015 Zbl 1130.53302

[Radó 1930] T. Radó, "The problem of the least area and the problem of Plateau", *Math. Z.* **32**:1 (1930), 763–796. MR 1545197 JFM 56.0436.01

[Ripoll and Telichevesky 2015] J. B. Ripoll and M. Telichevesky, "Regularity at infinity of Hadamard manifolds with respect to some elliptic operators and applications to asymptotic Dirichlet problems", *Trans. Amer. Math. Soc.* **367**:3 (2015), 1523–1541. MR 3286491 Zbl 1308.58011

[Ripoll and Tomi 2014] J. B. Ripoll and F. Tomi, "On solutions to the exterior Dirichlet problem for the minimal surface equation with catenoidal ends", *Adv. Calc. Var.* **7**:2 (2014), 205–226. MR 3187916 Zbl 1292.35110

[Sá Earp and Rosenberg 1989] R. Sá Earp and H. Rosenberg, "The Dirichlet problem for the minimal surface equation on unbounded planar domains", *J. Math. Pures Appl.* (9) **68**:2 (1989), 163–183. MR 90m:35072 Zbl 0696.49069

Received February 5, 2015. Revised September 17, 2015.

MIRIAM TELICHEVESKY
INSTITUTO DE MATEMÁTICA
UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
91509-900 PORTO ALEGRE, RS
BRAZIL
miriam.telichevesky@ufrgs.br

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 281 No. 1 March 2016