Pacific Journal of Mathematics

Volume 276 No. 1 July 2015

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 2015 is US $420/year for the electronic version, and $570/year for print and electronic.
Subscriptions, requests for back issues and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.1

ON THE DEGREE OF CERTAIN LOCAL L-FUNCTIONS

U. K. ANANDAVARDHANAN AND AMIYA KUMAR MONDAL

Let π be an irreducible supercuspidal representation of $\mathrm{GL}_n(F)$, where F is a p-adic field. By a result of Bushnell and Kutzko, the group of unramified self-twists of π has cardinality n/e, where e is the $\mathfrak{o}_F$-period of the principal $\mathfrak{o}_F$-order in $M_n(F)$ attached to π. This is the degree of the local Rankin–Selberg L-function $L(s, \pi \times \pi^\vee)$. In this paper, we compute the degree of the Asai, symmetric square, and exterior square L-functions associated to π. As an application, assuming p is odd, we compute the conductor of the Asai lift of a supercuspidal representation, where we also make use of the conductor formula for pairs of supercuspidal representations due to Bushnell, Henniart, and Kutzko (1998).

1. Introduction

Let F be a p-adic field. Let $\mathfrak{o}_F$ denote its ring of integers and let $\mathfrak{p}_F$ be the unique maximal ideal of $\mathfrak{o}_F$. Let q denote the cardinality of the residue field $\mathfrak{o}_F/\mathfrak{p}_F$. Let W_F' denote the Weil–Deligne group of F. For a reductive algebraic group G defined over F, let ${}^LG$ be its Langlands dual. Given a Langlands parameter $\rho : W_F' \to {}^LG$ and a finite-dimensional representation $r : {}^LG \to \mathrm{GL}(V)$, we have an L-function $L(s, \rho, r)$ defined as follows. If N is the nilpotent endomorphism of V associated to $r \circ \rho$, then

$$L(s, \rho, r) = \frac{1}{\det\bigl(1 - (r \circ \rho)(\mathrm{Frob})|_{(\mathrm{Ker}\, N)^I}\, q^{-s}\bigr)}$$

where Frob is the geometric Frobenius and I is the inertia subgroup of the Weil group of F. Thus, $L(s, \rho, r) = P(q^{-s})^{-1}$ for some polynomial $P(X)$ with $P(0) = 1$, and by the degree of $L(s, \rho, r)$ we mean the degree of $P(X)$. If $\pi = \pi(\rho)$ denotes the L-packet of irreducible admissible representations of $G(F)$ corresponding to ρ under the conjectural Langlands correspondence, then its Langlands L-function, denoted by $L(s, \pi, r)$, is expected to coincide with $L(s, \rho, r)$. In many cases, candidates for $L(s, \pi, r)$ can also be obtained either via the Rankin–Selberg method

MSC2010: primary 22E50; secondary 11F33, 11F70, 11F85.
Keywords: Asai L-function, symmetric square L-function, exterior square L-function, degree of a local L-function.

of integral representations or by the Langlands–Shahidi method, and in several instances it is known that all these approaches lead to the same L-function [Shahidi 1984; 1990; Anandavardhanan and Rajan 2005; Henniart 2010; Matringe 2011; Kewat and Raghunathan 2012].

Let $G = \mathrm{GL}(n_1) \times \mathrm{GL}(n_2)$. If π_i is an irreducible admissible representation of $\mathrm{GL}_{n_i}(F)$ $(i = 1, 2)$, and if r is the tensor product representation of ${}^L G = \mathrm{GL}_{n_1}(\mathbb{C}) \times \mathrm{GL}_{n_2}(\mathbb{C})$ on $\mathbb{C}^n \otimes \mathbb{C}^n$ given by $r((a, b)) \cdot (x \otimes y) = ax \otimes by$, then the resulting L-function is the Rankin–Selberg L-function $L(s, \pi_1 \times \pi_2)$ [Jacquet et al. 1983; Shahidi 1984]. If we assume that both π_1 and π_2 are supercuspidal representations, then we know that $L(s, \pi_1 \times \pi_2) \equiv 1$ unless $n_1 = n_2$ and $\pi_2^\vee \cong \pi_1 \otimes \chi \circ \det$ for an unramified character χ of $F^\times$. Here, $\pi^\vee$ denotes the representation contragredient to π. Moreover, in the latter case, the degree of $L(s, \pi_1 \times \pi_2)$ is equal to the degree of $L(s, \pi_1 \times \pi_1^\vee)$, which in turn equals the cardinality of the group

$$\{\eta : F^\times \to \mathbb{C}^\times \mid \pi_1 \otimes \eta \circ \det \cong \pi_1, \eta \text{ unramified}\}.$$

The result of Bushnell and Kutzko mentioned in the abstract computes the cardinality of the above group of unramified self-twists of $\pi = \pi_1$ [Bushnell and Kutzko 1993, Lemma 6.2.5]. In order to state the result, let $[\mathfrak{A}, m, 0, \beta]$ be the simple stratum defining a maximal simple type occurring in the irreducible supercuspidal representation π. Here, $\mathfrak{A}$ is a principal $\mathfrak{o}_F$-order in $M_n(F)$, $m \geq 0$ is an integer called the level of π, and $\beta \in M_n(F)$ is such that $F[\beta]$ is a field with $F[\beta]^\times$ normalizing $\mathfrak{A}$. Let $e = e(\mathfrak{A}|\mathfrak{o}_F)$ be the $\mathfrak{o}_F$-period of $\mathfrak{A}$; this quantity in fact equals the ramification index $e(F[\beta]/F)$ of $F[\beta]/F$. Then e divides n, and the cardinality of the group of unramified self-twists of π is n/e. We mention in passing that the level m of π is related to the conductor $f(\pi)$ of π by $f(\pi) = n(1 + m/e)$.

The aim of the present work is to analogously compute the degree of some other local L-functions in the supercuspidal case. Investigating the supercuspidal case would suffice as the L-function of any irreducible admissible representation can usually be built out of L-functions associated to supercuspidal representations. The L-functions that we study in this paper are the Asai L-function, the symmetric square L-function, and the exterior square L-function.

For the Asai L-function, take $G = \mathrm{Res}_{E/F}\, \mathrm{GL}(n)$, the Weil restriction of $\mathrm{GL}(n)$, where E is a quadratic extension of F. Thus, $G(F) = \mathrm{GL}_n(E)$. In this case, the dual group is ${}^L G = \mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C}) \rtimes \mathrm{Gal}(E/F)$, where the nontrivial element σ of the Galois group $\mathrm{Gal}(E/F)$ acts by $\sigma \cdot (a, b) = (b, a)$. The representation r is the Asai representation, also known as the twisted tensor representation, of ${}^L G$ on $\mathbb{C}^n \otimes \mathbb{C}^n$ given by $r((a, b)) \cdot (x \otimes y) = ax \otimes by$ and $r(\sigma) \cdot (x \otimes y) = y \otimes x$. The Asai L-function can be studied both by the Rankin–Selberg method (see [Flicker 1993, Appendix; Kable 2004]) and by the Langlands–Shahidi method [Shahidi 1990]. It is also known that all three definitions match [Anandavardhanan and Rajan 2005; Henniart 2010; Matringe 2011].

For the symmetric square L-function (resp. the exterior square L-function), take $G = \mathrm{GL}(n)$ and let r be the symmetric square (resp. the exterior square) of the standard representation of ${}^L G = \mathrm{GL}_n(\mathbb{C})$. The Langlands–Shahidi theory of these L-functions is satisfactorily understood [Shahidi 1990; 1992] and this definition is known to match with the one via the Langlands formalism [Henniart 2010]. For the Rankin–Selberg theory of these L-functions, we refer to [Jacquet and Shalika 1990; Bump and Ginzburg 1992; Kewat and Raghunathan 2012].

These L-functions are ubiquitous in number theory and the degree of $L(s, \pi, r)$ often has several meaningful and important interpretations. For instance, these L-functions detect functorial lifts from classical groups. In particular, by the work of Shahidi [1992] and Goldberg [1994], for an irreducible supercuspidal representation π, the degree of $L(s, \pi, r)$ is either the number of unramified twists or half the number of unramified twists of π which are functorial lifts from classical groups (see [Shahidi 1992, Theorem 7.7] and [Goldberg 1994, Theorems 5.1 and 5.2]). We refer to Section 2 for some more details in this regard. Since reducibility of parabolic induction is understood in terms of poles of these L-functions, the degree of $L(s, \pi, r)$ when π is self-dual if $r = \mathrm{Sym}^2$ or $\bigwedge^2$, or when π is conjugate self-dual if $r = \mathrm{Asai}$, counts the number of unramified twists or half the number of unramified twists of π such that the parabolically induced representation to the relevant classical group is irreducible (see [Shahidi 1992, Theorem 7.6] and [Goldberg 1994, Theorem 6.5]).

These L-functions are also related to the theory of distinguished representations. If π is a supercuspidal representation of $\mathrm{GL}_n(E)$, then the degree of its Asai L-function is the number of unramified characters μ of $F^\times$ for which π is μ-distinguished with respect to $\mathrm{GL}_n(F)$ (see [Anandavardhanan et al. 2004, Corollary 1.5]). Similarly, if π is a supercuspidal representation of $\mathrm{GL}_n(F)$, the degree of its exterior square L-function is half the number of unramified characters μ of $F^\times$ such that $\pi \otimes \mu \circ \det$ admits a Shalika functional (see [Jiang et al. 2008, Theorem 5.5]).

Our main theorem computes the degree of $L(s, \pi, r)$, when π is a supercuspidal representation, in terms of the simple stratum $[\mathfrak{A}, m, 0, \beta]$ defining a maximal simple type occurring in the irreducible supercuspidal representation π. Note that π is a supercuspidal representation of $\mathrm{GL}_n(E)$, with E/F a quadratic extension, in the Asai case, whereas otherwise it is a supercuspidal representation of $\mathrm{GL}_n(F)$. As before, let e denote the $\mathfrak{o}$-period of $\mathfrak{A}$ where $\mathfrak{o} = \mathfrak{o}_E$ in the Asai case and $\mathfrak{o} = \mathfrak{o}_F$ otherwise.

Let $\omega = \omega_{E/F}$ be the quadratic character of $F^\times$ associated to the extension E/F and let κ be an extension of ω to $E^\times$. For the purposes of this paper, let us say that a supercuspidal representation, and more generally a discrete series representation, π of $\mathrm{GL}_n(E)$ is distinguished (resp. ω-distinguished) if its Asai L-function $L(s,\pi,r)$ (resp. $L(s,\pi\otimes\kappa,r)$) has a pole at $s=0$. Strictly speaking, this is not how distinction is usually defined, but the property above does characterize distinction for the pair $(\mathrm{GL}_n(E),\mathrm{GL}_n(F))$ (see [Anandavardhanan et al. 2004, Corollary 1.5]). It follows that a supercuspidal representation, and more generally a discrete series representation, cannot be both distinguished and ω-distinguished because of the identity

$$L(s,\pi\times\pi^\sigma)=L(s,\pi,r)L(s,\pi\otimes\kappa,r).$$

Here, σ is the nontrivial element of the Galois group $\mathrm{Gal}(E/F)$.

Recall also that a supercuspidal representation π, and more generally a discrete series representation, of $\mathrm{GL}_n(F)$ which is self-dual is said to be orthogonal (resp. symplectic) if its symmetric square L-function $L(s,\pi,\mathrm{Sym}^2)$ (resp. its exterior square L-function $L(s,\pi,\bigwedge^2)$) has a pole at $s=0$. Thus, a supercuspidal representation, and more generally a discrete series representation, cannot be both orthogonal and symplectic, since we have the factorization

$$L(s,\pi\times\pi)=L(s,\pi,\mathrm{Sym}^2)L(s,\pi,\textstyle\bigwedge^2).$$

Thanks to the above factorizations, if π is a supercuspidal representation of $\mathrm{GL}_n(E)$, we can conclude that

$$\deg L(s,\pi,r)+\deg L(s,\pi\otimes\kappa,r)=\begin{cases}2n/e & \text{if } E/F \text{ is unramified,}\\ n/e & \text{if } E/F \text{ is ramified.}\end{cases}$$

Similarly, if π is a supercuspidal representation of $\mathrm{GL}_n(F)$, then

$$\deg L(s,\pi,\mathrm{Sym}^2)+\deg L(s,\pi,\textstyle\bigwedge^2)=n/e$$

by the result of Bushnell and Kutzko mentioned earlier. Our main results assert that if both the degrees on the left-hand side of the above identities are nonzero, then they are equal.

To state the result more precisely, we introduce the following notion. Let $[\pi]$ denote the inertial equivalence class of π; thus $[\pi]$ consists of all the unramified twists of π. We say that $[\pi]$ is μ-distinguished (resp. orthogonal, symplectic) if there is an unramified twist of π which is μ-distinguished (resp. orthogonal, symplectic). Now we state the main results of this paper.

Theorem 1.1. *Let π be a supercuspidal representation of* $\mathrm{GL}_n(E)$, *with E/F a quadratic extension. Let e be the $\mathfrak{o}_E$-period of the principal $\mathfrak{o}_E$-order in $M_n(E)$ attached to π. Let $L(s,\pi,r)$ be the Asai L-function of π.*

(1) *Suppose E/F is unramified. Then the degree of $L(s, \pi, r)$ is*

$$d(\text{Asai}) = \begin{cases} 0 & \text{if } [\pi] \text{ is not distinguished,} \\ n/e & \text{if } [\pi] \text{ is distinguished.} \end{cases}$$

(2) *Suppose E/F is ramified. Then the degree of $L(s, \pi, r)$ is*

$$d(\text{Asai}) = \begin{cases} 0 & \text{if } [\pi] \text{ is not distinguished,} \\ n/2e & \text{if } [\pi] \text{ is both distinguished and } \omega\text{-distinguished,} \\ n/e & \text{if } [\pi] \text{ is distinguished but not } \omega\text{-distinguished.} \end{cases}$$

Theorem 1.2. *Let π be a supercuspidal representation of $\mathrm{GL}_n(F)$. Let e be the $\mathfrak{o}_F$-period of the principal $\mathfrak{o}_F$-order in $M_n(F)$ attached to π. Then the degree of its symmetric square L-function $L(s, \pi, \mathrm{Sym}^2)$ is*

$$d(\mathrm{Sym}^2) = \begin{cases} 0 & \text{if } [\pi] \text{ is not orthogonal,} \\ n/2e & \text{if } [\pi] \text{ is both orthogonal and symplectic,} \\ n/e & \text{if } [\pi] \text{ is orthogonal but not symplectic.} \end{cases}$$

Theorem 1.3. *Let π be a supercuspidal representation of $\mathrm{GL}_n(F)$. Let e be the $\mathfrak{o}_F$-period of the principal $\mathfrak{o}_F$-order in $M_n(F)$ attached to π. Then the degree of its exterior square L-function $L(s, \pi, \bigwedge^2)$ is*

$$d(\textstyle\bigwedge^2) = \begin{cases} 0 & \text{if } [\pi] \text{ is not symplectic,} \\ n/2e & \text{if } [\pi] \text{ is both symplectic and orthogonal,} \\ n/e & \text{if } [\pi] \text{ is symplectic but not orthogonal.} \end{cases}$$

Remark. As mentioned earlier, a consequence of Theorems 1.2 and 1.3 is that

$$\deg L(s, \pi, \mathrm{Sym}^2) = \deg L(s, \pi, \textstyle\bigwedge^2)$$

if both these L-functions are not identically 1. In this context, we also refer to the remark following Theorem 2.1 in Section 2, which places the above observation in the framework of the work of Shahidi [1992].

Finally, in Section 6, we prove the following theorem. We stress that the assumption of odd residue characteristic is essential in its proof.

Theorem 1.4. *Let E/F be a quadratic extension of p-adic fields. If it is ramified, assume also that $p \neq 2$. Let κ be a character of $E^\times$ which restricts to the quadratic character $\omega_{E/F}$ of $F^\times$ associated to E/F. Let π be an irreducible supercuspidal representation of $\mathrm{GL}_n(E)$ and let $r(\pi)$ be its Asai lift to $\mathrm{GL}_{n^2}(F)$. Then*

$$f(r(\pi)) + \deg L(s, \pi, r) = f(r(\pi) \otimes \omega_{E/F}) + \deg L(s, \pi \otimes \kappa, r).$$

Remark. The conductor formula of Bushnell, Henniart, and Kutzko [1998, Theorem 6.5] gives an explicit formula for $f(\pi \times \pi^\sigma)$ (see Section 5). Thus, together with Theorem 1.1 and this explicit conductor formula for pairs of supercuspidal representations of general linear groups, Theorem 1.4 in fact produces an explicit conductor formula for the Asai lift. Since the statement of such an explicit formula involves introducing further notations, we leave the precise formula to Section 6 (see Theorem 6.1).

2. Results of Shahidi and Goldberg

We recall the results of [Shahidi 1992; Goldberg 1994] to place our Theorems 1.1, 1.2, and 1.3 in context. For the unexplained definitions in the following, we refer to [Shahidi 1992, Definitions 7.4 and 7.5].

Theorem 2.1 [Shahidi 1992, Theorem 7.7]. *Let π be an irreducible supercuspidal representation of* $\mathrm{GL}_n(F)$.

(1) *The L-function $L(s,\pi,\bigwedge^2)$ is identically* 1 *unless some unramified twist of π is self-dual. Assume π is self-dual. Let S be the (possibly empty) set of all the unramified characters η, no two of which have equal squares, for which $\pi\otimes\eta\circ\det$ comes from $\mathrm{SO}_{n+1}(F)$. Then*
$$L(s,\pi,\textstyle\bigwedge^2)=\displaystyle\prod_{\eta\in S}\left(1-\eta^2(\varpi)q^{-s}\right)^{-1}.$$

(2) *The L-function $L(s,\pi,\mathrm{Sym}^2)$ is identically* 1 *unless some unramified twist of π is self-dual. Assume π is self-dual. If π comes from $\mathrm{Sp}_{n-1}(F)$, then*
$$L(s,\pi,\mathrm{Sym}^2)=(1-q^{-rs})^{-1},$$
where r is the number of unramified self-twists of π. Otherwise, let S' be the (possibly empty) set of all the unramified characters η, no two of which have equal squares, for which $\pi\otimes\eta\circ\det$ comes from $\mathrm{SO}_n^*(F)$. Then
$$L(s,\pi,\mathrm{Sym}^2)=\prod_{\eta\in S'}\left(1-\eta^2(\varpi)q^{-s}\right)^{-1}.$$

Remark. A consequence of Theorem 1.2 and Theorem 1.3 is that S and S' have the same cardinality if both these sets are nonempty.

Next we state Theorems 5.1 and 5.2 of [Goldberg 1994]. Here, E/F is a quadratic extension of p-adic fields and σ denotes the nontrivial element of $\mathrm{Gal}(E/F)$. For an irreducible admissible representation of $\mathrm{GL}_n(E)$, let $L(s,\pi,r)$ denote its Asai L-function. In the following, $q=q_F$ is the residue cardinality of F. For the unexplained definitions in the following two theorems, we refer to Definitions 1.11 and 1.12 of [Goldberg 1994].

Theorem 2.2. *Let n be odd. Suppose that π is an irreducible supercuspidal representation of $\mathrm{GL}_n(E)$ such that $\pi^\vee\cong\pi^\sigma$. Let S be the set of all unramified* *characters η of $E^\times$, no two of which have equal squares, such that $\pi \otimes \eta \circ \det$ is a stable lift from $U(n, E/F)$.*

(1) *Suppose E/F is ramified. Then*

$$L(s, \pi, r) = \prod_{\eta \in S} \bigl(1 - \eta(\varpi_F) q^{-s}\bigr)^{-1}.$$

(2) *Suppose E/F is unramified. Then*

$$L(s, \pi, r) = \prod_{\eta \in S} \bigl(1 - \eta^2(\varpi_F) q^{-s}\bigr)^{-1}.$$

Theorem 2.3. *Let n be even. Suppose that π is an irreducible supercuspidal representation of $\mathrm{GL}_n(E)$ such that $\pi^\vee \cong \pi^\sigma$. Let S be the set of all unramified characters η of $E^\times$, no two of which have equal value at ϖ_F, such that $\pi \otimes \eta \circ \det$ is an unstable lift from $U(n, E/F)$. Then*

$$L(s, \pi, r) = \prod_{\eta \in S} \bigl(1 - \eta(\varpi_F) q^{-s}\bigr)^{-1}.$$

Remark. Theorem 1.1 computes explicitly the cardinality of S in Theorems 2.2 and 2.3.

3. The Asai lift

We collect together various results on the Asai representation in this section.

Let H be a subgroup of index two in a group G. Let ρ be a finite dimensional representation of H of dimension n. Its Asai lift, which we do not define here, is a representation of G of dimension n^2. Let $r(\rho)$ denote the Asai lift of ρ to G. The following proposition summarizes the key properties of the Asai lift (see [Prasad 1999; Murty and Prasad 2000]).

Proposition 3.1. *The Asai lift satisfies*:

(1) $r(\rho_1 \otimes \rho_2) \cong r(\rho_1) \otimes r(\rho_2)$.

(2) $r(\rho)^\vee \cong r(\rho^\vee)$.

(3) *$r(\chi)$ for a character χ is $\chi \circ \mathrm{tr}$, where tr is the transfer map from G to the abelianization of H.*

(4) *$r(\rho^\sigma) \cong r(\rho)$, where σ is the nontrivial element of G/H.*

(5) $r(\rho)|_H \cong \rho \otimes \rho^\sigma$.

(6) *For a representation τ of G, we have $r(\tau|_H) \cong \mathrm{Sym}^2 \tau \oplus \omega_{G/H} \bigwedge^2 \tau$, where $\omega_{G/H}$ is the nontrivial character of G/H.*

(7) *Let $\mathrm{Ind}_H^G \rho$ denote the representation of G induced from ρ. Then*:

(a) $\mathrm{Sym}^2(\mathrm{Ind}_H^G \rho) \cong \mathrm{Ind}_H^G \mathrm{Sym}^2 \rho \oplus r(\rho)$.
(b) $\bigwedge^2(\mathrm{Ind}_H^G \rho) \cong \mathrm{Ind}_H^G \bigwedge^2 \rho \oplus r(\rho) \otimes \omega_{G/H}$.

Remark. We have assumed $[G : H] = 2$ since that is the case of interest to us. The Asai lift can be more generally defined when H is of any finite index in G.

4. Proofs of Theorems 1.1–1.3

We now prove Theorems 1.1, 1.2, and 1.3. We first prove (1) of Theorem 1.1, use this to prove Theorems 1.2 and 1.3, and finally prove (2) of Theorem 1.1. We will appeal to a result mentioned in Section 1, which we formally state now for ease of reference.

Theorem 4.1 [Bushnell and Kutzko 1993, Lemma 6.2.5]. *Let π be an irreducible supercuspidal representation of $\mathrm{GL}_n(E)$. Let $[\mathfrak{A}, m, 0, \beta]$ be the simple stratum defining a maximal simple type occurring in π, where $\mathfrak{A}$ is a principal $\mathfrak{o}_E$-order in $M_n(E)$, $m \geq 0$ is the level of π, and $\beta \in M_n(E)$ is such that $E[\beta]$ is a field with $E[\beta]^\times$ normalizing $\mathfrak{A}$. Let $e = e(\mathfrak{A}|\mathfrak{o}_E)$ be the $\mathfrak{o}_E$-period of $\mathfrak{A}$ (which is the same as the ramification index $e(E[\beta]/E)$ of $E[\beta]/E$). Then e divides n, and the cardinality of the group of unramified self-twists of π is n/e.*

Proof of Theorem 1.1(1). Let E/F be quadratic unramified. Let π be a supercuspidal representation of $\mathrm{GL}_n(E)$. Let $\rho_\pi : W_E \to \mathrm{GL}_n(\mathbb{C})$ be its Langlands parameter. We assume that its Asai lift $r(\rho_\pi) : W_F \to \mathrm{GL}_{n^2}(\mathbb{C})$ contains the trivial character of W_F, which in particular implies that $\rho_\pi^\sigma \cong \rho_\pi^\vee$. Since $\omega = \omega_{E/F}$ is unramified, clearly the number of unramified characters in $r(\rho_\pi)$ and $r(\rho_\pi) \otimes \omega$ is the same. Since

$$\deg L(s, \pi, r) + \deg L(s, \pi \otimes \kappa, r) = \deg L(s, \pi \times \pi^\vee) = 2n/e$$

by Theorem 4.1, item (1) of Theorem 1.1 is immediate. $\square$

Proof of Theorems 1.2 and 1.3. Let π be a supercuspidal representation of $\mathrm{GL}_n(F)$. Let $\rho_\pi : W_F \to \mathrm{GL}_n(\mathbb{C})$ be its Langlands parameter. We assume that $r(\rho_\pi)$ contains the trivial character of W_F, which in particular implies that $\rho_\pi \cong \rho_\pi^\vee$. Here, r is either the symmetric square representation or the exterior square representation of $\mathrm{GL}_n(\mathbb{C})$. Thus, the dimension of $r(\rho_\pi)$ is either $n(n+1)/2$ or $n(n-1)/2$. We have the identity

$$L(s, \pi \times \pi) = L(s, \pi, \mathrm{Sym}^2) L(s, \pi, \textstyle\bigwedge^2),$$

and we know that the left-hand side L-function has degree n/e by Theorem 4.1.

If $n/e = 1$, then the trivial character of W_F is the only unramified character appearing in $\rho_\pi \otimes \rho_\pi^\vee$ and hence in $r(\rho_\pi)$. Therefore, in this case there is nothing to prove. Otherwise, there is a nontrivial unramified character $\chi : W_F \to \mathbb{C}^\times$ such that $\rho_\pi \otimes \chi \cong \rho_\pi$. Thus,

$$\rho_\pi = \mathrm{Ind}_{W_{F'}}^{W_F} \tau$$

for some irreducible representation τ of $W_{F'}$, where F'/F is the unramified extension of degree n/e. Let σ denote a generator of $\mathrm{Gal}(F'/F)$.

We know that

$$\rho_\pi \otimes \rho_\pi = \mathrm{Ind}_{W_{F'}}^{W_F}(\tau \otimes \tau) \oplus \mathrm{Ind}_{W_{F'}}^{W_F}(\tau \otimes \tau^\sigma) \oplus \cdots \oplus \mathrm{Ind}_{W_{F'}}^{W_F}\left(\tau \otimes \tau^{\sigma^{n/e-1}}\right).$$

If n/e is an odd integer, then observe that each summand other than the first one on the right-hand side of the above identity appears twice. This is indeed the case since

$$\mathrm{Ind}_{W_{F'}}^{W_F}\left(\tau \otimes \tau^{\sigma^a}\right) = \mathrm{Ind}_{W_{F'}}^{W_F}\left(\tau \otimes \tau^{\sigma^{n/e-a}}\right)$$

for every $1 \le a \le n/e$. Since the trivial character of W_F appears exactly once on the left-hand side, it follows that

$$1 \in \mathrm{Ind}_{W_{F'}}^{W_F}(\tau \otimes \tau)$$

when n/e is odd. Therefore, precisely one of $\mathrm{Sym}^2 \tau$ or $\bigwedge^2 \tau$ contains the trivial character of $W_{F'}$, and hence precisely one of $\mathrm{Ind}_{W_{F'}}^{W_F}(\mathrm{Sym}^2 \tau)$ or $\mathrm{Ind}_{W_{F'}}^{W_F}(\bigwedge^2 \tau)$ contains all the unramified self-twists of ρ_π. Thus, Theorems 1.2 and 1.3 follow in the case when n/e is an odd integer.

If n/e is an even integer, we proceed by induction on $\dim \rho_\pi$. We start by writing $\rho_\pi = \mathrm{Ind}_{W_E}^{W_F} \tau$ for an irreducible representation τ of W_E, where E is the quadratic unramified extension of F. This can always be done because an unramified extension of even degree necessarily has the quadratic unramified subextension. By (7) of Proposition 3.1, we have

$$r(\rho_\pi) \cong \begin{cases} \mathrm{Ind}_{W_E}^{W_F} \mathrm{Sym}^2 \tau \oplus \mathrm{Asai}(\tau) & \text{if } r = \mathrm{Sym}^2, \\ \mathrm{Ind}_{W_E}^{W_F} \bigwedge^2 \tau \oplus \mathrm{Asai}(\tau) \otimes \omega_{E/F} & \text{if } r = \bigwedge^2. \end{cases}$$

Now either $\tau \cong \tau^\vee$ or $\tau^\sigma \cong \tau^\vee$ but not both, since ρ_π is an irreducible representation of W_F. Here, σ is the element of order two in $\mathrm{Gal}(E/F)$. We claim that $\mathrm{Asai}(\tau)$ (resp. $\mathrm{Asai}(\tau) \otimes \omega_{E/F}$) contains an unramified character of W_F only if $\mathrm{Sym}^2 \tau$ (resp. $\bigwedge^2 \tau$) does not contain an unramified character of W_F. Indeed, if $\mathrm{Asai}(\tau)$ contains an unramified character of W_F, the total number of unramified characters in

$$\mathrm{Asai}(\tau) \oplus \mathrm{Asai}(\tau) \otimes \omega_{E/F}$$

is $n/2e + n/2e = n/e$, by applying part (1) of Theorem 1.1 to the representation τ which has dimension $n/2$, and by observing that $\omega_{E/F}$ is unramified. Note also that $e = e(\rho_\pi) = e(\tau)$, since the extension E/F is unramified. Since this number equals the number of unramified characters contained in

$$\rho_\pi \otimes \rho_\pi = \mathrm{Sym}^2 \rho_\pi \oplus \bigwedge^2 \rho_\pi,$$

the claim follows.

Therefore, if $\mathrm{Asai}(\tau)$ contains an unramified character, the proof is complete by appealing to part (1) of Theorem 1.1. Otherwise, since $\dim\tau = \frac{1}{2}\dim\rho_\pi$, the proof is complete by appealing to the induction hypothesis. Note that the base case of the induction is easily verified since there are at most two unramified characters to consider when $\dim\rho_\pi = 2$, i.e., when $\dim\tau = 1$. $\square$

Proof of Theorem 1.1(2). Now let E/F be a ramified quadratic extension, and let π be a supercuspidal representation of $\mathrm{GL}_n(E)$. Let $\rho_\pi : W_E \to \mathrm{GL}_n(\mathbb{C})$ be its Langlands parameter. We may assume that $r(\rho_\pi) \ni 1$, where r denotes the Asai lift from W_E to W_F. Note that this implies that $r(\rho_\pi)$ does not contain $\omega_{E/F}$, the nontrivial character of W_F/W_E. In what follows, we use this assumption many times to reduce the number of cases that we need to analyze.

Consider the $2n$-dimensional representation $\mathrm{Ind}_{W_E}^{W_F}\rho_\pi$ of W_F. We have

$$\mathrm{Sym}^2(\mathrm{Ind}_{W_E}^{W_F}\rho_\pi) \cong \mathrm{Ind}_{W_E}^{W_F}\mathrm{Sym}^2\rho_\pi \oplus r(\rho_\pi), \tag{1}$$

$$\textstyle\bigwedge^2(\mathrm{Ind}_{W_E}^{W_F}\rho_\pi) \cong \mathrm{Ind}_{W_E}^{W_F}\bigwedge^2\rho_\pi \oplus r(\rho_\pi)\otimes\omega_{E/F}. \tag{2}$$

We divide the proof into two cases.

First, we assume that $\pi \ncong \pi^\sigma$ so that $\mathrm{Ind}_{W_E}^{W_F}\rho_\pi$ is irreducible. Let $\mathrm{Ind}_E^F\pi$ denote the corresponding supercuspidal representation of $\mathrm{GL}_{2n}(F)$. Note that by our assumption that $r(\rho_\pi)\ni 1$, $\mathrm{Ind}_E^F\pi$ is orthogonal and not symplectic by (1). Therefore, it follows from Theorems 1.2 and 1.3 that

$$x = \deg L(s, \mathrm{Ind}_E^F\pi, \mathrm{Sym}^2) - \deg L(s, \mathrm{Ind}_E^F\pi, \textstyle\bigwedge^2)$$

is given by

$$x = \begin{cases} \deg L(s, \mathrm{Ind}_E^F\pi, \mathrm{Sym}^2) & \text{if } [\mathrm{Ind}_E^F\pi] \text{ is orthogonal but not symplectic,} \\ 0 & \text{if } [\mathrm{Ind}_E^F\pi] \text{ is orthogonal and symplectic.} \end{cases} \tag{3}$$

Since the extension E/F is ramified, the period associated to $\mathrm{Ind}_E^F\pi$ may be e or $2e$, and thus the degree of $L(s, \mathrm{Ind}_E^F\pi, \mathrm{Sym}^2)$ is either $2n/e$ or n/e.

On the other hand, the difference

$$y = \deg L(s,\pi,\mathrm{Sym}^2) - \deg L(s,\pi,\textstyle\bigwedge^2) \tag{4}$$

could be, a priori, n/e or 0 or $-n/e$.

Now we do a case-by-case analysis to list all the possible candidates for the pair (x, y). To this end, note that:

(i) In (1) and (2), possible values for the degree of the first summand on the right-hand side are 0, n/e, and $n/2e$ (by Theorems 1.2 and 1.3).

(ii) In (1) and (2), the second summand on the right-hand side cannot have degree more than n/e (since $\deg L(s,\pi,r) + \deg L(s,\pi\otimes\kappa, r) = n/e$).

(iii) In addition, in (1), the second summand on the right hand side has nonzero degree (by the assumption that $r(\rho_\pi) \ni 1$).

We have already observed, using (3), that when $x \neq 0$, it is either $2n/e$ or n/e, and the degree of $L(s, \operatorname{Ind}_E^F \pi, \bigwedge^2)$ is 0. In particular, when $x \neq 0$, all the terms in (2) have degree 0. When $x = 2n/e$, both the summands in (1) have degree n/e by (i) and (ii), and thus $y = n/e$. When $x = n/e$, we claim that the degree of the first summand in (1) is 0 (and that of the second summand is n/e), and thus $y = 0$. Indeed, if the first summand had nonzero degree it would have to be either n/e or $n/2e$ by (i). But it cannot be n/e by (iii), and it cannot be $n/2e$ since this would imply that the second summand in (2), which we know to be 0, would have degree $n/2e$ as well.

When $x = 0$, the degrees of the left-hand sides in both (1) and (2) are equal by (3), and are either n/e or $n/2e$. When this degree is n/e, the degree of the first summand in (1) is either 0 or $n/2e$ by (iii). Note that the degree of the first summand in (2) would then be either n/e or $n/2e$ respectively, and thus $y = -n/e$ or 0 respectively. In the preceding argument, we have made use of the identity

$$\deg L(s, \pi, r) + \deg L(s, \pi \otimes \kappa, r) = n/e. \tag{5}$$

When the degrees of the left-hand sides in both (1) and (2) are $n/2e$, the degrees of the first summands are both 0. Thus $y = 0$, once again by arguing with (i), (iii), and (5).

Observe that since E/F is ramified, the number of unramified characters in $\operatorname{Ind}_{W_E}^{W_F} \operatorname{Sym}^2 \rho_\pi$ (resp. in $\operatorname{Ind}_{W_E}^{W_F} \bigwedge^2 \rho_\pi$) is the same as the number of unramified characters in $\operatorname{Sym}^2 \rho_\pi$ (resp. in $\bigwedge^2 \rho_\pi$). It follows that

$$\deg L(s, \pi, r) - \deg L(s, \pi \otimes \kappa, r) = x - y$$

is either n/e or 0. This proves (2) of Theorem 1.1 in this case.

Next, suppose that $\pi \cong \pi^\sigma \cong \pi^\vee$. Since $\pi \cong \pi^\sigma$, it follows that

$$\rho_\pi \cong \tau|_{W_E}$$

for an irreducible representation τ of W_F. In this case,

$$\operatorname{Ind}_{W_E}^{W_F} \rho_\pi \cong \tau \oplus \tau \otimes \omega_{E/F}.$$

Thus, we get

$$\operatorname{Sym}^2 \tau \oplus \operatorname{Sym}^2 \tau \oplus \tau \otimes \tau \otimes \omega_{E/F} \cong \operatorname{Sym}^2(\operatorname{Ind}_{W_E}^{W_F} \rho_\pi) \cong \operatorname{Ind}_{W_E}^{W_F} \operatorname{Sym}^2 \rho_\pi \oplus r(\rho_\pi) \tag{6}$$

and

$$\textstyle\bigwedge^2 \tau \oplus \bigwedge^2 \tau \oplus \tau \otimes \tau \otimes \omega_{E/F} \cong \bigwedge^2(\operatorname{Ind}_{W_E}^{W_F} \rho_\pi) \cong \operatorname{Ind}_{W_E}^{W_F} \bigwedge^2 \rho_\pi \oplus r(\rho_\pi) \otimes \omega_{E/F}. \tag{7}$$

By our assumption that $r(\rho_\pi) \ni 1$, we conclude that the irreducible representation τ is not symplectic. This is because if $\bigwedge^2 \tau \ni 1$, then the left-hand side of (7) contains the trivial character at least twice whereas the right-hand side can contain the trivial character at most once since $r(\rho_\pi) \otimes \omega_{E/F} \not\ni 1$.

As before, we now do a case-by-case analysis to list all possible pairs (a, b) where

$$a = \deg L(s, \tau, \mathrm{Sym}^2) - \deg L(s, \tau, \textstyle\bigwedge^2), \tag{8}$$

$$b = \deg L(s, \rho_\pi, \mathrm{Sym}^2) - \deg L(s, \rho_\pi, \textstyle\bigwedge^2), \tag{9}$$

and we verify that

$$\deg L(s, \rho_\pi, r) - \deg L(s, \rho_\pi \otimes \kappa, r) = 2a - b$$

is either n/e or 0.

Since we have observed that the irreducible representation τ is not symplectic, $a \geq 0$ and it is either n/e or 0 by Theorems 1.2 and 1.3. Now the possible values for b could be, a priori, n/e or 0 or $-n/e$.

When $a = n/e$, considering the sum of (6) and (7), we can conclude that all the terms in (7) are of degree 0. Also, note that both the terms on the right-hand side of (6) will have degree n/e, and in particular $b = n/e$. When $a = 0$, the left-hand sides of both (6) and (7) are each of total degree n/e. Since $r(\rho_\pi) \ni 1$, the degree of $L(s, \rho_\pi, \mathrm{Sym}^2)$ is either 0 or $n/2e$. It follows that the value of b is either $-n/e$ or 0 respectively. Thus, in all cases $2a - b$ is n/e or 0, and the result follows. □

5. The conductor formula of Bushnell, Henniart, and Kutzko

We state the explicit conductor formula for pairs of supercuspidal representation due to Bushnell, Henniart, and Kutzko. This section closely follows [Bushnell et al. 1998, § 6].

Let π be a supercuspidal representation of $\mathrm{GL}_n(F)$. Following [Bushnell and Kutzko 1993], let $[\mathfrak{A}, m, 0, \beta]$ be a simple stratum of a maximal simple type occurring in π. Here, $\mathfrak{A}$ is a principal $\mathfrak{o}_F$-order in $M_n(F)$, m is the level of π, and $\beta \in M_n(F)$ is such that $E = F[\beta]$ is a field with $E^\times$ normalizing $\mathfrak{A}$. If e denotes the $\mathfrak{o}_F$-period of $\mathfrak{A}$, then the number of unramified self-twists of π is n/e by Theorem 4.1. As mentioned in the introduction, the conductor $f(\pi)$ of π is given by

$$f(\pi) = n\Big(1 + \frac{m}{e}\Big).$$

Let π_i be two supercuspidal representations of $\mathrm{GL}_{n_i}(F)$ for $i = 1, 2$. There are three distinct possibilities: (i) π_1 and π_2 are unramified twists of each other, (ii) π_1 and π_2 are *completely distinct*, and (iii) π_1 and π_2 admit a *common approximation*. We do not get into defining these notions and refer to [Bushnell et al. 1998, § 6] instead. Suffice to say that when π_1 and π_2 admit a common approximation, there is a best common approximation and this is an object of the form $([\Lambda, m, 0, \gamma], l, \vartheta)$, where the stratum $[\Lambda, m, 0, \gamma]$ is determined by π_1 and π_2, $0 \le l < m$ is an integer, and ϑ is a character of a compact group attached to the data coming from π_1 and π_2.

Another ingredient in the conductor formula is an integer $\mathfrak{c}(\beta)$ associated to β. This comes from the "generalized discriminant", say $C(\beta)$, associated to the exact sequence

$$0 \longrightarrow E \longrightarrow \mathrm{End}_F(E) \xrightarrow{a_\beta} \mathrm{End}_F(E) \xrightarrow{s_\beta} E \longrightarrow 0,$$

where s_β is a tame corestriction relative to E/F [Bushnell and Kutzko 1993, § 1.3] and a_β is the adjoint map $x \mapsto \beta x - x\beta$. The constant $\mathfrak{c}(\beta)$ is defined such that

$$C(\beta) = q^{\mathfrak{c}(\beta)}.$$

Now we state the conductor formula of [Bushnell et al. 1998].

Theorem 5.1 (Bushnell, Henniart, and Kutzko). *For $i = 1, 2$, let π_i be an irreducible supercuspidal representation of $\mathrm{GL}_{n_i}(F)$. Define quantities m_i, e_i, β_i as above. Let $e = \mathrm{lcm}(e_1, e_2)$ and $m/e = \max\{m_1/e_1, m_2/e_2\}$.*

(1) *Suppose that $n_1 = n_2 = n$ and π_1 and π_2 are unramified twists of each other. Let $\beta = \beta_1$ and $d = [F[\beta] : F]$. Then*

$$f(\pi_1^\vee \times \pi_2) = n^2\Big(1 + \frac{\mathfrak{c}(\beta)}{d^2}\Big) - \deg L(s, \pi_1^\vee \times \pi_2).$$

(2) *Suppose that π_1 and π_2 are completely distinct. Then*

$$f(\pi_1^\vee \times \pi_2) = n_1 n_2\Big(1 + \frac{m}{e}\Big).$$

(3) *Suppose that π_2 is not equivalent to an unramified twist of π_1, but that π_1 and π_2 are not completely distinct. Let $([\Lambda, m, 0, \gamma], l, \vartheta)$ be a best common approximation to the π_i, and assume that the stratum $[\Lambda, m, l, \gamma]$ is simple. Put $d = [F[\gamma] : F]$. Then*

$$f(\pi_1^\vee \times \pi_2) = n_1 n_2\Big(1 + \frac{\mathfrak{c}(\gamma)}{d^2} + \frac{l}{de}\Big).$$

Remark. Observe that in (2) and (3), $\deg L(s, \pi_1^\vee \times \pi_2) = 0$.

6. Conductor of the Asai lift

Let E/F be a quadratic extension of p-adic fields. Let π be a supercuspidal representation of $\mathrm{GL}_n(E)$. Let $\rho_\pi : W_E \to \mathrm{GL}_n(\mathbb{C})$ be its Langlands parameter. Let $r(\rho_\pi) : W_F \to \mathrm{GL}_{n^2}(\mathbb{C})$ be the Asai lift of ρ_π. In this section, we compute the

Artin conductor of $r(\rho_\pi)$. Throughout this section, we assume that p is odd. For a representation τ of the Weil–Deligne group, let $f(\tau)$ denote its Artin conductor.

Our formula for the Asai lift is a consequence of the conductor formula for pairs of supercuspidal representations due to Bushnell, Henniart, and Kutzko [1998, Theorem 6.5]. Since

$$r(\rho_\pi)|_{W_E} \cong \rho_\pi \otimes \rho_\pi^\sigma,$$

it follows that

$$f(\rho_\pi \otimes \rho_\pi^\sigma) = \begin{cases} f(r(\rho_\pi)) & \text{if } E/F \text{ is unramified,} \\ f(r(\rho_\pi)) + f(r(\rho_\pi) \otimes \omega_{E/F}) - n^2 & \text{if } E/F \text{ is ramified.} \end{cases}$$

In the second case of the above, we have made use of the fact that E/F is tamely ramified, which is true since p is odd by our assumption. Since the formula of Bushnell, Henniart, and Kutzko [1998] computes the left hand side, in order to derive a formula for the Asai lift, it suffices to compute $f(r(\rho_\pi)) - f(r(\rho_\pi) \otimes \omega_{E/F})$.

Let

$$r(\rho_\pi) \cong \bigoplus\nolimits_i \rho_i$$

be the direct sum decomposition of $r(\rho_\pi)$ into irreducible representations. Now

$$r(\rho_\pi) \otimes \omega_{E/F} \cong \bigoplus\nolimits_i \rho_i \otimes \omega_{E/F},$$

and since the Artin conductor is additive, it follows that

$$f(r(\rho_\pi)) - f(r(\rho_\pi) \otimes \omega_{E/F}) = \bigoplus\nolimits_i [f(\rho_i) - f(\rho_i \otimes \omega_{E/F})].$$

We know that

$$f(\rho \otimes \chi) \leq \max\{f(\rho), \dim \rho \cdot f(\chi)\},$$

with equality in the above identity if $f(\rho) \neq \dim \rho \cdot f(\chi)$. Thus,

$$f(\rho_i \otimes \omega_{E/F}) = f(\rho_i)$$

unless ρ_i is a one-dimensional character with Artin conductor one, in which case $f(\rho_i \otimes \omega_{E/F})$ can be 0 or 1.

Observe that the contribution to

$$f(r(\rho_\pi)) - f(r(\rho_\pi) \otimes \omega_{E/F})$$

from tamely ramified characters ρ_i in $r(\rho_\pi)$ such that $\rho_i \otimes \omega_{E/F}$ is unramified is

$$\deg L(s, \pi \otimes \kappa, r),$$

whereas the contribution from unramified characters ρ_i in $r(\rho_\pi)$ such that $\rho_i \otimes \omega_{E/F}$ is tamely ramified is

$$-\deg L(s, \pi, r).$$

Therefore, it follows that

$$f(r(\rho_\pi)) - f(r(\rho_\pi) \otimes \omega_{E/F}) = \deg L(s, \pi \otimes \kappa, r) - \deg L(s, \pi, r).$$

Now making use of Theorem 5.1, we get the following conductor formula for the Asai lift.

Theorem 6.1. *Let E/F be a quadratic extension of p-adic fields, where p is odd, with ramification index $e(E/F)$. Let σ denote the nontrivial element of* $\mathrm{Gal}(E/F)$. *Let π be a supercuspidal representation of* $\mathrm{GL}_n(E)$. *Let e be the $\mathfrak{o}_E$-period of the principal $\mathfrak{o}_E$-order in $M_n(E)$ attached to π. Let $r(\pi)$ be its Asai lift to* $\mathrm{GL}_{n^2}(F)$ *and let $L(s, \pi, r)$ be the Asai L-function attached to π.*

(1) *Suppose $\pi^\vee$ and π^σ are unramified twists of each other. Then*

$$f(r(\pi)) = n^2\Big(1 + \frac{\mathfrak{c}(\beta)}{e(E/F)d^2}\Big) - \deg L(s, \pi, r).$$

(2) *Suppose $\pi^\vee$ and π^σ are completely distinct. Then*

$$f(r(\pi)) = n^2\Big(1 + \frac{m}{e(E/F)e}\Big).$$

(3) *Suppose that $\pi^\vee$ is not equivalent to an unramified twist of π^σ and that they are not completely distinct. Let $([\Lambda, m, 0, \gamma], l, \vartheta)$ be a best common approximation to $\pi^\vee$ and π^σ, and assume that the stratum $[\Lambda, m, l, \gamma]$ is simple. Set $d = [F[\gamma] : F]$. Then*

$$f(r(\pi)) = n^2\Big(1 + \frac{\mathfrak{c}(\gamma)}{e(E/F)d^2} + \frac{l}{e(E/F)de}\Big).$$

Remark. Together with Theorem 1.1, Theorem 6.1 gives an explicit conductor formula for the Asai lift. As in the case of Theorem 5.1, $\deg L(s, \pi, r) = 0$ in cases (2) and (3).

Acknowledgements

The authors thank Dipendra Prasad for several useful discussions. They thank the anonymous referee for carefully reading the manuscript and for making several useful suggestions. The second named author would like to acknowledge a grant (07IR001) from the Industrial Research and Consultancy Centre, Indian Institute of Technology Bombay, during the period of the present work.

References

[Anandavardhanan and Rajan 2005] U. K. Anandavardhanan and C. S. Rajan, "Distinguished representations, base change, and reducibility for unitary groups", *Int. Math. Res. Not.* **2005**:14 (2005), 841–854. MR 2006g:22013 Zbl 1070.22011

[Anandavardhanan et al. 2004] U. K. Anandavardhanan, A. C. Kable, and R. Tandon, "Distinguished representations and poles of twisted tensor L-functions", *Proc. Amer. Math. Soc.* **132**:10 (2004), 2875–2883. MR 2005g:11080 Zbl 1122.11033

[Bump and Ginzburg 1992] D. Bump and D. Ginzburg, "Symmetric square L-functions on $\mathrm{GL}(r)$", *Ann. of Math.* (2) **136**:1 (1992), 137–205. MR 93i:11058 Zbl 0753.11021

[Bushnell and Kutzko 1993] C. J. Bushnell and P. C. Kutzko, *The admissible dual of* $\mathrm{GL}(N)$ *via compact open subgroups*, Annals of Mathematics Studies **129**, Princeton University Press, 1993. MR 94h:22007 Zbl 0787.22016

[Bushnell et al. 1998] C. J. Bushnell, G. M. Henniart, and P. C. Kutzko, "Local Rankin–Selberg convolutions for GL_n: explicit conductor formula", *J. Amer. Math. Soc.* **11**:3 (1998), 703–730. MR 99h:22022 Zbl 0899.22017

[Flicker 1993] Y. Z. Flicker, "On zeroes of the twisted tensor L-function", *Math. Ann.* **297**:2 (1993), 199–219. MR 95c:11065 Zbl 0786.11030

[Goldberg 1994] D. Goldberg, "Some results on reducibility for unitary groups and local Asai L-functions", *J. Reine Angew. Math.* **448** (1994), 65–95. MR 95g:22031 Zbl 0815.11029

[Henniart 2010] G. M. Henniart, "Correspondance de Langlands et fonctions L des carrés extérieur et symétrique", *Int. Math. Res. Not.* **2010**:4 (2010), 633–673. MR 2011c:22028 Zbl 1184.22009

[Jacquet and Shalika 1990] H. Jacquet and J. A. Shalika, "Exterior square L-functions", pp. 143–226 in *Automorphic forms, Shimura varieties, and L-functions* (Ann Arbor, MI, 1988), vol. 2, edited by L. Clozel and J. S. Milne, Perspectives in Mathematics **11**, Academic Press, Boston, 1990. MR 91g:11050 Zbl 0695.10025

[Jacquet et al. 1983] H. Jacquet, I. I. Piatetskii-Shapiro, and J. A. Shalika, "Rankin–Selberg convolutions", *Amer. J. Math.* **105**:2 (1983), 367–464. MR 85g:11044 Zbl 0525.22018

[Jiang et al. 2008] D. Jiang, C. Nien, and Y. Qin, "Local Shalika models and functoriality", *Manuscripta Math.* **127**:2 (2008), 187–217. MR 2010b:11057 Zbl 1167.11021

[Kable 2004] A. C. Kable, "Asai L-functions and Jacquet's conjecture", *Amer. J. Math.* **126**:4 (2004), 789–820. MR 2005g:11083 Zbl 1061.11023

[Kewat and Raghunathan 2012] P. K. Kewat and R. Raghunathan, "On the local and global exterior square L-functions of GL_n", *Math. Res. Lett.* **19**:4 (2012), 785–804. MR 3008415 Zbl 06165853

[Matringe 2011] N. Matringe, "Distinguished generic representations of $\mathrm{GL}(n)$ over p-adic fields", *Int. Math. Res. Not.* **2011**:1 (2011), 74–95. MR 2012f:22032 Zbl 1223.22015

[Murty and Prasad 2000] V. K. Murty and D. Prasad, "Tate cycles on a product of two Hilbert modular surfaces", *J. Number Theory* **80**:1 (2000), 25–43. MR 2000m:14028 Zbl 0955.14016

[Prasad 1999] D. Prasad, "Some remarks on representations of a division algebra and of the Galois group of a local field", *J. Number Theory* **74**:1 (1999), 73–97. MR 99m:11138 Zbl 0931.11054

[Shahidi 1984] F. Shahidi, "Fourier transforms of intertwining operators and Plancherel measures for $\mathrm{GL}(n)$", *Amer. J. Math.* **106**:1 (1984), 67–111. MR 86b:22031 Zbl 0567.22008

[Shahidi 1990] F. Shahidi, "A proof of Langlands' conjecture on Plancherel measures; complementary series for p-adic groups", *Ann. of Math.* (2) **132**:2 (1990), 273–330. MR 91m:11095 Zbl 0780.22005

[Shahidi 1992] F. Shahidi, "Twisted endoscopy and reducibility of induced representations for p-adic groups", *Duke Math. J.* **66**:1 (1992), 1–41. MR 93b:22034 Zbl 0785.22022

Received March 31, 2014. Revised November 9, 2014.

U. K. ANANDAVARDHANAN
DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY BOMBAY
MUMBAI 400076
INDIA
anand@math.iitb.ac.in

AMIYA KUMAR MONDAL
DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY BOMBAY
MUMBAI 400076
INDIA
amiya@math.iitb.ac.in

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.19

TORUS ACTIONS AND TENSOR PRODUCTS OF INTERSECTION COHOMOLOGY

ASILATA BAPAT

Given certain intersection cohomology sheaves on a projective variety with a torus action, we relate the cohomology groups of their tensor product to the cohomology groups of the individual sheaves. We also prove a similar result in the case of equivariant cohomology.

1. Introduction

Let X be a smooth complex projective variety together with an action of a complex algebraic torus T with isolated fixed points. We fix a regular algebraic one-parameter subgroup $\lambda\colon \mathbb{C}^* \to T$, which means that the set of λ-fixed points on X equals the set of T-fixed points on X (denoted X^T). Consider the Białynicki-Birula decomposition [1973] of X: for each $w \in X^T$ define the *plus* and *minus* cells to be respectively

$$U_w = U_w^+ = \{x \in X \mid \lim_{t\to 0} \lambda(t)\cdot x = w\}, \quad t \in \mathbb{C}^*, \text{ and}$$
$$U_w^- = \{x \in X \mid \lim_{t\to \infty} \lambda(t)\cdot x = w\}, \quad t \in \mathbb{C}^*.$$

Each plus or minus cell is a λ-stable affine space, and hence the decompositions $X = \coprod_{w\in X^T} U_w$ and $X = \coprod_{w\in X^T} U_w^-$ are cell decompositions. For the purposes of this paper, we make the following additional assumptions on the T-action on X.

Assumption 1.1. The cell decompositions $X = \coprod_{w\in X^T} U_w$ and $X = \coprod_{w\in X^T} U_w^-$ are algebraic stratifications of X. In particular, the closure of every plus cell is a union of plus cells, and analogously for minus cells.

Assumption 1.2. For each $w \in X^T$, there is a one-parameter subgroup $\lambda_w\colon \mathbb{C}^* \to T$ and a neighborhood V_w of w such that $\lim_{t\to 0} \lambda_w(t)\cdot v = w$ for every $v \in V_w$ and $t \in \mathbb{C}^*$.

In this paper, we use the words *sheaf* and *complex of sheaves* interchangeably to mean an object in $D^b_{c,\mathrm{BB}}(X, \mathbb{C})$, the bounded derived category of sheaves of $\mathbb{C}$-vector spaces on X that are constructible with respect to the Białynicki-Birula

MSC2010: 14F05, 14F43, 14L30, 55N33.
Keywords: intersection cohomology, torus action.

stratification. (Here we make use of Assumption 1.1.) Moreover all functors are derived, so for ease of notation we omit the decorations R and L.

For each $w \in X^T$, let IC_w denote the intersection cohomology sheaf on the closure of the cell U_w, extended by zero to all of X. The main theorem of the paper describes the cohomology of the tensor products of a collection of IC_w, in terms of the tensor products of the cohomologies of the individual IC_w.

Main result. Let $\Delta\colon X \to X^m$ be the diagonal embedding. Consider any sheaves $\mathcal{F}_1, \dots, \mathcal{F}_m$ in $D^b_{c,\mathrm{BB}}(X, \mathbb{C})$. Then their (derived) tensor product is also a sheaf in $D^b_{c,\mathrm{BB}}(X, \mathbb{C})$, and will be denoted by $\mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_m$. Recall that

$$\mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_m = \Delta^{-1}(\mathcal{F}_1 \boxtimes \cdots \boxtimes \mathcal{F}_m).$$

For any sheaf $\mathcal{F}$, its cohomology $H^\bullet(\mathcal{F}) = H^\bullet(X, \mathcal{F})$ is a graded vector space. There is a natural cup product $\cup\colon H^\bullet(\mathcal{F}_1) \otimes \cdots \otimes H^\bullet(\mathcal{F}_m) \to H^\bullet(\mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_m)$, defined on page 22.

Let $\underline{\mathbb{C}}$ denote the constant sheaf on X. For any sheaf $\mathcal{F}$, its cohomology $H^\bullet(\mathcal{F})$ is naturally a (graded) left and right module over the (graded) ring $H(X) = H^\bullet(X, \underline{\mathbb{C}})$, as follows:

$$\cup\colon H(X) \otimes H^\bullet(\mathcal{F}) \to H^\bullet(\underline{\mathbb{C}} \otimes \mathcal{F}) \xrightarrow{\cong} H^\bullet(\mathcal{F}),$$
$$\cup\colon H^\bullet(\mathcal{F}) \otimes H(X) \to H^\bullet(\mathcal{F} \otimes \underline{\mathbb{C}}) \xrightarrow{\cong} H^\bullet(\mathcal{F}).$$

Moreover, the cup product descends to a morphism

$$H^\bullet(\mathcal{F}_1) \underset{H(X)}{\otimes} \cdots \underset{H(X)}{\otimes} H^\bullet(\mathcal{F}_m) \to H^\bullet(\mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_m).$$

Theorem 1.3. *Let $(p_1, \dots, p_m)$ be an m-tuple of T-fixed points of X, and suppose that Assumptions 1.1 and 1.2 hold. Then the cup product map*

$$H^\bullet(\mathrm{IC}_{p_1}) \underset{H(X)}{\otimes} \cdots \underset{H(X)}{\otimes} H^\bullet(\mathrm{IC}_{p_m}) \to H^\bullet(\mathrm{IC}_{p_1} \otimes \cdots \otimes \mathrm{IC}_{p_m}) \tag{1-1}$$

is an isomorphism.

As X is a T-space, each IC sheaf IC_{p_j} carries a canonical T-equivariant structure, and so does the tensor product $\mathrm{IC}_{p_1} \otimes \cdots \otimes \mathrm{IC}_{p_m}$. Let $H_T(X) = H^\bullet_T(X, \mathbb{C})$ be the T-equivariant cohomology of X. For any T-equivariant sheaf $\mathcal{F}$ on X, its T-equivariant cohomology $H^\bullet_T(\mathcal{F}) = H^\bullet_T(X, \mathcal{F})$ is a graded $H_T(X)$-module. As before, there is a cup product map for T-equivariant cohomology, which factors through $H_T(X)$.

Theorem 1.4. *Under Assumptions 1.1 and 1.2, the cup product map*

$$H^\bullet_T(\mathrm{IC}_{p_1}) \underset{H_T(X)}{\otimes} \cdots \underset{H_T(X)}{\otimes} H^\bullet_T(\mathrm{IC}_{p_m}) \to H^\bullet_T(\mathrm{IC}_{p_1} \otimes \cdots \otimes \mathrm{IC}_{p_m})$$

is an isomorphism.

Remark 1.5. Even though our results are stated using IC sheaves, it is possible that they generalize to parity sheaves (defined and discussed by Juteau, Mautner, and Williamson in [Juteau et al. 2014]). Our results and proof methods are similar to the main theorem from [Ginzburg 1991]. Achar and Rider [2014, Theorem 4.1] prove a version of Ginzburg's theorem for parity sheaves on generalized flag varieties of a Kac–Moody group. Similar generalizations may work in our case as well.

2. Setup

The Białynicki-Birula stratification. One can find (see, e.g., [Sumihiro 1974] or [Kambayashi 1966]) a T-equivariant projective embedding of X into some $\mathbb{P}^N$, such that the action of T on $\mathbb{P}^N$ is linear. Consider the following standard Morse–Bott function on $\mathbb{P}^N$:

$$[z_0 : \cdots : z_N] \mapsto \frac{\sum_{i=0}^N c_i |z_i|^2}{\sum_{i=0}^N |z_i|^2},$$

where c_i are the weights of the λ-action on $\mathbb{P}^N$. The critical sets of this function are precisely the T-fixed points on $\mathbb{P}^N$. The Morse–Bott cells of this function are locally closed algebraic subvarieties of $\mathbb{P}^N$. Since X has isolated T-fixed points, one can show that the composition $f : X \to \mathbb{P}^N \to \mathbb{R}$ is a Morse function with critical set X^T (see, e.g., [Audin 2004]). Each cell of the Morse decomposition under f is a preimage of a Morse–Bott cell of $\mathbb{P}^N$. Hence it is a locally closed algebraic subvariety of X. Moreover, each cell of the Morse decomposition is known to be a union of Białynicki-Birula plus cells. A discussion of this may also be found in [Chriss and Ginzburg 1997, Section 2.4].

The collection of fixed points of the λ-action carries a partial order, where $v < w$ if $U_v \subset \overline{U_w}$. By the previous discussion, we see that $v < w$ if and only if $f(v) < f(w)$. Fix a weakly increasing enumeration $\{0, 1, \ldots, N\}$ of the points of X^T (sometimes denoted $\{w_0, \ldots, w_N\}$), and set $X_n = \bigcup_{i \leq n} U_i$. Since the closure of every plus cell is a union of plus cells, it follows from the previous discussion that each X_n is a closed subvariety of X.

Similarly, set $X_n^- = \bigcup_{i \geq n} U_i^-$. By using the Morse function $(-f)$ instead of f, we see that each X_n^- is a closed subvariety of X. Hence we obtain two increasing filtrations of X by closed subvarieties: $X_0 \subset \cdots \subset X_N = X$ and $X_N^- \subset \cdots \subset X_0^- = X$.

We have the following inclusions:

$$X_n \overset{i_n}{\hookrightarrow} X, \quad X_{n-1} \overset{v}{\hookrightarrow} X_n \overset{u}{\hookleftarrow} U_n.$$

For any point $p \in X_n^-$, we have $f(w_n) \leq f(p)$, with equality only if $p \in X^T$. For any point $p \in X_n$, we have $f(p) \leq f(w_n)$, with equality only if $p \in X^T$. Hence if $p \in X_n^- \cap X_n$, then $f(p) = f(w_n)$, and $p \in X^T$. But $X_n^- \cap X_n \cap X^T = \{w_n\}$, and it follows that $p = w_n$. Hence for every n, the subvarieties X_n^- and X_n intersect transversally in the single point w_n.

Let $c_n \in H^\bullet(X)$ be the Poincaré dual to the homology class of X_n^-. As a vector space, $H^\bullet(X)$ is generated by the collection $\{c_n\}$. Finally, fix an m-tuple $(p_1, \ldots, p_m)$ of T-fixed points of X, and set $L_{j,n} = i_n^{-1}\,\mathrm{IC}_{p_j}$ for each j and n.

The cup product in cohomology. Let $\pi\colon X \to \mathrm{pt}$ be the unique morphism to a point. For any sheaf $\mathscr{F}$ on X, its cohomology $H^\bullet(\mathscr{F})$ is a graded vector space, and may be thought of as $\pi_*\mathscr{F}$. We use this to define the cup product map.

Recall that the functors (π^{-1}, π_*) form an adjoint pair, which has a counit $\pi^{-1} \circ \pi_* \to \mathrm{id}$. Let $\mathscr{F}_1, \ldots, \mathscr{F}_m$ be sheaves on X. Tensoring the counit maps together, we have a map

$$\pi^{-1} \circ \pi_*(\mathscr{F}_1) \otimes \cdots \otimes \pi^{-1} \circ \pi_*(\mathscr{F}_m) \to \mathscr{F}_1 \otimes \cdots \otimes \mathscr{F}_m\,.$$

The left hand side is canonically isomorphic to $\pi^{-1}(\pi_*\mathscr{F}_1 \otimes \cdots \otimes \pi_*\mathscr{F}_m)$. Using the (π^{-1}, π_*) adjunction once more, we obtain the *cup product*:

$$\cup\colon \pi_*\mathscr{F}_1 \otimes \cdots \otimes \pi_*\mathscr{F}_m \to \pi_*(\mathscr{F}_1 \otimes \cdots \otimes \mathscr{F}_m).$$

The cup product gives each $H^\bullet(\mathscr{F}_i)$ the structure of a left and right module over $H(X)$. This module structure induces the following map, also called the cup product:

$$H^\bullet(\mathscr{F}_1) \underset{H(X)}{\otimes} \cdots \underset{H(X)}{\otimes} H^\bullet(\mathscr{F}_m) \to H^\bullet(\mathscr{F}_1 \otimes \cdots \otimes \mathscr{F}_m).$$

Proposition 2.1. *For every n, the cup product map*

$$H^\bullet(L_{1,n}) \underset{H(X)}{\otimes} \cdots \underset{H(X)}{\otimes} H^\bullet(L_{m,n}) \to H^\bullet(L_{1,n} \otimes \cdots \otimes L_{m,n}) \tag{2-1}$$

is an isomorphism.

When $X_n = X$, we have $L_{j,n} = \mathrm{IC}_{p_j}$ for each j. Hence Theorem 1.3 follows from this proposition, and we now focus on proving the proposition.

3. Proof of the isomorphism

We prove Proposition 2.1 by induction on the nth filtered piece of $X_0 \subset \cdots \subset X_N$. In the base case of $n = 0$, the space X_0 is zero-dimensional. Hence each sheaf $L_{j,0}$ is isomorphic to its cohomology. In this case the cup product map (2-1) reduces to the identity map, which is an isomorphism.

Now we prove the induction step on the filtered piece X_n. We mainly use the following distinguished triangles:

$$u_!u^{-1}L_{j,n} \to L_{j,n} \to v_*v^{-1}L_{j,n}, \tag{3-1}$$

$$v_!v^!L_{j,n} \to L_{j,n} \to u_*u^{-1}L_{j,n}. \tag{3-2}$$

After taking cohomology, each of the above distinguished triangles produces a long exact sequence. In our case, all connecting homomorphisms of these long exact sequences vanish (see, e.g., [Soergel 1990, Lemma 20] and [Ginzburg 1991, Proposition 3.2]).

For brevity, we will use the following notation through the remainder of the paper.

$$\begin{aligned} M_{m,n} &= L_{2,n} \otimes \cdots \otimes L_{m,n}, \\ A_{m,n} &= H^\bullet(L_{2,n}) \underset{H(X)}{\otimes} \cdots \underset{H(X)}{\otimes} H^\bullet(L_{m,n}), \\ B_{m,n} &= H^\bullet(u_*u^{-1}L_{2,n}) \underset{H(X)}{\otimes} \cdots \underset{H(X)}{\otimes} H^\bullet(u_*u^{-1}L_{m,n}). \end{aligned} \tag{3-3}$$

The following two lemmas prove the proposition on the open part U_n in X_n.

Lemma 3.1. *Let $\mathcal{F}$ and $\mathcal{G}$ be any complexes of sheaves on U_n with locally constant cohomology sheaves. Then the cup product map*

$$\cup\colon H^\bullet(u_!\mathcal{F}) \otimes H^\bullet(u_*\mathcal{G}) \to H^\bullet(u_!\mathcal{F} \otimes u_*\mathcal{G})$$

is an isomorphism. Since $\cup$ factors through the surjection

$$H^\bullet(u_!\mathcal{F}) \otimes H^\bullet(u_*\mathcal{G}) \twoheadrightarrow H^\bullet(u_!\mathcal{F}) \underset{H(X)}{\otimes} H^\bullet(u_*\mathcal{G}),$$

the induced cup product

$$\cup\colon H^\bullet(u_!\mathcal{F}) \underset{H(X)}{\otimes} H^\bullet(u_*\mathcal{G}) \to H^\bullet(u_!\mathcal{F} \otimes u_*\mathcal{G})$$

is also an isomorphism.

Proof. Consider the following commutative diagram, where π is the projection to a point.

$$\begin{tikzcd} U_n \arrow[r, "u"] \arrow[dr, "p=\pi\circ u"'] & X_n \arrow[d, "\pi"] \\ & \mathrm{pt} \end{tikzcd}$$

Recall that if A and B are any two complexes on X, then the cup product is induced by adjunction from the natural map

$$\pi^{-1}(\pi_* A \otimes \pi_* B) \cong \pi^{-1}\pi_* A \otimes \pi^{-1}\pi_* B \to A \otimes B,$$

which may be broken up as follows:

$$\pi^{-1}\pi_* A \otimes \pi^{-1}\pi_* B \to A \otimes \pi^{-1}\pi_* B \to A \otimes B.$$

Therefore the cup product map may be broken up as follows:

$$\pi_* A \otimes \pi_* B \to \pi_*(A \otimes \pi^{-1}\pi_* B) \to \pi_*(A \otimes B).$$

In our case, this becomes the following sequence of maps:

$$\pi_* u_! \mathscr{F} \otimes \pi_* u_* \mathscr{G} \xrightarrow{\mu_1} \pi_*(u_! \mathscr{F} \otimes \pi^{-1}\pi_* u_* \mathscr{G}) \xrightarrow{\mu_2} \pi_*(u_! \mathscr{F} \otimes u_* \mathscr{G}).$$

Since π is a proper map, we know that $\pi_* \cong \pi_!$, and hence μ_1 is an isomorphism by the projection formula. It remains to show that μ_2 is an isomorphism.

The pair of adjoint functors (π^{-1}, π_*) gives the counit morphism $p^{-1}p_* \mathscr{G} \to u^{-1}u_* \mathscr{G}$. The key observation is that this map is an isomorphism, because $\mathscr{G}$ is a direct sum of its cohomology sheaves on the affine space U_n. Now consider the following commutative diagram.

$$\begin{tikzcd} u_! \mathscr{F} \otimes \pi^{-1}\pi_* u_* \mathscr{G} \arrow[r, "\cong", "\text{(proj.)}"'] \arrow[d, "\mu_2"', "\text{(counit)}"] & u_!(\mathscr{F} \otimes p^{-1}p_* \mathscr{G}) \arrow[d, "\cong"', "\text{(counit)}"] \\ u_! \mathscr{F} \otimes u_* \mathscr{G} \arrow[r, "\cong", "\text{(proj.)}"'] & u_!(\mathscr{F} \otimes u^{-1}u_* \mathscr{G}) \end{tikzcd} \tag{3-4}$$

The map μ_2 is obtained by applying the functor π_* to the left vertical map in (3-4) above. The diagram shows that this map is an isomorphism, and hence μ_2 is also an isomorphism. □

Lemma 3.2. *The cup product map induces an isomorphism*

$$H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} B_{m,n} \xrightarrow{\cong} H_c^\bullet(u^{-1}(L_{1,n} \otimes M_{m,n})).$$

Proof. Using Lemma 3.1 with complexes of sheaves $\mathscr{F} = u^{-1}L_{1,n}$ and $\mathscr{G} = u^{-1}L_{2,n}$, we obtain an isomorphism

$$H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} H^\bullet(u_* u^{-1} L_{2,n}) \xrightarrow{\cong} H^\bullet(u_! u^{-1} L_{1,n} \otimes u_* u^{-1} L_{2,n}).$$

Moreover, $u^{-1}u_* u^{-1} L_{2,n} \cong u^{-1}L_{2,n}$. Using this fact and the projection formula,

$$\begin{aligned} H^\bullet(u_! u^{-1} L_{1,n} \otimes u_* u^{-1} L_{2,n}) &\cong H^\bullet(u_!(u^{-1} L_{1,n} \otimes u^{-1} u_* u^{-1} L_{2,n})) \\ &\cong H^\bullet(u_! u^{-1}(L_{1,n} \otimes L_{2,n})). \end{aligned}$$

All together, we get an isomorphism

$$H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} H^\bullet(u_* u^{-1} L_{2,n}) \xrightarrow{\cong} H^\bullet(u_! u^{-1}(L_{1,n} \otimes L_{2,n})),$$

which can be written in our previously introduced notation as

$$H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} B_{2,n} \xrightarrow{\cong} H^\bullet(u_! u^{-1}(L_{1,n} \otimes M_{2,n})).$$

Now we can successively tensor the above map over $H(X)$ with the spaces $H^\bullet(u_* u^{-1} L_{i,n})$, with i ranging from 3 to m. Each time, we apply Lemma 3.1 for $\mathscr{F} = u^{-1}(L_{1,n} \otimes M_{i-1,n})$ and $\mathscr{G} = u^{-1} L_{i,n}$ and use the argument above. Ultimately this construction yields

$$\begin{aligned} H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} B_{m,n} &\xrightarrow{\cong} H^\bullet(u_! u^{-1}(L_{1,n} \otimes M_{m-1,n})) \underset{H(X)}{\otimes} H^\bullet(u_* u^{-1} L_{m,n}) \\ &\xrightarrow{\cong} H^\bullet(u_!(u^{-1}(L_{1,n} \otimes M_{m,n}))) \\ &\cong H_c^\bullet(u^{-1}(L_{1,n} \otimes M_{m,n})). \end{aligned}$$ □

The next lemma is a refinement of a standard cohomology exact sequence to our particular case.

Lemma 3.3. *There is an exact sequence*

$$H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} B_{m,n} \to H^\bullet(L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \to H^\bullet(v_* v^{-1} L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \to 0.$$

Proof. Consider the distinguished triangle (3-1) for the sheaf $L_{1,n}$. Taking cohomology and applying the functor $(-) \underset{H(X)}{\otimes} A_{m,n}$, we obtain the right-exact sequence

$$H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \xrightarrow{f} H^\bullet(L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \xrightarrow{g} H^\bullet(v_* v^{-1} L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \to 0.$$

Using the distinguished triangles (3-2) for each of the sheaves $L_{j,n}$ for $j \geq 2$, we have surjective morphisms

$$H^\bullet(L_{j,n}) \twoheadrightarrow H^\bullet(u_* u^{-1} L_{j,n}).$$

Taking the tensor product of all of these along with $H^\bullet(u_! u^{-1} L_{1,n})$, we obtain a surjective morphism

$$H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \overset{h}{\twoheadrightarrow} H^\bullet(u_! u^{-1} L_{1,n}) \underset{H(X)}{\otimes} B_{m,n}.$$

We now show that the map f factors through the map h, by showing that $f(\ker h) = 0$. Since all boundary maps in the cohomology long exact sequence of the triangles (3-2) vanish, the following set generates $\ker h$:

$$\{a_1 \otimes a_2 \otimes \cdots \otimes a_n \mid a_j \in H^\bullet(v_* v^! L_{j,n}) \text{ for some } 2 \leq j \leq m\}.$$

Consider any element $a_1 \otimes a_2 \otimes \cdots \otimes a_n \in \ker h$. Suppose that $a_j \in H^\bullet(v_* v^! L_{j,n})$. Recall the commutative diagram (3.8a) from [Ginzburg 1991], reproduced below.

$$\begin{tikzcd}
H^\bullet(v_* v^! L_{j,n}) \arrow[r, hook] & H^\bullet(L_{j,n}) \arrow[r, two heads] \arrow[d, "c_n"'] & H^\bullet(u^{-1} L_{j,n}) \arrow[d, "c_n"', "\cong"] \\
& H^\bullet(L_{j,n}) & H_c^\bullet(u^{-1} L_{j,n}) \arrow[l, hook']
\end{tikzcd}$$

From this diagram it follows that $c_n a_j = 0$, and that $a_1 \in c_n H^\bullet(L_{1,n})$. Since all tensor products are over $H(X)$, the image of $h(a_1 \otimes \cdots \otimes a_n)$ under f must be zero. Therefore f factors through h, and we obtain the desired short exact sequence. □

Finally, we use the induction hypothesis to tackle the right side of the right-exact sequence from the previous lemma.

Lemma 3.4. *The cup product map induces an isomorphism*

$$H^\bullet(v_* v^{-1} L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \xrightarrow{\cong} H^\bullet(L_{1,n-1} \otimes M_{m,n-1}).$$

Proof of lemma. The cup product map on the left hand side is the following composition:

$$H^\bullet(v_* v^{-1} L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \to H^\bullet(v_* v^{-1} L_{1,n}) \underset{H(X)}{\otimes} H^\bullet(M_{m,n}) \to H^\bullet(v_* v^{-1} L_{1,n} \otimes M_{m,n}),$$

where the first map is the cup product on the last $(m-1)$ factors, and the second map is the cup product of the first factor with the rest. The projection formula also shows that

$$H^\bullet(v_* v^{-1} L_{1,n} \otimes M_{m,n}) \cong H^\bullet(v^{-1} L_{1,n} \otimes v^{-1} M_{m,n}) \cong H^\bullet(L_{1,n-1} \otimes M_{m,n-1}).$$

By induction on m, we may assume that the cup product $A_{m,n} \to H^\bullet(M_{m,n})$ is an isomorphism, and hence the first map above is an isomorphism. It remains to show that the following map is an isomorphism:

$$H^\bullet(v_* v^{-1} L_{1,n}) \underset{H(X)}{\otimes} H^\bullet(M_{m,n}) \to H^\bullet(v_* v^{-1} L_{1,n} \otimes M_{m,n})$$

Since $L_{1,n-1}$ is supported on X_{n-1}, the element $c_n \in H$ acts on $H^\bullet(v_* L_{1,n-1})$ by zero. Recall from [op. cit.] that the cokernel of c_n on $H^\bullet(M_{m,n})$ is just $H^\bullet(M_{m,n-1})$. Hence

$$H^\bullet(v_* v^{-1} L_{1,n}) \underset{H(X)}{\otimes} H^\bullet(M_{m,n}) \cong H^\bullet(L_{1,n-1}) \underset{H(X)}{\otimes} H^\bullet(M_{m,n-1}).$$

Therefore, the map above can be rewritten as the cup product map

$$H^\bullet(L_{1,n-1}) \underset{H(X)}{\otimes} H^\bullet(M_{m,n-1}) \to H^\bullet(L_{1,n-1} \otimes M_{m,n-1}),$$

which is an isomorphism by the induction hypothesis. □

We now apply Saito's theory [1990; 1988] of mixed Hodge modules to obtain another short exact sequence, as follows. Every IC-sheaf has the additional structure of a pure mixed Hodge module, which induces a mixed Hodge structure on tensor products of the $L_{i,n}$.

Lemma 3.5. (i) *The cohomology $H^\bullet(L_{1,n} \otimes M_{m,n})$ is pure.*

(ii) *There is a short exact sequence*

$$0 \to H_c^\bullet(u^{-1}(L_{1,n} \otimes M_{m,n})) \to H^\bullet(L_{1,n} \otimes M_{m,n}) \to H^\bullet(L_{1,n-1} \otimes M_{m,n-1}) \to 0.$$

Proof. The proof is by induction on n. When $n = 0$, we have $X_{-1} = \varnothing$ and $U = X_0$. The open inclusion u is the identity map, and the closed inclusion v is the zero map, hence (ii) is clear in the base case.

The set X_0 consists of a single T-fixed point of X. Call this point w. By Assumption 1.2, there exists a neighborhood V_w of w and a one-parameter subgroup $\lambda_w : \mathbb{C}^* \to T$ that contracts V_w to w. Let i_w denote the inclusion of $\{w\}$ into the corresponding V_w. Let j_w denote the inclusion of V_w into X. By applying [Springer 1984, Corollary 1] or [Braden 2003, Lemma 6] to the sheaves $j_w^{-1}\,\mathrm{IC}_{p_i}$ for each i, we see that

$$H^\bullet(V_w, j_w^{-1}\,\mathrm{IC}_{p_i}) \cong H^\bullet(i_w^{-1} j_w^{-1}\,\mathrm{IC}_{p_i}) = H^\bullet(L_{i,0}).$$

The functor $H^\bullet(V_w, j_w^{-1}(-))$ weakly increases weights; on the other hand, the functor $H^\bullet(i_w^{-1} j_w^{-1}(-))$ weakly decreases weights. Hence $H^\bullet(L_{i,0})$ is pure for each i. Taking the tensor product, we see that $H^\bullet(L_{1,0}) \otimes \cdots \otimes H^\bullet(L_{m,0})$ is pure. Since w is a single point, we can naturally make the following identification:

$$H^\bullet(L_{1,0}) \otimes \cdots \otimes H^\bullet(L_{m,0}) \cong H^\bullet(L_{1,0} \otimes \cdots \otimes L_{m,0}) = H^\bullet(L_{1,0} \otimes M_{m,0}).$$

Hence $H^\bullet(L_{1,0} \otimes M_{m,0})$ is pure, and (i) is proved in the base case. A similar argument has been used in [Ginzburg 1991, Lemma 3.5].

For the induction step, consider the distinguished triangle (3-1) for $L_{1,n}$. Apply the functor $(- \otimes L_{2,n} \otimes \cdots \otimes L_{m,n})$, which may be written as $(- \otimes M_{m,n})$ in the notation of (3-3). This yields the following distinguished triangle:

$$u_! u^{-1} L_{1,n} \otimes M_{m,n} \to L_{1,n} \otimes M_{m,n} \to v_* v^{-1} L_{1,n} \otimes M_{m,n}.$$

By a repeated application of the projection formula, we may write the first term of this triangle as

$$u_! u^{-1} L_{1,n} \otimes M_{m,n} \cong u_!(u^{-1} L_{1,n} \otimes \cdots \otimes u^{-1} L_{m,n}) = u_! u^{-1}(L_{1,n} \otimes M_{m,n}),$$

and the third term of this triangle as

$$v_* v^{-1} L_{1,n} \otimes M_{m,n} \cong v_*(v^{-1} L_{1,n} \otimes \cdots \otimes v^{-1} L_{m,n}) = v_*(L_{1,n-1} \otimes M_{m,n-1}).$$

Taking cohomology, we obtain the following long exact sequence:

$$\cdots \to H_c^\bullet(u^{-1}(L_{1,n} \otimes M_{m,n})) \to H^\bullet(L_{1,n} \otimes M_{m,n}) \\ \to H^\bullet(L_{1,n-1} \otimes M_{m,n-1}) \to \cdots .$$

The term $H^\bullet(L_{1,n-1} \otimes M_{m,n-1})$ is pure by the induction hypothesis.

From Lemma 3.2, we know that

$$H_c^\bullet(u^{-1}(L_{1,n} \otimes M_{m,n})) \cong H_c^\bullet(u^{-1}L_{1,n}) \underset{H(X)}{\otimes} H^\bullet(u^{-1}L_{2,n}) \underset{H(X)}{\otimes} \cdots \underset{H(X)}{\otimes} H^\bullet(u^{-1}L_{m,n}).$$

Recall that U_n is the Białynicki-Birula plus cell for the fixed point w_n. Hence the λ-action contracts U_n to w_n. By [Springer 1984, Corollary 2], we know that $H_c^\bullet(u^{-1}L_{1,n})$ is isomorphic to the costalk of $u^{-1}L_{1,n}$ at w_n, which is isomorphic to a shift of the stalk of IC_{p_1} at w_n. For any $i > 1$, we know by [Springer 1984, Corollary 1] that $H^\bullet(u^{-1}L_{i,n})$ is isomorphic to the stalk of $u^{-1}L_{i,n}$ at w_n, which is equal to the stalk of IC_{p_i} at w_n. By using Assumption 1.2 and the argument used earlier in this proof, we know that the stalk of each IC_{p_i} at any T-fixed point is pure, and hence the spaces $H_c^\bullet(u^{-1}L_{1,n})$ as well as $H^\bullet(u^{-1}L_{i,n})$ for $i > 1$ are all pure. Therefore the tensor product $H_c^\bullet(u^{-1}(L_{1,n} \otimes M_{m,n}))$ is pure.

Since the terms on either side of the long exact sequence are pure, the connecting homomorphisms are zero, and hence $H^\bullet(L_{1,n} \otimes M_{m,n})$ is also pure. This argument completes the induction step, and hence completes the proof. □

Putting together the exact sequences from Lemmas 3.3 and 3.5, we obtain the following commutative diagram, where the vertical maps are induced by cup products. In particular, the middle map b is just the map from Proposition 2.1.

$$\begin{tikzcd} H^\bullet(u_!u^{-1}L_{1,n}) \underset{H(X)}{\otimes} B_{m,n} \arrow[r] \arrow[d, "a"] & H^\bullet(L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \arrow[r, two heads] \arrow[d, "b"] & H^\bullet(v_*v^{-1}L_{1,n}) \underset{H(X)}{\otimes} A_{m,n} \arrow[d, "c"] \\ H_c^\bullet(u^{-1}(L_{1,n} \otimes M_{m,n})) \arrow[r, hook] & H^\bullet(L_{1,n} \otimes M_{m,n}) \arrow[r, two heads] & H^\bullet(L_{1,n-1} \otimes M_{m,n-1}) \end{tikzcd} \tag{3-5}$$

The leftmost map a is an isomorphism by Lemma 3.2. The rightmost map c is an isomorphism by Lemma 3.4. By the snake lemma, the middle map b is an isomorphism as well, and Proposition 2.1 is proved.

4. Computation of equivariant cohomology

Consider a smooth complex projective variety X with the same assumptions as in Section 1. The goal of this section is to prove Theorem 1.4.

First, recall some constructions in equivariant cohomology, following [Bernstein and Lunts 1994] and [Goresky et al. 1998]. Fix a universal principal T-bundle

$ET \to BT$, where ET is the direct limit over m of algebraic approximations ET_m and analogously for BT and BT_m. Consider the following diagram, where the map p is the second projection, and the map q is the quotient by the diagonal T-action.

$$
\begin{tikzcd}
& ET \times X \arrow[dl, "p"'] \arrow[dr, "q"] & \\
X & & ET \times_T X
\end{tikzcd}
$$

Since each stratum U_n is a locally closed T-invariant affine subvariety of X, the trivial local system on U_n gives rise to a canonically defined sheaf $\overline{\mathrm{IC}_n}$ on $ET \times_T X$ and a canonical isomorphism $\beta\colon\ p^{-1}\,\mathrm{IC}_n \xrightarrow{\cong} q^{-1}\overline{\mathrm{IC}_n}$ (see, e.g., [Bernstein and Lunts 1994]). The triple $(\mathrm{IC}_n, \overline{\mathrm{IC}_n}, \beta)$ is called the equivariant IC sheaf corresponding to U_n.

Equivariant homology and cohomology. For a variety Y equipped with a T-action, the cohomology of $ET \times_T Y$ is called the *equivariant cohomology* of Y, and is denoted by $H^\bullet_T(Y)$. In particular, since $ET \times_T \mathrm{pt} \cong BT$, we have $H^\bullet_T(\mathrm{pt}) \cong H^\bullet(BT)$. The space $H^\bullet_T(Y)$ is a ring under cup product and is also an $H_T(X)$-module via pullback under the projection $Y \to \mathrm{pt}$. For convenience, we will denote $H^\bullet_T(X)$ by $H_T(X)$. In our case, $H_T(X)$ is isomorphic to $H^\bullet(X) \otimes H^\bullet(BT)$ as an $H_T(X)$-module (see, e.g., [Goresky et al. 1998, Theorem 14.1]). Similarly, the equivariant cohomology of any T-equivariant sheaf on X also carries an $H_T(X)$-module structure.

One can define the T-equivariant Borel–Moore homology of X, denoted $H^T_\bullet(X)$. Every T-equivariant closed subvariety Y of X defines a class $[Y]_T$ of degree $2\dim_{\mathbb{C}} Y$ in $H^T_\bullet(X)$. If X is smooth, then every class $[Y]_T$ has an equivariant Poincaré dual cohomology class in $H^\bullet_T(X)$. More details can be found in [Graham 2001] and [Brion 2000].

Proof of the equivariant case. Consider an m-tuple $(p_1, \dots, p_m)$ of T-fixed points of X. Then $\mathrm{IC}_{p_1}, \dots, \mathrm{IC}_{p_m}$ are the IC sheaves corresponding to $U_{p_1}, \dots, U_{p_m}$ respectively. Let $L_{j,n} = i_n^{-1}\,\mathrm{IC}_{p_j}$ for each j and n.

Proposition 4.1. *Under Assumptions 1.1 and 1.2, the cup product maps*

$$H^\bullet_T(L_{1,n}) \underset{H_T(X)}{\otimes} \cdots \underset{H_T(X)}{\otimes} H^\bullet_T(L_{m,n}) \to H^\bullet_T(L_{1,n} \otimes \cdots \otimes L_{m,n})$$

are isomorphisms for each n.

When $X_n = X$, we have $L_{j,n} = \mathrm{IC}_{p_j}$ for each j. Hence this proposition implies Theorem 1.4. To prove the proposition, we first state two general lemmas about T-equivariant cohomology of sheaves.

Lemma 4.2. *Consider the fiber bundle $ET \times_T X \to BT$, with fiber X. Let IC_w be the (T-equivariant) IC sheaf on the closure of a stratum X_w, extended by zero to all of X. Then the Leray spectral sequence for the computation of $H_T^\bullet(X; \mathrm{IC}_w) = H^\bullet(ET \times_T X; \overline{\mathrm{IC}_w})$ collapses at the E_2 page. Hence $H_T^\bullet(\mathrm{IC}_w)$ is isomorphic to $H^\bullet(\mathrm{IC}_w) \otimes H^\bullet(BT)$ as a graded $H^\bullet(BT)$-module.*

Proof. See [Goresky et al. 1998, Theorem 14.1]. The proof uses the fact that the cohomology of $BT \cong (\mathbb{C}P^\infty)^{\dim T}$ is pure. $\square$

Lemma 4.3. *Let Y be any T-space, and let $\mathcal{F}$ be a T-equivariant sheaf on Y such that the space $H^\bullet(Y; \mathcal{F})$ is pure. Then $H_T^\bullet(Y; \mathcal{F})$ is pure as well.*

Proof. Recall that $H_T^\bullet(Y, \mathcal{F}) = H^\bullet(ET \times_T X, \overline{\mathcal{F}})$. The result follows from computing the Leray spectral sequence for the fiber bundle $ET \times_T Y \to BT$, and by using that $H^\bullet(BT)$ and $H^\bullet(Y, \mathcal{F})$ are pure. $\square$

We also record some equivariant analogues of results stated in Section 3. First note that the boundary maps in the long exact sequences of T-equivariant cohomology for the distinguished triangles (3-1) and (3-2) vanish. The proof is analogous to the nonequivariant case, using Lemma 4.3.

The following lemma is an analogue of Lemma 3.1.

Lemma 4.4. *Let $U = X_n \backslash X_{n-1}$. Let $\mathcal{F}$ and $\mathcal{G}$ be any T-equivariant complexes of sheaves on U. Then the cup product map*

$$\cup\colon H_T^\bullet(u_! \mathcal{F}) \underset{H^\bullet(BT)}{\otimes} H_T^\bullet(u_* \mathcal{G}) \to H_T^\bullet(u_! \mathcal{F} \otimes u_* \mathcal{G})$$

is an isomorphism. Since $\cup$ factors through the surjection

$$H_T^\bullet(u_! \mathcal{F}) \underset{H^\bullet(BT)}{\otimes} H_T^\bullet(u_* \mathcal{G}) \twoheadrightarrow H_T^\bullet(u_! \mathcal{F}) \underset{H_T(X)}{\otimes} H_T^\bullet(u_* \mathcal{G}),$$

the induced cup product

$$H_T^\bullet(u_! \mathcal{F}) \underset{H_T(X)}{\otimes} H_T^\bullet(u_* \mathcal{G}) \to H_T^\bullet(u_! \mathcal{F} \otimes u_* \mathcal{G})$$

is also an isomorphism.

Proof. Consider the fiber bundle $ET \times_T X_n \to BT$, with fiber X_n. The E_2 pages of the Leray spectral sequences for $u_! \mathcal{F}$ and $u_* \mathcal{G}$ are as follows:

$$H^p(BT, H^q(u_! \mathcal{F})) \Longrightarrow H_T^{p+q}(u_! \mathcal{F}),$$
$$H^r(BT, H^s(u_* \mathcal{G})) \Longrightarrow H_T^{r+s}(u_* \mathcal{G}).$$

On the E_2 page, the cup product map can be written as the composition of the following two maps. The first map is the cup product with local coefficients, and the second is the fiberwise cup product on the local systems.

$$H^p(BT, H^q(u_!\mathcal{F})) \underset{H^\bullet(BT)}{\otimes} H^r(BT, H^s(u_*\mathcal{G})) \to H^{p+r}(BT, H^q(u_!\mathcal{F}) \otimes H^s(u_*\mathcal{G})),$$
$$H^{p+r}(BT, H^q(u_!\mathcal{F}) \otimes H^s(u_*\mathcal{G})) \to H^{p+r}(BT, H^{q+s}(u_!\mathcal{F} \otimes u_*\mathcal{G})).$$

Since the local systems $H^q(u_!\mathcal{F})$ and $H^s(u_*\mathcal{G})$ are constant on BT, the first map yields isomorphisms

$$H^\bullet(BT, H^q(u_!\mathcal{F})) \underset{H^\bullet(BT)}{\otimes} H^\bullet(BT, H^s(u_*\mathcal{G})) \xrightarrow{\cong} H^\bullet(BT, H^q(u_!\mathcal{F}) \otimes H^s(u_*\mathcal{G})).$$

Finally, we know from Lemma 3.1 that $H^\bullet(u_!\mathcal{F}) \otimes H^\bullet(u_*\mathcal{G}) \xrightarrow{\cong} H^\bullet(u_!\mathcal{F} \otimes u_*\mathcal{G})$ via the cup product map. Altogether, the cup product maps on the E_2 page yield an isomorphism

$$H^\bullet(BT, H^\bullet(u_!\mathcal{F})) \underset{H^\bullet(BT)}{\otimes} H^\bullet(BT, H^\bullet(u_*\mathcal{G})) \xrightarrow{\cong} H^\bullet(BT, H^\bullet(u_!\mathcal{F} \otimes u_*\mathcal{G})).$$

The left hand side is a tensor product of two free $H^\bullet(BT)$-modules over $H^\bullet(BT)$. Hence it converges to

$$H_T^\bullet(u_!\mathcal{F}) \underset{H^\bullet(BT)}{\otimes} H_T^\bullet(u_*\mathcal{G}).$$

The right hand side converges to $H_T^\bullet(u_!\mathcal{F} \otimes u_*\mathcal{G})$. Since the E_2 pages of the left hand side and the right hand side are isomorphic via the cup product map, the following cup product map

$$H_T^\bullet(u_!\mathcal{F}) \underset{H^\bullet(BT)}{\otimes} H_T^\bullet(u_*\mathcal{G}) \to H_T^\bullet(u_!\mathcal{F} \otimes u_*\mathcal{G})$$

is an isomorphism. □

Let $\widetilde{c_n} \in H_T(X)$ be the equivariant Poincaré dual of $[X_n^-]_T$. Each $\widetilde{c_n}$ restricts to the class c_n under the map $H_T(X) \to H^\bullet(X)$, hence the collection $\{\widetilde{c_n}\}$ generates $H_T(X)$ over $H^\bullet(BT)$.

The following lemma (analogous to [Ginzburg 1991, (3.8a)]) describes the action of $\widetilde{c_n}$ on the equivariant cohomology of the sheaves $L_{j,n}$ on X.

Lemma 4.5. *For every j, the action of $\widetilde{c_n}$ on $H_T^\bullet(L_{j,n})$ fits into the following commutative diagram*:

$$\begin{tikzcd} H_T^\bullet(L_{j,n}) \arrow[r, two heads] \arrow[d, "\widetilde{c_n}"'] & H_T^\bullet(u^{-1}L_{j,n}) \arrow[d, "\widetilde{c_n}"', "\cong"] \\ H_T^\bullet(L_{j,n}) & H_{T,c}^\bullet(u^{-1}L_{j,n}) \arrow[l, hook'] \end{tikzcd}$$

Proof. Recall that the intersection of X_n and X_n^- lies away from X_{n-1}. Hence $\widetilde{c}_n$ restricts to zero on X_{n-1}, and cup product by $\widetilde{c}_n$ annihilates the cohomology of any sheaf supported on X_{n-1}. The kernel of $H_T^\bullet(L_{j,n}) \twoheadrightarrow H_T^\bullet(u^{-1}L_{j,n})$ and the cokernel of $H_{T,c}^\bullet(u^{-1}L_{j,n}) \to H_T^\bullet(L_{j,n})$ are both supported on X_{n-1}. So the map of multiplication by $\widetilde{c}_n$ from $H_T^\bullet(X_n)$ to $H_T^\bullet(X_n)$ factors as follows.

$$\begin{tikzcd}
H_T^\bullet(L_{j,n}) \arrow[r, two heads] \arrow[d, "\widetilde{c}_n"'] & H_T^\bullet(u^{-1}L_{j,n}) \arrow[d, "\widetilde{c}_n"'] \\
H_T^\bullet(L_{j,n}) & H_{T,c}^\bullet(u^{-1}L_{j,n}) \arrow[l, hook']
\end{tikzcd}$$

It remains to show that the vertical map on the right is an isomorphism. Since X_n and X_n^- intersect transversally in the single point w_n, the restriction of $\widetilde{c}_n$ to X_n is the image in $H_T^\bullet(X_n)$ of a generator of the local cohomology group $H_T^\bullet(X_n, X_n\backslash\{w_n\})$.

Since $w_n \in U_n$, we have $H_T^\bullet(X_n, X_n\backslash\{w_n\}) \cong H_T^\bullet(U_n, U_n\backslash\{w_n\})$ by excision. But U_n is an affine space that is T-equivariantly contractible to w_n, and hence $H_T^\bullet(U_n, U_n\backslash\{w_n\}) \cong H_{T,c}^\bullet(U_n)$. This shows that multiplication by $\widetilde{c}_n$ maps $H_T^\bullet(U_n)$ isomorphically to $H_{T,c}^\bullet(U_n)$.

Since $u^{-1}L_{j,n}$ is T-equivariant, the above argument applies to the cohomology of $u^{-1}L_{j,n}$ as well. This means that $\widetilde{c}_n$ maps $H_T^\bullet(u^{-1}L_{j,n})$ isomorphically to $H_{T,c}^\bullet(u^{-1}L_{j,n})$, and the proof is complete. □

Once again, let $M_{m,n}$ denote the sheaf $L_{2,n} \otimes \cdots \otimes L_{m,n}$. For brevity, we set up the following additional notation.

$$\begin{aligned}
\bar{A}_{m,n} &= H_T^\bullet(L_{2,n}) \underset{H_T(X)}{\otimes} \cdots \underset{H_T(X)}{\otimes} H_T^\bullet(L_{m,n}), \\
\bar{B}_{m,n} &= H_T^\bullet(u_*u^{-1}L_{2,n}) \underset{H_T(X)}{\otimes} \cdots \underset{H_T(X)}{\otimes} H_T^\bullet(u_*u^{-1}L_{m,n}).
\end{aligned}$$

The following two lemmas are analogues of Lemmas 3.3 and 3.5, respectively.

Lemma 4.6. *There is an exact sequence*

$$H_T^\bullet(u_!u^{-1}L_{1,n}) \underset{H_T(X)}{\otimes} \bar{B}_{m,n} \to H_T^\bullet(L_{1,n}) \underset{H_T(X)}{\otimes} \bar{A}_{m,n} \to H_T^\bullet(v_*v^{-1}L_{1,n}) \underset{H_T(X)}{\otimes} \bar{A}_{m,n} \to 0.$$

Proof. The proof is analogous to the proof of Lemma 3.3. We use the fact that $H_T^\bullet(X) \cong H^\bullet(X) \otimes H^\bullet(BT)$ and use Lemma 4.5 as a substitute for the commutative diagram (3.8a) in [Ginzburg 1991]. □

Lemma 4.7. (i) *The cohomology $H_T^\bullet(L_{1,n} \otimes M_{m,n})$ is pure.*

(ii) *There is a short exact sequence*

$$0 \to H_{T,c}^\bullet(u^{-1}(L_{1,n} \otimes M_{m,n})) \to H_T^\bullet(L_{1,n} \otimes M_{m,n}) \to H_T^\bullet(L_{1,n-1} \otimes M_{m,n-1}) \to 0.$$

Proof. The proofs are analogous to the proofs of their counterparts from Section 3, using the observation of Lemma 4.3 and the fact that $H^\bullet(BT)$ is pure. $\square$

We now complete the proof of Theorem 1.4.

Proof of Theorem 1.4. We obtain the following commutative diagram from the exact sequences of Lemmas 4.6 and 4.7.

$$
\begin{tikzcd}
H_T^\bullet(u_!u^{-1}L_{1,n}) \underset{H_T(X)}{\otimes} \bar{B}_{m,n} \arrow[r] \arrow[d, "a"] & H_T^\bullet(L_{1,n}) \underset{H_T(X)}{\otimes} \bar{A}_{m,n} \arrow[r, two heads] \arrow[d, "b"] & H_T^\bullet(v_*v^{-1}L_{1,n}) \underset{H_T(X)}{\otimes} \bar{A}_{m,n} \arrow[d, "c"] \\
H_T^\bullet(u_!u^{-1}L_{1,n} \otimes M_{m,n}) \arrow[r, hook] & H_T^\bullet(L_{1,n} \otimes M_{m,n}) \arrow[r, two heads] & H_T^\bullet(v_*v^{-1}L_{1,n} \otimes M_{m,n})
\end{tikzcd} \tag{4-1}
$$

First observe that the action of $H_T(X)$ on $H_T^\bullet(u_!u^{-1}L_{1,n})$ and on $\bar{B}_{m,n}$ factors through the map $H_T(X) \to H_T^\bullet(U) \cong H^\bullet(BT)$, so

$$H_T^\bullet(u_!u^{-1}L_{1,n}) \underset{H_T(X)}{\otimes} \bar{B}_{m,n} \cong H_T^\bullet(u_!u^{-1}L_{1,n}) \underset{H^\bullet(BT)}{\otimes} \bar{B}_{m,n}.$$

We prove by induction on m that the map a is an isomorphism. As in the proof of Lemma 3.2, the case of $m = 2$ is proved by Lemma 4.4, and the general case is proved by iterating the argument. An argument similar to the proof of Lemma 3.4 proves that the map c is an isomorphism.

Hence by the snake lemma, the middle map b is an isomorphism as well. Consequently, we obtain the following isomorphisms for every n:

$$H_T^\bullet(L_{1,n}) \underset{H_T(X)}{\otimes} \cdots \underset{H_T(X)}{\otimes} H_T^\bullet(L_{m,n}) \to H_T^\bullet(L_{1,n} \otimes \cdots \otimes L_{m,n}).$$

In particular when $X_n = X$, we see that the cup product map

$$H_T^\bullet(\mathrm{IC}_{p_1}) \underset{H_T(X)}{\otimes} \cdots \underset{H_T(X)}{\otimes} H_T^\bullet(\mathrm{IC}_{p_m}) \to H_T^\bullet(\mathrm{IC}_{p_1} \otimes \cdots \otimes \mathrm{IC}_{p_m})$$

is an isomorphism. $\square$

Acknowledgments. I am extremely grateful to Victor Ginzburg for suggesting the problem, and for his valuable advice and continued guidance throughout the project. I am grateful to the anonymous referee for providing detailed feedback and several improvements. I thank Quoc Ho for many inspiring mathematical conversations.

References

[Achar and Rider 2014] P. N. Achar and L. Rider, "Parity sheaves on the affine Grassmannian and the Mirković–Vilonen conjecture", preprint, 2014. arXiv 1305.1684

[Audin 2004] M. Audin, *Torus actions on symplectic manifolds*, 2nd ed., Progress in Mathematics **93**, Birkhäuser, Basel, 2004. MR 2005k:53158 Zbl 1062.57040

[Bernstein and Lunts 1994] J. Bernstein and V. Lunts, *Equivariant sheaves and functors*, Lecture Notes in Mathematics **1578**, Springer, Berlin, 1994. MR 95k:55012 Zbl 0808.14038

[Białynicki-Birula 1973] A. Białynicki-Birula, "Some theorems on actions of algebraic groups", *Ann. of Math.* (2) **98** (1973), 480–497. MR 51 #3186 Zbl 0275.14007

[Braden 2003] T. Braden, "Hyperbolic localization of intersection cohomology", *Transform. Groups* **8**:3 (2003), 209–216. MR 2004f:14037 Zbl 1026.14005

[Brion 2000] M. Brion, "Poincaré duality and equivariant (co)homology", *Michigan Math. J.* **48** (2000), 77–92. MR 2001m:14032 Zbl 1077.14523

[Chriss and Ginzburg 1997] N. Chriss and V. Ginzburg, *Representation theory and complex geometry*, Birkhäuser, Boston, 1997. Reprinted Springer, 2010. MR 98i:22021 Zbl 0879.22001

[Ginzburg 1991] V. Ginsburg, "Perverse sheaves and $\mathbb{C}^*$-actions", *J. Amer. Math. Soc.* **4**:3 (1991), 483–490. MR 92d:14013 Zbl 0760.14008

[Goresky et al. 1998] M. Goresky, R. Kottwitz, and R. MacPherson, "Equivariant cohomology, Koszul duality, and the localization theorem", *Invent. Math.* **131**:1 (1998), 25–83. MR 99c:55009 Zbl 0897.22009

[Graham 2001] W. Graham, "Positivity in equivariant Schubert calculus", *Duke Math. J.* **109**:3 (2001), 599–614. MR 2002h:14083 Zbl 1069.14055

[Juteau et al. 2014] D. Juteau, C. Mautner, and G. Williamson, "Parity sheaves", *J. Amer. Math. Soc.* **27**:4 (2014), 1169–1212. MR 3230821 Zbl 06355534

[Kambayashi 1966] T. Kambayashi, "Projective representation of algebraic linear groups of transformations", *Amer. J. Math.* **88** (1966), 199–205. MR 34 #5826 Zbl 0141.18303

[Saito 1988] M. Saito, "Modules de Hodge polarisables", *Publ. Res. Inst. Math. Sci.* **24**:6 (1988), 849–995. MR 90k:32038 Zbl 0691.14007

[Saito 1990] M. Saito, "Mixed Hodge modules", *Publ. Res. Inst. Math. Sci.* **26**:2 (1990), 221–333. MR 91m:14014 Zbl 0727.14004

[Soergel 1990] W. Soergel, "Kategorie $\mathbb{O}$, perverse Garben und Moduln über den Koinvarianten zur Weylgruppe", *J. Amer. Math. Soc.* **3**:2 (1990), 421–445. MR 91e:17007 Zbl 0747.17008

[Springer 1984] T. A. Springer, "A purity result for fixed point varieties in flag manifolds", *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **31**:2 (1984), 271–282. MR 86c:14034 Zbl 0581.20048

[Sumihiro 1974] H. Sumihiro, "Equivariant completion", *J. Math. Kyoto Univ.* **14**:1 (1974), 1–28. MR 49 #2732 Zbl 0277.14008

Received July 8, 2014. Revised December 12, 2014.

ASILATA BAPAT
DEPARTMENT OF MATHEMATICS
THE UNIVERSITY OF CHICAGO
5734 S UNIVERSITY AVENUE
CHICAGO, IL 60637
UNITED STATES

asilata@math.uchicago.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.35

CYCLICITY IN DIRICHLET-TYPE SPACES AND EXTREMAL POLYNOMIALS II: FUNCTIONS ON THE BIDISK

CATHERINE BÉNÉTEAU, ALBERTO A. CONDORI, CONSTANZE LIAW, DANIEL SECO AND ALAN A. SOLA

We study Dirichlet-type spaces $\mathfrak{D}_\alpha$ of analytic functions in the unit bidisk and their cyclic elements. These are the functions f for which there exists a sequence $(p_n)_{n=1}^{\infty}$ of polynomials in two variables such that $\|p_n f - 1\|_\alpha \to 0$ as $n \to \infty$. We obtain a number of conditions that imply cyclicity, and obtain sharp estimates on the best possible rate of decay of the norms $\|p_n f - 1\|_\alpha$, in terms of the degree of p_n, for certain classes of functions using results concerning Hilbert spaces of functions of one complex variable and comparisons between norms in one and two variables.

We give examples of polynomials with no zeros on the bidisk that are not cyclic in $\mathfrak{D}_\alpha$ for $\alpha > 1/2$ (including the Dirichlet space); this is in contrast with the one-variable case where all nonvanishing polynomials are cyclic in Dirichlet-type spaces that are not algebras ($\alpha \le 1$). Further, we point out the necessity of a capacity zero condition on zero sets (in an appropriate sense) for cyclicity in the setting of the bidisk, and conclude by stating some open problems.

1. Introduction

Dirichlet-type spaces on the bidisk. We consider a scale of Hilbert spaces of holomorphic functions on the bidisk

$$\mathbb{D}^2 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1| < 1, |z_2| < 1\},$$

indexed by a parameter $\alpha \in (-\infty, \infty)$. A holomorphic function $f : \mathbb{D}^2 \to \mathbb{C}$ belongs

---

Liaw is partially supported by the NSF grant DMS-1261687. Seco is supported by ERC Grant 2011-ADG-20110209 from EU programme FP2007-2013, and by MEC/MICINN Project MTM2011-24606. Sola acknowledges support from the EPSRC under grant EP/103372X/1.

MSC2010: primary 32A37; secondary 32A36, 47A16.

Keywords: cyclicity, Dirichlet-type spaces, optimal approximation, norm restrictions.

to the *Dirichlet-type space* $\mathfrak{D}_\alpha$ if its power series expansion

$$f(z_1, z_2) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_{k,l} z_1^k z_2^l$$

satisfies

$$\|f\|_\alpha^2 = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} (k+1)^\alpha (l+1)^\alpha |a_{k,l}|^2 < \infty. \tag{1-1}$$

Recall that a function of two complex variables is said to be *holomorphic* if it is holomorphic in each variable separately. A review of the definitions and basic properties such as power series expansions can be found in [Hörmander 1990, Chapter 2]. Since zero sets on the boundary of functions $f \in \mathfrak{D}_\alpha$ will play a role later on, we point out that the topological boundary of the bidisk is much larger than the so-called *distinguished boundary*

$$\mathbb{T}^2 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1| = |z_2| = 1\},$$

which is still large enough to support standard integral representations and the maximum principle on the bidisk.

The spaces $\mathfrak{D}_\alpha$ are a natural generalization to two variables of the classical Dirichlet-type spaces D_α, $-\infty < \alpha < \infty$, consisting of functions

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

that are analytic in the unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and satisfy

$$\|f\|_{D_\alpha}^2 = \sum_{k=0}^{\infty} (k+1)^\alpha |a_k|^2 < \infty;$$

see, for instance, [Taylor 1966; Brown and Shields 1984], and the references therein. As a remark on notation, we will continue to use $\|\cdot\|_\alpha$ for the norm of two variable functions in $\mathfrak{D}_\alpha$ while $\|\cdot\|_{D_\alpha}$ will denote the norm of one variable functions in D_α. We point out that the particular choice $\alpha = 0$ in D_α and $\mathfrak{D}_\alpha$ leads to the classical Hardy spaces H^2 on the disk and bidisk, respectively, while

$$D_{-1} = A^2(\mathbb{D}) \quad \text{and} \quad \mathfrak{D}_{-1} = A^2(\mathbb{D}^2)$$

are the canonical Bergman spaces of the disk and bidisk, and D_1 and $\mathfrak{D}_1$ are the Dirichlet spaces of the disk and bidisk, respectively.

The spaces $\mathfrak{D}_\alpha$ were studied in detail by Jupiter and Redett [2006]. Spaces of this type appear in the earlier work of Kaptanoğlu [1994], which focuses on Möbius invariance and boundary behavior in Dirichlet-type spaces, and Hedenmalm [1988], which concentrates on closed ideals in function algebras. We note here (compare [Kaptanoğlu 1994, p. 343; Hedenmalm 1988, Section 4]) that an equivalent norm for $\mathfrak{D}_\alpha$ is given by

$$\begin{aligned}\|f\|_\alpha^2 = |f(0,0)|^2 &\\ &+ \int_{\mathbb{D}} |\partial_{z_1} f(z_1,0)|^2 (1-|z_1|^2)^{1-\alpha}\, dA(z_1) \\ &+ \int_{\mathbb{D}} |\partial_{z_2} f(0,z_2)|^2 (1-|z_2|^2)^{1-\alpha}\, dA(z_2) \\ &+ \int_{\mathbb{D}^2} |\partial_{z_2}\partial_{z_1} f(z_1,z_2)|^2 (1-|z_1|^2)^{1-\alpha}(1-|z_2|^2)^{1-\alpha}\, dA(z_1)\, dA(z_2),\end{aligned}$$

where $dA(z) = \pi^{-1}\, dx\, dy$ denotes area measure. The proof involves computations with power series, and is omitted.

Extending the earlier one-variable work of G. D. Taylor [1966] and Stegenga [1980], Jupiter and Redett identified multipliers on $\mathfrak{D}_\alpha$ and studied restriction properties of these spaces. It was also shown in [Jupiter and Redett 2006] that evaluation at a point in $\mathbb{D}^2$ is a bounded linear functional, and hence $\mathfrak{D}_\alpha$ is a *reproducing kernel Hilbert space* for all α. When $\alpha > 1$, the spaces $\mathfrak{D}_\alpha$ are actually *algebras* (viz. the proof of [op. cit., Theorem 3.10]) that are contained (as sets) in $H^\infty(\mathbb{D}^2)$, the algebra of bounded holomorphic functions.

It is clear from the definition of the norm in (1-1) that any polynomial $p = p(z_1, z_2)$ belongs to $\mathfrak{D}_\alpha$. Moreover, any $f \in D_\alpha$ lifts to $\mathfrak{D}_\alpha$ when regarded as constant in one of the variables. In fact, if $g \in D_\alpha$ and $h \in D_\alpha$, then the function

$$f(z_1,z_2) = g(z_1)h(z_2), \quad (z_1,z_2) \in \mathbb{D}^2,$$

is analytic in the bidisk and belongs to $\mathfrak{D}_\alpha$ [op. cit., Proposition 4.7], and so $\mathfrak{D}_\alpha$ certainly contains nontrivial holomorphic functions.

Shift operators and cyclic functions. In this paper, we are interested in a natural pair $\{S_1, S_2\}$ of bounded linear operators acting on the spaces $\mathfrak{D}_\alpha$. The *shift operators* S_1 and S_2 are defined by setting, for $f \in \mathfrak{D}_\alpha$,

$$S_1 f(z_1,z_2) = z_1 f(z_1,z_2) \quad \text{and} \quad S_2 f(z_1,z_2) = z_2 f(z_1,z_2).$$

It is then clear that S_1 and S_2 are linear, and it follows from (1-1) that, for every α, $\{S_1, S_2\}$ forms a pair of bounded operators mapping $\mathfrak{D}_\alpha$ into itself.

It is a standard problem of operator theory to describe the invariant subspaces of an operator. In the present context, we are interested in closed subspaces $\mathcal{M} \subset \mathfrak{D}_\alpha$ such that

$$S_1\mathcal{M} \subset \mathcal{M} \quad \text{and} \quad S_2\mathcal{M} \subset \mathcal{M}.$$

As a first step towards understanding the invariant subspaces of the pair $\{S_1, S_2\}$, we seek conditions under which a function $f \in \mathfrak{D}_\alpha$ is *cyclic*, that is,

$$[f] = \overline{\operatorname{span}\{z_1^k z_2^l f : k = 0, 1, \dots; l = 0, 1, \dots\}} = \mathfrak{D}_\alpha.$$

It is easy to see that there exists at least one cyclic function in each $\mathfrak{D}_\alpha$, namely the function $f(z_1, z_2) = 1$. This follows from the fact that polynomials in two variables are dense in $\mathfrak{D}_\alpha$. On the other hand, since norm convergence implies uniform convergence on compact subsets, every $g \in [f]$ inherits any zeros f may have inside $\mathbb{D}^2$, and so a necessary condition for cyclicity is that $f(z_1, z_2) \neq 0$, $(z_1, z_2) \in \mathbb{D}^2$. Note that since $g \in [f]$ implies $[g] \subset [f]$, an equivalent condition for f to be cyclic in $\mathfrak{D}_\alpha$ is that there exists a sequence of polynomials $(p_n)_{n=1}^\infty$ of two variables with

$$\|p_n f - 1\|_\alpha \to 0, \quad n \to \infty.$$

Since point evaluation is a bounded linear functional, this latter condition is equivalent to the existence of a sequence of polynomials (p_n) such that

$$p_n(z_1, z_2) f(z_1, z_2) - 1 \to 0, \quad (z_1, z_2) \in \mathbb{D}^2,$$

and

$$\|p_n f - 1\|_\alpha \leq C.$$

When $\alpha > 1$ the spaces D_α and $\mathfrak{D}_\alpha$ are algebras, and cyclic functions have to be nonvanishing on $\overline{\mathbb{D}}$ and $\overline{\mathbb{D}^2}$, respectively.

In one variable, Beurling characterized the cyclic vectors of $H^2(\mathbb{D})$: a function f is cyclic if and only if it is outer. In the bidisk, one can show that if $f \in H^2(\mathbb{D}^2)$ or indeed if f belongs to the Nevanlinna class, then f has (nonzero) radial limits f^* at almost every $(\zeta_1, \zeta_2) \in \mathbb{T}^2$. Thus, we can declare $f \in H^2(\mathbb{D}^2)$ to be *outer* if

$$\log|f(z_1, z_2)| = \int_{\mathbb{T}^2} \log|f^*(e^{i\theta}, e^{i\eta})| \, P((z_1, z_2); (e^{i\theta}, e^{i\eta})) \, d\theta d\eta;$$

here, P is the product Poisson kernel

$$P((z_1, z_2); (e^{i\theta}, e^{i\eta})) = P_{|z_1|}(\arg z_1 - \theta) \, P_{|z_2|}(\arg z_2 - \eta),$$

where $(z_1, z_2) \in \mathbb{D}^2$ and $\theta, \eta \in [0, 2\pi)$. As usual,

$$P_r(\theta) = \frac{1 - r^2}{(r^2 - 2r\cos(\theta) + 1)^2}$$

denotes the Poisson kernel of the unit disk.

The cyclicity of $f \in H^2(\mathbb{D}^2)$ does imply that f is an outer function. But this condition is no longer sufficient: there are outer functions that are not cyclic [Rudin 1969, Theorem 4.4.6]; this is another example of how the higher-dimensional theory is somewhat different. (See, however, [Mandrekar 1988; Douglas and Yang 2000; Redett and Tung 2010] for some positive results.)

Polynomials in two variables with no zeros in $\mathbb{D}^2$ are outer functions, and are therefore candidates for being cyclic in $\mathfrak{D}_\alpha$ for $\alpha \geq 0$. Indeed, Gelca [1995] proved that polynomials f with $\mathcal{Z}(f) \cap \mathbb{D}^2 = \varnothing$ are cyclic in $H^2(\mathbb{D}^2)$, the Hardy space of the bidisk, and hence in $\mathfrak{D}_\alpha$ for all $\alpha \leq 0$.

Overview of results. In [Bénéteau et al. 2015], the problem of cyclicity in Dirichlet-type spaces in the unit disk was studied. More specifically, the authors identified some subclasses of cyclic functions and derived sharp estimates on the rate of decay of the norms $\|p_n f - 1\|_\alpha$ for such $f \in D_\alpha$. It seems natural to investigate to what extent these results can be extended to functions $f \in \mathfrak{D}_\alpha$.

To make the notion of best possible norm decay precise, we let $\mathfrak{P}_n$, $n = 1, 2, \ldots,$ be the subspaces of $\mathfrak{D}_\alpha$ consisting of polynomials of two variables of the form

$$p_n = \sum_{k=0}^{n} \sum_{l=0}^{n} c_{k,l} z_1^k z_2^l.$$

Note that we regard a monomial $z_1^k z_2^l$ in two variables as having degree $k + l$, meaning that members of $\mathfrak{P}_n$ are polynomials of degree at most $2n$. Similarly, we denote by $\mathcal{P}_n$ the space of polynomials of one complex variable having degree at most n. We now make the following definition.

Definition 1.1. Let $f \in \mathfrak{D}_\alpha$. We say that a polynomial $p_n \in \mathfrak{P}_n$ is an *optimal approximant* of order n to $1/f$ if p_n minimizes $\|pf - 1\|_\alpha$ among all polynomials $p \in \mathfrak{P}_n$. We call $\|p_n f - 1\|_\alpha$ the *optimal norm* of order n associated with f.

Stated differently, p_n is an optimal approximant to $1/f$ if we have

$$\|p_n f - 1\|_\alpha = \operatorname{dist}_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n);$$

here, $\operatorname{dist}_X(x, A) = \inf\{\|x - a\|_X : a \in A\}$ is the usual distance function between a point and a subset $A \subset X$ of a normed space X.

Sharp estimates on the unit disk analog of $\operatorname{dist}_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n)$ were obtained for certain classes of functions in [Bénéteau et al. 2015]. To state these estimates, we define $\varphi_1(s) = \log^+(s)$ for $s \in [0, \infty)$ and, when $\alpha < 1$,

$$\varphi_\alpha(s) = s^{1-\alpha}, \quad s \in [0, \infty).$$

Theorem 1.2 [Bénéteau et al. 2015, Theorem 3.6]. *Let $\alpha \leq 1$. If f is a function admitting an analytic continuation to the closed unit disk and whose zeros lie in $\mathbb{C} \setminus \mathbb{D}$, then there exists a constant $C = C(\alpha, f)$ such that*

$$\operatorname{dist}^2_{D_\alpha}(1, f \cdot \mathcal{P}_m) \leq C \varphi_\alpha^{-1}(m + 1)$$

holds for all sufficiently large m. This estimate is sharp in the sense that if such a function f has at least one zero on $\mathbb{T}$, there exists a constant $\tilde{C} = \tilde{C}(\alpha, f)$ such that

$$\tilde{C}\varphi_\alpha^{-1}(m+1) \leq \operatorname{dist}^2_{D_\alpha}(1, f \cdot \mathcal{P}_m).$$

In this paper, we obtain analogous theorems for certain subclasses of functions in $\mathfrak{D}_\alpha$. We begin Section 2 with some general remarks concerning cyclicity in $\mathfrak{D}_\alpha$. For instance, if f is cyclic, then each slice function f_{z_j} obtained when fixing the variable z_j, $j = 1$ or 2, has to be cyclic in D_α. Then the problem of cyclicity and rates associated with optimal approximants is addressed for separable functions, i.e., for functions f of the form $f(z_1, z_2) = g(z_1)h(z_2)$. We prove that such a function is cyclic if and only if the factors g and h are cyclic in the one-variable space D_α, and then obtain, in Theorem 2.6, sharp estimates on $\operatorname{dist}_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n)$ under the assumption that g and h admit analytic continuation to the closed disk and have no zeros in $\mathbb{D}$.

In Section 3, we turn our attention to functions of the form $f(z_1, z_2) = f(z_1^M \cdot z_2^N)$, for integers $M, N \geq 1$, and again obtain cyclicity results and sharp estimates in Theorem 3.1. Our proofs are based on the fact that certain restriction operators furnish isomorphisms between our subclasses of functions in $\mathfrak{D}_\alpha$ and the one-variable spaces $D_{2\alpha}$, and on comparisons between the associated norms.

In [Bénéteau et al. 2015], a key role was played by certain Riesz-type means of the power series expansion of $1/f$, which turned out to produce optimal, or near optimal, approximants to $1/f$. The one-variable construction extends to the bidisk setting as follows. Suppose $1/f$ has formal power series expansion

$$\frac{1}{f(z_1, z_2)} = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} b_{k,l} z_1^k z_2^l.$$

We then set

$$p_n(z_1, z_2) = \sum_{k=0}^{n}\sum_{l=0}^{n}\left(1 - \frac{\varphi_\alpha(\max\{k, l\})}{\varphi_\alpha(n+1)}\right) b_{k,l} z_1^k z_2^l. \tag{1-2}$$

Note that when $\alpha = 0$, the polynomials p_n are simply the n-th Cesàro means of the Taylor series of $1/f$:

$$\begin{aligned} C_n(1/f)(z_1, z_2) &= \sum_{k=0}^{n}\sum_{l=0}^{n}\left(1 - \frac{\max\{k, l\}}{n+1}\right) b_{k,l} z_1^k z_2^l \\ &= \frac{1}{n+1}\sum_{m=0}^{n} t_m(1/f)(z_1, z_2), \end{aligned}$$

where t_m denotes the m-th order Taylor polynomial. In Section 4, we take a closer look at some concrete polynomials in two variables, and show that in some cases the polynomials (1-2) are indeed close to optimal.

Recall that in the case of the unit disk, any polynomial that is zero-free in $\mathbb{D}$ is cyclic in D_α for all $\alpha \le 1$. However, the analogous statement for the bidisk need not hold. In fact, we give examples of polynomials whose zero sets lie in $\mathbb{T}^2$ that are noncyclic for $\alpha > \frac{1}{2}$, and also polynomials with zeros on the boundary of the bidisk that are cyclic for all $\alpha \le 1$; in fact, such polynomials can have zero sets that intersect $\mathbb{T}^2$, and extend into $\partial\mathbb{D}^2 \setminus \mathbb{T}^2$.

The existence of noncyclic polynomials in Hilbert spaces of analytic functions in higher dimensions has also been observed by Richter and Sundberg in the setting of the Drury–Arveson space in the unit ball of $\mathbb{C}^d$ when $d \ge 4$; see [Richter and Sundberg 2012] for this and other results on cyclic vectors in that context.

Many of our results and arguments carry over to the d-dimensional polydisk $\mathbb{D}^d$, but as notation becomes much more cumbersome, we restrict our attention to functions on the bidisk.

2. Classes of cyclic vectors in $\mathfrak{D}_\alpha$

In this section, we present some examples of cyclic functions in the bidisk. As a preliminary example, we have already observed that $f(z_1, z_2) = 1$ is cyclic in $\mathfrak{D}_\alpha$ for all α, and that cyclic functions cannot vanish inside the bidisk. Moreover, it is not difficult to see that if both f and $1/f$ extend to a larger bidisk, then f is nonvanishing on the closure $\overline{\mathbb{D}^2}$, and f is cyclic; indeed, if (p_n) is a sequence of polynomials such that $\|p_n - 1/f\|_\alpha$ tends to 0, the estimate

$$\|p_n f - 1\|_\alpha \le \|f\|_{M(\mathfrak{D}_\alpha)} \|p_n - 1/f\|_\alpha,$$

where $\|\cdot\|_{M(\mathfrak{D}_\alpha)}$ denotes the multiplier norm, implies that $1 \in [f]$ and so f is cyclic.

However, there do exist cyclic functions in $\mathfrak{D}_\alpha$ that vanish on the boundary of the bidisk, as in the one variable case. In this section, we focus on three different ways of building functions in the bidisk from one variable functions in the unit disk, and explore the relationship between the cyclicity in two variables versus that in one variable. First, let us make some preliminary remarks.

Slices of a function. For a function $f = f(z_1, z_2)$ in the bidisk, we can fix the variable z_2, say, and consider the *slice*

$$f_{z_2}(z_1) = f(z_1, z_2), \quad z_1 \in \mathbb{D},$$

as a function in the unit disk. The slice f_{z_1} is defined in an analogous manner.

Proposition 2.1. *If f is cyclic in $\mathfrak{D}_\alpha$, then the slices f_{z_2} and f_{z_1} are cyclic in D_α.*

Proof. As a consequence of the Cauchy–Schwarz inequality applied to the coefficients of f_{z_2} we obtain

$$\|f_{z_2}\|_{D_\alpha} \leq \|k_{z_2}\|_{D_\alpha} \cdot \|f\|_\alpha,$$

where k_{z_2} denotes the reproducing kernel at z_2 for D_α. Therefore, for any polynomial $p = p(z_1, z_2)$, we get

$$\|p_{z_2} f_{z_2} - 1\|_{D_\alpha} \leq \|k_{z_2}\|_{D_\alpha} \cdot \|pf - 1\|_\alpha.$$

If f is cyclic in $\mathfrak{D}_\alpha$, then this last norm tends to 0 as the degree of p approaches ∞, and therefore for fixed z_2, $\|p_{z_2} f_{z_2} - 1\|_{D_\alpha}$ approaches 0 as well. Consequently, the slice f_{z_2} is cyclic in D_α. An analogous argument applies to the slices in z_1, and thus the result follows. $\square$

Note that the converse of the above statement does not hold: consider, for example, $f(z_1, z_2) = 1 - z_1 z_2$. Then each slice f_{z_2} and f_{z_1} is nonvanishing in the closed unit disk (for a fixed z_2 and a fixed z_1, respectively), and thus each is cyclic in every D_α, but it turns out that f is only cyclic in $\mathfrak{D}_\alpha$ for $\alpha \leq \frac{1}{2}$; see Remark 3.2.

Let us now consider three different natural ways to construct a one variable function from a two variable function and examine issues of cyclicity.

Diagonal restrictions. The *restriction to the diagonal* of a holomorphic function on the bidisk produces a function on the disk, and it turns out that these functions often inherit properties that allow us to transfer information between one and two variable spaces; see, e.g., [Horowitz and Oberlin 1975; Rudin 1969]. For instance, Massaneda and Thomas [2013] were able to use restriction arguments to show that it is not possible to characterize cyclic functions in $H^2(\mathbb{D}^2)$ in terms of decay at the boundary.

We define the restriction operator R_{diag} on $f \in \mathfrak{D}_\alpha$ by

$$R_{\text{diag}} : f \mapsto (\oslash f)(z) = f(z, z), \quad z \in \mathbb{D}.$$

To rigorously define which spaces this restriction operator acts on, we define the map

$$\beta(\alpha) = \begin{cases} \alpha - 1 & \text{for } \alpha \geq 0, \\ 2\alpha - 1 & \text{for } \alpha < 0. \end{cases}$$

In order to shorten notation, we use the abbreviation $\beta = \beta(\alpha)$. In the context of the Dirichlet-type spaces, the following restriction estimate holds.

Proposition 2.2. *For all* $f \in \mathfrak{D}_\alpha$,

$$\|\oslash f\|_{D_\beta} \leq \|f\|_\alpha.$$

This result is probably known to the experts, and can be proved by appealing to the theory of reproducing kernels. For the convenience of the reader, we give an elementary proof.

Proof of Proposition 2.2. Let $f(z_1, z_2) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_{k,l} z_1^k z_2^l$, which converges absolutely for every $|z_1| < 1$ and $|z_2| < 1$. Then

$$\oslash f(z) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_{k,l} z^{k+l}$$

converges absolutely for every $|z| < 1$, hence can be rewritten as $\oslash f(z) = \sum_{n=0}^{\infty} b_n z^n$, where $b_n = \sum_{k+l=n} a_{k,l} = \sum_{k=0}^{n} a_{k,n-k}$. Thus,

$$\| \oslash f \|_{D_\beta}^2 = \sum_{n=0}^{\infty} |b_n|^2 (n+1)^\beta = \sum_{n=0}^{\infty} \left| \sum_{k=0}^{n} a_{k,n-k} \right|^2 (n+1)^\beta$$

and

$$\| f \|_\alpha^2 = \sum_{n=0}^{\infty} \sum_{k=0}^{n} |a_{k,n-k}|^2 (k+1)^\alpha (n-k+1)^\alpha.$$

By the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\left| \sum_{k=0}^{n} a_{k,n-k} \right|^2 &\le \left(\sum_{k=0}^{n} |a_{k,n-k}|^2 (k+1)^\alpha (n-k+1)^\alpha \right) \left(\sum_{k=0}^{n} (k+1)^{-\alpha} (n-k+1)^{-\alpha} \right) \\
&\le \left(\sum_{k=0}^{n} |a_{k,n-k}|^2 (k+1)^\alpha (n-k+1)^\alpha \right) (n+1)^{-\beta}.
\end{aligned}$$

In summary, our observations yield, as required,

$$\begin{aligned}
\| \oslash f \|_{D_\beta}^2 &= \sum_{n=0}^{\infty} \left| \sum_{k=0}^{n} a_{k,n-k} \right|^2 (n+1)^\beta \\
&\le \sum_{n=0}^{\infty} \sum_{k=0}^{n} |a_{k,n-k}|^2 (k+1)^\alpha (n-k+1)^\alpha = \| f \|_\alpha^2. \qquad \square
\end{aligned}$$

This result implies that a function $g \in D_\beta$ that arises as the restriction to the diagonal of a cyclic function in $\mathfrak{D}_\alpha$ is itself cyclic. Viewed differently, a function of two variables cannot be cyclic in $\mathfrak{D}_\alpha$ unless its restriction $\oslash f$ is cyclic in D_β, though it can happen that $\oslash f$ is cyclic, and $f \in \mathfrak{D}_\alpha$ is not: the functions considered in the examples in Section 4 are not cyclic in $\mathfrak{D}_2$, but their restrictions $\oslash f$ are cyclic in the Dirichlet space D (see also [Massaneda and Thomas 2013] for a discussion in the context of $H^2(\mathbb{D}^2)$). Moreover, together with the second assertion in Theorem 1.2, Proposition 2.2 immediately implies a lower bound for the decay rate of $\| p_n f - 1 \|_\alpha^2$ for certain "nice" functions f:

Corollary 2.3. *Let $\alpha \leq 2$. Suppose $f \in \mathfrak{D}_\alpha$ is such that the diagonal restriction $\oslash f$ satisfies the hypotheses of Theorem 1.2. Then,*

$$\|p_n f - 1\|_\alpha^2 \geq C\varphi_\beta^{-1}(n+1), \quad \textit{for all } p_n \in \mathfrak{P}_n.$$

Here, we have used that $\varphi_\beta^{-1}(2n+1)$ is comparable to $\varphi_\beta^{-1}(n+1)$. We will see later (see Proposition 2.4 and Theorem 3.1) that this decay rate is not optimal in general. Note that the diagonal restrictions of the functions $f(z_1, z_2) = 1 - z_1 z_2$, $f(z_1, z_2) = (1 - z_1)(1 - z_2)$, and $f(z_1, z_2) = 1 - z_1$ all satisfy the hypotheses.

The above remarks show how, given a cyclic function of two variables, one can easily obtain examples of cyclic functions of one variable (although we might need to change the index α of the space in which cyclicity is being considered!) In the next two subsections we examine how to obtain some classes of cyclic functions of two variables from cyclic functions of one variable, and we obtain *sharp* rates of decay in some cases.

Separable functions. Let us now consider functions of two variables that can be written as products of two functions of one variable:

$$f(z_1, z_2) = g(z_1)h(z_2). \tag{2-1}$$

We shall refer to such functions as *separable*. Note that for such products, it follows from (1-1) that $\|f\|_\alpha = \|g\|_{D_\alpha} \|h\|_{D_\alpha}$.

Proposition 2.4. *Let $\alpha \in \mathbb{R}$ and f be defined as in* (2-1), *where $g, h \in D_\alpha$. Then, f is cyclic in $\mathfrak{D}_\alpha$ if and only if g and h are cyclic in D_α.*

Proof. First notice that by Proposition 2.1, if f is cyclic in $\mathfrak{D}_\alpha$, then g and h are constant multiples (with respect to the fixed variable) of the slices of f, and thus are cyclic in D_α.

For the converse, suppose both g and h are cyclic in D_α. Let (p_n) and (q_n) be sequences of polynomials such that $\|p_n g - 1\|_{D_\alpha} \to 0$ and $\|q_n h - 1\|_{D_\alpha} \to 0$, respectively. Since the expression $p_n gh - h = (p_n(z_1)g(z_1) - 1)h(z_2)$ is separable, we obtain

$$\|p_n f - h\|_\alpha = \|p_n g - 1\|_{D_\alpha} \|h\|_{D_\alpha}.$$

Hence, we get that $h \in [f]$, where $[\,\cdot\,]$ denotes the cyclicity class in $\mathfrak{D}_\alpha$, and so $[h] \subset [f]$. Since $\|q_n h - 1\|_\alpha = \|q_n h - 1\|_{D_\alpha}$, the function h is cyclic in $\mathfrak{D}_\alpha$ and D_α simultaneously, and the assertion follows. $\square$

It seems natural to ask whether the growth of the extremal polynomials for separable functions is the same as for functions in the unit disk. As we will see in Theorem 2.6, this is indeed the case. Let us first prove a lemma that will help to establish the sharp growth restrictions.

Lemma 2.5. *Suppose $f = g \cdot h \in \mathfrak{D}_\alpha$ for $g, h \in \mathfrak{D}_\alpha$, and suppose that g admits a nonvanishing analytic continuation to the closed bidisk. Then, there exists a constant C, independent of n, such that*

$$\operatorname{dist}_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n) \geq C \operatorname{dist}_{\mathfrak{D}_\alpha}(1, h \cdot \mathfrak{P}_{2n}).$$

Proof. Notice first that since the power series for g converges in a larger polydisk than the unit bidisk, there exists $R > 1$ such that if g_n are the Taylor polynomials of degree n approximating g, the multiplier norm $\|g - g_n\|_{M(\mathfrak{D}_\alpha)}$ decays exponentially like $R^{-(n+1)}$. Moreover, since in addition g has no zeros in the closed disk, the multiplier norm $\|1/g\|_{M(\mathfrak{D}_\alpha)}$ is bounded.

Now let $p_n(z_1, z_2)$ be the optimal approximant to $1/f$ of degree n. Then by the above remarks, we have

$$\|p_n h - 1/g\|_\alpha \leq \|1/g\|_{M(\mathfrak{D}_\alpha)} \|p_n f - 1\|_\alpha,$$

which goes to 0 as $n \to \infty$, and therefore, in particular, the norms $\|p_n h\|_\alpha$ are bounded by some constant C_1. Moreover,

$$\begin{aligned}\|p_n f - 1\|_\alpha &= \|p_n h(g - g_n) + g_n p_n h - 1\|_\alpha \\ &\geq \|g_n p_n h - 1\|_\alpha - \|p_n h\|_\alpha \|g - g_n\|_{M(\mathfrak{D}_\alpha)}.\end{aligned}$$

Since $\|p_n h\|_\alpha$ is bounded and $\|g - g_n\|_{M(\mathfrak{D}_\alpha)}$ decays exponentially, we obtain that there exists a constant C such that

$$\|p_n f - 1\|_\alpha \geq C \operatorname{dist}_{\mathfrak{D}_\alpha}(1, h \cdot \mathfrak{P}_{2n}). \qquad \square$$

Using Lemma 2.5, we obtain sharp estimates on the decay of norms.

Theorem 2.6. *Let $\alpha \leq 1$ and $g, h \in D_\alpha$. Suppose that g and h admit analytic continuations to $\overline{\mathbb{D}}$ and have no zeros in $\mathbb{D}$. Define $f(z_1, z_2) = g(z_1)h(z_2)$. Then there exists a constant $C = C(g, h, \alpha)$ such that*

$$\operatorname{dist}^2_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n) \leq C \varphi_\alpha^{-1}(n + 1),$$

for all sufficiently large n. Moreover, this estimate is sharp in the sense that if h has at least one zero on $\mathbb{T}$ and g has no zeros in the closed disk $\mathbb{D}$ (or vice versa), then there exists a constant $\tilde{C} = \tilde{C}(g, h, \alpha)$ such that

$$\tilde{C} \varphi_\alpha^{-1}(n + 1) \leq \operatorname{dist}^2_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n).$$

Proof. By Theorem 1.2, for any polynomials $p_n(z_1)$ and $q_n(z_2)$ of degree less than or equal to n, there exist constants C_1 and C_2 such that

$$\begin{aligned}\|p_n(z_1)g(z_1) - 1\|_{D_\alpha} &\leq C_1 \varphi_\alpha^{-1/2}(n + 1), \\ \|q_n(z_2)h(z_2) - 1\|_{D_\alpha} &\leq C_2 \varphi_\alpha^{-1/2}(n + 1).\end{aligned}$$

Therefore,

$$\begin{aligned}
&\|p_n(z_1)q_n(z_2)g(z_1)h(z_2)-1\|_\alpha \\
&\qquad\qquad \le \|q_n(z_2)h(z_2)(p_n(z_1)g(z_1)-1)\|_\alpha + \|q_n(z_2)h(z_2)-1\|_\alpha \\
&\qquad\qquad \le \|q_n h\|_\alpha \, \|p_n g-1\|_\alpha + \|q_n h-1\|_\alpha \\
&\qquad\qquad = \|q_n h\|_{D_\alpha} \, \|p_n g-1\|_{D_\alpha} + \|q_n h-1\|_{D_\alpha} \\
&\qquad\qquad \le (\|q_n h-1\|_{D_\alpha}+1) \, \|p_n g-1\|_{D_\alpha} + \|q_n h-1\|_{D_\alpha} \\
&\qquad\qquad \le C_2 C_1 \varphi_\alpha^{-1}(n+1) + (C_1+C_2)\varphi_\alpha^{-1/2}(n+1) \\
&\qquad\qquad \le C\varphi_\alpha^{-1/2}(n+1)
\end{aligned}$$

for some constant C. Therefore,

$$\operatorname{dist}^2_{\mathfrak{D}_\alpha}(1, f\cdot\mathfrak{P}_n) \le C\varphi_\alpha^{-1}(n+1),$$

for all sufficiently large n, as desired.

Moreover, the inequality is sharp. To see this, suppose h has at least one zero on $\mathbb{T}$ and g has no zeros in the closed unit disk. Then, by Lemma 2.5, there exists a constant C_1 such that

$$\operatorname{dist}_{\mathfrak{D}_\alpha}(1, f\cdot\mathfrak{P}_n) \ge C_1 \operatorname{dist}_{\mathfrak{D}_\alpha}(1, h\cdot\mathfrak{P}_{2n}). \tag{2-2}$$

Note that $h=h(z_2)$, and so, by orthogonality of monomials in $\mathfrak{D}_\alpha$, the quantity $\operatorname{dist}_{\mathfrak{D}_\alpha}(1, h\cdot\mathfrak{P}_{2n})$ is bounded from below by $\operatorname{dist}_{\mathfrak{D}_\alpha}(1, h\cdot\mathcal{P}_{2n}) = \operatorname{dist}_{D_\alpha}(1, h\cdot\mathcal{P}_{2n})$. Now, by Theorem 1.2 applied to h, and again since $\varphi_\alpha(2n+1)$ is comparable to $\varphi_\alpha(n+1)$, there exists a constant C_2 such that

$$\operatorname{dist}^2_{D_\alpha}(1, h\cdot\mathcal{P}_n) \ge C_2\varphi_\alpha^{-1}(n+1). \tag{2-3}$$

Thus, the inequalities in (2-2) and (2-3) imply the desired result. □

3. Norm comparisons and sharp decay of norms for the subspaces $\mathcal{J}_{\alpha,M,N}$

Let us now consider a third way of relating two variable cyclic functions to one variable cyclic functions. In particular, we shall show that the polynomials in (1-2) furnish optimal approximants for a certain subclass of functions.

The subspaces $\mathcal{J}_{\alpha,M,N}$. In order to formulate our results, we need some notation. For $-\infty<\alpha<\infty$ and integers $M, N\ge 1$, we consider the closed subspaces

$$\mathcal{J}_{\alpha,M,N} = \left\{ f\in\mathfrak{D}_\alpha : f=\sum_{k=0}^{\infty} a_k z_1^{Mk} z_2^{Nk} \right\}.$$

For instance, $\mathcal{J}_\alpha = \mathcal{J}_{\alpha,1,1}$ consists of the functions f whose Taylor coefficients $(a_{k,l})$ vanish off the diagonal $k = l$, meaning that $f(z_1, z_2) = f(z_1 \cdot z_2)$.

We shall write D_{α,z_1} for the set of functions in D_α in the variable z_1, viewed as a subspace of $\mathfrak{D}_\alpha$.

Theorem 3.1. *Let $\alpha \le \frac{1}{2}$ and suppose that $f \in \mathcal{J}_{\alpha,M,N}$ has the property that $R(f)(z) = f(z^{1/M}, 1)$ is a function that admits an analytic continuation to the closed unit disk, whose zeros lie in $\mathbb{C} \setminus \mathbb{D}$. Then, f is cyclic in $\mathfrak{D}_\alpha$, and there exists a constant $C = C(\alpha, f, M, N)$ such that*

$$\operatorname{dist}^2_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n) \le C\varphi^{-1}_{2\alpha}(n+1).$$

This result is sharp in the sense that, if $R(f)$ has at least one zero on $\mathbb{T}$, then there exists a constant $c = c(\alpha, f, M, N) > 0$ such that, for large n,

$$c\varphi^{-1}_{2\alpha}(n+1) \le \operatorname{dist}^2_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n).$$

The same conclusions remain valid for $f \in D_{\alpha,z_1}$, with the rate $\varphi^{-1}_{2\alpha}$ replaced by φ^{-1}_{α}.

We should point out that the hypotheses of Theorem 3.1 imply that f is nonvanishing in $\mathbb{D}^2$. Suppose $f \in \mathcal{J}_{\alpha,M,N}$ and $f(z_1, z_2) = 0$, for some $(z_1, z_2) \in \mathbb{D}^2$. Then, the function $R(f)$ will have a zero at $z = z_1^M z_2^N \in \mathbb{D}$.

Remark 3.2. It is straightforward to check that functions like $f(z_1, z_2) = 1 - z_1$, $f(z_1, z_2) = (1 - z_1 z_2)^N$, $N \in \mathbb{N}$, and $f(z_1, z_2) = z_1^2 z_2^2 - 2\cos\theta z_1 z_2 + 1$, $\theta \in \mathbb{R}$, satisfy the assumptions of Theorem 3.1.

The arguments used in the proof of Theorem 3.1 imply a function $f \in \mathcal{J}_{\alpha,M,N}$ can *fail* to be cyclic in $\mathfrak{D}_\alpha$ when $\alpha > \frac{1}{2}$. For instance, the function $f(z_1, z_2) = 1 - z_1 z_2$ is cyclic if and only if $\alpha \le \frac{1}{2}$ (see Example 2 below), and the Riesz polynomials (1-2) are optimal approximants to $1/f$ when $\alpha \le \frac{1}{2}$.

Liftings, restrictions, and norm comparisons. The proof of Theorem 3.1 ultimately relies on Theorem 1.2, and comparison between the norm of $\mathfrak{D}_\alpha$ and that of $D_{2\alpha}$.

Suppose that for some real α, the function $F = \sum_{k=0}^{\infty} a_k z^k$ belongs to D_α, a Dirichlet-type space on the unit disk. We define $E : D_\alpha \to \mathfrak{D}_\alpha$ by

$$E(F)(z_1, z_2) = F(z_1).$$

In addition, if $f \in D_{\alpha,z_1}$, the mapping $C : \mathfrak{D}_\alpha \to D_\alpha$ given by $C(f)(z) = f(z, 1)$ is well-defined, and we have $E \circ C|_{D_{\alpha,z_1}} = \mathrm{id}_{D_{\alpha,z_1}}$. Moreover, it is immediate that

$$\|E(F)\|_\alpha = \|F\|_{D_\alpha}, \quad F \in D_\alpha$$

and

$$\|f\|_\alpha = \|C(f)\|_{D_\alpha}, \quad f \in D_{\alpha,z_1}.$$

Another embedding is the following one. For $\alpha \in \mathbb{R}$ fixed, define the mappings

$$L_{M,N} : D_{2\alpha} \to \mathfrak{D}_\alpha \quad \text{via} \quad L_{M,N}(F)(z_1, z_2) = F(z_1^M \cdot z_2^N),$$

and

$$R_{M,N} : \mathcal{J}_{\alpha,M,N} \to D_{2\alpha} \quad \text{via} \quad R_{M,N}(f)(z) = f(z^{1/M}, 1).$$

We initially view $f(z^{1/M}, 1)$ as a formal expression, but the assumption that

$$\sum_k (k+1)^{2\alpha} |a_k|^2 < \infty$$

implies that $f(z_1^{1/M}, 1)$ is actually a well-defined holomorphic function on $\mathbb{D}$; this will become apparent below. By definition, we again have $L \circ R|_{\mathcal{J}_{\alpha,M,N}} = \mathrm{id}_{\mathcal{J}_{\alpha,M,N}}$.

Lemma 3.3. *For $F \in D_{2\alpha}$ and $f \in \mathcal{J}_{\alpha,M,N}$, there are constants $c_1 = c_1(\alpha, M, N)$ and $c_2 = c_2(\alpha, M, N)$ such that*

$$\|L_{M,N}(F)\|_\alpha \le c_1 \|F\|_{D_{2\alpha}} \quad \textit{and} \quad c_2 \|R(f)\|_{D_{2\alpha}} \le \|f\|_\alpha.$$

In particular, if $f \in \mathcal{J}_{\alpha,M,N}$, then

$$c_2 \|R(f)\|_{D_{2\alpha}} \le \|f\|_\alpha \le c_1 \|R(f)\|_{D_{2\alpha}}. \tag{3-1}$$

Proof. We provide the proof of the second inequality; the proof of the first is analogous.

We first observe that for any $\alpha \in \mathbb{R}$ and $M \ge 1$, there exist constants $c_1(\alpha, M)$ and $c_2(\alpha, M)$ such that

$$c_1(\alpha, M)(k+1)^\alpha \le (Mk+1)^\alpha \le c_2(\alpha, M)(k+1)^\alpha,$$

for any $k \in \mathbb{N}$. Thus, writing $R(f)(z) = \sum_{k=0}^\infty a_k z^k$, we have

$$\begin{aligned}
\|R(f)\|_{D_{2\alpha}}^2 &= \sum_{k=0}^\infty (k+1)^{2\alpha} |a_k|^2 = \sum_{k=0}^\infty (k+1)^\alpha (k+1)^\alpha |a_k|^2 \\
&\le [c_1(\alpha, M) c_1(\alpha, N)]^{-1} \sum_{k=0}^\infty (Mk+1)^\alpha (Nk+1)^\alpha |a_k|^2 \\
&= [c_1(\alpha, M) c_1(\alpha, N)]^{-1} \|f\|_\alpha^2,
\end{aligned}$$

which proves the assertion. The two-sided bound (3-1) follows from the one-sided bounds and the fact that $f = L(R(f))$. □

In particular, we see from the proof of Lemma 3.3 that in the case $M = N = 1$, the equalities

$$\|L(F)\|_\alpha = \|F\|_{D_{2\alpha}} \quad \text{and} \quad \|R(f)\|_{D_{2\alpha}} = \|f\|_\alpha$$

hold; hence, R is an isometric isomorphism between $\mathcal{J}_\alpha$ and $D_{2\alpha}$.

Sharpness of norm decay. We shall use Lemma 3.3, along with the following lemma, to prove Theorem 3.1.

Lemma 3.4. *Suppose that $f \in \mathcal{J}_{\alpha,M,N}$ for some $\alpha \in \mathbb{R}$ and some integers $M, N \geq 1$. Let $r_n = \sum_{k=0}^{n} \sum_{l=0}^{n} c_{k,l} z_1^k z_2^l$ be an arbitrary polynomial, and let s_n be its projection onto $\mathcal{J}_{\alpha,M,N}$,*

$$s_n = \sum_{\{k:\, Mk, Nk \leq n\}} c_{Mk,Nk} z_1^{Mk} z_2^{Nk}.$$

Then,

$$\|r_n f - 1\|_\alpha \geq \|s_n f - 1\|_\alpha.$$

Proof. We begin by noting again that monomials of the form $\{z_1^k z_2^l\}$ form an orthogonal basis for $\mathfrak{D}_\alpha$. Next, setting $\tilde{s}_n = r_n - s_n$, we have $s_n f \in \mathcal{J}_{\alpha,M,N}$, and $\tilde{s}_n f \notin \mathcal{J}_{\alpha,M,N}$. Then, by the previous observation, $s_n f - 1 \perp \tilde{s}_n f$.

This means that

$$\begin{aligned} \|r_n f - 1\|_\alpha^2 &= \|s_n f - 1 + \tilde{s}_n f\|_\alpha^2 \\ &= \|s_n f - 1\|_\alpha^2 + \|\tilde{s}_n f\|_\alpha^2 \\ &\geq \|s_n f - 1\|_\alpha^2. \end{aligned}$$ ◻

An analogous result holds for functions in the subspace D_{α,z_1}.

Proof of Theorem 3.1. We present the details for functions $f \in \mathcal{J}_\alpha$; the same type of arguments work for $\mathcal{J}_{\alpha,M,N}$, with the appropriate inequalities from Lemma 3.3 in place of equalities, and also for $f \in D_{\alpha,z_1}$.

We begin by establishing the lower bound. Let $r_n = \sum_k \sum_l c_{k,l} z_1^k z_2^l$ be any polynomial, and extract the diagonal part s_n from r_n as in the preceding lemma. Note that by construction, $s_n f - 1 \in \mathcal{J}_\alpha$ for each α. By Lemma 3.4 and the norm inequality (3-1), we obtain

$$\|r_n f - 1\|_\alpha \geq \|s_n f - 1\|_\alpha = \|R(s_n f - 1)\|_{D_{2\alpha}} = \|R(s_n)R(f) - 1\|_{D_{2\alpha}}.$$

It is assumed that $R(f)$ satisfies the hypotheses of Theorem 1.2; the theorem then asserts that $\operatorname{dist}^2_{D_{2\alpha}}(1, R(f)\cdot \mathcal{P}_n) \geq \tilde{C}\varphi_{2\alpha}^{-1}(n+1)$. In particular, this yields a lower bound for $\|R(s_n)R(f) - 1\|_{D_{2\alpha}}$, and the lower bound on $\operatorname{dist}_{\mathfrak{D}_\alpha}(1, f \cdot \mathfrak{P}_n)$ follows.

To obtain the upper bound, it is enough to exhibit a concrete sequence (p_n) of polynomials having $\|p_n f - 1\|_\alpha^2 \leq C(\alpha, f)\varphi_{2\alpha}^{-1}(n+1)$. However, since $R(f)$ satisfies the hypotheses of Theorem 1.2, there exists a sequence (q_n) of polynomials in one variable that achieves

$$\|q_n R(f) - 1\|_{D_{2\alpha}}^2 \leq C(\alpha, f)\varphi_{2\alpha}^{-1}(n+1)$$

for large enough n. But then we can define $p_n = L(q_n) \in \mathcal{J}_\alpha$, and the desired estimate follows since

$$\|L(q_n)f - 1\|_\alpha^2 = \|R(L(q_n))R(f) - 1\|_{D_{2\alpha}}^2 = \|q_n R(f) - 1\|_{D_{2\alpha}}^2$$

by Lemma 3.3. □

Note that if $R(f)$ is a polynomial with only simple zeros on the unit circle $\mathbb{T}$, then it is shown in [Bénéteau et al. 2015, Section 3] that the one-variable Riesz polynomials achieve the norm decay obtained above. In the situation $M = N = 1$ then, we have $L(q_n)(z_1, z_2) = p_n(z_1, z_2)$, where p_n are the Riesz-type polynomials defined in (1-2).

4. Polynomials with zeros on $\partial\mathbb{D}^2$ and measures of finite energy

Let us now examine the relationship between cyclicity and boundary zero sets of functions in $\mathfrak{D}_\alpha$. Surprisingly, some functions with large zero sets in some sense *are* cyclic while others with smaller zero sets are not.

Examples. Let us examine a few simple examples.

Example 1. Set $f(z_1, z_2) = 1 - z_1$. Then f has zero set

$$\mathcal{Z}(f) = \{1\} \times \overline{\mathbb{D}},$$

a (real) 2-dimensional subset of the topological boundary of $\mathbb{D}^2$ which meets the distinguished boundary along the 1-dimensional curve $\{1\} \times \mathbb{T}$. Note that f is an example of a function of the product type $g(z_1)h(z_2)$ with $g(z_1) = 1 - z_1$ and $h(z_2) = 1$, and therefore by Proposition 2.4, f is cyclic in $\mathfrak{D}_\alpha$ if and only if $\alpha \le 1$.

Example 2. Consider the function $f(z_1, z_2) = 1 - z_1 z_2$. The part of the zero set of f that lies on the boundary of the bidisk,

$$\mathcal{Z}(f) = \{(e^{i\theta}, e^{-i\theta}) : \theta \in [0, 2\pi)\},$$

can be seen as a 1-dimensional real curve contained in the distinguished boundary $\mathbb{T}^2$. One verifies that all the points in $\mathcal{Z}(f)$ are simple zeros. Since

$$\frac{1}{f(z_1, z_2)} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \delta_{k,l} z_1^k z_2^l = \sum_{k=0}^{\infty} z_1^k z_2^k,$$

we have $\|1/f\|_{-1}^2 = \sum_{k=0}^{\infty} (1+k)^{-2} < \infty$ but $\|1/f\|_0^2 = \sum_{k=0}^{\infty} 1 = +\infty$, so f is invertible in the Bergman space, and indeed in $\mathfrak{D}_\alpha$ whenever $\alpha < -\frac{1}{2}$, but not in the Hardy space of the bidisk.

Nevertheless, by Theorem 3.1, f is cyclic in $\mathfrak{D}_\alpha$ if and only if $\alpha \le \frac{1}{2}$. Note in particular that this function is *not* cyclic in the classical Dirichlet space of the bidisk!

Explicit computations with the Riesz polynomials in (1-2) recover the upper bound in Theorem 3.1. Namely, we have

$$p_n(z_1, z_2)f(z_1, z_2) - 1 = -\frac{1}{\varphi_\alpha(n+1)} \sum_{k=1}^{n+1} [\varphi_\alpha(k) - \varphi_\alpha(k-1)](z_1 z_2)^k,$$

and then, since $|\varphi_\alpha(k) - \varphi_\alpha(k-1)|^2 \le C(\alpha)(k-1)^{-2\alpha}$, we obtain

$$\|p_n f - 1\|_\alpha^2 \le \frac{C_1(\alpha)}{(n+1)^{1-2\alpha}}.$$

Thus $\|p_n f - 1\|_\alpha^2 \to 0$ as $n \to \infty$ and f is cyclic, provided that $\alpha \le \frac{1}{2}$.

In fact, considering functions of the form $f = 1 - z_1^M z_2^N$ for integer $M, N \ge 1$ instead, and performing the analogous computations, we obtain

$$\|p_n f - 1\|_\alpha^2 \le \frac{C_1(\alpha, M, N)}{(n+1)^{1-2\alpha}} \tag{4-1}$$

with a constant $C_1(\alpha, M, N)$ which does not depend on n.

Example 3. We examine $f(z_1, z_2) = 1 - z_1 - z_2 + z_1 z_2 = (1 - z_1)(1 - z_2)$. The zero set of f is

$$\mathcal{Z}(f) = (\{1\} \times \overline{\mathbb{D}}) \cup (\overline{\mathbb{D}} \times \{1\}),$$

a 2-dimensional set that extends into the topological boundary of the bidisk. Its intersection with $\mathbb{T}^2$ consists of the curves

$$\mathcal{Z}(f) = (\{1\} \times \mathbb{T}) \cup (\mathbb{T} \times \{1\}).$$

All zeros of f are simple, except the point $(1, 1)$, which has order 2. Since

$$\frac{1}{f(z_1, z_2)} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} z_1^k z_2^l,$$

it follows that $1/f \notin A^2(\mathbb{D}^2)$. Note that again, f is separable with $g(z_1) = 1 - z_1$ and $h(z_2) = 1 - z_2$, and therefore f is cyclic in $\mathfrak{D}_\alpha$ if and only if $\alpha \le 1$.

In this case, computing with the Riesz polynomials leads to misleading estimates. Defining polynomials p_n, as before, via (1-2), we compute

$$\begin{aligned} p_n f = &- \frac{1}{(n+1)^{1-\alpha}} \sum_{k=1}^{n+1} [k^{1-\alpha} - (k-1)^{1-\alpha}](z_1^k + z_2^k) \\ &+ \frac{1}{(n+1)^{1-\alpha}} \sum_{k=1}^{n+1} [k^{1-\alpha} - (k-1)^{1-\alpha}] z_1^k z_2^k. \end{aligned}$$

We use the estimates from the previous example, and exploit the one-variable estimates from [Bénéteau et al. 2015], to obtain

$$\begin{aligned}\|p_n f - 1\|^2_{\mathfrak{D}_\alpha} &= \frac{2}{(n+1)^{2-2\alpha}} \sum_{k=1}^{n+1} (k+1)^\alpha [k^{1-\alpha} - (k-1)^{1-\alpha}]^2 \\ &\quad + \frac{1}{(n+1)^{2-2\alpha}} \sum_{k=1}^{n+1} (k+1)^{2\alpha} [k^{1-\alpha} - (k-1)^{1-\alpha}]^2 \\ &\leq \frac{c_1(\alpha)}{(n+1)^{1-\alpha}} + \frac{c_2(\alpha)}{(n+1)^{1-2\alpha}}.\end{aligned}$$

The first term in the right-hand side dominates when $\alpha < 0$, whereas the second is larger when $\alpha > 0$. In particular, the estimate does show that f is cyclic in $\mathfrak{D}_\alpha$ provided $\alpha \leq \frac{1}{2}$. However, as we have seen, the rate is not optimal, and f remains cyclic when $\alpha > \frac{1}{2}$.

Note the interesting contrast between Example 2 and Example 3: the function in Example 2 is not cyclic in the (classical) Dirichlet space of the bidisk, and yet in some sense has a much smaller zero set than the function in Example 3, which is cyclic! On the other hand, as a kind of dual phenomenon, $f = 1 - z_1 z_2$ exhibits a faster rate of decay of norms $\|p_n f - 1\|_\alpha$ for $\alpha < 0$ than does $f = (1 - z_1)(1 - z_2)$.

Example 4. The polynomial $f(z_1, z_2) = 1 - (z_1 + z_2)/2$ has no zeros in $\mathbb{D}^2$, and vanishes at a single boundary point: $\mathcal{Z}(f) = \{(1, 1)\} \subset \mathbb{T}^2$.

In [Hedenmalm 1988, Section 4], it is proved that if $f \in \mathfrak{D}_2$ has $\mathcal{Z}(f) = \{(1, 1)\}$, and both $f(\,\cdot\,, 1)$ and $f(1, \cdot\,)$ are outer functions, then the closure of the principal ideal generated by f coincides with the closed ideal

$$\mathcal{I}\big(\{(1, 1)\}\big) = \{f \in \mathfrak{D}_2 : f(1, 1) = 0\}.$$

(Hedenmalm's norm is defined using the weights $(1 + k^2)(1 + l^2)$ but is equivalent to the norm in $\mathfrak{D}_2$.) Since the norm of $\mathfrak{D}_1$ is weaker than that of $\mathfrak{D}_2$, it follows that such functions are cyclic in $\mathfrak{D}_\alpha$ for $\alpha \leq 1$ as the $\mathfrak{D}_1$-closure of the invariant subspace $\mathcal{I}\big(\{(1, 1)\}\big) \subset \mathfrak{D}_2$ coincides with $[f]$, and contains the cyclic function $1 - z_1$.

In particular, the polynomial $f(z_1, z_2) = 1 - (z_1 + z_2)/2$ is cyclic in $\mathfrak{D}_\alpha$, for all $\alpha \leq 1$. (An independent proof of this fact has been given by T. J. Ransford [personal communication, 2014].) Computing with polynomials of the form

$$p_n(z_1, z_2) = \sum_{k=0}^{n} \left(1 - \frac{\varphi_\alpha(k)}{\varphi_\alpha(n+1)}\right) \frac{(z_1 + z_2)^k}{2^k}$$

and using the fact that $(z_1+z_2)^{k_1} \perp (z_1+z_2)^{k_2}$ when $k_1 \neq k_2$, one finds that

$$\|p_n f - 1\|_\alpha^2 = \sum_{k=1}^{n+1} 4^{-k} \left(\frac{k^{1-\alpha} - (k-1)^{1-\alpha}}{(n+1)^{1-\alpha}}\right)^2 \sum_{j=0}^{k} \binom{k}{j}^2 (j+1)^\alpha (k-j+1)^\alpha.$$

Using the bound

$$(j+1)^\alpha (k-j+1)^\alpha \le C(k+1)^{2\alpha}, \quad 0 \le j \le k,$$

together with the identity

$$\sum_{j=0}^{k} \binom{k}{j}^2 = \binom{2k}{k}$$

and standard estimates on binomial coefficients, we obtain the estimate

$$\operatorname{dist}^2_{\mathfrak{D}_\alpha}(1, (2 - z_1 - z_2) \cdot \mathfrak{P}_n) \le C\varphi_{2\alpha-1/2}(n+1).$$

Unfortunately, we have not been able to obtain a sharp estimate, but the above bound shows that the optimal rate is different from the two rates we have seen previously.

Measures of finite energy. It would be interesting to understand the relationship between cyclicity and boundary zero sets — in particular, given a function f, to find a measure whose support lies on the zero set of the boundary values of f that relates to the cyclicity properties of f.

We now give a necessary condition for a function to be cyclic. This condition involves the notion of capacity, and represents a straightforward generalization of results of Brown and Shields in the one-variable case.

Definition 4.1. Let $E \subset \mathbb{T}^2$ be a Borel set. We say that a probability measure μ supported in E has *finite logarithmic energy* if

$$I[\mu] = \int_{\mathbb{T}^2} \int_{\mathbb{T}^2} \log \frac{e}{|e^{i\theta_1} - e^{i\vartheta_1}|} \log \frac{e}{|e^{i\theta_2} - e^{i\vartheta_2}|} \, d\mu(\theta_1, \theta_2)\, d\mu(\vartheta_1, \vartheta_2) < \infty.$$

If E supports no such measure, we say that E has *logarithmic capacity* 0.

The integral defining the energy $I[\mu]$ can be seen as a convolution with kernel

$$h(s,t) = \log \frac{e}{|1 - e^{is}|} \log \frac{e}{|1 - e^{it}|}.$$

Replacing the logarithmic product kernel in the definitions above with

$$h_\alpha(s,t) = \frac{1}{|1 - e^{is}|^{1-\alpha}} \frac{1}{|1 - e^{it}|^{1-\alpha}},$$

one obtains the notions of *Riesz energy*, denoted by $I_\alpha[\mu]$, and *Riesz capacity* of order $0 < \alpha < 1$.

The α-energy of μ can be expressed in terms of its Fourier coefficients

$$\hat{\mu}(k,l) = \int_{\mathbb{T}^2} e^{-i(k\theta_1 + l\theta_2)}\, d\mu(\theta_1,\theta_2), \quad k,l \in \mathbb{Z}.$$

Namely, we have (compare [El-Fallah et al. 2014, Chapter 2], for instance)

$$I_\alpha[\mu] = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \hat{h}_\alpha(k,l)|\hat{\mu}(k,l)|^2.$$

Computing the Fourier coefficients $\hat{h}(k,l)$ (see [Brown and Shields 1984, p. 294] for details), we find that

$$I[\mu] = 1 + \sum_{k=1}^{\infty} \frac{|\hat{\mu}(k,0)|^2}{k} + \sum_{l=1}^{\infty} \frac{|\hat{\mu}(0,l)|^2}{l} + \frac{1}{2} \sum_{k\in\mathbb{Z}\setminus\{0\}} \sum_{l=1}^{\infty} \frac{|\hat{\mu}(k,l)|^2}{|k|l}. \tag{4-2}$$

Similarly, one can show (again see [El-Fallah et al. 2014, Chapter 2]) that the Fourier coefficients of h_α satisfy

$$c_1(|k|+1)^{-\alpha}(|l|+1)^{-\alpha} \le |\hat{h}_\alpha(k,l)| \le c_2(|k|+1)^{-\alpha}(|l|+1)^{-\alpha}$$

for some constants $0 < c_1 < c_2 < \infty$.

The notion of energy now allows us to identify some noncyclic $f \in \mathfrak{D}_\alpha$ by looking at their boundary zero sets. To make this notion precise, we note that one can show that functions $f \in \mathfrak{D}_\alpha$ have radial limits

$$f^*(e^{i\theta_1}, e^{i\theta_2}) = \lim_{r\to 1^-} f(re^{i\theta_1}, re^{i\theta_2})$$

quasi-everywhere with respect to the appropriate capacity. That is, the limit exists for all points outside a set of capacity 0, and hence it makes sense to speak of the capacity of the set $\mathcal{Z}(f^*)$. (In fact, Kaptanoğlu considers more general approach regions in [Kaptanoğlu 1994], but we do not need this here.)

Proposition 4.2. *If $f \in \mathfrak{D}$ and $\mathcal{Z}(f^*)$ has positive logarithmic capacity, then f is not cyclic.*

Proof. The proof is completely analogous to that of [Brown and Shields 1984, Theorem 5]; we refer the reader to the paper of Brown and Shields for details and present the arguments in condensed form here.

The key idea is to identify the Bergman space $A^2(\mathbb{D}^2)$ with the dual of $\mathfrak{D}$ via the pairing

$$\langle f, g\rangle = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} a_{k,l} b_{k,l},$$

where $f = \sum_{k,l} a_{k,l} z_1^k z_2^l \in \mathfrak{D}$ and $g = \sum_{k,l} b_{k,l} z_1^k z_2^l \in A^2(\mathbb{D}^2)$. We then consider the

Cauchy integral

$$C[\mu] = \int_{\mathbb{T}^2} (1 - e^{i\theta_1} z_1)^{-1} (1 - e^{i\theta_2} z_2)^{-1}\, d\mu(\theta_1, \theta_2)$$

of μ, a measure of finite logarithmic energy with $\operatorname{supp}(\mu) \subset \mathcal{Z}(f^*)$. A comparison with (4-2) then reveals that

$$\|C[\mu]\|^2_{A^2(\mathbb{D}^2)} = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{|\hat{\mu}(k,l)|^2}{(k+1)(l+1)} < \infty$$

so that $C[\mu]$ induces a nontrivial element of $\mathfrak{D}^*$. On the other hand, since the measure μ is supported on $\mathcal{Z}(f^*)$ by assumption, the functional induced by $C[\mu]$ annihilates $[f]$, and so f is not cyclic. □

For $0 < \alpha < 1$, the same result holds once we replace logarithmic capacity with Riesz capacity and make the identification $(\mathfrak{D}_\alpha)^* = \mathfrak{D}_{-\alpha}$ in the proof.

The argument used in the proof of Proposition 4.2 can be used to give another proof of the noncyclicity of the function $f(z_1, z_2) = 1 - z_1 z_2$ in $\mathfrak{D}$. Namely, consider the probability measure $\mu_{\mathcal{Z}}$ on $\mathbb{T}^2$ induced by the (normalized) *integration current* associated with the variety $\mathcal{Z}(1 - z_1 z_2) \cap \mathbb{T}^2$ (see [Lelong and Gruman 1986, Chapter 2] for the relevant definitions). A quick computation reveals that $\hat{\mu}_{\mathcal{Z}}(k, l) = \delta_{kl}$, so that $C[\mu_{\mathcal{Z}}](z_1, z_2) = 1/(1 - z_1 z_2)$, a function in the Bergman space of the bidisk which satisfies

$$\langle z_1^k z_2^l f, C[\mu_{\mathcal{Z}}] \rangle = 0, \quad \text{for all } k, l \geq 0.$$

In fact, $\mathcal{Z}(1 - z_1 z_2) \cap \mathbb{T}^2$ has positive Riesz capacity precisely when $\alpha > \frac{1}{2}$.

5. Concluding remarks and open problems

It appears to be a difficult task to characterize the cyclic elements of $\mathfrak{D}_\alpha$ for $\alpha \leq 1$, and many basic questions remain. For instance, it is natural to ask whether the *Brown–Shields conjecture* is true for functions on the bidisk.

Problem 5.1. Is the condition that $f \in \mathfrak{D}$ is outer and $\mathcal{Z}(f^*)$ has logarithmic capacity 0 sufficient for f to be cyclic?

This question remains open for the Dirichlet space of the unit disk, and is widely considered to be a challenging problem. A first step towards understanding cyclic functions in $\mathfrak{D}_\alpha$ might be to solve the following natural problem.

Problem 5.2. Characterize the *cyclic polynomials* $f \in \mathfrak{D}_\alpha$ for each $\alpha \in (0, 1]$.

An obvious necessary condition for f to be cyclic is that $\mathcal{Z}(f) \cap \mathbb{D}^2 = \varnothing$, and if f is a polynomial that does not vanish in $\overline{\mathbb{D}^2}$, then f is cyclic because both f and $1/f$ extend analytically to a larger polydisk. But the problem appears to be open for polynomials with $\mathcal{Z}(f) \cap \partial\mathbb{D}^2 \neq \varnothing$: we would at least like to identify the polynomials whose zero sets have positive capacity. We have proved that polynomials that are products of polynomials in one variable are cyclic, and so the zero sets associated with such functions must all have zero capacity.

As we have seen in our examples, it can happen that a polynomial with a larger zero set, in the topological sense and in the sense of measure, is cyclic in $\mathfrak{D}_\alpha$ for some α, while a polynomial with a smaller zero set is not. We have also noted that a polynomial that *fails* to be cyclic in $\mathfrak{D}_\alpha$ when $\alpha > \frac{1}{2}$ can be "more" cyclic in $\mathfrak{D}_\alpha$, for $\alpha < 0$, than polynomials that are cyclic in all $\mathfrak{D}_\alpha$. We mean this in the sense that

$$\operatorname{dist}^2_{\mathfrak{D}_\alpha}(1, (1 - z_1 z_2) \cdot \mathfrak{P}_n) \asymp C\varphi^{-1}_{2\alpha}(n+1)$$

while

$$\operatorname{dist}^2_{\mathfrak{D}_\alpha}(1, (1 - z_1)(1 - z_2) \cdot \mathfrak{P}_n) \asymp C\varphi^{-1}_{\alpha}(n+1).$$

It would be interesting to develop a rigorous understanding of this phenomenon.

Acknowledgments

The authors wish to thank Thomas Ransford and Stefan Richter for interesting conversations and helpful suggestions. Thanks are due to the referee for a careful reading of the paper.

References

[Bénéteau et al. 2015] C. Bénéteau, A. A. Condori, C. Liaw, D. Seco, and A. A. Sola, "Cyclicity in Dirichlet-type spaces and extremal polynomials", *J. Analyse Math.* **126** (2015), 259–286. arXiv 1301.4375

[Brown and Shields 1984] L. Brown and A. L. Shields, "Cyclic vectors in the Dirichlet space", *Trans. Amer. Math. Soc.* **285**:1 (1984), 269–303. MR 86d:30079 Zbl 0517.30040

[Douglas and Yang 2000] R. G. Douglas and R. Yang, "Operator theory in the Hardy space over the bidisk (I)", *Integral Equations Operator Theory* **38**:2 (2000), 207–221. MR 2002m:47006 Zbl 0970.47016

[El-Fallah et al. 2014] O. El-Fallah, K. Kellay, J. Mashreghi, and T. Ransford, *A primer on the Dirichlet space*, Cambridge Tracts in Mathematics **203**, Cambridge University Press, 2014. MR 3185375 Zbl 1304.30002

[Gelca 1995] R. Gelca, "Rings with topologies induced by spaces of functions", *Houston J. Math.* **21**:2 (1995), 395–405. MR 96h:46075 Zbl 0849.46036 arXiv funct-an/9604006

[Hedenmalm 1988] H. Hedenmalm, "Outer functions in function algebras on the bidisc", *Trans. Amer. Math. Soc.* **306**:2 (1988), 697–714. MR 90c:32007 Zbl 0655.32017

[Hörmander 1990] L. Hörmander, *An introduction to complex analysis in several variables*, 3rd ed., North-Holland Mathematical Library **7**, North-Holland, Amsterdam, 1990. MR 91a:32001 Zbl 0685.32001

[Horowitz and Oberlin 1975] C. Horowitz and D. M. Oberlin, "Restriction of H^p functions to the diagonal of U^n", *Indiana Univ. Math. J.* **24**:8 (1975), 767–772. MR 50 #7583 Zbl 0282.32001

[Jupiter and Redett 2006] D. Jupiter and D. Redett, “Multipliers on Dirichlet type spaces”, *Acta Sci. Math.* (*Szeged*) **72**:1-2 (2006), 179–203. MR 2007f:47023 Zbl 1174.46320 arXiv math/0510193

[Kaptanoğlu 1994] H. T. Kaptanoğlu, “Möbius-invariant Hilbert spaces in polydiscs”, *Pacific J. Math.* **163**:2 (1994), 337–360. MR 94m:46044 Zbl 0791.32015

[Lelong and Gruman 1986] P. Lelong and L. Gruman, *Entire functions of several complex variables*, Grundlehren der Mathematischen Wissenschaften **282**, Springer, Berlin, 1986. MR 87j:32001 Zbl 0583.32001

[Mandrekar 1988] V. Mandrekar, “The validity of Beurling theorems in polydiscs”, *Proc. Amer. Math. Soc.* **103**:1 (1988), 145–148. MR 90c:32008 Zbl 0658.47033

[Massaneda and Thomas 2013] X. Massaneda and P. J. Thomas, “Non cyclic functions in the Hardy space of the bidisc with arbitrary decrease”, preprint, 2013. arXiv 1301.2622

[Redett and Tung 2010] D. Redett and J. Tung, “Invariant subspaces in Bergman space over the bidisc”, *Proc. Amer. Math. Soc.* **138**:7 (2010), 2425–2430. MR 2011d:47021 Zbl 1193.47013

[Richter and Sundberg 2012] S. Richter and C. Sundberg, “Cyclic vectors in the Drury–Arveson space”, slides, 2012, http://math.utk.edu/~richter/talk/Richter_Cyclic_Drury_Arveson2012.pdf.

[Rudin 1969] W. Rudin, *Function theory in polydiscs*, W. A. Benjamin, New York, 1969. MR 41 #501 Zbl 0177.34101

[Stegenga 1980] D. A. Stegenga, “Multipliers of the Dirichlet space”, *Illinois J. Math.* **24**:1 (1980), 113–139. MR 81a:30027 Zbl 0432.30016

[Taylor 1966] G. D. Taylor, “Multipliers on D_α”, *Trans. Amer. Math. Soc.* **123** (1966), 229–240. MR 34 #6514 Zbl 0166.12003

Received October 15, 2013. Revised September 10, 2014.

CATHERINE BÉNÉTEAU
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTH FLORIDA
4202 E. FOWLER AVENUE
TAMPA, FL 33620-5700
UNITED STATES
cbenetea@usf.edu

ALBERTO A. CONDORI
DEPARTMENT OF MATHEMATICS
FLORIDA GULF COAST UNIVERSITY
10501 FGCU BOULEVARD SOUTH
FORT MYERS, FL 33965-6565
UNITED STATES
acondori@fgcu.edu

CONSTANZE LIAW
CASPER AND DEPARTMENT OF MATHEMATICS
BAYLOR UNIVERSITY
ONE BEAR PLACE #97328
WACO, TX 76798-7328
UNITED STATES
constanze_liaw@baylor.edu

DANIEL SECO
MATHEMATICS INSTITUTE
UNIVERSITY OF WARWICK
ZEEMAN BUILDING
COVENTRY
CV4 7AL
UNITED KINGDOM
d.seco@warwick.ac.uk

ALAN A. SOLA
CENTRE FOR MATHEMATICAL SCIENCES
UNIVERSITY OF CAMBRIDGE
WILBERFORCE ROAD
CAMBRIDGE
CB3 0WB
UNITED KINGDOM
a.sola@statslab.cam.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.59

COMPACTNESS RESULTS FOR SEQUENCES OF APPROXIMATE BIHARMONIC MAPS

CHRISTINE BREINER AND TOBIAS LAMM

We will prove energy quantization for approximate (intrinsic and extrinsic) biharmonic maps into spheres where the approximate map is in $L \log L$. Moreover, we demonstrate that if the $L \log L$ norm of the approximate maps does not concentrate, the images of the bubbles are connected without necks.

1. Introduction

Critical points to the Dirichlet energy

$$E(u) := \frac{1}{2} \int_{\Omega} |Du|^2 \, dx$$

are called *harmonic maps*, and the compactness theory for such a sequence in two dimensions is well understood. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain and N a smooth, compact Riemannian manifold. For a sequence of harmonic maps $u_k \in W^{1,2}(\Omega, N)$ with uniform energy bounds, Sacks and Uhlenbeck [1981] proved that a subsequence u_k converges weakly to a harmonic u_∞ on Ω and $u_k \to u_\infty$ in $C^\infty(\Omega \setminus \{x_1, \ldots, x_\ell\})$ for some finite ℓ depending on the energy bound. For each x_i, they showed that there exist some number of "bubbles", maps $\phi_{ij} : \mathbb{S}^2 \to N$, that result from appropriate conformal scalings of the sequence u_k near x_i. In dimension 2, $E(u)$ is conformally invariant and thus one can ask whether any energy is lost in the limit. Jost [1991] proved that in fact the energy is quantized; there is no unaccounted energy loss:

$$\lim_{k \to \infty} E(u_k) = E(u_\infty) + \sum_{i=1}^{\ell} \sum_{j=1}^{\ell_i} E(\phi_{ij}).$$

Parker [1996] provided the complete description of the C^0 limit or "bubble tree". In particular, he demonstrated that the images of the limiting map u_∞ and the bubbles ϕ_{ij} are connected without necks. Around the same time, various authors

Breiner was supported in part by NSF grant DMS-1308420 and an AMS-Simons Travel Grant.
MSC2010: primary 35J60, 35J48; secondary 58E20.
Keywords: harmonic maps, biharmonic maps, bubbling, energy quantization.

proved energy quantization and the no-neck property for approximate harmonic maps [Ding and Tian 1995; Wang 1996; Qing and Tian 1997; Lin and Wang 1998; Chen and Tian 1999].

In this paper, we are interested in an analogous compactness problem for a scale-invariant energy in four dimensions. Let (M^4, g) and (N^k, h) be compact Riemannian manifolds without boundary, with N^k isometrically embedded in some $\mathbb{R}^n$. Consider the energy functional

$$E_{\mathrm{ext}}(u) := \int_M |\Delta u|^2 \, dx$$

for $u \in W^{2,2}(M, N)$, where Δ is the Laplace–Beltrami operator. Critical points to this functional are called *extrinsic biharmonic maps*, and the Euler–Lagrange equation satisfied by such maps is of fourth order. Clearly, this functional depends upon the immersion of N into $\mathbb{R}^n$. To avoid such a dependence, one may instead consider critical points to the functional

$$E_{\mathrm{int}}(u) := \int_M \left|(\Delta u)^T\right|^2 dx,$$

where $(\Delta u)^T$ is the projection of Δu onto $T_u N$. Critical points to this functional are called *intrinsic biharmonic maps*. The Euler–Lagrange equations satisfied by extrinsic and intrinsic biharmonic maps have been computed (see, for instance, [Wang 2004b]). We will be interested in approximate critical points.

Definition 1.1. Let $u \in W^{2,2}(B_1, N)$, where $B_1 \subset \mathbb{R}^4$ and N is a C^3 closed submanifold of some $\mathbb{R}^n$. Let $f \in L \log L(B_1, \mathbb{R}^n)$. Then u is an *f-approximate biharmonic map* if

$$\Delta^2 u - \Delta(A(u)(Du, Du)) - 2d^*\langle \Delta u, DP(u)\rangle + \langle \Delta(P(u)), \Delta u\rangle = f.$$

We call u an *f-approximate intrinsic biharmonic map* if

$$\begin{aligned}\Delta^2 u - \Delta(A(u)(Du, Du)) &- 2d^*\langle \Delta u, DP(u)\rangle \\ &+ \langle \Delta(P(u)), \Delta u\rangle - P(u)(A(u)(Du, Du) D_u A(u)(Du, Du)) \\ &\qquad - 2A(u)(Du, Du)A(u)(Du, DP(u)) = f.\end{aligned}$$

Here A is the second fundamental form of $N \hookrightarrow \mathbb{R}^n$ and $P(u) : \mathbb{R}^n \to T_u N$ is the orthogonal projection from $\mathbb{R}^n$ to the tangent space of N at u.

Recently, Hornung and Moser [2012], Laurain and Rivière [2013], and Wang and Zheng [2012] determined the energy quantization result for sequences of intrinsic biharmonic maps, approximate intrinsic and extrinsic biharmonic maps, and approximate extrinsic biharmonic maps, respectively. (In fact, the result of

[Laurain and Rivière 2013] applies to a broader class of solutions to scaling-invariant variational problems in dimension four.)

As a first result, we demonstrate that when the target manifold is a sphere, the energy quantization result extends to f-approximate biharmonic maps with $f \in L \log L$. For the definition of this Banach space, see the appendix.

Theorem 1.2. *Let $f_k \in L \log L(B_1, \mathbb{R}^{n+1})$ and $u_k \in W^{2,2}(B_1, \mathbb{S}^n)$ a sequence of f_k-approximate biharmonic maps with*

$$\|D^2 u_k\|_{L^2(B_1)} + \|Du_k\|_{L^4(B_1)} + \|f_k\|_{L \log L(B_1)} \leq \Lambda < \infty. \tag{1-1}$$

If $u_k \rightharpoonup u$ weakly in $W^{2,2}(B_1, \mathbb{S}^n)$, there exists $\{x_1, \ldots, x_\ell\} \subset B_1$ such that $u_k \to u$ in $W^{2,2}_{\mathrm{loc}}(B_1 \setminus \{x_1, \ldots, x_\ell\}, \mathbb{S}^n)$.

Moreover, for each $1 \leq i \leq \ell$ there exists an $\ell_i \in \mathbb{N}$ and nontrivial, smooth biharmonic maps $\omega_{ij} \in C^\infty(\mathbb{R}^4, \mathbb{S}^n)$ with finite energy $(1 \leq j \leq \ell_i)$ such that

$$\lim_{k\to\infty} \int_{B_{r_i}(x_i)} |D^2 u_k|^2 = \int_{B_{r_i}(x_i)} |D^2 u|^2 + \sum_{j=1}^{\ell_i} \int_{\mathbb{R}^4} |D^2 \omega_{ij}|^2,$$

$$\lim_{k\to\infty} \int_{B_{r_i}(x_i)} |Du_k|^4 = \int_{B_{r_i}(x_i)} |Du|^4 + \sum_{j=1}^{\ell_i} \int_{\mathbb{R}^4} |D\omega_{ij}|^4.$$

Here $r_i = \frac{1}{2} \min_{1\leq j\leq \ell, j\neq i}\{|x_i - x_j|, \operatorname{dist}(x_i, \partial B_1)\}$.

As a second result, we demonstrate the no-neck property for approximate biharmonic maps with the approximating functions $L \log L$ norm not concentrating.

Theorem 1.3. *Let $f_k \in L \log L$ such that the $L \log L$ norm does not concentrate. For u_k a sequence of f_k-approximate biharmonic maps satisfying* (1-1), *the images of u and the maps ω_{ij} described above are connected in $\mathbb{S}^n$ without necks.*

In particular, if $f_k \in \phi(L)$, an Orlicz space such that $\lim_{t\to\infty} \phi(t)/(t \log t) = \infty$, the theorem holds. For a definition of an Orlicz space, see the appendix.

Remark 1.4. The theorems also hold for u_k a sequence of f_k-approximate intrinsic biharmonic maps. We will prove the theorems in detail for f_k-approximate biharmonic maps, and point out the necessary changes one must make to prove the intrinsic case.

We consider biharmonic maps into spheres because the symmetry of the target provides structure for the equation that can be exploited to prove higher regularity. For an f-approximate biharmonic map into $\mathbb{S}^n$, the structural equations takes the form (see [Wang 2004a])

$$d^*(D\Delta u \wedge u - \Delta u \wedge Du) = f \wedge u, \tag{1-2}$$

and, for an f-approximate intrinsic biharmonic u,

$$d^*(D\Delta u \wedge u - \Delta u \wedge Du + 2|Du|^2 Du \wedge u) = f \wedge u. \tag{1-3}$$

The structure of the equation for harmonic maps from a compact Riemann surface into $\mathbb{S}^n$ was determined independently by Chen [1989] and Shatah [1988]. They demonstrated that u satisfies the conservation law

$$d^*(Du \wedge u) = 0.$$

Hélein [1990] used the structure of this equation and Wente's inequality [1969] to determine that any weakly harmonic $u \in W^{1,2}$ was in fact C^∞.

Li and Zhu [2011] used this additional structure to determine energy quantization for approximate harmonic maps. In their setting, the equation takes the form $d^*(Du \wedge u) = \tau \wedge u$ for $\tau \in L \log L$. Our proof of energy quantization is similar in spirit to their work and to the recent small-energy compactness result of Sharp and Topping [2013]. Of critical importance are the energy estimates we prove in Section 2. The first estimates, from Proposition 2.1, are used in two ways. First, the L^p estimates of (2-2), (2-3) provide sufficient control to determine a small-energy compactness result away from the bubbles. Second, we use the Lorentz space duality to prove energy quantization and thus require uniform bounds on the appropriate Lorentz energies as in (2-1). In Section 3 we prove the energy quantization result. We point out that since the oscillation bound contains an energy term of the form $\|D\Delta u_k\|_{L^{4/3}}$, we must also prove this energy is quantized. This point justifies the necessity of the estimate (2-4). We prove the energy quantization result, under the presumption of the occurrence of one bubble, in Proposition 3.4.

We next use this stronger energy quantization result for maps into spheres to prove a no-neck property. Zhu [2012] showed the no-neck property for approximate harmonic maps with τ in a space essentially between L^p with $p > 1$ and $L \log L$. For w a cutoff function of the approximate harmonic map u, Zhu considered a Hodge decomposition of the 1-form $\beta := Dw \wedge u$. (This is actually a matrix of 1-forms, but we gloss over that point for now.) He bounded $\|\beta\|_{L^{2,1}}$ by bounding each component of the decomposition, and used this to bound $\|Dw\|_{L^{2,1}}$ by $\|Du\|_{L^2}$ plus a norm of the torsion term, τ. Using ε-compactness and a simple duality argument, he showed the oscillation of u is controlled by $\|Dw\|_{L^{2,1}}$, which in turn implies the desired result.

Like Zhu, we prove the no-neck property by demonstrating that the oscillation of an f-approximate biharmonic map is controlled by norms that tend to zero in the neck region. Using a duality argument, we first determine that the oscillation of u on an annular region is bounded by quantized energy terms plus a third derivative of a cutoff function w. Our main work is in determining an appropriate estimate for $\|D\Delta w\|_{L^{4/3,1}}$. We determine this bound by considering the 1-form

$\beta = D\Delta w \wedge u - \Delta w \wedge Du$, and we bound $D\Delta w$ by bounding β via its Hodge decomposition. In particular, we take advantage of the divergence structure of the equation for biharmonic maps into spheres to show that β not only has good $L^{4/3}$ estimates but in fact has good estimates in $L^{4/3,1}$. This second estimate allows us to prove the necessary oscillation lemma. The proof of the oscillation lemma constitutes the work of Section 4. Coupling the oscillation lemma with energy quantization, we prove Theorem 1.3 in Section 5.

Finally, the arguments we use require a familiarity with Lorentz spaces and the appropriate embedding theorems relevant in dimension four. In the appendix, we describe the various Banach spaces and collect the necessary embeddings and estimates.

Many steps of the proof require the use of cutoff functions, so we set:

Definition 1.5. Let $\phi \in C_0^\infty(B_2)$ with $\phi \equiv 1$ in B_1. For all $r > 0$ set $\phi_r(x) = \phi(x/r)$.

Note added in proof: As we finalized the paper, we noticed a somewhat related preprint [Liu and Yin 2013], in which the authors claim that the no-neck property holds for sequences of biharmonic maps into general targets. Their methods are quite different from ours and we believe our results are of independent interest.

2. Energy estimates

To establish strong convergence away from points of energy concentration, we first prove the necessary energy estimates. The small-energy compactness result relies on the fact that in both (2-2) and (2-3) there is an extra power of the energy on the right-hand side of the inequality. Thus, small energy implies that $\|Du_k\|_{L^4}$ and $\|D^2u_k\|_{L^2}$ must converge to zero on small balls. Measure-theoretic arguments in the next section will then imply strong convergence for these norms to some Du and D^2u respectively.

Proposition 2.1. *Let $u \in W^{2,2}(B_2, \mathbb{S}^n)$ be an f-approximate (intrinsic) biharmonic map, where $f \in L\log L(B_2, \mathbb{R}^{n+1})$. Then there exists $C > 0$ such that*

$$\begin{aligned}\text{(2-1)}\quad \|D^3u\|_{L^{4/3,1}(B_1)} + \|D^2u\|_{L^{2,1}(B_1)} + \|Du\|_{L^{4,1}(B_1)} \\ \le C(\|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^2(B_2)} + \|Du\|_{L^2(B_2)} + \|f\|_{L\log L(B_2)}).\end{aligned}$$

Moreover, there exists $\tilde{\varepsilon} > 0$ such that, if

$$\|D^2u\|_{L^2(B_2)} + \|Du\|_{L^4(B_2)} < \tilde{\varepsilon},$$

then, for every $0 < r < \frac{1}{2}$,

$$\begin{aligned}\text{(2-2)}\quad \|D^2u\|^2_{L^2(B_r)} \le Cr^2\|D^2u\|^2_{L^2(B_2)} \\ + C(\|D^2u\|^4_{L^2(B_2)} + \|Du\|^4_{L^4(B_2)} + \|f\|^2_{L^1(B_2)}\|f\|_{L\log L(B_2)}),\end{aligned}$$

$$\begin{aligned}(2\text{-}3)\quad \|Du\|^4_{L^4(B_r)} &\le Cr^4\|Du\|^4_{L^2(B_2)}\\ &\quad + C(\|D^2u\|^8_{L^2(B_2)} + \|Du\|^8_{L^4(B_2)} + \|f\|^3_{L^1(B_2)}\|f\|_{L\log L(B_2)}),\\ (2\text{-}4)\quad \|D\Delta u\|^{4/3}_{L^{4/3}(B_r)} &\le Cr^{4/3}\|D^2u\|^{4/3}_{L^2(B_2)}\\ &\quad + C(\|D^2u\|^{8/3}_{L^2(B_2)} + \|Du\|^{8/3}_{L^4(B_2)} + \|f\|^{1/3}_{L^1(B_2)}\|f\|_{L\log L(B_2)}).\end{aligned}$$

Remark 2.2. In point of fact, we do not need the full strength of (2-4) in application. We use instead the estimate

$$\|D\Delta u\|^{4/3}_{L^{4/3}(B_r)} \le C(\|D^2u\|^{4/3}_{L^2(B_{8r})} + \|Du\|^{4/3}_{L^4(B_{8r})} + \|f\|^{1/3}_{L^1(B_{8r})}\|f\|_{L\log L(B_{8r})}),$$

which can be immediately proven via the method outlined below.

Proof. First, find $v \in W^{1,2}_0(B_2, \mathrm{so}(n+1)) \cap W^{2,2}(B_2, \mathrm{so}(n+1))$ such that

$$\Delta v = \Delta u \wedge u.$$

Thus, for each $i, j \in \{1, \dots, n+1\}$, $\Delta v^{ij} = u^j \Delta u^i - u^i \Delta u^j$. It follows from (1-2) that

$$\Delta^2 v = \Delta(\Delta u \wedge u) = 2d^*(\Delta u \wedge Du) + f \wedge u.$$

Next we let $\phi \in W^{2,2}_0(B_2, \mathrm{so}(n+1) \otimes \Omega^1\mathbb{R}^4)$ be the solution of

$$\Delta^2\phi = d^*(2\Delta u \wedge Du).$$

Here $\mathrm{so}(n+1) \otimes \Omega^1\mathbb{R}^4$ denotes the space of 1-forms tensored with $(n+1)\times(n+1)$-antisymmetric matrices. Using Calderón–Zygmund theory coupled with interpolation, and using the estimates from Section A.2, we determine that

$$\begin{aligned}(2\text{-}5)\quad \|D^3\phi\|_{L^{4/3,1}(B_2)} + \|D^2\phi\|_{L^{2,1}(B_2)} + \|D\phi\|_{L^{4,1}(B_2)}&\\ \le c(\|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^2(B_2)})&.\end{aligned}$$

Moreover, letting $\psi \in W^{2,2}_0(B_2, \mathrm{so}(n+1))$ be the solution of

$$\Delta^2\psi = f \wedge u,$$

we conclude that

$$(2\text{-}6)\quad \|D\psi\|_{L^{4,1}(B_2)} + \|D^2\psi\|_{L^{2,1}(B_2)} + \|D^3\psi\|_{L^{4/3,1}(B_2)} \le c\|f\|_{L\log L(B_2)}.$$

Defining

$$B := v - \phi - \psi$$

and using the above equation for v, we conclude that each B^{ij} is a biharmonic function on B_2. Now every biharmonic function satisfies the mean value property

$$B(x) = c_1 \fint_{B_r(x)} B(y)\,dy - c_2 \fint_{B_{2r}(x)} B(y)\,dy,$$

for every $B_{2r}(x) \subset B_2$ (see, e.g., [Huilgol 1971]). Hence we estimate

$$\begin{aligned}
&\|D^2 B\|_{L^{2,1}(B_{3/2})} + \|D^3 B\|_{L^{4/3,1}(B_{3/2})}\\
&\qquad \le c\|DB\|_{L^2(B_2)}\\
&\qquad \le c(\|Dv\|_{L^2(B_2)} + \|f\|_{L\log L(B_2)} + \|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^2(B_2)}).
\end{aligned}$$

Since $v = 0$ on ∂B_2, we can use the divergence theorem and Cauchy–Schwarz to show that

$$\begin{aligned}
\int_{B_2} |Dv^{ij}|^2 = -\int_{B_2} v^{ij}\,\Delta v^{ij} &= -\int_{B_2} Dv^{ij}\cdot (Du\wedge u)^{ij}\\
&\le \frac{1}{2}\int_{B_2} |Dv^{ij}|^2 + C\int_{B_2} |Du|^2.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\|D^2 B\|_{L^{2,1}(B_{3/2})} + \|D^3 B\|_{L^{4/3,1}(B_{3/2})}\\
&\qquad \le c(\|Du\|_{L^2(B_2)} + \|f\|_{L\log L(B_2)} + \|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^2(B_2)}).
\end{aligned}$$

Now we observe that, since $\Delta v = \Delta u \wedge u$,

$$\Delta u = (\Delta u \wedge u).u + \langle \Delta u, u\rangle u = \Delta v.u - |Du|^2 u,$$

where here $\Omega.u$ represents matrix multiplication. Therefore,

$$\Delta^2 u = \Delta(\Delta v.u - |Du|^2 u) = d^*(D\Delta v.u + \Delta v.Du - D(|Du|^2 u)).$$

To get the second- and third-derivative estimates in (2-1), we first observe that

$$\begin{aligned}
&\|D^2 v\|_{L^{2,1}(B_{3/2})} + \|D^3 v\|_{L^{4/3,1}(B_{3/2})}\\
&\qquad \le c(\|Du\|_{L^2(B_2)} + \|f\|_{L\log L(B_2)} + \|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^2(B_2)}).
\end{aligned}$$

Using the previous estimates and Section A.2, we observe that the 1-form in the parentheses is in $L^{4/3,1}$. Lemma A.3 in [Lamm and Rivière 2008] implies that

$$\begin{aligned}
&\|D^2 u\|_{L^{2,1}(B_1)} + \|D^3 u\|_{L^{4/3,1}(B_1)}\\
&\qquad \le c(\|D^3 v\|_{L^{4/3,1}(B_{3/2})} + \|D^2v\|^2_{L^2(B_2)} + \|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^2(B_2)})\\
&\qquad \le c(\|Du\|_{L^2(B_2)} + \|f\|_{L\log L(B_2)} + \|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^2(B_2)}).
\end{aligned}$$

Finally, Sobolev embedding for Lorentz spaces implies that

$$\begin{aligned}\|Du\|_{L^{4,1}(B_1)} &\le c(\|D^2u\|_{L^{2,1}(B_2)} + \|Du\|_{L^{2,1}(B_2)})\\ &\le c(\|D^2u\|_{L^{2,1}(B_2)} + \|Du\|_{L^{2}(B_2)}).\end{aligned}$$

Combining this with the previous estimates finishes the proof of (2-1).

To prove the small-energy estimates, we observe that u satisfies (see, for instance, [Lamm and Rivière 2008, Equations 1.4 and 1.14])

$$\Delta^2 u = \Delta(V\cdot Du) + d^*(wDu) + W\cdot Du + f, \tag{2-7}$$

where $V^{ij} = u^i Du^j - u^j Du^i$, $w^{ij} = -d^*(V^{ij}) - 2|Du|^2\delta_{ij}$, and

$$W^{ij} = -D(d^*(V^{ij})) + 2(\Delta u^i Du^j - \Delta u^j Du^i).$$

Let $\mathcal{M}_m$ denote the space of $m\times m$ matrices and $\mathcal{M}_m \otimes \Omega^k\mathbb{R}^4$ the space of k-forms tensored with $m\times m$ matrices. Then $V\in W^{1,2}(B_2, \mathcal{M}_{n+1}\otimes\Omega^1\mathbb{R}^4)$, $w\in L^2(B_2,\mathcal{M}_{n+1})$, and $W\in W^{-1,2}(B_2,\mathcal{M}_{n+1}\otimes\Omega^1\mathbb{R}^4)$.

Without loss of generality we extend f by zero outside of B_2. The small-energy hypothesis implies (see, for instance, [Lamm and Rivière 2008]) that there exist $A\in L^\infty\cap W^{2,2}(B_1,\mathrm{GL}_{n+1})$ and $\widetilde{B}\in W^{1,4/3}(B_1,\mathcal{M}_{n+1}\otimes\Omega^2\mathbb{R}^4)$ such that

$$D\Delta A + \Delta A V - DAw + AW = D\widetilde{B}$$

and

$$\begin{aligned}&\Delta(A\Delta u)\\ &\quad = d^*(2DA\Delta u - \Delta A Du + AwD - DA(V\cdot Du) + AD(V\cdot Du) + \widetilde{B}\cdot Du)) + Af\\ &\quad := d^*(K) + Af.\end{aligned}$$

Moreover,

$$\begin{aligned}&\|DA\|_{W^{1,2}(B_1)} + \|\operatorname{dist}(A,\mathrm{SO}(n+1))\|_{L^\infty(B_1)} + \|\widetilde{B}\|_{W^{1,4/3}(B_1)}\\ &\qquad\qquad \le c(\|D^2u\|_{L^2(B_2)} + \|Du\|_{L^4(B_2)}).\end{aligned}$$

First, we determine $E, F\in W_0^{1,2}(B_1)$ such that

$$\Delta E = d^*(K), \quad \Delta F = Af.$$

Interpolating on standard L^p theory, we get the estimates

$$\begin{aligned}\|E\|_{L^{2,1}(B_1)} + \|DE\|_{L^{4/3,1}(B_1)} &\le c\|K\|_{L^{4/3,1}(B_2)}\\ &\le c(\|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^4(B_2)}).\end{aligned}$$

Note that the estimate on K comes from considering the form of (2-7) and the estimates on V, w, W and consequently those on A, $\widetilde{B}$.

To determine estimates on F, we first observe that the estimates of Section A.2 imply that for G the fundamental solution to $\Delta^2 G = \delta_0$,

$$\begin{aligned}\|F\|_{L^{2,\infty}(B_1)} &\le c\|D^2 G * (Af)\|_{L^{2,\infty}(B_1)} \le c\|f\|_{L^1(B_2)},\\ \|DF\|_{L^{4/3,\infty}(B_1)} &\le c\|D^3 G\|_{L^{4/3,\infty}(B_2)}\|f\|_{L^1(B_2)}.\end{aligned}$$

Also, since $\Delta F = Af \in \mathcal{H}^1(\mathbb{R}^4)$, standard theory implies that $D^2 F \in L^1(\mathbb{R}^4)$ and thus, by the embedding of $W^{1,1}$ into $L^{4/3,1}$ and Sobolev embeddings in $\mathbb{R}^4$,

$$\|F\|_{L^{2,1}(B_1)} + \|DF\|_{L^{4/3,1}(B_1)} \le c\|f\|_{L\log L(B_2)}.$$

Using a duality argument, we conclude that

$$\begin{aligned}\|F\|^2_{L^2(B_1)} &\le c\|F\|_{L^{2,\infty}(B_1)}\|F\|_{L^{2,1}(B_1)}\\ &\le c\|f\|_{L^1(B_2)}\|f\|_{L\log L(B_2)},\\ \|DF\|^{4/3}_{L^{4/3}(B_1)} &\le c\|(DF)^{1/3}\|_{L^{4,\infty}(B_1)}\|DF\|_{L^{4/3,1}(B_1)}\\ &\le c\|DF\|^{1/3}_{L^{4/3,\infty}(B_2)}\|f\|_{L\log L(B_2)}\\ &\le c\|f\|^{1/3}_{L^1(B_2)}\|f\|_{L\log L(B_2)}.\end{aligned}$$

Now, set $H = A\Delta u - E - F$. Then $\Delta H = 0$ in B_1, and, using standard estimates on harmonic functions, we determine that for all $0 < r < \frac{1}{2}$

$$\|H\|_{L^2(B_r)} + \|DH\|_{L^{4/3}(B_r)} \le cr\|H\|_{W^{1,\infty}(B_{1/2})} \le cr\|H\|_{L^2(B_1)}.$$

The previous estimates imply that

$$\|H\|^2_{L^2(B_1)} \le c(\|D^2 u\|^2_{L^2(B_2)} + \|Du\|^4_{L^4(B_2)} + \|f\|_{L^1(B_2)}\|f\|_{L\log L(B_2)}).$$

Since

$$\Delta u = A^{-1}(E + F + H),$$

the estimates for $D^2 u$ now follow from a standard cutoff argument and the previous estimates.

We estimate $\|D\Delta u\|_{L^{4/3}(B_r)}$ by using the previous estimates and noting that

$$\begin{aligned}&\|D(A^{-1}(E + F + H))\|_{L^{4/3}(B_r)}\\ &\qquad \le C(\|E + F + H\|_{L^2(B_r)}\|DA\|_{L^4(B_r)} + \|D(E + F + H)\|_{L^{4/3}(B_r)}).\end{aligned}$$

To estimate Du, we first consider $\alpha \in W^{2,2}(B_1)$, $\beta \in W^{1,2}_0 \cap W^{2,2}(B_1, \Omega^1\mathbb{R}^4)$ such that

$$ADu = d\alpha + d^*\beta.$$

Then

$$\Delta^2\alpha = \Delta d^*(ADu) = \Delta(A\Delta u + DA.Du) = d^*(\widetilde{K}) + Af \quad \text{on } B_1$$

and

$$\Delta\beta = DA \wedge Du \quad \text{on } B_1.$$

Here $\widetilde{K}$ is the appropriate modification of K to include the additional term. We first observe that

$$\|D\beta\|_{L^4(B_r)} \le c(\|D^2\beta\|_{L^2(B_1)} + \|D\beta\|_{L^2(B_1)}).$$

Standard L^p theory implies that

$$\|D^2\beta\|_{L^2(B_2)} \le c\|DA\|_{W^{1,2}(B_1)}\|Du\|_{W^{1,2}(B_1)}.$$

Moreover, using a weighted Cauchy–Schwarz inequality and the Poincaré inequality, we note that

$$\begin{aligned}\int_{B_1} |D\beta^{ij}|^2 &= -\int_{B_1} \beta^{ij}(DA \wedge Du)^{ij}\\ &\le c\|DA\|^2_{L^4(B_1)}\|Du\|^2_{L^4(B_1)} + \tfrac{1}{2}\|D\beta\|^2_{L^2(B_1)}.\end{aligned}$$

Combining this with previous estimates implies that

$$\|D\beta\|_{L^4(B_r)} \le c(\|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^4(B_2)}).$$

For the α term, we follow the ideas used to prove (2-1). Indeed, first determine $\phi, \psi \in W^{2,2}_0(B_2)$ such that $\Delta^2\phi = d^*(K)$ and $\Delta^2\psi = Af$. Then by (2-5), (2-6), and appropriate duality arguments, we conclude that, for any $0 < r < 1$,

$$\begin{aligned}\|D\phi\|_{L^4(B_r)} &\le c(\|D^2u\|^2_{L^2(B_2)} + \|Du\|^2_{L^4(B_2)}),\\ \|D\psi\|^4_{L^4(B_r)} &\le c\|f\|^3_{L^1(B_2)}\|f\|_{L\log L(B_2)}.\end{aligned}$$

Setting $B = \alpha - \psi - \phi$, we have $\Delta^2 B = 0$ on B_1, and we use the mean value property to show that for any $0 < r < \frac{1}{2}$

$$\|DB\|_{L^4(B_r)} \le cr\|DB\|_{L^\infty(B_{3/4})} \le cr\|DB\|_{L^4(B_{7/8})}.$$

Noting that

$$\begin{aligned}\|DB\|^4_{L^4(B_{7/8})} \le c\big(\|D\alpha\|^4_{L^4(B_{7/8})} + \|Du\|^4_{L^4(B_1)} + \|D^2u\|^8_{L^2(B_2)}\\ + \|Du\|^8_{L^4(B_2)} + \|f\|^3_{L^1(B_2)}\|f\|_{L\log L(B_2)}\big),\end{aligned}$$

we combine the previous estimates to get the result for Du. □

Remark 2.3. When u is intrinsic, the strategy is the same, except for two things. In the first part of the argument, the equation for u will have the additional term $-d^*(|Du|^2 Du \wedge u)$ on the right side. But this term doesn't change the estimates. In the second part of the argument, W^{ij} will include the term $|Du|^2(u^i Du^j - u^j Du^i)$. This gives the same value for $d^*(W^{ij})$, and all estimates going forward are the same.

We will prove the energy quantization results by appealing to Lorentz duality. In Proposition 2.1, we determined uniform estimates for Lorentz norms of the form $L^{p,1}$. The next lemma provides the necessary small-energy estimates for the $L^{p,\infty}$ norms on the annular region, presuming small energy on all dyadic annuli:

Lemma 2.4. *Let $u \in W^{2,2}(B_1, \mathbb{S}^n)$ be an f-approximate biharmonic map with $f \in L\log L(B_1, \mathbb{R}^{n+1})$. Given $\varepsilon > 0$, suppose that for all ρ such that $B_{2\rho} \setminus B_\rho \subset B_{2\delta} \setminus B_{t/2}$ we have*

$$\int_{B_{2\rho}\setminus B_\rho} |Du|^4 + |D^2u|^2 + |D\Delta u|^{4/3} < \varepsilon. \tag{2-8}$$

Then,

$$\|Du\|_{L^{4,\infty}(B_\delta\setminus B_t)} + \|D^2u\|_{L^{2,\infty}(B_\delta\setminus B_t)} + \|D\Delta u\|_{L^{4/3,\infty}(B_\delta\setminus B_t)} \leq C(\varepsilon^{\frac{1}{8}} + (\log(1/\delta))^{-1}).$$

Proof. Let $\widetilde{\phi}_k := \phi_{2^{k+2}t}(1-\phi_{2^{k-2}t})$ be the annular cutoff supported on $A_k := B_{2^{k+3}t} \setminus B_{2^{k-2}t}$ which is identically 1 on $B_{2^{k+2}t} \setminus B_{2^{k-1}t}$. Let G be the distribution such that $\Delta^2 G = \delta_0$ in $\mathbb{R}^4$. Then $|DG(x)| = C|x|^{-1}$. Note that operator bounds on $D^k G$ can be found in the appendix. Let $\bar{u}_k := ⨍_{A_k} u$. Set $\tilde{u}_k(x) := \widetilde{\phi}_k(u - \bar{u}_k)(x)$. Therefore on $B_{2^{k+1}t} \setminus B_{2^k t}$

$$\Delta^2\tilde{u}_k = (\Delta^2\widetilde{\phi}_k)(u-\bar{u}_k) + 4D\Delta\widetilde{\phi}_k \cdot D(u-\bar{u}_k) + 2\Delta\widetilde{\phi}_k\Delta u + 4D\widetilde{\phi}_k \cdot D\Delta u + \widetilde{\phi}_k\Delta^2 u.$$

Using the facts that $\Delta^2 u = \Delta(\Delta u \wedge u.u - |Du|^2 u)$ and that $\Delta^2 u \wedge u = f \wedge u$, we note that

$$\begin{aligned}\widetilde{\phi}_k\Delta^2 u = d^*\big(&\widetilde{\phi}_k(2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2))\big)\\ &- D\widetilde{\phi}_k \cdot (2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2))\\ &+ \widetilde{\phi}_k(f \wedge u.u - 2\Delta u \wedge Du.Du - \Delta u \wedge u.\Delta u).\end{aligned}$$

And thus,

$$\begin{aligned}\Delta^2\tilde{u}_k = (\Delta^2\widetilde{\phi}_k)&(u-\bar{u}_k) + 4D\Delta\widetilde{\phi}_k \cdot D(u-\bar{u}_k) + 2\Delta\widetilde{\phi}_k\Delta u + 4D\widetilde{\phi}_k \cdot D\Delta u\\ &- D\widetilde{\phi}_k \cdot (2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2))\\ &+ d^*\big(\widetilde{\phi}_k(2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2))\big)\\ &+ \widetilde{\phi}_k(f \wedge u.u - 2\Delta u \wedge Du.Du - \Delta u \wedge u.\Delta u).\end{aligned}$$

For ease of notation, we let I_k denote the first four terms above, and $\mathrm{II}_k, \mathrm{III}_k, \mathrm{IV}_k$ denote the last three terms, respectively. Then on each $B_{2^{k+1}t} \setminus B_{2^k t}$

$$\begin{aligned}|Du(x)| &= |D(\widetilde{\phi}_k(u-\bar{u}_k))(x)| = |\Delta^2 G * D(\widetilde{\phi}_k(u-\bar{u}_k))(x)|\\ &= |DG * \Delta^2(\widetilde{\phi}_k(u-\bar{u}_k))(x)| = |DG * (\mathrm{I}_k + \mathrm{II}_k + \mathrm{III}_k + \mathrm{IV}_k)(x)|.\end{aligned}$$

We consider each of these estimates separately. First, note that

$$\begin{aligned}&|DG * \mathrm{I}_k(x)|\\ &\quad \le C\bigg|\int_{(B_{2^{k+3}t}\setminus B_{2^{k+2}t})\cup(B_{2^{k-1}t}\setminus B_{2^{k-2}t})} \frac{1}{|x-y|}\\ &\qquad \times \big((2^k t)^{-4}(u-\bar{u}_k) + (2^k t)^{-3} D(u-\bar{u}_k) + (2^k t)^{-2}\Delta u + (2^k t)^{-1} D\Delta u\big)\, dy\bigg|\\ &\quad \le C\bigg|\int_{A_k} (2^k t)^{-1}\big((2^k t)^{-4}(u-\bar{u}_k) + (2^k t)^{-3} Du\\ &\qquad\qquad + (2^k t)^{-2}\Delta u + (2^k t)^{-1} D\Delta u\big)\, dy\bigg|\\ &\quad \le C\int_{A_k} (2^k t)^{-4}|Du| + (2^k t)^{-3}|D^2 u| + (2^k t)^{-2}|D\Delta u|\\ &\quad \le C(2^k t)^{-1}(\|Du\|_{L^4} + \|D^2 u\|_{L^2} + \|D\Delta u\|_{L^{4/3}})\\ &\quad \le C(\varepsilon^{1/4} + \varepsilon^{1/2} + \varepsilon^{3/4})|x|^{-1}.\end{aligned}$$

Using the same ideas as previously, we bound

$$\begin{aligned}&|DG * \mathrm{II}_k(x)|\\ &\qquad \le C(2^k t)^{-2}\int_{A_k} |2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2)|\\ &\qquad \le C(2^k t)^{-1}\|2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2)\|_{L^{4/3}(A_k)}\\ &\qquad \le C(2^k t)^{-1}(\|D^2 u\|_{L^2}\|Du\|_{L^4} + \|Du\|_{L^4}^3)\\ &\qquad \le C(\varepsilon^{1/8} + \varepsilon^{3/4})|x|^{-1}.\end{aligned}$$

Using the estimates from the appendix, we note that

$$\begin{aligned}&\|DG * \mathrm{III}_k\|_{L^{4,\infty}(A_k)}\\ &\qquad \le C\|D^2 G * \widetilde{\phi}_k(2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2)\|_{L^{4,\infty}(A_k)}\\ &\qquad \le C\|\widetilde{\phi}_k(2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2))\|_{L^{4/3}(A_k)}\end{aligned}$$

and

$$\|DG * \mathrm{IV}_k\|_{L^{4,\infty}(A_k)} \le C\|\widetilde{\phi}_k(f \wedge u.u - 2\Delta u \wedge Du.Du - \Delta u \wedge u.\Delta u)\|_{L^1(A_k)}.$$

Thus

$$
\begin{aligned}
&|\{x : |DG * (\mathrm{III}_k + \mathrm{IV}_k)(x)| > \lambda\}| \\
&\qquad \le \lambda^{-4} \| DG * (\mathrm{III}_k + \mathrm{IV}_k)\|^4_{L^{4,\infty}(\mathbb{R}^4)} \\
&\qquad \le C\lambda^{-4}\big(\|\widetilde{\phi}_k(f \wedge u.u - 2\Delta u \wedge Du.Du - \Delta u \wedge u.\Delta u)\|^4_{L^1(A_k)} \\
&\qquad\qquad + \|\widetilde{\phi}_k(2\Delta u \wedge Du.u + 2\Delta u \wedge u.Du - D(u|Du|^2))\|^4_{L^{4/3}(A_k)}\big) \\
&\qquad \le C\lambda^{-4}\left(\left(\int \widetilde{\phi}_k |D^2 u|^2\right)^2 \int \widetilde{\phi}_k |Du|^4 + \left(\int \widetilde{\phi}_k |Du|^4\right)^3\right) \\
&\qquad\qquad + C\lambda^{-4}\|\widetilde{\phi}_k(f \wedge u.u - 2\Delta u \wedge Du.Du - \Delta u \wedge u.\Delta u)\|^4_{L^1(A_k)}.
\end{aligned}
$$

Thus, if $\delta = 2^M t$, then (letting $S_k := B_{2^{k+1}t} \setminus B_{2^k t}$ for ease of notation)

$$
\begin{aligned}
&|\{x \in B_\delta \setminus B_t : |Du(x)| > 3\lambda\}| \\
&\qquad \le \sum_{k=0}^{M-1} |\{x \in S_k : |Du(x)| > 3\lambda\}| \\
&\qquad \le \sum_{k=0}^{M-1} |\{x \in S_k : |DG * \mathrm{I}_k| > \lambda\}| + \sum_{k=0}^{M-1} |\{x \in S_k : |DG * \mathrm{II}_k| > \lambda\}| \\
&\qquad\quad + \sum_{k=0}^{M-1} |\{x \in S_k : |DG * (\mathrm{III}_k + \mathrm{IV}_k)| > \lambda\}| \\
&\qquad \le \sum_{k=0}^{M-1} |\{|DG * (\mathrm{III}_k + \mathrm{IV}_k)| > \lambda\}| + \left|\left\{x \in B_1 : C\frac{\varepsilon^{\frac{1}{8}}}{|x|} > \lambda\right\}\right| \\
&\qquad \le C\lambda^{-4}\left(\varepsilon^{\frac{1}{2}} + \sum_{k=0}^{M-1} \|\widetilde{\phi}_k(f \wedge u.u)\|^4_{L^1(A_k)} + \sum_{k=0}^{M-1}\left(\left(\int \widetilde{\phi}_k |D^2 u|^2\right)^4\right.\right. \\
&\qquad\qquad\qquad \left.\left. + \left(\int \widetilde{\phi}_k |Du|^4\right)^4 + \left(\int \widetilde{\phi}_k |Du|^4\right)^3 + \left(\int \widetilde{\phi}_k |Du|^4\right)^2\right)\right) \\
&\qquad \le C\lambda^{-4}(\varepsilon^{\frac{1}{2}} + (\log(1/\delta))^{-4}\|f\|^4_{L\log L(B_{2\delta})} + \varepsilon^2).
\end{aligned}
$$

For the estimate on $\|f \wedge u.u\|_{L^1}$ we use Lemma A.2, and for the rest of the L^1 estimate we just use Cauchy–Schwarz. This proves the estimate for Du. The estimates for $D^2 u$ and $D\Delta u$ work in much the same way. In the case of $D^2 u$, the terms like III_k and IV_k require the fact that $D^3 G : L^{4/3} \to L^{2,\infty}$ and $D^2 G : L^1 \to L^{2,\infty}$ are bounded operators, where the operation is convolution. For the term $D\Delta u$ we observe that $D^3 G : L^1 \to L^{4/3,\infty}$ and $D^4 G : L^{4/3} \to L^{4/3,\infty}$ are also bounded operators. □

3. Energy quantization — proof of Theorem 1.2

We now determine a weak convergence result which will give small-energy compactness and help us complete the proof of the energy quantization. We follow the ideas of [Li and Zhu 2011; Sharp and Topping 2013], which in turn follow the arguments of [Evans 1990], with appropriate minor modifications. Throughout this lemma and its proof, we consider a measurable function f as both a function and a Radon measure.

Lemma 3.1. *Suppose $\{V_k\} \subset W^{1,4/3}(B_1)$ is a bounded sequence in $B_1 \subset \mathbb{R}^4$. Then there exist at most countable $\{x_i\} \subset B_1$ and $\{a_i > 0\}$ with $\sum_i a_i < \infty$ and $V \in W^{1,4/3}(B_1)$ such that, after passing to a subsequence,*

$$V_k^2 \rightharpoonup V^2 + \sum_i a_i \delta_{x_i}$$

weakly as measures.

Proof. As $W^{1,4/3}$ embeds continuously into L^2 in four dimensions, after taking a subsequence, by Rellich compactness there exists some $V \in L^2$ such that $V_k \to V$ strongly in L^p for $1 \leq p < 2$ and $V_k \rightharpoonup V$ weakly in L^2. Moreover, since $\{DV_k\}$ is uniformly bounded in $L^{4/3}$, it follows that $DV_k \rightharpoonup f \in L^{4/3}$ and f is necessarily DV.

Set $g_k := V_k - V$. Then $g_k \in L^2$ and $Dg_k \in L^{4/3}$ with uniform bounds. Thus, in the weak-$*$ topology, both $|Dg_k|^{4/3}$ and g_k^2 converge to nonnegative Radon measures with finite total mass. (We denote this space by $M(B)$). Then $g_k^2 \rightharpoonup \nu \in M(B)$ and $|Dg_k|^{4/3} \rightharpoonup \mu \in M(B)$ where ν, μ are both nonnegative. Now consider $\phi \in C_0^1(B_1)$, and observe that the Sobolev embedding of $W^{1,4/3}$ into L^2 implies that

$$\left(\int (\phi g_k)^2 \, dx\right)^{\frac{1}{2}} \leq C\left(\int |D(\phi g_k)|^{\frac{4}{3}} \, dx\right)^{\frac{3}{4}}.$$

Taking $k \to \infty$ and noting that $g_k \to 0$ in $L^{4/3}$, we use the weak convergence to observe that

$$\int \phi^2 \, d\nu \leq C\left(\int |\phi|^{\frac{4}{3}} \, d\mu\right)^{\frac{3}{2}}.$$

Let ϕ approximate $\chi_{B_r(x)}$ for $B_r(x) \subset B_1$. Then

$$\nu(B_r(x)) \leq C\big(\mu(B_r(x))\big)^{\frac{3}{2}}.$$

By standard results on the differentiation of measures (see [Evans and Gariepy 1992, Section 1.6]), for any Borel set E

$$\nu(E) = \int_E D_\mu \nu \, d\mu,$$

where

$$D_\mu \nu(x) = \lim_{r\to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} \quad \text{for } \mu\text{-a.e. } x \in \mathbb{R}^4.$$

Now, as μ is a finite, nonnegative Radon measure, there exist at most countably many $x_i \in B_1$ such that $\mu(\{x_i\}) > 0$. Moreover, for all $x \in B$ such that $\mu(\{x\}) = 0$, we note that

$$D_\mu \nu(x) = \lim_{r\to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} \le C \lim_{r\to 0} \mu(B_r(x))^{\frac{1}{2}} = 0.$$

For every x_j such that $\mu(\{x_j\}) > 0$, set $a_j = D_\mu \nu(x_j)\mu(\{x_j\})$. Then

$$\nu(E) = \int_E D_\mu \nu \, d\mu = \sum_{\{j : x_j \in E\}} a_j \quad \text{or} \quad \nu = \sum_j a_j \delta_{x_j}.$$

Since $g_k^2 \rightharpoonup \nu$ as measures, for $\phi \in C_0^0(B_1)$,

$$\sum_j a_j \phi(x_j) = \lim_{k\to\infty} \int_{B_1} g_k^2 \phi \, dx = \lim_{k\to\infty} \int_{B_1} (V_k - V)^2 \, dx.$$

Since $(V_k - V)^2 = V_k^2 - V^2 + 2V(V - V_k)$ and $V - V_k = g_k \rightharpoonup 0$ in L^2, we have the result. □

Corollary 3.2. *For $\{V_k\}$ as in Lemma 3.1, if*

$$\lim_{r\to 0} \limsup_{k\to\infty} \|V_k\|_{L^2(B_r(x))} = 0 \tag{3-1}$$

for all $x \in B$, then

$$V_k \to V \text{ strongly in } L^2_{\mathrm{loc}}(B).$$

Proof. Notice the condition (3-1) implies that $|V_k|^2 \rightharpoonup |V|^2$ weakly as bounded Radon measures. Then, by [Evans and Gariepy 1992, Section 1.9], for any $B_r(x) \subset B_1$, we have $\|V_k\|_{L^2(B_r(x))} \to \|V\|_{L^2(B_r(x))}$ strongly for all $B_r(x) \subset B_1$. Then, again using the fact that $(V_k - V)^2 = V_k^2 - V^2 + 2V(V - V_k)$ and

$$\int_{B_r(x)} V_k^2 - V^2 \, dx + \int_{B_r(x)} 2V(V - V_k) \, dx \to 0 \quad \text{as } k \to \infty,$$

we conclude that $V_k \to V$ strongly in $L^2_{\mathrm{loc}}(B_1)$. □

We now use the energy estimates of Proposition 2.1 to prove a small-energy compactness result:

Lemma 3.3. *Let u_k be a sequence of f_k-approximate biharmonic maps in B_2 with $f_k \in L \log L(B_2)$ satisfying* (1-1). *There exists $\varepsilon_0 > 0$ such that if*

$$\|Du_k\|_{L^4(B_2)} + \|D^2 u_k\|_{L^2(B_2)} < \varepsilon_0,$$

then there exists $u \in W^{2,2}_{\text{loc}}(B_2)$ such that

$$Du_k \to Du \text{ strongly in } L^4_{\text{loc}}(B_1) \quad \text{and} \quad D^2u_k \to D^2u \text{ strongly in } L^2_{\text{loc}}(B_1).$$

Proof. We will first prove convergence of Du_k to Du and D^2u_k to D^2u in L^2_{loc} and then use Gagliardo–Nirenberg interpolation to get the L^4 convergence.

Begin by choosing $0 < \varepsilon_0 < \tilde{\varepsilon}$ from Proposition 2.1. First note that the uniform bounds on u_k in $W^{2,2}(B_2)$ imply that there exists a $u \in W^{2,2}_{\text{loc}}(B_2)$ such that $u_k \rightharpoonup u$ in $W^{2,2}_{\text{loc}}(B_2)$. We now show strong convergence for the derivatives indicated.

Pick any $x_0 \in B_1$ and $2R \in (0, \frac{3}{4}]$. Then $B_{2R}(x_0) \subset B_2$. Let $\hat{u}_k(x) := u_k(x_0 + 2Rx)$ and $\hat{f}_k(x) := (2R)^4 f_k(x_0 + 2Rx)$. Then $\hat{u}_k$ is an $\hat{f}_k$-approximate biharmonic map on B_1. From (2-2), (2-3), we note that, for any $r \in (0, \frac{1}{2}]$,

$$\begin{aligned}
&\|D\hat{u}_k\|_{L^4(B_r)} + \|D^2\hat{u}_k\|_{L^2(B_r)} \\
&\quad \le Cr(\|D\hat{u}_k\|_{L^4(B_2)} + \|D^2\hat{u}_k\|_{L^2(B_2)}) \\
&\qquad + C\big(\|D\hat{u}_k\|^2_{L^4(B_2)} + \|D^2\hat{u}_k\|^2_{L^2(B_2)} + (\|\hat{f}_k\|_{L^1(B_2)}\|\hat{f}_k\|_{L\log L(B_2)})^{\frac{1}{2}} \\
&\qquad\qquad + (\|\hat{f}_k\|^3_{L^1(B_2)}\|\hat{f}_k\|_{L\log L(B_2)})^{\frac{1}{4}}\big).
\end{aligned}$$

Using the scaling relations listed in Section A.3 and Lemma A.3 we observe that

$$\begin{aligned}
&\|Du_k\|_{L^4(B_{r2R}(x_0))} + \|D^2u_k\|_{L^2(B_{r2R}(x_0))} \\
&\quad \le Cr(\|Du_k\|_{L^4(B_{2R}(x_0))} + \|D^2u_k\|_{L^2(B_{2R}(x_0))}) \\
&\qquad + C\big(\|Du_k\|^2_{L^4(B_{2R}(x_0))} + \|D^2u_k\|^2_{L^2(B_{2R}(x_0))} \\
&\qquad\qquad + (\|f_k\|_{L^1(B_{2R}(x_0))}\|f_k\|_{L\log L(B_{2R}(x_0))})^{\frac{1}{2}} \\
&\qquad\qquad\qquad + (\|f_k\|^3_{L^1(B_{2R}(x_0))}\|f_k\|_{L\log L(B_{2R}(x_0))})^{\frac{1}{4}}\big).
\end{aligned}$$

Lemma A.2 and (1-1) together imply that

$$\|f_k\|_{L^1(B_{2R}(x_0))} \le C\Big(\log\frac{1}{2R}\Big)^{-1}\|f_k\|_{L\log L(B_{2R}(x_0))} \le C\Lambda\Big(\log\frac{1}{2R}\Big)^{-1}.$$

Note that the right-hand side goes to zero as $R \to 0$. Therefore, the small-energy hypothesis implies that

$$\begin{aligned}
&\lim_{R\to 0}\lim_{r\to 0}\lim_{k\to\infty}(\|Du_k\|_{L^4(B_{r2R}(x_0))} + \|D^2u_k\|_{L^2(B_{r2R}(x_0))}) \\
&\qquad \le C\varepsilon_0 \lim_{R\to 0}\lim_{r\to 0}\lim_{k\to\infty}(\|Du_k\|_{L^4(B_{2R}(x_0))} + \|D^2u_k\|_{L^2(B_{2R}(x_0))}).
\end{aligned}$$

Decreasing ε_0, if necessary, so that $\varepsilon_0 < 1/C$, implies that

$$\lim_{r\to 0}\lim_{k\to\infty}(\|Du_k\|_{L^4(B_r(x_0))} + \|D^2u_k\|_{L^2(B_r(x_0))}) = 0$$

for all $x_0 \in B_1$. Let $V_k = D^2 u_k$ and $V = D^2 u$. Since $V_k \rightharpoonup V$ weakly in L^2 as measures and V_k satisfies the hypotheses of Lemma 3.1 and Corollary 3.2 on B_1, $V_k \to V$ strongly in $L^2_{\text{loc}}(B_1)$.

Since $Du_k \rightharpoonup Du$ weakly as measures in $L^2(B_2)$ and

$$\lim_{r\to 0}\lim_{k\to\infty} \|Du_k\|_{L^2(B_r(x_0))} \le \lim_{r\to 0}\lim_{k\to\infty} r\|Du_k\|_{L^4(B_r(x_0))} = 0$$

for all $x_0 \in B_1$, Corollary 3.2 again implies that $Du_k \to Du$ strongly in $L^2_{\text{loc}}(B_1)$.

Now, for any $B_r(x) \subset B_1$, we consider the functions

$$w_k := (u_k - u) - \fint_{B_r(x)} (u_k - u).$$

Then, $Dw_k = D(u_k - u)$ and $D^2 w_k = D^2(u_k - u)$. We apply the Gagliardo–Nirenberg interpolation inequality for w_k and then the Poincaré inequality for the L^2 estimates on w_k to conclude that

$$\|Dw_k\|_{L^4(B_r(x))} \le C\|D^2 w_k\|_{L^2(B_r(x))}\|Dw_k\|_{L^2(B_r(x))} + C\|Dw_k\|_{L^2(B_r(x))}.$$

Then, using the strong convergence of $D^2 u_k \to Du$ in L^2_{loc} and $Du_k \to Du$ in L^2_{loc}, we conclude $Du_k \to Du$ in $L^4_{\text{loc}}(B_1)$. □

Finally, we prove the energy quantization result under the presumption of one bubble at the origin.

Proposition 3.4. *Let $f_k \in L\log L(B_1, \mathbb{R}^{n+1})$, and let $u_k \in W^{2,2}(B_1, \mathbb{S}^n)$ be a sequence of f_k-approximate biharmonic maps with bounded energy such that*

$$\begin{aligned} u_k &\to u && \text{in } W^{2,2}_{\text{loc}}(B_1 \setminus \{0\}, \mathbb{S}^n),\\ \tilde{u}_k(x) := u_k(\lambda_k x) &\to \omega(x) && \text{in } W^{2,2}_{\text{loc}}(\mathbb{R}^4, \mathbb{S}^n). \end{aligned}$$

Presume further that ω is the only "bubble" at the origin. Let

$$A_k(\delta, R) := \{x : \lambda_k R \le |x| \le \delta\}.$$

Then

$$\lim_{R\to\infty}\lim_{\delta\to 0}\lim_{k\to\infty} \big(\|D^2 u_k\|_{L^2(A_k(\delta,R))} + \|Du_k\|_{L^4(A_k(\delta,R))} + \|D\Delta u_k\|_{L^{4/3}(A_k(\delta,R))}\big) = 0.$$

The proposition also holds if u_k is a sequence of f_k-approximate intrinsic biharmonic maps.

Proof. We first prove that for any $\varepsilon > 0$ there exists K sufficiently large and δ small so that, for all $k \geq K$ and $\rho_k > 0$ such that $B_{4\rho_k} \setminus B_{\rho_k/2} \subset A_k(\delta, R)$,

$$\|D^2 u_k\|_{L^2(B_{2\rho_k}\setminus B_{\rho_k})} + \|Du_k\|_{L^4(B_{2\rho_k}\setminus B_{\rho_k})} + \|D\Delta u_k\|_{L^{4/3}(B_{2\rho_k}\setminus B_{\rho_k})} < \varepsilon. \tag{3-2}$$

Since $\{0\}$ is the only point of energy concentration, the strong convergence of $D^2 u_k \to D^2 u$ in L^2 and $Du_k \to Du$ in L^4 implies that for any $\varepsilon > 0$ and any $m \in \mathbb{Z}^+$ and δ sufficiently small, there exists $K := K(m)$ sufficiently large such that, for all $k \geq K(m)$,

$$\|D^2 u_k\|_{L^2(B_{2\delta}\setminus B_{\delta 2^{-m-1}})} + \|Du_k\|_{L^4(B_{2\delta}\setminus B_{\delta 2^{-m-1}})} \leq \frac{\varepsilon}{C\Gamma^{m+3}}. \tag{3-3}$$

Here C is an appropriately large constant determined by the bounds of Proposition 2.1 and Γ is the number of balls of radius $r/32$ needed to cover $B_r \setminus B_{r/2}$. By (2-4), for any $x \in B_{2\delta} \setminus B_{\delta 2^{-m-1}}$ and $0 < r < \delta 2^{-m-1}$,

$$\|D\Delta u_k\|_{L^{4/3}(B_{r/32}(x))} \leq C\big(\|D^2 u_k\|_{L^2(B_{r/2}(x))} + \|Du_k\|_{L^4(B_{r/2}(x))} + \|f_k\|^{1/4}_{L^1(B_{r/2}(x))}\|f_k\|^{3/4}_{L\log L(B_{r/2}(x))}\big) \tag{3-4}$$

Since Lemma A.2 and (1-1) imply that

$$\|f_k\|_{L^1(B_{r/2}(x))} \leq C\Big(\log\frac{1}{r}\Big)^{-1}\|f_k\|_{L\log L(B_{r/2}(x))}, \tag{3-5}$$

for sufficiently small δ, (3-3), (3-4), and (3-5) together imply that for $k \geq K(m)$

$$\|D\Delta u_k\|_{L^{4/3}(B_{2\delta}\setminus B_{\delta 2^{-m-1}})} + \|Du_k\|_{L^4(B_{2\delta}\setminus B_{\delta 2^{-m-1}})} + \|D^2 u_k\|_{L^2(B_{2\delta}\setminus B_{\delta 2^{-m-1}})} \leq \tfrac{1}{2}\varepsilon. \tag{3-6}$$

A similar argument (perhaps requiring a larger K) implies that

$$\|D\Delta u_k\|_{L^{4/3}(B_{2^m\lambda_k R}\setminus B_{\lambda_k R})} + \|Du_k\|_{L^4(B_{2^m\lambda_k R}\setminus B_{\lambda_k R})} + \|D^2 u_k\|_{L^2(B_{2^m\lambda_k R}\setminus B_{\lambda_k R})} \leq \tfrac{1}{2}\varepsilon. \tag{3-7}$$

Now suppose there exists a sequence t_k with $\lambda_k R < t_k < \delta$ such that

$$\|D^2 u_k\|_{L^2(B_{2t_k}\setminus B_{t_k})} + \|Du_k\|_{L^4(B_{2t_k}\setminus B_{t_k})} + \|D\Delta u_k\|_{L^{4/3}(B_{2t_k}\setminus B_{t_k})} \geq \varepsilon.$$

By (3-6) and (3-7), $t_k \to 0$ and $B_{\delta/t_k} \setminus B_{\lambda_k R/t_k} \to \mathbb{R}^4\setminus\{0\}$. Define $v_k(x) = u_k(t_k x)$ and $\tilde{f}_k(x) := t_k^4 f_k(t_k x)$. Then v_k is an $\tilde{f}_k$-approximate biharmonic map, defined on

$B_{t_k^{-1}}$. We first observe that $v_k \to v_\infty$ weakly in $W^{2,2}_{\mathrm{loc}}(\mathbb{R}^4, \mathbb{S}^n)$. Notice for any $R > 0$

$$\begin{aligned}\int_{B_R} |\tilde{f}_k(x)|\,dx &= \int_{B_{Rt_k}} |f_k(s)|\,ds \\ &\le \int_0^{|B_{Rt_k}|} (f_k)^*(t)\,dt \\ &\le c\left(\log\left(2+\frac{1}{Rt_k}\right)\right)^{-1}\int_0^\infty (f_k)^*(t)\log\left(2+\frac{1}{t}\right)dt \\ &= c\left(\log\left(2+\frac{1}{Rt_k}\right)\right)^{-1}\|f_k\|_{L\log L(B_1)}.\end{aligned}$$

By (1-1), $\tilde{f}_k \to 0$ in $L^1_{\mathrm{loc}}(\mathbb{R}^4)$. Moreover, for all k,

$$\|D^2 v_k\|_{L^2(B_2\setminus B_1)} + \|Dv_k\|_{L^4(B_2\setminus B_1)} + \|D\Delta v_k\|_{L^{4/3}(B_2\setminus B_1)} \ge \varepsilon.$$

If $v_k \to v_\infty$ strongly in $W^{2,2}(B_{16}\setminus B_{1/16}, \mathbb{S}^n)$, then v_∞ is a nonconstant biharmonic map into $\mathbb{S}^n$. Note that by Proposition 2.1 we get

$$\|D^2 v_\infty\|_{L^2(B_2\setminus B_1)} + \|Dv_\infty\|_{L^4(B_2\setminus B_1)} > 0.$$

This contradicts the fact that there is only one bubble at $\{0\}$. If the convergence is not strong, then Lemma 3.3 implies that the energy must concentrate. That is, there exists a subsequence v_k such that $\|D^2 v_k\|_{L^2(B_r(x))} + \|Dv_k\|_{L^4(B_r(x))} \ge \varepsilon_0^2$ for all $r > 0$. This also contradicts the existence of only one bubble. Thus, (3-2) holds.

Using the duality of Lorentz spaces and the estimates of Section A.2, we get the bounds

$$\begin{aligned}\|D^2 u_k\|_{L^2}^2 &\le C\|D^2 u_k\|_{L^{2,\infty}}\|D^2 u_k\|_{L^{2,1}}, \\ \|Du_k\|_{L^4}^4 &\le C\|\,|Du_k|^3\|_{L^{4/3,\infty}}\|Du_k\|_{L^{4,1}} \\ &\le C\|Du_k\|_{L^{4,\infty}}^3\|Du_k\|_{L^{4,1}}, \\ \|D\Delta u_k\|_{L^{4/3}}^{4/3} &\le C\|(D\Delta u_k)^{1/3}\|_{L^{4,\infty}}\|D\Delta u_k\|_{L^{4/3,1}} \\ &\le C\|D\Delta u_k\|_{L^{4/3,\infty}}^{1/3}\|D\Delta u_k\|_{L^{4/3,1}}.\end{aligned} \tag{3-8}$$

Using (1-1) and (2-1), we observe that

$$\|D^2 u_k\|_{L^{2,1}} + \|Du_k\|_{L^{4,1}} + \|D\Delta u_k\|_{L^{4/3,1}} \le C\Lambda.$$

Since (3-2) allows us to apply Lemma 2.4, appealing to (3-8) implies the result. □

The full proof of Theorem 1.2 now follows immediately from the uniform energy bounds of (1-1), the small-energy compactness results of this section, and standard induction arguments on the bubbles.

4. Oscillation bounds

The proof of the following oscillation lemma will constitute the work of this section:

Lemma 4.1. *Let $u \in W^{2,2}(B_1, \mathbb{S}^n)$ be an f-approximate biharmonic map for $f \in L\log L(B_1, \mathbb{R}^{n+1})$ with*

$$\|D^2 u\|_{L^2(B_1)} + \|Du\|_{L^4(B_1)} + \|f\|_{L\log L(B_1)} \le \Lambda < \infty.$$

Then for $0 < 2t < \delta/2 < 1/16$,

$$\begin{aligned}\sup_{x,y\in B_{\delta/2}\setminus B_{2t}} &|u(x)-u(y)| \\ &\le C\big(\|D^2u\|_{L^2(B_{2\delta}\setminus B_t)} + \|Du\|_{L^4(B_{2\delta}\setminus B_t)} + \|f\|_{L\log L(B_{2\delta})} \\ &\qquad + \|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)} + \|D\Delta u\|_{L^{4/3,1}(B_{2t}\setminus B_t)} + |B_{4\delta}|\big).\end{aligned}$$

The lemma also holds if u is an f-approximate intrinsic biharmonic map.

Consider the map $u_1 : B_1 \to \mathbb{R}^{n+1}$ such that $u_1(\boldsymbol{x}) = \boldsymbol{b} + A\boldsymbol{x}$, where $\boldsymbol{b} \in \mathbb{R}^{n+1}$ and A is an $(n+1)\times 4$ matrix with

$$A := \fint_{B_{2t}\setminus B_t} Du \quad \text{and} \quad \boldsymbol{b} := \fint_{B_{2t}\setminus B_t} (u(\boldsymbol{x}) - A\boldsymbol{x})\, d\mathrm{Vol}(\boldsymbol{x}).$$

Then by construction

$$\fint_{B_{2t}\setminus B_t} u - u_1 = 0, \quad \fint_{B_{2t}\setminus B_t} Du - Du_1 = 0, \quad D^k u_1 \equiv 0 \text{ for all } k \ge 2.$$

Set $w = (1-\phi_t)(u-u_1)$. Let $w_1 : B_1 \to \mathbb{R}^{n+1}$ such that $w_1(\boldsymbol{x}) = \boldsymbol{m} + N\boldsymbol{x}$, where

$$N := \fint_{B_\delta\setminus B_{\delta/2}} Dw \quad \text{and} \quad \boldsymbol{m} := \fint_{B_\delta\setminus B_{\delta/2}} (w(\boldsymbol{x}) - N\boldsymbol{x})\, d\mathrm{Vol}(\boldsymbol{x}).$$

Let $\widetilde{w} = (w - w_1)\phi_{\delta/2}$, so $\widetilde{w} = w - w_1$ on $B_{\delta/2}$ and the support of $\widetilde{w}$ is contained in B_δ.

By definition,

$$\begin{aligned}\sup_{x,y\in B_{\delta/2}\setminus B_{2t}} |u(x)-u(y)| &= \sup_{x,y\in B_{\delta/2}\setminus B_{2t}} |w(x)-w(y)+u_1(x)-u_1(y)| \\ &= \sup_{x,y\in B_{\delta/2}\setminus B_{2t}} |(\widetilde{w}+u_1+w_1)(x)-(\widetilde{w}+u_1+w_1)(y)| \\ &\le 2 \sup_{x\in B_{\delta/2}\setminus B_{2t}} |\widetilde{w}(x)-\widetilde{w}(0)+(A+N)x|.\end{aligned}$$

We first observe that, outside of B_{2t}, $w = u - u_1$ so the definition of N implies that

$$A + N = A + \fint_{B_\delta\setminus B_{\delta/2}} Du - \fint_{B_\delta\setminus B_{\delta/2}} A = \fint_{B_\delta\setminus B_{\delta/2}} Du.$$

Thus, for $x \in B_{\delta/2}$, Hölder's inequality implies that

$$|(A+N)x| \le C\delta^{-3} \int_{B_\delta \setminus B_{\delta/2}} |Du| \le C\|Du\|_{L^4(B_\delta \setminus B_{\delta/2})}.$$

As before, let G be the distribution in $\mathbb{R}^4$ such that $\Delta^2 G = \delta_0$. Then $G(x) = C \log|x|$, and recall that $DG \in L^{4,\infty}(\mathbb{R}^4)$. It is enough to show that:

Claim 4.2. $$\left|\widetilde{w}(x) - \fint_{\mathbb{R}^4} \widetilde{w}\right| \le C\|D\Delta\widetilde{w}\|_{L^{4/3,1}(\mathbb{R}^4)}.$$

Since all of the above quantities are translation-invariant, we may assume $x = 0$. Then

$$\begin{aligned}\left|\widetilde{w}(0) - \fint \widetilde{w}\right| &= \left|\int_{\mathbb{R}^4} \Delta^2 G(y)\left(\widetilde{w}(y) - \fint \widetilde{w}\right) dV(y)\right| \\ &= \left|\int_{\mathbb{R}^4} DG(y) D\Delta\widetilde{w}(y)\, dV(y)\right| \\ &\le C\|DG\|_{L^{4,\infty}(\mathbb{R}^4)}\|D\Delta\widetilde{w}\|_{L^{4/3,1}(\mathbb{R}^4)}.\end{aligned}$$

Using the definition of $\widetilde{w}$,

$$\begin{aligned}&\|D\Delta\widetilde{w}\|_{L^{4/3,1}(\mathbb{R}^4)} \\ &\quad\le C\|(\delta^{-3}|w-w_1| + \delta^{-2}|D(w-w_1)| + \delta^{-1}|D^2 w|)\|_{L^{4/3,1}(B_\delta \setminus B_{\delta/2})} \\ &\qquad\qquad + C\|D\Delta w\|_{L^{4/3,1}(B_\delta)}.\end{aligned}$$

Interpolation techniques and Poincaré's inequality imply that

$$\begin{aligned}\|\delta^{-3}(w-w_1)\|_{L^{4/3,1}(B_\delta \setminus B_{\delta/2})} &\le C\|\delta^{-2} D(w-w_1)\|_{L^{4/3,1}(B_\delta \setminus B_{\delta/2})} \\ &\le C\|\delta^{-1} D^2 w\|_{L^{4/3,1}(B_\delta \setminus B_{\delta/2})}.\end{aligned}$$

Moreover, the embedding theorems for Lorentz spaces imply that

$$\|\delta^{-1} D^2 w\|_{L^{4/3,1}(B_\delta \setminus B_{\delta/2})} \le C\|D^2 w\|_{L^2(B_\delta \setminus B_{\delta/2})}.$$

Therefore,

$$\|D\Delta\widetilde{w}\|_{L^{4/3,1}(\mathbb{R}^4)} \le C\|D^2 w\|_{L^2(B_\delta \setminus B_{\delta/2})} + C\|D\Delta w\|_{L^{4/3,1}(B_\delta)}. \tag{4-1}$$

Since $D^2 w = D^2 u$ on $B_\delta \setminus B_{2t}$, we conclude that

$$\begin{aligned}&\operatorname{osc}_{B_{\delta/2} \setminus B_{2t}} u \\ &\qquad\le C(\|D\Delta w\|_{L^{4/3,1}(B_\delta)} + \|D^2 u\|_{L^2(B_\delta \setminus B_{\delta/2})} + \|Du\|_{L^4(B_\delta \setminus B_{\delta/2})}).\end{aligned} \tag{4-2}$$

The remainder of the proof will be devoted to bounding the $D\Delta w$ term.

We define $\beta = D\Delta w \wedge u - \Delta w \wedge Du$. Then

$$\beta^{ij} := u^j D\Delta w^i - u^i D\Delta w^j - \Delta w^i Du^j + \Delta w^j Du^i \in \Omega^1 \mathbb{R}^4$$

for $i, j = 1, \dots, n+1$. By definition $\beta = D\Delta u \wedge u - \Delta u \wedge Du$ in $B_\delta \setminus B_{2t}$ and thus $d^*\beta = f \wedge u$ in $B_\delta \setminus B_{2t}$. We will require an $L^{4/3}$ bound for β, and to that end note that

$$\begin{aligned} \|\beta\|_{L^{4/3}(B_{2\delta})} &\le C(\|D\Delta w\|_{L^{4/3}(B_{2\delta})} + \|\Delta w \wedge Du\|_{L^{4/3}(B_{2\delta})}) \\ &\le C(\|D\Delta w\|_{L^{4/3}(B_{2\delta})} + \|\Delta w\|_{L^2(B_{2\delta})}\|Du\|_{L^4(B_{2\delta})}) \\ &\le C(\|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)} + \|D^2 u\|_{L^2(B_{2\delta}\setminus B_t)}). \end{aligned} \tag{4-3}$$

For the last inequality, $\|Du\|_{L^4(B_{2\delta})}$ is bounded and is absorbed into the constant. In addition, we use the definition of w and repeated applications of Poincaré and Hölder to determine

$$\begin{aligned} \|D\Delta w\|_{L^{4/3}(B_{2\delta})} &\le C(\|D^2 u\|_{L^2(B_{2t}\setminus B_t)} + \|(1-\phi_t)D\Delta u\|_{L^{4/3}(B_{2\delta})}), \\ \|\Delta w\|_{L^2(B_{2\delta})} &\le C\|D^2 u\|_{L^2(B_{2\delta}\setminus B_t)}. \end{aligned}$$

Set

$$\gamma := d^*(D\Delta(w-u) \wedge u - \Delta(w-u) \wedge Du).$$

Then

$$\begin{aligned} d^*\beta = f \wedge u + \gamma, \qquad\qquad d\beta = -2D\Delta w \wedge Du, \\ \Delta\beta = (dd^* + d^*d)\beta = d(f \wedge u + \gamma) + d^*(-2D\Delta w \wedge Du). \end{aligned}$$

We consider a decomposition $\beta^{ij} = H^{ij} + d\Psi^{ij} + d^*\Phi^{ij}$ for each component β^{ij}, where H^{ij} is a harmonic 1-form and Φ, Ψ satisfy appropriate partial differential equations. Our objective is to bound $\|D\Delta w\|_{L^{4/3,1}}$ by $\|\beta\|_{L^{4/3,1}}$, and to that end we determine such bounds for $d\Psi$, $d^*\Phi$, and H.

Remark 4.3. For the intrinsic case, we modify a few definitions. Let $\beta_I := \beta + 2|Du|^2 Dw_I \wedge u$, where $w_I = (1-\phi_t)(u - \boldsymbol{d})$ and $\boldsymbol{d} := \fint_{B_{2t}\setminus B_t} u$. Using the definition of w_I, we get the bound $\|\beta_I\|_{L^{4/3}(B_{2\delta})} \le \|\beta\|_{L^{4/3}(B_{2\delta})} + C\|Du\|_{L^4(B_{2\delta}\setminus B_t)}$ by using Hölder's inequality and Poincaré's inequality. We then define $\gamma_I := \gamma + d^*(2|Du|^2 D(w_I - u) \wedge u)$, and thus

$$d^*\beta_I = f \wedge u + \gamma_I \quad \text{and} \quad d\beta_I = d\beta + D(|Du|^2)Dw_I \wedge u - |Du|^2 Dw_I \wedge Du.$$

We now continue with the proof for the extrinsic case:

Proposition 4.4. *Let Ψ^{ij} be a function on $B_{2\delta}$ satisfying*

$$\begin{cases} \Delta\Psi^{ij} = f^i u^j - f^j u^i + \gamma^{ij} & \text{in } B_{2\delta}, \\ \Psi^{ij} = 0 & \text{on } \partial B_{2\delta}. \end{cases}$$

Then

$$\begin{aligned}\|d\Psi^{ij}\|&_{L^{4/3,1}(B_{2\delta})}\\ &\le C\big(\|D^2u\|_{L^2(B_{2t}\setminus B_t)}+\|Du\|_{L^4(B_{2t}\setminus B_t)}+\|D\Delta u\|_{L^{4/3}(B_{2t}\setminus B_t)}\\ &\qquad\qquad+\|f\|_{L\log L(B_{2\delta})}+|B_{4\delta}|\big).\end{aligned}$$

Proof. We decompose $\Psi^{ij}=\Psi_1^{ij}+\Psi_2^{ij}$ so that

$$\begin{cases}\Delta\Psi_1^{ij}=\gamma^{ij} & \text{in } B_{2\delta},\\ \quad\Psi_1^{ij}=0 & \text{on } \partial B_{2\delta}.\end{cases}$$

Following classical arguments,

$$\|D^2\Psi_1^{ij}\|_{L^1(B_{2\delta})}\le C\|\gamma^{ij}\|_{\mathcal{H}^1(B_{2\delta})}.$$

Thus the embedding theorems imply that $\|D\Psi_1^{ij}\|_{L^{4/3,1}(B_{2\delta})}\le C\|\gamma^{ij}\|_{\mathcal{H}^1(B_{2\delta})}$. Now we consider the $\mathcal{H}^1$ norm of γ^{ij}. By definition,

$$\begin{aligned}\gamma^{ij}&=d^*(D\Delta(w^i-u^i)u^j-D\Delta(w^j-u^j)u^i-[\Delta(w^i-u^i)Du^j-\Delta(w^j-u^j)Du^i])\\ &=\Delta^2(w^i-u^i)u^j-\Delta^2(w^j-u^j)u^i-(\Delta(w^i-u^i)\Delta u^j-\Delta(w^j-u^j)\Delta u^i).\end{aligned}$$

Recall that $w:=(1-\phi_t)(u-u_1)$. So

$$\begin{aligned}\Delta(w^j-u^j)&=-\Delta\phi_t(u^j-u_1^j)-2D\phi_t\cdot D(u^j-u_1^j)-\phi_t\Delta u^j,\\ \Delta^2(w^j-u^j)&=-\Delta^2\phi_t(u^j-u_1^j)-\Delta\phi_t\Delta u^j-2D\Delta\phi_t D(u^j-u_1^j)\\ &\quad-2\Delta(D\phi_t\cdot D(u^j-u_1^j))-\Delta\phi_t\Delta u^j-2D\phi_t D\Delta u^j-\phi_t\Delta^2u^j.\end{aligned}$$

Combining all of the terms, we estimate

$$\begin{aligned}|\gamma^{ij}|&\le C|D^4\phi_t|\,|u-u_1|+C|D^3\phi_t|\,|D(u-u_1)|+C|D^2\phi_t|\,|D^2u|\\ &\quad+C|D\phi_t|(|D\Delta u|+|D(u-u_1)|\,|\Delta u|)+|\phi_t||u^i\Delta^2u^j-u^j\Delta^2u^i|.\end{aligned}$$

The definition of γ^{ij} implies that $\gamma^{ij}=0$ on $\mathbb{R}^4\setminus B_{2t}$ and

$$\int_{\mathbb{R}^4}\gamma^{ij}=\int_{\partial B_{2t}}(D\Delta(w-u)\wedge u-\Delta(w-u)\wedge Du)^{ij}\cdot\boldsymbol{n}=0.$$

The estimate from Lemma A.1 implies that

$$\begin{aligned}\|\gamma^{ij}\|_{\mathcal{H}^1(B_{2\delta})}&\le c\big(t\|\gamma^{ij}-\phi_t(u^j\Delta^2u^i-u^i\Delta^2u^j)\|_{L^{4/3}(B_{2t})}\\ &\qquad+\|\phi_t(u^j\Delta^2u^i-u^i\Delta^2u^j)\|_{L\log L(B_{2\delta})}+|B_{4\delta}|\big).\end{aligned}$$

Repeating techniques used previously, we bound the first three terms of $|\gamma^{ij}|$:

$$\begin{aligned} t^{-4}\|u-u_1\|_{L^{4/3}(B_{2t}\setminus B_t)} &\le Ct^{-3}\|D(u-u_1)\|_{L^{4/3}(B_{2t}\setminus B_t)} \\ &\le Ct^{-2}\|D^2u\|_{L^{4/3}(B_{2t}\setminus B_t)} \le Ct^{-1}\|D^2u\|_{L^2(B_{2t}\setminus B_t)}. \end{aligned}$$

We will preserve the term

$$t^{-1}\|D\Delta u\|_{L^{4/3}(B_{2t}\setminus B_t)},$$

as our energy quantization result implies that this term will vanish when taking limits. Hölder's inequality and the fact that $\|D(u-u_1)\|_{L^4(B_{2t}\setminus B_t)} \le C\|Du\|_{L^4(B_{2t}\setminus B_t)}$ imply that

$$\begin{aligned} \|D(u-u_1)\Delta u\|_{L^{4/3}(B_{2t}\setminus B_t)} &\le C\|D(u-u_1)\|_{L^4(B_{2t}\setminus B_t)}\|D^2u\|_{L^2(B_{2t}\setminus B_t)} \\ &\le C\|D^2u\|_{L^2(B_{2t}\setminus B_t)}. \end{aligned}$$

For the last term, since u is an f-approximate biharmonic map into $\mathbb{S}^n$,

$$\|\phi_t(\Delta^2u\wedge u)\|_{L\log L(B_{2\delta})} \le \|f\wedge u\|_{L\log L(B_{2\delta})} \le \|f\|_{L\log L(B_{2\delta})}.$$

All of the above estimates imply that

$$\begin{aligned} \|\gamma^{ij}\|_{\mathcal{H}^1(B_{2\delta})} \le C\big(&\|D^2u\|_{L^2(B_{2t}\setminus B_t)} + \|Du\|_{L^4(B_{2t}\setminus B_t)} \\ &+ \|D\Delta u\|_{L^{4/3}(B_{2t}\setminus B_t)} + \|f\|_{L\log L(B_{2\delta})} + |B_{4\delta}|\big). \end{aligned}$$

Finally, consider

$$\begin{cases} \Delta\Psi_2^{ij} = f^iu^j - u^if^j & \text{in } B_{2\delta}, \\ \Psi_2^{ij} = 0 & \text{on } \partial B_{2\delta}. \end{cases}$$

Then classical results give $\|\Psi_2^{ij}\|_{W^{2,1}(B_{2\delta})} \le C\|f\|_{\mathcal{H}^1(B_{2\delta})} \le C\|f\|_{L\log L(B_{2\delta})}$. Thus

$$\|d\Psi_2^{ij}\|_{W^{1,1}(B_{2\delta})} \le C\|f\|_{L\log L(B_{2\delta})},$$

and the embedding theorems in $\mathbb{R}^4$ imply that

$$\|d\Psi_2^{ij}\|_{L^{4/3,1}(B_{2\delta})} \le C\|f\|_{L\log L(B_{2\delta})}. \qquad \square$$

Remark 4.5. For the intrinsic case, we define

$$\begin{aligned} \gamma_I &= \gamma + d^*(2|Du|^2D(w_I-u)\wedge u) \\ &= \gamma - 2\phi_t d^*(|Du|^2Du\wedge u) + 2|Du|^2(\Delta\phi_t(\boldsymbol{d}-u)\wedge u - D\phi_t\cdot Du\wedge(\boldsymbol{d}+u)) \\ &\qquad + 2D|Du|^2\cdot D\phi_t(\boldsymbol{d}-u)\wedge u. \end{aligned}$$

We bound $\|\gamma_I\|_{\mathcal{H}^1}$ by making some observations: First, $-2\phi_t d^*(|Du|^2Du\wedge u)$ is added to the term $-\phi_t\Delta^2u\wedge u$ that appears in the expansion of γ. We then make the substitution $-\phi_t f\wedge u$ as in the extrinsic case. Second, using Poincaré's inequality, Hölder's inequality, and the global energy bound for u, the $L^{4/3}$ norm of what remains is bounded by $Ct^{-1}(\|Du\|_{L^4(B_{2t}\setminus B_t)} + \|D^2u\|_{L^2(B_{2t}\setminus B_t)})$. Finally, observe that, by construction, γ_I is supported on B_{2t} and $\int_{\mathbb{R}^4}\gamma_I = 0$, so the estimate used for $\|\gamma\|_{\mathcal{H}^1}$ still applies.

Proposition 4.6. *Let $\Phi^{ij} \in \Omega^2\mathbb{R}^4$ be the solution to the system*

$$\begin{cases} \Delta\Phi^{ij} = -2(D\Delta w^i\, Du^j - D\Delta w^j\, Du^i) & \text{in } B_{2\delta}, \\ \Phi^{ij} = 0 & \text{on } \partial B_{2\delta}. \end{cases}$$

Then

$$\|d^*\Phi^{ij}\|_{L^{4/3,1}(B_{2\delta})} \le C(\|D^2u\|_{L^2(B_{2t}\setminus B_t)} + \|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)}). \tag{4-4}$$

Proof. Using the same techniques and estimates as in the previous proposition, we note that

$$\begin{aligned} \|d\Phi^{ij}\|_{L^{4/3,1}(B_{2\delta})} &\le C\|D\Delta w \wedge Du\|_{\mathcal{H}^1(B_{2\delta})} \\ &\le C\|D\Delta w\|_{L^{4/3}(B_{2\delta})}\|Du\|_{L^4(B_{2\delta})} \\ &\le C(\|D^2u\|_{L^2(B_{2t}\setminus B_t)} + \|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)}). \qquad \square \end{aligned}$$

Remark 4.7. In the intrinsic setting the steps of the proof are the same, though the equation for $\Delta\Phi^{ij}_I$ includes the terms $D(|Du|^2)Dw_I \wedge u - |Du|^2 Dw_I \wedge Du$. Since $\|Dw_I\|_{L^4(B_{2\delta})} \le C\|Du\|_{L^4(B_{2\delta}\setminus B_t)}$, one can quickly show the intrinsic bound has the form

$$\|d^*\Phi_I\|_{L^{4/3,1}(B_{2\delta})} \le \|d^*\Phi\|_{L^{4/3,1}(B_{2\delta})} + C\|Du\|_{L^4(B_{2\delta}\setminus B_t)}.$$

Now consider the harmonic 1-form

$$H^{ij} = \beta^{ij} - d^*\Phi^{ij} - d\Psi^{ij}.$$

Propositions 4.4 and 4.6, along with (4-3), imply that

$$\begin{aligned} \|H\|_{L^{4/3}(B_{2\delta})} &\le \|\beta\|_{L^{4/3}(B_{2\delta})} + \|d^*\Phi\|_{L^{4/3}(B_{2\delta})} + \|d\Psi\|_{L^{4/3}(B_{2\delta})} \\ &\le C\big(\|D^2u\|_{L^2(B_{2\delta}\setminus B_t)} + \|Du\|_{L^4(B_{2\delta}\setminus B_t)} \\ &\qquad + \|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)} + \|f\|_{L\log L(B_{2\delta})} + |B_{4\delta}|\big). \end{aligned}$$

The mean value property and Hölder's inequality together imply that

$$\begin{aligned} \|H^{ij}\|_{C^0(B_\delta)} \le \frac{C}{\delta^3}\big(&\|D^2u\|_{L^2(B_{2\delta}\setminus B_t)} + \|Du\|_{L^4(B_{2\delta}\setminus B_t)} \\ &+ \|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)} + \|f\|_{L\log L(B_{2\delta})} + |B_{4\delta}|\big). \end{aligned}$$

Moreover, a straightforward calculation implies that

$$\|H^{ij}\|_{L^{4/3,1}(B_\delta)} \le C\delta^3\|H^{ij}\|_{C^0(B_\delta)}.$$

Thus,

$$\|\beta\|_{L^{4/3,1}(B_\delta)} \le C\big(\|D^2u\|_{L^2(B_{2\delta}\setminus B_t)} + \|Du\|_{L^4(B_{2\delta}\setminus B_t)} + \|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)} + \|f\|_{L\log L(B_{2\delta})} + |B_{4\delta}|\big).$$

Using the appropriate harmonic 1-form H_I, we produce an identical estimate for β_I.

We now use the definitions of w and β to determine a bound on $\|D\Delta w\|_{L^{4/3,1}(B_\delta)}$. First we consider the function on B_{2t}

$$\begin{aligned}\|D\Delta w\|_{L^{4/3,1}(B_{2t})} &\le \|C(t^{-3}|u-u_1| + t^{-2}|D(u-u_1)| + t^{-1}|D^2u|)\|_{L^{4/3,1}(B_{2t}\setminus B_t)} \\ &\qquad + \|(1-\phi_t)D\Delta u\|_{L^{4/3,1}(B_{2t})} \\ &\le C\|D^2u\|_{L^2(B_{2t}\setminus B_t)} + C\|D\Delta u\|_{L^{4/3,1}(B_{2t}\setminus B_t)}.\end{aligned}$$

On $B_\delta \setminus B_{2t}$, $w = u - u_1$ so $D\Delta w \equiv D\Delta u$. We first decompose $D\Delta u$ into tangential and normal parts with tangency relative to the target manifold $\mathbb{S}^n$. Then

$$D\Delta u = D\Delta u^T + D\Delta u^N = D\Delta u \wedge u.u + \langle D\Delta u, u\rangle u.$$

Here we define $\langle Dv, u\rangle := \sum_{i,k}(\partial v^k/\partial x_i)u^k\, dx_i$. On $B_\delta \setminus B_{2t}$, $D\Delta u \wedge u = \beta + \Delta u \wedge Du$, and thus

$$|(D\Delta u)^T| \le |\beta| + |\Delta u|\,|Du|.$$

Since

$$\begin{aligned}\langle D\Delta u, u\rangle = D\langle \Delta u, u\rangle - \langle \Delta u, Du\rangle &= D(d^*\langle Du, u\rangle - |Du|^2) - \langle \Delta u, Du\rangle \\ &= -D|Du|^2 - \langle \Delta u, Du\rangle,\end{aligned}$$

we estimate

$$\begin{aligned}\|D\Delta w\|_{L^{4/3,1}(B_\delta\setminus B_{2t})} &\le C\|\beta\|_{L^{4/3,1}(B_\delta)} + C\|D^2u\|_{L^2(B_\delta\setminus B_{2t})}\|Du\|_{L^4(B_\delta\setminus B_{2t})} \\ &\le C\big(\|D^2u\|_{L^2(B_{2\delta}\setminus B_t)} + \|Du\|_{L^4(B_{2\delta}\setminus B_t)} + \|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)} \\ &\qquad + \|f\|_{L\log L(B_{2\delta})} + |B_{4\delta}|\big).\end{aligned}$$

Thus,

$$\begin{aligned}\|D\Delta w\|_{L^{4/3,1}(B_\delta)} \le C(&\|D^2u\|_{L^2(B_{2\delta}\setminus B_t)} + \|Du\|_{L^4(B_{2\delta}\setminus B_t)} + \|f\|_{L\log L(B_{2\delta})} \\ &+ \|D\Delta u\|_{L^{4/3,1}(B_{2t}\setminus B_t)} + \|D\Delta u\|_{L^{4/3}(B_{2\delta}\setminus B_t)} + |B_{4\delta}|).\end{aligned}$$

Inserting this inequality into (4-2) proves the oscillation lemma.

Remark 4.8. To complete the proof in the intrinsic case, observe that, on $B_\delta \setminus B_{2t}$, $D\Delta w \wedge u = D\Delta u \wedge u = \beta + \Delta u \wedge Du + 2|Du|^2 Du \wedge u$. This changes the L^∞ estimate for $|(D\Delta u)^T|$ on $B_\delta \setminus B_{2t}$, but using embedding theorems for Lorentz spaces we note that the $L^{4/3,1}$ estimate is unchanged.

5. No-neck property — proof of Theorem 1.3

The proof of the no-neck property now follows easily from combining the energy quantization and the oscillation bounds.

Proof. As we may use induction to deal with the case of multiple bubbles, we prove the theorem for one bubble. Let λ_k be such that $\tilde{u}_k(x) := u_k(\lambda_k x) \to \omega(x) \in W^{2,2}_{\mathrm{loc}}(\mathbb{R}^4, \mathbb{S}^n)$. Since each of the $u_k \in W^{2,2}(B_1, \mathbb{S}^n)$ are f_k-approximate biharmonic maps with $f_k \in L\log L(B_1, \mathbb{R}^{n+1})$ and have uniform energy bounds, Lemma 4.1 implies that

$$\begin{aligned}\sup_{x,y\in B_{\delta/2}\setminus B_{2\lambda_k R}} &|u_k(x)-u_k(y)| \\ &\le C\big(\|D^2u_k\|_{L^2(B_{2\delta}\setminus B_{\lambda_k R/2})} + \|Du_k\|_{L^4(B_{2\delta}\setminus B_{\lambda_k R/2})} \\ &\qquad + \|f_k\|_{L\log L(B_{2\delta})} + \|D\Delta u_k\|_{L^{4/3,1}(B_{\lambda_k R}\setminus B_{\lambda_k R/2})} \\ &\qquad\qquad + \|D\Delta u_k\|_{L^{4/3}(B_{2\delta}\setminus B_{\lambda_k R/2})} + |B_{4\delta}|\big).\end{aligned}$$

Theorem 1.2 implies that

$$\begin{aligned}\lim_{\delta\to 0}\lim_{R\to\infty}\lim_{k\to\infty}\big(\|D^2u_k\|_{L^2(B_{2\delta}\setminus B_{\lambda_k R/2})} &+ \|Du_k\|_{L^4(B_{2\delta}\setminus B_{\lambda_k R/2})} \\ &+ \|D\Delta u_k\|_{L^{4/3}(B_{2\delta}\setminus B_{\lambda_k R/2})}\big) = 0.\end{aligned}$$

Further, (2-1) and Hölder's inequality imply that

$$\begin{aligned}\|D\Delta u_k\|_{L^{4/3,1}(B_{\lambda_k R}\setminus B_{\lambda_k R/2})} &\le C(\|Du_k\|_{L^4(B_{2\lambda_k R}\setminus B_{\lambda_k R/4})} \\ &\quad + \|D^2u_k\|_{L^2((B_{2\lambda_k R}\setminus B_{\lambda_k R/4})} + \|f_k\|_{L\log L(B_{2\lambda_k R})}).\end{aligned}$$

Since we presume the $L\log L$ norm of f_k does not concentrate,

$$\lim_{\delta\to 0}\lim_{R\to\infty}\lim_{k\to\infty}\|f_k\|_{L\log L(B_{2\delta})} = 0.$$

Therefore,

$$\lim_{\delta\to 0}\lim_{R\to\infty}\lim_{k\to\infty}\|D\Delta u_k\|_{L^{4/3,1}(B_{\lambda_k R}\setminus B_{\lambda_k R/2})} = 0.$$

Taking all of the estimates together implies that

$$\lim_{\delta\to 0}\lim_{R\to\infty}\lim_{k\to\infty}\sup_{x,y\in B_{\delta/2}\setminus B_{2\lambda_k R}}|u_k(x)-u_k(y)| = 0.$$

Thus, no neck occurs in the blowup. □

Remark 5.1. For $f_k \in \phi(L)$, we use the estimate

$$
\begin{aligned}
&\|f_k\|_{L\log L(B_{2\delta})}\\
&\qquad = \int_{B_{2\delta}\cap\{|f_k|\le\delta^{-1}\}} |f_k|\log(2+|f_k|)\,dx + \int_{|f_k|>\delta^{-1}} |f_k|\log(2+|f_k|)\,dx\\
&\qquad \le C\delta^3\log(2+\delta^{-1}) + \sup_{t>\delta^{-1}} \frac{t\log(2+t)}{\phi(t)} \int_{|f_k|>\delta^{-1}} \phi(|f_k|)\,dx\\
&\qquad \le C\delta^3\log(2+\delta^{-1}) + \sup_{t>\delta^{-1}} \frac{t\log(2+t)}{\phi(t)}\Lambda.
\end{aligned}
$$

Since we presumed $\lim_{t\to\infty}\phi(t)/(t\log t) = \infty$, we determine

$$\lim_{\delta\to 0}\sup_k \|f_k\|_{L\log L(B_{2\delta})} = 0.$$

Appendix: Necessary background

A.1. ***Hardy spaces, Lorentz spaces, L*** **log** ***L, and Orlicz spaces.*** Let

$$T := \{\Phi \in C^\infty(\mathbb{R}^4) : \mathrm{spt}(\Phi) \subset B_1, \|\nabla\Phi\|_{L^\infty(\mathbb{R}^4)} \le 1\}.$$

For any $\Phi \in T$, let $\Phi_t(x) := t^{-4}\Phi(x/t)$. For each $f \in L^1(\mathbb{R}^4)$, let

$$f_*(x) = \sup_{\Phi\in T}\sup_{t>0} |(\Phi_t * f)(x)|.$$

Then f is in the *Hardy space* $\mathcal{H}^1(\mathbb{R}^4)$ if $f_* \in L^1(\mathbb{R}^4)$ and

$$\|f\|_{\mathcal{H}^1(\mathbb{R}^4)} = \|f_*\|_{L^1(\mathbb{R}^4)}.$$

Thus, one has the continuous embedding $\mathcal{H}^1 \hookrightarrow L^1$.

For a measurable function $f : \Omega \to \mathbb{R}$, let f^* denote the nonincreasing rearrangement of $|f|$ on $[0, |\Omega|)$ such that

$$|\{x \in \Omega : |f(x)| \ge s\}| = |\{t \in (0, |\Omega|) : f^*(t) \ge s\}|.$$

Let

$$f^{**}(t) := \frac{1}{t}\int_0^t f^*(s)\,ds.$$

For $p \in (1,\infty)$, let

$$\|f\|_{L^{p,q}} = \begin{cases} \int_0^\infty t^{1/p-1} f^{**}(t)\,dt & \text{if } q = 1,\\ \sup_{t>0} t^{1/p} f^{**}(t) & \text{if } q = \infty. \end{cases}$$

We will also occasionally exploit the fact that one may understand $\|f\|_{L^{p,\infty}}$ by understanding instead its seminorm

$$\|f\|^*_{L^{p,\infty}} := \sup_{\lambda>0} \lambda |\{x : |f(x)| > \lambda\}|^{1/p}.$$

We define the Banach spaces

$$L^{p,q} := \{f : \|f\|_{L^{p,q}} < \infty\}.$$

The spaces $L^{p,1}$ and $L^{p,\infty}$ are examples of *Lorentz spaces*, and can be thought of as interpolation spaces between the standard L^p spaces. For example, one observes that the following embeddings are all continuous

$$L^r(B_1) \hookrightarrow L^{p,1}(B_1) \hookrightarrow L^{p,p}(B_1) = L^p(B_1) \hookrightarrow L^{p,\infty}(B_1) \hookrightarrow L^q(B_1)$$

for all $q < p < r$ [Hélein 1990].

We define

$$L\log L := \left\{ f : \int |f(x)| \log(2 + |f(x)|)\, dx < \infty \right\}.$$

Since this is nonlinear, we will use the following seminorm which is equivalent to the norm for $L\log L$

$$\|f\|_{L\log L} := \int f^*(t) \log\Big(2 + \frac{1}{t}\Big)\, dt.$$

We also note that $L^p(B_1) \hookrightarrow L\log L(B_1) \hookrightarrow L^1(B_1)$ are continuous embeddings for all $p > 1$. Finally, we say f is in $\mathcal{H}^1(B_1)$ if

$$\left(f - \fint_{B_1} f(x)\, dx \right) \chi_{B_1} \in \mathcal{H}^1(\mathbb{R}^4).$$

We record here the often-used estimate

$$\|f\|_{\mathcal{H}^1(B_1)} \le C \|f\|_{L\log L(B_1)}. \tag{A-1}$$

Finally, for any increasing function $\phi : [0,\infty) \to [0,\infty)$ we define the Orlicz space

$$\phi(L) := \left\{ f : \|f\|_{\phi(L)} := \int \phi(|f|)\, dx < \infty \right\}.$$

Examples include the L^p spaces for $\phi(t) = t^p$ and $L\log L$ when $\phi(t) = t\log(2+t)$.

A.2. ***Embeddings and estimates for Lorentz spaces.*** We will frequently use the following facts about Lorentz spaces:

(1) $L^{p,q} \cdot L^{p',q'}$ continuously embeds into $L^{r,s}$ for $1/p + 1/p' \leq 1$ where
$$\frac{1}{r} = \frac{1}{p} + \frac{1}{p'} \quad \text{and} \quad \frac{1}{s} = \frac{1}{q} + \frac{1}{q'},$$
with
$$\|fg\|_{L^{r,s}} \leq C\|f\|_{L^{p,q}}\|g\|_{L^{p',q'}}.$$

(2) For $f \in L^2$ and $g \in W^{1,2}$,
$$\|fg\|_{L^{4/3,1}} \leq C\|f\|_{L^2}\|g\|_{W^{1,2}}.$$

(3) $W^{1,1}(\mathbb{R}^4) \hookrightarrow L^{4/3,1}(\mathbb{R}^4)$ and $W^{1,2}(\mathbb{R}^4) \hookrightarrow L^{4,2}(\mathbb{R}^4)$ are continuous embeddings.

(4) $L^{2,1}$ and $L^{2,\infty}$ are dual spaces, as are $L^{4,\infty}$, $L^{4/3,1}$ and $L^{4,1}$, $L^{4/3,\infty}$.

(5) For all $0 < p, r < \infty$ and $0 < q \leq \infty$ (see [Grafakos 2008], Section 1.4.2),
$$\|f^r\|_{L^{p,q}} = \|f\|^r_{L^{pr,qr}}.$$

(6) Let $f \in L^{p,q}(\mathbb{R}^4)$ and $g \in L^{p',q'}(\mathbb{R}^4)$ with $1/p + 1/p' > 1$. Then $h = f * g \in L^{r,s}(\mathbb{R}^4)$ where $1/r = 1/p + 1/p' - 1$ and s is a number such that $1/q + 1/q' \geq 1/s$. Moreover,
$$\|h\|_{L^{r,s}(\mathbb{R}^4)} \leq c\|f\|_{L^{p,q}(\mathbb{R}^4)}\|g\|_{L^{p',q'}(\mathbb{R}^4)}.$$

For a proof, see [Ziemer 1989].

Let G be the distribution such that $\Delta^2 G = \delta_0$. Then, $D^2G \in L^{2,\infty}(\mathbb{R}^4)$ and $D^3G \in L^{4/3,\infty}(\mathbb{R}^4)$. Moreover, $DG \in L^{4,\infty}(\mathbb{R}^4)$.

Using (6), and considering D^2G, D^3G as operators by convolution, we have:

(7) $D^2G : L^{4/3,1}(\mathbb{R}^4) \to L^{4,1}(\mathbb{R}^4)$ and $D^3G : L^{4/3,1}(\mathbb{R}^4) \to L^{2,1}(\mathbb{R}^4)$ are bounded operators.

A.3. ***Scaling and estimates for L* log *L and*** $\mathcal{H}^1$**.** We first prove an essential but technical lemma that is probably well known, though we have not found a reference in the literature. (We prove the lemma for our particular setting, though a more general result is true.)

Lemma A.1. *Let $f = f_1 + f_2$, where $f_1 \in L^{4/3}(B_R)$ and $f_2 \in L\log L(B_R)$, be a compactly supported function with* $\operatorname{spt}(f) \subset B_R$ *and* $\int_{\mathbb{R}^4} f(x)\,dx = 0$. *Then $f \in \mathcal{H}^1(B_R)$ and there exists $C > 0$ such that*

$$\|f\|_{\mathcal{H}^1(B_R)} \leq C(R\|f_1\|_{L^{4/3}(B_R)} + \|f_2\|_{L\log L(B_R)} + |B_{2R}|). \tag{A-2}$$

Proof. First note that

$$\|f_*\|_{L^1} = \int_{B_{2R}} f_*(x)\,dx + \int_{\mathbb{R}^4\setminus B_{2R}} f_*(x)\,dx. \tag{A-3}$$

Since $f_1 \in L^{4/3}(\mathbb{R}^4)$ and $f_2 \in L\log L(\mathbb{R}^4)$, we see that $f \in L^1_{\text{loc}}(\mathbb{R}^4)$ and therefore $f_*(x) \le cMf(x)$ for every $x \in \mathbb{R}^4$. Here $Mf : \mathbb{R}^4 \to \mathbb{R}$ is the maximal function defined by

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\,dy.$$

Using the above, Hölder's inequality and the estimates $\|Mf_1\|_{L^{4/3}} \le c\|f_1\|_{L^{4/3}}$ and $\|Mf_2\|_{L^1(B_{2R})} \le c\|f_2\|_{L\log L(B_{2R})} + c|B_{2R}|$,

$$\begin{aligned}\int_{B_{2R}} f_*(x)\,dx \le& cR\|(f_1)_*\|_{L^{4/3}} + \|(f_2)_*\|_{L^1}\\ \le& cR\|Mf_1\|_{L^{4/3}} + c\|Mf_2\|_{L^1}\\ \le& cR\|f_1\|_{L^{4/3}} + c\|f_2\|_{L\log L} + c|B_{2R}|.\end{aligned} \tag{A-4}$$

Now we calculate for $\phi \in T$ and $x \in \mathbb{R}^4$:

$$\begin{aligned}|\phi_t \star f(x)| &= \left|\int_{B_R} \phi_t(x-y) f(y)\,dy\right|\\ &= \left|\int_{B_R} (\phi_t(x-y) - \phi_t(x)) f(y)\,dy\right|\\ &\le \|\nabla\phi_t\|_{L^\infty} \int_{B_R} |y||f(y)|\,dy,\end{aligned}$$

where we used the mean value theorem and the cancellation property $\int_{\mathbb{R}^4} f(y)\,dy = 0$. Since $\|\nabla\phi_t\|_{L^\infty} \le 1/t^5$, for $t > 0$, we estimate

$$\begin{aligned}|\phi_t \star f(x)| &\le \frac{R}{t^5} \int_{B_R} |f(y)|\,dy\\ &\le \frac{cR^2}{t^5}\|f_1\|_{L^{4/3}} + \frac{cR}{t^5}\|f_2\|_{L\log L}.\end{aligned} \tag{A-5}$$

Assuming now that $|x| \ge 2R$, we can apply a technical result to get

$$f_*(x) = \sup_{\phi\in T}\ \sup_{t>|x|/2} |\phi_t \star f(x)| \le \frac{cR^2}{|x|^5}\|f_1\|_{L^{4/3}} + \frac{cR}{|x|^5}\|f_2\|_{L\log L}. \tag{A-6}$$

Inserting (A-4) and (A-6) into (A-3), we conclude that

$$
\begin{aligned}
\|f_*\|_{L^1} &\le cR\|f_1\|_{L^{4/3}} + c\|f_2\|_{L\log L} + c|B_{2R}| \\
&\quad + (cR^2\|f_1\|_{L^{4/3}} + cR\|f_2\|_{L\log L}) \int_{\mathbb{R}^4\setminus B_{2R}} \frac{1}{|x|^5}\,dx \\
&\le cR\|f_1\|_{L^{4/3}} + c\|f_2\|_{L\log L} + c|B_{2R}|.
\end{aligned}
\tag{A-7}
$$

This concludes the proof. □

We also note two important inequalities (with proofs following those of [Sharp and Topping 2013]):

Lemma A.2. *Let $f \in L\log L(B_r(x_0))$ for $r \in (0, 1/2]$. There exists $C > 0$ such that*

$$
\|f\|_{L^1(B_r(x_0))} \le C(\log(1/r))^{-1}\|f\|_{L\log L(B_r(x_0))}. \tag{A-8}
$$

Proof. Start by observing that

$$
\begin{aligned}
0 &\le r^4 \int_0^{|B_1|} f^*(r^4 t)\log\left(2+\frac{1}{t}\right) dt \\
&= \int_0^{|B_r(x_0)|} f^*(s)\log\left(2+\frac{r^4}{s}\right) ds \\
&= \int_0^{|B_r(x_0)|} f^*(s)\log(r^4)\,ds + \int_0^{|B_r(x_0)|} f^*(s)\log\left(\frac{2}{r^4}+\frac{1}{s}\right) ds \\
&\le -4\log(1/r)\|f\|_{L^1(B_r(x_0))} + C\|f\|_{L\log L(B_r(x_0))}.
\end{aligned}
$$

The last inequality follows from the fact that there exists a fixed C such that

$$
\frac{2}{r^4} + \frac{1}{s} \le \frac{2\omega_4 + 1}{s} \le \left(2 + \frac{1}{s}\right)^C
$$

for all $s \le \omega_4 r^4$. □

Let u be an f-approximate biharmonic map on B_1 with $f \in L\log L(B_1)$. For $x_0 \in B_1$ and $R > 0$ such that $B_R(x_0) \subset B_1$, define $\hat{u}(x) := u(x_0 + Rx)$ and $\hat{f}(x) := R^4 f(x_0 + Rx)$. Then $\hat{u}$ is an $\hat{f}$-approximate biharmonic map. Moreover, we note that for any $r \in (0,1)$, $p \ge 1$, and $k = 1, 2, 3$:

(1) $\|D^k\hat{u}\|_{L^{4/k}(B_r)} = \|D^k u\|_{L^{4/k}(B_{rR}(x_0))}$.

(2) $\|\hat{f}\|_{L^p(B_r)} = R^{4(1-1/p)}\|f\|_{L^p(B_{rR}(x_0))}$.

Lemma A.3. *Let $f \in L\log L(B_r(x_0))$, where $r \in (0, 1/2]$ and define $\hat{f}(x) := r^4 f(x_0 + rx)$. Then there exists $C > 0$ such that*

$$
\|\hat{f}\|_{L\log L(B_1)} \le C\|f\|_{L\log L(B_r(x_0))}.
$$

Proof. First note that, using the definition of $\hat{f}$, one can immediately show that $\hat{f}^*(t) = r^4 f^*(r^4 t)$. Thus,

$$\begin{aligned}\int_0^{|B_1|} \hat{f}^*(t)\log\left(2+\frac{1}{t}\right)dt &= \int_0^{|B_1|} r^4 f^*(r^4 t)\log\left(2+\frac{1}{t}\right)dt\\ &= \int_0^{|B_r(x_0)|} f^*(s)\log\left(2+\frac{r^4}{s}\right)ds\\ &\le \int_0^{|B_r(x_0)|} f^*(s)\log\left(2+\frac{1}{s}\right)ds. \qquad \square\end{aligned}$$

References

[Chen 1989] Y. M. Chen, "The weak solutions to the evolution problems of harmonic maps", *Math. Z.* **201**:1 (1989), 69–74. MR 90i:58030 Zbl 0685.58015

[Chen and Tian 1999] J. Chen and G. Tian, "Compactification of moduli space of harmonic mappings", *Comment. Math. Helv.* **74**:2 (1999), 201–237. MR 2001k:58024 Zbl 0958.53047

[Ding and Tian 1995] W. Ding and G. Tian, "Energy identity for a class of approximate harmonic maps from surfaces", *Comm. Anal. Geom.* **3**:3-4 (1995), 543–554. MR 97e:58055 Zbl 0855.58016

[Evans 1990] L. C. Evans, *Weak convergence methods for nonlinear partial differential equations*, CBMS Regional Conference Series in Mathematics **74**, Amer. Math. Soc., Providence, RI, 1990. MR 91a:35009 Zbl 0698.35004

[Evans and Gariepy 1992] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992. MR 93f:28001 Zbl 0804.28001

[Grafakos 2008] L. Grafakos, *Classical Fourier analysis*, 2nd ed., Graduate Texts in Mathematics **249**, Springer, New York, 2008. MR 2011c:42001 Zbl 1220.42001

[Hélein 1990] F. Hélein, "Régularité des applications faiblement harmoniques entre une surface et une sphère", *C. R. Acad. Sci. Paris Sér. I Math.* **311**:9 (1990), 519–524. MR 92a:58034 Zbl 0728.35014

[Hornung and Moser 2012] P. Hornung and R. Moser, "Energy identity for intrinsically biharmonic maps in four dimensions", *Anal. PDE* **5**:1 (2012), 61–80. MR 2957551 Zbl 1273.58007

[Huilgol 1971] R. R. Huilgol, "On Liouville's theorem for biharmonic functions", *SIAM J. Appl. Math.* **20** (1971), 37–39. MR 43 #552 Zbl 0217.10502

[Jost 1991] J. Jost, *Two-dimensional geometric variational problems*, Wiley, Chichester, 1991. MR 92h:58045 Zbl 0729.49001

[Lamm and Rivière 2008] T. Lamm and T. Rivière, "Conservation laws for fourth order systems in four dimensions", *Comm. Partial Differential Equations* **33**:1-3 (2008), 245–262. MR 2009h:35095 Zbl 1139.35328

[Laurain and Rivière 2013] P. Laurain and T. Rivière, "Energy quantization for biharmonic maps", *Adv. Calc. Var.* **6**:2 (2013), 191–216. MR 3043576 Zbl 1275.35098

[Li and Zhu 2011] J. Li and X. Zhu, "Small energy compactness for approximate harmomic mappings", *Commun. Contemp. Math.* **13**:5 (2011), 741–763. MR 2847227 Zbl 1245.58008

[Lin and Wang 1998] F. Lin and C. Wang, "Energy identity of harmonic map flows from surfaces at finite singular time", *Calc. Var. Partial Differential Equations* **6**:4 (1998), 369–380. MR 99k:58047 Zbl 0908.58008

[Liu and Yin 2013] L. Liu and H. Yin, "Neck analysis for biharmonic maps", preprint, 2013. arXiv 1312.4600v1

[Parker 1996] T. H. Parker, "Bubble tree convergence for harmonic maps", *J. Differential Geom.* **44**:3 (1996), 595–633. MR 98k:58069 Zbl 0874.58012

[Qing and Tian 1997] J. Qing and G. Tian, "Bubbling of the heat flows for harmonic maps from surfaces", *Comm. Pure Appl. Math.* **50**:4 (1997), 295–310. MR 98k:58070 Zbl 0879.58017

[Sacks and Uhlenbeck 1981] J. Sacks and K. Uhlenbeck, "The existence of minimal immersions of 2-spheres", *Ann. of Math.* (2) **113**:1 (1981), 1–24. MR 82f:58035 Zbl 0462.58014

[Sharp and Topping 2013] B. Sharp and P. Topping, "Decay estimates for Rivière's equation, with applications to regularity and compactness", *Trans. Amer. Math. Soc.* **365**:5 (2013), 2317–2339. MR 3020100 Zbl 1270.35152

[Shatah 1988] J. Shatah, "Weak solutions and development of singularities of the SU(2) σ-model", *Comm. Pure Appl. Math.* **41**:4 (1988), 459–469. MR 89f:58044 Zbl 0686.35081

[Wang 1996] C. Wang, "Bubble phenomena of certain Palais–Smale sequences from surfaces to general targets", *Houston J. Math.* **22**:3 (1996), 559–590. MR 98h:58053 Zbl 0879.58019

[Wang 2004a] C. Wang, "Remarks on biharmonic maps into spheres", *Calc. Var. Partial Differential Equations* **21**:3 (2004), 221–242. MR 2005e:58026 Zbl 1060.58011

[Wang 2004b] C. Wang, "Stationary biharmonic maps from $\mathbb{R}^m$ into a Riemannian manifold", *Comm. Pure Appl. Math.* **57**:4 (2004), 419–444. MR 2005e:58027 Zbl 1055.58008

[Wang and Zheng 2012] C. Wang and S. Zheng, "Energy identity of approximate biharmonic maps to Riemannian manifolds and its application", *J. Funct. Anal.* **263**:4 (2012), 960–987. MR 2927401 Zbl 1257.58010

[Wente 1969] H. C. Wente, "An existence theorem for surfaces of constant mean curvature", *J. Math. Anal. Appl.* **26** (1969), 318–344. MR 39 #4788 Zbl 0181.11501

[Zhu 2012] X. Zhu, "No neck for approximate harmonic maps to the sphere", *Nonlinear Anal.* **75**:11 (2012), 4339–4345. MR 2921993 Zbl 1243.58011

[Ziemer 1989] W. P. Ziemer, *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*, Graduate Texts in Mathematics **120**, Springer, New York, 1989. Sobolev spaces and functions of bounded variation. MR 91e:46046 Zbl 0692.46022

Received December 30, 2013. Revised July 4, 2014.

CHRISTINE BREINER
DEPARTMENT OF MATHEMATICS
FORDHAM UNIVERSITY
BRONX, NY 10458
UNITED STATES
cbreiner@fordham.edu

TOBIAS LAMM
INSTITUTE FOR ANALYSIS
KARLSRUHE INSTITUTE OF TECHNOLOGY
KAISERSTRASSE 89-93
D-76133 KARLSRUHE
GERMANY
tobias.lamm@kit.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.93

CRITERIA FOR VANISHING OF TOR OVER COMPLETE INTERSECTIONS

OLGUR CELIKBAS, SRIKANTH B. IYENGAR, GREG PIEPMEYER AND ROGER WIEGAND

We exploit properties of Dao's η-pairing (see *Trans. Amer. Math. Soc.* 365:6 (2013), 2803–2821), as well as techniques of Huneke, Jorgensen, and Wiegand (*J. Algebra* 238:2 (2001), 684–702), to study the vanishing of $\mathrm{Tor}_i(M, N)$ for finitely generated modules M, N over complete intersections. We prove vanishing of $\mathrm{Tor}_i(M, N)$ for all $i \geq 1$ under depth conditions on M, N, and $M \otimes N$. Our arguments improve a result of Dao and establish a new connection between the vanishing of Tor and the depth of tensor products.

1. Introduction

In a seminal paper, Auslander [1961] proved that if R is a local ring and M and N are nonzero finitely generated R-modules such that $\mathrm{pd}(M) < \infty$ and $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq 1$, then

$$\text{(1.0.1)} \qquad \operatorname{depth}(M) + \operatorname{depth}(N) = \operatorname{depth}(R) + \operatorname{depth}(M \otimes_R N),$$

that is, the *depth formula* holds. Huneke and Wiegand [1994, Theorem 2.5] established the depth formula for Tor-independent modules (not necessarily of finite projective dimension) over complete intersection rings. Christensen and Jorgensen [2015] extended that result to AB rings [Huneke and Jorgensen 2003], a class of Gorenstein rings strictly containing the class of complete intersections. The depth formula is important for the study of depths of tensor products of modules [Auslander 1961; Huneke and Wiegand 1994], as well as of complexes [Foxby 1980; Iyengar 1999]. We seek conditions on the modules M, N and $M \otimes_R N$ forcing such a formula to hold, in particular, conditions implying $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq 1$. The following conjecture — implicit in the work of Huneke, Jorgensen, and Wiegand — guides our search.

Iyengar was partly supported by NSF grant DMS-1201889 and a Simons Fellowship; Wiegand was partly supported by a Simons Collaboration Grant.
MSC2010: 13C40, 13D07.
Keywords: complete intersection, tensor product, torsion, vanishing of Tor.

Conjecture 1.1 [Huneke et al. 2001]. Let M, N be finitely generated modules over a complete intersection R of codimension c. If $M \otimes_R N$ is a $(c+1)$-st syzygy and M has rank, then $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$.

The conjecture is true if $c = 0$ or $c = 1$, by [Lichtenbaum 1966, Corollary 1] and [Huneke and Wiegand 1994, Theorem 2.7], respectively. Without the assumption of rank, there are easy counterexamples, e.g., $R = k[\![x, y]\!]/(xy)$ and $M = N = R/(x)$; M is an n-th syzygy for all n, but the odd index Tor modules are nonzero.

A finitely generated module over a complete intersection is an n-th syzygy of some finitely generated module if and only if it satisfies *Serre's condition* (S_n); see §2.6. Our methods yield a sharpening of the following theorem due to Dao:

Theorem 1.2 [Dao 2007]. *Let R be a complete intersection in an unramified regular local ring, of relative codimension c, and let M, N be finitely generated R-modules. Assume*

(i) *M and N satisfy* (S_c),

(ii) *$M \otimes_R N$ satisfies* (S_{c+1}), *and*

(iii) *$M_\mathfrak{p}$ is a free $R_\mathfrak{p}$-module for all prime ideals $\mathfrak{p}$ of height at most c.*

Then $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \geq 1$* (*and hence the depth formula holds*).

By analyzing Serre's conditions, we remove Dao's assumption that the ambient regular local ring be unramified; see Corollary 3.14. Even though complete intersections in unramified regular local rings suffice for many applications, our conclusion is of interest: Dao's proof uses the nonnegativity of partial Euler characteristics, but nonnegativity remains unknown for the ramified case; see [Dao 2007, Theorem 6.3 and the proof of Lemma 7.7].

If the ambient regular local ring *is* unramified, we can replace c with $c - 1$ in both hypotheses (i) and (ii), remove hypothesis (iii), and still conclude that $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$ *provided* that $\eta_c^R(M, N) = 0$; see § 3.1 for the definition of $\eta_c^R(-, -)$ and Theorem 3.10 for our result.

Moore, Piepmeyer, and Spiroff [Moore et al. 2013] and Walker [2014] have proved vanishing of the η-pairing in several important cases. These, in turn, yield results on vanishing of Tor. See Proposition 4.1, Theorem 4.2, and Corollary 4.3.

Our proofs rely on a reduction technique using *quasiliftings*; see §2.8. Quasiliftings were initially defined and studied in [Huneke et al. 2001]. The key ingredient for our argument is Lemma 3.9. It shows that if $R = S/(f)$ and S is a complete intersection of codimension $c - 1$, and if $\eta_c^R(M, N) = 0$, then $\eta_{c-1}^S(E, F) = 0$, where E and F are quasiliftings of M and N to S, respectively. By induction, we get that $\operatorname{Tor}_i^S(E, F) = 0$ for all $i \geq 1$. This allows us to prove the vanishing of $\operatorname{Tor}_i^R(M, N)$ from the depth and syzygy relations between the pairs E, F and M, N.

In the Appendices we revisit [Huneke and Wiegand 1994] and use our work to obtain one of the main results there. Moreover, we point out an oversight in [Miller 1998] and state the author's result in its corrected form as Corollary B.3.

2. Preliminaries

We review a few concepts and results, especially universal pushforwards and quasi-liftings [Huneke et al. 2001; Huneke and Wiegand 1994]. Throughout R will be a commutative noetherian ring.

Let $\nu_R(M)$ denote the minimal number of generators of the R-module M. If $(R, \mathfrak{m})$ is local, then the *codimension* of R is $\operatorname{codim}(R) := \nu_R(\mathfrak{m}) - \dim(R)$, a nonnegative integer. We have $\operatorname{codim}(\hat{R}) = \operatorname{codim}(R)$, where $\hat{R}$ is the $\mathfrak{m}$-adic completion of R.

2.1. Complete intersections. R is a *complete intersection in a local ring* $(Q, \mathfrak{n})$ if there a surjection $\pi\colon Q \twoheadrightarrow R$ with $\ker(\pi)$ generated by a Q-regular sequence in $\mathfrak{n}$; the length of this regular sequence is the *relative codimension of R in Q*. A *hypersurface in Q* is a complete intersection of relative codimension one in Q.

Assume $\hat{R}$ is a complete intersection in a regular local ring $(Q, \mathfrak{n})$, of relative codimension c. Then $\hat{R} = Q/(\underline{f})$ for a regular sequence $\underline{f} = f_1, \dots, f_c$, where $\operatorname{codim}(R) \le c$. Moreover, the codimension of R is c if and only if $(\underline{f}) \subseteq \mathfrak{n}^2$.

A ring is a *complete intersection* (resp., *hypersurface*) if it is local and its completion is a complete intersection (resp., hypersurface) in a regular local ring.

2.2. Ramified regular local rings. A regular local ring $(Q, \mathfrak{n}, k)$ is said to be *unramified* if either (i) Q is equicharacteristic, i.e., contains a field, or else (ii) $Q \supset \mathbb{Z}$, $\operatorname{char}(k) = p$, and $p \notin \mathfrak{n}^2$. In contrast, the regular local ring $R = V[x]/(x^2 - p)$, where V is the ring of p-adic integers, is *ramified*. Every localization, at a prime ideal, of an unramified regular local ring is again unramified; see [Auslander 1961, Lemma 3.4].

Let $(Q, \mathfrak{n}, k)$ be a d-dimensional complete regular local ring. If Q is ramified, then k has characteristic p. Further, there is a complete unramified discrete valuation ring (V, pV) such that $Q \cong T/(p - f)$, where $T = V[\![x_1, \dots, x_d]\!]$ and f is contained in the square of the maximal ideal of T; see for example [Bourbaki 2006, Chaper IX, §3]. Hence every complete regular local ring is a hypersurface in an unramified one. Consequently, when R is a complete intersection, $\hat{R}$ is a complete intersection in an unramified regular local ring Q such that

$$\operatorname{codim} R \le c \le \operatorname{codim} R + 1,$$

where c is the relative codimension of $\hat{R}$ in Q.

2.3. The depth formula [Huneke and Wiegand 1994, Theorem 2.5]. Let R be a complete intersection and let M, N be finitely generated R-modules. If

$\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq 1$, then the *depth formula* (1.0.1) holds, that is,

$$\operatorname{depth}(M) + \operatorname{depth}(N) = \operatorname{depth}(R) + \operatorname{depth}(M \otimes_R N).$$

Recall that $\operatorname{depth}(0) = \infty$, so the formula holds trivially if a zero module appears.

2.4. Torsion submodule. The *torsion submodule* $\top_R M$ of M is the kernel of the natural homomorphism $M \to \mathrm{Q}(R) \otimes_R M$, where $\mathrm{Q}(R) = \{\text{non-zerodivisors}\}^{-1} R$ is the total quotient ring of R. The module M is *torsion* if $\top_R M = M$, and *torsion-free* if $\top_R M = 0$. To restate, M is torsion-free if and only if every non-zerodivisor of R is a non-zerodivisor on M, that is, if and only if $\bigcup \operatorname{Ass} M \subseteq \bigcup \operatorname{Ass} R$. Similarly, M is torsion if and only if $M_{\mathfrak{p}} = 0$ for all $\mathfrak{p} \in \operatorname{Ass}(R)$. For notation, the inclusion $\top_R M \subseteq M$ has cokernel $\perp_R M$:

$$0 \longrightarrow \top_R M \longrightarrow M \longrightarrow \perp_R M \longrightarrow 0. \tag{2.4.1}$$

2.5. Torsionless and reflexive modules. Let M be a finitely generated R-module; M^* denotes its dual $\operatorname{Hom}_R(M, R)$. The module M is *torsionless* if it embeds in a free module, equivalently, the canonical map $M \to M^{**}$ is injective. Torsionless modules are torsion-free, and the converse holds if $R_{\mathfrak{p}}$ is Gorenstein for every associated prime $\mathfrak{p}$ of R; see [Vasconcelos 1968, Theorem A.1]. The module M is *reflexive* provided the map $M \to M^{**}$ is an isomorphism.

2.6. Serre's conditions (see [Leuschke and Wiegand 2012, Appendix A, §1] and [Evans and Griffith 1985, Theorem 3.8]). Let M be a finitely generated R-module and let n be a nonnegative integer. Then M is said to satisfy *Serre's condition* (S_n) provided that

$$\operatorname{depth}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) \geq \min\{n, \operatorname{height}(\mathfrak{p})\} \quad \text{for all } \mathfrak{p} \in \operatorname{Supp}(M).$$

A finitely generated module M over a local ring R is *maximal Cohen–Macaulay* if $\operatorname{depth}(M) = \dim(R)$; necessary for this equality is that $M \neq 0$.

If M satisfies (S_1), then M is torsion-free, and the converse holds if R has no embedded primes, e.g., is reduced or Cohen–Macaulay; see §2.4. If R is Gorenstein, then M satisfies (S_2) if and only if M is reflexive; see §2.5 and [Evans and Griffith 1985, Theorem 3.6]. Moreover, if R is Gorenstein, then M satisfies (S_n) if and only if M is an n-th syzygy module; see [Leuschke and Wiegand 2012, Corollary A.12].

A localization of a torsion-free module need not be torsion-free; see, for example, [Epstein and Yao 2012, Example 3.9]. However, over Cohen–Macaulay rings, we have the following.

Remark 2.7. Assume that R is Cohen–Macaulay and M is a finitely generated R-module. Let $\mathfrak{p}$ be a prime ideal of R. Note that, since $\top_R M$ is killed by a non-zerodivisor of R, $(\top_R M)_{\mathfrak{p}}$ is a torsion $R_{\mathfrak{p}}$-module. Next, $\perp_R M$ satisfies (S_1) as R is Cohen–Macaulay, and so $(\perp_R M)_{\mathfrak{p}}$ is a torsion-free $R_{\mathfrak{p}}$-module; see §2.6.

Localizing the exact sequence (2.4.1) at $\mathfrak{p}$, we see that $(\top_R M)_\mathfrak{p} \cong \top_{R_\mathfrak{p}}(M_\mathfrak{p})$. In particular, if M is a torsion-free R-module, then $M_\mathfrak{p}$ is a torsion-free $R_\mathfrak{p}$-module.

We recall a technique from [Huneke et al. 2001, §1] for lowering the codimension.

2.8. Pushforward and quasilifting [Huneke et al. 2001, §1]. Let R be a Gorenstein local ring and let M be a finitely generated torsion-free R-module. Choose a surjection $\varepsilon\colon R^{(\nu)} \twoheadrightarrow M^*$ with $\nu = \nu_R(M^*)$. Applying $\operatorname{Hom}(-, R)$ to this surjection, we obtain an injection $\varepsilon^*\colon M^{**} \hookrightarrow R^{(\nu)}$. Let M_1 be the cokernel of the composition $M \hookrightarrow M^{**} \hookrightarrow R^{(\nu)}$. The exact sequence

$$0 \to M \to R^{(\nu)} \to M_1 \to 0 \tag{2.8.1}$$

is called a *pushforward* of M. The extension (2.8.1) and the module M_1 are unique up to noncanonical isomorphism; see [Celikbas 2011, pp. 174–175]. We refer to such a module M_1 as the pushforward of M. Note $M_1 = 0$ if and only if M is free.

Assume $R = S/(f)$ where $(S, \mathfrak{n})$ is a local ring and f is a non-zerodivisor in $\mathfrak{n}$. Let $S^{(\nu)} \twoheadrightarrow M_1$ be the composition of the canonical map $S^{(\nu)} \twoheadrightarrow R^{(\nu)}$ and the map $R^{(\nu)} \twoheadrightarrow M_1$ in (2.8.1). The *quasilifting* of M to S is the module E in the exact sequence of S-modules:

$$0 \to E \to S^{(\nu)} \to M_1 \to 0\,. \tag{2.8.2}$$

The quasilifting of M is unique up to isomorphism of S-modules.

Proposition 2.9 is from [Huneke et al. 2001, Propositions 1.6 and 1.7]; while Proposition 2.10 is embedded in the proofs of [Huneke et al. 2001, Propositions 1.8 and 2.4] and is recorded explicitly in [Celikbas 2011, Proposition 3.2(3)(b)]. We will use Proposition 2.10 in the proofs of Theorem 3.10 and Theorem B.2 below.

Proposition 2.9 [Huneke et al. 2001]. *Let R be a Gorenstein local ring and let M be a finitely generated torsion-free R-module. Let M_1 denote the pushforward of M.*

(i) *Let $n \geq 0$. Then M satisfies* (S_{n+1}) *if and only if M_1 satisfies* (S_n).

(ii) *Let $\mathfrak{p}$ be a prime ideal. If $M_\mathfrak{p}$ is a maximal Cohen–Macaulay $R_\mathfrak{p}$-module, then $(M_1)_\mathfrak{p}$ is either zero or a maximal Cohen–Macaulay $R_\mathfrak{p}$-module.*

Proposition 2.10 [Huneke et al. 2001]. *Let $R = S/(f)$ where S is a complete intersection and f is a non-zerodivisor in S. Let N be a finitely generated torsion-free R-module such that $M \otimes_R N$ is reflexive. Assume* $\operatorname{Tor}_i^R(M, N)_\mathfrak{p} = 0$ *for all $i \geq 1$ and for all primes $\mathfrak{p}$ of R with* $\operatorname{height}(\mathfrak{p}) \leq 1$.

(i) *Then $M_1 \otimes_R N$ is torsion-free.*

(ii) *Let E and F denote the quasiliftings of M and N to S, respectively; see §2.8. Assume* $\operatorname{Tor}_i^S(E, F) = 0$ *for all $i \geq 1$. Then* $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \geq 1$.*

Serre's conditions (S_n) need not ascend along flat local homomorphisms. This can be problematic:

Example 2.11. The ring $\mathbb{C}[\![x, y, u, v]\!]/(x^2, xy)$ has depth two and therefore, by Heitmann's theorem [1993, Theorem 8], it is the completion $\hat{R}$ of a unique factorization domain $(R, \mathfrak{m})$. Then R, being normal, satisfies (S_2), but $\hat{R}$ does not even satisfy (S_1), since the localization at the height-one prime ideal (x, y) has depth zero.

For flat local homomorphisms between Cohen–Macaulay rings, and more generally when the fibers are Cohen–Macaulay, however, (S_n) *does* ascend and descend:

Lemma 2.12. *Let R be a local ring, $\mathfrak{p}$ a prime ideal of R, and let M be a finitely generated R-module.*

(1) *If M is reflexive, then so is the $R_\mathfrak{p}$-module $M_\mathfrak{p}$.*
(2) *Suppose R is Cohen–Macaulay. Then $(\top_R M)_\mathfrak{p} = \top_{R_\mathfrak{p}} M_\mathfrak{p}$; in particular, if M is torsion-free, then so is $M_\mathfrak{p}$.*
(3) *Suppose $R \to S$ is a flat local homomorphism. If $S \otimes_R M$ satisfies (S_n) as an S-module, then M satisfies (S_n) as an R-module; the converse holds when the fibers of the map $R \to S$ are Cohen–Macaulay.*

Proof. For part (1), localize the isomorphism $M \to M^{**}$. Part (2) is Remark 2.7. Part (3) can be proved along the same lines as [Matsumura 1989, Theorem 23.9]: For any $\mathfrak{q}$ in Spec S with $\mathfrak{p} = \mathfrak{q} \cap R$, it follows from [Matsumura 1989, Theorems 15.1 and 23.3] that

$$\operatorname{height}(\mathfrak{q}) = \operatorname{height}(\mathfrak{p}) + \dim(S_\mathfrak{q}/\mathfrak{p}S_\mathfrak{q}),$$
$$\operatorname{depth}_{S_\mathfrak{q}}(S \otimes_R M)_\mathfrak{q} = \operatorname{depth}_{R_\mathfrak{p}}(M_\mathfrak{p}) + \operatorname{depth}(S_\mathfrak{q}/\mathfrak{p}S_\mathfrak{q}).$$

When $S \otimes_R M$ satisfies (S_n), for $\mathfrak{q}$ minimal in $S/\mathfrak{p}S$, these equalities give

$$\operatorname{depth}_{R_\mathfrak{p}}(M_\mathfrak{p}) = \operatorname{depth}_{S_\mathfrak{q}}(S \otimes_R M)_\mathfrak{q} \geq \min\{n, \operatorname{height}(\mathfrak{q})\} = \min\{n, \operatorname{height}(\mathfrak{p})\}.$$

Thus M satisfies (S_n). Conversely, if $S_\mathfrak{q}/\mathfrak{p}S_\mathfrak{q}$ is Cohen–Macaulay and the R-module M satisfies (S_n), one gets

$$\operatorname{depth}_{S_\mathfrak{q}}(S \otimes_R M)_\mathfrak{q} \geq \min\{n, \operatorname{height}(\mathfrak{p})\} + \dim(S_\mathfrak{q}/\mathfrak{p}S_\mathfrak{q}) \geq \min\{n, \operatorname{height}(\mathfrak{q})\}.$$

This completes the proof of part (3). □

3. Main theorem

Our main result, Theorem 3.10, is here. We use the θ- and η-pairings introduced by Hochster [1981] and Dao [2007]. After preliminaries on these, we focus on complete intersections; see §2.1, the setting of our applications.

3.1. The θ- and η-pairings [Hochster 1981; Dao 2013a; Dao 2007]. Let R be a local ring and let M and N be finitely generated R-modules. Assume that there exists an integer f (depending on M and N), such that $\operatorname{Tor}_i^R(M, N)$ has finite length for all $i \geq f$.

If R is a hypersurface, then $\operatorname{Tor}_i^R(M, N) \cong \operatorname{Tor}_{i+2}^R(M, N)$ for all $i \gg 0$; see [Eisenbud 1980]. Hochster [1981] introduced the θ pairing for $n \gg 0$ by

$$\theta^R(M, N) = \operatorname{length}(\operatorname{Tor}_{2n}^R(M, N)) - \operatorname{length}(\operatorname{Tor}_{2n-1}^R(M, N))$$

When R is any complete intersection, Dao [2007, Definition 4.2.] made the definition

$$\eta_e^R(M, N) = \lim_{n\to\infty} \frac{1}{n^e} \sum_{i=f}^{n} (-1)^i \operatorname{length}(\operatorname{Tor}_i^R(M, N)).$$

The η-pairing is a natural extension to complete intersections of the θ-pairing. Moreover the following statements hold; see [Dao 2007, Theorem 4.3].

(i) $\eta_e^R(M, -)$ and $\eta_e^R(-, N)$ are additive on short exact sequences, provided η_e^R is defined on the pairs of modules involved.

(ii) If R is a hypersurface, then $\eta_1^R(M, N) = \frac{1}{2}\theta^R(M, N)$. Hence $\eta_1^R(M, N) = 0$ if and only if $\theta^R(M, N) = 0$.

Assume R is a complete intersection.

(iii) $\eta_e^R(M, N) = 0$ if $e \geq \operatorname{codim} R$ and either M or N has finite length.

(iv) η_e^R is finite when $e = \operatorname{codim}(R)$, and η_e^R is zero when $e > \operatorname{codim} R$.

The next result [Dao 2007, Theorem 6.3], on *Tor-rigidity*, shows the utility of the η-pairing.

Theorem 3.2 [Dao 2007]. *Let R be a local ring whose completion is a complete intersection, of relative codimension $c \geq 1$, in an* unramified *regular local ring. Let M, N be finitely generated R-modules. Assume* $\operatorname{Tor}_i^R(M, N)$ *has finite length for all $i \gg 0$, and that $\eta_c^R(M, N) = 0$. Then the pair M, N is c-Tor-rigid, that is, if $s \geq 0$ and* $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i = s, \ldots, s + c - 1$, then* $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \geq s$.*

The following conjectures have received quite a bit of attention:

Conjectures 3.3. Assume R is a local ring which is an isolated singularity, i.e., $R_{\mathfrak{p}}$ is a regular local ring for all nonmaximal prime ideals $\mathfrak{p}$ of R.

(i) [Dao 2013a, Conjecture 3.15] If R is an equicharacteristic hypersurface of even dimension, then $\eta_1^R(M, N) = 0$ for all finitely generated R-modules M and N.

(ii) [Moore et al. 2013, Conjecture 2.4] If R is a complete intersection of codimension $c \geq 2$, then $\eta_c^R(M, N) = 0$ for all finitely generated R-modules M and N.

Moore, Piepmeyer, Spiroff and Walker [2011] have settled Conjecture 3.3(i) in the affirmative for certain types of affine algebras. Polishchuk and Vaintrob [2012, Remark 4.1.5], as well as Buchweitz and Van Straten [2012, Main Theorem], have since given other proofs, in somewhat different contexts, of this result; see Theorem 4.2 for a recent result of Walker [2014] concerning Conjecture 3.3(ii), and Corollary 4.3 for an application of his result.

Our proofs of Lemma 3.6 and Theorem B.2 use the following (see [Auslander 1961, Lemma 3.1] or [Huneke and Wiegand 1994, Lemma 1.1]).

Remark 3.4. Let R be a local ring, and let M and N be nonzero finitely generated R-modules. Assume $M \otimes_R N$ is torsion-free. Then $M \otimes_R N \cong M \otimes \perp_R N$. Moreover, if $\operatorname{Tor}_1^R(M, \perp_R N) = 0$, then $\top_R N = 0$, and hence N is torsion-free.

We encounter the same hypotheses often enough to warrant a piece of notation.

Notation 3.5. Let c be a positive integer. A pair M, N of finitely generated modules over a ring R satisfies (SP_c) provided the following conditions hold:

(i) M and N satisfy Serre's condition (S_{c-1}).

(ii) $M \otimes_R N$ satisfies (S_c).

(iii) $\operatorname{Tor}_i^R(M, N)$ has finite length for all $i \gg 0$.

Hypersurfaces. We begin with a lemma analogous to [Dao 2008, Proposition 3.1]; however, we do not assume any depth properties on either M or N; see §2.1 and Notation 3.5.

Lemma 3.6. *Let R be a local ring whose completion is a hypersurface in an* unramified *regular local ring*, *and let M, N be finitely generated R-modules. Assume that the following hold*:

(i) $\dim(R) \geq 1$.

(ii) *The pair M, N satisfies* (SP_1).

(iii) $\operatorname{Supp}_R(\top_R N) \subseteq \operatorname{Supp}_R(M)$.

(iv) $\theta^R(M, N) = 0$.

Then $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \geq 1$, and N is torsion-free.*

Proof. Consider the following conditions for a prime ideal $\mathfrak{p}$ of R:

$$(\top_R N)_{\mathfrak{p}} \text{ has finite length over } R_{\mathfrak{p}} \quad \text{and} \quad \dim(R_{\mathfrak{p}}) \geq 1. \tag{3.6.1}$$

<u>Claim:</u> If $\mathfrak{p}$ is as in (3.6.1), then $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (\perp_R N)_{\mathfrak{p}}) = 0$ for all $i \geq 1$.

We may assume that $M_{\mathfrak{p}} \neq 0$. We know from (ii) that $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, N_{\mathfrak{p}})$ has finite length over $R_{\mathfrak{p}}$ for all $i \gg 0$. Since $(\top_R N)_{\mathfrak{p}}$ has finite length, the exact sequence (2.4.1) for N, localized at $\mathfrak{p}$, shows that $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (\perp_R N)_{\mathfrak{p}})$ has finite length over $R_{\mathfrak{p}}$ for all $i \gg 0$.

Using the additivity of $\theta^{R_{\mathfrak{p}}}$ along the same exact sequence, we see that

$$\theta^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (\perp_R N)_{\mathfrak{p}}) = -\theta^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (\top_R N)_{\mathfrak{p}}) = 0, \tag{3.6.2}$$

the last by §3.1.

Since $\perp_R N$ is a torsionless R-module (see §2.5), there exists an exact sequence

$$0 \to \perp_R N \to R^{(n)} \to Z \to 0. \tag{3.6.3}$$

Localizing this sequence at $\mathfrak{p}$, we see that, for $i \gg 0$, $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, Z_{\mathfrak{p}})$ has finite length and hence (since $\dim(R_{\mathfrak{p}}) \geq 1$) is torsion. Now Corollary A.2 forces $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, Z_{\mathfrak{p}})$ to be torsion for all $i \geq 1$.

From (3.6.3), we see that $\operatorname{Tor}_1^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, Z_{\mathfrak{p}})$ embeds into $M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} (\perp_R N)_{\mathfrak{p}}$. But $\operatorname{Tor}_1^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, Z_{\mathfrak{p}})$ is torsion, and (by Remarks 2.7 and Remark 3.4) $M_{\mathfrak{p}} \otimes_{R_{\mathfrak{p}}} (\perp_R N)_{\mathfrak{p}}$ is torsion-free; therefore $\operatorname{Tor}_1^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, Z_{\mathfrak{p}}) = 0$.

Next we note that $\theta^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, Z_{\mathfrak{p}}) = -\theta^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (\perp_R N)_{\mathfrak{p}}) = 0$; see (3.6.3) and (3.6.2). This implies, by Theorem 3.2, that $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, Z_{\mathfrak{p}}) = 0$ for all $i \geq 1$; see §3.1. The claim now follows from (3.6.3).

If $\top_R N \neq 0$, then there is a prime $\mathfrak{p}$, minimal in $\operatorname{Supp}_R(\top_R N)$, and so $(\top_R N)_{\mathfrak{p}}$ is a nonzero module of finite length. Moreover $\dim(R_{\mathfrak{p}}) \geq 1$: otherwise $\mathfrak{p} \in \operatorname{Ass}(R)$ and hence $(\top_R N)_{\mathfrak{p}} = 0$; see §2.4. Thus $\mathfrak{p}$ satisfies (3.6.1) and, by our claim, $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, (\perp_R N)_{\mathfrak{p}}) = 0$ for $i \geq 1$. The hypothesis (iii) on supports implies that $M_{\mathfrak{p}} \neq 0$, and now Remark 3.4 yields a contradiction. We conclude that $\top_R N = 0$.

Applying the claim to the maximal ideal $\mathfrak{p}$ of R yields the required vanishing. □

Remark 3.7. (i) The hypothesis (iii) of Lemma 3.6 holds when, for example, the support of N is contained in that of M. Moreover, if R is a domain and M and N are nonzero, then, since $M \otimes_R N$ is torsion-free, we see that $\operatorname{Supp}(M \otimes_R N) = \operatorname{Spec}(R)$, whence $\operatorname{Supp}(M) = \operatorname{Spec}(R)$.

(ii) Most of the hypotheses in Lemma 3.6 are essential; see the discussion after [Huneke and Wiegand 1997, Remark 1.5]. Notice, without the assumption that $\dim(R) \geq 1$, the lemma would fail. Take, for example, $R = \mathbb{C}[x]/(x^2)$ and $M = R/(x) = N$. The vanishing of θ is also essential: let $R = \mathbb{C}[\![x, y]\!]/(xy)$, $M = R/(x)$ and $N = R/(x^2)$. Then the pair M, N satisfies conditions (ii) and (iii) of Lemma 3.6. On the other hand $\operatorname{Tor}_{2i+1}^R(M, N) \cong k$ for all $i \geq 0$, and $\operatorname{Tor}_{2i}^R(M, N) = 0$ for all $i \geq 1$. (Thus $\theta^R(M, N) = -1$.)

The completion of any regular ring is a hypersurface in an unramified regular local ring; see §2.2. Hence the following consequence of Lemma 3.6 extends [Lichtenbaum 1966, Corollary 3], which in turn builds on [Auslander 1961, Theorem 3.2]; see C. Miller's result recorded as Corollary B.3 here.

Proposition 3.8. *Let $(R, \mathfrak{m})$ be a d-dimensional local ring whose completion is a hypersurface in an unramified regular local ring, with $d \geq 1$, and let M be a* *finitely generated R-module. Assume* $\operatorname{pd}_{R_\mathfrak{p}}(M_\mathfrak{p}) < \infty$ *for all prime ideals* $\mathfrak{p} \neq \mathfrak{m}$ *and that* $\theta^R(M, -) = 0$. *If* $\bigotimes_R^n M$ *is torsion-free for some integer* $n \geq 2$, *then* $\operatorname{pd}(M) \leq (d-1)/n$. *Consequently, if M is not free, then* $\bigotimes_R^n M$ *has torsion for each* $n \geq \max\{2, d\}$.

Proof. We may assume $M \neq 0$. Iterating Lemma 3.6 shows that $\bigotimes_R^p M$ is torsion-free for $p = 1, \dots, n$, and that $\operatorname{Tor}_i^R(M, \bigotimes_R^{p-1} M) = 0$ for all $i \geq 1$. Taking $p = 2$, we see from [Huneke and Wiegand 1997, Theorem 1.9] that $\operatorname{pd}(M) < \infty$. Since $\operatorname{depth}(\bigotimes_R^n M) \geq 1$, one obtains, using [Auslander 1961, Corollary 1.3] and the Auslander–Buchsbaum formula [1957, Theorem 3.7],

$$n \cdot \operatorname{pd}(M) = \operatorname{pd}\Big(\bigotimes_R^n M\Big) = d - \operatorname{depth}\Big(\bigotimes_R^n M\Big) \leq d - 1. \qquad \square$$

Complete intersections. Hypersurfaces in complete intersections give the inductive step for our proof of Theorem 3.10; see §2.8 on pushforwards.

Lemma 3.9. *Let* $(S, \mathfrak{n})$ *be a complete intersection, and let R be a hypersurface in S. Let M and N be finitely generated torsion-free R-modules, and let E and F be the quasiliftings of M and N, respectively, to S. Assume* $\operatorname{Tor}_i^R(M, N)$ *has finite length for all* $i \gg 0$. *Let e be an integer with* $e \geq \max\{2, \operatorname{codim}(S) + 1\}$. *Then*

(i) $\operatorname{Tor}_i^S(E, F)$ *has finite length for all* $i \gg 0$, *and*

(ii) $\eta_{e-1}^S(E, F) = 2e \cdot \eta_e^R(M, N)$.

Proof. By hypothesis, $R \cong S/(f)$, where f is a non-zerodivisor in S. The spectral sequence associated to the change of rings $S \to R$ yields the following exact sequence — see [Lichtenbaum 1966, pp. 223–224] or [Murthy 1963, p. 561] — for all $n \geq 1$:

$$\cdots \to \operatorname{Tor}_{n-1}^R(M, N) \to \operatorname{Tor}_n^S(M, N) \to \operatorname{Tor}_n^R(M, N) \to \cdots.$$

Consequently $\operatorname{Tor}_i^S(M, N)$ has finite length for $i \gg 0$. Let M_1 and N_1 be the pushforwards of M and N, respectively. Since $\operatorname{Tor}_i^S(R, -) = 0$ for all $i \geq 2$, the sequences (2.8.2) and (2.8.1) yield isomorphisms

$$\operatorname{Tor}_i^S(E, N) \cong \operatorname{Tor}_{i+1}^S(M_1, N) \cong \operatorname{Tor}_i^S(M, N) \text{ for all } i \geq 2.$$

Arguing in the same vein, one gets isomorphisms

$$\operatorname{Tor}_i^S(E, F) \cong \operatorname{Tor}_i^S(E, N) \text{ for all } i \geq 2.$$

Hence the length of $\operatorname{Tor}_i^S(E, F)$ is finite for all $i \gg 0$, and so (i) holds.

Similar arguments show the η-pairing, over both R and S, as appropriate, is defined for all pairs (X, Y) with $X \in \{M, M_1, E\}$ and $Y \in \{N, N_1, F\}$.

By hypothesis, $\operatorname{codim}(S) \le e-1$, and hence $\operatorname{codim}(R) \le e$; see §2.1. Additivity of η along the exact sequences (2.8.1) and (2.8.2) thus gives

$$\eta_e^R(M, N) = -\eta_e^R(M_1, N) = \eta_e^R(M_1, N_1),$$
$$\eta_{e-1}^S(E, F) = -\eta_{e-1}^S(M_1, F) = \eta_{e-1}^S(M_1, N_1).$$

Our assumption that $e \ge \max\{2, \operatorname{codim} S + 1\}$, together with Theorem 4.1(3) from [Dao 2007], allow us to invoke Theorem 4.3(3) from the same reference, which says that

$$2e \cdot \eta_e^R(M_1, N_1) = \eta_{e-1}^S(M_1, N_1).$$

This gives (ii), completing the proof. □

The next theorem is our main result. As its hypotheses are technical, several of its consequences are discussed in Section 4; see Section 2 for background.

Theorem 3.10. *Let R be a local ring whose completion is a complete intersection in an* unramified *regular local ring, of relative codimension $c \ge 1$. Let M, N be finitely generated R-modules. Assume the following hold*:

(i) $\dim(R) \ge c$.

(ii) *The pair (M, N) satisfies* (SP_c).

(iii) $\operatorname{Supp}_R(\top_R N) \subseteq \operatorname{Supp}_R(M)$.

(iv) $\eta_c^R(M, N) = 0$.

Then, $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \ge 1$.*

Proof. The case $c = 1$ is Lemma 3.6. For $c \ge 2$, proceed by induction on c. We can assume R is complete, so that $R = Q/(\underline{f})$, where Q is an unramified regular local ring and $\underline{f} = f_1, \dots, f_c$ is a Q-regular sequence; see §2.2 and Lemma 2.12. Let $R = S/(f)$, where $S = Q/(f_1, \dots, f_{c-1})$ and $f = f_c$.

Hypothesis (ii) implies $\operatorname{Tor}_i^R(M, N)$ has finite length for all $i \gg 0$; see 3.5. Hence Corollary A.3 implies that, for all primes $\mathfrak{p}$ with $\operatorname{height}(\mathfrak{p}) \le c-1$,

$$\operatorname{Tor}_i^R(M, N)_{\mathfrak{p}} = 0 \text{ for all } i \ge 1. \tag{3.10.1}$$

Condition (ii) also implies M and N are torsion-free since $c \ge 2$; see 3.5. Hence quasiliftings E and F of M and N to S, respectively, exist; see §2.8. Using the vanishing of Tor modules in (3.10.1) and [Huneke et al. 2001, Theorem 4.8] — compare [Celikbas 2011, Proposition 3.1(7)] — one gets that

$$E \otimes_S F \text{ satisfies } (\mathrm{S}_{c-1}) \text{ as an } S\text{-module.} \tag{3.10.2}$$

It follows from [Huneke et al. 2001, Propositions 1.6 and 1.7] (see also [Celikbas 2011, Propositions 3.1(2) and 3.1(6)]) that the assumptions in (i) of (SP_c) pass to

E and F; see Notation 3.5. So,

$$E \text{ and } F \text{ satisfy } (\mathrm{S}_{c-1}) \text{ as } S\text{-modules.} \tag{3.10.3}$$

Lemma 3.9 guarantees that $\mathrm{Tor}_i^S(E, F)$ has finite length for all $i \gg 0$ and that $\eta_{c-1}(E, F) = 0$. In particular the pair E, F satisfies (SP_{c-1}) over the ring S. Moreover, E and F, being syzygies, are torsion-free, so we indeed have that $\mathrm{Supp}_S(\top_S F) \subseteq \mathrm{Supp}_S(E)$. Now the inductive hypothesis implies that

$$\mathrm{Tor}_i^S(E, F) = 0 \text{ for all } i \geq 1. \tag{3.10.4}$$

Condition (ii) also implies that $M \otimes_R N$ is reflexive since $c \geq 2$; see §2.6. Furthermore, $\mathrm{Tor}_i^R(M, N)_{\mathfrak{p}} = 0$ for all $i \geq 1$ and for all $\mathfrak{p} \in \mathrm{Spec}(R)$ with $\mathrm{height}(\mathfrak{p}) \leq 1$; see (3.10.1). Thus Proposition 2.10 and (3.10.4) yield $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq 1$. □

Remark 3.11. In Theorem 3.10, if $c \geq 2$, hypothesis (ii) implies that N is torsion-free, i.e., $\top_R N = 0$; see §2.6 and Notation 3.5. Thus, when $c \geq 2$, hypothesis (iii) of Theorem 3.10 is redundant.

When $\dim(R) > c$, the equivalence of (i) and (ii) in the following corollary seems interesting; see also §2.3. Actually, in that case the equivalence of (ii) and (iii) holds without the assumption that $\eta_c^R(M, N) = 0$. See [Celikbas 2011, Corollary 2.4].

Corollary 3.12. *Let R be an isolated singularity whose completion is a complete intersection in an unramified regular local ring, of relative codimension c. Let M and N be maximal Cohen–Macaulay R-modules. Assume* $\dim(R) \geq c$. *Assume further that* $\eta_c^R(M, N) = 0$. *The following conditions are equivalent*:

(i) *$M \otimes_R N$ satisfies* (S_c).

(ii) *$M \otimes_R N$ is maximal Cohen–Macaulay.*

(iii) $\mathrm{Tor}_i^R(M, N) = 0$ *for all $i \geq 1$, and hence the depth formula holds.*

Over a complete intersection, vanishing of Ext is closely related to vanishing of Tor: $\mathrm{Ext}_R^i(M, N) = 0$ for all $i \gg 0$ if and only if $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \gg 0$; see [Avramov and Buchweitz 2000, Remark 6.3]. Our next example shows the hypotheses of Theorem 3.10 do *not* force the vanishing of $\mathrm{Ext}_R^i(M, N)$ for all $i \geq 1$.

Example 3.13. Let $(R, \mathfrak{m}, k)$ be a complete intersection with $\mathrm{codim}(R) = 2$ and $\dim(R) \geq 3$. Let N be the d-th syzygy of k, where $d = \dim(R)$, and let M be the second syzygy of $R/(\underline{x})$, where $\underline{x}$ is a maximal R-regular sequence.

Note that N is maximal Cohen–Macaulay, $\mathrm{depth}(M) = 2$, and $N_{\mathfrak{p}}$ is free over $R_{\mathfrak{p}}$ for all primes $\mathfrak{p} \neq \mathfrak{m}$. It follows, since $\mathrm{pd}(M) < \infty$, that $\eta_2^R(M, N) = 0$ and $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq 1$; see §3.1 and Theorem A.1. Therefore the depth formula §2.3 shows that $\mathrm{depth}(M \otimes_R N) = 2$. Since M is a second syzygy, it satisfies (S_2) and hence $M \otimes_R N$ satisfies (S_2); see §2.6. In particular, the pair M, N satisfies (SP_2); see 3.5. However $\operatorname{Ext}_R^{d-2}(M, N) = \operatorname{Ext}^d(R/(\underline{x}), N) \neq 0$; see, for example, [Matsumura 1989, Chapter 19, Lemma 1(iii)].

Here is the extension of Dao's theorem [2007, Theorem 7.7] promised in the introduction (compare Theorem 1.2):

Corollary 3.14. *Let R be a local ring that is a complete intersection, and let M and N be finitely generated R-modules. Assume that the following conditions hold for some integer $e \geq \operatorname{codim}(R)$:*

(i) *M and N satisfy* (S_e).

(ii) *$M \otimes_R N$ satisfies* (S_{e+1}).

(iii) *$M_{\mathfrak{p}}$ is a free for all prime ideals $\mathfrak{p}$ of R of height at most e.*

Then $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \geq 1$, and hence the depth formula holds.*

Proof. If $e = 0$ this is a theorem in [Auslander 1961] and [Lichtenbaum 1966, Corollary 2]. Assume now that $e \geq 1$. We use induction on $\dim R$. If $\dim R \leq e$, condition (iii) implies that M is free, and there is nothing to prove. Assuming $\dim R \geq e + 1$, we note that the hypotheses localize, so $\operatorname{Tor}_i^R(M, N)_{\mathfrak{p}} = 0$ for each $i \geq 1$ and each prime ideal $\mathfrak{p}$ in the punctured spectrum of R; that is to say, $\operatorname{Tor}_i^R(M, N)$ has finite length for all $i \geq 1$. Thus the pair M, N satisfies (SP_{e+1}). Moreover, since $\operatorname{codim} R < e + 1$, we have $\eta_{e+1}^R = 0$ by item (iv) of §3.1. The completion of R can be realized as a complete intersection, of relative codimension $e+1$, in an unramified regular local ring (see §2.2). Hence the desired result follows from Theorem 3.10. □

4. Vanishing of η

In this section we apply our results to situations where the η-pairing is known to vanish. We know, from Theorem 3.10, that, as long as the critical hypothesis $\eta_c^R(M, N) = 0$ holds, we can replace c with $c - 1$ in the hypotheses of Theorem 1.2 and still conclude the vanishing of Tor. Although it is not easy to verify vanishing of η (see Conjectures 3.3), there are several classes of rings R for which it is known that $\eta^R(M, N) = 0$ for all finitely generated R-modules M and N. For example, if R is an even-dimensional simple ("ADE") singularity in characteristic zero, then Dao observed [2013a, Corollary 3.16] that $\theta^R(M, N) = 0$; see [Dao 2013a, Corollary 3.6] and also [Dao 2013a, §3] for more examples.

Now we give a localized version of a vanishing theorem for graded rings, due to Moore, Piepmeyer, Spiroff, and Walker [2013].

Proposition 4.1. *Let k be a perfect field and $Q = k[x_1, \dots, x_n]$ the polynomial ring with the standard grading. Let $\underline{f} = f_1, \dots, f_c$ be a Q-regular sequence of* *homogeneous polynomials, with* $c \geq 2$. *Put* $A = Q/(\underline{f})$ *and* $R = A_{\mathfrak{m}}$, *where* $\mathfrak{m} = (x_1, \dots, x_n)$. *Assume that* $A_{\mathfrak{p}}$ *is a regular local ring for each* $\mathfrak{p}$ *in* $\operatorname{Spec}(A)\setminus\{\mathfrak{m}\}$. *Then* $\eta_c^R(M, N) = 0$ *for all finitely generated* R*-modules* M *and* N. *In particular, if* $n \geq 2c$ *and the pair* M, N *satisfies* (SP_c), *then* M *and* N *are Tor-independent.*

Proof. Choose finitely generated A-modules U and V such that $U_{\mathfrak{m}} \cong M$ and $V_{\mathfrak{m}} \cong N$. For any maximal ideal $\mathfrak{n} \neq \mathfrak{m}$, the local ring $A_{\mathfrak{n}}$ is regular, and hence $\operatorname{Tor}_i^A(U, V)_{\mathfrak{n}} = 0$ for $i \gg 0$. It follows that the map $\operatorname{Tor}_i^A(U, V) \to \operatorname{Tor}_i^R(M, N)$ induced by the localization maps $U \to M$ and $V \to N$ is an isomorphism for $i \gg 0$. Also, for any A-module supported at $\mathfrak{m}$, its length as an A-module is equal to its length as an R-module. In conclusion, $\eta_c^R(M, N) = \eta_c^A(U, V)$.

As k is perfect, the hypothesis on A implies that the k-algebra $A_{\mathfrak{p}}$ is smooth for each nonmaximal prime $\mathfrak{p}$ in A; see [Eisenbud 1995, Corollary 16.20]. Thus, the morphism of schemes $\operatorname{Spec}(R)\setminus\{\mathfrak{m}\} \to \operatorname{Spec}(k)$ is smooth. Now [Moore et al. 2013, Corollary 4.7] yields $\eta_c^A(U, V) = 0$, and hence $\eta_c^R(M, N) = 0$. It remains to note that if $n \geq 2c$, then $\dim R \geq c$, so Theorem 3.10 applies. □

Next, we quote a recent theorem due to Walker; it provides strong support for Conjectures 3.3, at least in equicharacteristic zero.

Theorem 4.2 [Walker 2014, Theorem 1.2]. *Let* k *be a field of characteristic zero, and let* Q *a smooth* k*-algebra. Let* $\underline{f} = f_1, \dots, f_c$ *be a* Q*-regular sequence, with* $c \geq 2$, *and put* $A = Q/(f_1, \dots, f_c)$. *Assume the singular locus* $\{\mathfrak{p} \in \operatorname{Spec}(A) :$ $A_{\mathfrak{p}}$ *is not regular*$\}$ *is a finite set of maximal ideals of* A. *Then* $\eta_c^A(U, V) = 0$ *for all finitely generated* A*-modules* U, V.

Corollary 4.3. *With* A *as in Theorem 4.2, put* $R = A_{\mathfrak{m}}$, *where* $\mathfrak{m}$ *is any maximal ideal of* A. *Then* $\eta_c^R(M, N) = 0$ *for all finitely generated* R*-modules* M *and* N. *In particular, if* $\dim R \geq c$ *and the pair* M, N *satisfies* (SP_c), *then* M *and* N *are Tor-independent.*

Proof. By inverting a suitable element of Q, we may assume that $A_{\mathfrak{p}}$ is a regular local ring for every prime ideal $\mathfrak{p} \neq \mathfrak{m}$. Now proceed as in the first paragraph of the proof of Proposition 4.1. □

Theorem 4.4. *Let* $(R, \mathfrak{m}, k)$ *be a two-dimensional, equicharacteristic, normal, excellent complete intersection of codimension* c, *with* $c \in \{1, 2\}$, *and let* M *and* N *be finitely generated* R*-modules. Assume* k *is contained in the algebraic closure of a finite field. Assume further that* M *and* N *satisfy conditions* (i) *and* (ii) *of* (SP_c). *Then* $\operatorname{Tor}_i^R(M, N) = 0$ *for all* $i \geq 1$.

Proof. The completion $\hat{R}$ is an isolated singularity because R is excellent; see [Leuschke and Wiegand 2012, Proposition 10.9], and so $\hat{R}$ is a normal domain. Replacing R by $\hat{R}$, we may assume that $R = S/(\underline{f})$, where $(S, \mathfrak{n}, k)$ is a regular local ring and $\underline{f}$ is a regular sequence in $\mathfrak{n}^2$ of length c. Let $\bar{k}$ be an algebraic closure of k, and choose a *gonflement* $S \hookrightarrow (\bar{S}, \bar{\mathfrak{n}}, \bar{k})$ lifting the field extension $k \hookrightarrow \bar{k}$; see [2012, Chapter 10, §3]. This is a flat local homomorphism and is an inductive limit of étale extensions. Moreover, $\mathfrak{n}\bar{S} = \bar{\mathfrak{n}}$, so $\bar{S}$ is a regular local ring. By [2012, Proposition 10.15], both $\bar{S}$ and $\bar{R} := \bar{S}/(\underline{f})$ are excellent, and $\bar{R}$ is an isolated singularity. Therefore $(\bar{R}, \bar{\mathfrak{m}}, \bar{k})$ is a normal domain. Finally, we pass to the completion $\hat{S}$ of $\bar{S}$ and put $\Lambda = \hat{S}/(\underline{f})$. This is still an isolated singularity, a normal domain, and a complete intersection of codimension c. Moreover, our hypotheses on M and N ascend along the flat local homomorphism $R \to \Lambda$; see Lemma 2.12. Since Λ is an isolated singularity, $\operatorname{Tor}_i^{\Lambda}(\Lambda \otimes_R M, \Lambda \otimes_R N)$ has finite length for $i \gg 0$; thus the pair $\Lambda \otimes_R M$, $\Lambda \otimes_R N$ satisfies (SP_c).

It follows from [Celikbas and Dao 2011, Proposition 2.5 and Remark 2.6] that $G(\Lambda)/L$ is torsion, where $G(\Lambda)$ is the Grothendieck group of Λ and L is the subgroup generated by classes of modules of finite projective dimension. This implies that $\eta_c^{\Lambda}(\Lambda \otimes_R M, \Lambda \otimes_R N) = 0$; see [Dao 2013a, Corollary 3.1] and the paragraph preceding it. Now Theorem 3.10 implies that $\operatorname{Tor}_i^{\Lambda}(\Lambda \otimes_R M, \Lambda \otimes_R N) = 0$ for all $i \geq 1$: the requirement on supports is automatically satisfied, since Λ is a domain; see Remark 3.7(i). Faithfully flat descent completes the proof. □

Appendix A: An application of pushforwards

In Theorem A.4 we use pushforwards to generalize [Celikbas 2011, Theorem 3.16]. We have two preparatory results. The first one is a special case of a theorem of Jorgensen:

Theorem A.1 [Jorgensen 1999, Theorem 2.1]. *Let R be a complete intersection and let M and N be finitely generated R-modules. Assume M is maximal Cohen–Macaulay. If* $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \gg 0$, then* $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \geq 1$.*

Corollary A.2. *Let R be a complete intersection and let M, N be finitely generated R-modules. If* $\operatorname{Tor}_i^R(M, N)$ *is torsion for all $i \gg 0$, then* $\operatorname{Tor}_i^R(M, N)$ *is torsion for all $i \geq 1$.*

Proof. Let $\mathfrak{p}$ be a minimal prime ideal of R. By §2.4, it suffices to prove that $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, N_{\mathfrak{p}}) = 0$ for all $i \geq 1$. For that we may assume $M_{\mathfrak{p}} \neq 0$. Then, since $R_{\mathfrak{p}}$ is artinian, it follows that $M_{\mathfrak{p}}$ is a maximal Cohen–Macaulay $R_{\mathfrak{p}}$-module. Therefore, Theorem A.1 gives the desired vanishing. □

Corollary A.3. *Let R be a complete intersection, and let M, N be finitely generated R-modules. Assume M satisfies* (S_w), *where w is a positive integer, and that* $\operatorname{Tor}_i^R(M, N)$ *has finite length for all $i \gg 0$. Let $\mathfrak{p}$ be a nonmaximal prime ideal of R such that* $\operatorname{height}(\mathfrak{p}) \leq w$. *Then* $\operatorname{Tor}_i^R(M, N)_{\mathfrak{p}} = 0$ *for all $i \geq 1$.*

Proof. Serre's condition (S_w) localizes, so $M_{\mathfrak{p}}$ is either zero or a maximal Cohen–Macaulay $R_{\mathfrak{p}}$-module; see §2.6. As $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, N_{\mathfrak{p}}) = 0$ for $i \gg 0$, Theorem A.1 implies that $\operatorname{Tor}_i^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, N_{\mathfrak{p}}) = 0$ for all $i \geq 1$. □

The next theorem generalizes [Celikbas 2011, Theorem 3.16; see also Theorems 3.4 and 3.15]; we emphasize that the ambient regular local ring in Theorem A.4 is allowed to be ramified.

Theorem A.4. *Let R be a complete intersection with* $\dim R \geq \operatorname{codim} R$, *and let M and N be finitely generated R-modules. Assume the pair M, N satisfies* (SP_c) *for some $c \geq$* $\operatorname{codim} R$. *If $c = 1$, assume further that M or N is torsion-free. If* $\operatorname{Tor}_1^R(M, N) = 0$, *then* $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i \geq 1$.*

Proof. Without loss of generality, one may assume that $c = \operatorname{codim} R$. When $c = 0$, the desired result is the rigidity theorem of Auslander [1961] and Lichtenbaum [1966], so in the remainder of the proof we assume that $c \geq 1$.

Assume first that $c = 1$. By hypotheses $\operatorname{Tor}_i^R(M, N)$ has finite length for $i \gg 0$ and $M \otimes_R N$ is torsion-free; see Notation 3.5. Moreover, we may assume N (say) is torsion-free. Tensoring M with the pushforward §2.8 for N gives the following:

$$\operatorname{Tor}_1^R(M, N_1) \hookrightarrow M \otimes_R N, \tag{A.4.1}$$

$$\operatorname{Tor}_i^R(M, N_1) \cong \operatorname{Tor}_{i-1}^R(M, N) \quad \text{for all } i \geq 2. \tag{A.4.2}$$

Equation (A.4.2) implies that $\operatorname{Tor}_i^R(M, N_1)$ has finite length for all $i \gg 0$. Therefore, since $\dim(R) \geq 1$, $\operatorname{Tor}_i^R(M, N_1)$ is torsion for all $i \gg 0$; see §2.4. Now Corollary A.2 implies that $\operatorname{Tor}_i^R(M, N_1)$ is torsion for all $i \geq 1$. As $M \otimes_R N$ is torsion-free, we deduce from (A.4.1) that $\operatorname{Tor}_1^R(M, N_1) = 0$. By (A.4.2) we have $\operatorname{Tor}_2^R(M, N_1) \cong \operatorname{Tor}_1^R(M, N) = 0$. Therefore $\operatorname{Tor}_2^R(M, N_1) = 0 = \operatorname{Tor}_1^R(M, N_1)$, and hence Murthy's rigidity theorem [1963, Theorem 1.6] implies that $\operatorname{Tor}_i^R(M, N_1) = 0$ for all $i \geq 1$. Now (A.4.2) completes the proof for the case $c = 1$.

Assume now that $c \geq 2$. We define a sequence $M_0, M_1, \ldots, M_{c-1}$ of finitely generated modules by setting $M_0 = M$, and M_n to be the pushforward of M_{n-1}, for all $n = 1, \ldots, c-1$. These pushforwards exist: M_0 satisfies (S_{c-1}) by Hypothesis 3.5(i), and so, by Proposition 2.9(i),

(1) each M_n satisfies (S_{c-n-1}).

For the desired result, it suffices to prove that $\operatorname{Tor}_i^R(M_{c-1}, N) = 0$ for all $i \geq c$. We will, in fact, prove this for all $i \geq 1$. To this end, we establish by induction that the following hold for $n = 0, \ldots, c-1$:

(2) $M_n \otimes_R N$ satisfies (S_{c-n});

(3) $\operatorname{Tor}_i^R(M_n, N)$ has finite length for all $i \gg 0$;

(4) $\operatorname{Tor}_i^R(M_n, N) = 0$ for $i = 1, \ldots, n+1$.

For $n = 0$, conditions (2) and (3) are part of Hypothesis 3.5, while (4) is from our hypothesis that $\operatorname{Tor}_1^R(M, N) = 0$; recall that $M_0 = M$. Assume that (2), (3) and (4) hold for some integer n with $0 \le n \le c-2$.

Tensor the pushforward of M_n with N — see §2.8 — to obtain

(A.4.3) $$\operatorname{Tor}_i^R(M_{n+1}, N) \cong \operatorname{Tor}_{i-1}^R(M_n, N) \text{ for all } i \ge 2,$$

and the following exact sequence in which F is finitely generated and free:

(A.4.4) $$0 \to \operatorname{Tor}_1^R(M_{n+1}, N) \to M_n \otimes_R N \to F \otimes_R N \to M_{n+1} \otimes_R N \to 0.$$

Induction and (A.4.3) imply that $\operatorname{Tor}_i^R(M_{n+1}, N)$ has finite length for all $i \gg 0$, so (3) holds; furthermore, by Corollary A.2, $\operatorname{Tor}_i^R(M_{n+1}, N)$ is torsion for all $i \ge 1$. (Recall that $\dim(R) \ge \operatorname{codim}(R) = c \ge 1$ so that finite length modules are torsion.) Since $n \le c-1$, condition (2) implies that $M_n \otimes_R N$ satisfies (S_1) and hence $M_n \otimes_R N$ is torsion-free; therefore the exact sequence (A.4.4) forces $\operatorname{Tor}_1^R(M_{n+1}, N)$ to vanish. Now (A.4.3) gives (4). It remains to verify (2), namely, that $M_{n+1} \otimes_R N$ satisfies (S_{c-n-1}). To that end, let $\mathfrak{p} \in \operatorname{Supp}(M_{n+1} \otimes_R N)$. We will verify that $\operatorname{depth}_{R_\mathfrak{p}}(M_{n+1} \otimes_R N)_\mathfrak{p} \ge \min\{c-n-1, \operatorname{height}(\mathfrak{p})\}$; see §2.6.

Suppose $\operatorname{height}(\mathfrak{p}) \ge c-n$. Recall, by Hypothesis 3.5(i), N satisfies (S_{c-1}). Hence $F \otimes_R N$, a direct sum of copies of N, satisfies (S_{c-n-1}). In particular it follows that $\operatorname{depth}_{R_\mathfrak{p}}(F \otimes_R N)_\mathfrak{p} \ge c-n-1$. Furthermore, by (2) of the induction hypothesis, we have that $\operatorname{depth}_{R_\mathfrak{p}}(M_n \otimes_R N)_\mathfrak{p} \ge c-n$. Recall that $\operatorname{Tor}_1^R(M_{n+1}, N) = 0$. Therefore, localizing the short exact sequence in (A.4.4) at $\mathfrak{p}$, we conclude by the depth lemma that $\operatorname{depth}_{R_\mathfrak{p}}(M_{n+1} \otimes_R N)_\mathfrak{p} \ge c-n-1$.

Next assume $\operatorname{height}(\mathfrak{p}) \le c-n-1$. We want to show that $(M_{n+1} \otimes_R N)_\mathfrak{p}$ is maximal Cohen–Macaulay. By the induction hypotheses, $\operatorname{Tor}_i^R(M_n, N)$ has finite length for all $i \gg 0$. As $n \ge 0$, we see that $\dim(R) \ge \operatorname{codim}(R) = c \ge c-n$, whence $\mathfrak{p}$ is not the maximal ideal. Thus $\operatorname{Tor}_i^R(M_n, N)_\mathfrak{p} = 0$ for all $i \gg 0$. Now, setting $w = c-n-1$ and using Corollary A.3 for the pair M_n, N, we conclude that $\operatorname{Tor}_i^R(M_n, N)_\mathfrak{p} = 0$ for all $i \ge 1$. Then (A.4.3) and the already established fact that $\operatorname{Tor}_1^R(M_{n+1}, N) = 0$ give that $\operatorname{Tor}_i^R(M_{n+1}, N)_\mathfrak{p} = 0$ for all $i \ge 1$. Thus, the depth formula holds — see §2.3:

$$\operatorname{depth}_{R_\mathfrak{p}}(M_{n+1})_\mathfrak{p} + \operatorname{depth}_{R_\mathfrak{p}}(N_\mathfrak{p}) = \operatorname{depth}(R_\mathfrak{p}) + \operatorname{depth}_{R_\mathfrak{p}}(M_{n+1} \otimes_R N)_\mathfrak{p}.$$

Since Serre's conditions localize, $N_\mathfrak{p}$ is maximal Cohen–Macaulay over $R_\mathfrak{p}$; see Hypothesis 3.5(i). Also, $(M_{n+1})_\mathfrak{p}$ is maximal Cohen–Macaulay whether or not $(M_n)_\mathfrak{p}$ is zero; see the pushforward sequence or Proposition 2.9(ii). By the depth formula, $(M_{n+1} \otimes_R N)_\mathfrak{p}$ is maximal Cohen–Macaulay. Thus $M_{n+1} \otimes_R N$ satisfies (2), and the induction is complete.

Now we parallel the argument for the case $c = 1$. At the end, $\operatorname{Tor}_i^R(M_{c-1}, N)$ has finite length for all $i \gg 0$, and is equal to 0 for $i = 1, \ldots, c$. Tensoring M_{c-1}

with the pushforward of N, we get

$$\operatorname{Tor}_i^R(M_{c-1}, N_1) \cong \operatorname{Tor}_{i-1}^R(M_{c-1}, N) \quad \text{for all } i \geq 2, \tag{A.4.5}$$

$$\operatorname{Tor}_1^R(M_{c-1}, N_1) \hookrightarrow M_{c-1} \otimes_R N. \tag{A.4.6}$$

In view of (A.4.5), it suffices to show that $\operatorname{Tor}_1^R(M_{c-1}, N_1) = 0$: this will imply $\operatorname{Tor}_i^R(M_{c-1}, N_1) = 0$ for all $i = 1, \ldots, c+1$, and hence Murthy's rigidity theorem [1963, Theorem 1.6] will yield that $\operatorname{Tor}_i^R(M_{c-1}, N_1) = 0$ for all $i \geq 1$, and consequently $\operatorname{Tor}_i^R(M_{c-1}, N) = 0$ for all $i \geq 1$ by (A.4.5). We know that $M_{c-1} \otimes_R N$ is torsion-free. Therefore we use (A.4.6) and Corollary A.2, and obtain $\operatorname{Tor}_1^R(M_{c-1}, N_1) = 0$, as we did in the case $c = 1$. □

Appendix B: Amending the literature

We use Theorem A.4 to give a different proof of an important result of Huneke and Wiegand; see Theorem B.2 and the ensuing paragraph. We also point out a missing hypothesis in a result of C. Miller [1998, Theorem 3.1], and state the corrected form of her theorem in Corollary B.3. At the end of the paper we indicate an alternative route to the proof of [Huneke and Wiegand 1994, Theorem 3.1], the main theorem in that reference.

Theorem B.1 [Huneke and Wiegand 1994]. *Let R be a hypersurface and let M, N be finitely generated R-modules. If M or N has rank and $M \otimes_R N$ is maximal Cohen–Macaulay, then both M and N are maximal Cohen–Macaulay, and either M or N is free.*

Theorem B.1 and its variations have been analyzed, used, and studied in the literature; see [Celikbas and Wiegand 2015] and [Dao 2013b] for some history and many consequences of the theorem. The following result [Huneke and Wiegand 1994, Theorem 2.7] played an important role in its proof.

Theorem B.2 [Huneke and Wiegand 1994]. *Let R be a hypersurface and let M, N be nonzero finitely generated R-modules. Assume $M \otimes_R N$ is reflexive and that N has rank. Then the following conditions hold*:

(i) $\operatorname{Tor}_i^R(M, N) = 0$ *for all* $i \geq 1$.

(ii) *M is reflexive, and N is torsion-free.*

Theorem B.2 was established in [Huneke and Wiegand 1994, Theorem 2.7]. However, the conclusion there was that *both* M and N are reflexive, and the proof of this stronger claim is flawed. Dao realized this, and subsequently Huneke and Wiegand corrected their oversight [2007]. A similar flaw can be found in [Miller 1998]; see Theorems 1.3 and 1.4 there and compare with our correction in Corollary B.3. The version stated above reflects our current understanding and is from [Celikbas and Piepmeyer 2014]. We do not yet know whether N is forced to be reflexive — that is, the question below remains open; cf. [Huneke and Wiegand 1994, Theorem 2.7] and [Miller 1998, Theorem 1.3].

Question. Let R be a hypersurface and M, N nonzero finitely generated R-modules. If N has rank and $M \otimes_R N$ is reflexive, must *both* M and N be reflexive?

This question has been recently studied in [Celikbas and Piepmeyer 2014], which gives partial answers using the New Intersection Theorem.

We now show how Theorem B.2 follows from Theorem A.4. In fact, one needs only the case $c = 1$ of Theorem A.4.

Proof of Theorem B.2 using Theorem A.4. Set $d = \dim R$. If $d = 0$, then N is free (since it has rank), so all is well. From now on assume $d \geq 1$. We remark at the outset that neither M nor N can be torsion, i.e., $\perp_R M \neq 0$ and $\perp_R N \neq 0$. Also, by the assumption of rank, $\mathrm{Supp}(N) = \mathrm{Spec}(R)$. Suppose first that both M and N are torsion-free; we will prove (i) by induction on $d = \dim R$. Let M_1 denote the pushforward of M; see §2.8. Then $\mathrm{Tor}_1^R(M_1, N)$ is torsion as N has rank. Since $M \otimes_R N$ is torsion-free, applying $- \otimes_R N$ to (2.8.1) shows that

$$\mathrm{Tor}_1^R(M_1, N) = 0. \tag{B.2.1}$$

Suppose for the moment that $d = 1$. Since N has rank, there is an exact sequence

$$0 \to N \to F \to C \to 0,$$

in which F is free and C is torsion; see [Huneke and Wiegand 1994, Lemma 1.3]. Note that C is of finite length since $d = 1$. Note also that $\mathrm{Tor}_2^R(M_1, C) \cong \mathrm{Tor}_1^R(M_1, N) = 0$; see (B.2.1). Therefore, Corollary 2.3 from that same reference implies that $\mathrm{Tor}_i^R(M_1, C) = 0$ for all $i \geq 2$, and hence $\mathrm{Tor}_i^R(M_1, N) = 0$ for all $i \geq 1$. Now (2.8.1) establishes (i).

Still assuming that both M and N are torsion-free, let $d \geq 2$. The inductive hypothesis implies that $\mathrm{Tor}_i^R(M, N)$ has finite length for all $i \geq 1$. In particular $\mathrm{Tor}_i^R(M, N)_{\mathfrak{q}} = 0$ for all prime ideals $\mathfrak{q}$ of R of height at most one. Therefore, Proposition 2.10 shows that $M_1 \otimes_R N$ is torsion-free, that is, $M_1 \otimes_R N$ satisfies (S_1); see §2.5 and §2.6. Furthermore, from the pushforward exact sequence (2.8.1), we see that $\mathrm{Tor}_i^R(M_1, N)$ has finite length for all $i \geq 2$. Consequently the pair M_1, N satisfies (SP_1). Now Theorem A.4, applied to M_1, N, shows that $\mathrm{Tor}_i^R(M_1, N) = 0$ for all $i \geq 1$. By (2.8.1), we see that $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq 1$. This proves (i) under the additional assumption that M and N are torsion-free.

Since $M \otimes_R N$ is torsion-free, by Remark 3.4, there are isomorphisms

$$M \otimes_R N \cong M \otimes_R \perp_R N \cong \perp_R M \otimes_R N \cong \perp_R M \otimes_R \perp_R N.$$

In particular, $\perp_R M \otimes_R \perp_R N$ is also reflexive. As noted before, neither M nor N is torsion, so $\perp_R M$ and $\perp_R N$ are nonzero. As N has rank so does $\perp_R N$, so the already established part of the result (applied to $\perp_R M$ and $\perp_R N$) yields that $\mathrm{Tor}_i^R(\perp_R M, \perp_R N) = 0$ for $i \geq 1$. Given this, since $\perp_R M \otimes_R N$ is torsion-free by the isomorphisms above, applying Remark 3.4 to the R-modules $\perp_R M$ and N gives $N = \perp_R N$; then applying Remark 3.4 to M and N yields $M = \perp_R M$. In conclusion, M and N are torsion-free, and hence $\mathrm{Tor}_i^R(M, N) = 0$ for all $i \geq 1$. From the last, the depth formula holds.

The remaining step is to prove that M is reflexive. Since $\mathrm{Supp}(N) = \mathrm{Spec}(R)$, we have $\mathrm{depth}(N_{\mathfrak{p}}) \leq \mathrm{height}(\mathfrak{p})$ for all primes $\mathfrak{p}$ of R. Localizing the depth formula §2.3 shows Serre's condition (S_2) on M; see §2.6. ☐

The next result is due to C. Miller [1998]. In the original formulation, the essential requirement — that M have rank — is missing: for example, the module $M = R/(x)$ over the node $k[\![x, y]\!]/(xy)$ is not free, yet $M \otimes_R M$, which is just M, is maximal Cohen–Macaulay and hence reflexive. We state her result here in its corrected form and include a proof for completeness.

Corollary B.3 [Miller 1998, Theorem 3.1]. *Let R be a d-dimensional hypersurface and let M be a finitely generated R-module with rank. If $\bigotimes_R^n M$ is reflexive for some $n \geq \max\{2, d-1\}$, then M is free.*

Proof. If $d \leq 2$, then $\bigotimes_R^n M$ is maximal Cohen–Macaulay, and Theorem B.1 gives the result. Assume now that $d \geq 3$. Applying Theorem B.2 and [Huneke and Wiegand 1997, Theorem 1.9] repeatedly, we conclude the following:

(i) $\bigotimes_R^r M$ is reflexive for all $r = 1, \dots, n$.

(ii) $\mathrm{Tor}_i^R(M, \bigotimes_R^{r-1} M) = 0$ for all $i \geq 1$ and all $r = 2, \dots, n$.

(iii) $\mathrm{pd}(M) < \infty$.

It follows from (i) that $\mathrm{depth}(\bigotimes_R^r M) \geq 2$ for all $r = 1, \dots, n$; see §2.6. Also, (ii) implies the depth formula

$$\mathrm{depth}(M) + \mathrm{depth}\Big(\bigotimes_R^{r-1} M\Big) = d + \mathrm{depth}\Big(\bigotimes_R^{r} M\Big),$$

for all $r = 2, \dots, n$. One checks by induction on r that

$$r \cdot \mathrm{depth}(M) = (r-1) \cdot d + \mathrm{depth}\Big(\bigotimes_R^{r} M\Big),$$

for $r = 2, \dots, n$. By setting $r = n$, and using the inequalities $n \geq d - 1$ and $\mathrm{depth}(\bigotimes_R^n M) \geq 2$, we obtain

$$n \cdot \mathrm{depth}(M) \geq (n-1) \cdot d + 2 = n \cdot (d-1) + n - d + 2 \geq n \cdot (d-1) + 1.$$

Therefore, $\operatorname{depth}(M) \geq d$, that is, M is maximal Cohen–Macaulay. Now (iii) and the Auslander–Buchsbaum formula [1957, Theorem 3.7] imply that M is free. $\square$

A consequence of Theorems B.1 and B.2 is the following result:

Proposition B.4 [Huneke and Wiegand 1997, Theorem 1.9]. *Suppose M and N are finitely generated modules over a hypersurface R, and assume that* $\operatorname{Tor}_i^R(M, N) = 0$ *for $i \gg 0$. Then at least one of the modules has finite projective dimension.*

At about the same time, Miller [1998] obtained the same result independently, by an elegant, direct argument. As Miller observed in that reference, one can turn things around and easily deduce Theorem B.1 from Proposition B.4 and the vanishing result Theorem B.2.

Acknowledgment

We would like to thank the referee for a careful reading of the paper.

References

[Auslander 1961] M. Auslander, "Modules over unramified regular local rings", *Illinois J. Math.* **5** (1961), 631–647. MR 31 #3460 Zbl 0104.26202

[Auslander and Buchsbaum 1957] M. Auslander and D. A. Buchsbaum, "Homological dimension in local rings", *Trans. Amer. Math. Soc.* **85** (1957), 390–405. MR 19,249d Zbl 0078.02802

[Avramov and Buchweitz 2000] L. L. Avramov and R.-O. Buchweitz, "Support varieties and cohomology over complete intersections", *Invent. Math.* **142**:2 (2000), 285–318. MR 2001j:13017 Zbl 0999.13008

[Bourbaki 2006] N. Bourbaki, *Éléments de mathématique: algèbre commutative, chapitres 8 et 9*, Springer, Berlin, 2006. Reprint of the 1983 original. MR 2007h:13001 Zbl 1103.13003

[Buchweitz and Van Straten 2012] R.-O. Buchweitz and D. Van Straten, "An index theorem for modules on a hypersurface singularity", *Mosc. Math. J.* **12**:2 (2012), 237–259. MR 2978754 Zbl 1269.32016

[Celikbas 2011] O. Celikbas, "Vanishing of Tor over complete intersections", *J. Commut. Algebra* **3**:2 (2011), 169–206. MR 2012f:13030 Zbl 1237.13031 arXiv 0904.1408

[Celikbas and Dao 2011] O. Celikbas and H. Dao, "Asymptotic behavior of Ext functors for modules of finite complete intersection dimension", *Math. Z.* **269**:3-4 (2011), 1005–1020. MR 2860275 Zbl 1235.13010

[Celikbas and Piepmeyer 2014] O. Celikbas and G. Piepmeyer, "Syzygies and tensor product of modules", *Math. Z.* **276**:1-2 (2014), 457–468. MR 3150213 Zbl 06259147

[Celikbas and Wiegand 2015] O. Celikbas and R. Wiegand, "Vanishing of Tor, and why we care about it", *J. Pure Appl. Algebra* **219**:3 (2015), 429–448. MR 3279364 Zbl 1301.13017 arXiv 1302.2170

[Christensen and Jorgensen 2015] L. W. Christensen and D. A. Jorgensen, "Vanishing of Tate homology and depth formulas over local rings", *J. Pure Appl. Algebra* **219**:3 (2015), 464–481. MR 3279366 Zbl 06371703

[Dao 2007] H. Dao, "Asymptotic behavior of Tor over complete intersections and applications", preprint, 2007. arXiv 0710.5818

[Dao 2008] H. Dao, "Some observations on local and projective hypersurfaces", *Math. Res. Lett.* **15**:2 (2008), 207–219. MR 2009c:13032 Zbl 1229.13014 arXiv math/0701881

[Dao 2013a] H. Dao, "Decent intersection and Tor-rigidity for modules over local hypersurfaces", *Trans. Amer. Math. Soc.* **365**:6 (2013), 2803–2821. MR 3034448 Zbl 1285.13018

[Dao 2013b] H. Dao, "Some homological properties of modules over a complete intersection, with applications", pp. 335–371 in *Commutative algebra*, edited by I. Peeva, Springer, New York, 2013. MR 3051378 Zbl 1262.13024

[Eisenbud 1980] D. Eisenbud, "Homological algebra on a complete intersection, with an application to group representations", *Trans. Amer. Math. Soc.* **260**:1 (1980), 35–64. MR 82d:13013 Zbl 0444.13006

[Eisenbud 1995] D. Eisenbud, *Commutative algebra: with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, New York, 1995. MR 97a:13001 Zbl 0819.13001

[Epstein and Yao 2012] N. Epstein and Y. Yao, "Criteria for flatness and injectivity", *Math. Z.* **271**:3-4 (2012), 1193–1210. MR 2945604 Zbl 1245.13009

[Evans and Griffith 1985] E. G. Evans and P. Griffith, *Syzygies*, London Mathematical Society Lecture Note Series **106**, Cambridge University Press, 1985. MR 87b:13001 Zbl 0569.13005

[Foxby 1980] H.-B. Foxby, "Homological dimensions of complexes of modules", pp. 360–368 in *Séminaire d'Algèbre Paul Dubreil et Marie-Paule Malliavin, 32ème année* (Paris, 1979), edited by M.-P. Malliavin, Lecture Notes in Mathematics **795**, Springer, Berlin, 1980. MR 82a:13001 Zbl 0438.13010

[Heitmann 1993] R. C. Heitmann, "Characterization of completions of unique factorization domains", *Trans. Amer. Math. Soc.* **337**:1 (1993), 379–387. MR 93g:13006 Zbl 0792.13011

[Hochster 1981] M. Hochster, "The dimension of an intersection in an ambient hypersurface", pp. 93–106 in *Algebraic geometry* (Chicago, 1980), edited by A. Libgober and P. Wagreich, Lecture Notes in Mathematics **862**, Springer, Berlin, 1981. MR 83g:13017 Zbl 0472.13005

[Huneke and Jorgensen 2003] C. Huneke and D. A. Jorgensen, "Symmetry in the vanishing of Ext over Gorenstein rings", *Math. Scand.* **93**:2 (2003), 161–184. MR 2004k:13039 Zbl 1062.13005

[Huneke and Wiegand 1994] C. Huneke and R. Wiegand, "Tensor products of modules and the rigidity of Tor", *Math. Ann.* **299**:3 (1994), 449–476. MR 95m:13008 Zbl 0803.13008

[Huneke and Wiegand 1997] C. Huneke and R. Wiegand, "Tensor products of modules, rigidity and local cohomology", *Math. Scand.* **81**:2 (1997), 161–183. MR 2000d:13027 Zbl 0908.13010

[Huneke and Wiegand 2007] C. Huneke and R. Wiegand, "Correction to 'Tensor products of modules and the rigidity of Tor'", *Math. Ann.* **338**:2 (2007), 291–293. MR 2007m:13018 Zbl 1122.13301

[Huneke et al. 2001] C. Huneke, D. A. Jorgensen, and R. Wiegand, "Vanishing theorems for complete intersections", *J. Algebra* **238**:2 (2001), 684–702. MR 2002h:13025 Zbl 1082.13504

[Iyengar 1999] S. B. Iyengar, "Depth for complexes, and intersection theorems", *Math. Z.* **230**:3 (1999), 545–567. MR 2000a:13027 Zbl 0927.13015

[Jorgensen 1999] D. A. Jorgensen, "Complexity and Tor on a complete intersection", *J. Algebra* **211**:2 (1999), 578–598. MR 99k:13014 Zbl 0926.13007

[Leuschke and Wiegand 2012] G. J. Leuschke and R. Wiegand, *Cohen–Macaulay representations*, Mathematical Surveys and Monographs **181**, American Mathematical Society, Providence, RI, 2012. MR 2919145 Zbl 1252.13001

[Lichtenbaum 1966] S. Lichtenbaum, "On the vanishing of Tor in regular local rings", *Illinois J. Math.* **10** (1966), 220–226. MR 32 #5688 Zbl 0139.26601

[Matsumura 1989] H. Matsumura, *Commutative ring theory*, 2nd ed., Cambridge Studies in Advanced Mathematics **8**, Cambridge University Press, 1989. MR 90i:13001 Zbl 0666.13002

[Miller 1998] C. Miller, “Complexity of tensor products of modules and a theorem of Huneke–Wiegand”, *Proc. Amer. Math. Soc.* **126**:1 (1998), 53–60. MR 98c:13022 Zbl 0886.13006

[Moore et al. 2011] W. F. Moore, G. Piepmeyer, S. Spiroff, and M. E. Walker, “Hochster’s theta invariant and the Hodge–Riemann bilinear relations”, *Advances in Math.* **226**:2 (2011), 1692–1714. MR 2011m:13029 Zbl 1221.13027

[Moore et al. 2013] W. F. Moore, G. Piepmeyer, S. Spiroff, and M. E. Walker, “The vanishing of a higher codimension analogue of Hochster’s theta invariant”, *Math. Z.* **273**:3-4 (2013), 907–920. MR 3030683 Zbl 1278.13013

[Murthy 1963] M. P. Murthy, “Modules over regular local rings”, *Illinois J. Math.* **7** (1963), 558–565. MR 28 #126 Zbl 0117.02701

[Polishchuk and Vaintrob 2012] A. Polishchuk and A. Vaintrob, “Chern characters and Hirzebruch–Riemann–Roch formula for matrix factorizations”, *Duke Math. J.* **161**:10 (2012), 1863–1926. MR 2954619 Zbl 1249.14001

[Vasconcelos 1968] W. V. Vasconcelos, “Reflexive modules over Gorenstein rings”, *Proc. Amer. Math. Soc.* **19** (1968), 1349–1355. MR 38 #5762 Zbl 0167.31201

[Walker 2014] M. E. Walker, “Chern characters for twisted matrix factorizations and the vanishing of the higher Herbrand difference”, preprint, 2014. arXiv 1404.0352

Received October 27, 2014. Revised December 17, 2014.

OLGUR CELIKBAS
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CONNECTICUT
STORRS, CT 06269
UNITED STATES
olgur.celikbas@uconn.edu

SRIKANTH B. IYENGAR
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
SALT LAKE CITY, UT 84112
UNITED STATES
iyengar@math.utah.edu

GREG PIEPMEYER
DEPARTMENT OF MATHEMATICS
COLUMBIA BASIN COLLEGE
PASCO, WA 99301
UNITED STATES
gpiepmeyer@columbiabasin.edu

ROGER WIEGAND
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NEBRASKA, LINCOLN
LINCOLN, NE 68588
UNITED STATES
rwiegand1@math.unl.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.117

CONVEX SOLUTIONS TO THE POWER-OF-MEAN CURVATURE FLOW

SHIBING CHEN

We prove some estimates for convex ancient solutions (the existence time for the solution starts at $-\infty$) to the power-of-mean curvature flow, when the power is strictly greater than $\frac{1}{2}$. As an application, we prove that in dimension two, the blow-down of an entire convex translating solution, namely $u_h = \frac{1}{h}u(h^{\frac{1}{1+\alpha}}x)$, locally uniformly converges to $\frac{1}{1+\alpha}|x|^{1+\alpha}$ as $h \to \infty$. Another application is that for the generalized curve shortening flow (convex curve evolving in its normal direction with speed equal to a power of its curvature), if the convex compact ancient solution sweeps the whole space $\mathbb{R}^2$, it must be a shrinking circle. Otherwise the solution must be defined in a strip region.

1. Introduction

Classifying ancient convex solutions to mean curvature flow is very important in studying the singularities of mean curvature flow. Translating solutions arise as a special case of ancient solutions when one uses a proper procedure to blow up the mean convex flow near type II singular points, and general ancient solutions arise at general singularities. Some important progress was made by Wang [2011], and Daskalopoulos, Hamilton and Sesum [Daskalopoulos et al. 2010]. Wang proved that in dimension $n = 2$, an entire convex translating solution to mean curvature flow must be rotationally symmetric in an appropriate coordinate system, which was a conjecture formulated explicitly by White [2000], but for $n \geq 3$ such solutions are not necessarily rotationally symmetric.

Wang also constructed some entire convex translating solutions with level sets neither spherical nor cylindrical in dimension greater or equal to 3. In the same paper, Wang also proved that if a convex ancient solution to the curve shortening flow sweeps the whole space $\mathbb{R}^2$, then it must be a shrinking circle — otherwise the convex ancient solution must be defined in a strip region, and he indeed constructed such solutions by a compactness argument. Daskalopoulos et al. [2010] showed that apart from the shrinking circle, the so called *Angenent oval* (a convex ancient

MSC2010: 35J60.
Keywords: mean curvature flow, convexity, translating solution, ancient solution.

solution of the curve shortening flow discovered by Angenent that decomposes into two translating solutions of the flow) is the only other embedded convex compact ancient solution of the curve shortening flow. That means that the corresponding curve shortening solution defined in a strip region constructed by Wang is exactly the "Angenent oval".

The power-of-mean curvature flow, in which a hypersurface evolves in its normal direction with speed equal to a power α of its mean curvature H, is well-studied [Andrews 1998; 2003; 2002; Schulze 2005; Chou and Zhu 2001; Sheng and Wu 2009]. Schulze [2005] called it H^α-flow. In the following, we will also call the one dimensional power-of-curvature flow the *generalized curve shortening flow*. It would be very interesting if one could classify the ancient convex solutions. In this paper, we use the method developed in [Wang 2011] to study the geometric asymptotic behavior of ancient convex solutions to H^α-flow. The general equation for H^α-flow is

$$\frac{\partial F}{\partial t} = -H^\alpha v,$$

where $F : M \times [0, T) \to \mathbb{R}^{n+1}$ is a time-dependent embedding of the evolving hypersurface, v is the unit normal vector to the hypersurface $F(M, t)$ in $\mathbb{R}^{n+1}$, and H is its mean curvature. If the evolving hypersurface can be represented as a graph of a function $u(x, t)$ over some domain in $\mathbb{R}^n$, then we can project the evolution equation to the $(n+1)$-st coordinate direction of $\mathbb{R}^{n+1}$ and the equation becomes

$$u_t = \sqrt{1+|Du|^2}\left(\operatorname{div}\frac{Du}{\sqrt{1+|Du|^2}}\right)^\alpha.$$

Then a translating solution to the H^α-flow will satisfy the equation

$$\sqrt{1+|Du|^2}\left(\operatorname{div}\frac{Du}{\sqrt{1+|Du|^2}}\right)^\alpha = 1,$$

which is equivalent to the special case $\sigma = 1$ of the following:

$$L_\sigma(u) = \left(\sqrt{\sigma+|Du|^2}\,\right)^{\frac{1}{\alpha}} \operatorname{div}\frac{Du}{\sqrt{\sigma+|Du|^2}} \tag{1}$$

$$= \left(\sigma+|Du|^2\right)^{\frac{1}{2\alpha}-\frac{1}{2}} \sum_{i,j=1}^{n}\left(\delta_{ij}-\frac{u_i u_j}{\sigma+|Du|^2}\right)u_{ij} \tag{2}$$

$$= 1, \tag{3}$$

where $\sigma \in [0, 1]$, $\alpha \in \left(\frac{1}{2}, \infty\right]$ is a constant, $n = 2$ is the dimension of $\mathbb{R}^2$. If u is a convex solution of (3), then $u + t$, as a function of $(x, t) \in \mathbb{R}^2 \times \mathbb{R}$, is a translating solution to the flow

$$u_t = \sqrt{\sigma + |Du|^2}\left(\operatorname{div} \frac{Du}{\sqrt{\sigma + |Du|^2}}\right)^{\alpha}. \tag{4}$$

When $\sigma = 1$, Equation (4) is the nonparametric power-of-mean curvature flow. When $\sigma = 0$, Equation (3) is the level set flow. That is, if u is a solution of (3) with $\sigma = 0$, then the level set $\{u = -t\}$, where $-\infty < t < -\inf u$, evolves by the power-of-mean curvature.

In the following we will assume $\sigma \in [0, 1]$, $\alpha \in \left(\frac{1}{2}, \infty\right]$, and the dimension $n = 2$, although some of the estimates do hold in higher dimension. The main results of this paper are the following theorems.

Theorem 1. *Let u be an entire convex solution of* (3). *Let*

$$u_h(x) = h^{-1} u(h^{\frac{1}{1+\alpha}} x).$$

Then, u_h locally uniformly converges to

$$\tfrac{1}{1+\alpha}|x|^{1+\alpha} \quad \text{as } h \to \infty.$$

Theorem 2. *Let u_σ be an entire convex solution of* (3). *Then,*

$$u_0(x) = \tfrac{1}{1+\alpha}|x|^{1+\alpha}$$

up to a translation of the coordinate system. When $\sigma \in (0, 1]$, if

$$|D^2 u(x)| = O(|x|^{\beta}) \quad \text{as } |x| \to \infty$$

for some fixed constant β satisfying $\beta < 3\alpha - 2$, then u_σ is rotationally symmetric after a proper translation of the coordinate system.

Corollary 3. *A convex compact ancient solution to the generalized curve shortening flow which sweeps the whole space $\mathbb{R}^2$ must be a shrinking circle.*

Remark 4. The condition $\alpha > \frac{1}{2}$ is necessary for our results. One can consider the translating solution $v(x)$ to (3) with $\sigma = 1$ in one dimension. In fact, when $\alpha \leq \frac{1}{2}$, the translating solution $v(x)$ is a convex function defined on the entire real line [Chou and Zhu 2001, p. 28]. Then one can construct a function $u(x, y) = v(x) - y$, defined on the entire plane, and u will satisfy (3) with $\sigma = 0$; it is obviously not rotationally symmetric. We can also let $u(x, y) = v(x)$, which is an entire solution to (3) with $\sigma = 1$, and it is not rotationally symmetric.

When the dimension is at least two, similar examples can be given: we can take an entire rotationally symmetric solution $v(x)$ to (3) with $n \geq 2$ and $\sigma = 1$, and again let $u(x, y) = v(x) - y$ (here, y is the $(n+1)$-st coordinate for $\mathbb{R}^{n+1}$). It is easy to see that u will satisfy (3) with n replaced by $n + 1$ and $\sigma = 0$, and the level set of u is neither a sphere nor a cylinder.

We would also like to point out that this elementary construction can be used to give a slight simplification of the proof of [Wang 2011, Theorem 2.1] (corresponding to our Corollary 10 for $\alpha = 1$). Let v_σ be an entire convex solution to (3) in dimension n with $\sigma \in (0, 1]$. Then $u(x, y) = v_\sigma(x) - \sqrt{\sigma}\, y$ will be an entire convex solution to (3) in dimension $n+1$ with $\sigma = 0$. Hence if one has proved the estimate in Corollary 10 for $\sigma = 0$ in all dimensions, the estimate for $\sigma \in (0, 1]$ follows immediately from the above construction. The remainder of the paper is divided into four sections. Sections 2 and 3 contain the proof of Theorem 1 and the first part of Theorem 2. Section 4 is devoted to the proof of Corollary 10, and the last section completes the proof of Theorem 2.

2. Power growth estimate

In this section, we prove a key estimate, which says that any entire convex solution u to (3) must satisfy

$$u(x) \le C(1 + |x|^{1+\alpha}),$$

where the constant C depends only on the upper bound of $u(0)$ and $|Du(0)|$. When $\alpha = 1$, the estimate was proved by Wang [2011, Theorem 2.1]. To apply Wang's method, the main difficulty is that now the speed function is nonlinear in the curvature. We overcome this difficulty by further exploiting some elementary convexity properties.

For any constant $h > 0$, we denote

$$\Gamma_h = \{x \in \mathbb{R}^n : u(x) = h\},$$
$$\Omega_h = \{x \in \mathbb{R}^n : u(x) < h\},$$

so that Γ_h is the boundary of Ω_h. Let κ be the curvature of the level curve Γ_h. We have

$$L_\sigma(u) = \left(\sigma + u_\gamma^2\right)^{\frac{1}{2\alpha} - \frac{1}{2}} \left(\kappa u_\gamma + \frac{\sigma u_{\gamma\gamma}}{\sigma + u_\gamma^2}\right) \tag{5}$$

$$\ge \kappa u_\gamma^{\frac{1}{\alpha}} = L_0(u), \tag{6}$$

where γ is the unit outward normal to Ω_h, and $u_{\gamma\gamma} = \gamma_i \gamma_j u_{ij}$.

Before starting the proof of our main results, we recall a well known convergence result for the generalized curve shortening flow.

Lemma 5 [Andrews 2003, Theorems 1.3, 1.4, 1.5]. *Let ℓ_t be a time-dependent family of closed curves in $\mathbb{R}^2$ evolving under the generalized curve shortening flow with $\alpha > \frac{1}{3}$. Suppose the initial curve ℓ_0 is convex. Then the curve converges in finite time T to a round point P in the sense that $((1+\alpha)(T-t))^{-\frac{1}{1+\alpha}}(\ell_t - P)$ is asymptotic to the unit circle.*

Next, we prove a lemma which will be used to control the shape of the level set of a complete convex solution to (3).

Lemma 6. *Let u be a complete convex solution of* (3). *Suppose that* $u(0) = 0$ *and that the infimum* $\inf\{|x| : x \in \Gamma_1\}$ *is attained at* $x_0 = (0, -\delta) \in \Gamma_1$ *for some* $\delta > 0$ *sufficiently small. Let D_1 be the projection of* Γ_1 *onto the axis* $\{x_2 = 0\}$. *Then, D_1 contains the interval* $(-R, R)$, *and when* $\alpha \leq 1$, *R satisfies*

$$R \geq C_1(-\log\delta - C_2)^{\frac{\alpha}{\alpha+1}}, \tag{7}$$

where $C_1, C_2 > 0$ *are independent of* δ; *when* $\alpha > 1$, $R \geq C$ *for some positive constant C.*

The proof of this lemma follows that of [Wang 2011, Lemma 2.4] with minor modifications; the for reader's convenience, we give some details here.

Proof. First, we prove the lemma when $\frac{1}{2} < \alpha \leq 1$. Suppose that near x_0, Γ_1 is given by $x_2 = g(x_1)$. Then, g is a convex function satisfying $g(0) = -\delta$ and $g'(0) = 0$. Let $b > 0$ be a constant such that $g'(b) = 1$. To prove (7), it suffices to prove

$$b \geq C_1(-\log\delta - C_2)^{\frac{\alpha}{\alpha+1}}. \tag{8}$$

For any $y = (y_1, y_2) \in \Gamma_1$, where $y_1 \in [0, b]$, as in the proof of [Wang 2011, Lemma 2.4] we have

$$u_\gamma(y) \geq \frac{\sqrt{1+g'^2}}{y_1 g' - y_2}, \tag{9}$$

where γ is the unit normal of the sublevel set Ω_1. Since $L_0 u \leq 1$, we have

$$\frac{g''}{(1+g'^2)^{\frac{3}{2}}}\,\frac{(1+g'^2)^{\frac{1}{2\alpha}}}{(y_1 g' - y^2)^{\frac{1}{\alpha}}} \leq \kappa u_\gamma^{\frac{1}{\alpha}} \leq 1, \tag{10}$$

where κ is the curvature of the level curve Γ_1. Hence,

$$g''(y_1) \leq (1+g'^2)^{\frac{3}{2}-\frac{1}{2\alpha}}(y_1 g' - y_2)^{\frac{1}{\alpha}} \tag{11}$$

$$\leq 10 y_1^{\frac{1}{\alpha}} g' + 10\delta \tag{12}$$

where $y_2 = g(y_1)$ and $g'(y_1) \leq 1$ for $y_1 \in (0, b)$. The inequality from (11) to (12) is trivial when $y_2 \geq 0$. When $y_2 \leq 0$, since $|y_2| \leq \delta$, we have either $y_1 g' \leq \delta$ or $y_1 g' > \delta$. For the former we have

$$(y_1 g' - y_2)^{\frac{1}{\alpha}} \leq (2\delta)^{\frac{1}{\alpha}} \leq 4\delta;$$

for the latter, since $g'(y_1) \leq 1$, we have

$$(y_1 g' - y_2)^{\frac{1}{\alpha}} \leq (2y_1 g')^{\frac{1}{\alpha}} \leq 4 y_1^{\frac{1}{\alpha}} g'.$$

We consider the equation

$$\rho''(t) = 10t^{\frac{1}{\alpha}}\rho' + 10\delta \tag{13}$$

with initial conditions $\rho(0) = -\delta$ and $\rho'(0) = 0$. Then for $t \in (0, b)$, we have

$$\rho'(t) = 10\delta e^{\frac{10\alpha}{\alpha+1}t^{\frac{\alpha+1}{\alpha}}}\int_0^t e^{-\frac{10\alpha}{\alpha+1}s^{\frac{\alpha+1}{\alpha}}}\,ds. \tag{14}$$

Since $\int_0^\infty e^{-\frac{10\alpha}{\alpha+1}s^{\frac{\alpha+1}{\alpha}}}\,ds$ is bounded above by some constant C, we have

$$1 \le \rho'(b) = 10\delta e^{\frac{10\alpha}{\alpha+1}b^{\frac{\alpha+1}{\alpha}}}\int_0^b e^{-\frac{10\alpha}{\alpha+1}s^{\frac{\alpha+1}{\alpha}}}\,ds \tag{15}$$

$$\le C_1\delta e^{\frac{10\alpha}{\alpha+1}b^{\frac{\alpha+1}{\alpha}}}, \tag{16}$$

from which (8) follows.

When $\alpha > 1$, the situation is different. First, we introduce a number a such that $g'(a) = \frac{1}{2}$. Then, we can follow the proof above until (11). For (12) the inequality becomes

$$g''(y_1) \le 10y_1^{\frac{1}{\alpha}}g' + 10\delta^{\frac{1}{\alpha}},$$

for $y_1 \in [a, b]$. Now (16) becomes

$$e^{-\frac{10\alpha}{\alpha+1}b^{\frac{\alpha+1}{\alpha}}}\rho'(b) - e^{-\frac{10\alpha}{\alpha+1}a^{\frac{\alpha+1}{\alpha}}}\rho'(a) \le C_1\delta^{\frac{1}{\alpha}}.$$

Then, it is easy to see that when δ is small, $b \ge C$, for some fixed constant C. $\square$

Remark 7. When $\alpha \le 1$, it follows from Lemma 6 that when δ is sufficiently small, by convexity and in view of Figure 1, we have that Ω_1 contains the shadowed region. Then it is easy to check that Ω_1 contains an ellipse

$$E = \left\{(x_1, x_2)\;\middle|\; \frac{x_1^2}{\left(\frac{R}{6}\right)^2} + \frac{\left(x_2 - \frac{7\delta^* - 5\delta}{12}\right)^2}{\left(\frac{\delta^*+\delta}{4}\right)^2} = 1\right\}, \tag{17}$$

where δ^* is a positive constant such that $u(0, \delta^*) = 1$ and R is defined in Lemma 5.

When $\alpha > 1$, if δ^* is very large, in the part $\{x : u(x) \le 1,\ x_1 \ge 0\}$, by convexity we can find an ellipse which has the length of short axis bounded from below and the length of long axis very large, and if we let the ellipse evolve under the generalized curve shortening flow, it will take time more than 1 for it to converge to a round point. When δ^* is less than some fixed constant, we need to consider two cases.

Case 1: The set $\{u \le 1\}$ is not compact. In this case when we project $\{u(x) = 1\}$ to the axis $\{x_2 = 0\}$, and denote the leftmost and rightmost points as $(-l, 0)$ and $(r, 0)$, respectively. Then either l or r is very large, which guarantees that one can

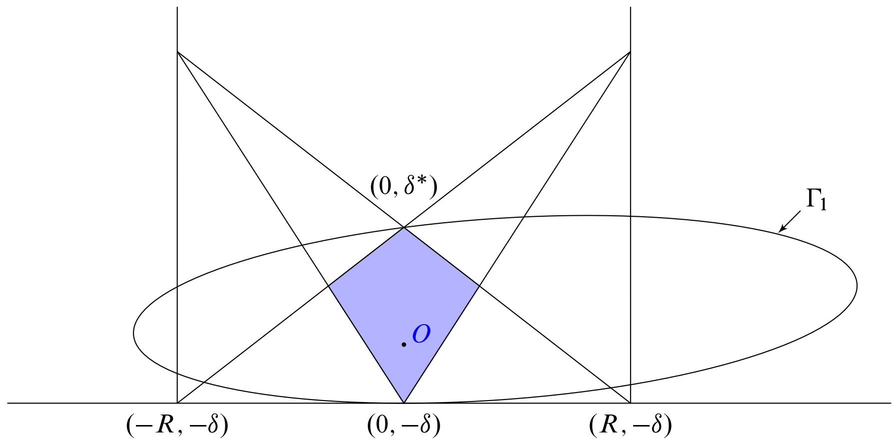

Figure 1. Γ_1 contains the shadow part.

still find an ellipse inside $\{x : u(x) \le 1,\ x_1 \le 0\}$ (or $\{x : u(x) \le 1,\ x_1 \ge 0\}$) with the similar property as before.

Case 2: The set $\{u \le 1\}$ is compact. For this case, we will always assume 0 is the minimum point of u, and $u(0) = 0$. We claim that when δ is very small, for the purpose of the proof of Corollary 10, we can assume one of l or r is very large. Indeed, if the claim is not true, then we have a sequence of functions u_i such that $\{u_i \le 1\}$ has width bounded by some constant independent of i, and $\text{dist}(0, \{u_i \le 1\}) \to 0$ as $i \to \infty$. In view of the following proof of Corollary 10, we can assume u_i satisfies (3) with $\sigma_i \to 0$. Then by passing to a subsequence, we can assume that $\{u_i = 1\}$ converges to a convex curve C_0 in hausdorff distance. Let C_0 evolve under the generalized curve shortening flow; by Lemma 5, it will converge to a point P, but by the above discussion we see that P is on C_0, which is clearly impossible. Once l or r is very large, we can find an ellipse with the similar property as in the case 1.

Remark 8. One can also establish a similar lemma in higher dimensions, which says that D_1 (a convex set with dimension greater than 1) contains a ball centered at the origin with radius

$$R \ge C_n(-\log\delta - C)^{\frac{\alpha}{\alpha+1}},$$

where C_n is a constant depending only on n and C is a positive constant independent of δ. The proof can be reduced to the two dimensional case; for the details, refer to the proof of [Wang 2011, Lemma 2.6].

Lemma 9. *Let u be a complete convex solution of* (3). *Suppose $u(0) = 0$, δ and δ^* are defined as in Lemma 6 and Remark 7. If δ and δ^* are sufficiently small, then u is defined in a strip region.*

When $\alpha = 1$, this lemma is proved by Wang [2011, Corollary 2.2]. The proof of Lemma 9 is based on a careful study of the shape of the level set of u. Before giving the proof, we will give an important corollary first.

Corollary 10. *Let u be an entire convex solution of* (3) *in* $\mathbb{R}^2$, *then*

$$u(x) \le C(1 + |x|^{1+\alpha}), \tag{18}$$

where the constant C depends only on the upper bound for $u(0)$ and $|Du(0)|$.

Proof. The proof of this Corollary follows the proof of [Wang 2011, Theorem 2.1]. We only record some necessary changes here. First, the rescaling $u_h(x) = \frac{1}{h}u(h^{\frac{1}{2}}x)$ used that proof should be replaced by $u_h(x) = \frac{1}{h}u(h^{\frac{1}{1+\alpha}}x)$. Note that u_h solves (3) with $\sigma = \sigma_h \to 0$ as $h \to \infty$. Second, the ellipse used in that proof when applying the comparison argument should be replaced by the one discussed in Remark 7. □

Proof of Lemma 9. By a rotation of coordinates we may assume that the axial directions of E in Remark 7 coincide with those of the coordinate system. Let $\mathcal{M}_u$ be the graph of u, which consists of two parts, $\mathcal{M}_u = \mathcal{M}^+ \cup \mathcal{M}^-$, where

$$\mathcal{M}^+ = \{(x, u(x)) \in \mathbb{R}^3 : \partial_{x_2}u \ge 0\} \quad \text{and} \quad \mathcal{M}^- = \{(x, u(x)) \in \mathbb{R}^3 : \partial_{x_2}u \le 0\}.$$

Then $\mathcal{M}^\pm$ can be represented as the graphs of functions $g^\pm$ of the form $x_2 = g^\pm(x_1, x_3)$, for $(x_1, x_3) \in D$ where D is the projection of $\mathcal{M}_u$ onto the plane $\{x_2 = 0\}$. The functions g^+ and g^- are concave and convex, respectively, and we have $x_3 = u(x_1, g^\pm(x_1, x_3))$. Set

$$g = g^+ - g^-. \tag{19}$$

Then g is a positive, concave function on D, vanishing on ∂D. For any $h > 0$ let $g_h(x_1) = g(x_1, h)$, $g_h^\pm(x_1) = g^\pm(x_1, h)$, and $D_h = \{x_1 \in \mathbb{R}^1 : (x_1, h) \in D\}$. Then g_h is a positive, concave function in D_h, vanishing on ∂D_h, and $D_h = (-\underline{a}_h, \bar{a}_h)$ is an interval containing the origin. Let $b_h = g_h(0)$. We consider the case $\sigma = 0$ first.

<u>Claim 1</u>: Suppose h is large, $g_1(0) = \delta^* + \delta$ is small, $b_h \le 4$, and $\underline{a}_h, \bar{a}_h \ge b_h$. Then,

$$\bar{a}_h \ge \frac{1}{1000}\frac{h}{b_h^\alpha} \text{ for } \alpha \le 1 \quad \text{and} \quad \bar{a}_h \ge \frac{1}{1000}\frac{h^{1/(2\alpha-1)}}{b_h^{1/(2\alpha-1)}} \text{ for } \alpha > 1.$$

Proof. Without loss of generality, we assume $\bar{a}_h \le \underline{a}_h$. Let $U_h = \Omega_h \cap \{x_1 > 0\}$. By the convexity of U_h and the assumption $\underline{a}_h, \bar{a}_h \ge b_h$, we have $\underline{a}_s, \bar{a}_s \ge \frac{1}{2}b_h$ for all $s \in \left(\frac{1}{2}h, h\right)$. Hence by the concavity of g,

$$\left|\frac{d}{dx_1}g_s(0)\right| \le 2 \quad \text{for } s \in \left(\tfrac{1}{2}h, h\right),$$

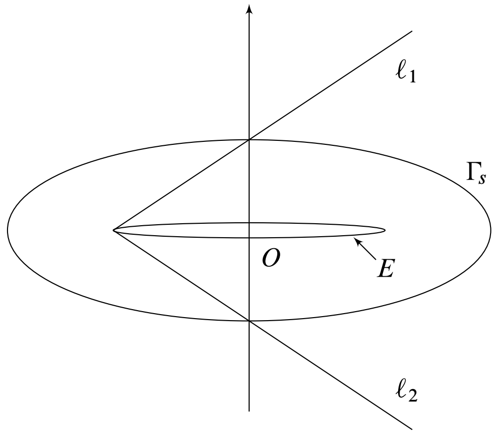

Figure 2. $\Gamma_s \cap \{x_1 > 0\}$ is trapped between two lines.

which means the arclength of the image of $\Gamma_s \cap \{x_1 > 0\}$ under the Gauss map is bigger then $\frac{\pi}{6}$. Notice that Ω_1 contains E, which was defined in Remark 7. When δ and δ^* are very small, E is very thin and long. The center of E is very close to the origin; in fact, for our purpose we can just pretend E is centered at the origin. By convexity of Ω_h and in view of Figure 2, we see that $\Gamma_s \cap \{x_1 > 0\}$ is trapped between two lines ℓ_1 and ℓ_2, and the slopes of ℓ_1 and ℓ_2 are very close to 0 when E is very long and thin. Then it is clear that the largest distance from the points on $\Gamma_s \cap \{x_1 > 0\}$ to the origin can not be bigger than $10\bar{a}_h$. By convexity of u,

$$u_\gamma(x) \geq \frac{h}{20\bar{a}_h} \quad \text{for } x \in \Gamma_s \cap \{x_1 > 0\}.$$

Since $\Gamma_s \cap \{x_1 > 0\}$ evolves under the generalized curve shortening flow, when $\alpha \leq 1$ we have the estimate

$$\frac{d}{ds}|U_s| = \int_{\Gamma_s \cap \{x_1 > 0\}} \kappa^\alpha \, d\xi \tag{20}$$

$$= \int_{\Gamma_s \cap \{x_1 > 0\}} u_\gamma^{\frac{1}{\alpha}-1} \kappa \, d\xi \tag{21}$$

$$\geq \frac{1}{50}\Big(\frac{h}{\bar{a}_h}\Big)^{\frac{1}{\alpha}-1} \frac{\pi}{6}, \tag{22}$$

where from (20) to (21) we used the equation $\kappa u_\gamma^{\frac{1}{\alpha}} = 1$. The claim follows by the simple fact that

$$\frac{3}{2} b_h \bar{a}_h \geq |U_h| \geq \frac{1}{50}\Big(\frac{h}{\bar{a}_h}\Big)^{\frac{1}{\alpha}-1} \frac{\pi}{6}\frac{h}{2}.$$

When $\alpha > 1$, let l_s denote the arclength of $\Gamma_s \cap \{x_1 > 0\}$. Then, by the above discussion, it is not hard to see that $l_s \approx C\bar{a}_h$. Then by a simple application of

Jensen's inequality,

$$\begin{aligned}\frac{d}{ds}(|U_s|) &= \int_{\Gamma_s\cap\{x_1>0\}} \kappa^\alpha \, d\xi \\ &= l_s \int_{\Gamma_s\cap\{x_1>0\}} \kappa^\alpha \frac{1}{l_s} \, d\xi \\ &\geq l_s \Big(\int_{\Gamma_s\cap\{x_1>0\}} \frac{\kappa}{l_s} \, d\xi \Big)^\alpha \geq C l_s^{1-\alpha} \geq C \bar a_h^{1-\alpha}.\end{aligned}$$

Again by the simple fact that $\frac{3}{2} b_h \bar a_h \geq |U_h|$, we can complete the proof in the same way as in the previous case. □

From here until (55) we will prove the lemma for the case $\frac{1}{2} < \alpha \leq 1$, and then we will give the details for the case $\alpha > 1$.

Claim 2: Let $h_k = 2^k$, $\bar a_k = \bar a_{h_k}$, $b_k = b_{h_k}$, $g_k = g_{h_k}$, and $D_k = D_{h_k}$. Then,

$$g_k(0) \leq g_{k-1}(0) + C_0 2^{-k/C} \quad \text{for all } k \text{ large,} \tag{23}$$

where C_0 is a fixed constant and C depends only on α.

Lemma 9 follows from Claims 1 and 2 in the following way. Let the convex set P be the projection of the graph of g onto the plane $\{x_3 = 0\}$. By Claim 2 and the fact that P contains x_1-axis (it follows from Claim 1), P must equal $I \times \mathbb{R}$ for some interval

$$I \subset \big[0, \lim_{k\to\infty} g_k(0)\big].$$

Then, by (19), $\mathcal{M}_u$ is also contained in a strip region as stated in Lemma 9.

Proof of Claim 2. To prove (23), observe that since g is positive and concave,

$$g_k(0) \leq h_k g_0(0) \leq 2^k(\delta + \delta^*).$$

Hence, we can start from sufficiently large k_0, satisfying $g_{k_0}(0) \leq 1$ and

$$g_{k_0} + C_0 \sum_{j=k_0}^{\infty} 2^{-j/C} \leq 2. \tag{24}$$

Suppose (23) holds up to k. Then by (24), we have $g_k(0) \leq 2$. By the concavity of g and and the fact that $g \geq 0$, we have $g_{k+1}(0) \leq 2g_k(0) \leq 4$. By Claim 1, we have $\bar a_{k+1} \geq \frac{1}{10000} h_k$. To prove (23) at $k+1$,

$$L_k = \Big\{ x_1 \in \mathbb{R}^1 : -\frac{C_1}{4} h_k < x_1 < \frac{C_1}{4} h_k \Big\}, \quad C_1 = \frac{1}{10000}, \tag{25}$$

$$Q_k = L_k \times [h_k, h_{k+1}] \subset D. \tag{26}$$

Since $g > 0$ and g is concave, we have the estimates

$$g(x_1, h) \le 8, \tag{27}$$

$$|\partial_h g(x_1, h)| \le \frac{16}{h_k}, \tag{28}$$

$$|\partial_{x_1} g(x_1, h)| \le \frac{16}{h_k} \quad \text{for all } (x_1, h) \in Q_k. \tag{29}$$

Let $\mathscr{X}^\pm = \{(x_1, h) \in Q_k : |\partial_{x_1x_1} g^\pm(x_1, h)| \ge h_k^{-\beta}\}$, where β is chosen such that $\frac{1}{\alpha} < \beta < 2$. For any $h \in (h_k, h_{k+1})$, by (29),

$$|\{x_1 \in L_k : (x_1, h) \in \mathscr{X}^+\}|\, h_k^{-\beta} \le \int_{L_k} \partial_{x_1x_1} g^+ \tag{30}$$

$$\le \int_{L_k} \partial_{x_1x_1} g \tag{31}$$

$$\le 2 \sup_{L_k} |\partial_{x_1} g| \tag{32}$$

$$\le \frac{C}{h_k}. \tag{33}$$

So, $|\mathscr{X}^+| \le C h_k^\beta$. Similarly, we have $|\mathscr{X}^-| \le C h_k^\beta$.

For any given $y_1 \in L_k$, let $\mathscr{X}^\pm_{y_1} = \mathscr{X}^+ \cap \{x_1 = y_1\}$. Then, by the estimate above, there is a set $\widetilde{L}^\pm \subset L_k$ with measure

$$|\widetilde{L}^\pm| \le C h_k^{\beta/2},$$

such that for any $y_1 \in L_k - \widetilde{L}^\pm$, we have $|\mathscr{X}^\pm_{y_1}| \le h_k^\beta / 2$. When k is large, we can always find $y_1 = C h^{\beta/2} \in L_k - \widetilde{L}^\pm$, where the constant C is under control. For such y_1, we have

$$\begin{aligned} g(y_1, h_{k+1}) - g(y_1, h_k) \\ = g^+(y_1, h_{k+1}) - g^+(y_1, h_k) + |g^-(y_1, h_{k+1}) - g^-(y_1, h_k)|. \end{aligned} \tag{34}$$

In the following, we will estimate $g^+(y_1, h_{k+1}) - g^+(y_1, h_k)$. The estimate for $|g^-(y_1, h_{k+1}) - g^-(y_1, h_k)|$ is analogous. By the same reason as that for [Wang 2011, §2.21], we have

$$\begin{cases} (\partial_h g^+)^{-1} = (1 + \varepsilon_1) u_\gamma, \\ \partial_{x_1x_1} g^+ = (1 + \varepsilon_2) \kappa. \end{cases} \tag{35}$$

Then, by the equation $u_\gamma^{\frac{1}{\alpha}} \kappa = 1$, we have

$$\partial_h g_h^+(y_1, h) \le C(\partial_{x_1x_1} g^+)^\alpha \le C h_k^{-\beta\alpha}. \tag{36}$$

Now,

$$g^+(y_1,h_{k+1})-g^+(y_1,h_k)=\int_{h_k}^{h_{k+1}}\partial_h g^+(y_1,h)\,dh \tag{37}$$

$$=\int_{\mathscr{X}^+_{y_1}}\partial_h g^+(y_1,h)\,dh+\int_{[h_k,h_{k+1}]-\mathscr{X}^+_{y_1}} dh \tag{38}$$

$$\le C_1 h_k^{\frac{\beta}{2}}\frac{1}{h_k}+C_2 h_k^{-\beta\alpha}h_k. \tag{39}$$

Recall that β satisfies $\frac{1}{\alpha}<\beta<2$, and we have $\eta:=\min\{1-\frac{\beta}{2},\beta\alpha-1\}>0$. From (34) and (39), we have the estimate

$$g(y_1,h_{k+1})-g(y_1,h_k)\le\frac{C}{h_k^\eta},$$

for some fixed constant C. Then, we will assume $\partial_{x_1}g(0,h_k)<0$ (otherwise we can replace x_1 by $-x_1$); therefore, by the above estimate,

$$g(y_1,h_{k+1})\le g(y_1,h_k)+\frac{C}{h_k^\eta}\le g(0,h_k)+\frac{C}{h_k^\eta}.$$

Since g is positive, concave, and defined on $[0,\bar{a}_{k+1}]$, with $\bar{a}_{k+1}\ge Ch_{k+1}$,

$$\frac{g_{k+1}(0)}{g_{k+1}(y_1)}\le\frac{\bar{a}_{k+1}}{\bar{a}_{k+1}-y_1}\le 1+Ch_{k+1}^{\frac{\beta}{2}-1}.$$

Therefore, by the two estimates above,

$$g_{k+1}(0)\le g_k(0)+Ch_k^{-\eta},$$

which implies (23) immediately. □

For the proof of Lemma 6 when $\sigma\in(0,1]$, we need to use (5) and (6). In fact, by (6) we see that Γ_h is moving at a velocity greater than or equal to its curvature to the power α. Hence, we still have the lower bound of $\frac{d}{ds}(|U_s|)$ as in the proof of Claim 1. Then we can follow the above proof for the case $\sigma=0$ until (37), replacing the equalities "$=$" in (20) and (21) with inequalities "$\ge$". As in [Wang 2011], when $\sigma=0$, in order to control the second integral in (38) we used the equation $\kappa u_\gamma^{1/\alpha}=1$. But when $\sigma\ne 0$, by (28) and (35) we have

$$u_\gamma\ge C(\partial_h g^+)^{-1}\ge Ch_k. \tag{40}$$

Hence, we may assume that u_γ is as large as we want, which means that in formula (5), the only important extra term is

$$(\sigma+u_\gamma^2)^{\frac{1}{2\alpha}-\frac{1}{2}}\frac{\sigma u_{\gamma\gamma}}{\sigma+u_\gamma^2}.$$

To handle this term, we divide the integral (39) into three parts:

$$g^+(y_1,h_{k+1})-g+(y_1,h_k)=\int_{h_k}^{h_{k+1}}\partial_h g^+(y_1,h)\,dh \tag{41}$$

$$=\left(\int_{I_1}+\int_{I_2}+\int_{I_3}\right)\partial_h g^+(y_1,h)\,dh, \tag{42}$$

where

$$I_1=\mathscr{X}^+_{y_1}, \tag{43}$$

$$I_2=\left\{h\in[h_k,h_{k+1}]-I_1\,:\,(\sigma+u_\gamma^2)^{\frac{1}{2\alpha}-\frac{1}{2}}\frac{\sigma u_{\gamma\gamma}}{\sigma+u_\gamma^2}\le\frac{1}{2}\right\}, \tag{44}$$

$$I_3=[h_k,h_{k+1}]-I_1\cup I_2. \tag{45}$$

For the first integral, we can do exactly the same thing as we have done from (38) to (39), namely,

$$\int_{I_1}\partial_h g^+(y_1,h)\,dh\le\frac{C}{h_k}h_k^{\frac{\beta}{2}}=Ch_k^{\frac{\beta}{2}-1}.$$

Note that the power $\frac{\beta}{2}-1$ is a negative number.

Then we estimate the second integral, note that when $(y_1,h)\in I_2$, we have

$$(\sigma+u_\gamma^2)^{\frac{1}{2\alpha}}\kappa u_\gamma\ge\tfrac{1}{2}.$$

By (40) u_γ is large, so we have $\kappa u_\gamma^{\frac{1}{\alpha}}\ge\frac{1}{4}$, hence by (35) we have

$$\partial_h g^+\le C(\partial_{x_1x_1}g^+)^\alpha\le Ch_k^{-\alpha\beta}. \tag{46}$$

Therefore,

$$\int_{I_2}\partial_h g^+(y_1,h)\,dh\le Ch_k^{-\beta\alpha}h_k=Ch_k^{1-\beta\alpha}.$$

Note that $1-\beta\alpha$ is a negative number. Observe that we can assume I_2 is on the right hand side of I_3, since by the concavity of g^+ we know that when $h\ge\inf I_2$, $\partial_h g^+(y_1,h)$ will satisfy the estimate (46).

For the third integral, notice that by the same argument as that for [Wang 2011, §2.24],

$$\begin{cases}u_\gamma(y_1,h)=u_{x_2}(1+\varepsilon_1),\\ u_{\gamma\gamma}(y_1,h)=u_{x_2x_2}(1+\varepsilon_2)+\varepsilon_3u_{x_2}.\end{cases} \tag{47}$$

Hence, by (47),

$$(\sigma+u_{x_2}^2)^{\frac{1}{2\alpha}-\frac{1}{2}}\frac{\sigma u_{x_2x_2}}{\sigma+u_\gamma^2}\ge\tfrac{1}{3}. \tag{48}$$

Since $\sigma \in [0, 1]$ and u_γ is large, we have

$$u'' = u_{x_2 x_2} \geq \tfrac{1}{4}(u')^{3-\frac{1}{\alpha}}. \tag{49}$$

By differentiating the equation $u(x_1, g^+(x_1, h)) = h$ twice with respect to h,

$$(g^+)'' = -u''(g^+)'^3 \leq -\tfrac{1}{4}(g^+)'^{\frac{1}{\alpha}-3}(g^+)'^3 = -\tfrac{1}{4}(g^+)'^{\frac{1}{\alpha}}. \tag{50}$$

Note that (50) is for points with corresponding $h \in I_3$. By the discussion after (46) we only need to estimate

$$\int_{[h_k + h_k^{\frac{\beta+2}{4}}, \inf I_2]} (g^+)'\, dh.$$

Therefore, by (50) and noticing that $(g^+)' \geq 0$,

$$\tfrac{\alpha}{\alpha-1}(g^+)'^{\frac{\alpha-1}{\alpha}}(h) \leq \tfrac{\alpha}{\alpha-1}(g^+)'^{\frac{\alpha-1}{\alpha}}(h_k) - \tfrac{1}{4}\big|I_3 \cap [h_k, h]\big|. \tag{51}$$

Hence, when $h \in [h_k^{(\beta+2)/4}, \inf I_2]$,

$$(g^+)'(h) \leq \big((g^+)'^{\frac{\alpha-1}{\alpha}}(h_k) + C(h - h_k)\big)^{\frac{\alpha}{\alpha-1}}. \tag{52}$$

Finally,

$$\int_{[h_k + h_k^{\frac{\beta+2}{4}}, \inf I_2]} (g^+)'\, dh \leq \int_{h_k}^{h_{k+1}} \big((g^+)'^{\frac{\alpha-1}{\alpha}}(h_k) + C(h-h_k)\big)^{\frac{\alpha}{\alpha-1}}\, dh \tag{53}$$

$$\leq \tfrac{\alpha-1}{2\alpha-1}\big((g^+)'^{\frac{\alpha-1}{\alpha}}(h_k) + C(h-h_k)\big)^{\frac{\alpha}{\alpha-1}+1}\Big|_{h_k}^{2h_k} \tag{54}$$

$$\leq C(g^+)'^{\frac{2\alpha-1}{\alpha}} \leq C h_k^{\frac{1-2\alpha}{\alpha}}. \tag{55}$$

Note that $\frac{1-2\alpha}{\alpha} < 0$ when $\alpha > \frac{1}{2}$, so we can complete the proof as in the $\sigma = 0$ case.

When $\alpha > 1$, we need to choose the constants and exponents more carefully. First of all, in view of the Lemma 9 for $\alpha > 1$, in order to have properties (35) and (47), we need only to replace the number 2 in (24) with some number much smaller than the constant C in Lemma 9. The definition of L_k in (25) should be modified to

$$L_k = \Big\{x_1 \in \mathbb{R}^1 : -\frac{C_1}{4} h_k^{\frac{1}{2\alpha-1}} < x_1 < \frac{C_1}{4} h_k^{\frac{1}{2\alpha-1}}\Big\}, \quad C_1 = \frac{1}{10000},$$

and the definition of Q_k in (26) remains the same. It is easy to see that we still have the estimates (27)–(28), but (29) becomes

$$|\partial_{x_1} g(x_1, h)| \leq 16 h_k^{-\frac{1}{2\alpha-1}} \quad \text{for all } (x_1, h) \in Q_k.$$

Then for the definition of

$$\mathscr{X}^\pm = \{(x_1, h) \in Q_k : |\partial_{x_1 x_1} g^\pm(x_1, h)| \geq h_k^{-\beta}\},$$

we need to choose the exponent β so that $\frac{1}{\alpha} < \beta < \frac{2}{2\alpha-1}$. By doing the same computation as (30)–(32),

$$\big|\{x_1 \in L_k : (x_1, h) \in \mathscr{X}^+\}\big| h_k^{-\beta} \le \int_{L_k} \partial_{x_1 x_1} g^+ \le C h_k^{-\frac{1}{2\alpha-1}}.$$

Hence,

$$|\mathscr{X}^+| \le C h_k^{1+\beta-\frac{1}{2\alpha-1}} \quad \text{and} \quad |\mathscr{X}^-| \le C h_k^{1+\beta-\frac{1}{2\alpha-1}}.$$

Then, by the above estimate there is a set $\widetilde{L}^{\pm} \subset L_k$ with measure

$$|\widetilde{L}^{\pm}| \le C h_k^{\beta+\varepsilon-\frac{1}{2\alpha-1}}$$

such that for any $y_1 \in L_k - \widetilde{L}^{\pm}$, we have $|\mathscr{X}^{\pm}_{y_1}| \le h_k^{1-\varepsilon}$, where ε is chosen such that $\beta + \varepsilon < \frac{2}{2\alpha-1}$. Now, (35)–(38) remain the same, and (39) becomes

$$g^+(y_1, h_{k+1}) - g^+(y_1, h_k) \le C_1 h_k^{1-\varepsilon} \frac{1}{h_k} + C_2 h_k^{-\beta\alpha} h_k.$$

By the choice of β, all the exponents of h_k are negative. We do not need to change anything from (40) to (49). Finally from (50) we need to replace the computation in the case $\alpha \le 1$ with the following computation.

First, we have $(g^+)'' \le -\frac{1}{4}(g^+)'^{\frac{1}{\alpha}} \le -\frac{1}{4}(g^+)'$, and we only need to bound

$$\int_{[h_k + h_k^{1-\varepsilon/2}, \inf I_2]} (g^+)' \, dh.$$

Note that $(g^+)' \ge 0$. By integrating the above differential inequality, we have

$$(g^+)'(h) \le (g^+)'(h_k) e^{-\frac{1}{4}|I_3|} \le (g^+)'(h_k) e^{\frac{1}{8}(h-h_k)}$$

when $h \in [h_k + h_k^{1-\varepsilon/2}, \inf I_2]$. Therefore, we have

$$\begin{aligned} \int_{[h_k + h_k^{1-\varepsilon/2}, \inf I_2]} (g^+)' \, dh &\le \int_{h_k}^{h_{k+1}} (g^+)'(h_k) e^{\frac{1}{8}(h-h_k)} \, dh \\ &\le C (g^+)'(h_k) \le \frac{C}{h_k}. \end{aligned}$$ □

3. Blow-down of an entire convex ancient solutions converges to a power function

In this section we prove that the blow-down of an entire convex solution to (3) converges to a power function.

Proof of Theorem 1 and the first part of Theorem 2. First, we prove that there is a subsequence of u_h converging to $\frac{1}{1+\alpha}|x|^{1+\alpha}$, where $u_h(x) = h^{-1} u(h^{\frac{1}{1+\alpha}} x)$.

By adding a constant we may suppose $u(0)=0$. Let $x_{n+1}=a\cdot x$ be the equation of the tangent plane of u at 0. By Corollary 10 and the convexity of u we have

$$a\cdot x\le u(x)\le C(1+|x|^{1+\alpha}).$$

Hence,

$$h^{-\frac{\alpha}{1+\alpha}}a\cdot x\le u_h(x)\le C\big(\tfrac{1}{h}+|x|^{1+\alpha}\big).$$

By convexity, Du_h is locally uniformly bounded. Hence, u_h subconverges to a convex function u_0 which satisfies $u_0(0)=0$, and

$$0\le u_0(x)\le C|x|^{1+\alpha}.$$

It is easy to check that u_0 is an entire convex viscosity solution to (3) with $\sigma=0$, and the comparison principle holds on any bounded domain.

Now we will prove that $\{u_0(x)=0\}=\{0\}$. In fact, if $\{x:u_0(x)=0\}$ is a bounded set, then $\{u_0(x)=h\}$ is a closed, bounded convex curve which evolves under the generalized curve shortening flow; from [Andrews 1998] it follows that $\{u_0(x)=0\}=\{0\}$. If $\{u_0(x)=0\}$ contains a straight line, say the line $\{(t,0):t\in\mathbb{R}\}$, then by convexity, u is independent of x_1, which is impossible. So we only need to rule out the possibility that $\{u_0(x)=0\}$ contains a ray but no straight lines. In this case, for fixed $h>0$, we can find an ellipse E inside $\{u_0(x)<h\}$, with the short axis bounded from below by a constant depending only on h and with the long axis as long as we want (one only needs to look at the asymptotic cone of $\{u_0(x)=h\}$), but since $\{u_0(x)=h\}$ evolves under the generalized curve shortening flow and $E\subset\{u_0(x)\le h\}$, this is impossible by comparison principle.

Then since $\{u_0(x)=0\}=\{0\}$, $\Gamma_{1,u_0}=\{u_0(x)=1\}$ is a bounded convex curve, and the level set $\{u_0(x)=-t\}$ evolves under the generalized curve shortening flow, with time $t\in(-\infty,0)$. From [Andrews 1998; 2003] we have the following asymptotic behavior of the convex solution u_0 of $L_0u=1$:

$$u_0(x)=\tfrac{1}{1+\alpha}|x|^{1+\alpha}+\varphi(x), \tag{56}$$

where $\varphi(x)=o(|x|^{\alpha+1})$ for $x\ne 0$ near the origin. In fact, if the initial level curve is in a sufficiently small neighborhood of circle, by Lemma 13, $|\varphi(x)|\le C|x|^{1+\alpha+\eta}$ for some small positive η, where C is a constant depending only on the initial closeness to the circle. Hence, given any $\epsilon>0$, for sufficiently small $h'>0$,

$$B_{(1-\epsilon)r}(0)\subset\Omega_{h',u_0}\subset B_{(1+\epsilon)r}(0),$$

where $r=((1+\alpha)h')^{\frac{1}{1+\alpha}}$. Hence, there is a sequence $h_m\to\infty$ such that

$$B_{(1-\frac{1}{m})r_{m,i}}(0)\subset\Omega_{h_m,u}\subset B_{(1+\frac{1}{m})r_{m,i}}(0),$$

where

$$r_{m,i}=((1+\alpha)ih_m)^{\frac{1}{1+\alpha}},\quad i=1,\dots,m.$$

Then u_{h_m} subconverges to $\frac{1}{1+\alpha}|x|^{1+\alpha}$.

Since u_0 is an entire convex solution to $L_0 u = 1$ (we still use the notation u_0, but it means an arbitrary entire convex solution), from the above argument, we can find a sequence h_m such that

$$u_{0h_m}(x) = \frac{1}{h_m} u_0\big(h_m^{\frac{1}{1+\alpha}} x\big)$$

locally uniformly converges to $\frac{1}{1+\alpha}|x|^{1+\alpha}$. Hence, the sublevel set $\Omega_{\frac{1}{1+\alpha}, u_{0hm}}$ satisfies

$$B_{1-\epsilon_m}(0) \subset \Omega_{\frac{1}{1+\alpha}, u_{0hm}} \subset B_{1+\epsilon_m}(0),$$

where $\epsilon_m \to 0$ as $m \to \infty$. By the discussion below (56),

$$u_{0h_m}(x) = \tfrac{1}{1+\alpha}|x|^{1+\alpha} + \varphi(x),$$

where $|\varphi(x)| \le C\,|x|^{1+\alpha+\eta}$ for some fixed small positive η, and the constant C is independent of m. Replacing x by $h_m^{-1/(1+\alpha)} x$ in the asymptotic formula above,

$$u_0(x) = \tfrac{1}{1+\alpha}|x|^{1+\alpha} + h_m \varphi\big(h_m^{-\frac{1}{1+\alpha}} x\big),$$

where for any fixed x, $h_m \varphi(h_m^{-\frac{1}{1+\alpha}} x) \to 0$. Hence $u_0(x) = \frac{1}{1+\alpha}|x|^{1+\alpha}$. So we have proved Theorem 1 and the first part of Theorem 2. □

4. One-dimensional entire convex ancient solution must be a shrinking circle

This section is devoted to the proof of Corollary 3, which is completed by combining the following lemma (corresponding to [Wang 2011, Lemma 4.1]) and Theorem 2.

Lemma 11. *Let Ω be a smooth, bounded, convex domain in $\mathbb{R}^2$. Let u be the solution of* (3) *with $\sigma = 0$, vanishing on $\partial\Omega$. Then for any constant h satisfying $\inf_\Omega u < h < 0$, the level set $\Gamma_{h,u} = \{u = h\}$ is convex. Moreover*, $\log(-u)$ *is a concave function.*

Proof. Observe that $\varphi := -\log(-u)$ satisfies

$$|D\varphi|^{\frac{1}{\alpha}-1} \sum_{i,j=1}^{2} \left(\delta_{ij} - \frac{\varphi_i \varphi_j}{|D\varphi|^2}\right) \varphi_{ij} = e^{\frac{1}{\alpha}\varphi}.$$

Since $\varphi(x) \to +\infty$ as $x \to \partial\Omega$, [Kawohl 1985, Theorem 3.13] implies that φ is convex. □

With the previous lemma and [Wang 2011, Lemma 4.4], we know that any convex compact ancient solution to the generalized curve shortening flow can be represented as a convex solution u to (3) with $\sigma = 0$, and if the solution to the flow sweeps the whole space, then the corresponding u will be an entire solution. Thus, Theorem 2 implies Corollary 3 immediately.

Remark 12. We can also use the method in [Wang 2011, Section 4] to construct a rotationally nonsymmetric convex compact ancient solution for generalized curve shortening flow with power $\alpha \in \left(\frac{1}{2}, 1\right)$. Indeed, for mean curvature flow, a rotationally nonsymmetric convex compact ancient solution is constructed in Lemmas 4.1–4.4 of that reference. By examining the proofs of these lemmas, we can see that they work well for the generalized curve shortening flow considered here.

5. Two-dimensional entire convex translating solution

In this section, by using the previous results and a delicate iteration argument, we prove that under some extra condition on the asymptotic behavior of the solution at infinity the translating solution must be rotationally symmetric.

First of all, we would like to point out that instead of using Gage and Hamilton's exponential convergence of the curve shortening flow [1986], we need to use the corresponding exponential convergence for the generalized curve shortening flow and we will state it as a lemma, corresponding to [Wang 2011, Lemma 3.2].

Lemma 13. *Let $\{\ell_t\}$ be a convex solution to the generalized curve shortening flow with initial curve $\{\ell_0\}$ uniformly convex. Suppose $\{\ell_0\}$ is in the δ_0-neighborhood of a unit circle, $\{\ell_t\}$ shrinks to the origin at $t = \frac{1}{1+\alpha}$. Let*

$$\tilde{\ell}_t = (1-(1+\alpha)t)^{-\frac{1}{1+\alpha}}\ell_t$$

be the normalization of ℓ_t. Then $\tilde{\ell}_t$ is in the δ_t-neighborhood of the unit circle centered at the origin,

$$\tilde{\ell}_t \subset N_{\delta_t} S^1,$$

with

$$\delta_t \le C\delta_0\left(\tfrac{1}{1+\alpha} - t\right)^{\iota}$$

for some small positive constant ι.

Remark 14. Exponential convergence of the standard curve shortening flow (when $\alpha = 1$) was proved by Gage and Hamilton [1986]. For the general case (when $\alpha > \frac{1}{3}$), as discussed in the following proof, Gage and Hamilton's method combined with Andrews' estimates [1998, Propositions II1.1 and II1.2] can still be used to prove the corresponding exponential convergence result.

Proof. The proof of Lemma 13 is similar to the proof of [Wang 2011, Lemma 3.2]. Since the initial curve ℓ_0 is uniformly convex and close to a unit circle, by [Andrews 1998, Propositions II1.1 and II1.2], the curvature of $\tilde{\ell}_t$ is bounded from below and from above by some constant depending only on δ_0, when $t \in \left(\frac{1}{4\alpha+4}, \frac{1}{2\alpha+2}\right)$.

Hence the evolution equation for $\tilde{\ell}_t$ is uniformly parabolic. Therefore, we can apply Schauder's estimates safely for $\alpha > \frac{1}{2}$, as in [Wang 2011], which says that

$$\|\tilde{\ell}_t - S^1\|_{C^k} \le C\delta_0 \quad \text{for } t \in \left(\tfrac{1}{4\alpha+4}, \tfrac{1}{2\alpha+2}\right).$$

Although the constant C will depend on the lower and upper bounds of the curvature of the initial curve, it is not a problem for our purpose, since when we blow down the solution for $\sigma = 0$, the norm of the gradient Du_h on the curve $\{u_h(x) = 1\}$ approaches 1.

By the equation $\kappa u_\gamma^{\frac{1}{\alpha}} = 1$, we see that the curvature κ is also very close to 1 on that curve. For the exponential decay rate of the derivative of curvature, one can imitate the proof in [Gage and Hamilton 1986, §§5.7.10–5.7.15], and our corresponding estimate will be $|\kappa'(\tau)| \le C\delta_0 e^{-\iota\tau}$ for some small positive number ι, where $\tau = -\frac{1}{1+\alpha}\log\left(\frac{1}{1+\alpha} - t\right)$. Indeed, in the case $\alpha > 1$, this is done by Chen and Huang [Huang 2011, Corollary 3.2], and it is easy to check that their computation also works for the case $\frac{1}{3} < \alpha < 1$ by taking ℓ small enough. This estimate immediately implies our lemma. □

In the following we will consider the case when $\sigma = 1$ and $\alpha > 1$. By translating and adding some constant we can assume $u(0) = \inf u$. Let

$$u_h(x) = \frac{1}{h} u(h^{\frac{1}{1+\alpha}} x).$$

Then u_h satisfies the equation $L_\sigma u_h = 1$ with $\sigma = h^{-\frac{2\alpha}{1+\alpha}}$. By Theorem 1, u_h converges to $\frac{1}{1+\alpha}|x|^{1+\alpha}$, and the level set $\Gamma_{\frac{1}{1+\alpha}, u_h}$ converges to the unit circle as $h \to \infty$.

Lemma 15 [Wang 2011, Lemma 3.3]. *The function u satisfies*

$$u(x) = \tfrac{1}{1+\alpha}|x|^{1+\alpha} + O(|x|^{1+\alpha-2\alpha\beta}), \tag{57}$$

where β is a constant, chosen such that $\frac{1}{2\alpha} < \beta < \min\{1, \frac{1+\alpha}{2\alpha}\}$.

Proof. For any given small $\delta_0 > 0$, take h sufficiently large such that

$$\Gamma_{\frac{1}{1+\alpha}, u_h} \subset N_{\delta_0}(S^1) \tag{58}$$

for the unit circle S^1 with center p_0. Note that when h is large, δ_0 is very close to 0. Then we will prove the following claim:

Claim 3: For small fixed τ,

$$\Gamma_{\tau, u_h} \subset ((1+\alpha)\tau)^{\frac{1}{1+\alpha}} N_{\delta_\tau}\left((1 + \tfrac{a_0}{\tau})^{\frac{1}{1+\alpha}} S^1\right) \tag{59}$$

with

$$\delta_\tau \le C_1(\tau)\sigma^\beta + C_2\delta_0\tau^\eta, \tag{60}$$

where η is a small positive constant, the constants C_1 and C_2 are independent of δ_0 and h, and C_2 is also independent of τ. Let u_0 be the solution of

$$L_0(u) = 1 \quad \text{in } \Omega_{\frac{1}{1+\alpha}, u_h},$$

satisfying

$$u_0 = u_h = \tfrac{1}{1+\alpha} \quad \text{on } \partial\Omega_{\frac{1}{1+\alpha}, u_h},$$

where $a_0 = |\inf u_0|$ and the center of

$$(1 + \tfrac{a_0}{\tau})^{\frac{1}{1+\alpha}} S^1$$

is the minimum point of u_0 multiplied by the factor $((1+\alpha)\tau)^{-\frac{1}{1+\alpha}}$.

Proof of Claim 3. We only need to prove that

$$\operatorname{dist}\big((1+\alpha)^{\frac{1}{1+\alpha}}(\tau + a_0)^{\frac{1}{1+\alpha}} S^1, \Gamma_{\tau,u}\big) \leq C_1(\tau)\sigma^\beta + C_2\delta_0\tau^{\frac{1}{1+\alpha}+\eta}, \tag{61}$$

where η is some small positive constant and C_2 is independent of τ. By Theorem 1 we know that u_h converges to $\frac{1}{1+\alpha}|x|^{1+\alpha}$ uniformly on any compact subset of $\mathbb{R}^2$. Then by the convexity of u_h, we have that $|Du_h|$ is bounded above and below by some constants depending on τ_0 for large h when

$$x \in \big\{x \in \Omega_{\frac{1}{1+\alpha}, u_h} : \tau_0 \leq u_h < \tfrac{1}{1+\alpha}\big\}.$$

Hence, by the growth condition for D^2u in Theorem 2, we have $\sigma(u_h)_{\gamma\gamma} \leq C\sigma^\beta$, where C is a constant depending on τ_0. Therefore, we have

$$\kappa(u_h)_\gamma^{1/\alpha} \approx 1 - C\sigma^\beta \quad \text{on } \big\{x \in \Omega_{\frac{1}{1+\alpha}, u_h} : \tau \leq u_h < \tfrac{1}{1+\alpha}\big\},$$

where C depends on τ_0. Let

$$\tilde{u}_0 = (1 - C\sigma^\beta)^\alpha (u_0 - \tfrac{1}{1+\alpha}) + \tfrac{1}{1+\alpha};$$

then

$$L_0(\tilde{u}_0) = 1 - C\sigma^\beta \quad \text{in } \Omega_{\frac{1}{1+\alpha}, u_h}$$

with

$$\tilde{u}_0 = u_h = \tfrac{1}{1+\alpha} \quad \text{on } \partial\Omega_{\frac{1}{1+\alpha}, u_h}.$$

Now by the comparison principle, $\Omega_{\tau,u_0} \supset \Omega_{\tau,u_h} \supset \Omega_{\tau,\tilde{u}_0}$, and by the asymptotic behavior of u_0,

$$\Gamma_{\tau,u_0} \subset N_\zeta\big((\tau + a_0)^{\frac{1}{1+\alpha}} S^1\big) \quad \text{and} \quad \Gamma_{\tau,\tilde{u}_0} \subset N_\zeta\big((\tau + a_0 - C\sigma^\beta)^{\frac{1}{1+\alpha}} S^1\big),$$

where $\zeta = C\delta_0(\tau + a_0)^\eta$. Let

$$\ell_1 = (\tau + a_0)^{\frac{1}{1+\alpha}} S^1 \quad \text{and} \quad \ell_2 = (\tau + a_0 - C\sigma^\beta)^{\frac{1}{1+\alpha}} S^1,$$

both centered at p_1, which is the minimum point of u_0. Hence

$$\operatorname{dist}((\tau+a_0)^{\frac{1}{1+\alpha}}S^1, \Gamma_{\tau,u_h}) \le \operatorname{dist}(\ell_1,\ell_2) + C\delta_0(\tau+a_0)^{\frac{1}{1+\alpha}+\eta}, \tag{62}$$

where $\operatorname{dist}(\ell_1,\ell_2)$ can be bounded by $C_1(\tau)\sigma^\beta$; hence (60) follows from the above discussion. □

Now, we will use an iteration argument to prove the following claim, which will enable us to simplify (59) and (60).

Claim 4: $a_0 \le C\sigma^\beta |\log(\sigma)|$

Proof. We fix a large constant A such that $\{u_{A/\tau} = \frac{1}{1+\alpha}\}$ is very close to a unit circle. Let u_{0,τ^k} solve $L_0 u = 1$ with boundary condition $u = \tau^k$ on $\{u_h = \tau^k\}$. Denote $a_k = |\inf u_{0,\tau^k}|$. From the proof of Claim 3, we see that

$$\{u_0 < \tau\} \supset \{u_{0,\tau} < \tau\} \supset \{\tilde{u}_0 < \tau\},$$

by the comparison principle, we have $\inf u_0 < \inf u_{0,\tau} < \inf \tilde{u}_0$. So by the construction of $\tilde{u}_0$ and a simple computation, we have $a_0 - a_1 \le \inf \tilde{u}_0 - \inf u_0 \le C\sigma^\beta$. When $\tau^k \ge \frac{A}{h}$, we can iterate this argument for u_{0,τ^k} and $u_{0,\tau^{k+1}}$ by rescaling them to

$$\tfrac{1}{1+\alpha}\tau^{-k} u_{0,\tau^k}\big((1+\alpha)^{\frac{1}{1+\alpha}}\tau^{\frac{k}{1+\alpha}}x\big) \quad \text{and} \quad \tfrac{1}{1+\alpha}\tau^{-k} u_{0,\tau^{k+1}}\big((1+\alpha)^{\frac{1}{1+\alpha}}\tau^{\frac{k}{1+\alpha}}x\big),$$

respectively. After rescaling back, we have $a_k - a_{k+1} \le C\sigma^\beta$. Note that the choice of A and the condition $\tau^k \ge \frac{A}{h}$ ensure the uniform gradient bound needed in the above argument. Let k_0 be an integer satisfying $\tau^{k_0} \ge \frac{A}{h} \ge \tau^{k_0+1}$. After k_0 steps we stop the iteration, and notice that

$$\{u_h = \tfrac{A}{h}\} = h^{-\frac{1}{1+\alpha}}\{u = A\}$$

is contained in a circle with radius $Ch^{-\frac{1}{1+\alpha}}$ for some constant C. Hence it takes at most time $Ch^{-1} = C\sigma^{\frac{1+\alpha}{2\alpha}}$ for $\{u_h = \frac{A}{h}\}$ to shrink to a point. Claim 4 follows from the above discussion. □

By omitting the lower order term we can rewrite (59) and (60) as

$$\Gamma_{\tau,u_h} \subset ((1+\alpha)\tau)^{\frac{1}{1+\alpha}} N_{\delta_\tau}(S^1)$$

with

$$\delta_\tau \le C_1(\tau)\sigma^\beta + C_2\delta_0\tau^\eta. \tag{63}$$

If we take τ small such that $C_2\tau^\eta \le \frac{1}{4}$, then (63) becomes

$$\delta_\tau \le C_1(\tau)\sigma^\beta + \tfrac{1}{4}\delta_0. \tag{64}$$

Now we can carry out an iteration argument similar to that in [Wang 2011]. We start at the level $\frac{1}{1+\alpha}\tau^{-k_0}$ for some sufficient large k_0. Let

$$\Omega_k = \tau^{\frac{k}{1+\alpha}} \Omega_{\frac{1}{1+\alpha}\tau^{-k},u} \quad \text{and} \quad \Gamma_k = \partial\Omega_k.$$

Note that Γ_k converges to a unit circle as $k \to \infty$. Suppose that Γ_k is in the δ_k neighborhood of S^1 centered at y_k, where $\delta_k \to 0$ as $k \to \infty$ and y_k is the minimum point of the solution of $L_0 u = 1$ in Ω_k with $u = \frac{1}{1+\alpha}$ on Γ_{k+1}. By (64) we have

$$\delta_{k-1} \le C_1(\tau)\tau^{(k-1)\frac{2\alpha\beta}{1+\alpha}} + \tfrac{1}{4}\delta_k \tag{65}$$

for $k = k_0, k_0+1, \ldots$. Then we have

$$\Gamma_j \subset N_{\delta_j}(S^1) \tag{66}$$

with

$$\delta_j \le C\tau^{j\frac{2\alpha\beta}{1+\alpha}} \tag{67}$$

It follows that

$$\Gamma_{\frac{1}{1+\alpha}\tau^{-j},u} \subset N_{\tilde{\delta}_j}(\tau^{\frac{-j}{1+\alpha}} S^1) \tag{68}$$

with

$$\tilde{\delta}_j \le C\tau^{\frac{2\alpha\beta-1}{1+\alpha}j}, \tag{69}$$

where $\tau^{\frac{-j}{1+\alpha}} S^1$ is centered at $z_j = \tau^{\frac{-j}{1+\alpha}} y_j$. From Lemma 13 and (64), it is not hard to see that

$$|z_j - z_{j-1}| \le C\tau^{\frac{2\alpha\beta-1}{1+\alpha}j}. \tag{70}$$

Let $z_0 = \lim_{j\to\infty} z_j$. Then

$$|z_j - z_0| \le C\tau^{\frac{2\alpha\beta-1}{1+\alpha}j}, \tag{71}$$

which means that in (68) we can assume the circle is centered at z_0 by changing the constant C a little bit. In fact when we choose different τ, the corresponding z_0 will not change, so we can assume $z_0 = 0$. Hence, for $h = \frac{1}{1+\alpha}\tau^{-j}$,

$$\Gamma_{h,u} \subset N_\delta\big((1+\alpha)^{\frac{1}{1+\alpha}} h^{\frac{1}{1+\alpha}} S^1\big),$$

where

$$\delta \le Ch^{\frac{1-2\alpha\beta}{1+\alpha}} \tag{72}$$

and S^1 is centered at the origin. By choosing different τ, we see that the estimate holds for all large h. Lemma 15 follows from the above estimates. $\square$

Remark 16. For the mean curvature flow (when $\alpha = 1$), I learned the proof of Claim 4 from Professor Xu-Jia Wang. Indeed, in this case one can do it as follows. Let v_k solve $L_0 w = 1$ with boundary condition $w = \tau^k$ on $\{u_h = \tau^k\}$, for $k = 1, 2, \ldots$. Let $a_k = |\inf v_k|$. Then by [Wang 2011, Lemma 3.1], we have $a_0 - a_1 \le C\sigma$. By rescaling invariance, we can iterate the argument to show that $a_k - a_{k+1} \le C\sigma$, provided $\tau^k \ge \frac{1}{h}$. Hence, we stop the iteration at k_0 when $\tau^{k_0} \ge \frac{1}{h} > \tau^{k_0+1}$. Notice that

$$\{u_h \le \tau^{k_0}\} = h^{-\frac{1}{2}}\{u = h\tau^{k_0}\} \subset h^{-\frac{1}{2}}\{u \le \tfrac{1}{\tau}\}.$$

So, it is easy to see that it takes at most time $C\sigma = \frac{C}{h}$ for $\{u_h \le \tau^{k_0}\}$ to shrink to a point, namely, $a_{k_0} \le C\sigma$. Therefore,

$$a_0 = a_{k_0} + \sum_{i=0}^{k_0-1} a_i - a_{i+1} \le C k_0 \sigma \le C\sigma|\log\sigma|.$$

In order to finish the proof of Theorem 2 we need to use the following fundamental Liouville theorem by Bernstein [Simon 1997, p. 245].

Lemma 17. *Let u be an entire solution to the elliptic equation*

$$\sum_{i,j=1}^{n} a_{ij}(x) u_{ij} = 0 \quad \text{in } \mathbb{R}^2.$$

If u satisfies the asymptotic estimate

$$|u(x)| = o(|x|) \quad \text{as } x \to \infty,$$

then u is a constant.

Proof of the second part of Theorem 2. Let u^* be the Legendre transform of u. Then u^* satisfies equation

$$G(x, D^2u^*) = \frac{\det D^2 u^*}{\left(\delta_{ij} - \frac{x_i x_j}{1+|x|^2}\right) F^{ij}(u^*)} = (1 + |x|^2)^{\frac{1}{2\alpha} - \frac{1}{2}}, \tag{73}$$

where $F^{ij}(u^*) = \frac{\partial \det r}{\partial r_{ij}}$ at $r = D^2u^*$. We have

$$u^*(x) = C(\alpha)|x|^{1+\alpha} + O\left(|x|^{\frac{1+\alpha-2\alpha\beta}{\alpha}}\right), \tag{74}$$

where $C(\alpha)$ is a constant depending only on α. In fact, for big h, by Lemma 15,

$$u_h(x) = \tfrac{1}{1+\alpha}|x|^{1+\alpha} + O\left(|h|^{\frac{-2\alpha\beta}{1+\alpha}}\right)$$

in $B_1(0)$. Denote by u_h^* the Legendre transform of u_h. Then,

$$u_h^*(x) = C(\alpha)|x|^{1+\frac{1}{\alpha}} + O\left(|h|^{\frac{-2\alpha\beta}{1+\alpha}}\right),$$

where $C(\alpha)$ is a constant depending only on α and it comes from the Legendre transform of the function $\frac{1}{1+\alpha}|x|^{1+\alpha}$. Note that $u_h^*(x) = h^{-1}u^*(h^{\frac{\alpha}{1+\alpha}}x)$, we obtain (74).

Let u_0 be the unique radial solution of (3) with $\sigma = 1$, and let u_0^* be the Legendre transform of u_0. Similar to (74) we have

$$u_0^*(x) = C(\alpha)|x|^{1+\alpha} + O(|x|^{\frac{1+\alpha-2\alpha\beta}{\alpha}}). \tag{75}$$

Since both u^* and u_0^* satisfy (73), $v = u^* - u_0^*$ satisfies the elliptic equation

$$\sum_{i,j=1}^{n} a_{ij}(x)v_{ij} = 0 \quad \text{in } \mathbb{R}^2,$$

where

$$a_{ij} = \int_0^1 G^{ij}(x, D^2u_0^* + t(D^2u^* - D^2u_0^*))\,dt.$$

Here,

$$G^{ij} = \frac{\partial G(x,r)}{\partial r_{ij}}$$

for any symmetric matrix r. Note that by the choice of β, $\frac{1+\alpha-2\alpha\beta}{\alpha} < 1$; hence, by (74) and (75),

$$v = O\big(|x|^{\frac{1+\alpha-2\alpha\beta}{\alpha}}\big) = o(|x|) \quad \text{as } |x| \to \infty.$$

By Lemma 17 we conclude that v is a constant. □

References

[Andrews 1998] B. Andrews, “Evolving convex curves”, *Calc. Var. PDE* **7**:4 (1998), 315–371. MR 99k:58038 Zbl 0931.53030

[Andrews 2002] B. Andrews, “Non-convergence and instability in the asymptotic behaviour of curves evolving by curvature”, *Comm. Anal. Geom.* **10**:2 (2002), 409–449. MR 2003e:53086 Zbl 1029.53079

[Andrews 2003] B. Andrews, “Classification of limiting shapes for isotropic curve flows”, *J. Amer. Math. Soc.* **16**:2 (2003), 443–459. MR 2004a:53083 Zbl 1023.53051

[Chou and Zhu 2001] K.-S. Chou and X.-P. Zhu, *The curve shortening problem*, Chapman & Hall/CRC, Boca Raton, FL, 2001. MR 2003e:53088 Zbl 1061.53045

[Daskalopoulos et al. 2010] P. Daskalopoulos, R. Hamilton, and N. Sesum, “Classification of compact ancient solutions to the curve shortening flow”, *J. Differential Geom.* **84**:3 (2010), 455–464. MR 2012d:53213 Zbl 1205.53070

[Gage and Hamilton 1986] M. Gage and R. S. Hamilton, “The heat equation shrinking convex plane curves”, *J. Differential Geom.* **23**:1 (1986), 69–96. MR 87m:53003 Zbl 0621.53001

[Huang 2011] R. L. Huang, “Blow-up rates for the general curve shortening flow”, *J. Math. Anal. Appl.* **383**:2 (2011), 482–489. MR 2012h:53152 Zbl 1220.53082

[Kawohl 1985] B. Kawohl, *Rearrangements and convexity of level sets in PDE*, Lecture Notes in Mathematics **1150**, Springer, Berlin, 1985. MR 87a:35001 Zbl 0593.35002

[Schulze 2005] F. Schulze, "Evolution of convex hypersurfaces by powers of the mean curvature", *Math. Z.* **251**:4 (2005), 721–733. MR 2006h:53065 Zbl 1087.53062

[Sheng and Wu 2009] W. Sheng and C. Wu, "On asymptotic behavior for singularities of the powers of mean curvature flow", *Chin. Ann. Math.* (*B*) **30**:1 (2009), 51–66. MR 2010h:53100 Zbl 1180.53066

[Simon 1997] L. Simon, "The minimal surface equation", pp. 239–266 in *Geometry V*, edited by R. Osserman, Encyclopaedia of Mathematical Sciences **90**, Springer, Berlin, 1997. MR 99b:53014 Zbl 0905.53003

[Wang 2011] X.-J. Wang, "Convex solutions to the mean curvature flow", *Ann. of Math.* (2) **173**:3 (2011), 1185–1239. MR 2800714 Zbl 1231.53058

[White 2000] B. White, "The size of the singular set in mean curvature flow of mean-convex sets", *J. Amer. Math. Soc.* **13**:3 (2000), 665–695. MR 2001j:53098 Zbl 0961.53039

Received February 19, 2013. Revised October 19, 2014.

SHIBING CHEN
DEPARTMENT OF MATHEMATICS
ZHEJIANG UNIVERSITY OF TECHNOLOGY
HANGZHOU, ON 310023
CHINA
Current address:
MATHEMATICAL SCIENCES INSTITUTE
THE AUSTRALIAN NATIONAL UNIVERSITY
JOHN DEDMAN BUILDING 27
CANBERRA ACT 2601
AUSTRALIA

chenshibing1982@hotmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.143

CONSTRUCTIONS OF PERIODIC MINIMAL SURFACES AND MINIMAL ANNULI IN SOL_3

CHRISTOPHE DESMONTS

We construct two one-parameter families of minimal properly embedded surfaces in the Lie group Sol_3 using a Weierstrass-type representation. These surfaces are not invariant by a one-parameter group of ambient isometries. The first one can be viewed as a family of helicoids, and the second one as a family of minimal annuli called catenoids. Finally we study limits of these catenoids, and in particular we show that one of these limits is a new minimal entire graph.

1. Introduction

The aim of this paper is to construct two one-parameter families of examples of properly embedded minimal surfaces in the Lie group Sol_3, endowed with its standard metric. This Lie group is a homogeneous Riemannian manifold with a 3-dimensional isometry group and is one of the eight Thurston geometries. There is no rotation in Sol_3, and so no surface of revolution.

The Hopf differential, which exists on surfaces in every 3-dimensional space form, has been generalized by Abresch and Rosenberg [2004; 2005] to every 3-dimensional homogeneous Riemannian manifold with 4-dimensional isometry group. This tool leads to a lot of works in the field of constant mean curvature (CMC) surfaces in Nil_3, $\widetilde{PSL}_2(\mathbb{R})$ and in the Berger spheres. More precisely, Abresch and Rosenberg [2005] proved that the generalized Hopf differential exists in a simply connected Riemannian 3-manifold if and only if its isometry group has at least dimension 4.

Berdinskii and Taimanov [2005] gave a representation formula for minimal surfaces in 3-dimensional Lie groups in terms of spinors, but they pointed out some difficulties for using this theory in the case of Sol_3. Nevertheless, some explicit simple examples of minimal surfaces in Sol_3 have been constructed in the past decade. Masaltsev [2006] and Daniel and Mira [2013] gave some basic examples of minimal graphs in Sol_3: $x_1 = ax_2 + b$, $x_1 = ae^{-x_3}$, $x_1 = ax_2e^{-x_3}$ and $x_1 = x_2e^{-2x_3}$ (and their images by ambient isometries). López and Munteanu

MSC2010: 53A10.
Keywords: periodic minimal surfaces, minimal annuli, Sol_3.

[2011; 2012; 2014], López [2014] and Masaltsev [2006] studied minimal surfaces in Sol_3 invariant by a one-parameter group of ambient isometries. Finally, Ana Menezes [2014] constructed singly and doubly periodic Scherk minimal surfaces in Nil_3 and Sol_3, and Minh Hoang Nguyen [2014] gave conditions for the Dirichlet problem for the minimal surface equation in Sol_3 to have solutions.

The method that we use in this paper is the one used by Daniel and Hauswirth [2009] in Nil_3 to construct minimal embedded annuli: We first construct a one-parameter family of embedded minimal surfaces called *helicoids* and we calculate its Gauss map g. A result of Inoguchi and Lee [2008] shows that this map is harmonic for a certain metric on $\overline{\mathbb{C}}$. Then we seek another family of maps g with separated variables that still satisfies the harmonic map equation, and we use a Weierstrass-type representation given by Inoguchi and Lee to construct a minimal immersion whose Gauss map is g. We prove that these immersions are periodic, so we get minimal annuli. As far as the authors know, these annuli are the first examples of nonsimply connected minimal surfaces with finite topology (that is, diffeomorphic to a compact surface without a finite number of points) in Sol_3.

The model we use for Sol_3 is described in Section 2. In the third section, we give some properties of the Gauss map of a conformal minimal immersion in Sol_3 (see [Daniel and Mira 2013]). In the fourth section, we construct the family $(\mathcal{H}_K)_{K\in]-1;1[}$ of helicoids, and finally we construct the family $(\mathcal{C}_\alpha)_{\alpha\in]-1;1[\setminus\{0\}}$ of embedded minimal annuli. The study of the limit case of the parameter of this family gives another example of a minimal surface in Sol_3, which is an entire graph. None of these surfaces is invariant by a one-parameter family of isometries.

Theorem. *There exists a one-parameter family* $(\mathcal{C}_\alpha)_{\alpha\in]-1;1[\setminus\{0\}}$ *of properly embedded minimal annuli in* Sol_3, *called* catenoids, *having the following properties*:

(1) *The intersection of* $\mathcal{C}_\alpha$ *with any plane* $\{x_3=\lambda\}$ *is a nonempty closed embedded convex curve.*

(2) *The annulus* $\mathcal{C}_\alpha$ *is conformally equivalent to* $\mathbb{C}\setminus\{0\}$.

(3) *The annulus* $\mathcal{C}_\alpha$ *has three symmetries fixing the origin*: *rotation by* π *around the* x_3*-axis*, *reflection in* $\{x_1=0\}$ *and reflection in* $\{x_2=0\}$.

2. The Lie group Sol_3

Definition. The Lie group Sol_3 is $\mathbb{R}^3$ with the multiplication $*$ defined by

$$(x_1,x_2,x_3)*(y_1,y_2,y_3)=(y_1e^{-x_3}+x_1,\,y_2e^{x_3}+x_2,\,x_3+y_3)$$

for all $(x_1,x_2,x_3),(y_1,y_2,y_3)\in\mathbb{R}^3$. The identity element is 0 and the inverse element of (x_1,x_2,x_3) is $(x_1,x_2,x_3)^{-1}=(-x_1e^{x_3},-x_2e^{-x_3},-x_3)$. The Lie group is noncommutative.

The left multiplication l_a by an element $a = (a_1, a_2, a_3) \in \mathbb{R}^3$ is given for all $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ by

$$\begin{aligned} l_a(x) = a * x &= (x_1 e^{-a_3} + a_1, x_2 e^{a_3} + a_2, a_3 + x_3) \\ &= a + M_a x, \end{aligned}$$

where

$$M_a = \begin{pmatrix} e^{-a_3} & 0 & 0 \\ 0 & e^{a_3} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For the metric $(\cdot, \cdot)$ on Sol$_3$ to be left-invariant, it has to satisfy

$$(M_a X, M_a Y)_{a*x} = (X, Y)_x$$

for all $a, x, X, Y \in \mathbb{R}^3$. We define a left-invariant Riemannian metric for $x, X, Y \in \mathbb{R}^3$ by the formula

$$(X, Y)_x = \langle M_{x^{-1}} X, M_{x^{-1}} Y \rangle, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ is the canonical scalar product on $\mathbb{R}^3$ and x^{-1} is the inverse element of x in Sol$_3$. The formula (1) leads to the expression of the previous metric

$$ds_x^2 = e^{2x_3}\, dx_1^2 + e^{-2x_3}\, dx_2^2 + dx_3^2, \tag{2}$$

where (x_1, x_2, x_3) are canonical coordinates of $\mathbb{R}^3$. Since the translations are isometries now, Sol$_3$ is a homogeneous manifold with this metric.

Remark. This metric is not the only possible left-invariant one on Sol$_3$. In fact, there exists a two-parameter family of nonisometric left-invariant metrics on Sol$_3$. One of these parameters is a homothetic one. The metrics that are nonhomothetic to (2) have no reflections; see [Meeks and Pérez 2012].

By setting

$$E_1(x) = e^{-x_3}\, \partial_1, \quad E_2(x) = e^{x_3}\, \partial_2, \quad \text{and} \quad E_3(x) = \partial_3,$$

we obtain a left-invariant orthonormal frame (E_1, E_2, E_3). Thus, we now have two frames to express the coordinates of a vector field on Sol$_3$; we will use brackets to denote the coordinates in the frame (E_1, E_2, E_3); then at a point $x \in$ Sol$_3$, we have

$$a_1\, \partial_1 + a_2\, \partial_2 + a_3\, \partial_3 = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{bmatrix} e^{x_3} a_1 \\ e^{-x_3} a_2 \\ a_3 \end{bmatrix}. \tag{3}$$

The following property holds (see [Daniel and Mira 2013]):

Proposition 1. *The isotropy group of the origin of* Sol_3 *is isomorphic to the dihedral group* $\mathcal{D}_4$ *and generated by orientation-reversing isometries*

$$\sigma : (x_1, x_2, x_3) \longmapsto (x_2, -x_1, -x_3) \quad \textit{and} \quad \tau : (x_1, x_2, x_3) \longmapsto (-x_1, x_2, x_3),$$

whose orders are 4 *and* 2, *respectively.*

While τ is simply a reflection in the plane $\{x_1 = 0\}$, the generator σ can be described as a rotation by $-\pi/3$ around E_3 composed with reflection in $\{x_3 = 0\}$. The cyclic group $\langle\sigma\rangle$ also contains $\sigma^3 = \sigma^{-1}$ and σ^2, the reflection in E_3 (rotation by π around E_3). The remaining nonidentity elements of the isotropy group of the origin are $\sigma\tau$ and $\sigma^3\tau$, which are respectively the reflections in the lines $\{(x_1, x_1, 0)\}$ and $\{(x_1, -x_1, 0)\}$, and $\sigma^2\tau$, which is reflection in the plane $\{x_2 = 0\}$.

We deduce the following theorem:

Theorem 2. *The isometry group of* Sol_3 *has dimension* 3.

Finally, we express the Levi-Civita connection ∇ of Sol_3 associated to the metric given by (2) in the frame (E_1, E_2, E_3). First, we calculate the Lie brackets of the vectors of the frame. The Lie bracket in the Lie algebra $\mathfrak{sol}_3$ of Sol_3 is given by

$$[X, Y] = (Y_3 X_1 - X_3 Y_1, X_3 Y_2 - Y_3 X_2, 0)$$

for all $X = (X_1, X_2, X_3)$ and $Y = (Y_1, Y_2, Y_3)$. Then we have

$$[E_1, E_2] = 0, \quad [E_1, E_3] = E_1, \quad [E_2, E_3] = -E_2.$$

Hence,

$$\begin{aligned} &\nabla_{E_1} E_1 = -E_3, && \nabla_{E_2} E_1 = 0, && \nabla_{E_3} E_1 = 0, \\ &\nabla_{E_1} E_2 = 0, && \nabla_{E_2} E_2 = E_3, && \nabla_{E_3} E_2 = 0, \\ &\nabla_{E_1} E_3 = E_1, && \nabla_{E_2} E_3 = -E_2, && \nabla_{E_3} E_3 = 0. \end{aligned}$$

3. The Gauss map

Let Σ be a Riemann surface and $z = u + iv$ local complex coordinates in Σ. Let $x : \Sigma \to \mathrm{Sol}_3$ be a conformal immersion. We set

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

and we define $\lambda \in \mathbb{R}_+^*$ by

$$2(x_z, x_{\bar{z}}) = \|x_u\|^2 = \|x_v\|^2 = \lambda.$$

Thus, a unit normal vector field is $N : \Sigma \to T\mathrm{Sol}_3$ defined by

$$N = -\frac{2i}{\lambda} x_z \wedge x_{\bar{z}} := \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix}.$$

Hence we define $\widehat{N} : \Sigma \to \mathbb{S}^2 \subset \mathbb{R}^3$ by the formula $M_{x^{-1}} N = \widehat{N}$, that is,

$$\widehat{N} = \begin{pmatrix} e^{x_3} & 0 & 0 \\ 0 & e^{-x_3} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} N_1 e^{-x_3} \\ N_2 e^{x_3} \\ N_3 \end{pmatrix} = \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}.$$

Definition. The Gauss map of the immersion x is the application

$$g = \sigma \circ \widehat{N} : \Sigma \longrightarrow \mathbb{C} \cup \{\infty\} = \overline{\mathbb{C}},$$

where σ is the stereographic projection with respect to the southern pole, i.e.,

$$N = \frac{1}{1+|g|^2} \begin{bmatrix} 2\Re(g) \\ 2\Im(g) \\ 1-|g|^2 \end{bmatrix}, \tag{4}$$

$$g = \frac{N_1 + iN_2}{1+N_3}. \tag{5}$$

The following result is due to [Inoguchi and Lee 2008]. It can be viewed as a Weierstrass representation in Sol_3.

Theorem 3. *Let $x : \Sigma \to \mathrm{Sol}_3$ be a conformal minimal immersion and $g : \Sigma \to \overline{\mathbb{C}}$ its Gauss map. Then, whenever g is neither real nor purely imaginary, it is nowhere antiholomorphic ($g_z \neq 0$ for every point for any local conformal parameter z on Σ), and it satisfies the second order elliptic equation*

$$g_{z\bar{z}} = \frac{2g g_z g_{\bar{z}}}{g^2 - \bar{g}^2}. \tag{6}$$

Moreover, the immersion $x = (x_1, x_2, x_3)$ can be expressed in terms of g by the representation formulas

$$x_{1z} = e^{-x_3} \frac{(\bar{g}^2 - 1) g_z}{g^2 - \bar{g}^2}, \quad x_{2z} = i e^{x_3} \frac{(\bar{g}^2 + 1) g_z}{g^2 - \bar{g}^2}, \quad x_{3z} = \frac{2\bar{g} g_z}{g^2 - \bar{g}^2} \tag{7}$$

whenever it is well-defined.

Conversely, given a map $g : \Sigma \to \overline{\mathbb{C}}$ defined on a simply connected Riemann surface Σ satisfying (6), *then the map $x : \Sigma \to \mathrm{Sol}_3$ given by the representation formulas* (7) *is a conformal minimal immersion with possibly branched points whenever it is well-defined, and its Gauss map is g.*

Remark. (1) There exists a similar result for the case of CMC H-surfaces ; see [Daniel and Mira 2013].

(2) Equation (6) is the harmonic map equation for maps $g : \Sigma \to (\overline{\mathbb{C}}, ds^2)$ equipped with the metric

$$ds^2 = \frac{|d\omega|^2}{|\omega^2 - \bar{\omega}^2|}.$$

This is a singular metric, not defined on the real and pure imaginary axes. See [Inoguchi and Lee 2008] for more details.

(3) Equation (6) can be only considered at points where $g \neq \infty$. But if g is a solution of (6), i/g is also a solution at points where $g \neq 0$. The conjugate map $\bar{g}$ and every $g \circ \phi$, with ϕ a locally injective holomorphic function, are solutions too. Moreover, if g is a nowhere antiholomorphic solution of (6), and x is the induced conformal minimal immersion, then ig and $1/g$ induce conformal minimal immersions given by σx and τx. Finally, $\bar{g}$ is the Gauss map of $\sigma^2 \tau x$ after a change of orientation.

Definition. The Hopf differential of the map g is the quadratic form

$$Q = q\, dz^2 = \frac{g_z \bar{g}_z}{g^2 - \bar{g}^2}\, dz^2.$$

Remark. (1) The function q depends on the choice of coordinates, whereas Q does not.

(2) As stated in the introduction, the Hopf differential (or its Abresch–Rosenberg generalization) is not defined on Sol_3. If we apply the definition of the Hopf differential of the harmonic maps on $(\overline{\mathbb{C}}, ds^2)$, we get

$$Q = \frac{g_z \bar{g}_z}{|g^2 - \bar{g}^2|}\, dz^2,$$

but this leads to a nonsmooth differential. Because $g^2 - \bar{g}^2$ is purely imaginary on each quarter of the complex plane, the definitions are related by multiplication by i or $-i$, depending on the quarter. Thus, this “Hopf differential” is defined and holomorphic only on each of the four quarters delimited by the real and purely imaginary axes.

4. Construction of the helicoids in Sol_3

In this section we construct a one-parameter family of helicoids in Sol_3: we define a helicoid to be a minimal surface containing the x_3-axis whose intersection with every plane $\{x_3 = \text{constant}\}$ is a straight line and which is invariant by left multiplication by an element of Sol_3 of the form $(0, 0, T)$ for some $T \neq 0$.

Theorem 4. *There exists a one-parameter family* $(\mathcal{H}_K)_{K\in]-1;1[\setminus\{0\}}$ *of properly embedded minimal helicoids in* Sol_3 *having the following properties*:

(1) *For all* $K \in]-1;1[\setminus\{0\}$, *the surface* $\mathcal{H}_K$ *contains the* x_3*-axis.*

(2) *For all* $K \in]-1;1[\setminus\{0\}$, *the intersection of* $\mathcal{H}_K$ *and any horizontal plane* $\{x_3 = \lambda\}$ *is a straight line.*

(3) *For all* $K \in]-1;1[\setminus\{0\}$, *there exists* T_K *such that* $\mathcal{H}_K$ *is invariant by left multiplication by* $(0, 0, T_K)$.

(4) *The helicoids* $\mathcal{H}_K$ *have three symmetries fixing the origin*: *rotation by* π *around the* x_3*-axis*, *rotation by* π *around the* $(x, x, 0)$*-axis and rotation by* π *around the* $(x, -x, 0)$*-axis.*

Let $K \in]-1, 1[$; we define a map $g : \mathbb{C} \to \overline{\mathbb{C}}$ by

$$g(z = u + iv) = e^{-u} e^{ib(v)} e^{-i\pi/4},$$

where b satisfies the ODE

$$b' = \sqrt{1 - K\cos(2b)}, \quad b(0) = 0. \tag{8}$$

Proposition 5. *The map b is well-defined and has the following properties*:

(1) *The function b is an increasing diffeomorphism from* $\mathbb{R}$ *onto* $\mathbb{R}$.

(2) *The function b is odd.*

(3) *There exists a real number* $W > 0$ *such that*

$$\forall v \in \mathbb{R}, \quad b(v + W) = b(v) + \pi.$$

(4) *The function b satisfies* $b(kW) = k\pi$, *for all* $k \in \mathbb{Z}$.

Proof. Since $K \in]-1, 1[$, there exists $r > 0$ such that $1 - K\cos(2b) \in]r, 2[$; the Cauchy–Lipschitz theorem can be applied, and b is well-defined. By (8) we have $b' > 0$ on its domain of definition, and $\sqrt{r} < b' < 2$. Since b' is bounded by two positive constants, b is defined on $\mathbb{R}$, and

$$\lim_{v\to\pm\infty} b(v) = \pm\infty.$$

The function $\hat{b} : v \mapsto -b(-v)$ satisfies (8) with $\hat{b}(0) = 0$; hence $\hat{b} = b$ and b is odd. Finally, there exists $W > 0$ such that $b(W) = \pi$; then the function $\tilde{b} : v \mapsto b(v + W) - \pi$ satisfies (8) with $\tilde{b}(0) = 0$; hence, $\tilde{b} = b$. □

Corollary 6. *We have* $b(kW/2) = k\pi/2$ *for all* $k \in 2\mathbb{Z} + 1$.

Proof. We have

$$b\Big(\frac{W}{2}\Big) = b\Big(-\frac{W}{2} + W\Big) = -b\Big(\frac{W}{2}\Big) + \pi,$$

which gives the formula for $k = 1$, and the corollary easily follows. □

Proposition 7. *The function g satisfies $(g^2 - \bar{g}^2)g_{z\bar{z}} = 2gg_z g_{\bar{z}}$, and its Hopf differential is*

$$Q = \frac{iK}{8}\,dz^2. \tag{9}$$

Proof. A direct calculation shows that g satisfies the equation. Hence, the Hopf differential is given by

$$Q = \frac{g_z \bar{g}_z}{g^2 - \bar{g}^2}\,dz^2 = \frac{i(1-b'^2)}{8\cos(2b)}\,dz^2 = \frac{iK}{8}\,dz^2. \qquad \square$$

Thus the map g induces a conformal minimal immersion $x = (x_1, x_2, x_3)$ such that

$$x_{1z} = e^{-x_3}\frac{(\bar{g}^2-1)g_z}{g^2-\bar{g}^2} = \frac{(1+ie^{-2u}e^{-2ib})(1-b')e^{ib}e^{i\pi/4}}{4e^{-u}\cos(2b)}e^{-x_3},$$

$$x_{2z} = ie^{x_3}\frac{(\bar{g}^2+1)g_z}{g^2-\bar{g}^2} = -\frac{(1-ie^{-2u}e^{-2ib})i(1-b')e^{ib}e^{i\pi/4}}{4e^{-u}\cos(2b)}e^{x_3},$$

$$x_{3z} = \frac{2\bar{g}g_z}{g^2-\bar{g}^2} = \frac{i(b'-1)}{2\cos(2b)}.$$

This map is an immersion since the metric induced by x is given by

$$dw^2 = \|x_u\|^2|dz|^2 = \frac{K^2}{(1+b')^2}\cosh^2 u|dz|^2.$$

We obtain immediately that x_3 is a one-variable function and satisfies

$$x_3'(v) = \frac{1-b'(v)}{\cos(2b(v))} = \frac{K}{1+b'(v)}.$$

Remark. For $K = 0$, we get x_3 is constant, and the image of x is a point. In the sequel, we will always exclude this case.

By setting $x_3(0) = 0$, we choose x_3 among the primitive functions.

Proposition 8. (1) *The function x_3 is defined on $\mathbb{R}$ and is bijective.*

(2) *The function x_3 is odd.*

(3) *The function x_3 satisfies*

$$x_3(v+W) = x_3(v) + x_3(W)$$

for all real numbers v.

Proof. The map x_3 is bijective on $\mathbb{R}$ since it is a primitive of a continuous function, and its derivative has the sign of K. Since the map b is odd, b' is even, so x_3' is even and x_3 is odd. Finally, we have $x_3'(v+W) = x_3'(v)$, and the result follows. $\square$

Hence, the functions

$$x_1(u+iv) = \tfrac{\sqrt{2}}{2}(\cos b(v) - \sin b(v))x_3' e^{-x_3} \sinh u,$$
$$x_2(u+iv) = \tfrac{\sqrt{2}}{2}(\cos b(v) + \sin b(v))x_3' e^{x_3} \sinh u,$$

satisfy the equations above.

Theorem 9. *Let K be a real number such that $|K| < 1$ and $K \neq 0$, and b the function defined by* (8). *We define the function x_3 by*

$$x_3' = \frac{K}{1+b'}, \quad x_3(0) = 0.$$

Then the map

$$x : u+iv \in \mathbb{C} \longmapsto \begin{pmatrix} \frac{\sqrt{2}}{2}(\cos b(v) - \sin b(v))x_3' e^{-x_3} \sinh u \\ \frac{\sqrt{2}}{2}(\cos b(v) + \sin b(v))x_3' e^{x_3} \sinh u \\ x_3(v) \end{pmatrix}$$

is a conformal minimal immersion whose Gauss map is

$$g : u+iv \in \mathbb{C} \longmapsto e^{-u} e^{ib(v)} e^{-i\pi/4}.$$

Moreover,

$$(0, 0, 2x_3(W)) * x(u+iv) = x(u+i(v+2W)) \tag{10}$$

for all $u, v \in \mathbb{R}$. The surface given by x is called a helicoid *of parameter K and will be denoted by $\mathcal{H}_K$.*

Proof. Equation (10) means that the helicoid is invariant by left multiplication by $(0, 0, 2x_3(W))$. Recall that we have the identity

$$x_3(v+2W) = x_3(v+W) + x_3(W) = x_3(v) + 2x_3(W)$$

for all real numbers v. Thus we get the result for the third coordinate and we prove in the same way that $e^{-2x_3(W)}x_1(u+iv) = x_1(u+i(v+2W))$ and $e^{2x_3(W)}x_2(u+iv) = x_2(u+i(v+2W))$. □

Remark. (1) The surface $\mathcal{H}_K$ is embedded because x_3 is bijective. It is easy to see that it is even properly embedded.

(2) The surfaces $\mathcal{H}_K$ and $\mathcal{H}_{-K}$ are related; if we denote by the indices K and K' the data describing $\mathcal{H}_K$ and $\mathcal{H}_{-K}$, we get

$$\begin{cases} b_{-K}(v) = b_K(v+W/2) - \pi/2, \\ x_{3\,-K}(v) = -x_{3\,K}(v+W/2) + x_{3\,K}(W/2). \end{cases}$$

In particular, ${x_3}_{-K}(W) = -{x_3}_K(W)$ and both surfaces have the same period $|{x_3}_K(W)|$. Finally,

$$x_{-K}(u+iv) = (0,0,{x_3}_K(W/2)) * \sigma^3 x_K(u+i(v+W/2)).$$

Thus, there exists an isometry of Sol_3 that puts $\mathcal{H}_{-K}$ on $\mathcal{H}_K$.

Proposition 10. *For every real number T, there exists a unique helicoid $\mathcal{H}_K$ (up to isometry, i.e., up to $K \leftarrow\!\!\dashv -K$) whose period is T.*

Proof. We noticed that the period of the helicoid $\mathcal{H}_K$ is

$$\begin{aligned} 2x_3(W) := 2{x_3}_K(W) &= 2\int_0^W \frac{K}{1+b'(s)}\,ds \\ &= 2K\int_0^\pi \frac{du}{\sqrt{1-K\cos(2u)}(1+\sqrt{1-K\cos(2u)})}, \end{aligned}$$

with the change of variables $u = b(s)$ and $b(W) = \pi$. Seeing ${x_3}_K(W)$ as a function of the variable K, we get

$$\frac{\partial {x_3}_K(W)}{\partial K} = \int_0^\pi \frac{1}{(1-K\cos(2u))^{3/2}}\,du$$

(valid for K in every compact set $[0,a] \subset [0,1[$, and so in $[0,1[$). Then the function $K \mapsto {x_3}_K(W)$ is injective. Moreover, we have ${x_3}_0(W) = 0$ and

$$\begin{aligned} {x_3}_1(W) &= \int_0^\pi \frac{1}{\sqrt{1-\cos(2u)}(1+\sqrt{1-\cos(2u)})}\,du \\ &= \int_0^\pi \frac{1}{\sqrt{2}\sin u(1+\sqrt{2}\sin u)}\,du \\ &= \frac{1}{\sqrt{2}}\int_0^\infty \frac{1+t^2}{1+2\sqrt{2}t+t^2}\,dt = +\infty, \end{aligned}$$

so ${x_3}_K(W)$ is a bijection from $]0,1[$ onto $]0,+\infty[$. □

The vector field defined by

$$\begin{aligned} N &= \frac{1}{1+|g|^2}\begin{bmatrix} 2\Re(g) \\ 2\Im(g) \\ 1-|g|^2 \end{bmatrix} \\ &= \frac{\sqrt{2}}{2\cosh u}\begin{bmatrix} \cos b + \sin b \\ \sin b - \cos b \\ \sqrt{2}\sinh u \end{bmatrix} \end{aligned}$$

is normal to the surface. We get

$$\nabla_{x_u} N = -\sin(2b)\frac{\sinh u}{\cosh u}x_u + \left(\frac{1+b'}{K\cosh^2 u} - \cos(2b)\right)x_v,$$
$$\nabla_{x_v} N = \left(\frac{1+b'}{K\cosh^2 u} - \cos(2b)\right)x_u + \sin(2b)\frac{\sinh u}{\cosh u}x_v,$$

and thus the Gauss curvature is given by

$$\mathcal{K} = -1 + \frac{1}{\cosh^2 u}\left(\frac{2(1+b')\cos(2b)}{K} - \frac{(1+b')^2}{K^2\cosh^2 u} + \sin^2(2b)\right).$$

In particular, the fundamental pieces of the helicoids have infinite total curvature since

$$\mathcal{K} dA = \left(-\frac{K^2}{(1+b')^2}\cosh^2 u + \frac{2K\cos(2b)}{1+b'} - \frac{1}{\cosh^2 u} + \frac{K^2\sin^2(2b)}{(1+b')^2}\right) du\, dv.$$

We notice that

$$x(-u+iv) = \begin{pmatrix} -x_1(u+iv) \\ -x_2(u+iv) \\ x_3(v) \end{pmatrix} = \sigma^2 x(u+iv),$$

where σ and τ are the isometries introduced in the first section: the helicoid $\mathcal{H}_K$ is symmetric by rotation by π around the x_3-axis, which is included in the helicoid as the image by x of the purely imaginary axis of $\mathbb{C}$. On this axis we have

$$g(0+iv) = -ie^{ib(v)}.$$

Hence, the straight line $\{(x,x,0) \mid x \in \mathbb{R}\}$ is included in the helicoid as the image by x of the real line. Along this line, we have

$$g(u+i0) = e^{-u}e^{-i\pi/4}.$$

Then we notice that

$$x(u-iv) = \begin{pmatrix} x_2(u+iv) \\ x_1(u+iv) \\ -x_3(v) \end{pmatrix} = \sigma\tau x(u+iv).$$

Thus, $\mathcal{H}_K$ is symmetric by rotation by π around the axis $\{(x,x,0) \mid x \in \mathbb{R}\}$.

Remark. The straight line $\{(x,x,0) \mid x \in \mathbb{R}\}$ is a geodesic of the helicoid. It's even a geodesic of Sol_3.

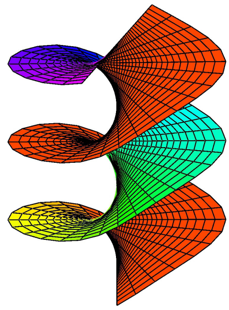

Figure 1. Helicoid for $K = 0.4$, created with Scilab.

Since the function sinh is odd, we deduce that

$$x(-u-iv) = \begin{pmatrix} -x_2(u+iv) \\ -x_1(u+iv) \\ -x_3(v) \end{pmatrix} = \sigma^3 \tau x(u+iv).$$

Thus, $\mathcal{H}_K$ is symmetric by rotation by π around the axis $\{(x,-x,0) \mid x \in \mathbb{R}\}$ (but this axis is not included in the surface).

The helicoid $\mathcal{H}_K$ has no more symmetry fixing the origin; indeed if it did, there would exist a diffeomorphism ϕ of $\mathbb{C}$ such that $x \circ \phi = \sigma^2 \circ x$ (we choose σ^2 as an example, but it is the same idea for the other elements of the isotropy group of the origin of Sol_3). By composition, the surface would have every symmetry of the isotropy group. But if $x \circ \phi = \tau x$, the decomposition $\phi = \phi_1 + i\phi_2$ leads to

$$\begin{pmatrix} x_1\big(\phi_1(u+iv) + i\phi_2(u+iv)\big) \\ x_2\big(\phi_1(u+iv) + i\phi_2(u+iv)\big) \\ x_3(\phi_2(u+iv)) \end{pmatrix} = \begin{pmatrix} -x_1(u+iv) \\ x_2(u+iv) \\ x_3(v) \end{pmatrix}.$$

Because x_3 is bijective, we get $\phi_2(u+iv) = v$ for all u, v, and then we get at the same time $\sinh(\phi_1(u+iv)) = \sinh u$ and $\sinh(\phi_1(u+iv)) = -\sinh u$, which is impossible.

5. Catenoids in Sol_3

In this section we construct examples of minimal annuli in Sol_3. Let $\alpha \in\ ]-1;1[$. We start from a map g defined on $\mathbb{C}$ by

$$g(z = u+iv) = -ie^{-u-\gamma(v)}e^{i\rho(v)},$$

where ρ satisfies the ODE

$$\rho' = \sqrt{1-\alpha^2 \sin^2(2\rho)}, \quad \rho(0) = 0, \tag{11}$$

and γ is defined by

$$\gamma' = -\alpha \sin(2\rho), \quad \gamma(0) = 0. \tag{12}$$

Proposition 11. *The map ρ is well-defined and has the following properties*:

(1) *The function ρ is an increasing diffeomorphism from $\mathbb{R}$ onto $\mathbb{R}$.*

(2) *The function ρ is odd.*

(3) *There exists a real number $V > 0$ such that*

$$\forall v \in \mathbb{R}, \quad \rho(v + V) = \rho(v) + \pi.$$

(4) *The function ρ satisfies $\rho(kV) = k\pi$ for all $k \in \mathbb{Z}$.*

Proof. Since $\alpha \in \,]-1, 1[$, there exists $r > 0$ such that $1 - \alpha^2 \sin^2(2\rho) \in \,]r, 1]$; the Cauchy–Lipschitz theorem can be applied, and ρ is well-defined. By (11) we have $\rho' > 0$ on its domain of definition, and $\sqrt{r} < \rho' < 1$. Since ρ' is bounded by two positive constants, ρ is defined on $\mathbb{R}$, and

$$\lim_{v \to \pm\infty} \rho(v) = \pm\infty.$$

The function $\hat{\rho} : v \mapsto -\rho(-v)$ satisfies (11) with $\hat{\rho}(0) = 0$; hence $\hat{\rho} = \rho$ and ρ is odd. Finally, there exists $V > 0$ such that $\rho(V) = \pi$; Then the function $\tilde{\rho} : v \mapsto \rho(v + V) - \pi$ satisfies (11) with $\tilde{\rho}(0) = 0$; hence $\tilde{\rho} = \rho$. ☐

Corollary 12. (1) *We have $\rho(kV/2) = k\pi/2$ for all $k \in 2\mathbb{Z} + 1$.*

(2) *We have $\rho(-v + V/2) = -\rho(v) + \frac{\pi}{2}$ for all $v \in \mathbb{R}$. In particular, $\rho(V/4) = \frac{\pi}{4}$ and $\rho(3V/4) = \frac{3\pi}{4}$.*

Proof. (1) We have

$$\rho\Big(\frac{V}{2}\Big) = \rho\Big(-\frac{V}{2} + V\Big) = -\rho\Big(\frac{V}{2}\Big) + \pi,$$

which gives the formula for $k = 1$, and part (1) easily follows.

(2) The functions $\rho^* : v \mapsto \pi/2 - \rho(-v + V/2)$ and ρ satisfy equation (11) with $\rho^*(0) = \rho(0) = 0$, so $\rho^* = \rho$ and

$$\rho(V/4) = \rho^*(V/4) = \tfrac{\pi}{2} - \rho\big(\tfrac{\pi}{2} - \tfrac{\pi}{4}\big),$$

and the result follows. ☐

Proposition 13. *The function g satisfies $(g^2 - \bar{g}^2) g_{z\bar{z}} = 2 g g_z g_{\bar{z}}$, and its Hopf differential is*

$$Q = -\frac{\alpha}{4}\, dz^2. \tag{13}$$

Proof. A direct calculation shows that g satisfies the equation. Hence, the Hopf differential is given by

$$\widetilde{Q} = \frac{i(1-\rho'^2-\gamma'^2-2i\gamma')}{8\sin(2\rho)}\,dz^2 = -\frac{\alpha}{4}\,dz^2. \qquad \square$$

Thus the map g induces a conformal minimal immersion $x = (x_1, x_2, x_3)$ such that

$$x_{1z} = e^{-x_3}\frac{(\bar{g}^2-1)g_z}{g^2-\bar{g}^2}, \quad x_{2z} = i\,e^{x_3}\frac{(\bar{g}^2+1)g_z}{g^2-\bar{g}^2}, \quad x_{3z} = \frac{2\bar{g}g_z}{g^2-\bar{g}^2}.$$

This application is an immersion since the metric induced by x is given by

$$\begin{aligned} dw^2 &= \|x_u\|^2\,|dz|^2 \\ &= (F'^2+\alpha^2)\cosh^2(u+\gamma)|dz|^2 \\ &= \left(\frac{\alpha^4\sin^2(2\rho)}{(1+\rho')^2}+\alpha^2\right)\cosh^2(u+\gamma)|dz|^2 \\ &= \frac{2\alpha^2}{1+\rho'}\cosh^2(u+\gamma)|dz|^2. \end{aligned}$$

In particular,

$$x_{3z} = \frac{i\rho'-\gamma'-i}{2\sin(2\rho)},$$

that is,

$$\begin{cases} x_{3u} = 2\Re(x_{3z}) = -\dfrac{\gamma'}{\sin(2\rho)} = \alpha, \\[2ex] x_{3v} = -2\Im(x_{3z}) = \dfrac{1-\rho'}{\sin(2\rho)} = \dfrac{\alpha^2\sin(2\rho)}{1+\rho'}. \end{cases}$$

Thus

$$x_3(u+iv) = \alpha u + \alpha^2\int^v \frac{\sin(2\rho(s))}{1+\rho'(s)}\,ds.$$

Here we have to choose an initial condition; we set

$$F(v) = \alpha^2\int_0^v \frac{\sin(2\rho(s))}{1+\rho'(s)}\,ds,$$

and define

$$x_3(u+iv) = \alpha u + F(v).$$

The function F is well-defined on $\mathbb{R}$.

Proposition 14. *The function F is even and V-periodic.*

Proof. The function F' is odd because ρ is odd and ρ' is even, so F is even. Then we get

$$F'(v+V) = \alpha^2 \frac{\sin(2\rho(v)+2\pi)}{1+\rho'(v)} = F'(v),$$

so there exists a constant C such that $F(v+V) = F(v) + C$ for all $v \in \mathbb{R}$. By evaluating at zero, we get $C = F(V)$, that is,

$$\begin{aligned} C &= \alpha^2 \int_0^V \frac{\sin(2\rho(s))}{1+\rho'(s)}\, ds = \alpha^2 \int_0^V H(s)\, ds \\ &= \alpha^2 \left\{ \int_0^{V/4} H(s)\, ds + \int_{V/4}^{3V/4} H(s)\, ds + \int_{3V/4}^{V} H(s)\, ds \right\} \\ &:= \sum_{k=0}^{2} L_k(\alpha). \end{aligned}$$

We can now do the change of variable $u = \sin(2\rho(s))$ in each integral $L_k(\alpha)$, with

$$du = 2\rho'(s)\cos(2\rho(s))\, ds = 2(-1)^k \sqrt{(1-\alpha^2u^2)(1-u^2)}\, ds.$$

Thus,

$$C = \alpha^2 \int_{-1}^{1} \frac{u\, du}{(1+\sqrt{1-\alpha^2u^2})\sqrt{(1-\alpha^2u^2)(1-u^2)}} = 0$$

and F is V-periodic. $\square$

Proposition 15. *The function γ is even and V-periodic.*

Proof. We prove the proposition in exactly the same way as for the function F. $\square$

The two other equations become

$$\begin{aligned} x_{1z} &= e^{-x_3} \frac{(e^{-u-\gamma-i\rho} + e^{u+\gamma+i\rho})(1-\rho'-i\gamma')}{4\sin(2\rho)}, \\ x_{2z} &= -e^{x_3} \frac{(e^{u+\gamma+i\rho} - e^{-u-\gamma-i\rho})(i - i\rho' + \gamma')}{4\sin(2\rho)}. \end{aligned}$$

Those equations lead to

$$\begin{aligned} x_1 &= e^{-\alpha u - F}\left(\frac{e^{u+\gamma}}{2(1-\alpha)}(F'\cos\rho - \alpha\sin\rho) - \frac{e^{-u-\gamma}}{2(1+\alpha)}(\alpha\sin\rho + F'\cos\rho) \right), \\ x_2 &= e^{\alpha u + F}\left(\frac{-e^{u+\gamma}}{2(1+\alpha)}(\alpha\cos\rho + F'\sin\rho) + \frac{e^{-u-\gamma}}{2(\alpha-1)}(\alpha\cos\rho - F'\sin\rho) \right). \end{aligned}$$

Remark. If $\alpha = 0$, then $x(\mathbb{C}) = \{0\}$. This case will be excluded in the sequel.

Theorem 16. *Let α be a real number such that $|\alpha| < 1$ and $\alpha \neq 0$, and ρ and γ the functions defined by* (11) *and* (12). *We define the function F by*

$$F(v) = \alpha^2 \int_0^v \frac{\sin(2\rho(s))}{1+\rho'(s)}\, ds.$$

Then the map $x : \mathbb{C} \to \mathrm{Sol}_3$ defined by

$$\begin{pmatrix} e^{-\alpha u - F}\left(\dfrac{e^{u+\gamma}}{2(1-\alpha)}(F'\cos\rho - \alpha\sin\rho) - \dfrac{e^{-u-\gamma}}{2(1+\alpha)}(\alpha\sin\rho + F'\cos\rho)\right) \\ e^{\alpha u + F}\left(\dfrac{-e^{u+\gamma}}{2(1+\alpha)}(\alpha\cos\rho + F'\sin\rho) + \dfrac{e^{-u-\gamma}}{2(\alpha-1)}(\alpha\cos\rho - F'\sin\rho)\right) \\ \alpha u + F \end{pmatrix}$$

is a conformal minimal immersion whose Gauss map is

$$g : u + iv \in \mathbb{C} \longmapsto -i\, e^{-u-\gamma(v)} e^{i\rho(v)}.$$

Moreover,

$$x(u + i(v + 2V)) = x(u + iv) \tag{14}$$

for all $u, v \in \mathbb{R}$. The surface given by x is called a catenoid *of parameter α and will be denoted by $\mathcal{C}_\alpha$.*

Proof. The periodicity of $\mathcal{C}_\alpha$ is an application of Propositions 11, 14 and 15. □

Remark. The surfaces $\mathcal{C}_\alpha$ and $\mathcal{C}_{-\alpha}$ are related; if we denote by the indices α and $-\alpha$ the data describing $\mathcal{C}_\alpha$ and $\mathcal{C}_{-\alpha}$, we get

$$\begin{cases} \rho_{-\alpha} = \rho_\alpha, \\ F_{-\alpha} = F_\alpha, \\ \gamma_{-\alpha} = -\gamma_\alpha. \end{cases}$$

Thus, we get

$$x_{-\alpha}(-u + iv) = \sigma^2 x_\alpha(u + iv).$$

In particular, there exists an orientation-preserving isometry of Sol_3 fixing the origin that sends $\mathcal{C}_\alpha$ on $\mathcal{C}_{-\alpha}$.

Now we show that the catenoids are embedded:

Proposition 17. *For all $\lambda \in \mathbb{R}$, the intersection of $\mathcal{C}_\alpha$ with the plane $\{x_3 = \lambda\}$ is a nonempty closed embedded convex curve.*

Proof. This intersection is nonempty: $x(\lambda/\alpha + i0) \in \mathcal{C}_\alpha \cap \{x_3 = \lambda\}$. We look at the curve in $\mathbb{C}$ defined by $x_3(u + iv) = \alpha u + F(v) = \lambda$, i.e., the curve

$$c : t \in \mathbb{R} \longmapsto \left(\frac{\lambda - F(t)}{\alpha}, t\right).$$

Its image by x is

$$\mathfrak{c} : t \in \mathbb{R} \longmapsto \begin{pmatrix} e^{-\lambda}\left(\dfrac{e^{\delta+\gamma}}{2(1-\alpha)}(F' \cos\rho - \alpha \sin\rho) - \dfrac{e^{-\delta-\gamma}}{2(1+\alpha)}(\alpha \sin\rho + F' \cos\rho)\right) \\ e^{\lambda}\left(\dfrac{-e^{\delta+\gamma}}{2(1+\alpha)}(\alpha \cos\rho + F' \sin\rho) + \dfrac{e^{-\delta-\gamma}}{2(\alpha-1)}(\alpha \cos\rho - F' \sin\rho)\right) \\ c \end{pmatrix}(t),$$

where $\delta = \frac{\lambda - F}{\alpha}$. The calculation leads to

$$\mathfrak{c}_1'(t) = \frac{e^{-\lambda}}{\alpha(1-\alpha^2)}\left(A(t) \cosh\left(\frac{\lambda - F}{\alpha} + \gamma\right) + B(t) \sinh\left(\frac{\lambda - F}{\alpha} + \gamma\right)\right),$$

with

$$\begin{cases} A = -F'^2 \cos\rho + \alpha\gamma' F' \cos\rho - \alpha^2 \rho' \cos\rho + \alpha^2 F' \sin\rho - \alpha^3 \gamma' \sin\rho \\ \qquad\qquad + \alpha^2 F'' \cos\rho - \alpha^2 F' \rho' \sin\rho, \\ B = \alpha F' \sin\rho - \alpha^2 \gamma' \sin\rho + \alpha F'' \cos\rho - \alpha F' \rho' \sin\rho - \alpha F'^2 \cos\rho \\ \qquad\qquad + \alpha^2 \gamma' F' \cos\rho - \alpha^3 \rho' \cos\rho. \end{cases}$$

We remark that $B \equiv 0$ after simplifications, and

$$A(t) = (F'^2(t) + \alpha^2)(\alpha^2 - 1) \cos\rho(t).$$

Finally,

$$\mathfrak{c}_1'(t) = -\frac{e^{-\lambda}}{\alpha}(F'^2(t) + \alpha^2) \cos\rho(t) \cosh\left(\frac{\lambda - F(t)}{\alpha} + \gamma(t)\right).$$

In the same way, we get

$$\mathfrak{c}_2'(t) = -\frac{e^{-\lambda}}{\alpha}(F'^2(t) + \alpha^2) \sin\rho(t) \cosh\left(\frac{\lambda - F(t)}{\alpha} + \gamma(t)\right).$$

Thus

$$\mathfrak{c}_1'^2 + \mathfrak{c}_2'^2 = \frac{e^{-2\lambda}}{\alpha^2}(F'^2(t) + \alpha^2)^2 \cosh^2\left(\frac{\lambda - F(t)}{\alpha} + \gamma(t)\right) > 0,$$

so the intersection $\mathcal{C}_\alpha \cap \{x_3 = \lambda\}$ is a smooth curve; moreover, it's closed since $\mathfrak{c}(t + 2V) = \mathfrak{c}(t)$ for all $t \in \mathbb{R}$.

The planes $\{x_3 = \lambda\}$ are flat: indeed, the metrics on these planes are $e^{2\lambda} dx_1^2 + e^{-2\lambda} dx_2^2$, so up to an affine transformation, we can work in euclidean coordinates, as we suppose in this proof since affinities preserve convexity.

To prove that $\mathfrak{c}$ is embedded and convex, we consider the part of $\mathfrak{c}$ corresponding to $t \in (-V/2, V/2)$. On $(-V/2, V/2)$, we have $\cos\rho(t) > 0$, thanks to Proposition 11

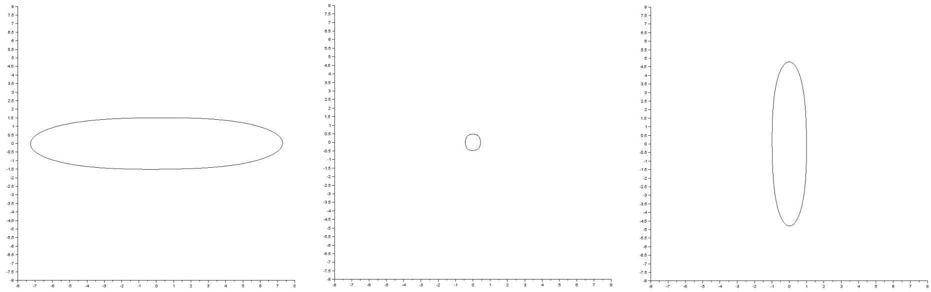

Figure 2. Sections with $\{x_3 = -1\}$, $\{x_3 = 0\}$, and $\{x_3 = 1\}$, created with Scilab.

and Corollary 12. So $\mathfrak{c}_1'(t) < 0$ if $\alpha > 0$ (and $\mathfrak{c}_1'(t) > 0$ if $\alpha < 0$) and $\mathfrak{c}_1$ is injective and decreasing if $\alpha > 0$ (and increasing if $\alpha < 0$). We get

$$\frac{d\mathfrak{c}_2}{d\mathfrak{c}_1} = \tan\rho(t),$$

so $d\mathfrak{c}_2/d\mathfrak{c}_1$ is an increasing function of t, and also of $\mathfrak{c}_1$ if $\alpha < 0$ (and a decreasing function of the decreasing function $\mathfrak{c}_1$ if $\alpha > 0$). In both cases, the curve is convex.

Then, the half of $\mathfrak{c}$ corresponding to $t \in (-V/2, V/2)$ is convex and embedded. Since $\mathfrak{c}(t+V) = -\mathfrak{c}(t)$, the entire curve is convex and embedded. □

Figure 2 shows sections of the catenoid $\alpha = -0.6$ with planes $\{x_3 = \text{constant}\}$.

Corollary 18. *The surface $\mathcal{C}_\alpha$ is properly embedded for all $\alpha \in \,]-1, 1[\setminus\{0\}$.*

Proposition 19. *For all $\alpha \in \,]-1, 1[\setminus\{0\}$, the surface $\mathcal{C}_\alpha$ is conformally equivalent to $\mathbb{C}\setminus\{0\}$.*

Proof. The map $x : \mathbb{C}/(2iV\mathbb{Z}) \to \mathcal{C}_\alpha$ is a conformal bijective parametrization of $\mathcal{C}_\alpha$. □

The vector field defined by

$$N = \frac{1}{\cosh u}\begin{bmatrix} e^{-\gamma}\sin\rho \\ -e^{-\gamma}\cos\rho \\ \sinh u \end{bmatrix}$$

is normal to the surface.

We have

$$x(u+i(v+V)) = \begin{pmatrix} -x_1(u+iv) \\ -x_2(u+iv) \\ x_3(u+iv) \end{pmatrix} = \sigma^2 x(u+iv).$$

Thus, the surface $\mathcal{C}_\alpha$ is symmetric by rotation by π around the x_3-axis.

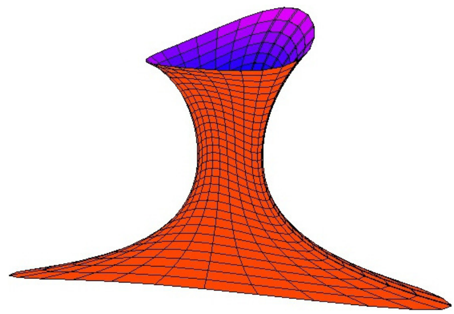

Figure 3. Catenoid for $\alpha = -0.6$, created with Scilab.

Remark. The x_3-axis is contained in the "interior" of $\mathcal{C}_\alpha$ since each curve $\mathcal{C}_\alpha \cap \{x_3 = \lambda\}$ is convex and symmetric with respect to this axis.

We also get

$$x(u-iv) = \begin{pmatrix} -x_1(u+iv) \\ x_2(u+iv) \\ x_3(u+iv) \end{pmatrix} = \tau x(u+iv),$$

and the surface $\mathcal{C}_\alpha$ is symmetric by reflection in the plane $\{x_1 = 0\}$, and finally we have

$$x(u+i(-v+V)) = \sigma^2 \tau x(u+iv),$$

and $\mathcal{C}_\alpha$ is symmetric by reflection in the plane $\{x_2 = 0\}$.

If $\mathcal{C}_\alpha$ had another symmetry fixing the origin, it would have every symmetry of the isotropy group of Sol$_3$, and we prove as for the helicoid that it is impossible.

6. Limits of catenoids

6.1. ***The case $\alpha = 0$.*** In this part we consider the limit surface of the catenoids $\mathcal{C}_\alpha$ when α goes to zero. For this, we do the change of parameters

$$\begin{cases} u' = u + \ln\alpha, \\ v' = v. \end{cases}$$

In these coordinates, the immersion x given in Theorem 16 takes the form

$$\begin{pmatrix} e^{\alpha\ln\alpha-\alpha u'-F}\left(\dfrac{e^{u'+\gamma}}{2\alpha(1-\alpha)}(\cos\rho F'-\alpha\sin\rho) - \dfrac{\alpha e^{-u'-\gamma}}{2(1+\alpha)}(\alpha\sin\rho+\cos\rho F')\right) \\ e^{-\alpha\ln\alpha+\alpha u'+F}\left(\dfrac{-e^{u'+\gamma}}{2\alpha(1+\alpha)}(\alpha\cos\rho+F'\sin\rho) + \dfrac{\alpha e^{-u'-\gamma}}{2(\alpha-1)}(\alpha\cos\rho-F'\sin\rho)\right) \\ -\alpha\ln\alpha+\alpha u'+F \end{pmatrix}.$$

Letting α go to zero, we get

$$\begin{cases} \rho \longrightarrow \mathrm{Id}, \\ F/\alpha \longrightarrow 0, \\ F'/\alpha \longrightarrow 0, \\ \gamma \longrightarrow 0, \end{cases}$$

and so the limit immersion is

$$\begin{pmatrix} -\dfrac{e^{u'}}{2}\sin v' \\ -\dfrac{e^{u'}}{2}\cos v' \\ 0 \end{pmatrix}.$$

Thus, we obtain a parametrization of the plane $\{x_3 = 0\}$, which is the limit of the family $(\mathcal{C}_\alpha)$ when $\alpha \to 0$.

6.2. *The case $\alpha = 1$.* We end by the study of the case $\alpha = 1$ (the case $\alpha = -1$ is exactly the same). We show that the limit surface is a minimal entire graph:

Proposition 20. *Let $\boldsymbol{x} : \mathbb{R}^2 \to \mathrm{Sol}_3$ be defined by*

$$\boldsymbol{x}(u + iv) = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \end{pmatrix} = \begin{pmatrix} -\dfrac{\tanh v}{2}(1 + e^{-2u}) \\ \dfrac{e^{2u}}{4} - \dfrac{u}{2} - \dfrac{\cosh(2v)}{4} \\ u + \ln(\cosh v) \end{pmatrix}.$$

Then $\boldsymbol{x}$ is a minimal immersion and there exists a C^∞-function f defined on $\mathbb{R}^2$ such that the image of x (called $\mathcal{S}$) is the $\boldsymbol{x}_2$-graph of f given by $\boldsymbol{x}_2 = f(\boldsymbol{x}_1, \boldsymbol{x}_3)$.

Proof. We show that this surface is (up to a translation) the limit surface of the family $(\mathcal{C}_\alpha)_{\alpha \in]-1,1[}$ when α goes to 1. For $\alpha = 1$, the Gauss map is still given by $g(z = u + iv) = -ie^{-u-\gamma(v)}e^{i\rho(v)}$, but ρ satisfies the ODE

$$\rho' = \cos(2\rho), \quad \rho(0) = 0, \tag{15}$$

and γ is still defined by

$$\gamma' = -\sin(2\rho), \quad \gamma(0) = 0. \tag{16}$$

We have explicit expressions for these functions, which are given by

$$\begin{aligned} \rho(v) &= \arctan e^{2v} - \pi/4 = \arctan(\tanh v), \\ \gamma(v) &= -\tfrac{1}{2}\ln(\cosh(2v)). \end{aligned}$$

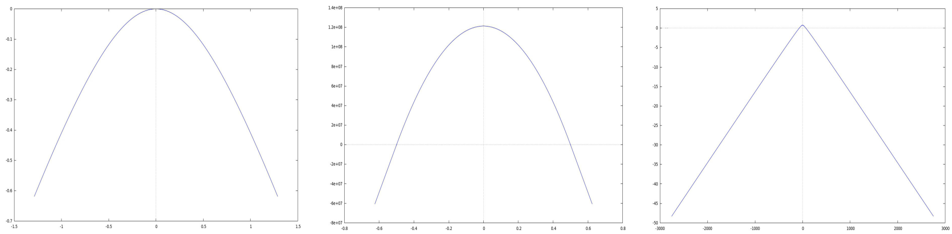

Figure 4. Sections with $\{x_3 = 0\}$, $\{x_3 = 10\}$, and $\{x_3 = -2\}$, created with Maxima.

Thus by setting

$$F(v) = \int_0^v \frac{\sin(2\rho(s))}{1+\cos(2\rho(s))}\,ds,$$

we obtain $F(v) = \ln(\cosh v)$. Then, the immersion x is given by

$$x = \begin{pmatrix} -\dfrac{e^{-2u}}{2}\tanh v + \dfrac{e^{-v}}{2\cosh v} \\ \dfrac{e^{2u}}{4} - \dfrac{u}{2} - \dfrac{\cosh(2v)}{4} \\ u + \ln(\cosh v) \end{pmatrix}.$$

A unit normal vector field is given by

$$N = \frac{1}{1+e^{-2u}\cosh(2v)} \begin{bmatrix} 2e^{-u}\sinh v \\ -2e^{-u}\cosh v \\ 1-e^{-2u}\cosh(2v) \end{bmatrix}.$$

Thus, we get

$$g(u+iv) = -ie^{-u}(\cosh v + i\sinh v),$$

which satisfies the harmonic equation (6). The metric induced by this immersion on the surface is

$$ds^2 = (e^{-4u}\tanh^2 v + e^{2u}\sinh^2 u + 1)\,|dz|^2.$$

This surface is symmetric by reflection in the plane $\{x_1 = 1/2\}$ since

$$x(u+iv) = \begin{pmatrix} \dfrac{1}{2} - \dfrac{\tanh v}{2}(1+e^{-2u}) \\ \dfrac{e^{2u}}{4} - \dfrac{u}{2} - \dfrac{\cosh(2v)}{4} \\ u + \ln(\cosh v) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + \tilde{x}_1(u,v) \\ x_2(u,v) \\ x_3(u,v) \end{pmatrix},$$

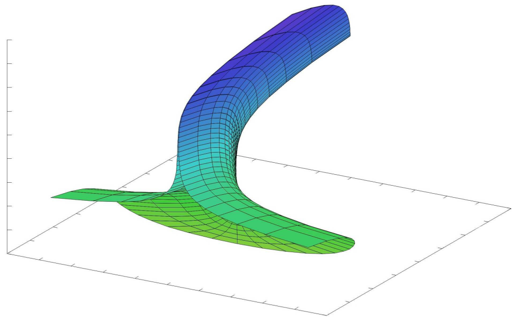

Figure 5. The surface $\mathcal{S}$, created with Maxima.

and so

$$x(u-iv)=\begin{pmatrix}\frac{1}{2}-\tilde{x}_1(u,v)\\ x_2(u,v)\\ x_3(u,v)\end{pmatrix}.$$

This property is equivalent to the property that the translated surface $(-1/2,0,0)*x(u+iv)$ is symmetric with respect to $\{x_1=0\}$. This translated surface is the image of the immersion $\boldsymbol{x}$ defined by

$$\boldsymbol{x}(u+iv)=(-1/2,0,0)*x(u+iv)=\begin{pmatrix}-\dfrac{\tanh v}{2}(1+e^{-2u})\\ \dfrac{e^{2u}}{4}-\dfrac{u}{2}-\dfrac{\cosh(2v)}{4}\\ u+\ln(\cosh v)\end{pmatrix}.$$

Then, this surface is analytic (like any minimal surface in Sol_3), so it is a local analytic $\boldsymbol{x}_2$-graph around every point where ∂_2 doesn't belong to the tangent plane, i.e., $\langle N,\partial_2\rangle\neq 0$. But

$$\langle N,\partial_2\rangle=0\iff \cosh v e^{-u}=0,$$

which is impossible. Thus, $\mathcal{S}$ is a local analytic $\boldsymbol{x}_2$-graph around every point. Then, we consider sections of the surface $\mathcal{S}$ with planes $\{\boldsymbol{x}_3=\text{constant}\}$: on the plane $\{\boldsymbol{x}_3=\lambda\}$, we get the curve

$$c_\lambda(t)=\begin{pmatrix}-\dfrac{\tanh t}{2}(1+e^{-2\lambda}\cosh^2 t)\\ \dfrac{e^{2\lambda}}{4\cosh^2 t}-\dfrac{\lambda}{2}+\dfrac{\ln(\cosh t)}{2}-\dfrac{\cosh(2t)}{4}\end{pmatrix}:=\begin{pmatrix}\boldsymbol{x}_{1\lambda}(t)\\ \boldsymbol{x}_{2\lambda}(t)\end{pmatrix}.$$

Then,

$$\boldsymbol{x}_{1\lambda}'(t)=\frac{\tanh^2 t-1}{2}-\frac{e^{-2\lambda}}{2}(\cosh^2 t+\sinh^2 t)<0$$

for all $t \in \mathbb{R}$. Thus, the curves are injective, so the surface $\mathcal{S}$ is embedded. Moreover, by the implicit function theorem, we deduce that for every $\lambda \in \mathbb{R}$, there exists a function f_λ such that $\boldsymbol{x}_{2\lambda} = f_\lambda(\boldsymbol{x}_{1\lambda})$. Because the function $\boldsymbol{x}_{1\lambda}$ is a decreasing diffeomorphism of $\mathbb{R}$, the function f_λ is defined on $\mathbb{R}$. This calculus is valid for all $\lambda \in \mathbb{R}$, so there exists a function $f : \mathbb{R}^2 \to \mathbb{R}$ such that $\boldsymbol{x}_2 = f(\boldsymbol{x}_1, \boldsymbol{x}_3)$.

Finally, this function f coincides around every point with the local C^∞-functions which give the local graphs, and so f is C^∞. □

As a conclusion, we can notice that, for a fixed $\boldsymbol{x}_3$,

- when $\boldsymbol{x}_1 \to +\infty$, $\boldsymbol{x}_2 \approx -\boldsymbol{x}_1 e^{2\boldsymbol{x}_3}$;
- when $\boldsymbol{x}_1 \to -\infty$, $\boldsymbol{x}_2 \approx \boldsymbol{x}_1 e^{2\boldsymbol{x}_3}$.

References

[Abresch and Rosenberg 2004] U. Abresch and H. Rosenberg, "A Hopf differential for constant mean curvature surfaces in $\mathbf{S}^2 \times \mathbb{R}$ and $\mathbf{H}^2 \times \mathbb{R}$", *Acta Math.* **193**:2 (2004), 141–174. MR 2006h:53003 Zbl 1078.53053

[Abresch and Rosenberg 2005] U. Abresch and H. Rosenberg, "Generalized Hopf differentials", *Mat. Contemp.* **28** (2005), 1–28. MR 2006h:53004 Zbl 1118.53036

[Berdinskii and Taimanov 2005] D. A. Berdinskii and I. A. Taimanov, "Поверхности в трехмерных пруппах Ли", *Sibirsk. Mat. Zh.* **46**:6 (2005), 1248–1264. Translated as "Surfaces in three-dimensional Lie groups" in *Sib. Math. J.* **46**:6 (2005), 1005–1019. MR 2006j:53087 Zbl 1117.53045

[Daniel and Hauswirth 2009] B. Daniel and L. Hauswirth, "Half-space theorem, embedded minimal annuli and minimal graphs in the Heisenberg group", *Proc. Lond. Math. Soc.* (3) **98**:2 (2009), 445–470. MR 2009m:53018 Zbl 1163.53036

[Daniel and Mira 2013] B. Daniel and P. Mira, "Existence and uniqueness of constant mean curvature spheres in Sol$_3$", *J. Reine Angew. Math.* **685** (2013), 1–32. MR 3181562 Zbl 06245199

[Inoguchi and Lee 2008] J.-I. Inoguchi and S. Lee, "A Weierstrass type representation for minimal surfaces in Sol", *Proc. Amer. Math. Soc.* **136**:6 (2008), 2209–2216. MR 2008k:53016 Zbl 1157.53007

[López 2014] R. López, "Invariant surfaces in Sol$_3$ with constant mean curvature and their computer graphics", *Adv. Geom.* **14**:1 (2014), 31–48. MR 3159090 Zbl 1283.53058

[López and Munteanu 2011] R. López and M. I. Munteanu, "Surfaces with constant mean curvature in Sol geometry", *Differential Geom. Appl.* **29**:S1 (2011), S238–S245. MR 2012k:53016 Zbl 1250.53058

[López and Munteanu 2012] R. López and M. I. Munteanu, "Minimal translation surfaces in Sol$_3$", *J. Math. Soc. Japan* **64**:3 (2012), 985–1003. MR 2965436 Zbl 1254.53033

[López and Munteanu 2014] R. López and M. I. Munteanu, "Invariant surfaces in the homogeneous space *Sol* with constant curvature", *Math. Nachr.* **287**:8-9 (2014), 1013–1024. MR 3219227 Zbl 1293.53076

[Masaltsev 2006] L. A. Masaltsev, "Minimal surfaces in standard three-dimensional geometry Sol^3", *Zh. Mat. Fiz. Anal. Geom.* **2**:1 (2006), 104–110. MR 2006j:53008 Zbl 1144.53016

[Meeks and Pérez 2012] W. H. Meeks, III and J. Pérez, "Constant mean curvature surfaces in metric Lie groups", pp. 25–110 in *Geometric analysis: partial differential equations and surfaces* (Granada, 2010), edited by J. Pérez and J. A. Gálvez, Contemporary Mathematics **570**, American Mathematical Society, Providence, RI, 2012. MR 2963596 Zbl 1267.53006

[Menezes 2014] A. Menezes, "Periodic minimal surfaces in semidirect products", *J. Aust. Math. Soc.* **96**:1 (2014), 127–144. MR 3177813 Zbl 1288.53052

[Nguyen 2014] M. H. Nguyen, "The Dirichlet problem for the minimal surface equation in Sol_3 with possible inifinite boundary data", preprint, 2014. arXiv 1312.6194

Received January 22, 2014. Revised August 1, 2014.

CHRISTOPHE DESMONTS
INSTITUT ELIE CARTAN DE LORRAINE
UNIVERSITÉ DE LORRAINE
B.P. 70239
54506 VANDOEUVRE-LÈS-NANCY CEDEX
FRANCE
christophe.desmonts@univ-lorraine.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.167

QUASI-EXCEPTIONAL DOMAINS

ALEXANDRE EREMENKO AND ERIK LUNDBERG

Exceptional domains are domains on which there exists a positive harmonic function, zero on the boundary and such that the normal derivative on the boundary is constant. Recent results classify (under some mild additional assumptions) exceptional domains as belonging to either a certain one-parameter family of simply periodic domains or one of its scaling limits.

We introduce quasi-exceptional domains by allowing the boundary values to be different constants on each boundary component. This relaxed definition retains the interesting property of being an *arclength quadrature domain*, and also preserves the connection to the hollow vortex problem in fluid dynamics. We give a partial classification of such domains in terms of certain abelian differentials. We also provide a new two-parameter family of periodic quasi-exceptional domains. These examples generalize the hollow vortex array found by Baker, Saffman, and Sheffield. A degeneration of regions of this family provides doubly connected examples.

1. Introduction

A domain $D \in \mathbb{R}^n$ is called exceptional if there is a positive function u (called a *roof function*) harmonic in D, zero on the boundary, and with

$$\frac{\partial}{\partial n} u(z) = 1, \quad z \in \partial D, \tag{1}$$

where the differentiation is along the normal pointing inwards into D and it is assumed that the boundary is smooth. Evident examples are exteriors of balls and half-spaces. For $n > 2$, the only other known examples are cylinders whose base is an exceptional domain in $\mathbb{R}^2$. If the smoothness assumption on the boundary is dropped, then there are also certain cones in higher dimensions and pathological "non-Smirnov" examples in the plane [Khavinson et al. 2013].

The problem of describing of all exceptional domains in the plane was stated in [Hauswirth et al. 2011] and settled in [Khavinson et al. 2013] under a topological

Supported by NSF grant DMS-1361836.
MSC2010: primary 30C20, 35R35; secondary 76B47, 31A05.
Keywords: quadrature domains, hollow vortices, elliptic functions, abelian differentials.

assumption which was removed in [Traizet 2014a] using an unexpected correspondence with minimal surfaces. The first nontrivial example was given in [Hauswirth et al. 2011]. This example appeared in another context, related to fluid dynamics, in [Longuet-Higgins 1988]. A second nontrivial example was noticed in [Khavinson et al. 2013] and [Traizet 2014a]. This example had also appeared previously in studies of fluid dynamics [Baker et al. 1976] (see also [Crowdy and Green 2011]).

Let us introduce *quasi-exceptional* domains, by relaxing the definition to allow the Dirichlet condition to be a different constant on each boundary component. Thus, a domain $D \in \mathbb{R}^n$ is called quasi-exceptional if there is a positive harmonic function u in D which is constant on each boundary component (but not necessarily the same constant) and the Neumann condition (1) holds. We will continue to call u a *roof function*. Again, we assume that each component of the boundary is smooth.

Added in press: In an interesting preprint, Martin Traizet [2014b] has considered an even more general problem, allowing the Neumann data to take different signs on different components. As with the current paper, that work is motivated by the hollow vortex problem from fluid dynamics. Extending [Traizet 2014a], a correspondence to minimal surfaces is given in that work, and techniques from minimal surface theory are used to produce new examples.

We summarize several interesting aspects of exceptional domains. These statements all hold true for quasi-exceptional domains.

- Fluid dynamics: As noted above, the two nontrivial examples first appeared in fluid dynamics [Longuet-Higgins 1988; Baker et al. 1976]. In general, one can interpret exceptional domains in terms of a *hollow vortex* problem. The level lines of u can be interpreted as stream lines of a two-dimensional stationary flow of ideal fluid, and condition (1) expresses the fact that the pressure is constant on the boundary. Such conditions may exist if the components of the complement of D are air bubbles in the surrounding liquid. Notice that the rotation of the fluid around all bubbles corresponding to exceptional domains is in the same direction because $\partial u/\partial n > 0$ on the boundary.

- Quadrature domains [Gustafsson 1987]: Exceptional domains provide examples of arclength null-quadrature domains, that is, domains for which integrals with respect to arclength over ∂D of every analytic function in the Smirnov class $E^1(D)$ vanish.

- Differentials on Riemann surfaces: By way of the connection to quadrature domains, the study [Gustafsson 1987] indicates a connection to half-order differentials. We make use of abelian differentials in Section 4 below.

- The Schwarz function of a curve: In [Khavinson et al. 2013], it was noticed that the function $u(z)$ satisfies

$$\partial_z u(z) = \sqrt{-S'(z)},$$

where $S(z)$ is the Schwarz function of $\partial\Omega$ and $\partial_z = \frac{1}{2}(\partial_x - i\partial_y)$ is the Cauchy–Riemann operator.

- Minimal surfaces: Traizet [2014a] established a nontrivial correspondence between exceptional domains and a special type of minimal surfaces called "minimal bigraphs". In [Traizet 2014b], this correspondence was extended to quasi-exceptional domains, but the minimal surfaces in that case need not be embedded. This prevents applying the results on complete embedded minimal surfaces that were used in [Traizet 2014a] to classify exceptional domains.

The classification results for exceptional domains of finite connectivity show that they are quite restricted; all examples can be conformally mapped onto a disk by elementary functions.

Problem A. Classify quasi-exceptional domains.

We begin to address this problem below, give a partial classification of periodic and finitely connected exceptional domains, and provide new periodic and doubly connected examples described in terms of elliptic functions. First, we explain the relation to arclength null-quadrature domains.

2. Arclength null-quadrature domains

A bounded domain $D \subset \mathbb{C}$ is a *quadrature domain* if it admits a formula expressing the area integral of every function f analytic and integrable in D as a finite sum of weighted point evaluations of the function and its derivatives, i.e.,

$$\int_D g(z)\, dA(z) = \sum_{m=1}^{N} \sum_{k=0}^{n_m} a_{m,k}\, g^{(k)}(z_m), \tag{2}$$

where the z_m are distinct points in D and the $a_{m,k}$ are constants independent of g.

A (necessarily unbounded) domain $D \subset \mathbb{C}$ is called a *null-quadrature domain* (NQD) if the area integral of every function g analytic and integrable in D vanishes:

$$\int_D g(z)\, dA(z) = 0. \tag{3}$$

M. Sakai [1981] completely classified NQDs in the plane.

Following [Khavinson et al. 2013], we refer to a domain $D \subset \mathbb{C}$ as an *arclength null-quadrature domain* (ALNQD) if the integral over ∂D of every function g in the Smirnov class $E^1(D)$ vanishes (in the case ∞ is an isolated point on ∂D, we take the restricted class of functions $g(z) \in E^1(D)$ vanishing at infinity):

$$\int_{\partial D} g(z)\, ds(z) = 0. \tag{4}$$

The Smirnov class $E^1(D)$ is not the same as the Hardy space $H^1(D)$. Namely, a function g analytic in D is said to belong to $E^1(D)$ if there exists a sequence of domains $D_1 \subset D_2 \subset \cdots$, with $\bigcup_k D_k = D$, and with rectifiable boundaries, such that:

$$\sup_k \int_{\partial D_k} |g(z)|\, |dz| < \infty.$$

One may also define quadrature domains in higher dimensions using a test class of harmonic functions, but we will restrict ourselves to the case of $n = 2$ dimensions.

Inspired by the successful classification of NQDs [Sakai 1981], the problem of classifying ALNQDs was suggested in [Khavinson et al. 2013]. We pose this problem again while stressing that it does not reduce to the classification of exceptional domains (whereas it might reduce to classification of *quasi-exceptional* domains).

Problem B. Classify ALNQDs.

The following proposition shows that quasi-exceptional domains are ALNQDs. Thus, the new examples (described in the last section) of quasi-exceptional domains also provide new ALNQDs. Problem B is closely related to Problem A, and if the converse of the proposition is true then the two problems are equivalent.

Proposition 1. *If D is a quasi-exceptional domain, then D is an ALNQD.*

Proof. Consider the complex analytic function $F(z) = u_x - iu_y$, where u is the roof function. We will need the following claim, which is proved in the next section (see Lemma 2).

Claim. *The roof function u of D satisfies $\nabla u(z) = O(1)$ in D, so $F(z)$ is bounded.*

Suppose that g is in the Smirnov space $E^1(D)$. Using the fact that $ds = i F(z)\, dz$,

$$\int_{\partial D} g(z)\, ds = \int_{\partial D} i g(z) F(z)\, dz. \tag{5}$$

As F is bounded, $gF \in E^1(D)$. If ∞ is not an isolated boundary point, then the integral equals zero by Cauchy's theorem.

If ∞ is an isolated boundary point, then we have $u(z) = \log|z| + \text{const} + O(1/z)$, so $F(z) = O(1/z)$, $z \to \infty$. Now $g(\infty) = 0$, so Fg has a zero of order at least 2 at ∞, and the integral is zero again. $\square$

3. A potential theoretic restriction on the roof function

We restrict ourselves to the case $n = 2$, and assume that the order of connectivity of D is finite, or that the roof function u is periodic and a fundamental region for D has finite connectivity.

Recall that a *Martin function* is a positive harmonic function M in a domain Ω with the property that for any positive harmonic function v in Ω the condition $v \le M$ implies that $v = cM$, where $c > 0$ is a constant. (Often, Martin functions are called minimal harmonic functions — see [Heins 1950].) Martin functions on finitely connected domains are simply Poisson kernels evaluated at points of the Martin boundary, the boundary under Carathéodory compactification (prime ends) of the domain (see [Brelot 1971]).

Any domain D of finite connectivity in $\mathbb{C}$ is conformally equivalent to a circular domain Ω. A circular domain is a domain whose boundary components are points or circles. For a circular domain, a Martin function M can be of two types:

(a) There is a component of $\partial\Omega$ which is a single point z_0, and M is proportional to the Green function of $\Omega \cup \{z_0\}$ with the singularity at z_0.

(b) There is a point $z_0 \in \partial\Omega$ which is not a component of $\partial\Omega$, and M has boundary value zero at all points of $\partial\Omega \setminus \{z_0\}$. The local behavior in this case is like $-\operatorname{Im}(1/z)$ in the upper half-plane near 0.

Let D be an exceptional domain, and u a harmonic function with the property (1). The following result was proved for exceptional domains by the current first author, but was communicated in [Khavinson et al. 2013, Theorem 4.2]. Here we repeat the proof with minor adjustments.

Lemma 2. *The roof function u of a quasi-exceptional domain satisfies the equation $\nabla u(z) = O(1)$ in D. Moreover, u is the sum of a bounded harmonic function and at most two Martin functions.*

Proof. We follow the second part of the proof from [Khavinson et al. 2013]. Let $R > 0$, and consider an auxiliary function

$$w_R = \frac{|\nabla u|}{u + R}.$$

A direct computation shows that

$$\Delta \log w_R = w_R^2, \tag{6}$$

and $w_R(z) = 1/(c_k + R) \le 1/R$ for $z \in \partial D$, where $c_k \ge 0$ are the constants taken in the Dirichlet condition. We claim that

$$w_R(z) \le 2/R, \quad z \in D, \tag{7}$$

from which the result follows by letting $R \to \infty$, which gives $|\nabla u| \leq 2$ in D.

Suppose, contrary to (7), that $w_R(z_0) > 2/R$ for some $z_0 \in D$. Let

$$v(z) = \frac{2R}{R^2 - |z - z_0|^2}, \quad z \in B(z_0, R) = \{z : |z - z_0| < R\}.$$

Obviously, $v(z) \geq 2/R$. A computation reveals that $\Delta \log v = v^2$. Let

$$K = \{z \in D \cap B(z_0, R) : w_R(z) > v(z)\}.$$

We have $z_0 \in K$, since $v(z_0) = 2/R$. Let K_0 be the component of K containing z_0. Then we have $w_R(z) = v(z)$ on ∂K_0, since $w_R(z) < v(z)$ on $\partial D \cap B(z_0, R)$ while $v(z) = +\infty$ on $\partial B(z_0, R)$. On the other hand,

$$\Delta(\log w_R - \log v) = w_R^2 - v^2 > 0 \quad \text{in } K_0.$$

So the subharmonic function $\log w_R - \log v$ is positive in K_0 and vanishes on the boundary — a contradiction.

This proves that $\nabla u = O(1)$. In order to see the second statement, we note that $\nabla u = O(1)$ implies that $u(z) = O(|z|)$ has order 1. The result then follows by first solving the Dirichlet problem (with a bounded function) having the same boundary values as u; subtracting this function, one may then apply [Kjellberg 1950, Theorem II]. □

4. Partial classification in terms of abelian differentials

Let D be a QE domain of one of the following types:

Type I: D is finitely connected.

Type II: D/Γ is finitely connected, where Γ is the group of transformations $z \mapsto z + n\omega$, and $u(z+\omega) = u(z)$ for some $\omega \in \mathbb{C} \setminus \{0\}$. We call this the periodic case. (As above, u is the roof function.)

In this section we give a classification of QE domains of these two types in terms of abelian differentials of a compact Riemann surface with an anticonformal involution.

If D is of type I, and ∞ is an isolated boundary point, then $D' = D \cup \{\infty\}$ is conformally equivalent to some bounded circular domain Ω, and we suppose that $p \in \Omega$ corresponds to ∞. If ∞ is not isolated, we put $D' = D$, and Ω is a bounded circular domain conformally equivalent to D'. In any case, we have a conformal map $\phi : \Omega \to D'$, which may have at most one simple pole at $p \in \Omega$.

If D is of type II, let $G = D/\Gamma$. The Riemann surface G is a finitely connected domain on the cylinder $\mathbb{C}/\Gamma$; this cylinder is conformally equivalent to the punctured plane, and we identify it with $\mathbb{C}^*$. Then $G \subset \mathbb{C}^*$ must have one or two punctures of $\mathbb{C}/\Gamma$ as isolated boundary points, and we denote by G' the union of G with these isolated boundary points. Then G' is conformally equivalent to a bounded circular domain of finite connectivity Ω in which there are one or two points a and b corresponding to the added punctures. We have a multivalued conformal map $\phi : \Omega \to D$.

The points a and b are logarithmic singularities of ϕ.

We pull back u on Ω; i.e., set $v = u \circ \phi$. As u is periodic, v is a single-valued positive harmonic function on $\Omega \setminus \{a, b\}$. Consider the differential on Ω

$$dv = v_z\, dz = \tfrac{1}{2}(v_x - i v_y)(dx + i dy) = g(z)\, dz.$$

This is well-defined on Ω: g is a single-valued meromorphic function in Ω with at most simple poles at p or a and b. Indeed, for a positive harmonic function, an isolated singularity is either removable or logarithmic. In the second case the gradient has a simple pole.

Next, we extend v as a multivalued function to a compact Riemann surface S. Let Ω' be the mirror image of Ω; we glue it to Ω in the standard way (along each circular boundary component) and obtain a compact Riemann surface S. We denote by $\sigma : z \mapsto z^*$ the anticonformal involution which fixes the boundary components of Ω. The Riemann surface S is of genus g, and the involution σ has fixed set corresponding to $\partial\Omega$, which consists of $n = g + 1$ ovals. Such involutions are called *involutions of maximal type*, meaning that the complement of the fixed set of the involution consists of two regions homeomorphic to planar regions.

Each branch of v is constant on each boundary component, so it extends through this boundary component by reflection to the double S of Ω. The extensions of various branches of v through different boundary components do not match: they differ by additive constants. On the other hand, the differential dv is well-defined on the double. Namely,

$$(dv)^* = -dv, \tag{8}$$

where $*$ is the action of involution on differentials. Thus we have a meromorphic differential dv on S.

Choose a basis of 1-homology in S so that the A-loops are simple closed curves in Ω, each homotopic to one boundary component of Ω, and the B loops are dual to the A-loops. For type I, all periods over A-loops are purely imaginary, because

$$v = \operatorname{Re} \int dv$$

is single-valued. For type II, these periods are imaginary except those which correspond to simple loops around one pole, a or b.

Now we discuss ϕ, or better, the differential $d\phi = \phi'(z)\,dz$. We have, from the condition that our domain is quasi-exceptional,

$$2|dv| = |d\phi| \quad \text{on } \partial\Omega.$$

The ratio of two differentials is a function. So we have a meromorphic function B on Ω such that

$$2B\,dv = d\phi. \tag{9}$$

This function has absolute value 1 on $\partial\Omega$. Therefore, it extends to S by symmetry. It has no zeros in Ω because $d\phi$ has no zeros. Its poles in Ω must match the zeros of dv, because $d\phi$ is zero-free (indeed, ϕ is univalent). In fact, B is a meromorphic function on S. To justify this claim when dv has a singularity on $\partial\Omega$, we observe that this singularity is removable for B, which follows from the next lemma:

Lemma 3. *Consider the equation*

$$\phi' = Bh,$$

where h is meromorphic in a neighborhood V of 0, B is holomorphic and zero-free in $V \setminus \{0\}$, $|B(z)| = 1$ for $z \in V \cap \mathbb{R} \setminus \{0\}$, and ϕ is univalent in $\{z \in V : \operatorname{Im} z > 0\}$. Then the singularity of B at 0 is removable.

Before proving the lemma, we note that in order to apply it in our setting we compose B with a linear fractional transformation that sends V to a neighborhood of the singularity we wish to remove such that the real line is mapped to the circular boundary component with 0 sent to the singularity.

Proof. In order to prove this by contradiction, assume that 0 is an essential singularity of B. By symmetry we have $B(\bar{z}) = 1/\overline{B(z)}$. We claim that there exists a sequence $z_k \to 0$ such that

$$\liminf_{k\to\infty} \bigl|z_k \log|B(z_k)|\bigr| > 0. \tag{10}$$

Indeed, suppose that this is not so. Then $\log|B(z)| = o(z^{-1})$, and the Phragmén–Lindelöf theorem (see, for example, [Levin 1980, Chapter I, Theorem 22]) implies that B has a limit as $z \to 0$. By choosing a subsequence and using symmetry, we can find a sequence in the upper half-plane with the property

$$\liminf_{k\to\infty} |z_k| \log|B(z_k)| > 0, \tag{11}$$

or with the property

$$\liminf_{k\to\infty} |z_k| \log|B(z_k)| < 0. \tag{12}$$

Distortion theorems for univalent functions imply that

$$c(\operatorname{Im} z)^3 \leq |\phi'(z)| \leq C(\operatorname{Im} z)^{-3}. \tag{13}$$

In addition to this, we have, for some integer m,

$$c|x|^m \leq |\phi'(x)| \leq C|x|^m, \quad z \in V \cap \mathbb{R}, \tag{14}$$

because h is meromorphic and $|B(x)| = 1$ for $z \in V \cap \mathbb{R}$. Taking $n = \max\{3, -m\}$, we obtain that the subharmonic function $u(z) = \log^+|z^n \phi'(z)/C|$ satisfies $u(x) = 0$ for $x \in V \cap \mathbb{R}$, and $u(re^{i\theta}) \leq \psi(\theta)$, where $\psi(\theta) = -3 \log \sin \theta$ for $\theta \in (0, \pi)$. As

$$\int_0^\pi \psi(\theta)\, d\theta < \infty,$$

we can apply Carleman's "log log" theorem [Carleman 1926; Rashkovskii 2009], and conclude that u is bounded from above in the intersection of V with the upper half-plane. This contradicts (11). If (12) holds, one applies the same argument to $1/\phi'$. This completes the proof of the lemma. □

We can thus restate the problem of finding QE domains (under the restrictions we impose) as follows:

Proposition 4. *All QE domains of types I and II are parametrized by triples* $(S, d\omega, B)$, *where S is a compact Riemann surface with an involution of maximal type*, *$d\omega$ is a meromorphic differential that enjoys the symmetry property* (8), *and B is a function that has the symmetry property*

$$B^*(z) := \overline{B(z^*)} = 1/B(z)$$

and has poles at the zeros of $d\omega$ on one half of S, *that is*, *in Ω. There is an additional condition*: *that*

$$\phi = 2\int B\, d\omega \tag{15}$$

is globally univalent and single-valued in type I, *and single-valued except the residues in type II.*

To recover D from a triple (S, dw, B), one takes one of the components $\Omega \subset S$ complementary to the fixed set of the involution. Then $D = \phi(\Omega)$, where ϕ is defined in (15).

In order to check the condition on the global univalence of ϕ, it is sufficient to verify that periods of $d\omega/B$ are zero on the boundary curves, and that these boundary curves are mapped by ϕ injectively.

The following is a general conclusion:

Proposition 5. *The boundary of a quasi-exceptional domain of type I or type II is parametrized by an abelian integral.*

Next we provide a partial classification of quasi-exceptional domains in terms of the data stated in the above formulation.

Theorem 6. *The differential dv has either two or four poles on S, counting multiplicity. Moreover, if dv has two poles in S, then D is the complement of either a disk or a half-plane.*

Remark. If $B \neq \text{const}$, then $1/B$ is an Ahlfors function of Ω.

Proof of Theorem 6. Let us first show that dv has some poles. Otherwise, dv is holomorphic, and thus u is bounded. Let $z_1 \in \partial D$ and $z_2 \in \partial D$ be the points where u assumes its maximal and minimal values. Then du/dn has opposite signs at these two points, which contradicts (1).

The differential dv has at most simple poles at p, a, and b (whichever of these points are present) and at their images σp, σa, and σb. In addition it may have double poles on $\partial\Omega$. The total number of poles (without multiplicity) in $\overline{\Omega}$ is at most two by Lemma 2. Thus, on S the differential dv has two or four poles, counting multiplicity.

Notice that v is constant on each boundary component, so the gradient is perpendicular to the boundary $\partial\Omega$, so the total rotation of this gradient as we traverse the boundary is the same as the total rotation of the tangent vector to the boundary. This is equal to $2\pi(2-n)$ because the outer boundary component is traversed counterclockwise and the rest clockwise, as parts of the boundary of Ω. So v_z, which is conjugate to the gradient, rotates $n-2$ times.

From this we can conclude how many zeros dv has in Ω. The number N of zeros of dv in Ω satisfies

$$n-2 = N - (\text{the number of poles in } \Omega), \tag{16}$$

where a double pole on $\partial\Omega$ is counted as a single pole in Ω. This formula is well known.

Suppose that dv has exactly two poles, counting multiplicity. This can occur in one of three ways:

Case (1): dv has a simple pole at p in Ω.

Case (2): dv has one double pole at $z_0 \in \partial\Omega$.

Case (3): dv has a simple pole at a in Ω (and b does not exist).

If Case (1) holds, then ∞ is an isolated point on ∂D, and, by Proposition 1, D is an arclength quadrature domain with quadrature point at ∞. It now follows from [Gustafsson 1987, Remark 6.1] that D is the exterior of a disk.

In Case (2), we will show that B is constant. First note that $d\phi$ has a double pole at z_0, so B does not have a zero or a pole at z_0. Since ϕ is a conformal map, it follows from (9) that B has no zeros and N poles in Ω (located at the zeros of dv).

Assume for the sake of contradiction that B is not constant. By Lemma 3, B is meromorphic in S, and, by Lemma 2, $1/|B|$ is bounded by a constant in Ω. Since $|B| = 1$ on $\partial\Omega$, B thus maps Ω to the exterior of the unit disk and maps each of the n components of $\partial\Omega$ to the unit circle. This implies that B has at least n poles in Ω. Combined with (16), this gives the contradiction $N = n - 1 \geq n$. We conclude that B is constant, which implies that the gradient of the roof function is constant. Thus, the roof function is linear, and D is a half-plane.

In Case (3), the behavior of ϕ at point a is logarithmic, so $d\phi$ has a simple pole at a and B does not have a zero or a pole at a. Arguing as before, we conclude that B is constant and that D is a half-plane. □

Corollary 7. *The only QE domains with compact boundary are exteriors of disks, and the only QE domains of types I or II with one unbounded boundary component are half-planes.*

If D is a quasi-exceptional domain that is not a disk or half-plane, then dv has four poles and, more precisely, we have the following two possibilities:

D is of *type I*: dv has two double poles on $\partial\Omega$. This implies that the boundary ∂D consists of two simple curves tending to ∞ in both directions and $n - 1$ bounded components. The unbounded components are the ϕ-images of two arcs of one boundary circle of Ω which contains both singularities of ϕ and v.

D is of *type II*: dv has two simple poles in Ω. In this case D must be periodic, all components of ∂D are compact, and there are n such components per period.

Note that the possibility that dv has one simple pole in Ω and one double pole on $\partial\Omega$ is excluded by Lemma 2: it is easy to see that in this case the number of Martin functions in the decomposition of u would be infinite.

We have thus described possible topologies of the QE domains satisfying the assumptions stated in the beginning of this section.

In the next section we construct the examples of types I and II with S of genus 1. We conjecture that there exist QE domains of types I and II with S of any genus.

5. New examples

Description of our examples requires elliptic functions (all known exceptional domains can be parametrized by elementary functions).

Example of type I. Let G be the rectangle with vertices $(0, 2\omega_1, 2\omega_1 + \omega_3, \omega_3)$, where $\omega_1 = 2\omega$, $\omega > 0$, and $\omega_3 = \omega'$, where $\omega' \in i\mathbb{R}$, $\omega'/i > \omega$. Let G' be the reflection of G in the real line. The union of G, G' and the interval $(0, 2\omega_1)$ make a fundamental domain of the lattice Λ generated by $2\omega_1, 2\omega_3$.

Let us consider the ω_1-periodic positive harmonic function h in G which is zero on the horizontal segments of the boundary ∂G, except for one singularity per period, at 0, where it behaves in the following way:

$$h(z) \sim -\operatorname{Im}(1/z), \quad z \to 0.$$

Note that the existence of h is clear as it can be expressed (through conformal mapping) in terms of the Poisson kernel of a ring domain.

The function h has two critical points in G: at w_1 and w_2 with $\operatorname{Re} w_1 = \omega_1/2$ and $\operatorname{Re} w_2 = 3\omega_1/2$, while the imaginary parts of w_1 and w_2 are equal. Let us choose real constants c_1 and c_2 such that $v = 2(h + c_1 y) + c_2$ is a positive harmonic function with critical points $\omega_1/2 + \omega_3/2$ and $3\omega_1/2 + \omega_3/2$. The existence of such constants c_1 and c_2 is evident by continuity.

The z-derivative $\partial_z v = (v_x - i v_y)/2$ is an elliptic function with periods $\omega_1, 2\omega_3$, and thus also elliptic with periods Λ. Asymptotics near 0 show that $\partial_z v \sim -i/z^2$, and, as this function has only one pole per period (with respect to the parallelogram $\omega_1, 2\omega_3$), we have $\partial_z v = -i\wp + i c_0$, where $\wp$ is the Weierstrass function corresponding to the lattice $(\omega_1, 2\omega_3)$. Zeros of $\partial_z v$ in $G \cup G'$ are $\omega_1/2 + \omega_3/2$, $3\omega_1/2 + \omega_3/2$ and their complex conjugates in G'.

Let B be an elliptic function with periods $2\omega_1, 2\omega_3$ having simple poles at $\omega_1/2+\omega_3/2$, $3\omega_1/2+\omega_3/2$, and zeros at complex conjugate points. Such a function exists by Abel's theorem: the sum of zeros minus the sum of poles equals $-2\omega_3$. This function is unique up to a constant factor. By symmetry, $B(\bar{z}) = c/\overline{B(z)}$, so on the real line $|B(x)|^2 = c$ and we can choose the constant factor in the definition of B so that $c = 1$. Thus

$$|B(x)| = 1, \quad x \in \mathbb{R}. \tag{17}$$

Then we have $B(x + \omega_3)\overline{B(x - \omega_3)} = 1$, but by periodicity we also have $B(x+\omega_3) = B(x - \omega_3)$, thus $|B(x + \omega_3)| = 1$. So

$$|B(z)| = 1 \quad \text{on the horizontal segments of } \partial G. \tag{18}$$

Now we consider the function

$$F = \frac{\partial v}{\partial z} B = (-i\wp + i c_0) B.$$

This function F is holomorphic and zero-free in G (the zeros of $\partial v/\partial z$ in G are exactly canceled by the poles of B). Let us show that

$$\int_0^{2\omega_1} F(x + iy)\, dx = 0, \quad y \in (0, \omega_3). \tag{19}$$

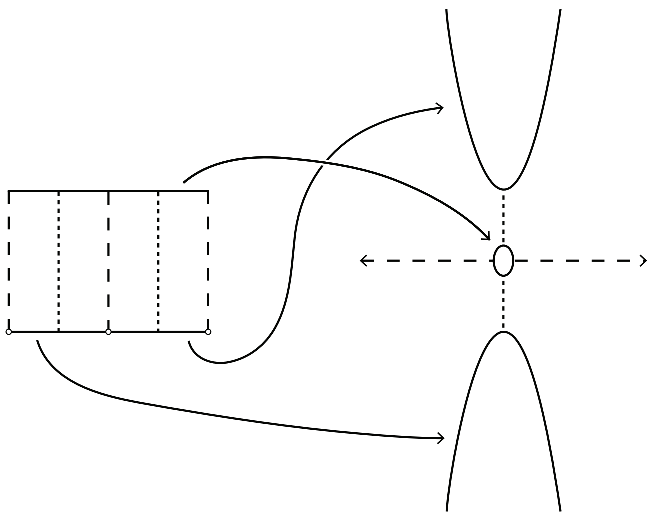

Figure 1. A doubly connected quasi-exceptional domain of type I mapped from the rectangle G.

This property follows from the fact that $B(z)$ and $B(z+\omega_1)$ have the same poles, but the residues at these poles are of opposite signs, because B has only two poles in the period parallelogram. Thus

$$B(z+\omega_1) = -B(z). \tag{20}$$

Property (20) and ω_1-periodicity of $\wp$ imply (19).

As F has no zeros, the primitive $f = \int F$ is locally univalent. Assuming for the moment that it is univalent, it maps G onto some region in the plane, and we have

$$|f'| = |F| = \left|\frac{\partial v}{\partial z}\right| |B|.$$

Define u by composing v with f^{-1}, so $u(f(z)) = v(z)$. Then u is positive and harmonic in $f(G)$. Taking into account (18), we conclude that u satisfies (1) so $f(G)$ is a quasi-exceptional domain. Note that, in accordance with the previous results in [Traizet 2014a], $f(G)$ is not an exceptional domain since the piecewise-constant Dirichlet data is not the same constant on each boundary component.

In order to show that f is in fact univalent, it is enough to show that it is one-to-one on the horizontal sides of G (since f is locally univalent). To this end, we make the following claims:

Claim 1: Re f is increasing along the segment $[\omega', \omega'+2\omega]$ and decreasing along the segment $[\omega'+2\omega, \omega'+4\omega]$.

Claim 2: $\operatorname{Im} f < \operatorname{Im} f(\omega')$ on the segment $(\omega', \omega' + 2\omega]$ and $\operatorname{Im} f > \operatorname{Im} f(\omega')$ on the segment $[\omega' + 2\omega, \omega' + 4\omega)$.

Claim 3: $\operatorname{Im} f$ achieves its minimum and maximum on the segment $[\omega', \omega' + 4\omega]$ at $\omega' + \omega$ and $\omega' + 3\omega$, respectively.

Claim 4: $\operatorname{Re} f$ is increasing along the segment $[0, 2\omega]$ and $\operatorname{Re} f$ is decreasing along the segment $[2\omega, 4\omega]$.

Claim 5: $\operatorname{Im} f$ attains its maximum on the segment $[0, 2\omega]$ at ω and its minimum on the segment $[2\omega, 4\omega]$ at 3ω.

Claim 6: $\operatorname{Im} f(\omega) < \operatorname{Im} f(\omega' + \omega) < \operatorname{Im} f(\omega' + 3\omega) < \operatorname{Im} f(3\omega)$.

Claim 1 implies that $\operatorname{Re} f$ is monotone along each of the named segments, and since $\operatorname{Im} f$ differs between the two segments by Claim 2, f must be one-to-one on the top side of G. Claim 4 implies that f is one-to-one on each of the two segments on the bottom side of G. Claims 3, 5, and 6 imply that the images of these three segments do not intersect each other. This shows that f is one-to-one on the horizontal sides of G.

The claims can be established by the properties of $f' = F = \partial_z v B$. First note that, since $v(z)$ is positive in G and vanishes on the horizontal sides of G, we have $\partial_x v(z) = 0$ on both sides, and for $x \in \mathbb{R}$ we have $\partial_y v(x + \omega_3) < 0$ and $\partial_y v(x) > 0$. In particular, $i\partial_z v(z) = i(\partial_x v - i\partial_y v)/2 = \partial_y v/2$ is real. The function $B(z)$ is a Jacobi sn function, whose properties are well known [Akhiezer 1990, Section 47]. $B(z)$ sends the top side of G to the unit circle, such that the four segments $[\omega', \omega' + \omega]$, $[\omega' + \omega, \omega' + 2\omega]$, $[\omega' + 2\omega, \omega' + 3\omega]$, and $[\omega' + 3\omega, \omega' + 4\omega]$ correspond to the fourth, third, second, and first quadrants of the unit circle, respectively. Multiplication by $\partial_z v(z)$ distorts this circle and rotates it by an angle of $\pi/2$ (since $\partial_z v(z)/i$ is positive), but preserves the two-fold symmetry. This determines the sign of the real and imaginary parts of f'. Since $dz = dx$ is purely real on the horizontal sides of G, this gives the monotonicity of $\operatorname{Re} f$ stated in Claim 1. Claims 2 and 3 follow from the sign of $\operatorname{Im} f'$ and the fact that $\operatorname{Im} f'$ is an odd function with respect to reflection in each of the points $\omega' + \omega$ and $\omega' + 3\omega$.

The four segments $[0, \omega]$, $[\omega, 2\omega]$, $[2\omega, 3\omega]$, and $[3\omega, 4\omega]$ on the bottom side of G are sent to the second, first, fourth, and third quadrants of the unit circle, respectively. Since $\partial_z v(z)/i$ is negative along the bottom side of G, under $f'(z)$ this becomes the first, fourth, third, and second quadrants, respectively. This establishes Claim 4, and, combined with the reflectional symmetry, also Claim 5. Claim 6 follows from the fact that $\partial_z v(z)B(z) > 0$ along the vertical segment $[\omega, \omega + \omega']$ and $\partial_z v(z)B(z) < 0$ along $[3\omega, 3\omega + \omega']$.

Figure 2. An example of type II with $\omega_1 = 2$, $\omega_3 = 2$ and $\epsilon = 0.5$.

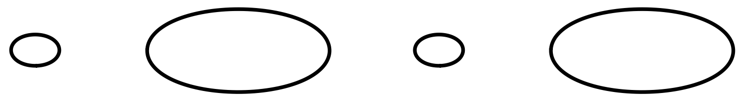

Figure 3. An example of type II with $\omega_1 = 2$, $\omega_3 = 1.5$, and $\epsilon = 0.4$.

Remark. For the purpose of plotting Figure 1, instead of the above construction, we expressed F as a ratio of Weierstrass sigma functions:

$$f'(z) = F(z) = \frac{\sigma(z-\omega+\omega'/2)^2 \cdot \sigma(z-3\omega+\omega'/2)^2}{\sigma(z)^2 \cdot \sigma(z-2\omega) \cdot \sigma(z-6\omega+2\omega')},$$

where σ is a Weierstrass sigma function with fundamental "periods" 4ω, $2\omega'$ (but recall that σ is not itself periodic). As usual, the shifts are chosen based on the zeros and poles of F, but one of the shifts must be replaced by an equivalent lattice point in a different rectangle in order to satisfy [Akhiezer 1990, Section 14, (1)]. This explains why one of the poles is placed at $6\omega - 2\omega'$.

Example of type II. Only small modifications of the previous example are needed. Using the same G, G', ω_1, ω_3, we define h as the ω_1-periodic function, positive and harmonic in G' except two logarithmic poles at $i\epsilon$ and $\omega_1 + i\epsilon$, where $\epsilon \in (0, \omega_3/2)$. Then we can find constants c_1 and c_2 such that $v = h + c_1 y + c_2$ has critical points at $\omega_1/2 + \omega_3/2$ and $3\omega_1/2 + \omega_3/2$.

Then v_z is an elliptic function with periods ω_1, $2\omega_3$ with two simple poles at $i\epsilon$ and $-i\epsilon$ per period parallelogram. This elliptic function has the form

$$\frac{-i\wp}{1 + c\wp} + ic_0$$

for some small real c. The rest of the construction is the same as in the previous example.

In a similar manner to the above, in order to plot Figures 2 and 3, we expressed F as a ratio of Weierstrass sigma functions:

$$f'(z) = F(z) = \frac{\sigma(z-\omega+\omega'/2)^2 \cdot \sigma(z-3\omega+\omega'/2)^2}{\sigma(z-i\epsilon) \cdot \sigma(z+i\epsilon) \cdot \sigma(z-2\omega-i\epsilon) \cdot \sigma(z-6\omega+i\epsilon+2\omega')}.$$

Note that we have displayed the figures horizontally in order to plot two periods.

6. Hollow vortex equilibria

Let G_j be smooth Jordan domains on the plane whose closures are disjoint, and

$$D = \mathbb{C} \setminus \bigcup_j G_j.$$

Let F be the complex potential of a flow of an ideal fluid which is divergence-free and locally irrotational in D. If the pressure (determined by $|F'|$ according to Bernoulli's law) is constant on ∂D then G_j can be interpreted as constant-pressure gas bubbles in the flow.

The first examples of this situation, with two bubbles, were constructed by Pocklington [1895]. Periodic exceptional domains give periodic examples with one bubble per period, with the flow on the surface on the bubbles rotating in the same direction [Baker et al. 1976] (see also [Crowdy and Green 2011]). Crowdy and Green [2011] constructed periodic examples with two bubbles per period rotating in the opposite direction. Our example of type II can be interpreted as a periodic flow with two bubbles per period rotating in the same direction.

The velocity at infinity in our examples is directed in the opposite directions on the two sides of the row of bubbles.

Acknowledgments

We are grateful to Dmitry Khavinson for many helpful discussions and to Razvan Teodorescu for a crucial observation regarding the construction of the examples of type I. We wish to thank Darren Crowdy for discussing with us the interesting connection to the hollow vortex problem. We also thank the referee who made many remarks which improved the paper.

References

[Akhiezer 1990] N. I. Akhiezer, *Elements of the theory of elliptic functions*, Translations of Mathematical Monographs **79**, American Mathematical Society, Providence, RI, 1990. MR 91k:33016 Zbl 0694.33001

[Baker et al. 1976] G. R. Baker, P. G. Saffman, and J. S. Sheffield, "Structure of a linear array of hollow vortices of finite cross-section", *J. Fluid Mech.* **74**:3 (1976), 469–476. Zbl 0343.76004

[Brelot 1971] M. Brelot, *On topologies and boundaries in potential theory*, Lecture Notes in Mathematics **175**, Springer, Berlin, 1971. MR 43 #7654 Zbl 0222.31014

[Carleman 1926] T. Carleman, "Extension d'un théorème de Liouville", *Acta Math.* **48**:3-4 (1926), 363–366. MR 1555232 JFM 52.0316.02

[Crowdy and Green 2011] D. G. Crowdy and C. C. Green, "Analytical solutions for von Kármán streets of hollow vortices", *Phys. Fluids* **23**:12 (2011), Article ID #126602.

[Gustafsson 1987] B. Gustafsson, "Application of half-order differentials on Riemann surfaces to quadrature identities for arc-length", *J. Analyse Math.* **49** (1987), 54–89. MR 89b:30032 Zbl 0652.30029

[Hauswirth et al. 2011] L. Hauswirth, F. Hélein, and F. Pacard, "On an overdetermined elliptic problem", *Pacific J. Math.* **250**:2 (2011), 319–334. MR 2012g:58046 Zbl 1211.35207

[Heins 1950] M. Heins, "A lemma on positive harmonic functions", *Ann. of Math.* (2) **52** (1950), 568–573. MR 12,259b Zbl 0045.18803

[Khavinson et al. 2013] D. Khavinson, E. Lundberg, and R. Teodorescu, "An overdetermined problem in potential theory", *Pacific J. Math.* **265**:1 (2013), 85–111. MR 3095114 Zbl 1279.31003

[Kjellberg 1950] B. Kjellberg, "On the growth of minimal positive harmonic functions in a plane region", *Ark. Mat.* **1** (1950), 347–351. MR 12,410f Zbl 0040.05502

[Levin 1980] B. Y. Levin, *Distribution of zeros of entire functions*, Translations of Mathematical Monographs **5**, American Mathematical Society, Providence, RI, 1980. Revised edition of 1964 original. MR 81k:30011 Zbl 0152.06703

[Longuet-Higgins 1988] M. S. Longuet-Higgins, "Limiting forms for capillary-gravity waves", *J. Fluid Mech.* **194** (1988), 351–375. MR 90f:76026 Zbl 0649.76005

[Pocklington 1895] H. C. Pocklington, "The configuration of a pair of equal and opposite hollow straight vortices, of finite cross-section, moving steadily through fluid", *Proc. Camb. Philos. Soc.* **8** (1895), 178–187. JFM 25.1468.02

[Rashkovskii 2009] A. Rashkovskii, "Classical and new loglog-theorems", *Expo. Math.* **27**:4 (2009), 271–287. MR 2010j:31001 Zbl 1177.31001

[Sakai 1981] M. Sakai, "Null quadrature domains", *J. Anal. Math.* **40** (1981), 144–154. MR 84e:30069 Zbl 0483.30002

[Traizet 2014a] M. Traizet, "Classification of the solutions to an overdetermined elliptic problem in the plane", *Geom. Funct. Anal.* **24**:2 (2014), 690–720. MR 3192039 Zbl 1295.35344

[Traizet 2014b] M. Traizet, "Hollow vortices and minimal surfaces", preprint, 2014. arXiv 1407.5308

Received June 4, 2014. Revised November 6, 2014.

ALEXANDRE EREMENKO
DEPARTMENT OF MATHEMATICS
PURDUE UNIVERSITY
150 N. UNIVERSITY STREET
WEST LAFAYETTE, IN 47907-2067
UNITED STATES
eremenko@math.purdue.edu

ERIK LUNDBERG
DEPARTMENT OF MATHEMATICAL SCIENCES
FLORIDA ATLANTIC UNIVERSITY
777 GLADES ROAD
BOCA RATON, FL 33431
UNITED STATES
elundber@fau.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.185

ENDOSCOPIC TRANSFER FOR UNITARY GROUPS AND HOLOMORPHY OF ASAI L-FUNCTIONS

NEVEN GRBAC AND FREYDOON SHAHIDI

The analytic properties of the complete Asai L-functions attached to cuspidal automorphic representations of the general linear group over a quadratic extension of a number field are obtained. The proof is based on the comparison of the Langlands–Shahidi method and Mok's endoscopic classification of automorphic representations of quasisplit unitary groups.

Introduction

In this paper we study the analytic properties of the complete Asai L-function attached to a cuspidal automorphic representation of the general linear group over a quadratic extension of a number field. The approach is based on the Langlands–Shahidi method, combined with the knowledge of the poles of Eisenstein series coming from a recent endoscopic classification of automorphic representations of the quasisplit unitary groups by Mok [2015].

In order to state the main result more precisely, we introduce some notation. Let E/F be a quadratic extension of number fields, and let θ be the unique nontrivial element in the Galois group $\mathrm{Gal}(E/F)$. Let $\mathbb{A}_E$ and $\mathbb{A}_F$ be the rings of adèles of E and F, respectively. Let $\hat{\delta}$ be any extension to $\mathbb{A}_E^\times/E^\times$ of the quadratic character of $\mathbb{A}_F^\times/F^\times$ attached to E/F by class field theory.

Neven Grbac is supported in part by Croatian Science Foundation under the project 9364 and by University of Rijeka research grant 13.14.1.2.02. Freydoon Shahidi is partially supported by NSF grant DMS1162299.

MSC2010: 11F66, 11F70, 11F72, 22E55.

Keywords: Asai L-function, automorphic representation, endoscopic classification, spectral decomposition, quasisplit unitary group.

For a cuspidal automorphic representation σ of $\mathrm{GL}_n(\mathbb{A}_E)$, we denote by σ^θ its Galois conjugate and by $\tilde{\sigma}$ its contragredient representation. We say that σ is Galois self-dual if $\sigma \cong \tilde{\sigma}^\theta$, that is, σ is isomorphic to its Galois conjugate contragredient.

Let $L(s, \sigma, r_A)$ denote the complete Asai L-function attached to σ and the Asai representation r_A via the Langlands–Shahidi method. See Section 2.A for a definition. Our main result on the holomorphy and nonvanishing of the Asai L-function $L(s, \sigma, r_A)$ is the following theorem.

Theorem 4.3. *Let σ be a cuspidal automorphic representation of* $\mathrm{GL}_n(\mathbb{A}_E)$. *Let $L(s, \sigma, r_A)$ (respectively, $L(s, \sigma \otimes \hat{\delta}, r_A)$) be the Asai (respectively, twisted Asai) L-function attached to σ, where $\hat{\delta}$ is any extension to $\mathbb{A}_E^\times/E^\times$ of the quadratic character of $\mathbb{A}_F^\times/F^\times$ attached to the extension E/F by class field theory.*

(1) *If σ is* not *Galois self-dual, i.e., if $\sigma \not\cong \tilde{\sigma}^\theta$, then $L(s, \sigma, r_A)$ is entire. It is nonzero for* $\mathrm{Re}(s) \geq 1$ *and* $\mathrm{Re}(s) \leq 0$.

(2) *If σ is Galois self-dual, i.e., if $\sigma \cong \tilde{\sigma}^\theta$, then*

  (a) *$L(s, \sigma, r_A)$ is entire, except for possible simple poles at $s = 0$ and $s = 1$, and nonzero for* $\mathrm{Re}(s) \geq 1$ *and* $\mathrm{Re}(s) \leq 0$;

  (b) *exactly one of the L-functions $L(s, \sigma, r_A)$ and $L(s, \sigma \otimes \hat{\delta}, r_A)$ has simple poles at $s = 0$ and $s = 1$, while the other is holomorphic at those points.*

The idea of the proof is to consider the Eisenstein series attached to σ on the quasisplit unitary group $U_{2n}(\mathbb{A}_F)$ defined by the quadratic extension E/F, where σ is viewed as a representation of the Levi factor of the Siegel maximal parabolic subgroup of U_{2n} in $2n$ variables. We look at the contribution of this Eisenstein series to the residual spectrum from two different points of view. On the one hand, by the Langlands–Shahidi method [2010], the poles of the Eisenstein series for the complex argument in the positive Weyl chamber are determined by certain ratio of the Asai L-functions. The residues at such a pole span a residual representation of $U_{2n}(\mathbb{A}_F)$. On the other hand, this residual representation should have an Arthur parameter, according to Mok's endoscopic classification [2015] of automorphic representations of quasisplit unitary groups (see also [Arthur 2005; 2013]). Comparing the possible Arthur parameters and poles of Asai L-functions, we are able to deduce the analytic properties of these L-functions.

Mok's work, as well as Arthur's, still depends on the stabilization of the twisted trace formula for the general linear group. Hence, our result is also conditional on this stabilization. This issue is considered by Waldspurger [2014a; 2014b; 2014c]. In our paper, we always make a remark when a partial result could have been obtained without using Mok's work. In fact, the crucial insight coming from endoscopic classification is holomorphy of the Asai L-function $L(s, \sigma, r_A)$ inside the critical strip $0 < \mathrm{Re}(s) < 1$.

This method was applied in [Grbac 2011] to the complete exterior and symmetric square L-functions attached to a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_F)$. It relies on Arthur's endoscopic classification of automorphic representations for split classical groups [Arthur 2013; 2005]. The result and the approach are of the same nature as the above theorem for Asai L-functions. The approach has also already been used as a part of a long term project to study endoscopy via descent by Jiang, Liu and Zhang [Jiang et al. 2013].

A different approach to describing the analytic properties of L-functions is that of integral representations. However, this approach usually gives holomorphy of *partial* L-functions, which is weaker than our result due to ramification or problems at archimedean places. For the exterior square L-function this was pursued in [Bump and Friedberg 1990; Kewat and Raghunathan 2012; Belt 2012], for the symmetric square L-function in [Bump and Ginzburg 1992] and more generally for twisted symmetric square in a series of papers [Takeda 2014b; 2014a; 2013], and for the Asai L-functions in [Flicker 1988; Flicker and Zinoviev 1995; Anandavardhanan and Rajan 2005].

The paper is organized as follows. In Section 1 we introduce the unitary group structure and fix the notation. In Section 2 the relation between poles of Eisenstein series on quasisplit unitary groups and the Asai L-functions is investigated. Section 3 provides a definition of Arthur parameters and packets for quasisplit unitary groups in terms of results of Mok. Finally, in Section 4 we prove the main result on the analytic properties of Asai L-functions.

1. The quasisplit unitary groups

1.A. ***Definition and basic structure.*** Let E/F be a quadratic extension of number fields. The nontrivial Galois automorphism in the Galois group $\mathrm{Gal}(E/F)$ is denoted by θ. Let $N_{E/F}$ denote the norm map from E to F. Let $\mathbb{A}_F$ and $\mathbb{A}_E$ be the rings of adèles of F and E, respectively, and $\mathbb{A}_F^\times$ and $\mathbb{A}_E^\times$ the corresponding groups of idèles.

The quadratic character of $\mathbb{A}_F^\times/F^\times$ attached to E/F by class field theory is denoted by $\delta_{E/F}$. We always identify $\delta_{E/F}$ with the corresponding character of the Weil group W_F of F under class field theory. Let $\hat{\delta}$ be any extension of $\delta_{E/F}$ to a character of $\mathbb{A}_E^\times/E^\times$. Such extension is not unique.

We denote by F_v the completion of F at the place v. If a place v of F does not split in E, we always denote by w the unique place of E lying over v. Then E_w/F_v is a quadratic extension of local fields. If v splits in E, we denote by w_1 and w_2 the two places of E lying over v. Then we have $E_{w_1} \cong E_{w_2} \cong F_v$. We use F_∞ to denote the product of F_v over archimedean places.

For an integer $N \geq 2$, we consider in this paper the F-quasisplit unitary group U_N in N variables defined by the extension E/F, viewed as an algebraic group over F. More precisely, U_N is a group scheme over F, whose functor of points is defined as follows. Consider θ as an element of the Galois group $\mathrm{Gal}(\overline{F}/F)$ trivial on $\overline{F}/E$, where $\overline{F}$ is a fixed algebraic closure of F. Let V be an N-dimensional vector space over E. We fix a form on V as in [Kim and Krishnamurthy 2004; 2005], that is, let

$$J_n = \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix} \quad \text{and} \quad J'_N = \begin{cases} \begin{pmatrix} 0 & J_n \\ -J_n & 0 \end{pmatrix} & \text{for } N = 2n, \\ \begin{pmatrix} 0 & 0 & J_n \\ 0 & 1 & 0 \\ -J_n & 0 & 0 \end{pmatrix} & \text{for } N = 2n+1. \end{cases}$$

Then the functor of points of U_N is given by

$$U_N(R) = \{g \in \mathrm{GL}_{E\otimes_F R}(V \otimes_F R) : {}^*g J'_N g = J'_N\}$$

for any F-algebra R, where ${}^*g = {}^tg^\theta$ is the conjugate transpose of g. In particular, the F-points of U_N are given as

$$U_N(F) = \{g \in \mathrm{GL}_N(E) : {}^*g J'_N g = J'_N\}.$$

Writing $N = 2n$ if N is even, and $N = 2n+1$ if N is odd, the F-rank of U_N in both cases equals $n \geq 1$. For $N = 1$, the unitary group U_1 in one variable is obtained by inserting $N = 1$ in the definition of U_N. Its F-points are nothing else than

$$U_1(F) = \{x \in E^\times : \theta(x)x = 1\},$$

which is the norm-one subgroup E^1 of $E^\times$.

For $m \geq 1$, let $G_m = \mathrm{Res}_{E/F}\,\mathrm{GL}_m$ be the algebraic group over F obtained from the general linear group GL_m over E by restriction of scalars from E to F. If $m \leq n$, it appears in the Levi factors of parabolic subgroups of U_N.

We fix the Borel subgroup P_0 of U_N consisting of upper-triangular matrices. Let $P_0 = M_0N_0$, where M_0 is a maximally split maximal torus of U_N (i.e., one containing a maximal split torus of U_N; see [Shahidi 2010, Chapter I]), and N_0 the unipotent radical of P_0. Then

$$M_0 \cong \begin{cases} G_1 \times \cdots \times G_1 & \text{for } N = 2n, \\ G_1 \times \cdots \times G_1 \times U_1 & \text{for } N = 2n+1, \end{cases}$$

with n copies of G_1, so that the F-points of M_0 are given by

$$M_0(F) = \begin{cases} \{\mathrm{diag}(t_1, \ldots, t_n, \theta(t_n)^{-1}, \ldots, \theta(t_1)^{-1}) : t_i \in E^\times\} & \text{for } N = 2n, \\ \{\mathrm{diag}(t_1, \ldots, t_n, t, \theta(t_n)^{-1}, \ldots, \theta(t_1)^{-1}) : t_i \in E^\times, t \in E^1\} & \text{for } N = 2n+1. \end{cases}$$

Let A_0 be a maximal F-split torus of U_N, which is a subtorus of M_0. Then

$$A_0(F) = \begin{cases} \{\mathrm{diag}(t_1, \dots, t_n, t_n^{-1}, \dots, t_1^{-1}) : t_i \in F^\times\} & \text{for } N = 2n, \\ \{\mathrm{diag}(t_1, \dots, t_n, 1, t_n^{-1}, \dots, t_1^{-1}) : t_i \in F^\times\} & \text{for } N = 2n+1. \end{cases}$$

The absolute root system $\Phi = \Phi(U_N, M_0)$ of U_N with respect to M_0 is of type A_{N-1}. The root system $\Phi_{\mathrm{red}} = \Phi(U_N, A_0)$ of U_N with respect to A_0 is a reduced root system. It is of type C_n for $N = 2n$ and of type BC_n for $N = 2n+1$. We make the choice of positive roots according to the fixed Borel subgroup P_0, and let Δ be the set of simple roots. We order the simple roots as in Bourbaki [1968].

Let P be the Siegel maximal proper standard parabolic F-subgroup of U_N. That is, it is defined, in a standard fashion, by a subset of simple roots obtained by removing the last simple root in the Bourbaki ordering (cf. [Bourbaki 1968] and [Shahidi 2010, Section 1.2]). Let $P = M_P N_P$ be the Levi decomposition of P, where

$$M_P \cong \begin{cases} G_n & \text{for } N = 2n, \\ G_n \times U_1 & \text{for } N = 2n+1, \end{cases}$$

is the Levi factor, and N_P the unipotent radical.

1.B. *L-groups.* The L-group of U_N is a semidirect product

$${}^L U_N = \mathrm{GL}_N(\mathbb{C}) \rtimes W_F,$$

where W_F is the Weil group of F. It is acting on the connected component ${}^L U_N^\circ = \mathrm{GL}_N(\mathbb{C})$ through the quotient $W_F/W_E \cong \mathrm{Gal}(E/F)$. The action of the nontrivial Galois automorphism $\theta \in \mathrm{Gal}(E/F)$ is given by

$$\theta(g) = J_N'^{-1}\, {}^t g^{-1} J_N'$$

for all $g \in \mathrm{GL}_N(\mathbb{C})$.

The L-group of the Levi factor M_P is a semidirect product

$${}^L M_P = \begin{cases} \mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C}) \rtimes W_F & \text{for } N = 2n, \\ \mathrm{GL}_n(\mathbb{C}) \times \mathrm{GL}_1(\mathbb{C}) \times \mathrm{GL}_n(\mathbb{C}) \rtimes W_F & \text{for } N = 2n+1, \end{cases}$$

where the Weil group W_F acts through the quotient $W_F/W_E \cong \mathrm{Gal}(E/F)$ on the connected component of the L-group, and $\theta \in \mathrm{Gal}(E/F)$ acts by interchanging the two $\mathrm{GL}_n(\mathbb{C})$ factors.

2. Eisenstein series and Asai L-functions

In this section we relate the analytic behavior of the Eisenstein series on the unitary group supported in the Siegel parabolic subgroup to a ratio of the Asai L-functions appearing in its constant term. For the study of analytic properties of the Asai L-functions, it is sufficient to consider the even quasisplit unitary group U_{2n}. However, for completeness and future reference, we also study the Eisenstein series in the odd case.

We retain all the notation of Section 1. So, P is the Siegel maximal proper standard parabolic F-subgroup of U_N, with the Levi factor $M_P \cong G_n$ if $N = 2n$ is even, and $M_P \cong G_n \times U_1$ if $N = 2n+1$ is odd, and the unipotent radical N_P. Recall that $G_n = \operatorname{Res}_{E/F} \mathrm{GL}_n$.

2.A. ***Asai L-functions.*** Let σ denote a cuspidal automorphic representation of $G_n(\mathbb{A}_F) \cong \mathrm{GL}_n(\mathbb{A}_E)$ and ν a character of $U_1(\mathbb{A}_F) \cong \mathbb{A}_E^1$ trivial on $U_1(F) \cong E^1$. To make a convenient normalization in the case of odd unitary groups, as in [Rogawski 1990] and [Goldberg 1994, Section 6], we denote by $\hat{\nu}$ a unitary character of $\mathrm{GL}_n(\mathbb{A}_E)$ given by

$$\hat{\nu}(g) = \nu(\det(g\,{}^*g^{-1}))$$

for all $g \in \mathrm{GL}_n(\mathbb{A}_E)$. Observe that $\det(g\,{}^*g^{-1})$ is of norm one. Then we define a cuspidal automorphic representation Σ of the Levi factor $M_P(\mathbb{A}_F)$ as

$$\Sigma = \begin{cases} \sigma & \text{for } N = 2n, \\ (\sigma\hat{\nu}) \otimes \nu & \text{for } N = 2n+1. \end{cases}$$

More precisely, in the case of odd unitary groups the action of Σ is given by

$$\Sigma(g, t) = \sigma(g)\nu(\det(g\,{}^*g^{-1}))\nu(t)$$

for $g \in \mathrm{GL}_n(\mathbb{A}_E)$ and $t \in \mathbb{A}_E^1$. We always assume that Σ is irreducible unitary and trivial on $A_P(F_\infty)^\circ$, the identity connected component of $A_P(F_\infty)$, where A_P is a maximal F-split torus in the center of M_P. The last condition is not restrictive. It is just a convenient normalization, obtained by twisting by a unitary character, which makes the poles of Eisenstein series real.

We define first the local L-functions. Let v be a place of F. By extension of scalars from F to F_v, we may view the unitary group U_N as an algebraic group over F_v. This algebraic group is denoted by U_N. Then we have the parabolic subgroup P_v of U_N defined over F_v with Levi decomposition $M_{P,v}N_{P,v}$, where $M_{P,v}$ is the Levi factor and $N_{P,v}$ the unipotent radical.

In the case of the even unitary group, say $N = 2n$, the adjoint representation r_v of the L-group ${}^L M_{P,v}$ on the Lie algebra ${}^L\mathfrak{n}_{P,v}$ of the L-group of $N_{P,v}$ is irreducible for all places v of F. If v does not split in E, then r_v is called the Asai representation, as it generalizes the case considered by Asai in [1977]. We denote it by $r_{A,v}$. This situation is labeled ${}^2A_{2n-1} - 2$ in the list of [Shahidi 1988, Section 4] and [2010, Appendix C]. Explicit action of $r_{A,v}$ is given in [Goldberg 1994, Section 3].

In the case of the odd unitary group, say $N = 2n+1$, the analogous adjoint representation is a direct sum $r_{1,v} \oplus r_{2,v}$ of two irreducible representations for all places v of F, ordered as in [Shahidi 1990]. If v does not split in E, then $r_{2,v}$ is the twisted Asai representation $r_{A,v} \otimes \delta_{E_w/F_v}$, where w is the unique place of E lying over v. This situation is labeled ${}^2A_{2n} - 3$ in the list of [1988, Section 4] and [2010, Appendix C].

For a cuspidal automorphic representation Σ of $M_P(\mathbb{A}_F)$, let $\Sigma \cong \otimes'_v \Sigma_v$ be a decomposition into a restricted tensor product over all places. Let R_v be one of the adjoint representations defined above. Then the local L-functions $L(s, \Sigma_v, R_v)$ attached to Σ_v and R_v are defined as follows.

- At archimedean places, they are the Artin L-functions attached to the Langlands parameter of Σ_v as in [Shahidi 1985] (see also [2010, Section 8.2], and [Langlands 1989] where the Langlands parametrization over reals was first introduced).
- At unramified nonarchimedean places, they are given in terms of Satake parameters of Σ_v (see [Shahidi 1988; 2010, Definition 2.3.5], and also [Harder et al. 1986] where Asai's name came up first).
- At the remaining nonarchimedean places, they are defined using the Langlands–Shahidi method [Shahidi 1990, Section 7] (see also [Shahidi 2010, Section 8.4]).

The corresponding global L-functions are defined as the analytic continuation from the domain of convergence of the product over all places of local L-functions $L(s, \Sigma_v, R_v)$. According to [Langlands 1971] (see also [Shahidi 2010, Section 2.5]), the product over all places defining the global L-functions converges absolutely in some right half-plane $\mathrm{Re}(s) > C$, where C is sufficiently large.

The global L-function obtained in this way from $\Sigma = \sigma \cong \otimes'_v \sigma_v$ and $R_v = r_{1,v}$ is denoted by $L(s, \sigma, r_A)$ and called the Asai L-function attached to σ. Its analytic properties are the main concern of this paper.

The global L-function obtained from $\Sigma \cong \otimes'_v \Sigma_v$ and $R_v = r_{2,v}$ is denoted by $L(s, \Sigma, r_A \otimes \delta_{E/F})$ and called the twisted Asai L-function attached to Σ. In fact, it is the same as the Asai L-function $L(s, \sigma \otimes \hat{\delta}, r_A)$ attached to $\sigma \otimes \hat{\delta}$ (see [Goldberg 1994]). Hence, the analytic properties of the twisted Asai L-function follow from the analytic properties of the Asai L-function attached to a twisted representation. Recall that $\hat{\delta}$ is any extension of the quadratic character $\delta_{E/F}$ to $\mathbb{A}_E^\times$.

Finally, as shown in [Goldberg 1994], the choice of the normalization of Σ in the case of odd unitary groups implies that the global L-function obtained from $\Sigma \cong \otimes'_v \Sigma_v$ and $R_v = r_{1,v}$ is the same as the principal L-function $L(s, \sigma)$ attached to σ by [Godement and Jacquet 1972]. Its analytic properties are well known. It is entire, unless $n = 1$ and σ is the trivial character $\mathbf{1}_{\mathbb{A}_E^\times}$ of $\mathbb{A}_E^\times$. In that case $L(s, \mathbf{1}_{\mathbb{A}_E^\times})$ is holomorphic except for simple poles at $s = 0$ and $s = 1$.

2.B. ***Eisenstein series.*** For $s \in \mathbb{C}$ and Σ a cuspidal automorphic representation of $M_P(\mathbb{A}_F)$ as above, let

$$I(s, \Sigma) = \begin{cases} \mathrm{Ind}_{P(\mathbb{A}_F)}^{U_N(\mathbb{A}_F)}(\sigma |\det|_E^s) & \text{for } N = 2n, \\ \mathrm{Ind}_{P(\mathbb{A}_F)}^{U_N(\mathbb{A}_F)}(\sigma \hat{\nu} |\det|_E^s \otimes \nu) & \text{for } N = 2n+1, \end{cases}$$

be the induced representation, where the induction is normalized. As in [Shahidi 2010, page 108], we realize $I(s, \Sigma)$ for all $s \in \mathbb{C}$ on the same space W_Σ of smooth functions

$$f : N_P(\mathbb{A}_F) M_P(F) A_P(F_\infty)^\circ \backslash U_N(\mathbb{A}_F) \to \mathbb{C},$$

K-finite with respect to a fixed maximal compact subgroup K of $U_N(\mathbb{A}_F)$ compatible to P (as in [Mœglin and Waldspurger 1995, Section I.1.4]), and such that the function on $M_P(\mathbb{A}_F)$ given by the assignment $m \mapsto f(mg)$ for $m \in M_P(\mathbb{A}_F)$ belongs to the space of Σ for all $g \in U_N(\mathbb{A}_F)$. The dependence on $s \in \mathbb{C}$ is hidden in the action of $U_N(\mathbb{A}_F)$.

Given $f \in W_\Sigma$ and $s \in \mathbb{C}$, set

$$f_s(g) = f(g) \exp\langle s + \rho_P, H_P(g)\rangle$$

for all $g \in U_N(\mathbb{A}_F)$. Here ρ_P is the half-sum of positive roots not being the roots of M_P, and H_P is a map

$$H_P : U_N(\mathbb{A}_F) \to \mathrm{Hom}(X(M_P)_F, \mathbb{R}),$$

where $X(M_P)_F$ denotes the group of F-rational characters of M_P, defined on $m = (m_v)_v \in M_P(\mathbb{A}_F)$ by the condition

$$\exp\langle \chi, H_P(m)\rangle = \prod_v |\chi(m_v)|_v$$

for every $\chi \in X(M_P)_F$, and extended via Iwasawa decomposition to $U_N(\mathbb{A}_F)$ trivially on the unipotent radical $N_P(\mathbb{A}_F)$ and the fixed maximal compact subgroup K (cf. [Shahidi 2010, Section 1.3]). Then the Eisenstein series is defined as the analytic continuation from the domain of convergence $\mathrm{Re}(s) > \rho_P$ of the series

$$E(f, s)(g) = \sum_{\gamma \in P(F)\backslash U_N(F)} f(\gamma g) \exp\langle s + \rho_P, H_P(\gamma g)\rangle = \sum_{\gamma \in P(F)\backslash U_N(F)} f_s(\gamma g)$$

for $g \in U_N(\mathbb{A}_F)$. The Eisenstein series $E(f, s)$ has a finite number of simple poles in the real interval $0 < s \leq \rho_P$, and all other poles have $\mathrm{Re}(s) < 0$ (cf. [Mœglin and Waldspurger 1995, Section IV.1.11 and IV.3.12]). The residue of the Eisenstein series $E(f, s)$ at $s > 0$ is a square-integrable automorphic form on $U_N(\mathbb{A}_F)$, but not cuspidal, thus belonging to the residual spectrum of $U_N(\mathbb{A}_F)$. In fact, such residues for all $f \in W_\Sigma$ span the summand of the residual spectrum of $U_N(\mathbb{A}_F)$ with cuspidal support in Σ (see [Mœglin and Waldspurger 1995, Section III.2.6] or [Franke and Schwermer 1998, Section 1] for the decomposition of the space of automorphic forms with respect to their cuspidal support).

2.C. ***Asai L-functions in the constant term.*** Now we prove that the poles of Eisenstein series $E(f,s)(g)$ for $\mathrm{Re}(s) > 0$ coincide with the poles for $\mathrm{Re}(s) > 0$ of the ratio of L-functions appearing in its constant term.

Theorem 2.1. *Let σ be a cuspidal automorphic representation of $G_n(\mathbb{A}_F) \cong \mathrm{GL}_n(\mathbb{A}_E)$ and ν a unitary character of $U_1(\mathbb{A}_F) \cong \mathbb{A}_E^1$ trivial on $U_1(F) \cong E^1$. As in Section 2.B, form a cuspidal automorphic representation Σ of the Levi factor $M_P(\mathbb{A}_F)$ in U_N. Then the poles with* $\mathrm{Re}(s) > 0$ *of the Eisenstein series $E(f,s)$ for some $f \in W_\Sigma$ coincide with the poles satisfying* $\mathrm{Re}(s) > 0$ *of*

$$\begin{cases} \dfrac{L(2s,\sigma,r_A)}{L(1+2s,\sigma,r_A)} & \text{if } N = 2n, \\[2ex] \dfrac{L(s,\sigma)}{L(1+s,\sigma)} \cdot \dfrac{L(2s,\sigma\otimes\hat{\delta},r_A)}{L(1+2s,\sigma\otimes\hat{\delta},r_A)} & \text{if } N = 2n+1, \end{cases}$$

where $\hat{\delta}$ is any extension to $\mathbb{A}_E^\times/E^\times$ of the quadratic character $\delta_{E/F}$ of $\mathbb{A}_F^\times/F^\times$ attached to E/F by class field theory.

Remark 2.2. Observe the factor 2 appearing in the argument $2s$ of the Asai L-function in the case of even unitary groups. The reason is that we have chosen, as in [Shahidi 1992], the determinant character to normalize the identification with $\mathbb{C}$ of the complex parameter s in the Eisenstein series, instead of the character $\tilde{\alpha}$ given in terms of the half-sum of positive roots and the coroot of the unique simple root α not being a root of M_P, as in [Shahidi 1990].

Proof of Theorem 2.1. This is an application of the Langlands spectral theory, using the Langlands–Shahidi method to normalize the intertwining operator.

The poles of the Eisenstein series $E(f,s)$ coincide with the poles of its constant term $E(f,s)_P$ along P. The constant term is defined as

$$E(f,s)_P(g) = \int_{N_P(F)\backslash N_P(\mathbb{A}_F)} E(f,s)(ng)\,dn,$$

where dn is a fixed Haar measure on $N_P(\mathbb{A}_F)$. On the other hand, the constant term can be written as

$$E(f,s)_P(g) = f_s(g) + \big(M(s,\Sigma,w_0)f\big)_{-s}(g),$$

where $M(s,\Sigma,w_0)$ is the standard intertwining operator. Here w_0 is the unique nontrivial Weyl group element such that $w_0(\alpha)$ is a simple root for all simple roots α except the last one in the ordering of [Bourbaki 1968].

As in [Shahidi 2010, page 109], the standard intertwining operator is defined as the analytic continuation from the domain of convergence of the integral

$$M(s,\Sigma,w_0)f(g)=\left(\int_{N_P(\mathbb{A}_F)} f_s(\dot{w}_0^{-1}ng)\,dn\right)\exp\langle s-\rho_P, H_P(g)\rangle,$$

where $\dot{w}_0$ is a fixed representative for w_0 in $U_N(F)$. For $s\in\mathbb{C}$ away from poles, the assignment $f\mapsto M(s,\Sigma,w_0)f$ defines a linear map on W_Σ, which depends on s. It intertwines the actions of $I(s,\Sigma)$ and $I(-s,\Sigma^{w_0})$. Let σ^θ denote σ conjugated by the nontrivial Galois automorphism $\theta\in\mathrm{Gal}(E/F)$, that is, $\sigma^\theta(m)=\sigma(m^\theta)$ for all $m\in\mathrm{GL}_n(\mathbb{A}_E)$. Note that in our case the conjugation by w_0 amounts to taking $\tilde{\sigma}^\theta$, where $\tilde{\sigma}$ is the contragredient of σ. In the case of odd unitary groups this means that $\Sigma^{w_0}\cong\tilde{\sigma}^\theta\hat{\nu}\otimes\nu$ (see [Goldberg 1994]).

It is clear from the expression for the constant term that the poles of the Eisenstein series are the same as those of the standard intertwining operator. We apply the Langlands–Shahidi method to normalize this operator. The normalizing factor in this situation, labeled ${}^2A_{2n-1}-2$ for the even unitary group and ${}^2A_{2n}-3$ for the odd unitary group in the list of [Shahidi 1988, Section 4] and [2010, Appendix C], is given in terms of L-functions and corresponding ε-factors as

$$r(s,\Sigma,w_0)=\begin{cases}\dfrac{L(2s,\sigma,r_A)}{L(1+2s,\sigma,r_A)\varepsilon(2s,\sigma,r_A)} & \text{for } N=2n,\\[2ex] \dfrac{L(s,\sigma)}{L(1+s,\sigma)\varepsilon(s,\sigma)}\cdot\dfrac{L(2s,\sigma\otimes\hat{\delta},r_A)}{L(1+2s,\sigma\otimes\hat{\delta},r_A)\varepsilon(2s,\sigma\otimes\hat{\delta},r_A)} & \text{for } N=2n+1.\end{cases}$$

The normalized intertwining operator

$$r(s,\Sigma,w_0)^{-1}M(s,\Sigma,w_0)$$

is holomorphic and *not* identically vanishing on $I(s,\Sigma)$ for $\mathrm{Re}(s)>0$. This is essentially a local fact proved in Lemma 2.3 below.

Assuming this fact, we now finish the proof. The holomorphy and nonvanishing of the normalized operator implies that the poles of $M(s,\Sigma,w_0)$ for $\mathrm{Re}(s)>0$ coincide with those of $r(s,\Sigma,w_0)$. Since the ε-factors are entire and nonvanishing for all $s\in\mathbb{C}$, these are the same as the poles of the ratios of L-functions given in the theorem. □

2.D. ***Holomorphy and nonvanishing of normalized intertwining operators.*** It remains to show the fact that $r(s,\Sigma,w_0)^{-1}M(s,\Sigma,w_0)$ is holomorphic and nonvanishing for $\mathrm{Re}(s)>0$. The notation is as in the proof of the previous theorem. This is essentially a local problem, because one can decompose over the places of F the action of the standard intertwining operator acting on a decomposable function using the fact that all ingredients are unramified at all but finitely many places.

Hence, the problem reduces to a finite number of ramified and archimedean places, which is solved for each place separately.

We introduce some local notation first. Let $\Sigma \cong \otimes'_v \Sigma_v$ be the decomposition of Σ into a restricted tensor product, where in the case of odd unitary groups $\Sigma_v = \sigma_v \hat{\nu}_v \otimes \nu_v$. We denote the local standard intertwining operator by $M(s, \Sigma_v, w_0)$. It is defined as the analytic continuation of the local analogue of the integral defining the global operator $M(s, \Sigma, w_0)$ (see the proof of Theorem 2.1). Let $r(s, \Sigma_v, w_0)$ be the local factor at v of $r(s, \Sigma, w_0)$. We show in the lemma below that the normalized local intertwining operator

$$N(s, \Sigma_v, w_0) = r(s, \Sigma_v, w_0)^{-1} M(s, \Sigma_v, w_0)$$

is holomorphic and not identically vanishing on the local induced representation $I(s, \Sigma_v)$ for $\mathrm{Re}(s) > 0$.

Lemma 2.3. *Let Σ_v be a local component of a cuspidal automorphic representation Σ of the Levi factor $M_P(\mathbb{A}_F)$ in the unitary group U_N. Then, for* $\mathrm{Re}(s) > 0$, *the normalized local intertwining operator $N(s, \Sigma_v, w_0)$ is holomorphic and not identically vanishing on the induced representation $I(s, \Sigma_v)$.*

Proof. Consider first the case in which the place v of F splits in E. Then $U_N(F_v)$ is isomorphic to $\mathrm{GL}_N(F_v)$, and the Levi factor

$$M_P(F_v) \cong \begin{cases} \mathrm{GL}_n(F_v) \times \mathrm{GL}_n(F_v) & \text{for } N = 2n, \\ \mathrm{GL}_n(F_v) \times \mathrm{GL}_1(F_v) \times \mathrm{GL}_n(F_v) & \text{for } N = 2n+1. \end{cases}$$

Hence, the normalized operator considered in the lemma is attached to a unitary representation of a Levi factor $M_P(F_v)$ in $\mathrm{GL}_N(F_v)$. The holomorphy and nonvanishing for $\mathrm{Re}(s) > 0$ follow from [Mœglin and Waldspurger 1989, Proposition I.10].

We consider now the case in which the place v of F does not split in E, and denote by w the unique place of E lying over v. Then E_w/F_v is a quadratic extension of local fields, and $U_N(F_v)$ is the quasisplit unitary group in N variables given by the extension E_w/F_v. The Levi factor $M_P(F_v)$ is isomorphic to

$$M_P(F_v) \cong \begin{cases} G_n(F_v) \cong \mathrm{GL}_n(E_w) & \text{for } N = 2n, \\ G_n(F_v) \times U_1(F_v) \cong \mathrm{GL}_n(E_w) \times E_w^1 & \text{for } N = 2n+1, \end{cases}$$

so that

$$\Sigma_v \cong \begin{cases} \sigma_w & \text{for } N = 2n, \\ (\sigma_w \hat{\nu}_w) \otimes \nu_w & \text{for } N = 2n+1, \end{cases}$$

where σ_w is the local component of a cuspidal automorphic representation σ of $\mathrm{GL}_n(\mathbb{A}_E)$ at the place w of E, and ν_w the local component of a unitary character ν of $\mathbb{A}_E^1$ trivial on E^1.

In particular, σ_w is unitary and generic, since it is a local component of a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_E)$. Hence, by [Tadić 1986] and [Vogan 1986], in the nonarchimedean and archimedean cases, respectively, there is

- a standard parabolic subgroup Q of GL_n such that the Levi factor M_Q of Q is isomorphic to $\mathrm{GL}_{d_1} \times \cdots \times \mathrm{GL}_{d_\ell}$, where $d_1 + \cdots + d_\ell = n$,
- unitary square-integrable representations δ_i of $\mathrm{GL}_{d_i}(E_w)$, for $i = 1, \ldots, \ell$, and
- real numbers α_i with $0 \le |\alpha_i| < \frac{1}{2}$, for $i = 1, \ldots, \ell$,

such that σ_w is isomorphic to the fully induced representation

$$\sigma_w \cong \mathrm{Ind}_{Q(E_w)}^{\mathrm{GL}_n(E_w)}(\delta_1 |\det|^{\alpha_1} \otimes \cdots \otimes \delta_\ell |\det|^{\alpha_\ell}).$$

Let R be the standard parabolic F-subgroup of U_N with the Levi factor

$$M_R \cong \begin{cases} G_{d_1} \times \cdots \times G_{d_\ell} & \text{for } N = 2n, \\ G_{d_1} \times \cdots \times G_{d_\ell} \times U_1 & \text{for } N = 2n+1, \end{cases}$$

so that $R \subset P$ and $M_R(F_v) = M_Q(E_w)$ for $N = 2n$ and $M_R(F_v) = M_Q(E_w) \times E_w^1$ for $N = 2n+1$. Let

$$\delta = \begin{cases} \delta_1 \otimes \cdots \otimes \delta_\ell & \text{for } N = 2n, \\ \delta_1 \hat{\nu}_1 \otimes \cdots \otimes \delta_\ell \hat{\nu}_\ell \otimes \nu & \text{for } N = 2n+1, \end{cases}$$

be a unitary square-integrable representation of $M_R(F_v)$, where $\hat{\nu}_i$ is the character of $\mathrm{GL}_{d_i}(E_w)$ given by $\hat{\nu}_i(h_i) = \nu(\det(h_i {}^*h_i^{-1}))$ for $h_i \in \mathrm{GL}_{d_i}(E_w)$.

By induction in stages, the intertwining operator $N(s, \Sigma_v, w_0)$ coincides with the intertwining operator

$$N((s+\alpha_1, \ldots, s+\alpha_\ell), \delta, w_0)$$

acting on the induced representation

$$\begin{cases} \mathrm{Ind}_{R(F_v)}^{U_N(F_v)}(\delta_1 |\det|^{s+\alpha_1} \otimes \cdots \otimes \delta_\ell |\det|^{s+\alpha_\ell}) & \text{for } N = 2n, \\ \mathrm{Ind}_{R(F_v)}^{U_N(F_v)}(\delta_1 \hat{\nu}_1 |\det|^{s+\alpha_1} \otimes \cdots \otimes \delta_\ell \hat{\nu}_\ell |\det|^{s+\alpha_\ell} \otimes \nu) & \text{for } N = 2n+1. \end{cases}$$

By Zhang's lemma [1997] (see also [Kim 2000, Lemma 1.7]), the holomorphy of this last operator at s implies nonvanishing. Hence, to show the lemma, it is sufficient to prove the holomorphy for $\mathrm{Re}(s) > 0$.

To prove the holomorphy for $\mathrm{Re}(s) > 0$, we decompose the intertwining operator into a product of intertwining operators as in [Shahidi 1981, Section 2.1]. If we show that each factor is holomorphic for $\mathrm{Re}(s) > 0$, then the product is holomorphic for $\mathrm{Re}(s) > 0$ as well, and the lemma is proved. The factors are normalized intertwining operators that can be viewed as intertwining operators on representations induced from appropriate maximal proper parabolic subgroups in certain reductive groups.

In our case these rank-one factors are normalized operators

$$N(2s+\alpha_i+\alpha_j,\, \delta_i\otimes\tilde{\delta}_j^{\theta}),$$

for $1\le i<j\le \ell$, acting on the induced representations

$$\mathrm{Ind}_{Q_{i,j}(E_w)}^{\mathrm{GL}_{d_i+d_j}(E_w)}\big(\delta_i|\det|^{s+\alpha_i}\otimes\tilde{\delta}_j^{\theta}|\det|^{-s-\alpha_j}\big),$$

where $Q_{i,j}$ is the maximal standard proper parabolic subgroup of $\mathrm{GL}_{d_i+d_j}$ with the Levi factor $\mathrm{GL}_{d_i}\times\mathrm{GL}_{d_j}$, and normalized operators

$$\begin{cases} N(s+\alpha_k,\, \delta_k) & \text{for } N=2n,\\ N(s+\alpha_k,\, (\delta_k\hat{\nu}_k)\otimes\nu) & \text{for } N=2n+1,\end{cases}$$

for $1\le k\le\ell$, acting on the induced representation

$$\begin{cases} \mathrm{Ind}_{Q_k(F_v)}^{U_{2d_k}(F_v)}(\delta_k|\det|^{s+\alpha_k}) & \text{for } N=2n,\\ \mathrm{Ind}_{Q_k(F_v)}^{U_{2d_k+1}(F_v)}(\delta_k\hat{\nu}_k|\det|^{s+\alpha_k}\otimes\nu) & \text{for } N=2n+1,\end{cases}$$

where Q_k is the maximal standard proper parabolic subgroup of U_{2d_k} with the Levi factor G_{d_k} if $N=2n$, and of U_{2d_k+1} with the Levi factor $G_{d_k}\times U_1$ if $N=2n+1$. We suppress the Weyl group element from the notation for these intertwining operators, because they are always determined by the maximal parabolic subgroup in question.

According to [Zhang 1997, Section 2], the rank-one normalized intertwining operator is holomorphic when the real part of its complex parameter is greater than the first negative point of reducibility of the induced representation on which it acts. For $\mathrm{Re}(s)>0$, using the bound on α_i, we have

$$\mathrm{Re}(s+\alpha_i+\alpha_j)>-1 \quad\text{and}\quad \mathrm{Re}(s+\alpha_k)>-\tfrac{1}{2}.$$

But these two bounds are precisely the first negative points of reducibility in the cases $Q_{i,j}\subset\mathrm{GL}_{d_i+d_j}$ and $Q_k\subset U_{2d_k}$ or U_{2d_k+1}. This essentially follows from the standard module conjecture, proved in [Vogan 1978] for any quasisplit real group, and thus for complex groups as well, and in [Muić 2001] for quasisplit classical groups over a p-adic field. In [Casselman and Shahidi 1998, Section 5] the reducibility points are determined in terms of local coefficients over any local field. A convenient reference making explicit the first reducibility points of such complementary series using local coefficients for any quasisplit classical group over a local field of characteristic zero is [Lapid et al. 2004, Lemma 2.6 and 2.7]. For the general linear group the reducibility is obtained in [Zelevinsky 1980] over a p-adic field, in [Speh 1981] over reals, and in [Wallach 1979] over complex numbers (see also [Kim 2000, Lemma 2.10]). For the unitary group over a nonarchimedean field, it is obtained in [Goldberg 1994, Section 3 and 6] by applying the general reducibility result of [Shahidi 1990], while at an archimedean place, the L-functions in the local coefficient that control reducibility are the L-functions of the restriction to $\mathbb{R}^\times$ of a character of $\mathbb{C}^\times$ (see [Lapid et al. 2004, Lemma 2.6]). Thus, the rank-one factors are all holomorphic and the lemma is proved. □

Remark 2.4. Kim and Krishnamurthy [2004; 2005] have proved the holomorphy and nonvanishing of normalized intertwining operators for a representation of the Levi factor of any maximal proper parabolic subgroup of U_N, which is a local component of a generic cuspidal automorphic representation. Since in our case all cuspidal automorphic representations of the Levi factor are generic, Lemma 2.3 follows from their work. Their proof uses their stable base change lift and bounds towards the Ramanujan conjecture obtained by Luo, Rudnick and Sarnak [1999] to bound the exponents on the unitary group. In our case these bounds are not required because our unitary factor in the Levi is either trivial or rank zero. This simplifies the proof.

3. Arthur parameters for unitary groups

Our next task is to introduce the notion of Arthur parameters and the endoscopic classification of automorphic representations for the quasisplit unitary group U_N in N variables. We consider both the even and odd case for completeness, although for the application to the analytic properties of the Asai L-functions only the even case is required.

In [Mok 2015], the results of [Arthur 2013] (see also [Arthur 2005]) are extended to the case of quasisplit unitary groups. As in [Arthur 2004], we avoid the conjectural Langlands group by describing the parameters in terms of irreducible constituents of the discrete spectrum of general linear groups. For quasisplit classical groups this approach was taken in [Mœglin 2008].

3.A. ***Arthur parameters.*** Let μ be a Galois self-dual cuspidal automorphic representation of $\mathrm{GL}_m(\mathbb{A}_E)$. One of the crucial results in Mok's proof of endoscopic classification of representations in the discrete spectrum for quasisplit unitary groups is the uniqueness (up to equivalence) of the twisted endoscopic datum associated to μ. This is the content of [Mok 2015, Theorem 2.4.2]. In fact, this unique endoscopic datum is simple, thus, determining a unique sign $\kappa \in \{\pm 1\}$ attached to μ. The parity of the endoscopic datum associated to μ is then defined as $\kappa(-1)^{m-1}$ (as in Section 2.4 of the same reference). Using parity we make the following definition as in [ibid., Theorem 2.5.4] (see also [Gan et al. 2012]).

Definition 3.1. Let μ be a Galois self-dual cuspidal automorphic representation of $\mathrm{GL}_m(\mathbb{A}_E)$. We say that μ is Galois orthogonal (resp. Galois symplectic), if the parity of the unique twisted endoscopic datum associated to μ is $+1$ (resp. -1).

It turns out, as also proved by Mok, that this definition can be rephrased in terms of poles at $s = 1$ of the Asai L-function $L(s, \mu, r_A)$ attached to μ.

Theorem 3.2 [Mok 2015, Theorem 2.5.4(a)]. *Let μ be a Galois self-dual cuspidal automorphic representation of $\mathrm{GL}_m(\mathbb{A}_E)$. Then μ is Galois orthogonal (resp. Galois symplectic) if and only if the Asai L-function $L(s, \mu, r_A)$ (resp. the twisted Asai L-function $L(s, \mu \otimes \hat{\delta}, r_A)$) has a pole at $s = 1$, where $\hat{\delta}$ is any extension to $\mathbb{A}_E^\times/E^\times$ of the quadratic character $\delta_{E/F}$ of $\mathbb{A}_F^\times/F^\times$ attached to E/F by class field theory.*

We are now ready to define global Arthur parameters for the quasisplit unitary group U_N in N variables. We in fact define the square-integrable Arthur parameters, which, according to [ibid., Theorem 2.5.2], parameterize global Arthur packets contributing to the discrete automorphic spectrum of $U_N(\mathbb{A}_F)$. These parameters depend on the choice of certain character of $\mathbb{A}_E^\times$, trivial on $E^\times$, that defines an L-embedding of the L-group of U_N into the L-group of G_N (cf. [ibid., Section 2.1]). Roughly speaking, this character determines whether we view parameters as the stable or twisted base change of a representation in the discrete spectrum. Of course, the decomposition of the discrete spectrum is independent of that choice, and we take it in this paper to be the trivial character of $\mathbb{A}_E^\times$, and suppress it from notation (see [ibid., Theorem 2.5.2]). The reason why Mok considers all possible characters is that they are all required for the induction argument in the proof of endoscopic classification.

Definition 3.3 (Arthur parameters). As before, let U_N be the quasisplit unitary group in N variables given by a quadratic extension E/F of number fields. The set $\Psi_2(U_N)$ of square-integrable global Arthur parameters for U_N is defined as the set of all *unordered* formal sums of formal tensor products of the form

$$\psi = (\mu_1 \boxtimes \nu(n_1)) \boxplus \cdots \boxplus (\mu_\ell \boxtimes \nu(n_\ell)),$$

such that

(i) μ_i is a Galois self-dual cuspidal automorphic representation of $\mathrm{GL}_{m_i}(\mathbb{A}_E)$, that is, $\mu_i \cong \tilde{\mu}_i^\theta$;

(ii) n_i is a positive integer, and $\nu(n_i)$ is the unique n_i-dimensional irreducible algebraic representation of $SL_2(\mathbb{C})$;

(iii) $m_1 n_1 + \cdots + m_\ell n_\ell = N$;

(iv) for $i \neq j$, we have $\mu_i \ncong \mu_j$ or $n_i \neq n_j$, that is, the formal sum ψ is multiplicity free;

(v) representation μ_i is Galois orthogonal (resp. Galois symplectic) if and only if integers n_i and N are of the same parity (resp. different parity).

According to Theorem 3.2, condition (v) is equivalent to the condition (v′) representation μ_i is such that the Asai L-function $L(s, \mu_i, r_A)$ (resp. the twisted Asai L-function $L(s, \mu_i \otimes \hat{\delta}, r_A)$) has a pole at $s = 1$ if and only if integers n_i and N are of the same parity (resp. different parity).

3.B. ***Arthur packets.*** We proceed, following [Mok 2015], to define the local and global Arthur packet associated to a global Arthur parameter $\psi \in \Psi_2(U_N)$. Every global Arthur parameter $\psi \in \Psi_2(U_N)$ gives rise, as in [ibid., Section 2.3], to a local Arthur parameter ψ_v for every place v of F. The local Arthur packet Π_{ψ_v} is a finite multiset of unitary irreducible representations of $U_N(F_v)$ associated to ψ_v in [ibid., Theorem 2.5.1] and the discussion following it. There is a canonical mapping from Π_{ψ_v} to the character group of a certain finite group $\mathcal{S}_{\psi_v}$ attached to ψ_v (for a definition see [ibid., Section 2.2]). For $\pi_v \in \Pi_{\psi_v}$, we denote the corresponding character by η_{π_v}. If $U_N(F_v)$ and π_v are unramified, then η_{π_v} is the trivial character. We are skipping here the details, because our main interest is only in unramified places.

The global Arthur packet Π_ψ associated to $\psi \in \Psi_2(U_N)$ is defined as

$$\Pi_\psi = \{\otimes'_v \pi_v : \pi_v \in \Pi_{\psi_v} \text{ and } \eta_{\pi_v} \text{ is trivial for almost all } v\}.$$

The global packets Π_ψ for all $\psi \in \Psi_2(U_N)$ contain all representations that can possibly appear in the decomposition of the discrete spectrum on $U_N(\mathbb{A}_F)$. There is a subtle further condition identifying elements of Π_ψ that indeed appear in the discrete spectrum (for a precise formulation see [Mok 2015, Theorem 2.5.2]). We do not recall this condition, because for our purposes it is sufficient to work with the full packets Π_ψ.

We now compare a representation in the discrete spectrum on $U_N(\mathbb{A}_F)$ and its Arthur parameter at unramified places. Through the application to residual representations supported in the Siegel maximal parabolic subgroup, this turns out to be crucial for the proof of holomorphy of the Asai L-function inside the critical strip. Given

$$\psi = (\mu_1 \boxtimes \nu(n_1)) \boxplus \cdots \boxplus (\mu_\ell \boxtimes \nu(n_\ell)) \in \Psi_2(U_N),$$

with notation as in Definition 3.3, let S be a finite set of places of F, containing all archimedean places and all nonarchimedean places ramified in E, and such that for all places w of E lying above some $v \notin S$ all $\mu_{i,w}$ are unramified. Then, for $v \notin S$, we attach to ψ a Frobenius–Hecke conjugacy class

$$c_v(\psi) = \begin{cases} \bigoplus_{i=1}^{\ell} \big(c(\mu_{i,w}) \otimes c_w(\nu(n_i))\big) & \text{if } v \text{ is inert and } v|w, \\ \Big(\bigoplus_{i=1}^{\ell} \big(c(\mu_{i,w_1}) \otimes c_{w_1}(\nu(n_i))\big), \bigoplus_{i=1}^{\ell} \big(c(\mu_{i,w_2}) \otimes c_{w_2}(\nu(n_i))\big)\Big) & \\ & \text{if } v \text{ splits into } w_1, w_2, \end{cases}$$

viewed as a semisimple conjugacy class in the L-group of G_N over F_v, where $c(\mu_{i,w}) \in \mathrm{GL}_{m_i}(\mathbb{C})$ is the Satake parameter, and

$$c_w(\nu(n_i)) = \mathrm{diag}\big(q_w^{(n_i-1)/2}, q_w^{(n_i-3)/2}, \dots, q_w^{-(n_i-1)/2}\big),$$

with q_w the cardinality of the residue field of E_w. Observe that $q_w = q_v^2$ if v is inert in E, and $q_{w_1} = q_{w_2} = q_v$ if v splits in E. The conjugacy classes $c_v(\psi)$ for $v \notin S$ may be viewed as the Satake parameters of the unramified constituents at places w of E lying above v of the induced representation

$$\begin{aligned}\mathrm{Ind}_{R(\mathbb{A}_E)}^{\mathrm{GL}_N(\mathbb{A}_E)}\Big(&\mu_1|\det|^{(n_1-1)/2} \otimes \mu_1|\det|^{(n_1-3)/2} \otimes \cdots \otimes \mu_1|\det|^{-(n_1-1)/2} \\ &\otimes\mu_2|\det|^{(n_2-1)/2} \otimes \mu_2|\det|^{(n_2-3)/2} \otimes \cdots \otimes \mu_2|\det|^{-(n_2-1)/2} \otimes \cdots \\ &\otimes\mu_\ell|\det|^{(n_\ell-1)/2} \otimes \mu_\ell|\det|^{(n_\ell-3)/2} \otimes \cdots \otimes \mu_\ell|\det|^{-(n_\ell-1)/2}\Big),\end{aligned}$$

where R is the standard parabolic subgroup of GL_N with the Levi factor $\mathrm{GL}_{m_1} \times \cdots \times \mathrm{GL}_{m_1} \times \mathrm{GL}_{m_2} \times \cdots \times \mathrm{GL}_{m_2} \times \cdots \times \mathrm{GL}_{m_\ell} \times \cdots \times \mathrm{GL}_{m_\ell}$ with n_i copies of GL_{m_i} in the product, and μ_i are unramified at v.

On the other hand, let $\pi \cong \otimes'_v \pi_v$ be an irreducible automorphic representation appearing in the discrete spectrum on $U_N(\mathbb{A}_F)$. Let S' be a finite set of places of F, containing all archimedean places, and such that for $v \notin S'$, we have that $U_N(F_v)$ and π_v are unramified. Then, for $v \notin S'$, the Satake isomorphism gives a Frobenius–Hecke conjugacy class $c(\pi_v)$ in the local L-group of U_N over F_v. However, we may view $c(\pi_v)$ as a conjugacy class in the local L-group of G_N through the stable base change map of L-groups. This is consistent with our choice of the trivial character in the definition of Arthur parameters.

According to the preliminary comparison of spectral sides of the trace formulas for U_N and the twisted trace formula for GL_N, carried out in [Mok 2015, Section 4.3] (see also [Arthur 2013, Section 3.4]), for every irreducible automorphic representation π of $U_N(\mathbb{A}_F)$ appearing in the discrete spectrum, there is a unique corresponding parameter $\psi \in \Psi_2(U_N)$ such that the Frobenius–Hecke conjugacy classes $c_v(\psi)$ attached to ψ coincide at almost all places with the classes $c(\pi_v)$ attached to π. This observation is the key to the following proposition.

Remark 3.4. Strictly speaking the preliminary comparison of trace formulas gives unique ψ in a larger set of parameters $\Psi(U_N)$ (see [Mok 2015] for a definition), but the full proof of endoscopic classification shows that such ψ belongs to $\Psi_2(U_N)$.

Proposition 3.5. *Let P be the Siegel maximal proper parabolic F-subgroup of U_{2n}. Let σ be a cuspidal automorphic representation of its Levi factor*

$$M_P(\mathbb{A}_F) \cong \mathrm{GL}_n(\mathbb{A}_{\mathbb{E}}).$$

If the induced representation

$$\mathrm{Ind}_{P(\mathbb{A}_F)}^{U_{2n}(\mathbb{A}_F)}\big(\sigma \otimes |\det|_{\mathbb{A}_E}^{s}\big)$$

has a constituent in the discrete spectrum of $U_{2n}(\mathbb{A}_F)$ for some $s > 0$, then its Arthur parameter is

$$\psi = \sigma \boxtimes \nu(2),$$

and in particular $s = \frac{1}{2}$ and σ is Galois self-dual.

Proof. Since an automorphic representation is unramified at almost all places, the local component of an irreducible constituent π of the induced representation

$$\mathrm{Ind}_{P(\mathbb{A}_F)}^{U_{2n}(\mathbb{A}_F)}\big(\sigma \otimes |\det|_{\mathbb{A}_E}^{s}\big)$$

belonging to the discrete spectrum is the unramified representation with the Satake parameter, viewed as a conjugacy class in the L-group of G_{2n} as above,

$$c(\pi_v) = \begin{cases} c(\sigma_w) \otimes \mathrm{diag}(q_w^{s}, q_w^{-s}) & \text{if } v \text{ is inert and } v|w, \\ \big(c(\sigma_{w_1}) \otimes \mathrm{diag}(q_{w_1}^{s}, q_{w_1}^{-s}), c(\sigma_{w_2}) \otimes \mathrm{diag}(q_{w_2}^{s}, q_{w_2}^{-s})\big) & \text{if } v \text{ splits into } w_1, w_2, \end{cases}$$

for almost all places v of F. Recall that $q_w = q_v^2$ if v is inert, and $q_{w_1} = q_{w_2} = q_v$ if v splits. We may also view $c(\pi_v)$ as the Satake parameter of the unramified constituent of the local components at places w of E lying over v of the induced representation

$$\mathrm{Ind}_{Q(\mathbb{A}_E)}^{\mathrm{GL}_{2n}(\mathbb{A}_E)}\big(\sigma |\det|_{\mathbb{A}_E}^{s} \otimes \sigma |\det|_{\mathbb{A}_E}^{-s}\big),$$

where Q is the standard parabolic subgroup of GL_{2n} with the Levi factor $\mathrm{GL}_n \times \mathrm{GL}_n$.

By the observation made just before the statement of the proposition, these Frobenius–Hecke conjugacy classes $c(\pi_v)$, viewed as conjugacy classes in the L-group of G_{2n}, should match at almost all places the conjugacy classes $c_v(\psi)$ attached to the Arthur parameter $\psi \in \Psi_2(U_N)$ parameterizing π. As mentioned above, these $c_v(\psi)$ may be viewed as Satake parameters of the unramified constituent at v of certain induced representation of $\mathrm{GL}_{2n}(\mathbb{A}_E)$. However, by the strong multiplicity one for general linear groups [Jacquet and Shalika 1981, Theorem 4.4], matching of Satake parameters at almost all places for induced representations of $\mathrm{GL}_N(\mathbb{A}_E)$ implies that the inducing data for these representations are associate. Since Q is self associate, this means that the parabolic subgroup R determined by ψ as above must be Q, and thus that ψ is of the form

$$\psi = \sigma \boxtimes \nu(k),$$

where $k = 2s + 1$. Since $k = 2$ by condition (iii) in Definition 3.3, it follows that $s = \frac{1}{2}$. As σ appears in ψ it is necessarily Galois self-dual. □

4. Holomorphy and nonvanishing of Asai L-functions

In this section we prove the analytic properties of the Asai L-functions as a consequence of Mok's endoscopic classification [2015] of automorphic representations of a quasisplit unitary group.

4.A. ***Analytic properties of Eisenstein series.*** The first task is to determine the poles of Eisenstein series $E(f, s)$ for $\mathrm{Re}(s) > 0$. We now consider only the case of even quasisplit unitary group U_{2n}.

Recall that for a cuspidal automorphic representation σ of $\mathrm{GL}_n(\mathbb{A}_E)$, we let σ^θ denote σ conjugated by the nontrivial Galois automorphism $\theta \in \mathrm{Gal}(E/F)$. We say that σ is Galois self-dual if it is isomorphic to $\tilde{\sigma}^\theta$, where $\tilde{\sigma}$ is the contragredient of σ.

Theorem 4.1. *Let σ be a cuspidal automorphic representation of the Levi factor $M_P(\mathbb{A}_F) \cong \mathrm{GL}_n(\mathbb{A}_E)$ in U_{2n}. Then the Eisenstein series $E(f, s)$ on $U_{2n}(\mathbb{A}_F)$, constructed as in Section 2.B from functions f in the representation space W_σ on which induced representations $I(s, \sigma)$ are realized for all s, is*

(1) *holomorphic for* $\mathrm{Re}(s) \geq 0$, *if σ is not Galois self-dual*,

(2) *holomorphic for* $\mathrm{Re}(s) \geq 0$, *except for a possible simple pole at $s = \frac{1}{2}$, if σ is Galois self-dual.*

Proof. The Eisenstein series is holomorphic on the imaginary axis $\mathrm{Re}(s) = 0$ (see [Mœglin and Waldspurger 1995, Section IV.1.11]). Hence, we may assume $\mathrm{Re}(s) > 0$. Suppose that the Eisenstein series $E(f, s)$ on $U_{2n}(\mathbb{A}_F)$ has a pole at $s = s_0 > 0$ for some $f \in W_\sigma$ in the notation of Section 2. Since $s_0 > 0$, the residues at $s = s_0$ of $E(f, s)$ when $f \in W_\sigma$ span a residual automorphic representation of $U_{2n}(\mathbb{A}_F)$. But this residual representation is a constituent of the induced representation

$$\mathrm{Ind}_{P(\mathbb{A}_F)}^{U_{2n}(\mathbb{A}_F)}\big(\sigma \otimes |\det|_E^{s_0}\big).$$

By Proposition 3.5, its Arthur parameter is

$$\psi = \sigma \boxtimes \nu(2),$$

where σ is Galois self-dual and $s_0 = \frac{1}{2}$. Therefore, the Eisenstein series $E(f, s)$ is holomorphic for $\mathrm{Re}(s) > 0$, except for a possible pole at $s = \frac{1}{2}$ if σ is Galois self-dual, as claimed. The possible pole is simple, by the general theory of Eisenstein series [Mœglin and Waldspurger 1995, Section IV.1.11] □

Remark 4.2. A significant part of Theorem 4.1 can be proved in a different way, without using Mok's work on the Arthur classification for unitary groups [Mok 2015], which is based on the trace formula, and still depends on the stabilization of the twisted trace formula for GL_n.

For instance, if σ is not Galois self-dual, the following general argument provides holomorphy of the Eisenstein series for $\mathrm{Re}(s) > 0$. By [Harish-Chandra 1968], see also [Mœglin and Waldspurger 1995, Section IV.3.12], a necessary condition for the Eisenstein series $E(f, s)$ to have a pole for $\mathrm{Re}(s) > 0$ and some $f \in W_\sigma$ is that $\sigma^{w_0} \cong \sigma$. But in our case, $\sigma^{w_0} = \tilde{\sigma}^\theta$, so that $E(f, s)$ is holomorphic for $\mathrm{Re}(s) > 0$ and all $f \in W_\sigma$ if σ is not Galois self-dual.

If σ is Galois self-dual, there is a unitarity argument, which gives the analytic behavior of the Eisenstein series for $\mathrm{Re}(s) \geq \frac{1}{2}$. However, the critical strip $0 < \mathrm{Re}(s) < \frac{1}{2}$ remains out of reach. For completeness, we include this argument in Section 4.C below.

4.B. *Analytic properties of Asai L-functions.* The following theorem describes completely the analytic properties of the Asai L-functions attached to a cuspidal automorphic representation σ of $\mathrm{GL}_n(\mathbb{A}_E)$. It is the main result of the paper.

Theorem 4.3. *Let σ be a cuspidal automorphic representation of $\mathrm{GL}_n(\mathbb{A}_E)$. Let $L(s, \sigma, r_A)$ (respectively, $L(s, \sigma \otimes \hat{\delta}, r_A)$) be the Asai (respectively, twisted Asai) L-function attached to σ, where $\hat{\delta}$ is any extension to $\mathbb{A}_E^\times/E^\times$ of the quadratic character of $\mathbb{A}_F^\times/F^\times$ attached to the extension E/F by class field theory.*

(1) *If σ is* not *Galois self-dual, that is, $\sigma \ncong \tilde{\sigma}^\theta$, then $L(s, \sigma, r_A)$ is entire. It is nonzero for* $\mathrm{Re}(s) \geq 1$ *and* $\mathrm{Re}(s) \leq 0$.

(2) *If σ is Galois self-dual, that is, $\sigma \cong \tilde{\sigma}^\theta$, then*

   (a) *$L(s, \sigma, r_A)$ is entire, except for possible simple poles at $s = 0$ and $s = 1$, and nonzero for* $\mathrm{Re}(s) \geq 1$ *and* $\mathrm{Re}(s) \leq 0$;

   (b) *exactly one of the L-functions $L(s, \sigma, r_A)$ and $L(s, \sigma \otimes \hat{\delta}, r_A)$ has simple poles at $s = 0$ and $s = 1$, while the other is holomorphic at those points.*

Proof. The idea of the proof goes back to [Shahidi 1981; 1988]. The proof of holomorphy is based on Theorem 2.1, which relates the poles of Eisenstein series to the Asai L-functions, and Theorem 4.1 providing the analytic behavior of the Eisenstein series. The nonvanishing, on the other hand, follows from considering the nonconstant term of the Eisenstein series as in [Shahidi 1981] (see also [Shahidi 2010, Section 7]), and using Theorem 4.1 again. It is sufficient to prove the claims for $\mathrm{Re}(s) \geq \frac{1}{2}$, due to the functional equation for Asai L-functions.

We begin with the proof of holomorphy. Consider first the case of σ not Galois self-dual. According to Theorem 4.1, the Eisenstein series attached to σ is holomorphic for $\mathrm{Re}(s) > 0$. Assume that $L(s, \sigma, r_A)$ has a pole for $s = s_0 > 0$. Since the poles of $E(f, s)$ for $\mathrm{Re}(s) > 0$ coincide, according to Theorem 2.1, with the poles of the ratio

$$(*) \qquad \frac{L(2s, \sigma, r_A)}{L(1 + 2s, \sigma, r_A)},$$

the pole of the numerator at $2s = s_0 > 0$ should be canceled by a pole in the denominator. Thus, $L(z, \sigma, r_A)$ should have a pole at $z = s_0 + 1$. Repeating this argument, we obtain a sequence of poles of the Asai L-function of the form $s_0 + M$, where M is any nonnegative integer. This is a contradiction, because $L(s, \sigma, r_A)$ is holomorphic in the right half-plane of absolute convergence of the defining product. Thus, we proved that $L(s, \sigma, r_A)$ is entire.

Consider now the case of σ Galois self-dual. By Theorem 4.1, the Eisenstein series $E(f, s)$ attached to σ is holomorphic for $\mathrm{Re}(s) > 0$, except for a possible simple pole at $s = \frac{1}{2}$. The same argument as in the previous case implies that $L(z, \sigma, r_A)$ is holomorphic for $\mathrm{Re}(z) > 0$, except for $z = 1$ if the Eisenstein series has a pole at $s = \frac{1}{2}$.

To prove that a possible pole of $L(z, \sigma, r_A)$ at $z = 1$ is at most simple, we again apply a similar argument. Suppose $E(f, s)$ has a pole at $s = \frac{1}{2}$. It is simple by Theorem 4.1. If $L(z, \sigma, r_A)$ had a higher order pole at $z = 2s = 1$, then Theorem 2.1 would imply that there is a pole in the denominator of the ratio of Asai L-functions in $(*)$. But this would mean that the Asai L-function has a pole at $z + 1 = 2$. The Eisenstein series is holomorphic at $s = 1$, so that the same argument as before gives a sequence of poles at all positive integers, which is a contradiction.

For nonvanishing, consider the nonconstant term $E(f, s)_\psi$ of the Eisenstein series $E(f, s)$ with respect to a fixed nontrivial additive character ψ of $F\backslash\mathbb{A}_F$. According to [Shahidi 2010, Theorem 7.1.2], we have

$$E(f, s)_\psi(e) = \frac{1}{L^S(1 + 2s, \sigma, r_A)} \cdot \prod_{v \in S} W_v(e_v),$$

where e and e_v are the identity matrices, W_v is the ψ_v-Whittaker function attached to f via a Jacquet integral, S is a finite set of places, containing all archimedean places, outside which $U_{2n}(F_v)$, σ_v and ψ_v are all unramified, and $L^S(z, \sigma, r_A)$ is the partial Asai L-function attached to σ. As in [Shahidi 2010, Section 7.2], there is a choice of $f \in W_\sigma$ such that $W_v(e_v) \neq 0$ for all $v \in S$. Thus, every zero of $L^S(1 + 2s, \sigma, r_A)$ for $\mathrm{Re}(s) \geq 0$, equivalently $\mathrm{Re}(1 + 2s) \geq 1$, would give a pole of the nonconstant term $E(f, s)_\psi$. However, by Theorem 4.1, the Eisenstein series $E(f, s)$, and thus $E(f, s)_\psi$ as well, is holomorphic for $\mathrm{Re}(s) \geq 0$, except for a possible pole at $s = \frac{1}{2}$, which may occur only if σ is Galois self-dual. Hence, $L^S(z, \sigma, r_A)$ has no zeroes for $\mathrm{Re}(z) \geq 1$, except possibly for $z = 1$. Since the local L-functions are nonvanishing, the same holds for the complete Asai L-function $L(z, \sigma, r_A)$.

For σ Galois self-dual, the nonvanishing of $L(z, \sigma, r_A)$ at the remaining point $z = 1$ follows from the identity

$$L(s, \sigma \times \sigma^\theta) = L(s, \sigma, r_A) L(s, \sigma, r_A \otimes \delta_{E/F}), \tag{$**$}$$

where $L(s, \sigma \times \sigma^\theta)$ is the Rankin–Selberg L-function, and recall that the twisted Asai L-function equals

$$L(s, \sigma, r_A \otimes \delta_{E/F}) = L(s, \sigma \otimes \hat{\delta}, r_A).$$

See [Goldberg 1994] for these identities. The poles of the Rankin–Selberg L-function $L(s, \sigma \times \sigma^\theta)$ are known from [Jacquet and Shalika 1981]. For σ Galois self-dual it has a simple pole at $s = 1$. Since $\sigma \otimes \hat{\delta}$ is Galois self-dual as well, we already proved that both Asai L-functions on the right-hand side of $(**)$ have at most a simple pole at $s = 1$. Hence, they are both nonzero at $s = 1$, and exactly one of them has a simple pole at $s = 1$, as claimed. □

Remark 4.4. Once the holomorphy of the Asai and twisted Asai L-function is known at some s_0 with $\mathrm{Re}(s_0) > 0$, the argument using the Rankin–Selberg L-function at the end of this proof can be applied directly to obtain nonvanishing. However, the result of Jacquet and Shalika [1981] providing analytic properties of the Rankin–Selberg L-functions is very deep, and we preferred to give an argument using nonconstant term of the Eisenstein series whenever possible.

4.C. ***Holomorphy of Eisenstein series using a unitarity argument.*** We now give a different proof that the Eisenstein $E(f, s)$, attached to a Galois self-dual cuspidal automorphic representation σ of $\mathrm{GL}_n(\mathbb{A}_F)$ as above, is holomorphic for $\mathrm{Re} \geq \frac{1}{2}$, except for a possible simple pole at $s = \frac{1}{2}$.

It is sufficient to prove that $E(f, s)$ is holomorphic for $\mathrm{Re}(s) > \frac{1}{2}$. Indeed, since we always normalize σ to be trivial on $A_P(F_\infty)^\circ$, the poles of the Eisenstein series are real. Hence, the only possible pole for $\mathrm{Re}(s) = \frac{1}{2}$ is at $s = \frac{1}{2}$. It is at most simple pole, because all poles of Eisenstein series inside the closure of the positive Weyl chamber are without multiplicity [Mœglin and Waldspurger 1995, Section IV.1.11].

Suppose that there is a simple pole of $E(f, s)$ at $s = s_0 > \frac{1}{2}$. We follow an idea of Kim [2000] based on the fact that residual representations are unitary. The space of residues of $E(f, s)$ at $s = s_0$ is a residual representation of $U_{2n}(\mathbb{A}_F)$, which is a constituent of the induced representation

$$I(s_0, \sigma) = \mathrm{Ind}_{P(\mathbb{A}_F)}^{U_{2n}(\mathbb{A}_F)}\big(\sigma |\det|_E^{s_0}\big).$$

In particular, this residual representation is unitary, so that the induced representation should have a unitary constituent. But then the local induced representation at every place v should have a unitary subquotient. Let v be a split nonarchimedean place of F such that σ_v is unramified. The local induced representation at v is isomorphic to

$$I(s_0, \sigma_v) \cong \mathrm{Ind}_{P(F_v)}^{\mathrm{GL}_{2n}(F_v)}\big(\sigma_{w_1} |\det|_{F_v}^{s_0} \otimes \tilde{\sigma}_{w_2} |\det|_{F_v}^{-s_0}\big),$$

where w_1 and w_2 are the two places of E lying above v. Since σ_{w_1} and σ_{w_2} are unramified unitary generic representations of $\mathrm{GL}_n(F_v)$, according to [Tadić 1986], they are fully induced representations of the form

$$\sigma_{w_1} \cong \mathrm{Ind}_{B_n(F_v)}^{\mathrm{GL}_n(F_v)}\big(\mu_1|\ |^{\alpha_1} \otimes \cdots \otimes \mu_k|\ |^{\alpha_k} \otimes \chi_1 \otimes \cdots \otimes \chi_l \otimes \mu_k|\ |^{-\alpha_k} \otimes \cdots \otimes \mu_1|\ |^{-\alpha_1}\big)$$

and

$$\tilde{\sigma}_{w_2} \cong \mathrm{Ind}_{B_n(F_v)}^{\mathrm{GL}_n(F_v)}\big(\mu'_1|\ |^{\beta_1} \otimes \cdots \otimes \mu'_{k'}|\ |^{\beta_{k'}} \otimes \chi'_1 \otimes \cdots \otimes \chi'_{l'} \otimes \mu'_{k'}|\ |^{-\beta_{k'}} \otimes \cdots \otimes \mu'_1|\ |^{-\beta_1}\big),$$

where B_n is a Borel subgroup of GL_n, the exponents satisfy $0 < \alpha_k < \cdots < \alpha_1 < \frac{1}{2}$ and $0 < \beta_{k'} < \cdots < \beta_1 < \frac{1}{2}$, and μ_i, μ'_i, χ_j, χ'_j are unramified unitary characters of $F_v^\times$. Hence,

$$\begin{aligned} I(s_0, \sigma_v) \cong \mathrm{Ind}_{B_{2n}(F_v)}^{\mathrm{GL}_{2n}(F_v)}\big(\mu_1|\ |^{s_0+\alpha_1} \otimes \cdots \otimes \mu_k|\ |^{s_0+\alpha_k} \otimes \chi_1|\ |^{s_0} \otimes \cdots \otimes \chi_l|\ |^{s_0} \\ \otimes \mu_k|\ |^{s_0-\alpha_k} \otimes \cdots \otimes \mu_1|\ |^{s_0-\alpha_1} \\ \otimes \mu'_1|\ |^{-s_0+\beta_1} \otimes \cdots \otimes \mu'_{k'}|\ |^{-s_0+\beta_{k'}} \otimes \chi'_1|\ |^{-s_0} \otimes \cdots \\ \otimes \chi'_{l'}|\ |^{-s_0} \otimes \mu'_{k'}|\ |^{-s_0-\beta_{k'}} \otimes \cdots \otimes \mu'_1|\ |^{-s_0-\beta_1}\big). \end{aligned}$$

According to the description of the unitary dual of $\mathrm{GL}_{2n}(F_v)$ [Tadić 1986], this representation would have a unitary subquotient, only if all the exponents whose absolute value is not smaller than $\frac{1}{2}$, induced with another character to a representation of $\mathrm{GL}_2(F_v)$, give a reducible representation with a unitary quotient of Speh type. However, this is possible only if for every such exponent that is not less than $\frac{1}{2}$ in absolute value, there is another exponent such that their difference is exactly 1.

Having this in mind, consider the largest exponent in the above induced representation. We write this exponent as $s_0+\alpha_1$, and allow the possibility $\alpha_1 = 0$, which happens in the case $k = 0$ as there are no α_i's. There should be another exponent of the form $-s_0 \pm \beta$, where $\beta = \beta_j$ for some j or $\beta = 0$, such that

$$(s_0+\alpha_1) - (-s_0 \pm \beta) = 1.$$

But this implies

$$2s_0 + \alpha_1 \mp \beta = 1,$$

which is possible for $s_0 > \frac{1}{2}$ only if the sign of β is minus and $\alpha_1 < \beta$. As β is certainly not greater than the largest of β_j's, it follows that necessarily $\alpha_1 < \beta_1$. However, considering the smallest exponent in the induced representation, that is, $-s_0 - \beta_1$, where again β_1 is set to zero if $l = 0$, we obtain the opposite inequality, $\beta_1 < \alpha_1$. This is a contradiction, proving that $I(s_0, \sigma_v)$ has not a unitary subquotient for $s_0 > \frac{1}{2}$, and therefore, the Eisenstein series $E(f, s)$ has no pole for $\mathrm{Re}(s) > \frac{1}{2}$, as claimed.

Acknowledgements

This paper was finalized during our visit to the Erwin Schrödinger Institute in Vienna for the workshop *Advances in the Theory of Automorphic Forms and their L-Functions*. We are grateful for the invitation and the ever stimulating atmosphere of the Erwin Schrödinger Institute. The first named author would like to thank Goran Muić for some useful discussions.

References

[Anandavardhanan and Rajan 2005] U. K. Anandavardhanan and C. S. Rajan, "Distinguished representations, base change, and reducibility for unitary groups", *Int. Math. Res. Not.* **2005**:14 (2005), 841–854. MR 2006g:22013 Zbl 1070.22011

[Arthur 2004] J. Arthur, "Automorphic representations of GSp(4)", pp. 65–81 in *Contributions to automorphic forms, geometry, and number theory* (Baltimore, MD, 2002), edited by H. Hida et al., Johns Hopkins University Press, Baltimore, MD, 2004. MR 2005d:11074 Zbl 1080.11037

[Arthur 2005] J. Arthur, "An introduction to the trace formula", pp. 1–263 in *Harmonic analysis, the trace formula, and Shimura varieties*, edited by J. Arthur et al., Clay Math. Proc. **4**, Amer. Math. Soc., Providence, RI, 2005. MR 2007d:11058 Zbl 1152.11021

[Arthur 2013] J. Arthur, *The endoscopic classification of representations: orthogonal and symplectic groups*, Amer. Math. Soc. Colloq. Publ. **61**, Amer. Math. Soc., Providence, RI, 2013. MR 3135650 Zbl 06231010

[Asai 1977] T. Asai, "On certain Dirichlet series associated with Hilbert modular forms and Rankin's method", *Math. Ann.* **226**:1 (1977), 81–94. MR 55 #2761 Zbl 0326.10024

[Belt 2012] D. D. Belt, *On the holomorphy of exterior-square L-functions*, thesis, Purdue University, West Lafayette, IN, 2012, available at http://search.proquest.com/docview/1237150822. MR 3103748

[Bourbaki 1968] N. Bourbaki, *Éléments de mathématique, Fasc. XXXIV: Groupes et algèbres de Lie, Chapitre IV: Groupes de Coxeter et systèmes de Tits. Chapitre V: Groupes engendrés par des réflexions. Chapitre VI: Systèmes de racines*, Actualités Scientifiques et Industrielles **1337**, Hermann, Paris, 1968. Translated as *Lie groups and Lie algebras: chapters 4–6*, Springer, Berlin, 2002. MR 39 #1590 Zbl 0186.33001

[Bump and Friedberg 1990] D. Bump and S. Friedberg, "The exterior square automorphic L-functions on $\mathrm{GL}(n)$", pp. 47–65 in *Festschrift in honor of I. I. Piatetski-Shapiro on the occasion of his sixtieth birthday, II: Papers in analysis, number theory and automorphic L-functions* (Tel Aviv, 1989), edited by S. Gelbart et al., Israel Math. Conf. Proc. **3**, Weizmann, Jerusalem, 1990. MR 93d:11050 Zbl 0712.11030

[Bump and Ginzburg 1992] D. Bump and D. Ginzburg, "Symmetric square L-functions on $\mathrm{GL}(r)$", *Ann. of Math.* (2) **136**:1 (1992), 137–205. MR 93i:11058 Zbl 0753.11021

[Casselman and Shahidi 1998] W. Casselman and F. Shahidi, "On irreducibility of standard modules for generic representations", *Ann. Sci. École Norm. Sup.* (4) **31**:4 (1998), 561–589. MR 99f:22028 Zbl 0947.11022

[Flicker 1988] Y. Z. Flicker, "Twisted tensors and Euler products", *Bull. Soc. Math. France* **116**:3 (1988), 295–313. MR 89m:11049 Zbl 0674.10026

[Flicker and Zinoviev 1995] Y. Z. Flicker and D. Zinoviev, "On poles of twisted tensor L-functions", *Proc. Japan Acad. Ser. A Math. Sci.* **71**:6 (1995), 114–116. MR 96f:11075 Zbl 0983.11023

[Franke and Schwermer 1998] J. Franke and J. Schwermer, "A decomposition of spaces of automorphic forms, and the Eisenstein cohomology of arithmetic groups", *Math. Ann.* **311**:4 (1998), 765–790. MR 99k:11077 Zbl 0924.11042

[Gan et al. 2012] W. T. Gan, B. H. Gross, and D. Prasad, "Symplectic local root numbers, central critical L values, and restriction problems in the representation theory of classical groups", pp. 1–109 in *Sur les conjectures de Gross et Prasad*, vol. I, Astérisque **346**, Société Mathématique de France, Paris, 2012. MR 3202556 Zbl 1280.22019 arXiv 0909.2999

[Godement and Jacquet 1972] R. Godement and H. Jacquet, *Zeta functions of simple algebras*, Lecture Notes in Math. **260**, Springer, Berlin, 1972. MR 49 #7241 Zbl 0244.12011

[Goldberg 1994] D. Goldberg, "Some results on reducibility for unitary groups and local Asai L-functions", *J. Reine Angew. Math.* **448** (1994), 65–95. MR 95g:22031 Zbl 0815.11029

[Grbac 2011] N. Grbac, "On the residual spectrum of split classical groups supported in the Siegel maximal parabolic subgroup", *Monatsh. Math.* **163**:3 (2011), 301–314. MR 2012k:11063 Zbl 1247.11080

[Harder et al. 1986] G. Harder, R. P. Langlands, and M. Rapoport, "Algebraische Zyklen auf Hilbert–Blumenthal-Flächen", *J. Reine Angew. Math.* **366** (1986), 53–120. MR 87k:11066 Zbl 0575.14004

[Harish-Chandra 1968] Harish-Chandra, *Automorphic forms on semisimple Lie groups*, Lecture Notes in Math. **62**, Springer, Berlin, 1968. MR 38 #1216 Zbl 0186.04702

[Jacquet and Shalika 1981] H. Jacquet and J. A. Shalika, "On Euler products and the classification of automorphic forms, II", *Amer. J. Math.* **103**:4 (1981), 777–815. MR 82m:10050b Zbl 0491.10020

[Jiang et al. 2013] D. Jiang, B. Liu, and L. Zhang, "Poles of certain residual Eisenstein series of classical groups", *Pacific J. Math.* **264**:1 (2013), 83–123. MR 3079762 Zbl 06203663

[Kewat and Raghunathan 2012] P. K. Kewat and R. Raghunathan, "On the local and global exterior square L-functions of GL_n", *Math. Res. Lett.* **19**:4 (2012), 785–804. MR 3008415 Zbl 06165853

[Kim 2000] H. H. Kim, "Langlands–Shahidi method and poles of automorphic L-functions, II", *Israel J. Math.* **117** (2000), 261–284. MR 2001i:11059a Zbl 1041.11035

[Kim and Krishnamurthy 2004] H. H. Kim and M. Krishnamurthy, "Base change lift for odd unitary groups", pp. 116–125 in *Functional analysis VIII* (Dubrovnik, 2003), edited by D. Bakić et al., Various Publ. Ser. (Aarhus) **47**, Aarhus University, Aarhus, 2004. MR 2006a:11154 Zbl 1146.11313

[Kim and Krishnamurthy 2005] H. H. Kim and M. Krishnamurthy, "Stable base change lift from unitary groups to GL_n", *Int. Math. Res. Pap.* **2005**:1 (2005), 1–52. MR 2006d:22028 Zbl 1146.11314

[Langlands 1971] R. P. Langlands, *Euler products* (New Haven, CT, 1967), Yale Mathematical Monographs **1**, Yale University Press, New Haven, CT, 1971. MR 54 #7387 Zbl 0231.20016

[Langlands 1989] R. P. Langlands, "On the classification of irreducible representations of real algebraic groups", pp. 101–170 in *Representation theory and harmonic analysis on semisimple Lie groups*, edited by P. J. Sally, Jr. and D. A. Vogan, Jr., Math. Surveys Monogr. **31**, Amer. Math. Soc., Providence, RI, 1989. MR 91e:22017 Zbl 0741.22009

[Lapid et al. 2004] E. Lapid, G. Muić, and M. Tadić, "On the generic unitary dual of quasisplit classical groups", *Int. Math. Res. Not.* **2004**:26 (2004), 1335–1354. MR 2005b:22021 Zbl 1079.22015

[Luo et al. 1999] W. Luo, Z. Rudnick, and P. Sarnak, "On the generalized Ramanujan conjecture for $\mathrm{GL}(n)$", pp. 301–310 in *Automorphic forms, automorphic representations, and arithmetic* (Fort

Worth, TX, 1996), edited by R. S. Doran et al., Proc. Sympos. Pure Math. **66**, Amer. Math. Soc., Providence, RI, 1999. MR 2000e:11072 Zbl 0965.11023

[Mœglin 2008] C. Mœglin, "Formes automorphes de carré intégrable non cuspidales", *Manuscripta Math.* **127**:4 (2008), 411–467. MR 2010i:11071 Zbl 05520091

[Mœglin and Waldspurger 1989] C. Mœglin and J.-L. Waldspurger, "Le spectre résiduel de $\mathrm{GL}(n)$", *Ann. Sci. École Norm. Sup.* (4) **22**:4 (1989), 605–674. MR 91b:22028 Zbl 0696.10023

[Mœglin and Waldspurger 1995] C. Mœglin and J.-L. Waldspurger, *Spectral decomposition and Eisenstein series: a paraphrase of the Scriptures*, Cambridge Tracts in Mathematics **113**, Cambridge University Press, 1995. MR 97d:11083 Zbl 0846.11032

[Mok 2015] C. P. Mok, "Endoscopic classification of representations of quasi-split unitary groups", American Mathematical Society, Providence, RI, 2015. Mem. Amer. Math. Soc.

[Muić 2001] G. Muić, "A proof of Casselman–Shahidi's conjecture for quasi-split classical groups", *Canad. Math. Bull.* **44**:3 (2001), 298–312. MR 2002f:22015 Zbl 0984.22007

[Rogawski 1990] J. D. Rogawski, *Automorphic representations of unitary groups in three variables*, Annals of Mathematics Studies **123**, Princeton University Press, 1990. MR 91k:22037 Zbl 0724.11031

[Shahidi 1981] F. Shahidi, "On certain L-functions", *Amer. J. Math.* **103**:2 (1981), 297–355. MR 82i: 10030 Zbl 0467.12013

[Shahidi 1985] F. Shahidi, "Local coefficients as Artin factors for real groups", *Duke Math. J.* **52**:4 (1985), 973–1007. MR 87m:11049 Zbl 0674.10027

[Shahidi 1988] F. Shahidi, "On the Ramanujan conjecture and finiteness of poles for certain L-functions", *Ann. of Math.* (2) **127**:3 (1988), 547–584. MR 89h:11021 Zbl 0654.10029

[Shahidi 1990] F. Shahidi, "A proof of Langlands' conjecture on Plancherel measures; complementary series for p-adic groups", *Ann. of Math.* (2) **132**:2 (1990), 273–330. MR 91m:11095 Zbl 0780.22005

[Shahidi 1992] F. Shahidi, "Twisted endoscopy and reducibility of induced representations for p-adic groups", *Duke Math. J.* **66**:1 (1992), 1–41. MR 93b:22034 Zbl 0785.22022

[Shahidi 2010] F. Shahidi, *Eisenstein series and automorphic L-functions*, Amer. Math. Soc. Colloq. Publ. **58**, Amer. Math. Soc., Providence, RI, 2010. MR 2012d:11119 Zbl 1215.11054

[Speh 1981] B. Speh, "The unitary dual of $\mathrm{Gl}(3, \mathbf{R})$ and $\mathrm{Gl}(4, \mathbf{R})$", *Math. Ann.* **258**:2 (1981), 113–133. MR 83i:22025 Zbl 0483.22005

[Tadić 1986] M. Tadić, "Classification of unitary representations in irreducible representations of general linear group (non-Archimedean case)", *Ann. Sci. École Norm. Sup.* (4) **19**:3 (1986), 335–382. MR 88b:22021 Zbl 0614.22005

[Takeda 2013] S. Takeda, "On a certain metaplectic Eisenstein series and the twisted symmetric square l-function", preprint, University of Missouri, Columbia, MO, 2013, available at http://www.math.missouri.edu/~takedas/analytic_Eisen.pdf.

[Takeda 2014a] S. Takeda, "Metaplectic tensor products for automorphic representations of $\tilde{GL}(r)$", preprint, 2014. arXiv 1303.2785

[Takeda 2014b] S. Takeda, "The twisted symmetric square L-function of GL(r)", *Duke Math. J.* **163**:1 (2014), 175–266. MR 3161314 Zbl 06279927

[Vogan 1978] D. A. Vogan, Jr., "Gel'fand–Kirillov dimension for Harish-Chandra modules", *Invent. Math.* **48**:1 (1978), 75–98. MR 58 #22205 Zbl 0389.17002

[Vogan 1986] D. A. Vogan, Jr., "The unitary dual of $\mathrm{GL}(n)$ over an Archimedean field", *Invent. Math.* **83**:3 (1986), 449–505. MR 87i:22042 Zbl 0598.22008

[Waldspurger 2014a] J.-L. Waldspurger, "Stabilisation de la formule des traces tordue, I: endoscopie tordue sur un corps local", preprint, Institut de Mathématiques de Jussieu, Paris, 2014, available at http://webusers.imj-prg.fr/~jean-loup.waldspurger/stabilisationIfinal.pdf.

[Waldspurger 2014b] J.-L. Waldspurger, "Stabilisation de la formule des traces tordue, II: intégrales orbitales et endoscopie sur un corps local non-archimédien; définitions et énoncés des résultats", preprint, Paris, 2014, available at http://webusers.imj-prg.fr/~jean-loup.waldspurger/stabilisationIIfinal.pdf.

[Waldspurger 2014c] J.-L. Waldspurger, "Stabilisation de la formule des traces tordue, III: intégrales orbitales et endoscopie sur un corps local non-archimédien; réductions et preuves", preprint, Paris, 2014, available at http://webusers.imj-prg.fr/~jean-loup.waldspurger/stabilisationIIIfinal.pdf.

[Wallach 1979] N. R. Wallach, "Representations of reductive Lie groups", pp. 71–86 in *Automorphic forms, representations and L-functions, 1* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, RI, 1979. MR 80m:22024 Zbl 0421.22006

[Zelevinsky 1980] A. V. Zelevinsky, "Induced representations of reductive $\mathfrak{p}$-adic groups, II: On irreducible representations of GL(n)", *Ann. Sci. École Norm. Sup.* (4) **13**:2 (1980), 165–210. MR 83g:22012 Zbl 0441.22014

[Zhang 1997] Y. Zhang, "The holomorphy and nonvanishing of normalized local intertwining operators", *Pacific J. Math.* **180**:2 (1997), 385–398. MR 98k:22076 Zbl 1073.22502

Received February 17, 2014. Revised April 17, 2014.

NEVEN GRBAC
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF RIJEKA
RADMILE MATEJCIC 2
HR-51000 RIJEKA
CROATIA

neven.grbac@math.uniri.hr

FREYDOON SHAHIDI
DEPARTMENT OF MATHEMATICS
PURDUE UNIVERSITY
150 N UNIVERSITY STREET
WEST LAFAYETTE, IN 47907-2067
UNITED STATES

shahidi@math.purdue.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.213

QUASICONFORMAL HARMONIC MAPPINGS BETWEEN DINI-SMOOTH JORDAN DOMAINS

DAVID KALAJ

Let D and Ω be Jordan domains with Dini-smooth boundaries. We prove that if $f : D \to \Omega$ is a harmonic homeomorphism and f is quasiconformal, then f is Lipschitz. This extends some recent results, where stronger assumptions on the boundary are imposed. Our result is optimal in that it coincides with the best condition for Lipschitz behavior of conformal mappings in the plane and conformal parametrizations of minimal surfaces.

1. Introduction and statement of the main result

Quasiconformal mappings. By definition, K-quasiconformal mappings (or qc mappings for short) are orientation-preserving homeomorphisms $f : D \to \Omega$ between domains $D, \Omega \subset \mathbb{C}$ that are contained in the Sobolev class $W^{1,2}_{\text{loc}}(D)$ and for which the differential matrix and its determinant are coupled in the distortion inequality

$$|Df(z)|^2 \le K \det Df(z)\,, \quad \text{where } |Df(z)| = \max_{|\xi|=1} |Df(z)\xi|, \tag{1-1}$$

for some $K \ge 1$. Here $\det Df(z)$ is the determinant of the formal derivative $Df(z)$, which will be denoted in the sequel by $J_f(z)$. Note that condition (1-1) can be written in complex notation as

$$(|f_z| + |f_{\bar z}|)^2 \le K(|f_z|^2 - |f_{\bar z}^2|) \quad \text{a.e. on } D, \tag{1-2}$$

or, what is the same,

$$|f_{\bar z}| \le k|f_z| \quad \text{a.e. on } D, \qquad \text{where } k = \frac{K-1}{K+1}, \text{ i.e., } K = \frac{1+k}{1-k}.$$

Harmonic mappings and the Hilbert transform. A mapping f is called *harmonic* in a region D if it has the form $f = u + iv$, where u and v are real-valued harmonic functions in D. If D is simply connected, then there are two analytic functions h and g defined on D such that f has the representation

$$f = h + \bar g.$$

MSC2010: primary 30C55; secondary 31C05.
Keywords: planar harmonic mappings, quasiconformal, Dini-smooth.

If f is a harmonic univalent function, then by Lewy's theorem [1936], f has a nonvanishing Jacobian and therefore is a diffeomorphism by the inverse mapping theorem.

Let

$$P(r, x-\varphi) = \frac{1-r^2}{2\pi(1-2r\cos(x-\varphi)+r^2)}$$

denote the Poisson kernel. If $F \in L^1(\mathbb{T})$, where $\mathbb{T}$ is the unit circle, we define the Poisson integral $\mathcal{P}[F]$ of F by

$$\mathcal{P}[F](z) = \int_0^{2\pi} P(r, x-\varphi)F(e^{ix})\,dx, \quad |z|<1,\ z = re^{i\varphi}. \tag{1-3}$$

The function $f(z) = \mathcal{P}[F](z)$ is a harmonic mapping in the unit disk $\mathbb{U} = \{z : |z| < 1\}$, which belongs to the Hardy space $h^1(\mathbb{U})$. The mapping f is bounded in $\mathbb{U}$ if and only if $F \in L^\infty(\mathbb{T})$. Standard properties of the Poisson integral show that $\mathcal{P}[F]$ extends by continuity to F on $\overline{\mathbb{U}}$, provided that F is continuous. For these facts and standard properties of harmonic Hardy spaces, we refer to [Axler et al. 1992, Chapter 6; Duren 1970]. With the additional assumption that F is an orientation-preserving homeomorphism of this circle onto a convex Jordan curve γ, $\mathcal{P}[F]$ is an orientation-preserving diffeomorphism of the open unit disk onto the region bounded by γ. This is indeed the celebrated theorem of Choquet–Radó–Kneser [Choquet 1945; Duren 2004]. This theorem is not true for nonconvex domains, but does hold under some additional assumptions. It has been extended in various directions (see for example [Jost 1981; Kalaj 2011b; Duren and Hengartner 1997]).

If $f = u + iv$ is a harmonic function defined in a Dini-smooth Jordan domain D then a harmonic function $\tilde{f} = \tilde{u} + i\tilde{v}$ is called the harmonic conjugate of f if $u + i\tilde{u}$ and $v + i\tilde{v}$ are analytic functions. Notice that $\tilde{f}$ is uniquely determined up to an additive constant. Let $\Phi : D \to \mathbb{U}$ be a conformal mapping, and let $G \in L^1(\partial D)$. Then the Poisson integral of G with respect to the domain D is defined by

$$\mathcal{P}_D[G](z) = \frac{1}{2\pi}\int_{\partial D} \frac{1-|\Phi(z)|^2}{|\Phi(z)-\Phi(\zeta)|^2} G(\zeta)|\Phi'(\zeta)|\,d\zeta.$$

Let χ be the boundary value of f and assume that $\tilde{\chi}$ is the boundary value of $\tilde{f}$. Then $\tilde{\chi}$ is called the Hilbert transform of χ and we also write it as $H(\chi)$. Assume that $\tilde{\chi} \in L^1(\partial D)$. In particular, the Hilbert transform of a function $\chi \in L^1(\mathbb{T})$ is defined by the formula

$$\tilde{\chi}(\tau) = H(\chi)(\tau) = -\frac{1}{\pi}\int_{0^+}^{\pi} \frac{\chi(\tau+t)-\chi(\tau-t)}{2\tan(t/2)}\,dt. \tag{1-4}$$

Here $\int_{0^+}^{\pi} \Phi(t)\,dt := \lim_{\epsilon\to 0^+}\int_{\epsilon}^{\pi} \Phi(t)\,dt$. This integral is improper and converges for a.e. $\tau \in [0, 2\pi]$. This and other facts concerning the operator H used in this paper can be found in [Zygmund 1959, Chapter VII]. Assume that χ, $\tilde{\chi}$ are in $L^1(\mathbb{T})$. Then

$$\mathcal{P}[\tilde{\chi}] = (\mathcal{P}[\chi])^{\sim}, \tag{1-5}$$

where $(k)^{\sim}$ is the harmonic conjugate of k (see for instance [Pavlović 2004, Theorem 6.1.3]).

If $f = h + \bar{g} : \mathbb{U} \to \Omega$ is a harmonic mapping then the radial and tangential derivatives at $z = re^{it}$ are defined by

$$\partial_r f(z) = \frac{1}{r}(h' + \bar{g}') \quad \text{and} \quad \partial_t f(z) = i(h' - \bar{g}').$$

So $r\partial_r f$ is the harmonic conjugate of $\partial_t f$. We generalize this definition for a mapping $f = h + \bar{g}$ defined in a Jordan domain D. In order to do so, let $\Phi = Re^{i\Theta}$ be a conformal mapping of the domain D onto the unit disk. Then the radial derivative and tangent derivative of f in a point $w \in D$ are defined by

$$\partial_R f(w) = \frac{1}{|\Phi(w)|} Df(w)\left(\frac{\Phi(w)}{\Phi'(w)}\right) \quad \text{and} \quad \partial_\Theta f(w) = Df(w)\left(i\frac{\Phi(w)}{\Phi'(w)}\right).$$

Here $\Phi(w)/\Phi'(w)$ and $i(\Phi(w)/\Phi'(w))$ are treated as two vectors from $\mathbb{R}^2 \cong \mathbb{C}$. Then it is easy to show that

$$R\partial_R f(w) = \frac{h'(w)}{\Phi'(w)} + \frac{\overline{g'(w)}}{\overline{\Phi'(w)}} \quad \text{and} \quad \partial_\Theta f(w) = i\left(\frac{h'(w)}{\Phi'(w)} - \frac{\overline{g'(w)}}{\overline{\Phi'(w)}}\right).$$

This implies that $R\partial_R f(w)$ and $\partial_\Theta f(w)$ are harmonic functions in D and $R\partial_R f(w)$ is the harmonic conjugate of $\partial_\Theta f(w)$. Notice also that these derivatives are uniquely determined up to a conformal mapping Φ. Assume further that D and Ω have Dini-smooth boundaries. If $F : \partial D \to \partial\Omega$ is the boundary function of f, and if $\partial_\Theta f(w)$ is a bounded harmonic function, then

$$\lim_{w \to w_0} \partial_\Theta f(w) = F'(w_0),$$

where the limit is nontangential. Here

$$F'(w_0) := \frac{\partial(F \circ \Phi^{-1})(e^{it})}{\partial t},$$

where $\Phi(w_0) = e^{it}$. If $F' \in L^1(\partial D)$, then the harmonic function $R\partial_R f(w)$ has nontangential limits in almost every point of ∂D and its boundary value is the Hilbert transform of F', namely

$$H(F')(w_0) = \lim_{w \to w_0} R\partial_R f(w).$$

From now on the boundary value of f will be denoted by F. We will focus on orientation-preserving harmonic quasiconformal mappings between smooth domains and investigate their Lipschitz character up to the boundary. For future reference, we will say that a qc mapping $f : \mathbb{U} \to \Omega$ of the unit disk onto the Jordan domain Ω with rectifiable boundary is *normalized* if $f(1) = w_0$, $f(e^{2\pi i/3}) = w_1$ and $f(e^{4\pi i/3}) = w_2$, where w_0w_1, w_1w_2 and w_2w_0 are arcs of $\gamma = \partial\Omega$ having the same length $|\gamma|/3$.

Background. Let Ω be a Jordan domain with rectifiable boundary, and let γ be an arc-length parametrization of $\partial\Omega$. We say that $\partial\Omega$ is C^1 if $\gamma \in C^1$. Then $\arg\gamma'$ is continuous and we let ω be its modulus of continuity. If ω satisfies

$$\int_0^\delta \frac{\omega(t)}{t}\,dt < \infty, \qquad \delta > 0, \tag{1-6}$$

we say that $\partial\Omega$ is Dini-smooth. Denote by $C^{1,\varpi}$ the class of all Dini-smooth Jordan curves. The derivative of a conformal mapping f of the unit disk onto Ω is continuous and nonvanishing in $\overline{D}$ [Pommerenke 1975, Theorem 10.2] (see also [Warschawski 1961]). This implies that f is bi-Lipschitz continuous. For later reference we refer to this result as Kellogg's theorem, see [Kellogg 1912; Goluzin 1969, p. 374]. Kellogg was the first to consider this type of result for $C^{1,\alpha}$ domains, where $0 < \alpha < 1$. Warschawski [1970] proved the same result for a conformal parametrization of a minimal surface.

If f is merely quasiconformal and maps the unit disk onto itself, then Mori's theorem implies that $|f(z) - f(w)| \le M_1(K)|z - w|^{1/K}$. The constant $1/K$ is the best possible. If f is a conformal mapping of the unit disk onto a Jordan domain with a C^1 boundary, then the function f is not necessarily Lipschitz (see for example [Lesley and Warschawski 1978, p. 277]). This is why we need to add some assumption, other than quasiconformality, as well as some smoothness of the image curve that is better than C^1 in order to obtain that the resulting mapping is Lipschitz or bi-Lipschitz.

Since every conformal mapping in the plane is harmonic and quasiconformal, it is an interesting question to ask to what extent the smoothness of the boundary of a Jordan domain Ω implies that a quasiconformal harmonic mapping of the unit disk onto Ω is Lipschitz. The first study of harmonic quasiconformal mappings of the unit disk onto itself was done by O. Martio [1968]. This paper has been generalized in [Kalaj 2004] for qc mappings from the unit disk onto a convex Jordan domain. Pavlović [2002] proved in a very interesting way that every qc harmonic mapping of the unit disk onto itself is Lipschitz. Kalaj [2008] proved that every qc harmonic mapping between two Jordan domains with $C^{1,\alpha}$ boundary is Lipschitz. This result has its counterpart for non-Euclidean metrics [Kalaj and Mateljević 2006]. For a generalization of the last result to the several-dimensional case we refer to [Kalaj 2013]. The problem of bi-Lipschitz continuity of a quasiconformal mapping of the unit disk onto a Jordan domain with C^2 boundary has been solved in [Kalaj 2011a]. The object of this paper is to extend some of these results.

New results. The following theorem is such an extension in which the Hölder continuity is replaced by the more general Dini condition.

Theorem 1.1. *Let $f = \mathcal{P}[F](z)$ be a harmonic normalized K-quasiconformal mapping between the unit disk and the Jordan domain Ω with $\gamma = \partial\Omega \in C^{1,\varpi}$. Then there exists a constant $C' = C'(\gamma, K)$ such that*

$$\Big|\frac{\partial F(e^{i\varphi})}{\partial\varphi}\Big| \le C' \quad \text{for almost every } \varphi \in [0, 2\pi], \tag{1-7}$$

and

$$|f(z_1) - f(z_2)| \le KC'|z_1 - z_2| \quad \text{for } z_1, z_2 \in \mathbb{U}. \tag{1-8}$$

By using Theorem 1.1, we obtain the following improvement of [Kalaj 2008, Theorem 3.1].

Theorem 1.2. *Let D and Ω be Jordan domains such that ∂D and $\partial\Omega$ are contained in $C^{1,\varpi}$ and let $f : D \mapsto \Omega$ be a harmonic homeomorphism. The following statements hold true.*

(a) *If f is qc, then f is Lipschitz.*

(b) *If Ω is convex and f is qc, then f is bi-Lipschitz.*

(c) *If Ω is convex, then f is qc if and only* $\log|F'|$ *and $H(F')$ are in $L^\infty(\partial D)$.*

Proof of Theorem 1.2. (a) Choose a conformal mapping $\Phi : \mathbb{U} \to D$ so that the qc mapping $f_1 = f \circ \Phi$ is normalized. Then f_1 is a qc harmonic mapping of the unit disk onto Ω that satisfies the conditions of Theorem 1.1. This implies in particular that f_1 is Lipschitz. In view of Kellogg's theorem, the mapping Φ is bi-Lipschitz. Thus $f = f_1 \circ \Phi^{-1}$ is Lipschitz.

(b) If Ω is a convex domain, and if $D = \mathbb{U}$, then by [Kalaj 2003], we have that

$$|Df(z)| \ge \tfrac{1}{4}\,\mathrm{dist}(f(0), \partial\Omega)$$

for $z \in \mathbb{U}$. If D is not the unit disk, then we make use of the conformal mapping $\Phi : \mathbb{U} \to D$ as in the proof of (a). Then we obtain

$$|Df(z)| = |Df_1(z)|/|\Phi'(z)| \ge c.$$

Now by using the quasiconformality of f, we have that

$$|Df(z)|^2 \le K J_f(z).$$

Therefore

$$J_{f^{-1}(f(z))} = \frac{1}{J_f(z)} \le \frac{K}{c^2}.$$

Since f^{-1} is K-quasiconformal, we have further that

$$|Df^{-1}(w)|^2 \le K J_{f^{-1}}(w) \le \frac{K^2}{c^2}.$$

This implies that f^{-1} is Lipschitz. This finishes the proof of (b).

(c) If f is harmonic and quasiconformal, then by (b) it is bi-Lipschitz, and so its boundary function F is bi-Lipschitz. Furthermore, $R\partial_R f$ is a bounded harmonic function and this is equivalent with the fact that $\log|F'| \in L^\infty(\partial D)$. Since $H(F')$ is its boundary function, it is bounded, i.e., it belongs to $L^\infty(\partial D)$.

We now prove the opposite implication. Since

$$\partial_\Theta f = \mathcal{P}_D[F'] \quad \text{and} \quad R\partial_R f = \mathcal{P}_D[H(F')],$$

it follows that $\partial_\Theta f$ and $R\partial_R f$ are bounded harmonic functions. This means that $|Df|$ is bounded by a constant M. In order to show that f is quasiconformal, it is enough to show that the Jacobian of f is bigger than a positive constant in D. Let $f_1 = f \circ \Phi^{-1}$, and let $\delta = \operatorname{dist}(f_1(0), \partial\Omega)$ and $\kappa = \min|\partial_t f_1(e^{it})|$. Then by [Kalaj 2004, Corollary 2.9], we have

$$J_f(\Phi(w))|\Phi'(w)|^2 = J_{f_1}(w) \ge \frac{\kappa\delta}{2}.$$

So

$$J_f(z) \ge c > 0, \quad z \in D.$$

We conclude that

$$\frac{|Df(z)|^2}{J_f(z)} \le \frac{M^2}{c}. \qquad \square$$

2. Preliminary results

Definition 2.1. Let $\xi : [a, b] \to \mathbb{C}$ be a continuous function. The modulus of continuity of ξ is

$$\omega(t) = \omega_\xi(t) = \sup_{|x-y|\le t} |\xi(x) - \xi(y)|.$$

The function ξ is called Dini-continuous if

$$\int_0^{b-a} \frac{\omega_\xi(t)}{t}\, dt < \infty. \tag{2-1}$$

Let γ be a C^1 Jordan curve γ with the length $l = |\gamma|$ and assume that $g : [0, l] \to \gamma$ is its arc-length parametrization . We say that γ is Dini-smooth if g' is Dini-continuous on $[0,l]$. If $\omega(t)$ is the modulus of continuity of g' for $0 \le t \le l$, then we extend ω by $\omega(t) = \omega(l)$ for $t \ge l$.

A function $F : \mathbb{T} \to \gamma$ is called Dini-smooth if the function $\Phi(t) = F(e^{it})$ is Dini-smooth, i.e.,

$$|\Phi'(t) - \Phi'(s)| \le \omega(|t-s|),$$

where ω is Dini-continuous. Observe that every smooth $C^{1,\alpha}$ Jordan curve is Dini-smooth.

Let

$$\mathcal{K}(s,t) = \operatorname{Re}\,[\overline{(g(t)-g(s))} \cdot i g'(s)] \tag{2-2}$$

be a function defined on $[0,l] \times [0,l]$. By $\mathcal{K}(s \pm l, t \pm l) = \mathcal{K}(s,t)$ we extend it to $\mathbb{R} \times \mathbb{R}$. Suppose now that $\Psi : \mathbb{R} \mapsto \gamma$ is an arbitrary 2π-periodic Lipschitz function such that $\Psi|_{[0,2\pi)} : [0,2\pi) \mapsto \gamma$ is an orientation-preserving bijective function. Then there exists an increasing continuous function $\psi : [0,2\pi] \mapsto [0,l]$ such that

$$\Psi(\tau) = g(\psi(\tau)). \tag{2-3}$$

We have for a.e. $e^{i\tau} \in \mathbb{T}$ that

$$\Psi'(\tau) = g'(\psi(\tau)) \cdot \psi'(\tau),$$

and therefore

$$|\Psi'(\tau)| = |g'(\psi(\tau))| \cdot |\psi'(\tau)| = \psi'(\tau).$$

Along with the function $\mathcal{K}$ we will also consider the function $\mathcal{K}_F$ defined by

$$\mathcal{K}_F(t,\tau) = \operatorname{Re}\,[\overline{(\Psi(t)-\Psi(\tau))} \cdot i\Psi'(\tau)].$$

Here $F(e^{it}) = \Psi(t)$. It is easy to see that

$$\mathcal{K}_F(t,\tau) = \psi'(\tau)\mathcal{K}(\psi(t),\psi(\tau)). \tag{2-4}$$

Lemma 2.2. *Let γ be a Dini-smooth Jordan curve and let $g : [0,l] \mapsto \gamma$ be a natural parametrization of a Jordan curve with g' having modulus of continuity ω. Assume further that $\Psi : [0,2\pi] \mapsto \gamma$ is an arbitrary parametrization of γ and let $F(e^{it}) = \Psi(t)$. Then*

$$|\mathcal{K}(s,t)| \le \int_0^{\min\{|s-t|,\,l-|s-t|\}} \omega(\tau)\,d\tau \tag{2-5}$$

and

$$|\mathcal{K}_F(\varphi,x)| \le |\psi'(\varphi)| \int_0^{d_\gamma(\Psi(\varphi),\Psi(x))} \omega(\tau)\,d\tau. \tag{2-6}$$

Here $d_\gamma(\Psi(\varphi),\Psi(x)) := \min\{|s(\varphi)-s(x)|, (l-|s(\varphi)-s(x)|)\}$ is the (shortest) *distance between* $\Psi(\varphi)$ *and* $\Psi(x)$ *along* γ, *and it satisfies*

$$|\Psi(\varphi)-\Psi(x)| \le d_\gamma(\Psi(\varphi), \Psi(x)) \le B_\gamma|\Psi(\varphi)-\Psi(x)|.$$

Proof. Note that the estimate (2-5) has been proved in [Kalaj 2011b, Lemma 2.3]. Now (2-6) follows from (2-5) and (2-4). □

A closed rectifiable Jordan curve γ satisfies a B-chord-arc condition for some constant $B > 1$ if for all $z_1, z_2 \in \gamma$ we have

$$d_\gamma(z_1, z_2) \le B|z_1 - z_2|. \tag{2-7}$$

Here $d_\gamma(z_1, z_2)$ is the length of the shorter arc of γ with endpoints z_1 and z_2. It is clear that if $\gamma \in C^1$, then γ satisfies a chord-arc condition for some $B_\gamma > 1$. The following lemma is proved in [Kalaj 2012].

Lemma 2.3. *Assume that* γ *satisfies a chord-arc condition for some* $B > 1$. *Then for every normalized K-qc mapping* f *between the unit disk* $\mathbb{U}$ *and the Jordan domain* $\Omega = \operatorname{int}\gamma$ *we have*

$$|f(z_1) - f(z_2)| \le \Lambda_\gamma(K)|z_1 - z_2|^\alpha, \quad z_1, z_2 \in \mathbb{T},$$

where

$$\alpha = \frac{2}{K(1+2B)^2}, \quad \Lambda_\gamma(K) = 4\cdot 2^\alpha(1+2B)\sqrt{\frac{2\pi K|\Omega|}{\log 2}}.$$

Next we recall some estimates for the Jacobian of a harmonic univalent function.

Lemma 2.4 [Kalaj 2011b, Lemma 3.1]. *Suppose* $f = \mathcal{P}[F]$ *is a harmonic mapping such that* F *is a Lipschitz homeomorphism from the unit circle onto a Dini-smooth Jordan curve* γ. *Let* g *be an arc-length parametrization of* γ, *let* $\psi(t) = g^{-1}(F(e^{it}))$, *and define* $\Psi(t) = F(e^{it}) = g(\psi(t))$. *Then for almost every* $\tau \in [0, 2\pi]$, *the limit*

$$J_f(e^{i\tau}) := \lim_{r\to 1} J_f(re^{i\tau})$$

exists and we have

$$J_f(e^{i\tau}) = \psi'(\tau)\int_0^{2\pi} \frac{\operatorname{Re}\big[\overline{(g(\psi(t)) - g(\psi(\tau)))}\cdot ig'(\psi(\tau))\big]}{2\sin^2((t-\tau)/2)}\frac{dt}{2\pi}. \tag{2-8}$$

From Lemma 2.2 and Lemma 2.4 we obtain

Lemma 2.5. *Under the conditions and notation of Lemma 2.4 we have*

$$J_f(e^{i\varphi}) \le \frac{\pi}{4}|\Psi'(\varphi)|\int_{-\pi}^{\pi}\frac{1}{x^2}\int_0^{d_\gamma(F(e^{i(\varphi+x)}), F(e^{i\varphi}))}\omega(\tau)\,d\tau\,dx \tag{2-9}$$

for a.e. $e^{i\varphi} \in \mathbb{T}$. *Here* ω *is the modulus of continuity of* g'.

Lemma 2.6. *Let $f = \mathcal{P}[F](z)$ be a harmonic mapping between the unit disk $\mathbb{U}$ and the Jordan domain Ω, with $F \in C^{1,\varpi}(\mathbb{T})$. Then the partial derivatives of f have a continuous extension to the boundary of the unit disk.*

Proof. In the proof of this lemma we denote $\partial_t \Psi(e^{it})$ by $\Psi'(t)$. If F is Lipschitz-continuous, then $\Phi = \Psi' \in L^\infty(\mathbb{T})$, and by the famous Marcel Riesz theorem (see for example [Garnett 1981, Theorem 2.3]) there is a constant A_p such that

$$\|H(\Psi')\|_{L^p(\mathbb{T})} \le A_p \|\Psi'\|_{L^p(\mathbb{T})}$$

for $1 < p < \infty$. It follows that $\tilde{\Phi} = H(\Psi') \in L^1$. Since rf_r is the harmonic conjugate of f_τ, we have $rw_r = \mathcal{P}[H(\Psi')]$ according to (1-5). By again using Fatou's theorem, we have

$$\lim_{r\to 1^-} f_r(re^{i\tau}) = H(\Psi')(\tau) \quad \text{a.e.} \tag{2-10}$$

By (1-4), and by following the proof of Privaloff's theorem [Zygmund 1959], we obtain that if $|\Psi'(x) - \Psi'(y)| \le \omega(|x-y|)$ for the Dini-continuous function, then

$$|H(\Psi')(x+h) - H(\Psi')(x)| \le A\int_0^{2h} \frac{\omega(t)}{t}\,dt + Bh\int_h^{2\pi} \frac{\omega(t)}{t^2}\,dt + C\omega(h),$$

for some absolute constants A, B and C. The detailed proof of the last fact can be found in [Garnett 1981, Theorem III 1.3.]. This implies that $rw_r(re^{it})$ and $f_t(re^{it})$ have continuous extensions to the boundary and this is what we needed to prove. □

We now prove the following lemma needed in the sequel.

Lemma 2.7. *Let A be a positive integrable function in $[0, B]$ and assume that $q, Q > 0$. Then there exists a continuous increasing function χ of $(0, +\infty)$ into itself, depending on A, B, q and Q, such that the following hold: $\lim_{x\to\infty} \chi(x) = \infty$, the function $g(x) = x\chi(x)$ is convex, and*

$$\int_0^B A(x)\chi(Qx^{-q})\,dx \le 4\int_0^B A(x)dx.$$

Proof. First define inductively a sequence $x_0 = B$, $x_k > 0$ for $k > 0$, such that $x_{k+1} < x_k/2$, and

$$\int_0^{x_k} A(x)\,dx \le M2^{-k} \quad \text{where } M = \int_0^B A(x)dx.$$

This is possible because A is integrable.

Then define a continuous function ξ in $[0, B]$ by $\xi(x_k) = k$, and by extending it linearly on each interval $[x_{k+1}, x_k]$, that is

$$\xi(x) = k + \frac{x_k - x}{x_k - x_{k+1}}, \quad x \in [x_{k+1}, x_k].$$

It is easy to see that this function is convex, decreasing and tends to $+\infty$ as $x \to \infty$. Moreover

$$\int_0^B A(x)\xi(x)\,dx \le M \sum_{k=0}^{\infty}(k+1)2^{-k} = 4M.$$

Now set $\chi(x) = \xi((Q/x)^\tau)$ for $\tau = 1/q$. It remains to verify that $x\chi(x)$ is convex. This we do by differentiation:

$$(x\chi(x))' = \xi(Q^\tau x^{-\tau}) - Q^\tau \tau x^{-\tau}\xi'(Q^\tau x^{-\tau}).$$

Since both summands are increasing, $x\chi(x)$ is convex. □

3. The proof of Theorem 1.1

By assumption of the theorem, the derivative of an arc-length parametrization g' has a Dini-continuous modulus of continuity ω. We consider two cases.

(i) $F(e^{it}) = \Psi(t) \in C^{1,\varpi}(\mathbb{T})$. Then by Lemma 2.6 the mapping $f(z) = \mathcal{P}[F](z)$ is C^1 up to the boundary. First we notice that for $L = \sup|\Psi'(t)|$, it is clear that $L < \infty$. We will prove more. We will show that L is bounded by a constant not depending a priori on F. According to Lemma 2.6 and to (1-1), we have

$$\begin{aligned}
|Df(e^{i\varphi})|^2 &= (|f_z(e^{i\varphi})| + |f_{\bar z}(e^{i\varphi})|)^2\\
&= \lim_{z\to e^{i\varphi}} (|f_z(z)| + |f_{\bar z}(z)|)^2\\
&\le K \lim_{z\to e^{i\varphi}} (|f_z(z)|^2 - |f_{\bar z}(z)|^2)\\
&= K(|f_z(e^{i\varphi})|^2 - |f_{\bar z}(e^{i\varphi})|^2) = K J_f(e^{i\varphi}).
\end{aligned} \tag{3-1}$$

Furthermore, we have

$$|Df(re^{i\varphi})| = \sup_{|\xi|=1} |Df(re^{i\varphi})\xi| \ge |Df(re^{i\varphi})(ie^{i\varphi})| = |\partial_\varphi f(re^{i\varphi})|. \tag{3-2}$$

This implies that

$$|Df(e^{i\varphi})|^2 \ge |\partial_\varphi f(e^{i\varphi})|^2 = |\Psi'(\varphi)|^2. \tag{3-3}$$

From (2-9), (3-3) and (3-1), we obtain:

$$|\Psi'(\varphi)|^2 \le KC_1|\Psi'(\varphi)| \int_{-\pi}^{\pi} \frac{1}{x^2}\int_0^{\rho(x,\varphi)} \omega(\tau)\,d\tau\,dx,$$

where

$$\rho(x,\varphi) = d_\gamma\big(F(e^{i(\varphi+x)}),\, F(e^{i\varphi})\big),$$

which is the same as

$$|\Psi'(\varphi)| \le KC_1 \int_{-\pi}^{\pi} \frac{\rho(\varphi, x)}{x^2} \int_0^1 \omega(\tau\rho(\varphi, x))\, d\tau\, dx.$$

Thus

$$|\Psi'(\varphi)| \le KC_1 \int_{-\pi}^{\pi} \frac{\rho(\varphi, x)}{x^2} \omega(\rho(\varphi, x))\, dx.$$

Let

$$L := \max_{x\in[0,2\pi]} |\Psi'(x)| = \max_{x\in[0,2\pi]} \psi'(x) = \psi'(\varphi). \tag{3-4}$$

Then

$$L \le KC_1 \int_{-\pi}^{\pi} \frac{\rho(\varphi, x)}{x^2} \omega(\rho(\varphi, x))\, dx.$$

Furthermore, we have

$$M := \frac{L}{2\pi KC_1} \le \int_{-\pi}^{\pi} M(x, \varphi) \frac{dx}{2\pi},$$

where

$$M(x, \varphi) = \frac{\rho(\varphi, x)}{x^2} \omega(\rho(\varphi, x)).$$

The idea is to make use of Lemma 2.7 with a convex function depending only on K to be found below.

Assume that $\chi : \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous increasing function to be determined in the sequel such that the function $\Phi(t) = t\chi(t)$ is convex. By using Jensen's inequality to the previous integral with respect to the convex function Φ, we obtain

$$\Phi(M) \le \int_{-\pi}^{\pi} \Phi(M(x, \varphi))\, \frac{dx}{2\pi},$$

or equivalently,

$$M\chi(M) \le \int_{-\pi}^{\pi} M(x, \varphi)\chi(M(x, \varphi))\, \frac{dx}{2\pi}. \tag{3-5}$$

From (2-7) and (3-4) we deduce that

$$\rho(\varphi, x) \le B_\gamma L|x|. \tag{3-6}$$

On the other hand, since f is a normalized qc mapping, we have by Lemma 2.3 that

$$\rho(\varphi, x) \le B_\gamma \Lambda_\gamma(K)|x|^\alpha. \tag{3-7}$$

Notice that this time we used the boundary normalization. This implies that

$$M(x,\varphi)=\frac{\rho(\varphi,x)}{x^2}\omega(\rho(\varphi,x))\le\frac{B_\gamma L}{x}\omega(B_\gamma\Lambda_\gamma(K)|x|^\alpha), \tag{3-8}$$

and

$$M(x,\varphi)=\frac{\rho(\varphi,x)}{x^2}\omega(\rho(\varphi,x))\le\frac{B_\gamma\Lambda_\gamma(K)}{x^{2-\alpha}}\omega(B_\gamma\Lambda_\gamma(K)|x|^\alpha). \tag{3-9}$$

So, in view of Definition 2.1 we have

$$M(x,\varphi)\le\frac{B_\gamma\Lambda_\gamma(K)}{x^{2-\alpha}}\omega(|\gamma|). \tag{3-10}$$

From (3-5) and (3-8), we obtain

$$\begin{aligned}\chi\left(\frac{L}{2\pi KC_1}\right)&\le\int_{-\pi}^{\pi}\frac{KC_1B_\gamma}{x}\omega(B_\gamma\Lambda_\gamma(K)|x|^\alpha)\chi\left(\frac{B_\gamma\Lambda_\gamma(K)\omega(|\gamma|)}{|x|^{2-\alpha}}\right)dx\\&=2\int_{0}^{\pi}\frac{KC_1B_\gamma}{x}\omega(B_\gamma\Lambda_\gamma(K)|x|^\alpha)\chi\left(\frac{B_\gamma\Lambda_\gamma(K)\omega(|\gamma|)}{|x|^{2-\alpha}}\right)dx\\&=\frac{2KC_1B_\gamma}{B_\gamma\Lambda_\gamma(K)\alpha}\int_0^B\frac{\omega(y)}{y}\chi(Qy^{1-2/\alpha})\,dy,\end{aligned} \tag{3-11}$$

where

$$B=B_\gamma\Lambda_\gamma(K)\pi^\alpha\quad\text{and}\quad Q=\omega(|\gamma|)(B_\gamma\Lambda_\gamma(K))^{2-2/\alpha}.$$

In view of the last term of (3-11), now is the time to determine the function χ. Lemma 2.7 with $q=2/\alpha-1$ and $A(y)=\omega(y)/y$, provides us with a function χ such that Φ is convex and such that the estimate

$$\int_0^B\frac{\omega(y)}{y}\chi(Qy^{1-2/\alpha})dy\le4\int_0^B\frac{\omega(y)}{y}dy$$

holds. From (3-11), we have

$$\chi\left(\frac{L}{2\pi KC_1}\right)\le\frac{8KC_1B_\gamma}{B_\gamma\Lambda_\gamma(K)\alpha}\int_0^B\frac{\omega(y)}{y}dy=:\Upsilon(K,\Omega).$$

Since χ is increasing, we infer finally that

$$L\le2\pi KC_1\cdot\chi^{-1}(\Upsilon(K,\Omega))=\frac{\pi^2}{2}K\cdot\chi^{-1}(\Upsilon(K,\Omega)). \tag{3-12}$$

By the maximum principle, for $z=re^{i\varphi}$, we further have

$$|\partial_\varphi f(z)|\le L.$$

Since f is K-quasiconformal, we have

$$|Dw(z)|\le K|\partial_\varphi f(z)|.$$

This and the mean value inequality imply that

$$|f(z) - f(z')| \le KL|z - z'|, \quad |z| < 1, |z'| < 1. \tag{3-13}$$

(ii) $F \notin C^{1,\varpi}(\mathbb{T})$. In order to deal with nonsmooth F, we make use of an approximation argument. We begin by this definition.

Definition 3.1. Let G be a domain in $\mathbb{C}$ and let $a \in \partial G$. We will say that $G_a \subset G$ is a neighborhood of a if there exists a disk $D(a, r) := \{z : |z - a| < r\}$ such that $(D(a, r) \cap G) \subset G_a$.

Let $t = e^{ix} \in \mathbb{T}$. Then $F(t) = \Psi(x) \in \partial\Omega$. Let g be an arc-length parametrization of $\partial\Omega$ with $g(\psi(x)) = F(e^{ix})$, where $\psi : [0, 2\pi] \to [0, |\gamma|]$ is as in the first part of the proof. Put $s = \psi(x)$. Since the modulus of continuity of g' is a Dini-continuous function ω, there exists a neighborhood Ω_t of $\Psi(t)$ such that the derivative of its arc-length parametrization g_t' has modulus of continuity $C_t \cdot \omega$. Moreover, there exist positive numbers r_t and R_t such that

$$\Omega_t^\tau := \Omega_t + ig'(s) \cdot \tau \subset \Omega, \quad \tau \in (0, R_t), \tag{3-14}$$

$$\partial\Omega_t^\tau \subset \Omega, \quad \tau \in (0, R_t), \tag{3-15}$$

$$g[s - r_t, s + r_t] \subset \partial\Omega_t. \tag{3-16}$$

An example of a family Ω_t^τ such that $\partial\Omega_t^\tau \in C^{1,\alpha}$ for $0 < \alpha < 1$ with property (3-14) has been given in [Kalaj 2008]. The same construction yields the family $\partial\Omega_t^\tau$ with the above mentioned properties.

Take $U_\tau = f^{-1}(\Omega_t^\tau)$. Let η_t^τ be a conformal mapping of the unit disk onto U_τ with normalized boundary condition: $\eta_t^\tau(e^{i2k\pi/3}) = f^{-1}(\zeta_k)$ for $k = 0, 1, 2$, where $\zeta_0, \zeta_1, \zeta_2$ are three points of $\partial\Omega_t^\tau$ of equal distance. Then the mapping

$$f_t^\tau(z) := f(\eta_t^\tau(z)) - ig'(s) \cdot \tau$$

is a harmonic K-quasiconformal mapping of the unit disk onto Ω_t satisfying the boundary normalization. Moreover,

$$f_t^\tau = \mathcal{P}[F_t^\tau] \in C^1(\overline{\mathbb{U}})$$

for some function $F_t^\tau \in C^1(\mathbb{T})$.

Since $[0, l]$ is compact, there exists a finite family of Jordan arcs

$$\gamma_j = g(s_j - r_{s_j}/2, s_j + r_{s_j}/2), \quad j = 1, \dots, n,$$

covering γ. Assume that $F(t_j) = s_j$. Let

$$F_{j,\tau} := F_{t_j}^\tau, \quad a_{j,\tau} := \eta_{t_j}^\tau \quad \text{and} \quad f_{j,\tau} := f_{t_j}^\tau.$$

Using the case $F \in C^{1,\varpi}$, it follows that there exists a constant $C'_j = C'(K, \gamma_j)$ such that

$$|\partial_\varphi F'_{j,\tau}(e^{i\varphi})| \le C'_j$$

and

$$|f_{j,\tau}(z_1) - f_{j,\tau}(z_2)| \le K C'_j |z_1 - z_2|. \tag{3-17}$$

Since $a_{j,\tau}(z)$ converges uniformly on compact subsets of $\mathbb{U}$ to the function $a_{j,0}(z)$ when $\tau \to 0$, and since $f_{j,\tau} = f \circ a_{j,\tau}$, inequality (3-17) implies

$$|f_j(z_1) - f_j(z_2)| \le K C'_j |z_1 - z_2| \quad \text{for } z_1, z_2 \in \overline{\mathbb{U}}, \tag{3-18}$$

where $f_j = f \circ a_{j,0} = \mathcal{P}[F_j]$. For $z_1 = e^{it}$ and $z_2 = e^{i\varphi}$ for $t \to \varphi$, we obtain that $|\partial_\varphi F_j(e^{i\varphi})| \le K C'_j$ a.e. Since the mapping $b_j = a_{0,j}^{-1}$ can be extended conformally across the arc $S_j = f^{-1}(\lambda_j)$, where $\lambda_j = g(s_j - t_{s_j}, s_j + t_{s_j})$, there exists a constant L_j such that $|b_j(z)| \le L_j$ on $S'_j = \mathbb{T} \cap f^{-1}(\gamma_j)$ for $j = 1, \dots, n$. Hence $|\partial_\varphi F(e^{i\varphi})| \le K C'_j \cdot L_j$ on S'_j. Let $C' = \max\{K C'_j \cdot L_j : j = 1, \dots, n\}$. Inequalities (1-7) and (1-8) easily follow from $\mathbb{T} = \bigcup_{j=1}^n S'_j$.

Notice that we can now repeat the first part of the proof for a Lipschitz $f = \mathcal{P}[F]$ in order to obtain a more concrete Lipschitz constant, i.e., the constant L satisfying (3-12). The proof is complete.

Acknowledgement

I am grateful to professor Alexandre Eremenko, whose idea is used in the proof of Lemma 2.7, for his numerous useful suggestions and corrections that have substantially improved this paper. I am grateful to the anonymous referee for some useful suggestions and corrections.

References

[Axler et al. 1992] S. Axler, P. Bourdon, and W. Ramey, *Harmonic function theory*, Graduate Texts in Mathematics **137**, Springer, New York, 1992. MR 93f:31001 Zbl 0765.31001

[Choquet 1945] G. Choquet, "Sur un type de transformation analytique généralisant la représentation conforme et définie au moyen de fonctions harmoniques", *Bull. Sci. Math.* (2) **69** (1945), 156–165. MR 8,93a Zbl 0063.00851

[Duren 1970] P. L. Duren, *Theory of H^p spaces*, Pure and Applied Mathematics **38**, Academic Press, New York, 1970. MR 42 #3552 Zbl 0215.20203

[Duren 2004] P. Duren, *Harmonic mappings in the plane*, Cambridge Tracts in Mathematics **156**, Cambridge Univ. Press, 2004. MR 2005d:31001 Zbl 1055.31001

[Duren and Hengartner 1997] P. Duren and W. Hengartner, "Harmonic mappings of multiply connected domains", *Pacific J. Math.* **180**:2 (1997), 201–220. MR 99f:30051 Zbl 0885.30020

[Garnett 1981] J. B. Garnett, *Bounded analytic functions*, Pure and Applied Mathematics **96**, Academic Press, New York, 1981. MR 83g:30037 Zbl 0469.30024

[Goluzin 1969] G. M. Goluzin, *Geometric theory of functions of a complex variable*, Translations of Mathematical Monographs **26**, Amer. Math. Soc., Providence, RI, 1969. MR 40 #308 Zbl 0183.07502

[Jost 1981] J. Jost, "Univalency of harmonic mappings between surfaces", *J. Reine Angew. Math.* **324** (1981), 141–153. MR 82h:58013 Zbl 0453.53036

[Kalaj 2003] D. Kalaj, "On harmonic diffeomorphisms of the unit disc onto a convex domain", *Complex Var. Theory Appl.* **48**:2 (2003), 175–187. MR 2003m:30043 Zbl 1041.30006

[Kalaj 2004] D. Kalaj, "Quasiconformal harmonic functions between convex domains", *Publ. Inst. Math.* (*Beograd*) (*N.S.*) **76**:90 (2004), 3–20. MR 2005j:30032 Zbl 1220.30032

[Kalaj 2008] D. Kalaj, "Quasiconformal and harmonic mappings between Jordan domains", *Math. Z.* **260**:2 (2008), 237–252. MR 2009e:30038 Zbl 1151.30014

[Kalaj 2011a] D. Kalaj, "Harmonic mappings and distance function", *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **10**:3 (2011), 669–681. MR 2012j:30054 Zbl 1252.30018

[Kalaj 2011b] D. Kalaj, "Invertible harmonic mappings beyond the Kneser theorem and quasiconformal harmonic mappings", *Studia Math.* **207**:2 (2011), 117–136. MR 2012k:30065 Zbl 1279.30032

[Kalaj 2012] D. Kalaj, "On boundary correspondences under quasiconformal harmonic mappings between smooth Jordan domains", *Math. Nach.* **285**:2-3 (2012), 283–294. MR 2881282 Zbl 1251.30032

[Kalaj 2013] D. Kalaj, "A priori estimate of gradient of a solution of a certain differential inequality and quasiconformal mappings", *J. Anal. Math.* **119**:1 (2013), 63–88. MR 3043147 Zbl 1270.35148

[Kalaj and Mateljević 2006] D. Kalaj and M. Mateljević, "Inner estimate and quasiconformal harmonic maps between smooth domains", *J. Anal. Math.* **100** (2006), 117–132. MR 2008b:30033 Zbl 1173.30311

[Kellogg 1912] O. D. Kellogg, "Harmonic functions and Green's integral", *Trans. Amer. Math. Soc.* **13**:1 (1912), 109–132. MR 1500909 JFM 43.0889.01

[Lesley and Warschawski 1978] F. D. Lesley and S. E. Warschawski, "On conformal mappings with derivative in VMOA", *Math. Z.* **158**:3 (1978), 275–283. MR 57 #6393 Zbl 0374.30007

[Lewy 1936] H. Lewy, "On the non-vanishing of the Jacobian in certain one-to-one mappings", *Bull. Amer. Math. Soc.* **42**:10 (1936), 689–692. MR 1563404 Zbl 0015.15903

[Martio 1968] O. Martio, "On harmonic quasiconformal mappings", *Ann. Acad. Sci. Fenn. Ser. A I No.* **425** (1968), 10. MR 38 #4678 Zbl 0162.37902

[Pavlović 2002] M. Pavlović, "Boundary correspondence under harmonic quasiconformal homeomorphisms of the unit disk", *Ann. Acad. Sci. Fenn. Math.* **27**:2 (2002), 365–372. MR 2003g:30030 Zbl 1017.30014

[Pavlović 2004] M. Pavlović, *Introduction to function spaces on the disk*, Posebna Izdanja **20**, Matematički Institut SANU, Belgrade, 2004. MR 2006d:30001 Zbl 1107.30001

[Pommerenke 1975] C. Pommerenke, *Univalent functions*, Studia Mathematica/Mathematische Lehrbücher **25**, Vandenhoeck & Ruprecht, Göttingen, 1975. MR 58 #22526 Zbl 0298.30014

[Warschawski 1961] S. E. Warschawski, "On differentiability at the boundary in conformal mapping", *Proc. Amer. Math. Soc.* **12** (1961), 614–620. MR 24 #A1374 Zbl 0100.28803

[Warschawski 1970] S. E. Warschawski, "Boundary derivatives of minimal surfaces", *Arch. Rational Mech. Anal.* **38** (1970), 241–256. MR 41 #7549 Zbl 0209.41802

[Zygmund 1959] A. Zygmund, *Trigonometric series, I*, 2nd ed., Cambridge Univ. Press, 1959. MR 21 #6498 Zbl 0085.05601

Received February 7, 2014. Revised December 17, 2014.

David Kalaj
Faculty of Natural Sciences and Mathematics
University of Montenegro
Dzordza Vasingtona b.b.
81000 Podgorica
Montenegro
davidk@ac.me

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.229

SEMISIMPLE SUPER TANNAKIAN CATEGORIES WITH A SMALL TENSOR GENERATOR

THOMAS KRÄMER AND RAINER WEISSAUER

We consider semisimple super Tannakian categories generated by an object whose symmetric or alternating tensor square is simple up to trivial summands. Using representation theory, we provide a criterion to identify the corresponding Tannaka super groups that applies in many situations. As an example we discuss the tensor category generated by the convolution powers of an algebraic curve inside its Jacobian variety.

1. Introduction

The goal of this paper is to classify reductive super groups with a representation which is *small* in the sense that its symmetric or alternating square is irreducible or splits into an irreducible plus a trivial representation. This discussion fits into the general framework of small objects in tensor categories over an algebraically closed field k of characteristic zero, where by definition a *tensor category over* k is a rigid symmetric monoidal k-linear abelian category $\mathbf{C}$ whose unit object $\mathbf{1} \in \mathbf{C}$ satisfies $\mathrm{End}(\mathbf{1}) = k$. Recall that the structure of a monoidal category is given by a k-linear exact bifunctor $- \otimes - : \mathbf{C} \times \mathbf{C} \to \mathbf{C}$ together with a unit object and associativity constraints $a_{U,V,W} : U \otimes (V \otimes W) \xrightarrow{\sim} (U \otimes V) \otimes W$ for $U, V, W \in \mathbf{C}$ such that the usual compatibilities hold. A monoidal category is called symmetric if it is equipped with symmetry constraints $s_{U,V} : U \otimes V \xrightarrow{\sim} V \otimes U$ which are compatible with the previous structure and satisfy $s_{V,U} \circ s_{U,V} = \mathrm{id}$. It is called rigid if to every $V \in \mathbf{C}$ one may functorially attach an object $V^\vee \in \mathbf{C}$ with natural isomorphisms

$$\mathrm{Hom}(U \otimes V, W) \xrightarrow{\sim} \mathrm{Hom}(U, V^\vee \otimes W).$$

Tensor categories are ubiquitous in many areas of mathematics like representation theory, topology and algebraic geometry [Deligne et al. 1982; Gabber and Loeser 1996; Krämer and Weissauer 2015; Krämer 2014]. The typical example is the category $\mathbf{C} = \mathrm{Rep}_k(G)$ of finite-dimensional algebraic super representations of an affine super group scheme G over k. Here the representation spaces are super vector

MSC2010: primary 20G05, 20C15; secondary 18D10.
Keywords: Tannakian category, tensor generator, symmetric square, alternating square, irreducible representation, reductive super group.

spaces, i.e., $\mathbb{Z}/2\mathbb{Z}$-graded vector spaces $V = V_0 \oplus V_1$ of finite dimension over k, and the symmetry constraints are defined by the sign rule $s_{U,V}(u \otimes v) = (-1)^{\alpha\beta}\, v \otimes u$ for $u \in U_\alpha$, $v \in V_\beta$ and $\alpha, \beta \in \mathbb{Z}/2\mathbb{Z}$.

We say that a tensor category $\mathbf{C}$ over k is *algebraic* if any object $V \in \mathbf{C}$ is of finite length $\ell(V)$ and if the length $\ell(V^{\otimes n})$ of tensor powers grows at most polynomially in n. Over an algebraically closed field k of characteristic zero, Deligne has shown [2002] that every algebraic tensor category is equivalent as a tensor category over k to the category $\mathrm{Rep}_k(G, \varepsilon)$ of finite-dimensional algebraic super representations V of an affine super group scheme G over k with the property that a certain element $\varepsilon \in G(k)$ acts via the parity automorphism on V. Here the parity automorphism of a super vector space $V = V_0 \oplus V_1$ is given by $(-1)^\alpha$ on V_α for $\alpha \in \mathbb{Z}/2\mathbb{Z}$. Note that the above framework includes the usual representation categories of algebraic groups G by taking $\varepsilon = 1$. Since by definition $\mathrm{Rep}_k(G, \varepsilon)$ is a full subcategory of the algebraic tensor category $\mathrm{Rep}_k(G)$ of all algebraic super representations, for the study of small objects it suffices to consider the latter.

If such an algebraic tensor category $\mathbf{C} = \mathrm{Rep}_k(G)$ has a *tensor generator* X in the sense that any object is a subquotient of a tensor power $(X \oplus X^\vee)^{\otimes r}$ for some $r \in \mathbb{N}$, then the super group scheme G is of finite type over k and will be called the *Tannaka super group* of the category. We then have a faithful algebraic super representation $G \hookrightarrow \mathrm{GL}(V)$ on the finite-dimensional super vector space V associated to X. In what follows, by an *algebraic super group* over k we mean an affine super group scheme of finite type over k. Coming back to the general case, any algebraic tensor category $\mathbf{C}$ is the direct limit of tensor subcategories with a tensor generator, so the corresponding affine super group scheme G is an inverse limit of algebraic super groups over k, and for the study of small objects it suffices to consider algebraic super groups. Unfortunately, in contrast to the situation for ordinary algebraic groups, the representation theory of algebraic super groups is hardly understood. Even for the general linear super groups $G = \mathrm{GL}_{m|n}(k)$ over k the categories $\mathrm{Rep}_k(G)$ are not semisimple, and their tensor structure seems to be rather complicated. For example, in general the tensor product of irreducible objects is not a direct sum of irreducible objects. This often makes it desirable to replace $\mathbf{C}$ by some quotient category with simpler properties.

For any algebraic tensor category $\mathbf{C}$ over k, a general construction due to André and Kahn [2002, Section 8] together with the above result by Deligne implies that there is a universal k-linear (though in general not exact) quotient functor $\pi : \mathbf{C} \to \mathbf{C}^{\mathrm{red}}$ of algebraic tensor categories such that $\mathbf{C}^{\mathrm{red}}$ is semisimple. An indecomposable object $V \in \mathbf{C}$ becomes isomorphic to zero in the quotient category $\mathbf{C}^{\mathrm{red}}$ if and only if its super dimension $\dim(V_0) - \dim(V_1)$ is zero. Furthermore the functor π maps indecomposable objects to irreducible or zero objects, so it maps small objects to small objects. Bearing this in mind, we say an affine super group scheme G over k is *reductive* if the category $\mathrm{Rep}_k(G)$ is semisimple. The reductive algebraic super groups over an algebraically closed field k of characteristic zero have been classified in [Weissauer 2009]. In particular, they are all isogenous to products of ordinary reductive groups and orthosymplectic super groups $\mathrm{OSp}_{1|2m}(k)$, and their representation theory may be understood in terms of the representation theory of ordinary connected reductive groups and finite groups. In general, an affine super group scheme over k is reductive if and only if it is an inverse limit of reductive algebraic super groups. The above construction then associates to any affine super group scheme G over k a reductive super group scheme G^{red} over k, where $\mathbf{C}^{\mathrm{red}} = \mathrm{Rep}_k(G^{\mathrm{red}})$ for the category $\mathbf{C} = \mathrm{Rep}_k(G)$.

While from a theoretical point of view this seems to give a rather satisfying picture, in concrete applications the algebraic tensor categories arising from the construction of André and Kahn are often hard to approach. The case of a classical algebraic group G over k, where $G^{\mathrm{red}} = G/U$ for the unipotent radical $U \trianglelefteq G^0$ of the connected component, is not typical. In general there may be no simple relation between G^{red} and G. For the general linear super groups $G = \mathrm{GL}_{m|n}(k)$ with $m, n > 1$ the associated reductive super group schemes G^{red} are not even of finite type over k. One of the motivations for studying small objects in algebraic tensor categories is to get a better understanding of the construction of André and Kahn in such situations.

Apart from examples in representation theory, this is useful also in algebraic geometry, especially in the context of Brill–Noether sheaves [Krämer and Weissauer 2013; Weissauer 2007; Weissauer 2008]. For a smooth complex projective variety X, the image of X in its Albanese variety defines a distinguished object V of a semisimple algebraic tensor category $\mathbf{C} = \mathbf{C}(X)$ which is constructed via convolutions of perverse sheaves, see [Krämer and Weissauer 2015]. The corresponding Tannaka super group $G = G_X$ is a classical reductive complex algebraic group which is an intrinsic invariant of the variety X. If the object $V \in \mathbf{C}$ is small, our main result (Theorem 1.1) gives a criterion to determine this group. In Section 6 we illustrate this for a smooth curve X of genus $g \geq 1$. It has been shown in [Weissauer 2007] that in this case

$$G_X = \begin{cases} \mathrm{Sp}_{2g-2}(\mathbb{C}) & \text{if } X \text{ is hyperelliptic,} \\ \mathrm{SL}_{2g-2}(\mathbb{C}) & \text{otherwise;} \end{cases}$$

our criterion leads to a very short and much simpler proof of this result.

Returning to representation theory, let k again be an algebraically closed field of characteristic zero. The main goal of this paper is to classify all reductive super groups G over k that arise as the Tannaka super group of a semisimple tensor category with a small tensor generator; see Theorem 1.1. For simplicity, in what follows the term *representation* refers to a representation on a super vector space in the case of true super groups, but to an ordinary representation otherwise. For V in $\mathrm{Rep}_k(G)$ we denote by

$$T_\epsilon(V) = \begin{cases} S^2(V) & \text{for } \epsilon = +1, \\ \Lambda^2(V) & \text{for } \epsilon = -1, \end{cases}$$

the symmetric and the alternating squares with respect to the symmetry constraint for super vector spaces. If $T_\epsilon(V)$ is irreducible or a direct sum of an irreducible and a one-dimensional trivial representation $\mathbf{1}$, we say V is *ϵ-small* (or just *small*). Small representations are irreducible. If the trivial direct summand $\mathbf{1}$ occurs in $T_\epsilon(V)$, then V is isomorphic to its dual $V^\vee$ and hence carries a nondegenerate symmetric or alternating bilinear form. We say that V is *very small* if both $S^2(V)$ and $\Lambda^2(V)$ are irreducible. Since $\dim_k(\mathrm{End}_G(V \otimes V)) = \dim_k(\mathrm{End}_G(V \otimes V^\vee))$, this is the case if and only if $V \otimes V^\vee \cong W \oplus \mathbf{1}$ for some irreducible representation $W \in \mathrm{Rep}_k(G)$.

By definition a super group is *quasisimple* if it is a perfect central extension of a (nonabelian) simple super group. For the finite quasisimple groups G very small and self-dual small faithful representations have been classified by Magaard, Malle and Tiep [2002, Theorem 7.14], using earlier results of Magaard and Malle [1998]. In a more general setup the list of very small representations has been extended by Guralnick and Tiep [2005, Theorem 1.5] to arbitrary reductive groups. In particular, except for the standard representation of the special linear group, very small representations of G only exist if the quotient $G/Z(G)$ by the center $Z(G)$ is finite. The class of small representations is much richer and contains several cases with $\dim(G/Z(G)) > 0$.

To state our main result we use the following notation. For super groups G_i and representations $V_i \in \mathrm{Rep}_k(G_i)$, define $G_1 \otimes G_2 \subset \mathrm{GL}(V_1 \boxtimes V_2)$ to be the image of the exterior tensor product representation. If a group of automorphisms of $G_1 \otimes G_2$ contains elements that interchange the two subgroups $G_1 \otimes \{1\}$ and $\{1\} \otimes G_2$, we say that it *flips the two factors*. If a group acts transitively on a set X and if the action on the set of 2-element subsets of X is still transitive, we say that the group acts 2-*homogeneously* on X. If for $V \in \mathrm{Rep}_k(G)$ the restriction $V|_K$ to some normal abelian subgroup $K \trianglelefteq G$ splits into a direct sum of pairwise distinct characters that are permuted 2-homogeneously and faithfully by the adjoint action of G/K, we say that the representation V is 2-*homogeneous monomial*. Finally, a finite p-group E is called *extraspecial* if $E/Z(E)$ is elementary abelian and $Z(E) = [E, E]$ is cyclic of order p. Then $|E| = p^{1+2n}$ for some $n \in \mathbb{N}$, and for any nontrivial character ω of $Z(E) \cong \mathbb{Z}/p\mathbb{Z}$ there is a unique irreducible representation $V_\omega \in \mathrm{Rep}_k(E)$ with dimension p^n on which $Z(E)$ acts via ω [Dornhoff 1971, Theorem 31.5].

Theorem 1.1. *Let G be a reductive super group and $V \in \mathrm{Rep}_k(G)$ an ϵ-small faithful representation of super dimension $d > 0$. Then one of the following holds*:

(a) *The connected component $G^0 \subseteq G$ is quasisimple and the restriction $V|_{G^0}$ remains ϵ-small. In this case the possible Dynkin types of G^0 and the highest weights of $V|_{G^0}$ are given in Theorem 4.1.*

(b) $(G^0, V|_{G^0}) \cong (G_1 \otimes G_1, W \boxtimes W)$ *where* $G_1 \in \{\mathrm{SL}_m(k), \mathrm{GL}_m(k)\}$ *and where W is the m-dimensional standard representation or its dual. Here G flips the two factors so that $G \cong G^0 \rtimes \mathbb{Z}/2\mathbb{Z}$ and $\epsilon = -1$.*

(c) *There exists an embedding* $G \hookrightarrow \mathrm{GO}_4(k)$ *such that V is the restriction of the four-dimensional orthogonal standard representation*, *and $\epsilon = +1$.*

(d) *The representation V is* 2*-homogeneous monomial*; *then $\epsilon = -1$ unless V has* (*nonsuper*) *dimension* $\dim_k(V) \leq 2$.

(e) *The group $G = Z(G) \cdot S$ is a* (*not necessarily direct*, *but commuting*) *product of its center and some finite subgroup $S \subseteq G$. Furthermore we have an exact sequence* $0 \to H \to S \to \mathrm{Out}(H)$ *where*

  (e_1) *either H is quasisimple*,

  (e_2) *or* $(H, V|_H) \cong (G_1 \otimes G_1, W \boxtimes W)$ *for some very small* $W \in \mathrm{Rep}_k(G_1)$, *in which case S flips the two factors and $\epsilon = -1$*,

  (e_3) *or H is a finite p-group for some prime p and contains a G-stable extraspecial subgroup E of order p^{2n+1} for some $n \in \mathbb{N}$. In this case $V|_E$ is irreducible with dimension p^n.*

By definition of the symmetry constraint, the parity flip $W = \Pi V$ with $W_0 = V_1$ and $W_1 = V_0$ satisfies $S^2(W) = \Lambda^2(V)$ and $\Lambda^2(W) = S^2(V)$. This parity flip changes the sign of the super dimension; since the super dimension of an irreducible representation of a reductive super group is always nonzero [Weissauer 2009, Lemma 15], this explains why we assumed $d > 0$ in Theorem 1.1.

Note that for any faithful irreducible $V \in \mathrm{Rep}_k(G)$, Schur's lemma implies that the center $Z(G)$ acts on V via scalar matrices. So either $Z(G) = \mathbb{G}_m$ or $Z(G)$ is a finite cyclic group. If the restriction $V|_{G^0}$ to the connected component remains irreducible, then the conclusion of Schur's lemma also holds with the center of G replaced by the centralizer $Z_G(G^0) \subseteq G$. Thus in the situation of case (a) the group of connected components is easily controlled since $G/(G^0 \cdot Z_G(G^0)) \hookrightarrow \mathrm{Out}(G^0)$ must be a subgroup of outer automorphisms fixing the isomorphism type of the representation $V|_{G^0}$ in the table of Theorem 4.1.

For the converse of Theorem 1.1 one readily checks that all representations V in case (a), (b), (e_2) are small. Concerning (c), recall that the group of orthogonal similitudes $\mathrm{GO}_4(k)$ is the product of its center with $\mathrm{GSO}_4(k) \cong \mathrm{GL}_2(k) \otimes \mathrm{GL}_2(k)$, and that for the latter any small representation must be a product of two very small ones. As a typical example of (d), for any 2-homogeneous subgroup F of the symmetric group $\mathfrak{S}_d$ we have the 2-homogeneous monomial small representation of $G = (\mathbb{G}_m)^d \rtimes F$ on $V = k^d$ with the natural action. Apart from a single extra case, the 2-homogeneous permutation groups on $d \geq 4$ letters are precisely the doubly transitive ones [Kantor 1969, Proposition 3.1; 1972], and the finite doubly transitive groups have been classified by Huppert, Hering and others [Dixon and Mortimer 1996, Section 7.7]. In the extraspecial case (e_3) the analysis of the smallness condition is more subtle and we postpone it to the remarks after the proof of Proposition 3.1. Thus altogether Theorem 1.1 gives an essentially complete picture except for the case (e_1) of finite quasisimple groups, which would require a close analysis of the representations of finite groups of Lie type generalizing the methods of Guralnick, Magaard, Malle and Tiep.

For the sake of brevity, in what follows the term *group* will always be taken to include super groups. However, until Section 4 the term *dimension* will still refer to the ordinary dimension (as opposed to the super dimension).

2. Clifford–Mackey theory

Let us say that $V \in \mathrm{Rep}_k(G)$ is *strongly irreducible* if for any noncentral normal subgroup $H \trianglelefteq G$ of finite index the restriction $V|_H$ is irreducible.

Proposition 2.1. *For any faithful ϵ-small representation $V \in \mathrm{Rep}_k(G)$ one of the following cases occurs*:

(a) *The representation V is strongly irreducible.*

(b) *V is a 2-homogeneous monomial representation. In this case $\epsilon = -1$ or V has dimension $\dim_k(V) \leq 2$.*

(c) *There exists an embedding $G \hookrightarrow \mathrm{GO}_4(k)$ such that V is the restriction of the four-dimensional orthogonal standard representation.*

Proof. Let $H \trianglelefteq G$ be a normal subgroup. If the restriction $V|_H$ is not isotypic, let $V|_H = W_1 \oplus \cdots \oplus W_n$ be its isotypic decomposition. Then $V \cong \mathrm{Ind}_{H_1}^G(W_1)$ is induced from a representation of the stabilizer $H_1 \leq G$ of W_1, and we get a splitting into two G-stable summands

$$T_\epsilon(V) \cong \mathrm{Ind}_{H_1}^G(T_\epsilon(W_1)) \oplus \left[\bigoplus_{i \neq j} W_i \otimes W_j\right]_\epsilon,$$

where the subscript ϵ in the second summand indicates the ϵ-eigenspace of the symmetry constraint which flips the two factors of the tensor product. Since in the nonisotypic case we have $n > 1$, ϵ-smallness implies that $\dim_k(W_1) = 1$, and $\epsilon = -1$ or $\dim_k(V) = n = 2$. All W_i have dimension one, so $V|_H$ splits as a sum of pairwise distinct characters. Now G acts by conjugation on the set X of these characters, and the kernel K of this permutation representation of G is a normal subgroup which is abelian since V is faithful. So (b) holds.

Now suppose that $V|_H$ is isotypic. Then, as in [Dornhoff 1971, Theorem 25.9], there are projective representations U_1, U_2 of G such that $V \cong U_1 \otimes U_2$, where the restriction $U_1|_H$ is irreducible and where every $h \in H$ acts as the identity on U_2. Then

$$T_\pm(V) \cong (T_+(U_1) \otimes T_\pm(U_2)) \oplus (T_-(U_1) \otimes T_\mp(U_2)),$$

and since V is small, one of the summands $T_{\epsilon_1}(U_1) \otimes T_{\epsilon_2}(U_2)$ must have dimension at most one. By direct inspection this can happen only if either $d_i = \dim_k(U_i) = 1$ for some $i \in \{1, 2\}$, or $d_1 = d_2 = 2$. Now $V|_H \cong U_1 \oplus \cdots \oplus U_1 = d_2 \cdot U_1$ so that for $d_1 = 1$ the group H is contained in the center $Z(G)$, which acts on V via scalar matrices. For $d_2 = 1$ the restriction $V|_H$ remains irreducible. For $d_1 = d_2 = 2$ case (c) occurs since $U_1, U_2 \in \mathrm{Rep}_k(H)$ extend to projective representations of the whole group G whose image then is contained in the product of its center with the special orthogonal similitude group $\mathrm{GL}_2(k) \otimes \mathrm{GL}_2(k) \cong \mathrm{GSO}_4(k)$. □

3. Reduction to the quasisimple case

Next we study strongly irreducible $V \in \mathrm{Rep}_k(G)$. To treat the case of finite groups simultaneously with the case of positive-dimensional reductive groups, recall from [Aschbacher 2000, Section 31] that for finite groups S the *generalized Fitting subgroup* $F^*(S)$ plays a role very similar to the one which for a reductive algebraic group is played by the derived group of the connected component. By definition $F^*(S) \leq S$ is the subgroup of S generated by the largest nilpotent normal subgroup together with the subnormal quasisimple subgroups. Here a subgroup $N \leq S$ is called *subnormal* if there is a chain $N = N_1 \trianglelefteq N_2 \trianglelefteq \cdots \trianglelefteq N_m = S$ of subgroups where each member of the chain is a normal subgroup of the next member. To make the role of the generalized Fitting subgroup more precise, let us temporarily call a group *basic* if it is either quasisimple or a finite p-group for some prime p. For a given group G we define $H \trianglelefteq G$ as follows:

- If $G^0 \subseteq Z(G)$, then $G = Z(G) \cdot S$ for some finite normal subgroup $S \trianglelefteq G$, and fixing such a subgroup we take $H = F^*(S)$.
- Otherwise we take $H = [G^0, G^0]$ to be the derived group of the connected component. The theory of reductive groups then implies $G^0 = Z(G^0) \cdot H$.

In both cases we can find a central isogeny $\tilde{H} = H_1 \times \cdots \times H_n \twoheadrightarrow H$ such that the image of each H_i is normal in G. Choosing the labeling in a suitable way, we may furthermore assume that for each i we have a central isogeny $\tilde{H}_i = (G_i)^{s_i} \twoheadrightarrow H_i$ for s_i copies of a suitable basic group G_i and that the images of these s_i copies are permuted transitively by the adjoint action of G.

Proposition 3.1. *For any faithful ϵ-small strongly irreducible $V \in \mathrm{Rep}_k(G)$ with dimension $\dim_k(V) > 1$ one of the following cases occurs*:

(a) *The group H is quasisimple.*

(b) $(H, V|_H) \cong (G_1 \otimes G_1, W \boxtimes W)$ *for some very small* $W \in \mathrm{Rep}_k(G_1)$, H *flips the two factors, and we have* $\epsilon = -1$.

(c) H *contains an extraspecial* G*-stable subgroup* E *of order* p^{2n+1} *for some prime* p *such that* $V|_E$ *is irreducible of dimension* p^n.

(d) *We have an embedding* $G \hookrightarrow \mathrm{GO}_4(k)$ *such that* V *is the restriction of the four-dimensional standard representation.*

Proof. We first claim that $H \not\subseteq Z(G)$. Indeed, for the finite group case recall that the generalized Fitting subgroup contains its own centralizer [Aschbacher 2000], so $H \subseteq Z(G)$ would imply $S = H$ and then $G = Z(G)$ would be abelian. In the infinite case where G^0 is not central, the strong irreducibility implies that $V|_{G^0}$ is irreducible so that the connected reductive group G^0 cannot be a torus. Thus indeed $H \not\subseteq Z(G)$.

Hence we can assume that the image of each H_i in G is a *noncentral* subgroup by discarding any occurring central components and saturating the other components with the center. Since $V|_{\tilde{H}} \cong U_1 \boxtimes \cdots \boxtimes U_n$ with irreducible $U_i \in \mathrm{Rep}_k(H_i)$, we get $n = 1$ by strong irreducibility. Hence $\tilde{H} \cong (G_1)^s$ for $s = s_1$ and again we get a decomposition $V|_{\tilde{H}} \cong W_1 \boxtimes \cdots \boxtimes W_s$ with irreducible $W_i \in \mathrm{Rep}_k(G_1)$, but now the adjoint action of G permutes the s factors G_1 transitively so that all W_i are isomorphic to a single $W \in \mathrm{Rep}_k(G_1)$. In the decomposition

$$T_\epsilon(V)|_H \cong \bigoplus_{r=0}^{s} T_{r,\epsilon} \quad \text{with } T_{r,\epsilon} = \bigoplus_{\substack{\epsilon_1 \cdots \epsilon_s = \epsilon \\ \#\{i \mid \epsilon_i = +1\} = r}} T_{\epsilon_1}(W) \boxtimes \cdots \boxtimes T_{\epsilon_s}(W)$$

each summand $T_{r,\epsilon}$ is stable under the action of G. By smallness it then follows that $s \leq 2$, and for $s = 2$ the conclusions of (b) or (d) hold.

So we may assume $s = 1$ and $H = G_1$ is a basic group. If case (a) does not occur, then H is a finite p-group for some prime p. Consider then a minimal G-stable noncentral subgroup $M \trianglelefteq H$. By minimality the subgroup $[M, M]$ is contained in $A := M \cap Z(H)$ so that the quotient $U := M/A$ is abelian. Looking at the p-torsion part of this quotient one obtains, again by minimality, that U is elementary abelian. The commutator induces a bilinear map $[\cdot, \cdot] : U \times U \to A$, and if we identify A with a subgroup of $\mathbb{G}_m$ via Schur's lemma, $p \cdot U = 0$ implies that $[M, M]$ is contained in the subgroup $\mu_p \subseteq A$ of p-th roots of unity. So M/μ_p is abelian and in fact elementary abelian: Otherwise by minimality its p-torsion subgroup would lie in the cyclic group A/μ_p so that the abelian p-group M/μ_p would be cyclic as well. But then M would be abelian, and this is impossible since it admits the faithful irreducible representation $V|_M$ of dimension $d > 1$.

Thus M/μ_p is elementary abelian, and we claim that the extraspecial case (c) occurs. Indeed, either $A = \mu_p$ or $A = \mu_{p^2}$. For $A = \mu_p$ the subgroup $E = M$ satisfies our requirements, so suppose that $A = \mu_{p^2}$. Since M/μ_p is elementary abelian, the Frattini subgroup is $\Phi(M) = \mu_p$ by [Aschbacher 2000, (23.2)]. The Frattini subgroup is the intersection of all maximal subgroups, so it follows that there exists a maximal subgroup $E \leq M$ which contains μ_p but not μ_{p^2}. Then $M = \mu_{p^2} \cdot E$, and $E \leq M$ is an extraspecial subgroup. We will be done if we can show this subgroup is stable under the group $\mathrm{Aut}_A(M)$ of automorphisms of M that are trivial on A. But this follows from the observation that every automorphism of E which is trivial on μ_p extends uniquely to an element of $\mathrm{Aut}_A(M)$, which gives a natural identification $\mathrm{Aut}_A(M) \cong \mathrm{Aut}_{\mu_p}(E)$ compatible with the actions on M and E. □

We remark that the only instance of case (b) in Proposition 3.1 with $\dim(H) > 0$ is $G_1 \cong \mathrm{SL}_m(k)$, acting on $W \cong k^m$ either via the standard representation or via its dual. Indeed this will follow from Theorem 4.1 below, applied to the very small representation W of the Lie algebra of G_1. Alternatively one could use [Guralnick and Tiep 2005].

In case (c) where H contains a G-stable extraspecial p-group E, put $|E| = p^{1+2n}$ with $n \in \mathbb{N}$. For any nontrivial character $\omega : Z(E) \cong \mathbb{Z}/p\mathbb{Z} \to \mathbb{G}_m$ there exists a unique irreducible representation $V_\omega \in \mathrm{Rep}_k(E)$ of dimension p^n on which $Z(E)$ acts via the character ω, and these are already all the irreducible representations of dimension > 1 by [Dornhoff 1971, Theorem 31.5]. Hence in case (c) we have $V|_E \cong V_\omega$ for a uniquely determined character ω. To decide which of the occurring representations are small, note that for the finite group S such that $H = F^*(S)$, we have a natural homomorphism $S \to \mathrm{Out}(E)$. We now distinguish two cases depending on p.

For $p > 2$ we have $\omega^2 \neq \mathbf{1}$, so $T_\epsilon(V)|_E$ is an isotypic multiple of V_{ω^2}. Then Mackey theory [Dornhoff 1971, Theorem 25.9] gives a tensor product decomposition $T_\epsilon(V) \cong U \otimes W_\epsilon$ where U and W_ϵ are projective representations of the group S such that $U|_E \cong V_{\omega^2}$ and such that every element of E acts trivially on W_ϵ. Via the nondegenerate alternating bilinear form defined by the commutator on $E/Z(E) \cong (\mathbb{F}_p)^{2n}$ we can identify the image of S in $\mathrm{Out}(E)$ with a subgroup of the symplectic group $\mathrm{Sp}_{2n}(\mathbb{F}_p)$. Looking at dimensions one then obtains from [Tiep and Zalesskii 1996, Theorem 5.2] that W_ϵ must be one of the two Weil representations of dimension $(p^n + \epsilon)/2$. Hence V is ϵ-small if and only if the image of S inside $\mathrm{Sp}_{2n}(\mathbb{F}_p)$ acts irreducibly on this Weil representation.

For $p = 2$ on the other hand, $\omega^2 = \mathbf{1}$, so that the restriction $T_\epsilon(V)|_E$ is a sum of characters. By [Winter 1972] we can identify $\mathrm{Out}(E)$ with an orthogonal group $O^\pm_{2n}(\mathbb{F}_2)$ where the type $\pm$ of the quadratic form depends on E. Recall that a nondegenerate quadratic form on $(\mathbb{F}_2)^{2n}$ has type $\pm$ if and only if there are precisely $2^{n-1}(2^n \pm 1)$ isotropic vectors for this form. One then obtains the following identifications:

- If the quadratic form has $+$ type, the isotropic vectors in $(\mathbb{F}_2)^{2n}$ correspond precisely to the characters in $T_+(V)|_E$.
- If the quadratic form has $-$ type, the isotropic vectors in $(\mathbb{F}_2)^{2n}$ correspond precisely to the characters in $T_-(V)|_E$.

A similar interpretation holds for the anisotropic vectors. Hence it follows that V is small if and only if the image of S inside $O_{2n}^{\pm}(\mathbb{F}_2)$ acts transitively on the nonzero isotropic resp. anisotropic vectors. Note that the set of isotropic vectors always includes the zero vector as a single orbit, corresponding to the trivial summand $\mathbf{1} \hookrightarrow T_\epsilon(V)$.

4. Lie super algebras

It remains to determine all small $V \in \mathrm{Rep}_k(G)$ when $H = [G^0, G^0]$ is quasisimple and $V|_H$ is irreducible. By the classification of reductive super groups in [Weissauer 2009], the Lie super algebra $\mathfrak{g}$ of H must then either be an ordinary simple Lie algebra or an orthosymplectic Lie super algebra $\mathfrak{osp}_{1|2m}(k)$ with $m \in \mathbb{N}$. Note that $\mathrm{Rep}_k(H)$ is a full subcategory of $\mathrm{Rep}_k(\mathfrak{g})$, where the latter denotes the category of all Lie algebra representations of the Lie super algebra $\mathfrak{g}$ on finite-dimensional super vector spaces over k. In particular $V|_H$ defines an irreducible representation of $\mathfrak{g}$.

The passage to representations of Lie algebras leads to a seemingly weaker notion of smallness. By the comments after Theorem 1.1 we know that $G/(G^0 \cdot Z_G(G^0))$ is a subgroup of $\mathrm{Out}(G^0)$ such that conjugation by any element φ of this subgroup fixes the isomorphism type of $V|_H$. For an irreducible summand $W \hookrightarrow T_\epsilon(V)$ in $\mathrm{Rep}_k(G)$ it may happen that the restriction $W|_H$ splits into several irreducible summands, but all these summands must be conjugate via automorphisms φ as above. Abstracting from this situation, let us now denote by $\mathfrak{g}$ any ordinary simple Lie algebra or $\mathfrak{osp}_{1|2m}(k)$ with $m \in \mathbb{N}$. We say that a representation $V \in \mathrm{Rep}_k(\mathfrak{g})$ is *ϵ-small* if either $T_\epsilon(V) \cong W$ or $T_\epsilon(V) \cong W \oplus \mathbf{1}$, where W is a sum of irreducible representations which are all conjugate to each other via automorphisms $\varphi \in \mathrm{Aut}(\mathfrak{g})$ fixing the isomorphism type of V. To finish the proof of Theorem 1.1 we classify all irreducible small representations in this sense. For a uniform treatment the terms dimension, vector space, trace and Lie algebra will from now on be taken in the super sense for $\mathfrak{osp}_{1|2m}(k)$ but in the ordinary sense otherwise.

We denote by $\varpi_1, \dots, \varpi_m$ the fundamental dominant weights of $\mathfrak{g}$ with respect to some fixed system of simple positive roots; see [Rittenberg and Scheunert 1982, Section 2.1] for the orthosymplectic Lie algebra $\mathfrak{g} = \mathfrak{osp}_{1|2m}(k)$ whose Dynkin type we abbreviate by BC_m. Put

$$\beta_i = \begin{cases} 2\varpi_m & \text{if } \mathfrak{g} = \mathfrak{osp}_{1|2m}(k) \text{ and } i = m, \\ \varpi_i & \text{otherwise.} \end{cases}$$

The irreducible finite-dimensional representations of $\mathfrak{g}$ are parametrized by highest weights $\lambda = \sum_{i=1}^m a_i \beta_i$ with $a_i \in \mathbb{N}_0$, see [Djoković 1976b, Theorem 6]. For any such λ we denote by V_λ the associated positive-dimensional irreducible representation. Note that, in the super case, negative-dimensional irreducible representations are obtained by the parity flip $W_\lambda = \Pi V_\lambda$ with $\dim(W_\lambda) = -\dim(V_\lambda)$ and $S^2(W_\lambda) \cong \Lambda^2(V_\lambda)$.

Theorem 4.1. *A positive-dimensional irreducible representation $V_\lambda \in \mathrm{Rep}_k(\mathfrak{g})$ is ϵ-small if and only if its highest weight λ appears in the following table*:

		λ	$\epsilon = +1$	$\epsilon = -1$
A_m	$m \geq 1$	$\beta_1,\ \beta_m$	$\star$	$\star$
	$m = 1$	$2\beta_1$	$\circ$	$\star$
		$3\beta_1$	$-$	$\circ$
	$m \geq 2$	$2\beta_1,\ 2\beta_m$	$-$	$\star$
	$m = 3$	β_2	$\circ$	$\star$
	$m \geq 4$	$\beta_2,\ \beta_{m-1}$	$-$	$\star$
	$m = 5$	β_3	$-$	$\circ$
B_m	$m \geq 2$	β_1	$\circ$	$\star$
	$m = 2$	β_2	$\star$	$\circ$
	$m = 3$	β_3	$\circ$	$-$
C_m	$m \geq 3$	β_1	$\star$	$\circ$
	$m = 3$	β_3	$-$	$\circ$
D_m	$m \geq 4$	β_1	$\circ$	$\star$
	$m = 4$	$\beta_3,\ \beta_4$	$\circ$	$\star$
	$m = 5$	$\beta_4,\ \beta_5$	$-$	$\star$
	$m = 6$	$\beta_5,\ \beta_6$	$-$	$\circ$
BC_m	$m \geq 1$	β_1	$\star$	$\circ$
E_6		$\beta_1,\ \beta_6$	$-$	$\star$
E_7		β_7	$-$	$\circ$
G_2		β_1	$\circ$	$-$

Here the label $\star$ means that $T_\epsilon(V_\lambda)$ is irreducible, $\circ$ *means that $T_\epsilon(V_\lambda) = W \oplus \mathbf{1}$ with W irreducible*, *and $-$ means that V_λ is not ϵ-small.*

Note that for $\mathfrak{g} = \mathfrak{sl}_2(k)$ with its two-dimensional standard representation st, any irreducible representation is a symmetric power $V_\lambda = S^n(st)$ of weight $\lambda = n\beta_1$ for some $n \in \mathbb{N}$. In this case Theorem 4.1 holds by direct inspection. A similar argument also works for $\mathfrak{g} = \mathfrak{osp}_{1|2}(k)$. Here we know from [Djoković 1976b, Theorems 7 and 11] that for $\lambda = n\beta_1$ the even subalgebra $\mathfrak{g}_0 = \mathfrak{sl}_2(k) \subset \mathfrak{g}$ acts on $V_\lambda = V_0 \oplus V_1$ via $V_0 = S^n(st)$ and $V_1 = S^{n-1}(st)$. A short computation yields the action on the even and odd parts of the tensor square $T_\epsilon(V)$ and Theorem 4.1 also holds in this case. Note that $\dim(V) = 1$ for all irreducible representations V of $\mathfrak{osp}_{1|2}(k)$. For all other cases we have:

Lemma 4.2. *For $\mathfrak{g} \neq \mathfrak{osp}_{1|2}(k)$ one has $\dim(V_\lambda) \leq \dim(\mathfrak{g})$ if and only if the highest weight λ appears among those listed in Tables 1 or 2.*

Proof. See [Andreev et al. 1967] for the ordinary case. For $\mathfrak{g} = \mathfrak{osp}_{1|2m}(k)$ with $m \geq 2$ we use the Kac–Weyl formula in [Tsohantjis and Cornwell 1990, Equation 11]. We embed the root system BC_m into a Euclidean space with standard basis $\epsilon_1, \dots, \epsilon_m$ such that $\beta_i = \epsilon_1 + \cdots + \epsilon_i$ for all i. The irreducible representations of $\mathfrak{osp}_{1|2m}(k)$ are parametrized by weights which in our basis are written $\lambda = (\lambda_1, \dots, \lambda_m)$ with integers $\lambda_1 \geq \cdots \geq \lambda_m \geq 0$. The Kac–Weyl formula gives

$$\dim(V_\lambda) = \prod_{1 \leq i < j \leq m} \left(\frac{\lambda_i - \lambda_j}{j - i} + 1 \right) \cdot \prod_{1 \leq i < j \leq m} \left(\frac{\lambda_i + \lambda_j}{2m + 1 - i - j} + 1 \right).$$

For $\lambda_1 \geq 2$ the second product is ≥ 2. Then the classical Weyl formula for the first product shows that $\dim(V_\lambda)$ is at least twice the dimension of the irreducible representation of $\mathfrak{sl}_m(k)$ with highest weight $\mu = (\lambda_1 - \lambda_m, \dots, \lambda_{m-1} - \lambda_m)$. Using that $\dim(\mathfrak{sl}_m(k)) \geq 2 \dim(\mathfrak{osp}_{1|2m}(k))$, it follows that μ is in the list for A_{m-1} in Table 1. Since $\lambda = \mu + \lambda_m \cdot \beta_m$ and since increasing the weight by β_m increases the dimension, this leaves only finitely many cases. For $\lambda_1 = 1$ we have $\lambda = \beta_r$

		λ	$S^2(V_\lambda)$	$\Lambda^2(V_\lambda)$
A_m	$m = 1$	β_1	$V_{2\beta_1}$	$\mathbf{1}$
	$m \geq 2$	β_1	$V_{2\beta_1}$	V_{β_2}
		β_m	$V_{2\beta_m}$	$V_{\beta_{m-1}}$
	$m \geq 2$	$2\beta_1$	$V_{4\beta_1} \oplus V_{2\beta_2}$	$V_{2\beta_1+\beta_2}$
		$2\beta_m$	$V_{4\beta_m} \oplus V_{2\beta_{m-1}}$	$V_{2\beta_m+\beta_{m-1}}$
	$m \geq 4$	β_2	$V_{2\beta_2} \oplus V_{\beta_4}$	$V_{\beta_1+\beta_3}$
		β_{m-1}	$V_{2\beta_{m-1}} \oplus V_{\beta_{m-3}}$	$V_{\beta_m+\beta_{m-2}}$
	$m = 3$	β_2	$V_{2\beta_2} \oplus \mathbf{1}$	$V_{\beta_1+\beta_3}$
	$m = 5$	β_3	$V_{2\beta_3} \oplus V_{\beta_1+\beta_5}$	$V_{\beta_2+\beta_4} \oplus \mathbf{1}$
	$m =$	β_3	$V_{2\beta_3} \oplus V_{\beta_1+\beta_5}$	$V_{\beta_2+\beta_4} \oplus V_{\beta_6}$
	$6, 7$	β_{m-2}	$V_{2\beta_{m-2}} \oplus V_{\beta_m+\beta_{m-4}}$	$V_{\beta_{m-1}+\beta_{m-3}} \oplus V_{\beta_{m-5}}$

Table 1. All λ with $1 < \dim(V_\lambda) < \dim(\mathfrak{g})$. For $\mathfrak{g} = \mathfrak{osp}_{1|2m}(k)$ we denote by $W_\mu = \Pi V_\mu$ the parity shifts of the highest weight modules. (Continues on the next page.)

B_m	$m \geq 2$	β_1	$V_{2\beta_1} \oplus \mathbf{1}$	V_{β_2}
	$m = 2$	β_2	$V_{2\beta_2}$	$V_{\beta_1} \oplus \mathbf{1}$
	$m = 3$	β_3	$V_{2\beta_3} \oplus \mathbf{1}$	$V_{\beta_1} \oplus V_{\beta_2}$
	$m = 4$	β_4	$V_{2\beta_4} \oplus V_{\beta_1} \oplus \mathbf{1}$	$V_{\beta_2} \oplus V_{\beta_3}$
	$m = 5$	β_5	$V_{2\beta_5} \oplus V_{\beta_2} \oplus V_{\beta_1}$	$V_{\beta_3} \oplus V_{\beta_4} \oplus \mathbf{1}$
	$m = 6$	β_6	$V_{2\beta_6} \oplus V_{\beta_3} \oplus V_{\beta_2}$	$V_{\beta_1} \oplus V_{\beta_4} \oplus V_{\beta_5} \oplus \mathbf{1}$
C_m	$m \geq 3$	β_1	$V_{2\beta_1}$	$V_{\beta_2} \oplus \mathbf{1}$
	$m \geq 4$	β_2	$V_{\beta_4} \oplus V_{2\beta_2} \oplus V_{\beta_2} \oplus \mathbf{1}$	$V_{2\beta_1} \oplus V_{\beta_1+\beta_3}$
	$m = 3$	β_2	$V_{2\beta_2} \oplus V_{\beta_2} \oplus \mathbf{1}$	$V_{2\beta_1} \oplus V_{\beta_1+\beta_3}$
		β_3	$V_{2\beta_3} \oplus V_{2\beta_1}$	$V_{2\beta_2} \oplus \mathbf{1}$
D_m	$m \geq 4$	β_1	$V_{2\beta_1} \oplus \mathbf{1}$	V_{β_2}
	$m = 4$	β_3	$V_{2\beta_3} \oplus \mathbf{1}$	V_{β_2}
		β_4	$V_{2\beta_4} \oplus \mathbf{1}$	V_{β_2}
	$m = 5$	β_4	$V_{2\beta_4} \oplus V_{\beta_1}$	V_{β_3}
		β_5	$V_{2\beta_5} \oplus V_{\beta_1}$	V_{β_3}
	$m = 6$	β_5	$V_{2\beta_5} \oplus V_{\beta_2}$	$V_{\beta_4} \oplus \mathbf{1}$
		β_6	$V_{2\beta_6} \oplus V_{\beta_2}$	$V_{\beta_4} \oplus \mathbf{1}$
	$m = 7$	β_6	$V_{2\beta_6} \oplus V_{\beta_3}$	$V_{\beta_1} \oplus V_{\beta_5}$
		β_7	$V_{2\beta_7} \oplus V_{\beta_3}$	$V_{\beta_1} \oplus V_{\beta_5}$
BC_m	$m \geq 2$	β_1	$V_{2\beta_1}$	$V_{\beta_2} \oplus \mathbf{1}$
	$m \geq 4$	β_2	$V_{2\beta_2} \oplus V_{\beta_2} \oplus V_{\beta_4} \oplus \mathbf{1}$	$V_{2\beta_1} \oplus V_{\beta_1+\beta_3}$
	$m = 2$	$\beta_1+\beta_2$	$V_{2\beta_1+2\beta_2} \oplus 2V_{2\beta_1+\beta_2} \oplus 2V_{2\beta_1}$ $\oplus W_{3\beta_1} \oplus W_{\beta_1+2\beta_2}$ $\oplus 2W_{\beta_1+\beta_2}$	$V_{4\beta_1} \oplus V_{2\beta_1+\beta_2} \oplus V_{3\beta_2} \oplus 2V_{2\beta_2}$ $\oplus 2V_{\beta_2} \oplus W_{3\beta_1+\beta_2} \oplus W_{\beta_1+2\beta_2}$ $\oplus W_{\beta_1+\beta_2} \oplus W_{\beta_1} \oplus \mathbf{1}$
		β_2	$V_{2\beta_2} \oplus V_{\beta_2} \oplus W_{\beta_1}$	$V_{2\beta_1} \oplus W_{\beta_1+\beta_2} \oplus \mathbf{1}$
	$m = 3$	β_2	$V_{2\beta_2} \oplus V_{\beta_2} \oplus W_{\beta_3}$	$V_{2\beta_1} \oplus V_{\beta_1+\beta_3} \oplus \mathbf{1}$
		β_3	$V_{2\beta_1} \oplus V_{\beta_1+\beta_3} \oplus V_{2\beta_3}$ $\oplus W_{\beta_1+\beta_2}$	$V_{2\beta_2} \oplus V_{\beta_2}$ $\oplus W_{\beta_1} \oplus W_{\beta_2+\beta_3} \oplus W_{\beta_3} \oplus \mathbf{1}$
	$m = 4$	β_4	$V_{2\beta_1} \oplus V_{\beta_2+\beta_4} \oplus V_{\beta_2} \oplus V_{2\beta_4}$ $\oplus V_{\beta_4} \oplus W_{\beta_1} \oplus W_{\beta_2+\beta_3}$ $\oplus W_{\beta_3} \oplus \mathbf{1}$	$V_{2\beta_2} \oplus V_{\beta_1+\beta_3} \oplus V_{2\beta_3}$ $\oplus W_{\beta_1+\beta_2} \oplus W_{\beta_1+\beta_4}$ $\oplus W_{\beta_3+\beta_4}$
	$m = 5$	β_5	$V_{2\beta_1} \oplus V_{\beta_1+\beta_3} \oplus V_{\beta_1+\beta_5}$ $\oplus V_{2\beta_3} \oplus V_{\beta_3+\beta_5} \oplus V_{2\beta_5}$ $\oplus W_{\beta_1+\beta_2} \oplus W_{\beta_1+\beta_4}$ $\oplus W_{\beta_3+\beta_4}$	$V_{2\beta_2} \oplus V_{\beta_2+\beta_4} \oplus V_{\beta_2} \oplus V_{2\beta_4}$ $\oplus V_{\beta_4} \oplus W_{\beta_1} \oplus W_{\beta_2+\beta_3}$ $\oplus W_{\beta_2+\beta_5} \oplus W_{\beta_3}$ $\oplus W_{\beta_4+\beta_5} \oplus W_{\beta_5} \oplus \mathbf{1}$
E_6		β_1	$V_{2\beta_1} \oplus V_{\beta_6}$	V_{β_3}
		β_6	$V_{2\beta_6} \oplus V_{\beta_1}$	V_{β_5}
E_7		β_7	$V_{2\beta_7} \oplus V_{\beta_1}$	$V_{\beta_6} \oplus \mathbf{1}$
F_4		β_4	$V_{2\beta_4} \oplus V_{\beta_4} \oplus \mathbf{1}$	$V_{\beta_3} \oplus V_{\beta_1}$
G_2		β_1	$V_{2\beta_1} \oplus \mathbf{1}$	$V_{\beta_1} \oplus V_{\beta_2}$

Table 1 (continued).

		λ	$S^2(V_\lambda)$	$\Lambda^2(V_\lambda)$
A_m	$m=1$	$2\beta_1$	$V_{4\beta_1} \oplus \mathbf{1}$	$V_{2\beta_1}$
	$m=2$	$\beta_1+\beta_2$	$V_{2\beta_1+2\beta_2} \oplus V_{\beta_1+\beta_2} \oplus \mathbf{1}$	$V_{3\beta_1} \oplus V_{3\beta_2} \oplus V_{\beta_1+\beta_2}$
	$m \geq 2$	$\beta_1+\beta_m$	$V_{2\beta_1+2\beta_m} \oplus V_{\beta_2+\beta_{m-1}}$ $\oplus V_{\beta_1+\beta_m} \oplus \mathbf{1}$	$V_{\beta_2+2\beta_m} \oplus V_{2\beta_1+\beta_{m-1}}$ $\oplus V_{\beta_1+\beta_m}$
B_m	$m=2$	$2\beta_2$	$V_{\beta_1} \oplus V_{2\beta_1} \oplus V_{4\beta_2} \oplus \mathbf{1}$	$V_{\beta_1+2\beta_2} \oplus V_{2\beta_2}$
	$m=3$	β_2	$V_{2\beta_1} \oplus V_{2\beta_2} \oplus V_{2\beta_3} \oplus \mathbf{1}$	$V_{\beta_1+2\beta_3} \oplus V_{\beta_2}$
	$m=4$	β_2	$V_{2\beta_1} \oplus V_{2\beta_2} \oplus V_{2\beta_4} \oplus \mathbf{1}$	$V_{\beta_1+\beta_3} \oplus V_{\beta_2}$
	$m \geq 5$	β_2	$V_{2\beta_1} \oplus V_{2\beta_2} \oplus V_{\beta_4} \oplus \mathbf{1}$	$V_{\beta_1+\beta_3} \oplus V_{\beta_2}$
C_m	$m \geq 3$	$2\beta_1$	$V_{4\beta_1} \oplus V_{2\beta_2} \oplus V_{\beta_2} \oplus \mathbf{1}$	$V_{2\beta_1} \oplus V_{2\beta_1+\beta_2}$
D_m	$m=4$	β_2	$V_{2\beta_1} \oplus V_{2\beta_2} \oplus V_{2\beta_3} \oplus V_{2\beta_4} \oplus \mathbf{1}$	$V_{\beta_2} \oplus V_{\beta_1+\beta_3+\beta_4}$
	$m=5$	β_2	$V_{2\beta_1} \oplus V_{2\beta_2} \oplus V_{\beta_4+\beta_5} \oplus \mathbf{1}$	$V_{\beta_2} \oplus V_{\beta_1+\beta_3}$
	$m \geq 6$	β_2	$V_{2\beta_1} \oplus V_{2\beta_2} \oplus V_{\beta_4} \oplus \mathbf{1}$	$V_{\beta_2} \oplus V_{\beta_1+\beta_3}$
BC_m	$m \geq 2$	$2\beta_1$	$V_{4\beta_1} \oplus V_{2\beta_2} \oplus V_{\beta_2} \oplus \mathbf{1}$	$V_{2\beta_1+\beta_2} \oplus V_{2\beta_1}$
	$m=4$	β_3	$V_{2\beta_1} \oplus V_{\beta_1+\beta_3} \oplus V_{2\beta_3}$ $\oplus W_{\beta_1+\beta_4}$	$V_{2\beta_2} \oplus V_{\beta_2} \oplus V_{\beta_2+\beta_4}$ $\oplus V_{\beta_4} \oplus W_{\beta_3} \oplus \mathbf{1}$
E_6		β_2	$V_{2\beta_2} \oplus V_{\beta_1+\beta_6} \oplus \mathbf{1}$	$V_{\beta_2} \oplus V_{\beta_4}$
E_7		β_1	$V_{2\beta_1} \oplus V_{\beta_6} \oplus \mathbf{1}$	$V_{\beta_1} \oplus V_{\beta 3}$
E_8		β_8	$V_{\beta_1} \oplus V_{2\beta_8} \oplus \mathbf{1}$	$V_{\beta_7} \oplus V_{\beta_8}$
F_4		β_1	$V_{2\beta_1} \oplus V_{2\beta_4} \oplus \mathbf{1}$	$V_{\beta_1} \oplus V_{\beta_2}$
G_2		β_2	$V_{2\beta_1} \oplus V_{2\beta_2} \oplus \mathbf{1}$	$V_{3\beta_1} \oplus V_{\beta_2}$

Table 2. All λ with $1 < \dim(V_\lambda) = \dim(\mathfrak{g})$. For the ordinary simple Lie algebras precisely the adjoint representations occur by the result of [Andreev et al. 1967].

with $r \leq m$, and $\dim(V_\lambda) = \binom{2m}{r} - \binom{2m}{r-1}$ by the description in [Djoković 1976b, Section 5]. $\square$

Corollary 4.3. *For* $\mathfrak{g} \neq \mathfrak{sl}_2(k)$, $\mathfrak{osp}_{1|2}(k)$ *and all weights* λ *one has* $\dim(V_\lambda) \geq 2$, *with equality holding only in the single case* $(\mathfrak{g}, \lambda) = (\mathfrak{osp}_{1|4}(k), \beta_2)$.

5. Proof of Theorem 4.1

Recall that $\mathfrak{g}$ admits a unique invariant nondegenerate bilinear form $(\cdot,\cdot)$ up to multiplication by a scalar [Scheunert 1979, p. 94]. Fixing any such form, we associate to any root α a coroot $\alpha^\vee = 2\alpha/(\alpha,\alpha)$. Let $\alpha_1, \dots, \alpha_m$ be a system of simple positive roots so that the fundamental weights ϖ_i satisfy $(\alpha_i^\vee, \varpi_j) = \delta_{ij}$. Then $\rho = \varpi_1 + \cdots + \varpi_m$ is half the sum of all positive roots, with the sign convention of [Tsohantjis and Cornwell 1990]. For the proof of Theorem 4.1 we consider the *index* of a representation $\varphi : \mathfrak{g} \to \mathfrak{gl}(V)$, i.e., the scalar $l(V)$ defined by $\mathrm{tr}(\varphi(X) \circ \varphi(Y)) = l(V) \cdot (X, Y)$.

Lemma 5.1. *The index has the following properties.*

(a) *For the symmetric or alternating square of a representation V it is given by the formula* $l(T_\epsilon(V)) = (\dim(V) + 2\epsilon) \cdot l(V)$.

(b) *There exists a constant* $\kappa \neq 0$ *such that* $\kappa \cdot l(V_\mu) = \dim(V_\mu) \cdot c(\mu)$ *for the scalar* $c(\mu) = (\mu, \mu) + 2(\mu, \rho) > 0$ *and for any highest weight* $\mu \neq 0$.

(c) *The index satisfies* $l(\mathbf{1}) = 0$, *and it is invariant under automorphisms and additive for direct sums in the sense that* $l(V \oplus V') = l(V) + l(V')$.

Proof. For (a) note that upon applying any tensor construction to V the index is multiplied by a constant depending only on $n = \dim(V)$. To compute this constant for $T_\epsilon(V)$, recall from [Scheunert 1979, p. 128] that $\mathfrak{sl}(V)$ is simple for $n \neq 0$. It then only remains to check that $\mathrm{tr}((T_\epsilon(X))^2) = (n + 2\epsilon)\,\mathrm{tr}(X^2)$ for a suitably chosen elementary matrix $X \in \mathfrak{sl}(V)$. For (b) one checks, by looking at the action on a highest weight vector, that the Casimir operator acts on V_μ by some fixed multiple of $c(\mu)$. The setting for $\mathfrak{osp}_{1|2m}(k)$ is described in [Djoković 1976b, p. 28; 1976a, p. 223]. One then has $\kappa = \dim(\mathrm{Ad}) \cdot c(\mathrm{Ad})$ for the adjoint representation Ad. Part (c) is obvious. □

Via these index computations, we may now complete the classification of ϵ-small representations for $\mathfrak{g} \neq \mathfrak{sl}_2(k)$, $\mathfrak{osp}_{1|2}(k)$ as follows.

Proof of Theorem 4.1. Suppose that V_λ is ϵ-small. By Corollary 4.3 we may assume that $n = \dim(V_\lambda) > 2$. Put $T_\epsilon(V_\lambda) = W \oplus \mathbf{1}^\delta$ where $\delta \in \{0, 1\}$ denotes the multiplicity with which the trivial representation enters. Note that by smallness all highest weights μ occurring in W are conjugate to each other. For any such μ Lemma 5.1(b)–(c) hence imply that $\kappa \cdot l(W) = \dim(W) \cdot c(\mu) = (n(n+\epsilon)/2 - \delta) \cdot c(\mu)$ and $\kappa \cdot l(V_\lambda) = n \cdot c(\lambda)$. So Lemma 5.1(a) shows

$$(\star) \qquad (n + 2\epsilon) \cdot n \cdot c(\lambda) = \tfrac{1}{2}\big(n(n + \epsilon) - 2\delta\big) \cdot c(\mu).$$

Now we distinguish between the symmetric and the alternating square. For $\epsilon = +1$ we may take $\mu = 2\lambda$. Then $c(\mu) = 4|\lambda|^2 + 4(\lambda, \rho)$. Since $c(\lambda) = |\lambda|^2 + 2(\lambda, \rho)$, Equation $(\star)$ easily gives

$$(n - 2\delta) \cdot |\lambda|^2 = 2(\lambda, \rho) \quad \text{and hence} \quad |\lambda| \leq \frac{2|\rho|}{n - 2\delta}$$

by the Cauchy–Schwartz inequality. Let Δ_0 be the set of simple positive roots of the even subalgebra $\mathfrak{g}_0$. Then

$$|(\lambda, \alpha^\vee)| \leq |\lambda| \cdot |\alpha^\vee| \leq \frac{2\,|\rho|\,|\alpha^\vee|}{n - 2\delta} < \frac{\dim(\mathfrak{g}) - 1}{n - 2\delta} \quad \text{for any} \quad \alpha \in \Delta_0,$$

where for the last inequality we have used the numerical values of $|\rho|^2$ and R in Table 3 and our assumption $\mathfrak{g} \neq \mathfrak{sl}_2(k)$, $\mathfrak{osp}_{1|2}(k)$. On the other hand $(\lambda, \alpha^\vee) \in \mathbb{Z}$

	$\lvert\rho\rvert^2$	R	$\dim(\mathfrak{g})$	r_i for $i = 1, \ldots, m$	$\lvert\mathrm{Out}(\mathfrak{g})\rvert$
A_m	$\frac{m(m+1)(m+2)}{12}$	$\sqrt{2}$	$m(m+2)$	$\frac{2(m+1)}{i(m+1-i)}$	2
B_m	$\frac{m(2m-1)(2m+1)}{12}$	2	$m(2m+1)$	$\frac{2}{i}(1+\delta_{im})$	1
C_m	$\frac{m(m+1)(2m+1)}{6}$	$\sqrt{2}$	$m(2m+1)$	$\frac{2}{i}(1+\delta_{im})$	1
D_m	$\frac{m(m-1)(2m-1)}{6}$	$\sqrt{2}$	$m(2m-1)$	$\frac{2}{i}$ if $i < m-1$	2 if $m \neq 4$
				$\frac{8}{m}$ if $i \in \{m-1, m\}$	6 if $m = 4$
BC_m	$\frac{m(2m-1)(2m+1)}{12}$	$\sqrt{2}$	$m(2m-1)$	$\frac{2}{i}(1+\delta_{im})$	1
E_6	78	$\sqrt{2}$	78	$\frac{3}{2}, 1, \frac{3}{5}, \frac{1}{3}, \frac{3}{5}, \frac{3}{2}$	2
E_7	$\frac{399}{2}$	$\sqrt{2}$	133	$1, \frac{4}{7}, \frac{1}{3}, \frac{1}{6}, \frac{4}{15}, \frac{1}{2}, \frac{4}{3}$	1
E_8	620	$\sqrt{2}$	248	$\frac{1}{2}, \frac{1}{4}, \frac{1}{7}, \frac{1}{15}, \frac{1}{10}, \frac{1}{6}, \frac{1}{3}, 1$	1
F_4	39	2	52	$1, \frac{1}{3}, \frac{1}{3}, 1$	1
G_2	14	$\sqrt{2}$	14	1	1

Table 3. Some numerical values. We put $r_i = \lvert\alpha_i\rvert^2/\lvert\beta_i\rvert^2$ and $R = \max_{\alpha\in\Delta_0} \lvert\alpha^\vee\rvert$ for the set Δ_0 of simple positive roots of $\mathfrak{g}_0$.

for all $\alpha \in \Delta_0$, and for $\lambda \neq 0$ at least one of these scalar products is nonzero. Thus we can find $\alpha \in \Delta_0$ with $\lvert(\lambda, \alpha^\vee)\rvert \geq 1$. This implies $n - 2\delta < \dim(\mathfrak{g}) - 1$. Hence λ is one of the highest weights in tables 1 and 2 by Lemma 4.2.

It remains to discuss the case $\epsilon = -1$. By smallness all highest weights in $\Lambda^2(V_\lambda)$ are conjugate to each other via automorphisms fixing λ. Hence Remark 5.2 below implies

$$(\star\star) \qquad \lambda = r \cdot (\beta_{i_1} + \cdots + \beta_{i_s}) \quad \text{for some } r \in \mathbb{N} \text{ and } i_1 < i_2 < \cdots < i_s,$$

and that for all $i \in \{i_1, \ldots, i_s\}$ the weight $\mu = 2\lambda - \alpha_i$ occurs as a highest weight in $\Lambda^2(V_\lambda)$. In what follows we fix $i \in \{i_1, \ldots, i_s\}$ with the smallest norm $\lvert\beta_i\rvert$. Since the norm of any simple positive root is given by the formula $\lvert\alpha_i\rvert^2 = 2\,(\alpha_i, \rho)$, we have $c(\mu)/2 = c(\lambda) + \lvert\lambda\rvert^2 - 2(\lambda, \alpha_i)$ so that $(\star)$ becomes

$$(n + 2\epsilon) \cdot n \cdot c(\lambda) = \big(n(n+\epsilon) - 2\delta\big) \cdot \big(c(\lambda) + \lvert\lambda\rvert^2 - 2(\lambda, \alpha_i)\big).$$

Now for $\epsilon = -1$ the first of the two factors on the right is $> (n + 2\epsilon) \cdot n$ since by assumption $n > 2$ and $\delta \in \{0, 1\}$. Hence

$$c(\lambda) + \lvert\lambda\rvert^2 - 2(\lambda, \alpha_i) < c(\lambda)$$

and therefore $2\,(\lambda, \alpha_i) > \lvert\lambda\rvert^2 \geq r^2 \cdot \lvert\beta_i\rvert^2 \cdot s$, where the second inequality comes from $(\star\star)$ together with the fact that all scalar products between $\beta_{i_1}, \ldots, \beta_{i_s}$ are nonnegative and β_i has the smallest norm among all these weights. On the other hand $2\,(\lambda, \alpha_i) = r \cdot 2\,(\beta_i, \alpha_i) = r \cdot |\alpha_i|^2$ by ($\star\star$). Hence $r_i := |\alpha_i|^2/|\beta_i|^2 > r \cdot s \geq 1$, which leaves only finitely many cases in view of Table 3. Note that for the Dynkin type A_m we may by duality assume $i < (m+1)/2$ so that $r_i \leq 4/i$. □

For the convenience of the reader we include a proof of the following basic fact used in the above argument; see also [Aslaksen 1994, Theorem 5].

Remark 5.2. *Let $\lambda = \sum_{i=1}^{m} a_i \beta_i$ with $a_i \in \mathbb{N}_0$. If $a_i > 0$, then the weight $2\lambda - \alpha_i$ appears as a highest weight in the alternating tensor square $\Lambda^2(V_\lambda)$.*

Proof. Let v be a highest weight vector of V_λ. For $a_i > 0$ let $X_\pm \in \mathfrak{g}_{\pm\alpha_i}$ be generators for the root spaces of the roots $\pm\alpha_i$ of $\mathfrak{g}$ and put $H = [X_+, X_-]$. It then follows from $X_+ v = 0$ that $X_+ X_- v = H v = (\alpha_i, \lambda) \cdot v \neq 0$. Since v and $X_- v$ have different weights (λ and $\lambda - \alpha_i$ respectively), this implies that $v \wedge X_- v \in \Lambda^2(V_\lambda)$ is a nonzero highest weight vector of weight $2\lambda - \alpha_i$. □

6. An application to Brill–Noether sheaves

In this independent section we briefly discuss an application of Theorem 1.1 to algebraic geometry. Let A be a complex abelian variety, and let $D(A) = D_c^b(A, \mathbb{C})$ denote the derived category of bounded constructible sheaf complexes on A in the sense of [Hotta et al. 1995]. For any sheaf complexes $K, L \in D(A)$ we may consider the exterior tensor product

$$K \boxtimes L \;=\; p_1^*(K) \otimes_{\mathbb{C}} p_2^*(L) \;\in\; D(A \times A),$$

where $p_1, p_2 : A \times A \to A$ denote the projections onto the two factors and where the tensor product on the right has to be taken in the derived sense. Passing to the direct image under the group law $a : A \times A \to A$ we then define the convolution product by

$$K * L = Ra_*(K \boxtimes L) \in D(A).$$

It has been shown in [Weissauer 2007; 2011] that with respect to this convolution product the category $D(A)$ is a rigid symmetric monoidal $\mathbb{C}$-linear category, though it is not abelian but only triangulated. Now for any perverse sheaf $K \in D(A)$ in the sense of [Hotta et al. 1995], the convolution powers of K generate an algebraic tensor category inside a certain natural symmetric monoidal quotient category $\overline{D}(A)$ of $D(A)$; see [Krämer and Weissauer 2015] for details. The Tannaka super group of this tensor category is an ordinary complex algebraic group $G(K)$ which is reductive if the perverse sheaf K is semisimple.

Now consider the special case where $A = \mathrm{Jac}(X)$ is the Albanese variety of a smooth complex projective curve X of genus $g \geq 1$. Fix an embedding $X \hookrightarrow A$, and denote by $\mathbb{C}_X$ the constant sheaf with support on the image curve. It will be more convenient to replace this constant sheaf by the sheaf complex $K = \mathbb{C}_X[1]$ placed in degree -1 since the degree shift by one leads to a complex which is a perverse sheaf. The group $G(K)$ depends on the chosen embedding $X \hookrightarrow A$, though one may show its commutator group does not. In what follows we choose the embedding so that the highest alternating convolution power $\Lambda^{*(2g-2)}(K)$ is represented in $\overline{D}(A)$ by the skyscraper sheaf $\mathbf{1}$ of rank one supported in the origin. We can achieve this via a suitable translation since by [loc. cit., Proposition 10.1] this alternating power is given in $\overline{D}(A)$ by a skyscraper sheaf of rank one. With this normalization of the embedding, the group $G_X = G(K)$ becomes an intrinsic invariant of X, and for $g > 2$ the classification in Theorem 1.1 leads to a very easy proof of the following result from [Weissauer 2007].

Theorem 6.1. *Let X be a smooth complex projective curve of genus $g \geq 1$ which is embedded into its Jacobian variety $A = \mathrm{Jac}(X)$ as above. Then*

$$G_X = \begin{cases} \mathrm{Sp}_{2g-2}(\mathbb{C}) & \text{if } X \text{ is hyperelliptic}, \\ \mathrm{SL}_{2g-2}(\mathbb{C}) & \text{otherwise}. \end{cases}$$

Proof for $g > 2$. For hyperelliptic curves X the Abel–Jacobi map $f : X^2 \to A$ is generically finite of degree two over its image, but blows down the hyperelliptic linear series g_2^1 to a point $a \in A(\mathbb{C})$. By our choice of the embedding $X \hookrightarrow A$ we can assume $a = 0$. Then one easily checks that the convolution square of the constant perverse sheaf $K = \mathbb{C}_X[1]$ has the form

$$K * K = Rf_*(\mathbb{C}_{X\times X}[2]) = \delta_+ \oplus \delta_- \oplus \mathbf{1}$$

for certain simple perverse sheaves $\delta_\pm$ and the rank one skyscraper sheaf $\mathbf{1}$ with support in the origin. The definition of the symmetry constraint in [Weissauer 2007] shows that $\mathbf{1}$ lies in the *alternating* convolution square of K. If $G = G_X$ denotes our Tannaka group and if $V \in \mathrm{Rep}_k(G)$ denotes the representation corresponding to the perverse sheaf K, it follows that the symmetric square $T_+(V)$ is irreducible and that $T_-(V)$ decomposes into an irreducible plus a trivial representation.

The ϵ-smallness of V for $\epsilon = +1$ rules out case (b) in Theorem 1.1. Case (d) is ruled out for the same reason because, by [Krämer and Weissauer 2015], the dimension of any representation of G is the Euler characteristic of the underlying perverse sheaf, which in our situation is $d = \dim_{\mathbb{C}}(V) = 2g-2 > 2$ for $g > 2$. Since $T_+(V)$ is irreducible whereas the symmetric square of the standard representation of the orthogonal group is not, case (c) is impossible. Case (e) is impossible since the group of connected components of the Tannaka group of a perverse sheaf is abelian [Weissauer 2012]. So case (a) occurs, and we look for entries in Theorem 4.1 with a $\star$ for $\epsilon = +1$ and a $\circ$ for $\epsilon = -1$. As we are dealing with ordinary groups, the only case is the standard representation of $\mathrm{Sp}_{2m}(\mathbb{C})$ where $2m = d = 2g-2$; for $g = 3$ notice $B_2 \cong C_2$. The nonhyperelliptic case is similar but here no summand $\mathbf{1}$ occurs. $\square$

References

[André and Kahn 2002] Y. André and B. Kahn, "Nilpotence, radicaux et structures monoïdales", *Rend. Sem. Mat. Univ. Padova* **108** (2002), 107–291. Correction in **113** (2005), 125–128. MR 2003m:18009 Zbl 1165.18300

[Andreev et al. 1967] E. M. Andreev, È. B. Vinberg, and A. G. Élashvili, "Орбиты наибольшей размерности полупростых линейных групп Ли", *Funkcional. Anal. i Priložen.* **1**:4 (1967), 3–7. Translated as "Orbits of greatest dimension in semi-simple linear Lie groups" in *Funct. Anal. Appl* **1**:4 (1967), 257–261. MR 42 #1942 Zbl 0176.30301

[Aschbacher 2000] M. Aschbacher, *Finite group theory*, 2nd ed., Cambridge Studies in Advanced Mathematics **10**, Cambridge University Press, 2000. MR 2001c:20001 Zbl 0997.20001

[Aslaksen 1994] H. Aslaksen, "Determining summands in tensor products of Lie algebra representations", *J. Pure Appl. Algebra* **93**:2 (1994), 135–146. MR 95b:17008 Zbl 0815.17006

[Deligne 2002] P. Deligne, "Catégories tensorielles", *Mosc. Math. J.* **2**:2 (2002), 227–248. MR 2003k:18010 Zbl 1005.18009

[Deligne et al. 1982] P. Deligne, J. S. Milne, A. Ogus, and K.-Y. Shih, *Hodge cycles, motives, and Shimura varieties*, Lecture Notes in Mathematics **900**, Springer, Berlin, 1982. MR 84m:14046 Zbl 0465.00010

[Dixon and Mortimer 1996] J. D. Dixon and B. Mortimer, *Permutation groups*, Graduate Texts in Mathematics **163**, Springer, New York, 1996. MR 98m:20003 Zbl 0951.20001

[Djoković 1976a] D. Ž. Djoković, "Classification of some 2-graded Lie algebras", *J. Pure Appl. Algebra* **7**:2 (1976), 217–230. MR 52 #10822 Zbl 0319.17005

[Djoković 1976b] D. Ž. Djoković, "Representation theory for symplectic 2-graded Lie algebras", *J. Pure Appl. Algebra* **9**:1 (1976), 25–38. MR 54 #10353 Zbl 0344.17005

[Dornhoff 1971] L. Dornhoff, *Group representation theory, A: Ordinary representation theory*, Pure and Applied Mathematics **7**, Marcel Dekker, New York, 1971. MR 50 #458a Zbl 0227.20002

[Gabber and Loeser 1996] O. Gabber and F. Loeser, "Faisceaux pervers l-adiques sur un tore", *Duke Math. J.* **83**:3 (1996), 501–606. MR 97i:14016 Zbl 0896.14009

[Guralnick and Tiep 2005] R. M. Guralnick and P. H. Tiep, "Decompositions of small tensor powers and Larsen's conjecture", *Represent. Theory* **9** (2005), 138–208. MR 2006a:20082 Zbl 1109.20040

[Hotta et al. 1995] R. Hotta, K. Takeuchi, and T. Tanisaki, *D-modules, perverse sheaves, and representation theory*, Springer, Tokyo, 1995. In Japanese; translated in Progress in Mathematics **236**, Birkhäuser, Boston, 2008. MR 2008k:32022 Zbl 1136.14009

[Kantor 1969] W. M. Kantor, "Automorphism groups of designs", *Math. Z.* **109** (1969), 246–252. MR 43 #71 Zbl 0212.03504

[Kantor 1972] W. M. Kantor, "k-Homogeneous groups", *Math. Z.* **124** (1972), 261–265. MR 46 #5422 Zbl 0232.20003

[Krämer 2014] T. Krämer, "Perverse sheaves on semiabelian varieties", *Rend. Sem. Mat. Univ. Padova* **132** (2014), 83–102. MR 3276828 Zbl 06379718

[Krämer and Weissauer 2013] T. Krämer and R. Weissauer, "On the Tannaka group attached to the theta divisor of a generic principally polarized abelian variety", 2013. To appear in *Math. Z.* arXiv 1309.3754

[Krämer and Weissauer 2015] T. Krämer and R. Weissauer, "Vanishing theorems for constructible sheaves on abelian varieties", *J. Alg. Geom.* **24**:3 (2015), 531–568.

[Magaard and Malle 1998] K. Magaard and G. Malle, "Irreducibility of alternating and symmetric squares", *Manuscripta Math.* **95**:2 (1998), 169–180. MR 99a:20011 Zbl 0919.20009

[Magaard et al. 2002] K. Magaard, G. Malle, and P. H. Tiep, "Irreducibility of tensor squares, symmetric squares and alternating squares", *Pacific J. Math.* **202**:2 (2002), 379–427. MR 2002m:20025 Zbl 1072.20013

[Rittenberg and Scheunert 1982] V. Rittenberg and M. Scheunert, "A remarkable connection between the representations of the Lie superalgebras $\mathrm{osp}(1,\ 2n)$ and the Lie algebras $\mathrm{o}(2n+1)$", *Comm. Math. Phys.* **83**:1 (1982), 1–9. MR 83c:17021 Zbl 0479.17001

[Scheunert 1979] M. Scheunert, *The theory of Lie superalgebras: an introduction*, Lecture Notes in Mathematics **716**, Springer, Berlin, 1979. MR 80i:17005 Zbl 0407.17001

[Tiep and Zalesskii 1996] P. H. Tiep and A. E. Zalesskii, "Minimal characters of the finite classical groups", *Comm. Algebra* **24**:6 (1996), 2093–2167. MR 97f:20018 Zbl 0901.20031

[Tsohantjis and Cornwell 1990] I. Tsohantjis and J. F. Cornwell, "Supercharacters and superdimensions of irreducible representations of $B(0/s)$ orthosymplectic simple Lie superalgebras", *Internat. J. Theoret. Phys.* **29**:4 (1990), 351–359. MR 91i:17007 Zbl 0702.17003

[Weissauer 2007] R. Weissauer, "Brill–Noether sheaves", 2007. arXiv math/0610923

[Weissauer 2008] R. Weissauer, "Torelli's theorem from the topological point of view", pp. 275–284 in *Modular forms on Schiermonnikoog*, edited by B. Edixhoven et al., Cambridge University Press, 2008. MR 2011g:14020 Zbl 1156.14303

[Weissauer 2009] R. Weissauer, "Semisimple algebraic tensor categories", 2009. arXiv 0909.1793

[Weissauer 2011] R. Weissauer, "A remark on rigidity of BN-sheaves", 2011. arXiv 1111.6095

[Weissauer 2012] R. Weissauer, "Why certain Tannaka groups attached to abelian varieties are almost connected", 2012. arXiv 1207.4039

[Winter 1972] D. L. Winter, "The automorphism group of an extraspecial p-group", *Rocky Mountain J. Math.* **2**:2 (1972), 159–168. MR 45 #6911 Zbl 0242.20023

Received March 5, 2014. Revised November 6, 2014.

THOMAS KRÄMER
CENTRE DE MATHÉMATIQUES LAURENT SCHWARTZ
ÉCOLE POLYTECHNIQUE
F-91128 PALAISEAU CEDEX
FRANCE

thomas.kraemer@polytechnique.edu

RAINER WEISSAUER
MATHEMATISCHES INSTITUT
RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG
IM NEUENHEIMER FELD 288
D-69120 HEIDELBERG
GERMANY

weissauer@mathi.uni-heidelberg.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 276, No. 1, 2015
dx.doi.org/10.2140/pjm.2015.276.249

ON MAXIMAL LINDENSTRAUSS SPACES

PETR PETRÁČEK AND JIŘÍ SPURNÝ

We solve a problem of Lacey (1973) by showing that there exist a metrizable compact space K and a closed space $\mathcal{H} \subset \mathcal{C}(K)$ containing constants with $\overline{\partial_{\mathcal{H}} K} = K$ such that $\mathcal{H}$ is maximal with respect to $\partial_{\mathcal{H}} K$ and $\mathcal{H}$ is not a Lindenstrauss space.

1. Introduction

Let X be a compact convex subset of a real locally convex space and let $\mathfrak{A}^c(X)$ denote the space of all affine continuous functions on X. Denote by $\operatorname{ext} X$ the set of all extreme points of X.

Let K be a compact Hausdorff topological space and $\mathcal{H} \subset \mathcal{C}(K)$ a closed subspace of $\mathcal{C}(K)$ containing constants and separating points of K. The space $\mathcal{H}$ can be identified with $\mathfrak{A}^c(X)$, where

$$X = \{s^* \in \mathcal{H}^* : s^*(1) = \|s^*\| = 1\}$$

with the weak* topology. Consider the set

$$\partial_{\mathcal{H}} K = \{x \in K : \varepsilon_x|_{\mathcal{H}} \text{ is an extreme point of the unit ball of } \mathcal{H}^*\},$$

where ε_x denotes the Dirac measure at $x \in K$. Then $\operatorname{ext} X$ is homeomorphic to $\partial_{\mathcal{H}} K$ via the evaluation mapping (see Theorem 2.1 and [LMNS 2010, Proposition 4.26]).

The space $\mathcal{H}$ is called *maximal with respect to* $\partial_{\mathcal{H}} K$ if for every closed space $\mathcal{G}$ with $\mathcal{H} \subset \mathcal{G} \subset \mathcal{C}(K)$ we have $\mathcal{H} = \mathcal{G}$ provided $\partial_{\mathcal{H}} K = \partial_{\mathcal{G}} K$.

(In [Lacey 1973], the property of separating points is not a part of the definition of a function space. Nevertheless, in our opinion, this property is necessary for $\partial_{\mathcal{H}} K$ to be homeomorphic to $\operatorname{ext} X$. Indeed, consider $\mathcal{H} = \operatorname{span}\{1\}$ on $[0, 1]$. Then X is a singleton, and thus $\partial_{\mathcal{H}}[0, 1] = [0, 1]$. Obviously, $[0, 1]$ is not homeomorphic to $\operatorname{ext} X$.)

It is shown in [Edwards and Vincent-Smith 1968] that $\mathcal{H}$ is maximal with respect to $\partial_{\mathcal{H}} K$ whenever $\mathcal{H}$ is a Lindenstrauss space; see Theorems 2.1 and 2.3 below.

The research was supported by grant GAČR P201/12/0290. Spurný was also supported by the Neuron Fund for Support of Science.
MSC2010: 46B25.
Keywords: Lindenstrauss space, L_1-predual, function space.

(A real Banach space X is called a *Lindenstrauss space*, or an L_1*-predual*, if its dual space X^* is isometric to a space $L_1(X, \mathcal{S}, \mu)$ for some measure space $(X, \mathcal{S}, \mu)$.) This result serves as a motivation for the following problem, stated as Question 5 in [Lacey 1973, p. 144] (see also [Lacey 1974, p. 198]).

Question 1.1. Let K be a compact space and $\mathcal{H} \subset \mathcal{C}(K)$ a closed subspace containing constants and separating points of K such that $\overline{\partial_{\mathcal{H}} K} = K$. Let $\mathcal{H}$ be maximal with respect to $\partial_{\mathcal{H}} K$. Is $\mathcal{H}$ then a Lindenstrauss space?

The aim of our paper is to show that the answer to Question 1.1 is in general negative by proving the following theorem.

Theorem 1.2. *There exist a metrizable compact space K and a closed space $\mathcal{H} \subset \mathcal{C}(K)$ containing constants and separating points of K with $\overline{\partial_{\mathcal{H}} K} = K$ such that $\mathcal{H}$ is maximal with respect to $\partial_{\mathcal{H}} K$ and $\mathcal{H}$ is not a Lindenstrauss space.*

2. Function spaces

Let K be a compact space (we consider all topological spaces as Hausdorff). We identify the dual of $\mathcal{C}(K)$ with the space $\mathcal{M}(K)$ of all signed Radon measures on K. By a positive Radon measure on K we mean a finite complete inner regular measure defined at least on all Borel subsets of K. Let $\mathcal{M}^1(K)$ denote the set of all probability Radon measures on K, $\mathcal{M}^+(K)$ the set of all positive Radon measures on K, and ε_x the Dirac measure at $x \in K$.

By a *function space* $\mathcal{H}$ on K we mean a subspace $\mathcal{H}$ of $\mathcal{C}(K)$ containing constants and separating points of K. Assuming $\mathcal{H}$ is a function space on K we assign to each $x \in K$ the set

$$\mathcal{M}_x(\mathcal{H}) = \{\mu \in \mathcal{M}^1(K) : \mu(h) = h(x), h \in \mathcal{H}\}$$

of all $\mathcal{H}$*-representing measures*. Clearly, $\varepsilon_x \in \mathcal{M}_x(\mathcal{H})$ for each $x \in K$. We call

$$\mathrm{Ch}_{\mathcal{H}}\, K = \{x \in K : \mathcal{M}_x(\mathcal{H}) = \{\varepsilon_x\}\}$$

the *Choquet boundary* of $\mathcal{H}$. If $h \in \mathcal{H}$ attains its strict minimum at some $x \in K$, we call h an $\mathcal{H}$*-exposing function* and x an $\mathcal{H}$*-exposed point*. It is easy to see that any $\mathcal{H}$-exposed point belongs to the Choquet boundary of $\mathcal{H}$.

We define the space $\mathcal{A}^c(\mathcal{H})$ of all continuous $\mathcal{H}$*-affine functions* to be the family of all continuous functions f on K satisfying

$$f(x) = \int_K f\, d\mu \quad \text{for each } x \in K \text{ and } \mu \in \mathcal{M}_x(\mathcal{H}).$$

$\mathcal{A}^c(\mathcal{H})$ is a closed function space containing $\mathcal{H}$ and satisfying $\mathcal{M}_x(\mathcal{H}) = \mathcal{M}_x(\mathcal{A}^c(\mathcal{H}))$ for every $x \in K$. Thus $\mathrm{Ch}_{\mathcal{H}}\, K = \mathrm{Ch}_{\mathcal{A}^c(\mathcal{H})}\, K$. We define the *state space* of $\mathcal{H}$ as

$$\boldsymbol{S}(\mathcal{H}) = \{s^* \in \mathcal{H}^* : s(1) = \|s\| = 1\}$$

endowed with the weak* topology. The state space $\boldsymbol{S}(\mathcal{H})$ is a compact convex set and K is homeomorphically embedded into $\boldsymbol{S}(\mathcal{H})$ via $\phi : K \to \boldsymbol{S}(\mathcal{H})$, where

$$\phi(x) : h \to h(x), \quad h \in \mathcal{H},\ x \in K.$$

Let $B_{\mathcal{H}^*}$ stand for the unit ball of $\mathcal{H}^*$. Following the notation in [Lacey 1973, p. 143] mentioned in the introduction,

$$\partial_{\mathcal{H}} K = \{x \in K : \phi(x) \in \operatorname{ext} B_{\mathcal{H}^*}\}.$$

The next assertion shows that our definition of the Choquet boundary coincides with Lacey's definition of $\partial_{\mathcal{H}} K$.

Theorem 2.1. *If $\mathcal{H}$ is a function space on a compact space K, then* $\operatorname{Ch}_{\mathcal{H}} K = \partial_{\mathcal{H}} K$.

Proof. By [LMNS 2010, Proposition 4.26(d)], $\phi(\operatorname{Ch}_{\mathcal{H}} K) = \operatorname{ext} \boldsymbol{S}(\mathcal{H})$. Since $\boldsymbol{S}(\mathcal{H})$ is a face of $B_{\mathcal{H}^*}$ (see [LMNS 2010, Section 2.3.A]), we have $\operatorname{ext} \boldsymbol{S}(\mathcal{H}) = \operatorname{ext} B_{\mathcal{H}^*} \cap \boldsymbol{S}(\mathcal{H})$. Thus, given any $x \in K$, we have $\phi(x) \in \operatorname{ext} \boldsymbol{S}(\mathcal{H})$ if and only if $\phi(x) \in \operatorname{ext} B_{\mathcal{H}^*}$. □

The *Choquet ordering* on $\mathcal{M}^+(K)$ is given as follows: $\mu \prec \nu$ if $\mu(k) \leq \nu(k)$ for each function k of the form $k = \max\{h_1, \dots, h_n\}$, where $n \in \mathbb{N}$ and $h_1, \dots, h_n \in \mathcal{H}$ (see [LMNS 2010, Definition 3.19 and Proposition 3.56]). A measure μ in $\mathcal{M}^+(K)$ is called *$\mathcal{H}$-maximal* if it is $\prec$-maximal. By [LMNS 2010, Theorem 3.65], there exists an $\mathcal{H}$-maximal measure $\mu \in \mathcal{M}_x(\mathcal{H})$ for every $x \in K$. Furthermore, if K is metrizable, the set $\operatorname{Ch}_{\mathcal{H}} K$ is G_δ (see [LMNS 2010, Theorem 3.42 and Proposition 3.43]) and $\mathcal{H}$-maximal measures are precisely those measures carried by $\operatorname{Ch}_{\mathcal{H}} K$ (see [LMNS 2010, Corollary 3.62]).

If for each $x \in K$ there exists only one $\mathcal{H}$-maximal measure in $\mathcal{M}_x(\mathcal{H})$, the function space $\mathcal{H}$ is called *simplicial* (see [LMNS 2010, Chapter 6]). A compact convex set X is called a *simplex* if the function space $\mathfrak{A}^c(X)$ is simplicial. The relation between simplicial function spaces and Lindenstrauss spaces is given by the following result.

Theorem 2.2. *Let $\mathcal{H}$ be a function space on a compact space K. Then $\mathcal{H}$ is simplicial if and only if the Banach space $\mathcal{A}^c(\mathcal{H})$ is a Lindenstrauss space.*

Proof. Let $\mathcal{A}^c(\mathcal{H})$ be a Lindenstrauss space. Since $\mathfrak{A}^c\big(\boldsymbol{S}(\mathcal{A}^c(\mathcal{H}))\big)$ is isometric to the space $\mathcal{A}^c(\mathcal{H})$ (see [LMNS 2010, Proposition 4.26]), it is a Lindenstrauss space as well. By [Fonf et al. 2001, Proposition 3.23], $\boldsymbol{S}(\mathcal{A}^c(\mathcal{H}))$ is a simplex. Thus it follows from [LMNS 2010, Theorem 6.54] that $\mathcal{H}$ is simplicial.

Conversely, if $\mathcal{H}$ is simplicial, $\boldsymbol{S}(\mathcal{A}^c(\mathcal{H}))$ is a simplex by [LMNS 2010, Theorem 6.54]. Using [Fonf et al. 2001, Proposition 3.23] we conclude that $\mathcal{A}^c(\mathcal{H})$, being isometric to $\mathfrak{A}^c\big(\boldsymbol{S}(\mathcal{A}^c(\mathcal{H}))\big)$, is a Lindenstrauss space. □

The next result asserts two important properties of closed function spaces that are Lindenstrauss spaces. As mentioned above, it can be considered a motivation for the question this paper aims to answer.

Theorem 2.3. *Let $\mathcal{H}$ be a closed function space on a compact space K such that $\mathcal{H}$ is a Lindenstrauss space. Then $\mathcal{H} = \mathcal{A}^c(\mathcal{H})$ and $\mathcal{H}$ is maximal with respect to* $\mathrm{Ch}_{\mathcal{H}} K$.

Proof. To prove the first assertion notice that $\mathfrak{A}^c(\boldsymbol{S}(\mathcal{H}))$, being isometric to $\mathcal{H}$ (see [LMNS 2010, Proposition 4.26]), is a Lindenstrauss space. By [Fonf et al. 2001, Proposition 3.23], $\boldsymbol{S}(\mathcal{H})$ is a simplex. This implies that $\mathfrak{A}^c(\boldsymbol{S}(\mathcal{H}))$ is simplicial and thus, by [LMNS 2010, Theorem 6.16(vi)], $\mathfrak{A}^c(\boldsymbol{S}(\mathcal{H}))$ has the so-called weak Riesz interpolation property. This, however, implies that $\mathcal{H}$ has the weak Riesz interpolation property according to [LMNS 2010, Proposition 4.26]. To finish the proof it is enough to consult [LMNS 2010, Exercise 6.78].

To prove the second assertion, let $\mathcal{G} \supset \mathcal{H}$ be a closed function space with $\mathrm{Ch}_{\mathcal{H}} K = \mathrm{Ch}_{\mathcal{G}} K$. Since $\mathcal{G} \subset \mathcal{A}^c(\mathcal{G})$ and $\mathrm{Ch}_{\mathcal{G}} K = \mathrm{Ch}_{\mathcal{A}^c(\mathcal{G})} K$, we can assume without loss of generality that $\mathcal{A}^c(\mathcal{G}) = \mathcal{G}$. Using [LMNS 2010, Theorem 10.60] we infer that $\mathcal{G} = \mathcal{A}^c(\mathcal{H})$. Since $\mathcal{H} = \mathcal{A}^c(\mathcal{H})$, we get $\mathcal{G} = \mathcal{H}$, finishing the proof. □

3. Proof of Theorem 1.2

We consider a compact subset of $\mathbb{R}^2$ defined as follows. Let $\{s, s^1, s^2, t^1, t^2\}$ be distinct points in $\mathbb{R}^2$. Let (s_n^i) and (t_n^i), $i = 0, 1, 2$, be sequences of points in $\mathbb{R}^2$ such that

- $s_n^0 \to s,\ t_n^0 \to s,$
- $s_n^i \to s^i,\ t_n^i \to t^i,\ i = 1, 2,$
- all the elements of these sequences are pairwise distinct and not contained in $\{s, s^1, s^2, t^1, t^2\}$.

Let $B(x, r)$ denote the closed ball in $\mathbb{R}^2$ with center $x \in \mathbb{R}^2$ and diameter $r > 0$. Let further $r_n > 0$, $n \in \mathbb{N}$, be numbers such that

- $r_n \to 0,$
- the family

$$\mathcal{K} = \big\{\{s\}, \{s^1\}, \{s^2\}, \{t^1\}, \{t^2\}\big\} \cup \{B(s_n^0, r_n) : n \in \mathbb{N}\} \cup \{B(t_n^0, r_n) : n \in \mathbb{N}\}$$

  is disjoint.

We define the compact space K as $K = \bigcup \mathcal{K}$. Furthermore, we set $\mathcal{H}$ to be

$$\mathcal{H} = \Big\{h \in \mathcal{C}(K) : h(s) = \tfrac{1}{2}\big(h(s^1) + h(s^2)\big) = \tfrac{1}{2}\big(h(t^1) + h(t^2)\big),$$
$$h(s_n^0) = \tfrac{1}{2}\big(h(s_n^1) + h(s_n^2)\big),\ h(t_n^0) = \tfrac{1}{2}\big(h(t_n^1) + h(t_n^2)\big),\ n \in \mathbb{N}\Big\}.$$

Lemma 3.1. *The space $\mathcal{H}$ is a well defined function space with $\mathcal{H} = \mathcal{A}^c(\mathcal{H})$. Let $L = \{s\} \cup \{s_n^0 : n \in \mathbb{N}\} \cup \{t_n^0 : n \in \mathbb{N}\}$. Then* $\operatorname{Ch}_{\mathcal{H}} K = K \setminus L$. *In particular*, $\operatorname{Ch}_{\mathcal{H}} K$ *is dense in K.*

Proof. Obviously, $\mathcal{H}$ contains constant functions. The fact that $\mathcal{H} = \mathcal{A}^c(\mathcal{H})$ follows immediately from the definition of $\mathcal{H}$. To verify that $\mathcal{H}$ separates points of K it is enough to consider elementary constructions of functions from $\mathcal{H}$. Given $n \in \mathbb{N}$ and $z \in B(s_n^0, r_n) \setminus \{s_n^0\}$, we consider a continuous function $g : B(s_n^0, r_n) \to [0, 1]$ attaining 0 precisely at z and 1 at s_n^0. Then the function

$$h_z(x) = \begin{cases} g(x) & \text{if } x \in B(s_n^0, r_n), \\ 1 & \text{otherwise} \end{cases}$$

separates z from the remaining points of K. It also $\mathcal{H}$-exposes z, and thus $z \in \operatorname{Ch}_{\mathcal{H}} K$.

We can further construct functions h_{s_n} and h_s in $\mathcal{H}$ as follows:

$$h_{s_n}(x) = \begin{cases} 0 & \text{if } x = s_n^1, \\ 2 & \text{if } x \in B(s_n^0, r_n), \\ 4 & \text{if } x = s_n^2, \\ 1 & \text{otherwise,} \end{cases}$$

$$h_s(x) = \begin{cases} 0 & \text{if } x = s^1, \\ 2 & \text{if } x = s^2, \\ \frac{1}{2n} & \text{if } x = s_n^1, \\ 2 - \frac{1}{2n} & \text{if } x = s_n^2, \\ 1 & \text{otherwise.} \end{cases}$$

The function h_{s_n} then separates the points s_n^1, s_n^2 from any point in K and it separates s_n^0 from any point in $K \setminus B(s_n^0, r_n)$. Its construction also shows that the points s_n^1, s_n^2 are $\mathcal{H}$-exposed and thus lie in $\operatorname{Ch}_{\mathcal{H}} K$. Similarly, the function h_s separates points s^1, s, s^2 from each other and it separates s from every point in $\{s_n^i : n \in \mathbb{N}, i \in \{1, 2\}\}$. Furthermore, the construction of h_s shows that the points s^1, s^2 are $\mathcal{H}$-exposed and thus belong to $\operatorname{Ch}_{\mathcal{H}} K$.

Analogously we can construct functions h_{t_n}, $\tilde{h}_s$ and h_y for any $n \in \mathbb{N}$ and $y \in B(t_n^0, r_n) \setminus \{t_n^0\}$ to show that $\mathcal{H}$ indeed separates points of K and that all points in

$$\{t_1, t_2\} \cup \{t_n^i : n \in \mathbb{N}, i \in \{1, 2\}\} \cup \bigcup_{n \in \mathbb{N}} (B(t_n^0, r_n) \setminus \{t_n^0\})$$

lie in $\operatorname{Ch}_{\mathcal{H}} K$.

Overall, we have

$$\{s^1, s^2, t^1, t^2\} \cup \{s_n^i : n \in \mathbb{N}, i \in \{1, 2\}\} \cup \{t_n^i : n \in \mathbb{N}, i \in \{1, 2\}\} \subset \operatorname{Ch}_{\mathcal{H}} K$$

and

$$\bigcup_{n \in \mathbb{N}} (B(s_n^0, r_n) \setminus \{s_n^0\}) \cup \bigcup_{n \in \mathbb{N}} (B(t_n^0, r_n) \setminus \{t_n^0\}) \subset \operatorname{Ch}_{\mathcal{H}} K.$$

Clearly, any point in L has a nontrivial $\mathcal{H}$-representing measure. This together with the inclusions above yields $\operatorname{Ch}_{\mathcal{H}} K = K \setminus L$. ☐

Lemma 3.2. *Let $n \in \mathbb{N}$. Then*

$$\mathcal{M}_{s_n^0}(\mathcal{H}) = \operatorname{conv}\left\{\varepsilon_{s_n^0}, \tfrac{1}{2}(\varepsilon_{s_n^1} + \varepsilon_{s_n^2})\right\}$$

and

$$\mathcal{M}_{t_n^0}(\mathcal{H}) = \operatorname{conv}\left\{\varepsilon_{t_n^0}, \tfrac{1}{2}(\varepsilon_{t_n^1} + \varepsilon_{t_n^2})\right\}.$$

Proof. Let $n \in \mathbb{N}$ and $\mu \in \mathcal{M}_{s_n^0}(\mathcal{H})$ be fixed. Pick a continuous function

$$g : B(s_n^0, r_n) \to [0, 1]$$

such that $g(s_n^0) = 0$ and $g(x) > 0$ otherwise. Using the function

$$h(x) = \begin{cases} 0 & \text{if } x \in \{s_n^1, s_n^2\}, \\ g(x) & \text{if } x \in B(s_n^0, r_n), \\ 1 & \text{otherwise,} \end{cases}$$

we infer that the support of μ is contained in $\{s_n^0, s_n^1, s_n^2\}$.

Further, let $a = \mu(\{s_n^0\})$. Assume first that $a = 0$, i.e., $\mu = b\varepsilon_{s_n^1} + (1 - b)\varepsilon_{s_n^2}$ for some $b \in [0, 1]$. Then the function

$$h(x) = \begin{cases} 0 & \text{if } x = s_n^1, \\ 1 & \text{if } x = s_n^0, \\ 2 & \text{if } x = s_n^2, \\ 1 & \text{otherwise} \end{cases}$$

shows that

$$1 = h(s_n^0) = \mu(h) = (1 - b)h(s_n^2) = (1 - b)2.$$

In other words, $b = \frac{1}{2}$ and $\mu = \frac{1}{2}(\varepsilon_{s_n^1} + \varepsilon_{s_n^2})$.

If $a \in (0, 1)$, then the measure $\nu = \mu - a\varepsilon_{s_n^0}$ satisfies

$$h(s_n^0) = \mu(h) = \nu(h) + ah(s_n^0), \quad h \in \mathcal{H}.$$

Hence $\dfrac{1}{1-a}\nu$ is in $\mathcal{M}_{s_n^0}(\mathcal{H})$ and is carried by $\{s_n^1, s_n^2\}$. By the first part of the proof,

$$\frac{1}{1-a}\nu = \tfrac{1}{2}(\varepsilon_{s_n^1} + \varepsilon_{s_n^2}).$$

Thus

$$\mu = \nu + a\varepsilon_{s_n^0} = (1-a)\frac{1}{1-a}\nu + a\varepsilon_{s_n^0} = (1-a)\tfrac{1}{2}(\varepsilon_{s_n^1} + \varepsilon_{s_n^2}) + a\varepsilon_{s_n^0}$$

is in $\operatorname{conv}\bigl\{\varepsilon_{s_n^0}, \tfrac{1}{2}(\varepsilon_{s_n^1} + \varepsilon_{s_n^2})\bigr\}$. If $a = 1$, obviously

$$\mu = \varepsilon_{s_n^0} \in \operatorname{conv}\bigl\{\varepsilon_{s_n^0}, \tfrac{1}{2}(\varepsilon_{s_n^1} + \varepsilon_{s_n^2})\bigr\}.$$

Thus

$$\mu \in \operatorname{conv}\bigl\{\varepsilon_{s_n^0}, \tfrac{1}{2}(\varepsilon_{s_n^1} + \varepsilon_{s_n^2})\bigr\}$$

holds in all cases.

The second part of the assertion can be proved analogously. □

Lemma 3.3. *The space $\mathcal{H}$ is not simplicial.*

Proof. The measures $\frac{1}{2}(\varepsilon_{s^1} + \varepsilon_{s^2})$, $\frac{1}{2}(\varepsilon_{t^1} + \varepsilon_{t^2})$ are different, they $\mathcal{H}$-represent s and, by Lemma 3.1, both are carried by $\operatorname{Ch}_{\mathcal{H}} K$. Hence there exist two $\mathcal{H}$-maximal measures representing s, which implies that $\mathcal{H}$ is not simplicial. □

Lemma 3.4. *The space $\mathcal{H}$ is maximal with respect to* $\operatorname{Ch}_{\mathcal{H}} K$. *That is*, $\mathcal{G} = \mathcal{H}$ *for any closed function space $\mathcal{H} \subset \mathcal{G}$ such that* $\operatorname{Ch}_{\mathcal{G}} K = \operatorname{Ch}_{\mathcal{H}} K$.

Proof. Fix an index $m \in \mathbb{N}$. Let $\tau \in \mathcal{M}_{s_m^0}(\mathcal{G})$ be a measure carried by $\operatorname{Ch}_{\mathcal{G}} K$. We aim to show that

$$\tau = \tfrac{1}{2}(\varepsilon_{s_m^1} + \varepsilon_{s_m^2}). \tag{3-1}$$

Since $\mathcal{M}_{s_m^0}(\mathcal{G}) \subset \mathcal{M}_{s_m^0}(\mathcal{H})$, we obtain by virtue of Lemma 3.2 that

$$\tau \in \operatorname{conv}\bigl\{\varepsilon_{s_m^0}, \tfrac{1}{2}(\varepsilon_{s_m^1} + \varepsilon_{s_m^2})\bigr\}.$$

This and the fact that τ is carried by $\operatorname{Ch}_{\mathcal{G}} K = \operatorname{Ch}_{\mathcal{H}} K \subset K \setminus \{s_m^0\}$ imply (3-1).

Pick $\mu_n \in \mathcal{M}_{s_n^0}(\mathcal{G})$, $n \in \mathbb{N}$, such that the measures μ_n are carried by $\operatorname{Ch}_{\mathcal{G}} K$ for all $n \in \mathbb{N}$. The sequence (s_n^0) converges to s, while the sequence (μ_n) converges to $\mu = \frac{1}{2}(\varepsilon_{s^1} + \varepsilon_{s^2})$. Thus $\mu \in \mathcal{M}_s(\mathcal{G})$. Analogously we infer that any measure ν_n in $\mathcal{M}_{t_n^0}(\mathcal{G})$ carried by $\operatorname{Ch}_{\mathcal{G}} K$ satisfies $\nu_n = \frac{1}{2}(\varepsilon_{t_n^1} + \varepsilon_{t_n^2})$, and thus $\nu = \frac{1}{2}(\varepsilon_{t^1} + \varepsilon_{t^2})$ is in $\mathcal{M}_s(\mathcal{G})$.

We want to show that $\mathcal{G} \subset \mathcal{H}$. To this end, let $g \in \mathcal{G}$ be given. We have to verify the conditions defining the space $\mathcal{H}$. Using the arguments above we get

$$g(s_n^0) = \mu_n(g) = \tfrac{1}{2}(g(s_n^1) + g(s_n^2)) \quad \text{and} \quad g(s) = \mu(g) = \tfrac{1}{2}(g(s^1) + g(s^2)),$$

while simultaneously

$$g(t_n^0) = \nu_n(g) = \tfrac{1}{2}(g(t_n^1) + g(t_n^2)) \quad \text{and} \quad g(s) = \nu(g) = \tfrac{1}{2}(g(t^1) + g(t^2)).$$

Hence $g \in \mathcal{H}$ by definition. This concludes the proof. □

Thus we have proved Theorem 1.2. Indeed, considering the compact space K and the closed function space $\mathcal{H} \subset \mathcal{C}(K)$ defined above, we have by Lemma 3.1 that $\operatorname{Ch}_{\mathcal{H}} K$ is dense in K. Furthermore, $\mathcal{H}$ is maximal with respect to $\operatorname{Ch}_{\mathcal{H}} K$ by Lemma 3.4. Since $\mathcal{H}$ is not simplicial according to Lemma 3.3, Theorem 2.2 asserts that $\mathcal{A}^c(\mathcal{H})$ is not a Lindenstrauss space. Since $\mathcal{H} = \mathcal{A}^c(\mathcal{H})$ by Lemma 3.1, it follows that $\mathcal{H}$ is not a Lindenstrauss space.

References

[Edwards and Vincent-Smith 1968] D. A. Edwards and G. Vincent-Smith, "A Weierstrass–Stone theorem for Choquet simplexes", *Ann. Inst. Fourier* (*Grenoble*) **18**:1 (1968), 261–282. MR 39 #6060 Zbl 0172.15604

[Fonf et al. 2001] V. P. Fonf, J. Lindenstrauss, and R. R. Phelps, "Infinite dimensional convexity", pp. 599–670 in *Handbook of the geometry of Banach spaces*, vol. I, edited by W. B. Johnson and J. Lindenstrauss, North-Holland, Amsterdam, 2001. MR 2003c:46014 Zbl 1086.46004

[Lacey 1973] H. E. Lacey, "On the classification of Lindenstrauss spaces", *Pacific J. Math.* **47** (1973), 139–145. MR 50 #5443 Zbl 0251.46030

[Lacey 1974] H. E. Lacey, *The isometric theory of classical Banach spaces*, Die Grundlehren der mathematischen Wissenschaften **208**, Springer, New York, 1974. MR 58 #12308 Zbl 0285.46024

[LMNS 2010] J. Lukeš, J. Malý, I. Netuka, and J. Spurný, *Integral representation theory: applications to convexity, Banach spaces and potential theory*, de Gruyter Studies in Mathematics **35**, de Gruyter, Berlin, 2010. MR 2011e:46002 Zbl 1216.46003

Received May 5, 2014.

PETR PETRÁČEK
DEPARTMENT OF MATHEMATICAL ANALYSIS
FACULTY OF MATHEMATICS AND PHYSICS
CHARLES UNIVERSITY
SOKOLOVSKÁ 83
186 75 PRAHA 8
CZECH REPUBLIC
petracek@karlin.mff.cuni.cz

JIŘÍ SPURNÝ
DEPARTMENT OF MATHEMATICAL ANALYSIS
FACULTY OF MATHEMATICS AND PHYSICS
CHARLES UNIVERSITY
SOKOLOVSKÁ 83
186 75 PRAHA 8
CZECH REPUBLIC
spurny@karlin.mff.cuni.cz

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 276 No. 1 July 2015

0030-8730(201507)276:1;1-Y