Pacific Journal of Mathematics

Volume 227 No. 1 September 2006

PACIFIC JOURNAL OF MATHEMATICS

http://www.pjmath.org

Founded in 1951 by

E. F. Beckenbach (1906–1982) F. Wolf (1904–1989)

See inside back cover or www.pjmath.org for submission instructions.

Regular subscription rate for 2006: $425.00 a year (10 issues). Special rate: $212.50 a year to individual members of supporting institutions. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840 is published monthly except July and August. 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.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS
at the University of California, Berkeley 94720-3840
A NON-PROFIT CORPORATION
Typeset in LaTeX

PACIFIC JOURNAL OF MATHEMATICS
Vol. 227, No. 1, 2006

ITERATED LOOP ALGEBRAS

BRUCE ALLISON, STEPHEN BERMAN AND ARTURO PIANZOLA

Iterated loop algebras are by definition obtained by repeatedly applying the loop construction, familiar from the theory of affine Kac–Moody Lie algebras, to a given base algebra. Our interest in this iterated construction is motivated by its use in the realization of extended affine Lie algebras, but the construction also appears naturally in the study of other classes of algebras. This paper consists of a detailed study of the basic properties of iterated loop algebras.

1. Introduction

Over the past 35 years affine Kac–Moody Lie algebras have been at the centre of a considerable amount of beautiful mathematics and theoretical physics. As of late, and perhaps influenced by some of the newest theories in physics, the need seems to have arisen for some "higher nullity" generalizations of affine Kac–Moody Lie algebras. It is still too early to decide what the correct final choice for these algebras will be, but it is fair to say notwithstanding, that Lie algebras graded by root systems and extended affine Lie algebras (EALAs) will play a prominent role in the process [Berman and Moody 1992; Benkart and Zelmanov 1996; Allison et al. 1997a; Saito and Yoshii 2000].

Recall that given a $\mathbb{Z}_m$-grading $\Sigma = \{\mathcal{A}_{\bar{i}}\}_{\bar{i}\in\mathbb{Z}_m}$ of an algebra $\mathcal{A}$ over a field k, the *loop algebra* of Σ based on $\mathcal{A}$ is the subalgebra

$$L(\mathcal{A}, \Sigma) := \bigoplus_{i\in\mathbb{Z}} \mathcal{A}_{\bar{i}} \otimes_k z^i$$

of $\mathcal{A}\otimes_k k[z, z^{-1}]$. Using this beautiful construction, V. Kac showed that (the derived algebra modulo its centre of) any complex affine Kac–Moody Lie algebras can be obtained as a loop algebra of a finite dimensional simple Lie algebra [Kac 1969]. The loop construction makes it clear, among other things, that the affine algebras

MSC2000: primary 17B65; secondary 17B67, 16S99, 17C99, 17D05, 17A01.
Keywords: loop algebra, Lie algebra, associative algebra, Jordan algebra.
The authors gratefully acknowledge the support of the Natural Sciences and Engineering Research Council of Canada.

are objects of nullity one in a sense that can be made precise. Indeed, in EALA theory, where the concept of nullity is well-defined, one finds that finite dimensional simple algebras are precisely the (tame) EALAs of nullity zero whereas affine algebras are precisely the (tame) EALAs of nullity one [Allison et al. 1997b].

It thus seems almost inevitable to ask whether, starting from an affine Kac–Moody Lie algebra and applying the loop construction, one obtains an extended affine Lie algebra of nullity 2. This and related questions have been investigated in some detail in [Wakimoto 1985; Pollmann 1994; Allison et al. 2002; Allison et al. 2004; van de Leur 2001]. In our work on this topic, as well as in [van de Leur 2001], it became clear that some advantages are to be had by thinking of loop algebras based on an affine algebra as being obtained from a finite dimensional Lie algebra by applying the loop construction *twice* (the advantages stemming from the fact that in this case the "bottom" algebra, namely the finite dimensional one, is much simpler than the affine algebra). As the reader will have surmised by now, the study of these "iterated loop algebras" took on a life of its own and became the subject of the present paper.

In general, if $\mathscr{A}$ is an (arbitrary) algebra over k, an *n-step iterated loop algebra* based on $\mathscr{A}$ is an algebra that can be obtained starting from $\mathscr{A}$ by a sequence of n loop constructions, each based on the algebra obtained at the previous step (see Definition 5.1). Far from being a mere generalization of the loop construction, iterated loop algebras seem to yield interesting mathematical objects in a natural way. Even when the resulting objects are known, the new point of view can be illuminating. As an example, we see in Example 9.8 that algebras representing elements of the Brauer group of the ring $k[t_1^{\pm 1}, t_2^{\pm 2}]$ are obtained as 2-step iterated loop algebras of $M_n(k)$. This information is not apparent if one thinks in terms of single loop algebras of $M_\ell(k[t_1^{\pm 1}])$.

This paper contains a detailed study of the basic properties of iterated loop algebras. We begin in Section 2 by recording some simple properties of the centroid of an algebra. In the rest of Section 2 and in Section 3 we define and give the basic properties of a very important class of algebras which for lack of a better name we have simply referred to as pfgc algebras (nonzero, perfect, and finitely generated as modules over their centroids). The property of being a prime pfgc algebra arises naturally in the study of iterated loop algebras since this property is carried over to a loop algebra (and hence to an iterated loop algebra) from its base. In contrast the property of finite dimensional central simplicity certainly does not carry over in the same way. After this discussion of pfgc algebras we establish in Section 4 some basic properties of (one step) loop algebras.

The main results of the paper appear in Sections 5, 6, 7 and 8 . These all deal with properties of an n-step iterated loop algebra $\mathscr{L}$ based on a pfgc algebra $\mathscr{A}$. First Theorem 5.5 establishes a long list of properties that carry over from $\mathscr{A}$ to $\mathscr{L}$.

In particular, it is shown (as mentioned above) that if $\mathscr{A}$ is a prime pfgc algebra then so is $\mathscr{L}$. Next Theorem 6.2 shows that the centroid $C(\mathscr{L})$ of $\mathscr{L}$ is itself an n-step iterated loop algebra of the centroid of $\mathscr{A}$. The same theorem describes a method of calculating $C(\mathscr{L})$ explicitly. Then Theorem 7.1 shows that $\mathscr{L}$ can be "untwisted" by a base ring extension of $C(\mathscr{L})$ that is free of finite rank. That is, the algebra $\mathscr{L}$ (after such a base ring extension) becomes isomorphic to the iterated loop algebra obtained using only the trivial gradings at each stage. Finally, Section 8 deals with the concept of type of an algebra (which is motivated by the concept of type in terms of root systems which exists in Lie theory). The main result, Theorem 8.16, states that type cannot change under the loop construction.

Each of the main results in Sections 6, 7 and 8 has several corollaries that are discussed in the respective sections. To give one important example, we show in Section 8 that if $\mathscr{L}$ is an n-step iterated loop algebra based on a finite dimensional split simple Lie algebra $\mathscr{A}$ over a field of characteristic 0 then both $\mathscr{A}$ and n are isomorphism invariants of $\mathscr{L}$ (see Corollary 8.19). This result will play a crucial role in our forthcoming work on the classification of the centreless cores of EALAs of nullity 2 [Allison et al. $\geq$ 2006].

In the last section, Section 9, we look closely at 2-step iterated loop algebras. If the base algebra is finite dimensional and central simple, these 2-step iterated loop algebras come in two kinds, depending on the structure of their centroids. We illustrate this fact along with many of the concepts discussed in the paper by describing two examples dealing respectively with Lie algebras and associative algebras.

2. Centroids and pfgc algebras

We record here some basic facts about centroids, and we define a class of algebras, which we call pfgc algebras, that will play an important role in the study of loop algebras. A good basic reference on the centroid is [Jacobson 1962, Ch. X, § 1].

Terminology and notation. A *ring* will mean a unital commutative associative ring. Homomorphisms, subrings and modules for rings will always be assumed to be unital.

A *base change* will mean a homomorphism $\upsilon : B \to B'$ of rings. This base change is said to be *free* (respectively *flat, faithfully flat*) if B' is a free (respectively flat, faithfully flat) B-module. Note that if $\upsilon : B \to B'$ is free and $B' \neq 0$, then υ is faithfully flat and hence flat [Bourbaki 1972, § I.3.1, Example 2]. An injective base change $\upsilon : B \to B'$ will be called an *extension of rings*, in which case we often identify B as a subring of B' and denote the extension by B'/B.

If B is a ring, an *algebra* over B will mean a B-module $\mathscr{A}$ together with a B-bilinear product (which is not necessarily associative, commutative or unital). If

$\mathcal{A}$ and $\mathcal{A}'$ are B-algebras, we use the notation

$$\mathcal{A} \simeq_B \mathcal{A}'$$

to mean that $\mathcal{A}$ and $\mathcal{A}'$ are isomorphic as B-algebras. If $\mathcal{A}$ is an algebra over B and $\upsilon : B \to B'$ is a base change, we will denote by $\mathcal{A} \otimes_B B'$ the (unique) B'-algebra which is obtained from $\mathcal{A}$ by base change [Bourbaki 1974, Ch. III, § 1.5].

For the rest of the section we assume that B is a ring, and that $\mathcal{A}$ is a B-algebra. Note that $\mathcal{A}$ can also be regarded as $\mathbb{Z}$-algebra under the natural action of $\mathbb{Z}$ on $\mathcal{A}$.

We now recall the definition of the centroid of $\mathcal{A}$ [Jacobson 1962, Ch. X, § 1].

Definition 2.1. (i) For $a \in \mathcal{A}$ consider the two maps from $\mathcal{A}$ to $\mathcal{A}$

$$a_L : x \mapsto ax \quad \text{and} \quad a_R : x \mapsto xa.$$

The *multiplication algebra* of $\mathcal{A}$ [Jacobson 1962, Ch. X, § 1] is defined to be the B-subalgebra $\mathrm{Mult}_B(\mathcal{A})$ of $\mathrm{End}_B(\mathcal{A})$ generated by $\{1\} \cup \{a_L \mid a \in \mathcal{A}\} \cup \{a_R \mid a \in \mathcal{A}\}$.

(ii) The set $C_B(\mathcal{A})$ of elements of $\mathrm{End}_B(\mathcal{A})$ that commute with the action of $\mathrm{Mult}_B(\mathcal{A})$ is called the *centroid* of $\mathcal{A}$. Equivalently

$$C_B(\mathcal{A}) := \{\chi \in \mathrm{End}_B(\mathcal{A}) : \chi(xy) = \chi(x)y = x\chi(y) \quad \text{for all} \quad x, y \in \mathcal{A}\}.$$

(The notation $\mathrm{Cent}_B(\mathcal{A})$ has been used for the centroid in some articles, for example in [Allison et al. 2004]. We are using the abbreviated notation $C_B(\mathcal{A})$ since it will arise frequently.) Clearly $C_B(\mathcal{A})$ is a B-subalgebra of $\mathrm{End}_B(\mathcal{A})$, and therefore $\mathcal{A}$ can be viewed in a natural way as a left $C_B(\mathcal{A})$-module by defining $\chi \cdot x = \chi(x)$.

(iii) For $b \in B$ we define $\lambda_{\mathcal{A}}(b) \in \mathrm{End}_B(\mathcal{A})$ by

$$\big(\lambda_{\mathcal{A}}(b)\big)(x) = b \cdot x.$$

Clearly $\lambda_{\mathcal{A}}(b) \in C_B(\mathcal{A})$ since $\mathcal{A}$ is a B-algebra. Then the map $\lambda_{\mathcal{A}} : B \to C_B(\mathcal{A})$ is a ring homomorphism, and $C_B(\mathcal{A})$ *is a unital associative B-algebra via this map*. Furthermore, if $\mathcal{A}$ is a faithful B-module then B can be identified with a subring of (the centre of) the centroid $C_B(\mathcal{A})$.

(iv) The B-algebra $\mathcal{A}$ is said to be *central* (or central over B) if $\lambda_{\mathcal{A}} : B \to C_B(\mathcal{A})$ is an isomorphism.

(v) The *centre* of $\mathcal{A}$ is defined to be the set $Z(\mathcal{A})$ of elements in $\mathcal{A}$ that commute and associate with all elements of $\mathcal{A}$. Then $Z(\mathcal{A})$ is a B-subalgebra of $\mathcal{A}$. If $\mathcal{A}$ is unital, the map which sends z to left multiplication by z is a B-algebra isomorphism of $Z(\mathcal{A})$ onto $C_B(\mathcal{A})$ [Erickson et al. 1975, § 1].

The following is clear:

Lemma 2.2. *Suppose that $\mathscr{A}$ and $\mathscr{A}'$ are B-algebras and $\rho : \mathscr{A} \to \mathscr{A}'$ is a B-algebra isomorphism. Then ρ induces a B-algebra isomorphism $C_B(\rho) : C_B(\mathscr{A}) \to C_B(\mathscr{A}')$ defined by $\chi \mapsto \rho\chi\rho^{-1}$.*

The formation of the centroid does not commute with base change. Nonetheless these two processes do commute in two important cases that we now describe. If $B \to B'$ is a homomorphism of rings, we define

$$\nu = \nu_{\mathscr{A},B,B'} : C_B(\mathscr{A}) \otimes_B B' \to C_{B'}(\mathscr{A} \otimes_B B')$$

to be the restriction of the canonical map $\mathrm{End}_B(\mathscr{A}) \otimes_B B' \to \mathrm{End}_{B'}(\mathscr{A} \otimes_B B')$. Then ν is a homomorphism, said to be canonical, of unital associative B'-algebras.

Lemma 2.3. *Suppose that $B \to B'$ is a homomorphism of rings. Then the map $\nu_{\mathscr{A},B,B'} : C_B(\mathscr{A}) \otimes_B B' \to C_{B'}(\mathscr{A} \otimes_B B')$ is an isomorphism of B'-algebras in the following cases*:

(a) *$\mathscr{A}$ is finitely generated as a module over its multiplication algebra* $\mathrm{Mult}_B(\mathscr{A})$ *and B' is a free B-module.*

(b) *B' is a finitely generated projective B-module.*

Proof. (a) Let $\{s_i\}_{i \in I}$ be a basis of the B-module B'.

It is clear that ν is injective. Indeed if $\sum \chi_i \otimes s_i$ is in the kernel of ν then $\sum \chi_i(x) \otimes s_i = 0$ for all x in $\mathscr{A}$ and so $\chi_i = 0$ for all i in I.

To see that ν is onto, let $\chi \in C_{B'}(\mathscr{A} \otimes_B B')$. Then for $x \in \mathscr{A}$ we can write $\chi(x \otimes_B 1_{B'})$ uniquely as

$$\chi(x \otimes 1_{B'}) = \sum \chi_i(x) \otimes s_i,$$

where $\chi_i(x) \in \mathscr{A}$ and only finitely many of these are nonzero. It is easy to see that for all $i \in I$ the map $\chi_i : \mathscr{A} \to \mathscr{A}$ given by $\chi_i : x \mapsto \chi_i(x)$ is an element of $C_B(\mathscr{A})$. Thus to see that χ is an image under ν it suffices to show that only finitely many of the maps χ_i are nonzero. For this let $\{x_1, \dots, x_n\}$ be a set of generators of $\mathscr{A}$ as a $\mathrm{Mult}_B(\mathscr{A})$-module. Then whenever χ_i vanishes on all x_j's we have

$$\chi_i(\mathscr{A}) = \chi_i\Big(\sum_{j=1}^{n} \mathrm{Mult}_B(\mathscr{A}) \cdot x_j\Big) = \sum_{j=1}^{n} \mathrm{Mult}_B(\mathscr{A}) \cdot \chi_i(x_j) = 0.$$

(b) Consider the unique B-module homomorphism

$$\varphi_{B,\mathscr{A}} : \mathrm{End}_B(\mathscr{A}) \to \mathrm{Hom}_B(\mathscr{A} \otimes_B \mathscr{A}, \mathscr{A} \oplus \mathscr{A})$$

satisfying

$$\varphi_{B,\mathscr{A}}(f)(a_1 \otimes_B a_2) = \big(f(a_1a_2) - f(a_1)a_2,\, f(a_1a_2) - a_1 f(a_2)\big).$$

By definition

$$\ker(\varphi_{B,\mathscr{A}}) = C_B(\mathscr{A}).$$

Also, by standard properties of projective modules we obtain the diagram

$$\begin{array}{ccccccc} 0 \to C_B(\mathscr{A})\otimes_B B' & \to & \mathrm{End}_B(\mathscr{A})\otimes_B B' & \to & \mathrm{Hom}_B(\mathscr{A}\otimes_B\mathscr{A},\, \mathscr{A}\oplus\mathscr{A})\otimes_B B' \\ \downarrow \nu & & \| & & \| \\ 0 \to C_{B'}(\mathscr{A}\otimes_B B') & \to & \mathrm{End}_{B'}(\mathscr{A}\otimes_B B') & \to & \mathrm{Hom}_{B'}\big((\mathscr{A}\otimes_B B')\otimes_{B'}(\mathscr{A}\otimes_B B'),\, \mathscr{A}\otimes_B B'\oplus\mathscr{A}\otimes_B B'\big) \end{array}$$

where the horizontal rows are exact. Indeed the exactness of the top row is by flatness of the B-module B' (every projective is flat). The two vertical isomorphisms come from B' being a finitely generated B-module which is projective [Bourbaki 1974, Ch. II, § 5.3, Prop. 7]. It follows that ν is an isomorphism. □

The following important fact is proved in [Jacobson 1962, Ch. X, § 1, Theorem 3]:

Lemma 2.4. *Suppose that B is a field and $\mathscr{A}$ is finite dimensional and central simple over B. If B'/B is a field extension, then $\mathscr{A}\otimes_B B'$ is finite dimensional and central simple over B'.*

Next we consider gradings on $C_B(A)$ that are induced by gradings on $\mathscr{A}$. For this suppose that $\mathscr{A}$ is Q-graded algebra over B where Q is a finite abelian group. Thus

$$\mathscr{A} = \bigoplus_{\alpha\in Q} \mathscr{A}_\alpha$$

for some B-submodules $\mathscr{A}_\alpha$ and $\mathscr{A}_\alpha\mathscr{A}_\beta \subset \mathscr{A}_{\alpha+\beta}$. Then, since Q is finite,

$$\mathrm{End}_B(\mathscr{A}) = \bigoplus_{\lambda\in Q} \mathrm{End}_B(\mathscr{A})_\lambda$$

is also a Q-graded B-algebra, where

$$\mathrm{End}_B(\mathscr{A})_\lambda = \{\theta \in \mathrm{End}_B(\mathscr{A}) \mid \theta(\mathscr{A}_\alpha) \subset \mathscr{A}_{\lambda+\alpha} \text{ for all } \alpha \in Q\}.$$

It is easy to check that $C_B(\mathscr{A})$ is a Q-graded B-subalgebra of $\mathrm{End}_B(\mathscr{A})$, and so we have:

Lemma 2.5. *Suppose that $\mathscr{A}$ is Q-graded algebra over B, where Q is a finite abelian group. Then*

$$C_B(\mathscr{A}) = \bigoplus_{\alpha\in Q} C_B(\mathscr{A})_\lambda$$

is a Q-graded algebra over B, where $C_B(\mathscr{A})_\lambda = C_B(\mathscr{A}) \cap \mathrm{End}_B(\mathscr{A})_\lambda$ for all $\lambda \in Q$.

Definition 2.6. If $\mathscr{I}$ and $\mathscr{J}$ are ideals of the B-algebra $\mathscr{A}$ we define

$$\mathscr{I}\mathscr{J} = \left\{ \sum x_i y_i : x_i \in \mathscr{I},\ y_i \in \mathscr{J} \right\}$$

(finite sums of course). Note that in general $\mathscr{I}\mathscr{J}$ is not an ideal of $\mathscr{A}$. We say that $\mathscr{A}$ is *perfect* if $\mathscr{A}\mathscr{A} = \mathscr{A}$.

Remark 2.7. It is clear that $\mathscr{A}$ is perfect as a B-algebra if and only if $\mathscr{A}$ is perfect as a $\mathbb{Z}$-algebra.

Lemma 2.8. *Assume $\mathscr{A}$ is perfect. Then*

(i) *$C_B(\mathscr{A})$ is commutative.*

(ii) *$C_B(\mathscr{A}) = C_{\mathbb{Z}}(\mathscr{A})$.*

Proof. (i) See [Jacobson 1962, Ch. X, § 1, Lemma 1].

(ii) We must show that any element $\chi \in C_{\mathbb{Z}}(\mathscr{A})$ is B-linear. Indeed if $x, y \in \mathscr{A}$ and $b \in B$ we have $\chi(b\cdot(xy)) = \chi(x(b\cdot y)) = \chi(x)(b\cdot y) = b\cdot(\chi(x)y) = b\cdot\chi(xy)$. □

We now introduce a convenient acronym, pfgc, that will be used throughout the paper.

Definition 2.9. A B-algebra $\mathscr{A}$ is said to be *pfgc* if it satisfies the following conditions

P0. $\mathscr{A} \neq (0)$

P1. $\mathscr{A}$ is perfect

P2. $\mathscr{A}$ is finitely generated as a module over its centroid $C_B(\mathscr{A})$.

Remark 2.10. The notion of pfgc algebra $\mathscr{A}$ is independent of the base ring under which $\mathscr{A}$ is viewed as an algebra. More precisely, if $\mathscr{A}$ is an algebra over B, it follows from Remark 2.7 and Lemma 2.8(ii) that $\mathscr{A}$ is a pfgc algebra over B if and only if $\mathscr{A}$ is a pfgc algebra over $\mathbb{Z}$

We now summarize the basic facts that we will need about pfgc algebras.

Proposition 2.11. *Suppose that $\mathscr{A}$ is a pfgc algebra over B. Then*

(i) *$C_B(\mathscr{A})$ is a nonzero unital commutative associative B-algebra.*

(ii) *$\mathscr{A}$ is finitely generated as a module over its multiplication algebra* $\mathrm{Mult}_B(\mathscr{A})$.

Proof. (i) Since $\mathscr{A}$ is perfect and nonzero, this follows from Lemma 2.8(i).

(ii) Let $C = C_B(\mathscr{A})$. Let $\{x_1, \dots, x_n\} \in \mathscr{A}$ be such that $\mathscr{A} = \sum Cx_i$. For each i we can write $x_i = \sum_j y_{ij} z_{ij}$ (finite sum) for some y_{ij} and z_{ij} in $\mathscr{A}$. Then

$$\mathscr{A} = \sum_i Cx_i = \sum_{i,j} C(y_{ij}z_{ij}) = \sum_{i,j} (Cy_{ij})z_{ij} \subset \sum_{i,j} \mathrm{Mult}_B(\mathscr{A})\cdot z_{ij},$$

which shows that $\mathscr{A}$ is generated by the z_{ij}'s as an $\mathrm{Mult}_B(\mathscr{A})$-module. □

3. Prime pfgc algebras

In this section, we recall some basic facts about prime algebras and consider in particular properties of prime pfgc algebras. A good basic reference on prime nonassociative algebras and their centroids is [Erickson et al. 1975].

We suppose again in this section that B is a ring and that $\mathcal{A}$ is B-algebra.

Definition 3.1. The B-algebra $\mathcal{A}$ is said to be *prime* if for all ideals $\mathcal{I}$ and $\mathcal{J}$ of the B-algebra $\mathcal{A}$ we have

$$\mathcal{I}\mathcal{J} = 0 \implies \mathcal{I} = 0 \text{ or } \mathcal{J} = 0.$$

On the other hand $\mathcal{A}$ is said to be *semiprime* if for all ideals $\mathcal{I}$ of $\mathcal{A}$ we have

$$\mathcal{I}\mathcal{I} = 0 \implies \mathcal{I} = 0.$$

The following lemma which is easily checked (see [Zhevlakov et al. 1982, Exercise 1, § 8.2]) tells us that the notion of $\mathcal{A}$ being prime (or semiprime) is independent of the base ring under which $\mathcal{A}$ is viewed as an algebra.

Lemma 3.2. *$\mathcal{A}$ is prime* (*resp. semiprime*) *as a B-algebra if and only if $\mathcal{A}$ is prime* (*resp. semiprime*) *as a $\mathbb{Z}$-algebra.*

The following is proved in [Erickson et al. 1975].

Lemma 3.3. *Assume $\mathcal{A}$ is a prime algebra over B. Then*

(i) *$C_B(\mathcal{A})$ is an integral domain and $\mathcal{A}$ is a torsion free $C_B(\mathcal{A})$-module.*

(ii) *If we denote the quotient field of $C_B(\mathcal{A})$ by $\widetilde{C_B(\mathcal{A})}$, then $\mathcal{A} \otimes_{C_B(\mathcal{A})} \widetilde{C_B(\mathcal{A})}$ is a prime algebra over $\widetilde{C_B(\mathcal{A})}$. Moreover, if $\mathcal{A}$ is finitely generated as a module over its multiplication algebra* $\mathrm{Mult}_B(\mathcal{A})$, *then $\mathcal{A} \otimes_{C_B(\mathcal{A})} \widetilde{C_B(\mathcal{A})}$ is central over $\widetilde{C_B(\mathcal{A})}$.*

Proof. (i) is Theorem 1.1(a) of [Erickson et al. 1975], whereas (ii) follows from Theorem 1.3(a) and (b) of [Erickson et al. 1975]. □

In a later section of the paper we will investigate the type of an iterated loop algebra. In that section, we will need the notion of central closure.

Definition 3.4. Let $\mathcal{A}$ be a prime pfgc algebra over B. Denote the quotient field of $C_B(\mathcal{A})$ by $\widetilde{C_B(\mathcal{A})}$, and form the $\widetilde{C_B(\mathcal{A})}$-algebra

$$\widetilde{\mathcal{A}} := \mathcal{A} \otimes_{C_B(\mathcal{A})} \widetilde{C_B(\mathcal{A})}.$$

We call $\widetilde{\mathcal{A}}$ the *central closure of* $\mathcal{A}$. (This is not apparently the same as the central closure defined in [Erickson et al. 1975, § II]. Here we are following the terminology in, for example, [McCrimmon and Zel'manov 1988, p. 154].) By Lemma 3.3(i), $\mathcal{A}$ is a torsion free $C_B(\mathcal{A})$-module, and so the map $a \mapsto a \otimes 1$ is an injection of $\mathcal{A}$ into $\widetilde{\mathcal{A}}$ which we regard as an identification. In this way $\mathcal{A}$ *is regarded as a subalgebra of its central closure* $\widetilde{\mathcal{A}}$.

We now summarize the main facts that we will need about the central closure:

Proposition 3.5. *Let $\mathcal{A}$ be a prime pfgc algebra over B. Then the central closure $\widetilde{\mathcal{A}}$ of $\mathcal{A}$ is a prime pfgc algebra over B. Moreover, $\widetilde{\mathcal{A}}$ is finite dimensional and central as an algebra over the field $\widetilde{C_B(\mathcal{A})}$.*

Proof. $\widetilde{\mathcal{A}}$ is prime by Lemma 3.3(ii). Next, since $\mathcal{A}$ is embedded as a subalgebra of $\widetilde{\mathcal{A}}$, we have $\widetilde{\mathcal{A}} \neq 0$. Also, since $\mathcal{A}$ is perfect, $\widetilde{\mathcal{A}}$ is perfect. Furthermore, since $\mathcal{A}$ is finitely generated as a $C_B(\mathcal{A})$-module, $\widetilde{\mathcal{A}}$ is finitely generated as a $\widetilde{C_B(\mathcal{A})}$-module and therefore also as a $C_B(\widetilde{\mathcal{A}})$-module (since $\lambda_{\widetilde{\mathcal{A}}}(\widetilde{C_B(\mathcal{A})}) \subset C_B(\widetilde{\mathcal{A}})$). Thus $\widetilde{\mathcal{A}}$ is pfgc.

We have just seen that $\widetilde{\mathcal{A}}$ is finite dimensional over $\widetilde{C_B(\mathcal{A})}$.

Finally, since $\mathcal{A}$ is pfgc, Proposition 2.11(ii) tells us that $\mathcal{A}$ is finitely generated as a $\mathrm{Mult}_B(\mathcal{A})$-module. Thus $\widetilde{\mathcal{A}}$ is central over $\widetilde{C_B(\mathcal{A})}$ by Lemma 3.3(ii). □

4. Loop algebras

Assumptions and notation: *For the rest of the article*, k will denote a fixed base field. Unless indicated to the contrary, the term algebra will mean algebra over k. For the sake of brevity, if $\mathcal{A}$ is an algebra (over k), we will often write

$$C(\mathcal{A}) := C_k(\mathcal{A}).$$

In this section we recall the definition of a loop algebra and derive some of its basic properties.

Throughout the section let m be a positive integer and let

$$\mathbb{Z}_m = \{\bar{\imath} : i \in \mathbb{Z}\}$$

be the group of integers modulo m, where $\bar{\imath} = i + m\mathbb{Z} \in \mathbb{Z}_m$ for $i \in \mathbb{Z}$. Let

$$R = k[t^{\pm 1}] \quad \text{and} \quad S = k[z^{\pm 1}]$$

be the algebras of Laurent polynomials in the variables t and z respectively, and we identify R as a subalgebra of S by identifying

$$t = z^m.$$

Observe that S is a free R-module of rank m with basis $\{1, z, \dots, z^{m-1}\}$, and hence the ring extension S/R is faithfully flat.

Recall that a $\mathbb{Z}_m$-*grading* of the algebra $\mathcal{A}$ is an indexed family $\Sigma = \{\mathcal{A}_{\bar{\imath}}\}_{\bar{\imath} \in \mathbb{Z}_m}$ of subspaces of $\mathcal{A}$ so that $\mathcal{A} = \bigoplus_{\bar{\imath} \in \mathbb{Z}_m} \mathcal{A}_{\bar{\imath}}$ and $\mathcal{A}_{\bar{\imath}}\mathcal{A}_{\bar{\jmath}} \subset \mathcal{A}_{\bar{\imath}+\bar{\jmath}}$ for $\bar{\imath}, \bar{\jmath} \in \mathbb{Z}_m$. The integer m is called the *modulus* of Σ.

Definition 4.1. Suppose that $\mathcal{A}$ is a k-algebra, and we are given a $\mathbb{Z}_m$-grading Σ of the algebra $\mathcal{A}$:

$$\mathcal{A} = \bigoplus_{\bar{\imath} \in \mathbb{Z}_m} \mathcal{A}_{\bar{\imath}}.$$

In $\mathcal{A} \otimes_k S$ we define

$$L(\mathcal{A}, \Sigma) := \bigoplus_{i \in \mathbb{Z}} \mathcal{A}_{\bar{\imath}} \otimes_k z^i = (\mathcal{A}_{\bar{0}} \otimes_k R) \oplus (\mathcal{A}_{\bar{1}} \otimes_k zR) \oplus \cdots \oplus (\mathcal{A}_{\overline{m-1}} \otimes_k z^{m-1} R).$$

Then $L(\mathcal{A}, \Sigma)$ is an R-subalgebra of $\mathcal{A} \otimes_k S$ that we call the *loop algebra of* Σ *based on* $\mathcal{A}$. Since $L(\mathcal{A}, \Sigma)$ is an algebra over R, it is also an algebra over k.

Remark 4.2. If we wish to emphasize the role of the variable z in the construction of the loop algebra we write $L(\mathcal{A}, \Sigma)$ as $L(\mathcal{A}, \Sigma, z)$.

Example 4.3. If $m = 1$, then $\mathbb{Z}_m = \{\bar{0}\}$, $\mathcal{A} = \mathcal{A}_{\bar{0}}$ and $L(\mathcal{A}, \Sigma) = \mathcal{A} \otimes_k S$ is called the *untwisted* loop algebra based on $\mathcal{A}$.

Remark 4.4. Suppose that k contains a primitive m-th root of unity ζ_m. In that case we can choose to work with finite order automorphisms of period m rather than $\mathbb{Z}_m$-gradings, provided that we fix the choice of ζ_m.

Indeed, suppose that $\mathcal{A}$ is an algebra. If σ is an algebra automorphism of period m of $\mathcal{A}$, we may define a $\mathbb{Z}_m$-grading $\Sigma = \{\mathcal{A}_{\bar{\imath}}\}_{\bar{\imath} \in \mathbb{Z}_m}$ of $\mathcal{A}$ by setting

$$\mathcal{A}_{\bar{\imath}} = \{ x \in \mathcal{A} \mid \sigma(x) = \zeta_m^i x \},$$

for $\bar{\imath} \in \mathbb{Z}_m$. We refer to this grading Σ as the *grading determined by* σ. It is clear that any $\mathbb{Z}_m$-grading is determined by a unique automorphism σ in this way. If Σ is the grading determined by σ, we denote the algebra $L(\mathcal{A}, \Sigma)$ by $L(\mathcal{A}, \sigma)$, or $L(\mathcal{A}, \sigma, z)$ if we want to emphasize the role of z. The algebra $L(\mathcal{A}, \sigma)$ can alternately be defined as the subalgebra of fixed points in $\mathcal{A} \otimes_k S$ of the automorphism $\sigma \otimes \eta_m^{-1}$, where $\eta_m \in \mathrm{Aut}_k(S)$ is defined by $\eta_m(z) = \zeta_m z$.

Remark 4.5. When $k = \mathbb{C}$, $\mathcal{A}$ is a finite dimensional simple Lie algebra over k and σ is a finite order automorphism of $\mathcal{A}$, the loop algebra $L(\mathcal{A}, \sigma)$ was used by V. Kac in [Kac 1969] to give realizations of all affine Kac–Moody Lie algebras and to classify finite order automorphisms of $\mathcal{A}$. (See [Kac 1990, Ch. 8] and [Helgason 1978, Ch. X, § 5] for more information about this.)

For the rest of the section, let $\mathcal{A}$ *a* k*-algebra, let* Σ *be a grading of* $\mathcal{A}$ *by* $\mathbb{Z}_m$*, and let*

$$\mathcal{L} = L(\mathcal{A}, \Sigma).$$

We next describe a useful canonical form for elements of $\mathcal{A} \otimes_k S$ in terms of elements of $L(\mathcal{A}, \Sigma)$. For this purpose note that $\mathcal{A} \otimes_k S$ is an S-module (with action denoted by "$\cdot$") and $L(\mathcal{A}, \Sigma)$ is contained in $\mathcal{A} \otimes_k S$. Thus we may write expressions like $\sum_{i=0}^{m-1} z^i \cdot x_i \in \mathcal{A} \otimes_k S$ if $x_0, \ldots, x_{m-1} \in L(\mathcal{A}, \Sigma)$.

Lemma 4.6. *Each element of $\mathcal{A} \otimes_k S$ can be written uniquely in the form*

$$\sum_{i=0}^{m-1} z^i \cdot x_i \tag{4–1}$$

where $x_0, \ldots, x_{m-1} \in \mathcal{L}$.

Proof. This fact was proved using a Galois cocycle argument in [Allison et al. 2004, Theorem 3.6 (b)] in the case when k contains a primitive m-th root of unity. We give a direct proof here instead. Let $x \in \mathcal{A} \otimes_k S$. Then x is the sum of elements of the form $a \otimes z^j$, where $j \in \mathbb{Z}$ and $a \in \mathcal{A}_{\bar{\ell}}$ for some $\ell \in \mathbb{Z}$. But, if we write $j - \ell = qm + i$, where $q \in \mathbb{Z}$ and $0 \le i \le m-1$, then $a \otimes z^j = z^i \cdot (a \otimes z^{j-i}) = z^i \cdot (a \otimes z^{qm+\ell})$ and $a \otimes z^{qm+\ell} \in \mathcal{L}$. So x can be expressed in the form (4–1). For uniqueness, suppose that $\sum_{i=0}^{m-1} z^i \cdot x_i = 0$, where $x_0, \ldots, x_{m-1} \in \mathcal{L}$. Write $x_i = \sum_{j \in \mathbb{Z}} a_{ij} \otimes z^j$, where $a_{ij} \in \mathcal{A}_{\bar{j}}$ for all j and only finitely many a_{ij} are nonzero. Then

$$\sum_{i=0}^{m-1} \sum_{j \in \mathbb{Z}} a_{ij} \otimes z^{i+j} = 0.$$

For $0 \le \ell \le m-1$, the $\mathcal{A}_{\bar{\ell}} \otimes_k S$-component of the expression on the left above must be zero. Thus we have

$$\sum_{i=0}^{m-1} \sum_{j \equiv \ell} a_{ij} \otimes z^{i+j} = 0$$

for $0 \le \ell \le m-1$, where $\equiv$ denotes congruence modulo m. The exponents $i + j$ appearing in this sum are all distinct and so we have $a_{ij} = 0$ for all i, j and hence $x_i = 0$ for all i. □

Next note that we have the canonical map $\xi = \xi_{\mathcal{A},\Sigma} : L(\mathcal{A}, \Sigma) \otimes_R S \to \mathcal{A} \otimes_k S$ defined by

$$\xi\left(x \otimes z^i\right) = z^i \cdot x$$

for $x \in L(\mathcal{A}, \Sigma)$, $i \in \mathbb{Z}$. As observed in [Allison et al. 2004, Theorem 3.6(b)], Lemma 4.6 has the following interpretation:

Lemma 4.7. *The map $\xi_{\mathcal{A},\Sigma} : L(\mathcal{A}, \Sigma) \otimes_R S \to \mathcal{A} \otimes_k S$ is an S-algebra isomorphism of $L(\mathcal{A}, \Sigma) \otimes_R S$ onto $\mathcal{A} \otimes_k S$.*

Proof. Clearly ξ is a homomorphism of S-algebras. Moreover, each element of $L(\mathcal{A}, \Sigma) \otimes_R S$ can be expressed in the form $\sum_{i=0}^{m-1} x_i \otimes z^i$, where $x_i \in L(\mathcal{A}, \Sigma)$ for each i, and so ξ is bijective by Lemma 4.6. □

Remark 4.8. Lemma 4.7 tells us that after base ring extension from R to S the loop algebra $L(\mathcal{A}, \Sigma)$ becomes isomorphic to the untwisted loop algebra $\mathcal{A} \otimes_k S$. In other words, $L(\mathcal{A}, \Sigma)$ is "untwisted" by the extension S/R. This fact is of great importance in the study of loop algebras since, among other things, it allows one to use the tools of Galois cohomology to study loop algebras [Allison et al. 2004; Pianzola 2002].

Lemma 4.9.

(i) *If* $\mathcal{A} \neq 0$, *then* $L(\mathcal{A}, \Sigma) \neq 0$.

(ii) *If* $\mathcal{A}$ *is perfect*, *then* $L(\mathcal{A}, \Sigma)$ *is perfect.*

Proof. Statement (i) is clear and statement (ii) is easily checked (see the argument in [Allison et al. 2004, Lemma 4.3]). □

We now examine the centroid of $\mathcal{L} = L(\mathcal{A}, \Sigma)$.

First note that since $\mathcal{L}$ is an R-algebra, $C_R(\mathcal{L})$ is naturally an R-algebra (see Definition 2.1(iii)). So since $C_R(\mathcal{L}) \subset C(\mathcal{L})$, it follows that $C(\mathcal{L})$ is also an R-algebra.

Next by Lemma 2.5 the centroid $C(\mathcal{A})$ inherits a $\mathbb{Z}_m$-grading that we denote by $C(\Sigma)$. Under this grading we have

$$C(\mathcal{A}) = \bigoplus_{\bar{i} \in \mathbb{Z}_m} C(\mathcal{A})_{\bar{i}},$$

where

(4–2) $$C(\mathcal{A})_{\bar{i}} = \{\, \chi \in C(\mathcal{A}) \mid \chi(\mathcal{A}_{\bar{j}}) \subset \mathcal{A}_{\bar{i}+\bar{j}} \text{ for } \bar{j} \in \mathbb{Z}_m \,\}.$$

Now let

$$\psi := \psi_{\mathcal{A},\Sigma} : L(C(\mathcal{A}), C(\Sigma)) \to C_R\big(L(\mathcal{A}, \Sigma)\big)$$

be the unique k-linear map so that

$$\big(\psi(\chi \otimes z^i)\big)(a \otimes z^j) = \chi(a) \otimes z^{i+j}$$

for $i, j \in \mathbb{Z}$, $\chi \in C(\mathcal{A})_{\bar{i}}$, $a \in \mathcal{A}_{\bar{j}}$. It is immediate from this definition that ψ is a homomorphism of R-algebras that we call canonical.

Lemma 4.10. *Assume* $\mathcal{A}$ *is finitely generated as a module over its multiplication algebra* $\mathrm{Mult}_k(\mathcal{A})$. *Then the map* $\psi_{\mathcal{A},\Sigma} : L\big(C(\mathcal{A}), C(\Sigma)\big) \to C_R\big(L(\mathcal{A}, \Sigma)\big)$ *is an* R*-algebra isomorphism.*

Proof. Since the ring extension S/R is faithfully flat, to show that ψ is an R-module isomorphism it suffices to show that ψ becomes an isomorphism of S-modules after the base change from R to S. That this is so follows from the commutative diagram

$$
\begin{array}{ccc}
L\big(C(\mathscr{A}), C(\Sigma)\big) \otimes_R S & \xrightarrow{\ \psi \otimes 1\ } & C_R(\mathscr{L}) \otimes_R S \\
 & & \Big\downarrow \nu_{\mathscr{L}} \\
\xi_C \Big\downarrow & & C_S\big(\mathscr{L} \otimes_R S\big) \\
 & & \Big\downarrow C_S(\xi) \\
C(\mathscr{A}) \otimes_k S & \xrightarrow[\ \nu_{\mathscr{A}}\ ]{} & C_S(\mathscr{A} \otimes_k S)
\end{array}
$$

in view of the fact that all vertical maps and the bottom row therein are S-isomorphisms. In this diagram $\xi_C = \xi_{C(\mathscr{A}),C(\Sigma)}$ as in Lemma 4.7, $\nu_{\mathscr{A}} = \nu_{\mathscr{A},k,S}$ as in Lemma 2.3(a), $\nu_{\mathscr{L}} = \nu_{\mathscr{L},R,S}$ as in Lemma 2.3(b), and $C_S(\xi)$ is the isomorphism induced by the isomorphism $\xi = \xi_{\mathscr{A},\Sigma} : \mathscr{L} \otimes_R S \to \mathscr{A} \otimes_k S$ (see Lemmas 2.2 and 4.7). □

The following proposition tells us that the centroid of a loop algebra based on a pfgc algebra $\mathscr{A}$ is isomorphic to the loop algebra of the centroid of $\mathscr{A}$.

Proposition 4.11. *Let $\mathscr{L} = L(\mathscr{A}, \Sigma)$ be a loop algebra based on a pfgc algebra $\mathscr{A}$. Then $C_R(\mathscr{L}) = C(\mathscr{L})$, and the canonical map*

$$\psi = \psi_{\mathscr{A},\Sigma} : L(C(\mathscr{A}), C(\Sigma)) \to C(\mathscr{L})$$

is an R-algebra isomorphism.

Proof. Since $\mathscr{L}$ is perfect by Lemma 4.9(ii), it follows that $C_R(\mathscr{L}) = C(\mathscr{L})$ by Lemma 2.8(ii). Also since $\mathscr{A}$ is pfgc, it follows from Proposition 2.11(ii) that $\mathscr{A}$ is finitely generated as a module over $\mathrm{Mult}_k(\mathscr{A})$. Thus, by Lemma 4.10, ψ is an R-algebra isomorphism from $L(C(\mathscr{A}), C(\Sigma))$ onto $C(\mathscr{L})$. □

Finally we want to show that a loop algebra based on a pfgc algebra is pfgc. For this we will use the following:

Lemma 4.12. *If $\mathscr{A}$ is finitely generated as a $C(\mathscr{A})$-module then $L(\mathscr{A}, \Sigma)$ is finitely generated as a $C_R\big(L(\mathscr{A}, \Sigma)\big)$-module.*

Proof. Let $\{a_1, \dots, a_p\}$ be a set of homogeneous elements of $\mathscr{A}$ that generates $\mathscr{A}$ as a $C(\mathscr{A})$-module. Fix integers $d_1, \dots, d_p$ so that $a_j \in \mathscr{A}_{\overline{d_j}}$. Let $\mathscr{M}$ be the $C_R(\mathscr{L})$-submodule of $\mathscr{L}$ generated by the elements $a_k \otimes z^{d_k}$. Since S/R is flat we may identify $\mathscr{M} \otimes_R S$ as an S-submodule of $\mathscr{L} \otimes_R S$, and since S/R is faithfully flat it is sufficient to show that $\mathscr{M} \otimes_R S = \mathscr{L} \otimes_R S$ [Bourbaki 1972, Ch. I, § 3.1, Proposition 2]. We do this by showing that $\xi(\mathscr{M} \otimes_R S) = \xi(\mathscr{L} \otimes_R S)$, where $\xi = \xi_{\mathscr{A},\Sigma} : \mathscr{L} \otimes_R S \longrightarrow \mathscr{A} \otimes_k S$ is the S-algebra isomorphism from Lemma 4.7.

Suppose that $i, j \in \mathbb{Z}$, $\chi \in C(\mathscr{A})_{\bar{\imath}}$ and $1 \le \ell \le p$. Then $\psi(\chi \otimes z^i)$ is an element of $C_R(\mathscr{L})$, where $\psi = \psi_{\mathscr{A},\Sigma}$. So $\big(\psi(\chi \otimes z^i)\big)(a_\ell \otimes z^{d_\ell}) \in \mathscr{M}$. But under ξ we have

$$\Big(\big(\psi(\chi \otimes z^i)\big)(a_\ell \otimes z^{d_\ell})\Big) \otimes z^j \mapsto \chi(a_\ell) \otimes z^{d_\ell + i + j}.$$

Since $\{a_1, \dots, a_p\}$ generates $\mathcal{A}$ as a $C(\mathcal{A})$-module, it follows $\xi(\mathcal{M}\otimes_R S) = \xi(\mathcal{L}\otimes_R S)$ as needed. □

Proposition 4.13. *Let $\mathcal{L} = L(\mathcal{A}, \Sigma)$ be a loop algebra based on a pfgc algebra $\mathcal{A}$. Then $\mathcal{L}$ is a pfgc algebra.*

Proof. $\mathcal{L} \neq (0)$ and $\mathcal{L}$ is perfect by Lemma 4.9. So P0 and P1 hold (see Definition 2.9). By Lemma 4.12, $\mathcal{L}$ is finitely generated as a $C_R(\mathcal{L})$-module. But by Proposition 4.11, we have $C_R(\mathcal{L}) = C(\mathcal{L})$. Thus $\mathcal{L}$ is finitely generated as a $C(\mathcal{L})$-module and so P2 holds. Hence $\mathcal{L}$ is pfgc. □

5. Iterated loop algebras

In this section we define iterated loop algebras and prove some of their basic properties.

Notation: *For the rest of this article, we fix some notation.* Let n be a positive integer. Let $z_1, \dots, z_n$ be a sequence of algebraically independent variables over k. For $0 \le p \le n$, let

$$S^{\otimes p} := k[z_1^{\pm 1}, \dots, z_p^{\pm 1}]$$

be the algebra of Laurent polynomials in the variables $z_1, \dots, z_p$ over k. (So $S^{\otimes 0} = k$.) We identify $S^{\otimes p} \otimes S^{\otimes q} = S^{\otimes (p+q)}$ in the natural fashion when $0 \le p, q \le n$ and $p + q \le n$. We also fix a sequence $m_1, \dots, m_n$ of positive integers, and we set

$$I_p := \{ (i_1, \dots, i_p) \in \mathbb{Z}^p \mid 0 \le i_j \le m_j - 1 \text{ for all } j \},$$

for $1 \le p \le n$.

Definition 5.1. Suppose that $\mathcal{A}$ is an algebra over k. An algebra $\mathcal{L}$ over k is called an *n-step loop algebra* or an *iterated loop algebra* based on $\mathcal{A}$ if there exists a sequence $\mathcal{L}_0, \mathcal{L}_1, \dots, \mathcal{L}_n$ of algebras so that $\mathcal{L}_0 = \mathcal{A}$, $\mathcal{L}_n = \mathcal{L}$ and

$$\mathcal{L}_p = L(\mathcal{L}_{p-1}, \Sigma_p, z_p),$$

for $1 \le p \le n$, where Σ_p is a $\mathbb{Z}_{m_p}$-grading of $\mathcal{L}_{p-1}$. (See Remark 4.2 for the notation used here.) In that case we write

$$\mathcal{L} = L(\mathcal{A}, \Sigma_1, \dots, \Sigma_n)$$

(suppressing in the notation the role of the variables $z_1, \dots, z_n$).

Remark 5.2. Suppose that $\mathcal{L}$ is an n-step loop algebra based on $\mathcal{A}$ and we have the notation from Definition 5.1.

(i) For $1 \le p \le n$, $\mathcal{L}_p = L(\mathcal{A}, \Sigma_1, \dots, \Sigma_p)$ is a p-step loop algebra based on $\mathcal{A}$.

(ii) Observe that $\mathcal{L}_p \subset \mathcal{L}_{p-1} \otimes_k k[z_p^{\pm 1}]$ for $1 \le p \le n$. Thus

$$\mathcal{L}_p \subset (\dots((\mathcal{A} \otimes_k k[z_1^{\pm 1}]) \otimes_k k[z_2^{\pm 1}]) \dots) \otimes_k k[z_p^{\pm 1}] = \mathcal{A} \otimes k[z_1^{\pm 1}, \dots, z_p^{\pm 1}],$$

for $0 \le p \le n$, where the last equality is the natural identification using the associativity of the tensor product and the identification $k[z_1^{\pm 1}] \otimes_k \ldots \otimes_k k[z_p^{\pm 1}] = S^{\otimes p}$. Consequently, *$\mathcal{L}_p$ is a subalgebra of $\mathcal{A} \otimes_k S^{\otimes p}$ for $0 \le p \le n$, and in particular $\mathcal{L}$ is a subalgebra of $\mathcal{A} \otimes_k S^{\otimes n}$.*

(iii) Suppose that k contains a primitive m_i-th root of unity ζ_{m_i} (which we fix) for $i = 1, \ldots, n$. Then for $p = 1, \ldots, n$, the grading Σ_p of $\mathcal{L}_{p-1}$ is determined by a unique automorphism σ_p of $\mathcal{L}_{p-1}$ of period m_p. We then denote the algebra $L(\mathcal{A}, \Sigma_1, \ldots, \Sigma_n)$ by $L(\mathcal{A}, \sigma_1, \ldots, \sigma_n)$.

Example 5.3. If $m_1 = \cdots = m_n = 1$ then $L(\mathcal{A}, \Sigma_1, \ldots, \Sigma_n) = \mathcal{A} \otimes_k S^{\otimes n}$ is called the *untwisted* n-step loop algebra based on $\mathcal{A}$.

Example 5.4 (Multiloop algebras). Suppose that k contains a primitive m_i-th root of unity ζ_{m_i} for $1 \le i \le n$. Let $\mathcal{A}$ be an algebra, and let $\sigma_1, \ldots, \sigma_n$ be commuting finite order automorphisms of $\mathcal{A}$ with periods $m_1, \ldots, m_n$ respectively. Let

$$\mathcal{A}_{\bar{\imath}_1, \ldots, \bar{\imath}_n} = \{ x \in \mathcal{A} \mid \sigma_j x = \zeta_{m_j}^{i_j} x \text{ for } 1 \le j \le n \}$$

for $(i_1, \ldots, i_n) \in \mathbb{Z}^n$, where $\bar{\imath}_j := i_j + m_j\mathbb{Z} \in \mathbb{Z}_{m_j}$ for $1 \le j \le n$. Then

$$\mathcal{A} = \bigoplus_{(i_1, \ldots, i_n) \in I_n} \mathcal{A}_{\bar{\imath}_1, \ldots, \bar{\imath}_n},$$

and we set

$$M(\mathcal{A}, \sigma_1, \ldots, \sigma_n) := \bigoplus_{(i_1, \ldots, i_n) \in \mathbb{Z}^n} \mathcal{A}_{\bar{\imath}_1, \ldots, \bar{\imath}_n} \otimes_k z_1^{i_1} \ldots z_n^{i_n}$$

in $\mathcal{A} \otimes_k S^{\otimes n}$. Then $M(\mathcal{A}, \sigma_1, \ldots, \sigma_n)$ is a subalgebra of $\mathcal{A} \otimes_k S^{\otimes n}$ that we call the n-step *multiloop algebra* of $\sigma_1, \ldots, \sigma_n$ based on $\mathcal{A}$.

Now the multiloop algebra $\mathcal{L} = M(\mathcal{A}, \sigma_1, \ldots, \sigma_n)$ can be interpreted as an iterated loop algebra. To see this, let $\mathcal{L}_0 = \mathcal{A}$ and let $\mathcal{L}_p = M(\mathcal{A}, \sigma_1, \ldots, \sigma_p)$ for $1 \le p \le n$. Then by definition we have $\mathcal{L}_0 = \mathcal{A}$ and $\mathcal{L}_n = \mathcal{L}$. Also, for $1 \le p \le n$, we may define a $\mathbb{Z}_{m_p}$-grading Σ_p on $\mathcal{L}_{p-1}$ by setting

$$(\mathcal{L}_{p-1})_{\bar{\imath}_p} = \bigoplus_{(i_1, \ldots, i_{p-1}) \in \mathbb{Z}^{p-1}} \mathcal{A}_{\bar{\imath}_1, \ldots, \bar{\imath}_p} \otimes_k z_1^{i_1} \ldots z_{p-1}^{i_{p-1}}$$

for $\bar{\imath}_p \in \mathbb{Z}_{m_p}$, in which case it is then clear that $L(\mathcal{L}_{p-1}, \Sigma_p, z_p) = \mathcal{L}_p$. So $\mathcal{L} = L(\mathcal{A}, \Sigma_1, \ldots, \Sigma_n)$.

We have just seen in Example 5.4 that any multiloop algebra is an iterated loop algebra. However we will see later in Example 9.7 that there are iterated loop algebras $\mathcal{A}$ that are not multiloop algebras.

For the rest of the section *we assume that*

$$\mathcal{L} = L(\mathcal{A}, \Sigma_1, \ldots, \Sigma_n)$$

is an n-step loop algebra based on an algebra $\mathcal{A}$ *over* k, and we use the notation $\mathcal{L}_0, \dots, \mathcal{L}_n$ of Definition 5.1.

Our first theorem describes some important basic algebraic properties that are inherited by a loop algebra from its base. In the last part of this theorem we will see how the Krull dimension of a loop algebra depends on the Krull dimension of its base. Here and subsequently we use

$$\operatorname{Dim} \mathcal{C}$$

to denote the *Krull dimension* of a unital commutative associative k-algebra $\mathcal{C}$ (when regarded as a ring). Note that if $\mathcal{C}$ is finitely generated as a k-algebra then $\operatorname{Dim} \mathcal{C}$ is finite [Kunz 1985, p. 52].

Theorem 5.5. *Let* $\mathcal{L} = L(\mathcal{A}, \Sigma_1, \dots, \Sigma_n)$.

(i) *If* $\mathcal{A} \neq 0$ *then* $\mathcal{L} \neq 0$.

(ii) *If* $\mathcal{A}$ *is perfect then* $\mathcal{L}$ *is perfect.*

(iii) *If* $\mathcal{A}$ *is pfgc then* $\mathcal{L}$ *is pfgc.*

(iv) *If* $\mathcal{A}$ *is prime then* $\mathcal{L}$ *is prime.*

(v) *If* $\mathcal{A}$ *is unital then* $\mathcal{L}$ *is a unital subalgebra of* $\mathcal{A} \otimes S^{\otimes n}$.

(vi) *If* $\mathcal{A}$ *is commutative then* $\mathcal{L}$ *is commutative.*

(vii) *If* $\mathcal{A}$ *is associative then* $\mathcal{L}$ *is associative.*

(viii) *If* $\mathcal{A}$ *is an integral domain then* $\mathcal{L}$ *is an integral domain.*

(ix) *If* $\mathcal{A}$ *is unital and finitely generated as a k-algebra then* $\mathcal{L}$ *is unital and finitely generated as a k-algebra.*

(x) *If* $\mathcal{A}$ *is unital, commutative, associative and finitely generated as a k-algebra, then* $\mathcal{L}$ *has the same properties and*

$$\operatorname{Dim} \mathcal{L} = \operatorname{Dim} \mathcal{A} + n. \tag{5–1}$$

Proof. Since $\mathcal{L}_{p+1}$ is a loop algebra based on $\mathcal{L}_p$ for $0 \leq p \leq n-1$, we can assume in the proof of each of these statements that $n = 1$. So we may use the notation of Section 4:

$$m = m_1, \; z = z_1, \; \Sigma = \Sigma_1, \; \mathcal{L} = L(\mathcal{A}, \Sigma, z), \; S = S^{\otimes 1} = k[z^{\pm 1}] \text{ and } R = k[z^{\pm m}].$$

Now (i) and (ii) follow from Lemma 4.9. (iii) follows from Proposition 4.13. (v) follows from that fact that $1_{\mathcal{A}} \in \mathcal{A}_{\bar{0}}$, since then $1_{\mathcal{A}} \otimes 1_S \in \mathcal{L}$. (vi), (vii) and (viii) follow from the fact that $\mathcal{L}$ is a subalgebra of $\mathcal{A} \otimes_k S \simeq_k \mathcal{A}[z^{\pm 1}]$. So we only need to prove (iv), (ix) and (x).

(iv) We show first that $\mathscr{A} \otimes_k S$ is prime. For this let $\mathscr{I}$ and $\mathscr{J}$ be ideals of the k-algebra $\mathscr{A} \otimes_k S$ such that $\mathscr{I}\mathscr{J} = 0$. For $m \in \mathbb{Z}$, let

$$\mathscr{I}_m = \{\, a \in \mathscr{A} \mid \exists\, a_i \in \mathscr{A} \text{ for } i \geq m \text{ with } a_m = a \text{ and } \textstyle\sum_{i \geq m} a_i \otimes z^i \in \mathscr{I} \,\},$$

in which case $\mathscr{I}_m$ is an ideal of $\mathscr{A}$. Similarly, using $\mathscr{J}$ instead of $\mathscr{I}$, we define an ideal $\mathscr{J}_n$ of $\mathscr{A}$ for $n \in \mathbb{Z}$. Furthermore, since $\mathscr{I}\mathscr{J} = 0$, we have $\mathscr{I}_m\mathscr{J}_n = 0$ for $m, n \in \mathbb{Z}$. Now suppose that $\mathscr{I} \neq 0$. Then $\mathscr{I}_m \neq 0$ for some $m \in \mathbb{Z}$. Thus, since $\mathscr{A}$ is prime, we have $\mathscr{J}_n = 0$ for all $n \in \mathbb{Z}$ and so $\mathscr{J} = 0$. Therefore $\mathscr{A} \otimes_k S$ is prime.

But $\mathscr{L} \otimes_R S \simeq_S \mathscr{A} \otimes_k S$ by Lemma 4.7. Hence $\mathscr{L} \otimes_R S$ is a prime algebra. To prove that $\mathscr{L}$ is prime (as a k-algebra), it is enough to show that $\mathscr{L}$ is a prime R-algebra (by Lemma 3.2). For this let $\mathscr{I}$ and $\mathscr{J}$ be ideals of the R-algebra $\mathscr{L}$ such that $\mathscr{I}\mathscr{J} = 0$. Since S/R is flat, we can identify $\mathscr{I} \otimes_R S$ and $\mathscr{J} \otimes_R S$ as ideals of the S-algebra $\mathscr{L} \otimes_R S$. Furthermore, we have $(\mathscr{I} \otimes_R S)(\mathscr{J} \otimes_R S) = 0$. Since $\mathscr{L} \otimes_R S$ is prime, either $\mathscr{I} \otimes_R S$ or $\mathscr{J} \otimes_R S$ is 0. Therefore $\mathscr{I} = 0$ or $\mathscr{J} = 0$ by the faithful flatness of S/R.

(ix) $\mathscr{L}$ is unital by (v). Let $\{a_1, \ldots, a_p\}$ be a set of homogeneous elements of $\mathscr{A}$ that generates $\mathscr{A}$ as a k-algebra, and fix integers $d_1, \ldots, d_p$ so that $a_j \in \mathscr{A}_{\bar{d}_j}$. One easily checks that the elements $a_1 \otimes z^{d_1}, \ldots, a_p \otimes z^{d_p}$ together with the elements $1_{\mathscr{A}} \otimes z^m$ and $1_{\mathscr{A}} \otimes z^{-m}$ generate $\mathscr{L}$ as a k-algebra.

(x) $\mathscr{L}$ is unital, commutative, associative and finitely generated as a k-algebra by (v), (vi), (vii) and (ix), and so the Krull dimensions of both $\mathscr{A}$ and $\mathscr{L}$ are finite. Now recall that $\mathscr{L}$ is a subalgebra of $\mathscr{A} \otimes_k S$ and, by Lemma 4.6, each element of $\mathscr{A} \otimes_k S$ can be written uniquely in the form

$$\sum_{i=0}^{m-1} z^i \cdot x_i = \sum_{i=0}^{m-1} x_i (1_{\mathscr{A}} \otimes z^i)$$

where $x_0, \ldots, x_{m-1} \in \mathscr{L}$. Thus $\mathscr{A} \otimes_k S$ is a free $\mathscr{L}$-module of rank m, and so in particular $\mathscr{A} \otimes_k S$ is a finitely generated $\mathscr{L}$-module. Hence $\mathscr{A} \otimes_k S/\mathscr{L}$ is an integral ring extension and so by [Kunz 1985, Corollary II.2.13],

$$\operatorname{Dim} \mathscr{L} = \operatorname{Dim} \big(\mathscr{A} \otimes_k S\big).$$

On the other hand since both $\mathscr{A}$ and S are finitely generated k-algebras

$$\operatorname{Dim} \big(\mathscr{A} \otimes_k S\big) = \operatorname{Dim} \mathscr{A} + \operatorname{Dim} S$$

[Kunz 1985, Corollary II.3.9]. Since $\operatorname{Dim} S = 1$, we obtain $\operatorname{Dim} \mathscr{L} = \operatorname{Dim} \mathscr{A} + 1$. □

Remark 5.6. It follows in particular from Theorem 5.5 that any n-step loop algebra based on a prime pfgc algebra is a prime pfgc algebra. The corresponding statement is not true for simple pfgc algebras. For example, an untwisted pfgc algebra $\mathscr{A} \otimes_k S^{\otimes n}$ is never simple (since $S^{\otimes n}$ is not simple). This is the reason why prime pfgc algebras are natural algebras to consider when studying loop algebras, even if one's main interest is in the case when the base algebras are simple.

We conclude this section with a generalization to iterated loop algebras of the canonical form described in Lemma 4.6. If $1 \le p \le n$, we use the usual convenient notation

$$z^{\mathbf{i}} = z_1^{i_1} \dots z_p^{i_p}$$

for $\mathbf{i} = (i_1, \dots, i_p) \in \mathbb{Z}^p$. Note that $\mathcal{A} \otimes_k S^{\otimes p}$ is an $S^{\otimes p}$-module (with action denoted by "·"), and $\mathcal{L}_p$ is contained in $\mathcal{A} \otimes_k S^{\otimes p}$. Thus, we can write expressions like $\sum_{\mathbf{i} \in I_p} z^{\mathbf{i}} \cdot x_{\mathbf{i}} \in \mathcal{A} \otimes_k S^{\otimes p}$, where $x_{\mathbf{i}} \in \mathcal{L}_p$ for all $\mathbf{i} \in I_p$.

Lemma 5.7. *If $1 \le p \le n$, each element in $\mathcal{A} \otimes_k S^{\otimes p}$ can be expressed uniquely in the form*

$$\sum_{\mathbf{i} \in I_p} z^{\mathbf{i}} \cdot x_{\mathbf{i}}, \tag{5–2}$$

where $x_{\mathbf{i}} \in \mathcal{L}_p$ for all $\mathbf{i}$.

Proof. We argue by induction on p. When $p = 1$, the statement follows from Lemma 4.6. So we suppose that the statement is true for p, where $1 \le p \le n-1$.

Let $x \in \mathcal{A} \otimes S^{\otimes(p+1)}$. To show that x can be expressed in the form (5–2), note first that x is a sum of elements of the form $x' \otimes z_{p+1}^j$, where $x' \in \mathcal{A} \otimes_k S^{\otimes p}$ and $j \in \mathbb{Z}$. But by the induction hypothesis, x' is the sum of elements of the form $z^{\mathbf{i}} \cdot x''$, where $\mathbf{i} \in I_p$ and $x'' \in \mathcal{L}_p$. Thus x is the sum of elements of the form

$$(z^{\mathbf{i}} \cdot x'') \otimes z_{p+1}^j = z^{\mathbf{i}} \cdot (x'' \otimes z_{p+1}^j).$$

But $x'' \otimes z_{p+1}^j \in \mathcal{L}_p \otimes_k k[z_{p+1}^{\pm 1}]$, and so, by Lemma 4.6, $x'' \otimes z_{p+1}^j$ is the sum of elements of the form $z_{p+1}^\ell \cdot x'''$, where $0 \le \ell \le m_{p+1} - 1$ and $x''' \in \mathcal{L}_{p+1}$. Thus x is the sum of elements of the form

$$z^{\mathbf{i}} \cdot (z_{p+1}^\ell \cdot x''') = (z^{\mathbf{i}} z_{p+1}^\ell) \cdot x'''$$

as desired.

For uniqueness, suppose that $\sum_{\mathbf{j} \in I_{p+1}} z^{\mathbf{j}} \cdot x_{\mathbf{j}} = 0$, where $x_{\mathbf{j}} \in \mathcal{L}_{p+1}$ for each $\mathbf{j} \in I_{p+1}$. Then

$$\sum_{\mathbf{i} \in I_p} \sum_{\ell=0}^{m_{p+1}-1} (z^{\mathbf{i}} z_{p+1}^\ell) \cdot x_{\mathbf{i},\ell} = 0,$$

where, if $\mathbf{i} = (i_1, \dots, i_p) \in I_p$ and $0 \le \ell \le m_{p+1} - 1$, we are using the notation $x_{\mathbf{i},\ell} := x_{(i_1,\dots,i_p,\ell)} \in \mathcal{L}_{p+1}$. So we have

$$\sum_{\mathbf{i}\in I_p} z^{\mathbf{i}} \cdot \Bigl(\sum_{\ell=0}^{m_{p+1}-1} z_{p+1}^{\ell} \cdot x_{\mathbf{i},\ell} \Bigr) = 0.$$

But for $\mathbf{i} \in I_p$, the element $\sum_{\ell=0}^{m_{p+1}-1} z_{p+1}^{\ell} \cdot x_{\mathbf{i},\ell}$ is in $\mathcal{L}_p \otimes_k k[z_{p+1}^{\pm 1}]$ and therefore we can write

$$\sum_{\ell=0}^{m_{p+1}-1} z_{p+1}^{\ell} \cdot x_{\mathbf{i},\ell} = \sum_{j\in\mathbb{Z}} y_{\mathbf{i},j} \otimes z_{p+1}^{j},$$

where each $y_{\mathbf{i},j}$ is in $\mathcal{L}_p$ and only finitely many of these elements are nonzero. Then $\sum_{\mathbf{i}\in I_p} z^{\mathbf{i}} \cdot \bigl(\sum_{j\in\mathbb{Z}} y_{\mathbf{i},j} \otimes z_{p+1}^{j} \bigr) = 0$, and so

$$\sum_{j\in\mathbb{Z}} \Bigl(\sum_{\mathbf{i}\in I_p} z^{\mathbf{i}} \cdot y_{\mathbf{i},j} \Bigr) \otimes z_{p+1}^{j} = 0.$$

Hence $\sum_{\mathbf{i}\in I_p} z^{\mathbf{i}} \cdot y_{\mathbf{i},j} = 0$ for each j and so by the induction hypothesis $y_{\mathbf{i},j} = 0$ for all $\mathbf{i} \in I_p$ and $j \in \mathbb{Z}$. So $\sum_{\ell=0}^{m_{p+1}-1} z_{p+1}^{\ell} \cdot x_{\mathbf{i},\ell} = 0$ for all $\mathbf{i} \in I_p$, and hence, by Lemma 4.6, $x_{\mathbf{i},\ell} = 0$ for all $\mathbf{i} \in I_p$ and $0 \le \ell \le m_{p+1} - 1$. □

If $\mathcal{A}$ is unital and associative, then $\mathcal{L} = L(\mathcal{A}, \Sigma_1, \dots, \Sigma_n)$ is a unital associative subalgebra of $\mathcal{A} \otimes_k S^{\otimes n}$ and hence $\mathcal{A} \otimes_k S^{\otimes n}$ is an $\mathcal{L}$-module (with action denoted by "·").

Corollary 5.8. *Suppose that $\mathcal{L} = L(\mathcal{A}, \Sigma_1, \dots, \Sigma_n)$ where $\mathcal{A}$ is unital and associative. Then $\mathcal{A} \otimes_k S^{\otimes n}$ is a free $\mathcal{L}$-module of rank $m_1 \dots m_n$ with basis*

$$\{ 1_{\mathcal{A}} \otimes z^{\mathbf{i}} \}_{\mathbf{i}\in I_n}.$$

Proof. This follows from Lemma 5.7 (with $p = n$) and the observation that

$$z^{\mathbf{i}} \cdot x = x \cdot (1_{\mathcal{A}} \otimes z^{\mathbf{i}})$$

for $x \in \mathcal{L}$ and $\mathbf{i} \in \mathbb{Z}^n$. (On the left of this equation · denotes the action of $S^{\otimes n}$ on $\mathcal{A} \otimes_k S^{\otimes n}$, whereas on the right · denotes the action of $\mathcal{L}$ on $\mathcal{A} \otimes_k S^{\otimes n}$.) □

6. The centroid of an iterated loop algebra

In this section, we give an explicit description of the centroid of an n-step loop algebra based on a pfgc algebra $\mathcal{A}$ as an n-step loop algebra based on $C(\mathcal{A})$.

Throughout the section *we assume that $\mathcal{L} = L(\mathcal{A}, \Sigma_1, \dots, \Sigma_n)$ is an n-step loop algebra based on a algebra $\mathcal{A}$ over k.* So we have algebras $\mathcal{L}_0, \dots, \mathcal{L}_n$ so that $\mathscr{L}_0 = \mathscr{A}$, $\mathscr{L}_n = \mathscr{L}$ and

$$\mathscr{L}_{p+1} = L(\mathscr{L}_p, \Sigma_{p+1}, z_{p+1})$$

for $0 \le p \le n-1$. As we observed in Remark 5.2, $\mathscr{L}_p$ is a subalgebra of $\mathscr{A} \otimes_k S^{\otimes p}$ for $0 \le p \le n$.

We next introduce some notation.

First let $0 \le p \le n$. Then $C(\mathscr{A}) \otimes_k S^{\otimes p}$ is a unital associative algebra and $\mathscr{A} \otimes_k S^{\otimes p}$ *is a* $C(\mathscr{A}) \otimes_k S^{\otimes p}$*-module* under the action "$\cdot$" defined by

$$(\chi \otimes z^{\mathbf{i}}) \cdot (a \otimes z^{\mathbf{j}}) = \chi(a) \otimes z^{\mathbf{i}+\mathbf{j}}.$$

We let $\bar{C}(\mathscr{L}_p)$ denote the stabilizer of $\mathscr{L}_p$ in $C(\mathscr{A}) \otimes_k S^{\otimes p}$ under this action. That is we let

$$\bar{C}(\mathscr{L}_p) := \{ u \in C(\mathscr{A}) \otimes_k S^{\otimes p} \mid u \cdot \mathscr{L}_p \subset \mathscr{L}_p \}.$$

Then $\bar{C}(\mathscr{L}_p)$ is a unital subalgebra of $C(\mathscr{A}) \otimes_k S^{\otimes p}$ and $\mathscr{L}_p$ is a $\bar{C}(\mathscr{L}_p)$-module. (For convenience, our notation suppresses the fact that $\bar{C}(\mathscr{L}_p)$ depends on $\mathscr{A}$, $\Sigma_1, \ldots, \Sigma_p$ and not just on the loop algebra $\mathscr{L}_p$.)

Next suppose that $0 \le p \le n-1$. Then Σ_{p+1} is a $\mathbb{Z}_{m_{p+1}}$-grading of the algebra $\mathscr{L}_p$ which we write as

$$\mathscr{L}_p = \bigoplus_{\bar{\imath} \in \mathbb{Z}_{m_{p+1}}} (\mathscr{L}_p)_{\bar{\imath}}.$$

We set

$$\bar{C}(\mathscr{L}_p)_{\bar{\imath}} := \{ u \in \bar{C}(\mathscr{L}_p) \mid u \cdot (\mathscr{L}_p)_{\bar{\jmath}} \subset (\mathscr{L}_p)_{\bar{\imath}+\bar{\jmath}} \text{ for all } \bar{\jmath} \in \mathbb{Z}_{m_{p+1}} \} \tag{6–1}$$

for $\bar{\imath} \in \mathbb{Z}_{m_{p+1}}$. We denote the collection $\{\bar{C}(\mathscr{L}_p)_{\bar{\imath}}\}_{\bar{\imath} \in \mathbb{Z}_{m_{p+1}}}$ by $\bar{C}(\Sigma_{p+1})$. We will see in Lemma 6.1(ii) below that $\bar{C}(\Sigma_{p+1})$ is a $\mathbb{Z}_{m_{p+1}}$-grading of $\bar{C}(\mathscr{L}_p)$.

Finally for $0 \le p \le n$ we define $\gamma_p : \bar{C}(\mathscr{L}_p) \to C(\mathscr{L}_p)$ by

$$\gamma_p(u)(x) = u \cdot x$$

for $u \in \bar{C}(\mathscr{L}_p)$, $x \in \mathscr{L}_p$, in which case γ_p is an algebra homomorphism. Note in particular that $\bar{C}(\mathscr{L}_0) = C(\mathscr{A})$ and γ_0 is the identity map.

Lemma 6.1. *Suppose that* $\mathscr{L} = L(\mathscr{A}, \Sigma_1, \ldots, \Sigma_n)$, *where* $\mathscr{A}$ *is a pfgc algebra.*

(i) *If* $0 \le p \le n$ *then* $\gamma_p : \bar{C}(\mathscr{L}_p) \to C(\mathscr{L}_p)$ *is an isomorphism of* k*-algebras.*

(ii) *If* $0 \le p \le n-1$ *then* $\bar{C}(\Sigma_{p+1})$ *is a* $\mathbb{Z}_{m_{p+1}}$*-grading of the algebra* $\bar{C}(\mathscr{L}_p)$ *and the map* γ_p *is an isomorphism of graded algebras.*

(iii) *If* $0 \le p \le n-1$ *then*

$$\bar{C}(\mathscr{L}_{p+1}) = L(\bar{C}(\mathscr{L}_p), \bar{C}(\Sigma_{p+1}), z_{p+1}). \tag{6–2}$$

Proof. (i) We first show that γ_p is injective for $0 \le p \le n$. To see this, suppose that $u \in \ker(\gamma_p)$. Then $u \cdot x = 0$ for all $x \in \mathcal{L}_p$, and so (since $\mathcal{L}_p$ spans $\mathcal{A} \otimes_k S^{\otimes p}$ over $S^{\otimes p}$ by Lemma 5.7) we have $u \cdot x = 0$ for all $x \in \mathcal{A} \otimes_k S^{\otimes p}$. This implies that $u = 0$.

Next we prove the bijectively of γ_p for $0 \le p \le n$ by induction on p. This is clear if $p = 0$ since γ_0 is the identity map. So we suppose that $0 \le p \le n-1$ and that γ_p is bijective. It is clear from this bijectivity and from the definitions of $\bar{C}(\mathcal{L}_{p+1})$ and $C(\mathcal{L}_{p+1})$ (see (6–1) and (4–2)) that

$$\gamma_p(\bar{C}(\mathcal{L}_{p+1})_{\bar{\imath}}) = C(\mathcal{L}_{p+1})_{\bar{\imath}}$$

for $\bar{\imath} \in \mathbb{Z}_{m_{p+1}}$. Hence $\bar{C}(\Sigma_{p+1})$ is a grading of the algebra $\bar{C}(\mathcal{L}_p)$ and $\gamma_p : \bar{C}(\mathcal{L}_p) \to C(\mathcal{L}_p)$ is a graded isomorphism. So γ_p induces an algebra isomorphism

$$L(\gamma_p) : L(\bar{C}(\mathcal{L}_p), \bar{C}(\Sigma_{p+1}), z_{p+1}) \to L(C(\mathcal{L}_p), C(\Sigma_{p+1}), z_{p+1}).$$

Consequently we have the composite algebra isomorphism
(6–3)

$$L(\bar{C}(\mathcal{L}_p), \bar{C}(\Sigma_{p+1}), z_{p+1}) \xrightarrow{L(\gamma_p)} L(C(\mathcal{L}_p), C(\Sigma_{p+1}), z_{p+1}) \xrightarrow{\psi_{\mathcal{L}_p,\Sigma_{p+1}}} C(\mathcal{L}_{p+1}),$$

where $\psi_{\mathcal{L}_p,\Sigma_{p+1}}$ is the isomorphism of Proposition 4.11. (Note that the proposition can be applied since $\mathcal{L}_p$ is a pfgc algebra by, Theorem 5.5(iii).) But

$$L(\bar{C}(\mathcal{L}_p), \bar{C}(\Sigma_{p+1}), z_{p+1}) \subset \bar{C}(\mathcal{L}_{p+1})$$

and one easily checks that the restriction

$$\gamma_{p+1}|_{L(\bar{C}(\mathcal{L}_p),\bar{C}(\Sigma_{p+1}),z_{p+1})} : L(\bar{C}(\mathcal{L}_p), \bar{C}(\Sigma_{p+1}), z_{p+1}) \to C(\mathcal{L}_{p+1}) \tag{6–4}$$

equals the composite map (6–3). Hence the restriction (6–4) of γ_{p+1} is bijective. Thus, since γ_{p+1} itself is injective, it follows that

$$L(\bar{C}(\mathcal{L}_p), \bar{C}(\Sigma_{p+1}), z_{p+1}) = \bar{C}(\mathcal{L}_{p+1})$$

and γ_{p+1} is bijective. So we have proved (i).

(ii) and (iii): These were proved in the argument for (i). □

Since $\mathcal{L} = \mathcal{L}_n$, we write $\bar{C}(\mathcal{L}) = \bar{C}(\mathcal{L}_n)$ and so

$$\bar{C}(\mathcal{L}) := \{u \in C(\mathcal{A}) \otimes_k S^{\otimes n} \mid u \cdot \mathcal{L} \subset \mathcal{L}\}.$$

Then $\bar{C}(\mathcal{L})$ is a unital subalgebra of $C(\mathcal{A}) \otimes_k S^{\otimes n}$, and $\mathcal{A} \otimes_k S^{\otimes n}$ is a $\bar{C}(\mathcal{L})$-module. We also write $\gamma_{\mathcal{L}} = \gamma_n$. Thus $\gamma_{\mathcal{L}} : \bar{C}(\mathcal{L}) \to C(\mathcal{L})$ is the k-algebra homomorphism (said to be canonical) defined by

$$\gamma_{\mathcal{L}}(u)(x) = u \cdot x$$

for $u \in \bar{C}(\mathcal{L})$, $x \in \mathcal{L}$.

Using Lemma 6.1 we can now give an explicit description of the centroid of an n-step loop algebra as an n-step loop algebra.

Theorem 6.2. *Suppose that $\mathcal{L} = L(\mathcal{A}, \Sigma_1, \ldots, \Sigma_n)$ is an n-step loop algebra based on a pfgc algebra $\mathcal{A}$. Then the canonical map $\gamma_{\mathcal{L}} : \bar{C}(\mathcal{L}) \to C(\mathcal{L})$ is an algebra isomorphism and we have*

$$\bar{C}(\mathcal{L}) = L(C(\mathcal{A}), \bar{C}(\Sigma_1), \ldots, \bar{C}(\Sigma_n)). \tag{6–5}$$

Proof. $\gamma_{\mathcal{L}}$ is an isomorphism by Lemma 6.1(i). Moreover (6–5) follows by repeated application of (6–2). □

Corollary 6.3. *Suppose that $\mathcal{L}$ is an n-step loop algebra based on a pfgc algebra $\mathcal{A}$. Then*

(i) *$C(\mathcal{A})$ and $C(\mathcal{L})$ are nonzero unital commutative associative algebras over k.*

(ii) *If $C(\mathcal{A})$ is an integral domain, then $C(\mathcal{L})$ is an integral domain.*

(iii) *If $C(\mathcal{A})$ is finitely generated as an algebra over k, then $C(\mathcal{L})$ is finitely generated as an algebra over k and* $\operatorname{Dim} C(\mathcal{L}) = \operatorname{Dim} C(\mathcal{A}) + n$.

Proof. (i) Since $\mathcal{A}$ is pfgc, we know that $\mathcal{L}$ is pfgc by Theorem 5.5(iii). Hence $C(\mathcal{A})$ and $C(\mathcal{L})$ are nonzero unital commutative associative algebras by Proposition 2.11(i).

(ii) and (iii): We know by Theorem 6.2 that $C(\mathcal{L})$ is isomorphic to an n-step loop algebra based on $C(\mathcal{A})$. Thus (ii) and (iii) follow from Theorem 5.5 (viii) and (x) respectively. □

If $\mathcal{A}$ is a finite dimensional central simple algebra over k, then $C(\mathcal{A}) = k$ and $\mathcal{A}$ is a pfgc algebra. Hence we have the following consequence of Corollary 6.3:

Corollary 6.4. *Suppose that $\mathcal{L}$ is an n-step loop algebra based on a finite dimensional central simple algebra $\mathcal{A}$ over k. Then $C(\mathcal{L})$ is an integral domain, $C(\mathcal{L})$ is finitely generated as an algebra over k, and* $\operatorname{Dim} C(\mathcal{L}) = n$. *Consequently, if $\mathcal{L}'$ is an n'-step loop algebra based on a finite dimensional central simple algebra $\mathcal{A}'$ over k, then*

$$\mathcal{L} \simeq_k \mathcal{L}' \implies n = n'.$$

Remark 6.5. Suppose that $\mathcal{L}$ is an n-step loop algebra based on a finite dimensional central simple algebra $\mathcal{A}$ over k. Then $C(\mathcal{A}) \otimes_k S^{\otimes n} = k \otimes_k S^{\otimes n} = S^{\otimes n}$ and so

$$C(\mathcal{L}) \overset{\gamma_{\mathcal{L}}}{\simeq_k} \bar{C}(\mathcal{L}) = \{ u \in S^{\otimes n} : u \cdot \mathcal{L} \subset \mathcal{L} \}.$$

This fact can be used to explicitly compute $C(\mathcal{L})$ in examples.

Corollary 6.6. *Suppose that $\mathscr{L} = M(\mathscr{A}, \sigma_1, \dots, \sigma_n)$ is a multiloop algebra based on a finite dimensional central simple algebra $\mathscr{A}$ over k, where $\sigma_1, \dots, \sigma_n$ are commuting finite order automorphisms of $\mathscr{A}$ with periods $m_1, \dots, m_n$ respectively. Then*

$$\bar{C}(\mathscr{L}) = k[(z_1^{m_1})^{\pm 1}, \dots, (z_n^{m_n})^{\pm 1}], \tag{6–6}$$

and so $C(\mathscr{L})$ is isomorphic to the algebra of Laurent polynomials in n-variables over k.

Proof. Recall (using the notation of Example 5.4) that

$$\mathscr{L} = \bigoplus_{(i_1,\dots,i_n)\in\mathbb{Z}^n} \mathscr{A}_{\bar{\imath}_1,\dots,\bar{\imath}_n} \otimes_k z_1^{i_1} \dots z_n^{i_n},$$

and so the inclusion "$\supset$" in (6–6) is clear. For the inclusion "$\subset$", let $u \in \bar{C}(\mathscr{L})$. Now $S^{\otimes n}$ is naturally $\mathbb{Z}^n$-graded and it is clear that $\bar{C}(\mathscr{L})$ is a graded subalgebra. Hence we can assume that $u = z_1^{j_1} \dots z_n^{j_n}$, where $(j_1, \dots, j_n) \in \mathbb{Z}^n$. But then $\mathscr{A}_{\bar{\imath}_1,\dots,\bar{\imath}_n} \subset \mathscr{A}_{\bar{\imath}_1+\bar{\jmath}_1,\dots,\bar{\imath}_n+\bar{\jmath}_n}$ for all $(i_1, \dots, i_n) \in \mathbb{Z}^n$ and so $(\bar{\jmath}_1, \dots, \bar{\jmath}_n) = (\bar{0}, \dots, \bar{0})$. □

7. Untwisting iterated loop algebras

In this section we show that any n-step loop algebra based on a pfgc algebra can be untwisted by an extension of the centroid of $\mathscr{L}$ that is free of finite rank.

Suppose again throughout the section that $\mathscr{L} = L(\mathscr{A}, \Sigma_1, \dots, \Sigma_n)$ *is an n-step loop algebra based on an algebra $\mathscr{A}$ over k.* We use the notation of the previous section.

It will be convenient to work with $\bar{C}(\mathscr{L})$ rather than $C(\mathscr{L})$ (although one could use $\gamma_{\mathscr{L}}$ to identify these algebras using Theorem 6.2 and avoid this distinction). Note that $\bar{C}(\mathscr{L})$ is a subalgebra of $C(\mathscr{A}) \otimes_k S^{\otimes n}$, and so $C(\mathscr{A}) \otimes_k S^{\otimes n}/\bar{C}(\mathscr{L})$ is a ring extension. This is the extension that we use to untwist $\mathscr{L}$.

We define

$$\omega_{\mathscr{L}} : \mathscr{L} \otimes_{\bar{C}(\mathscr{L})} (C(\mathscr{A}) \otimes_k S^{\otimes n}) \to \mathscr{A} \otimes_k S^{\otimes n}$$

by

$$\omega_{\mathscr{L}}(x \otimes u) = u \cdot x$$

for $x \in \mathscr{L}$ and $u \in C(\mathscr{A}) \otimes_k S^{\otimes n}$. One easily checks that $\omega_{\mathscr{L}}$ is a well-defined $C(\mathscr{A}) \otimes_k S^{\otimes n}$-algebra homomorphism which we call canonical.

Our untwisting theorem is the following:

Theorem 7.1. *Suppose that $\mathscr{L} = L(\mathscr{A}, \Sigma_1, \dots, \Sigma_n)$ is an n-step loop algebra based on a pfgc algebra $\mathscr{A}$, where Σ_p has modulus m_p for $1 \le p \le n$. Then*

(i) *$C(\mathscr{A})\otimes_k S^{\otimes n}$ is a free $\bar{C}(\mathscr{L})$-module of rank $m_1 \dots m_n$ with basis*

$$\{\, 1_{C(\mathscr{A})} \otimes z^{\mathbf{i}} \mid \mathbf{i} \in I_n \,\}.$$

(ii) *The canonical map $\omega_{\mathscr{L}}$ is an isomorphism and so*

$$\mathscr{L} \otimes_{\bar{C}(\mathscr{L})} (C(\mathscr{A}) \otimes_k S^{\otimes n}) \simeq_{C(\mathscr{A})\otimes_k S^{\otimes n}} \mathscr{A} \otimes_k S^{\otimes n}. \tag{7–1}$$

Proof. (i) Since $\bar{C}(\mathscr{L})$ is an n-step loop algebra based on $C(\mathscr{A})$ by Theorem 6.2, statement (i) follows from Corollary 5.8.

(ii) First $\mathscr{L}$ spans $\mathscr{A} \otimes_k S^{\otimes n}$ over $S^{\otimes n}$ by Lemma 5.7, and so $\mathscr{L}$ spans $\mathscr{A} \otimes_k S^{\otimes n}$ over $C(\mathscr{A}) \otimes S^{\otimes n}$. Thus $\omega_{\mathscr{L}}$ is surjective.

To show that $\omega_{\mathscr{L}}$ is injective, let $x \in \ker(\omega_{\mathscr{L}})$. Then, in particular, x is an element of $\mathscr{L} \otimes_{\bar{C}(\mathscr{L})} (C(\mathscr{A}) \otimes_k S^{\otimes n})$. Now since $\bar{C}(\mathscr{L})$ is an n-step loop algebra based on $C(\mathscr{A})$, it follows from Lemma 5.7 that every element of $C(\mathscr{A}) \otimes_k S^{\otimes n}$ can be written as a sum of elements of the form $z^{\mathbf{i}} \cdot u$, where $\mathbf{i} \in I_n$ and $u \in \bar{C}(\mathscr{L})$. But $z^{\mathbf{i}} \cdot u = u \cdot (1_{C(\mathscr{A})} \otimes z^{\mathbf{i}})$. Thus from the balanced property in the tensor product $\mathscr{L} \otimes_{\bar{C}(\mathscr{L})} (C(\mathscr{A}) \otimes_k S^{\otimes n})$, it follows that x can be written in the form

$$x = \sum_{\mathbf{i} \in I_n} x_{\mathbf{i}} \otimes (1_{C(\mathscr{A})} \otimes z^{\mathbf{i}}),$$

where $x_{\mathbf{i}} \in \mathscr{L}$ for all $\mathbf{i}$. Applying $\omega_{\mathscr{L}}$ to this expression yields $\sum_{\mathbf{i} \in I_n} z^{\mathbf{i}} \cdot x_{\mathbf{i}} = 0$, and so $x_{\mathbf{i}} = 0$ for all $\mathbf{i} \in I_n$ by Lemma 5.7. Thus $x = 0$ and $\omega_{\mathscr{L}}$ is injective. □

Remark 7.2. Suppose that $\mathscr{L}$ is an n-step loop algebra based on a pfgc algebra $\mathscr{A}$.

(i) We can use the canonical isomorphism $\gamma_{\mathscr{L}} : \bar{C}(\mathscr{L}) \to C(\mathscr{L})$ of Theorem 6.2 to identify the algebras $\bar{C}(\mathscr{L})$ and $C(\mathscr{L})$. This identification is compatible with the actions of these algebras on $\mathscr{L}$ and it gives $C(\mathscr{A}) \otimes_k S^{\otimes n}$ the structure of a $C(\mathscr{L})$-module. Then (7–1) can be restated as

$$\mathscr{L} \otimes_{C(\mathscr{L})} \big(C(\mathscr{A}) \otimes_k S^{\otimes n}\big) \simeq_{C(\mathscr{A})\otimes_k S^{\otimes n}} \mathscr{A} \otimes_k S^{\otimes n}. \tag{7–1′}$$

Since $\mathscr{A} \otimes_k S^{\otimes n}$ is the untwisted n-step loop algebra based on $\mathscr{A}$, Theorem 7.1 tells us that $\mathscr{L}$ *is untwisted by a free base ring extension of rank $m_1 \dots m_n$ of the centroid of $\mathscr{L}$.*

(ii) Also observe that the algebras $\mathscr{A} \otimes_k S^{\otimes n}$ and $\mathscr{A} \otimes_{C(\mathscr{A})} \big(C(\mathscr{A}) \otimes_k S^{\otimes n}\big)$ are canonically isomorphic as $C(\mathscr{A}) \otimes_k S^{\otimes n}$-algebras. Thus the isomorphism (7–1′) can be further restated as

$$\mathscr{L} \otimes_{C(\mathscr{L})} \big(C(\mathscr{A}) \otimes_k S^{\otimes n}\big) \simeq_{C(\mathscr{A})\otimes_k S^{\otimes n}} \mathscr{A} \otimes_{C(\mathscr{A})} \big(C(\mathscr{A}) \otimes_k S^{\otimes n}\big). \tag{7–1″}$$

Theorem 7.1 can be used to compare properties of an iterated loop algebra as a module or algebra over its centroid with corresponding properties of the base algebra over its centroid. We now indicate an example of this sort of argument.

Corollary 7.3. *Let $\mathcal{L}$ be an n-step loop algebra based on a pfgc algebra $\mathcal{A}$. If $\mathcal{A}$ is a projective $C(\mathcal{A})$-module then $\mathcal{L}$ is a finitely generated projective $C(\mathcal{L})$-module*

Proof. As in Remark 7.2, we identify $\bar{C}(\mathcal{L})$ and $C(\mathcal{L})$ using $\gamma_{\mathcal{L}}$. By axiom P2 of pfgc algebras and the present assumption, $\mathcal{A}$ is a finitely generated projective $C(\mathcal{A})$-module. Hence $\mathcal{A} \otimes_{C(\mathcal{A})} \big(C(\mathcal{A}) \otimes_k S^{\otimes n}\big)$ is a finitely generated projective $C(\mathcal{A}) \otimes_k S^{\otimes n}$-module. Thus by (7–1″), $\mathcal{L} \otimes_{C(\mathcal{L})} \big(C(\mathcal{A}) \otimes_k S^{\otimes n}\big)$ is a finitely generated projective $C(\mathcal{A}) \otimes_k S^{\otimes n}$-module. But the extension $C(\mathcal{A}) \otimes_k S^{\otimes n}/C(\mathcal{L})$ is free of finite rank by Theorem 7.1(i), and so it is faithfully flat. The result now follows from [Bourbaki 1972, Ch. I, § 3.6, Prop. 12] □

In the same spirit, we now describe an application of Theorem 7.1 for associative algebras. For this purpose we first recall some definitions and basic facts about Azumaya algebras. A unital associative algebra $\mathcal{D}$ over a ring B is called an *Azumaya algebra* over B if $\mathcal{D}$ is central and separable over B (see for example [Knus and Ojanguren 1974, §5]). If $\mathcal{D}$ is an Azumaya algebra over B, then $\mathcal{D}$ is a finitely generated projective B-module [Knus and Ojanguren 1974, Théorème 5.1], and so $\mathcal{D}_{\mathfrak{m}}$ is a free $B_{\mathfrak{m}}$-module of finite rank $r_{\mathfrak{m}}$ for each maximal ideal $\mathfrak{m}$ of B. $\mathcal{D}$ is said to have *constant rank* r over B if $r_{\mathfrak{m}} = r$ for all such $\mathfrak{m}$ [Bourbaki 1972, §II.5.3]. It is known that if $\mathcal{D}$ is a unital associative algebra over a ring B and ℓ is a positive integer then

(7–2) $\mathcal{D}$ is an Azumaya algebra of constant rank ℓ^2 over B if and only if there exists a faithfully flat extension B'/B of rings so that $\mathcal{D} \otimes_B B'$ is isomorphic as a B'-algebra to the algebra $M_\ell(B')$ of $\ell \times \ell$-matrices over B'.

In that case we will say that $\mathcal{D}$ is *split* by the extension B'/B. Indeed the implication "⇒" of (7–2) is Corollary 6.7 of [Knus and Ojanguren 1974]. For the converse, the algebra $M_\ell(B')$ is an Azumaya algebra of constant rank ℓ^2 over B', and hence $\mathcal{D}$ is an Azumaya algebra of constant rank ℓ^2 over B since the extension B'/B is faithfully flat (see Lemma 5.1.9(1) in [Knus 1991] and Exercise 8 in [Bourbaki 1972, § II.5]).

Corollary 7.4. *Suppose that $\mathcal{L}$ is an n-step loop algebra based on the matrix algebra $M_\ell(k)$ over k. Then $\mathcal{L}$ is a prime Azumaya algebra of finite rank ℓ^2 over its centroid $C(\mathcal{L})$ which is split by the extension $S^{\otimes n}/C(\mathcal{L})$.*

Proof. Let $\mathcal{A} = M_\ell(k)$. Then $\mathcal{A}$ is a prime unital associative algebra over k and hence so is $\mathcal{L}$ (by Theorem 5.5). Also $C(\mathcal{A}) = k$ and so $C(\mathcal{A}) \otimes_k S^{\otimes n} = S^{\otimes n}$ as in Remark 6.5. Thus by (7–1′) we have

$$\mathcal{L} \otimes_{C(\mathcal{L})} S^{\otimes n} \simeq_{S^{\otimes n}} M_\ell(k) \otimes_k S^{\otimes n} \simeq_{S^{\otimes n}} M_\ell(S^{\otimes n}).$$

Our conclusion now follows from (7–2), since the extension $S^{\otimes n}/C(\mathcal{L})$ is faithfully flat. □

8. Permanence of type

There is a classical notion of type for simple pfgc Lie algebras in characteristic zero (see Example 8.1 below). This notion can easily be extended using the central closure to include prime pfgc Lie algebras in characteristic 0. It will be a consequence of the results in this section that type is preserved under the loop construction (that is type is permanent).

An analogous notion of type can be defined for many other important classes of prime pfgc algebras besides Lie algebras in characteristic 0. Moreover, since algebras in these classes arise naturally as coordinate algebras in the study of Lie algebras, and in particular in the study of extended affine Lie algebras, it is desirable to include these classes in our discussion of type. This generality requires almost no extra effort once the appropriate definitions are made. That being said, the reader can safely choose to assume throughout this section that the base algebras, and hence their loop algebras, are Lie algebras in characteristic 0.

We begin by recalling the classical notion of type for simple pfgc Lie algebras in characteristic 0.

Example 8.1. Suppose that k has characteristic 0. Let $\mathcal{A}$ be a simple pfgc Lie algebra over k (or equivalently let $\mathcal{A}$ be a simple Lie algebra over k that is finitely generated as a module over its centroid). Then, since $\mathcal{A}$ is simple, it is easily checked that $C(\mathcal{A})$ is a field. Hence, if we let K be an algebraic closure of $C(\mathcal{A})$, the algebra $\mathcal{A}\otimes_{C(\mathcal{A})}K$ is a finite dimensional simple Lie algebra over K by Lemma 2.4. The *type* of $\mathcal{A}$ is defined in [Jacobson 1962, Ch. X, § 3] to coincide with the type of the root system of the K-algebra $\mathcal{A}\otimes_{C(\mathcal{A})}K$ relative to any Cartan subalgebra.

In order to extend this notion to other classes of algebras, we need to introduce some terminology.

Definition 8.2. Recall that a *variety* over k is a class $\mathbb{V}_k$ of algebras over k that is defined by a set of identities in the free nonassociative algebra $k_{\mathrm{na}}[x_1, x_2, \dots]$ in countably many symbols [Zhevlakov et al. 1982, § 1.2]. A variety $\mathbb{V}_k$ over k is said to be *homogeneous* if the ideal in $k_{\mathrm{na}}[x_1, x_2, \dots]$ consisting of all identities satisfied by all algebras in $\mathbb{V}_k$ is homogeneous. Algebras in a variety $\mathbb{V}_k$ over k will be simply called $\mathbb{V}_k$*-algebras*.

A very familiar example is the variety $\mathbb{V}_k$ of Lie algebras over k which is defined by the identities x_1x_1 and $(x_1x_2)x_3+(x_2x_3)x_1+(x_3x_1)x_2$. In that case $\mathbb{V}_k$ is homogeneous [Zhevlakov et al. 1982, § 1.4], and a $\mathbb{V}_k$-algebra is just a Lie algebra over k.

Remark 8.3. Suppose that $\mathbb{V}_k$ is a *variety* over k. Suppose that B is a unital associative commutative k-algebra. A *$\mathbb{V}_k$-algebra over B* will mean a B-algebra $\mathscr{A}$ with the property that $\mathscr{A}$, when regarded as an algebra over k, is in $\mathbb{V}_k$. In other words, a $\mathbb{V}_k$-algebra over B is a B-algebra that satisfies the identities (which are identities with coefficients from our fixed base field k) that define $\mathbb{V}_k$.

The homogeneity assumption is important for our purposes since homogeneous varieties are closed under base ring extension.

Lemma 8.4. *Suppose that $\mathbb{V}_k$ is a homogeneous variety over k and $B \to B'$ is a homomorphism of unital commutative associative k-algebras. If $\mathscr{A}$ is a $\mathbb{V}_k$-algebra over B then $\mathscr{A} \otimes_B B'$ is a $\mathbb{V}_k$-algebra over B'.*

Proof. This follows the proof of Theorem 6 in [Zhevlakov et al. 1982, § 1.4]. □

Corollary 8.5. *Suppose that $\mathbb{V}_k$ is a homogeneous variety over k. If $\mathscr{L}$ is an n-step loop algebra based on a $\mathbb{V}_k$-algebra $\mathscr{A}$, then $\mathscr{L}$ is a $\mathbb{V}_k$-algebra.*

Proof. By Lemma 8.4, $\mathscr{A} \otimes_k S^{\otimes n}$ is a $\mathbb{V}_k$-algebra. Hence so is its subalgebra $\mathscr{L}$. □

We will be interested in homogeneous varieties $\mathbb{V}_k$ that satisfy the following axiom:

(S) If K/k is a field extension and $\mathscr{A}$ is a finite dimensional semiprime $\mathbb{V}_k$-algebra over K then $\mathscr{A}$ is a direct sum of simple algebras over K.

Example 8.6. In each of the following cases, the variety $\mathbb{V}_k$ is homogeneous and satisfies axiom (S):

(a) $\operatorname{char}(k) = 0$, $\mathbb{V}_k$ is the variety of Lie algebras.

(b) $\mathbb{V}_k$ is the variety of associative algebras.

(c) $\mathbb{V}_k$ is the variety of commutative associative algebras.

(d) $\mathbb{V}_k$ is the variety of alternative algebras.

(e) $\operatorname{char}(k) \neq 2$, $\mathbb{V}_k$ is the variety of Jordan algebras.

Indeed the fact that these varieties are homogeneous is proved in [Zhevlakov et al. 1982, § 1.4]. Moreover axiom (S) follows from the structure theory for the variety $\mathbb{V}_k$ in each case. For example, in case (a), suppose that K/k is a field extension and $\mathscr{A}$ is a finite dimensional semiprime Lie algebra over K. If the radical $\mathscr{R}$ of $\mathscr{A}$ is nonzero, then the last nonzero term in the derived series for $\mathscr{R}$ has trivial multiplication, contrary to the assumption that $\mathscr{A}$ is semiprime. So $\mathscr{R} = 0$ and hence $\mathscr{A}$ is the direct sum of simple algebras [Jacobson 1962, § III.4]. Similarly we can use (for example) [Zhevlakov et al. 1982, §12.2, Theorem 3] in cases (b), (c) and (d) and [Jacobson 1968, § V.2, Lemma 2 and § V.5, Corollary 2] in case (e) to verify axiom (S).

The reason for our interest in Axiom (S) is that it allows us to prove the following proposition.

Proposition 8.7. *Let $\mathbb{V}_k$ be a homogeneous variety over k that satisfies axiom* (S)*. Suppose that $\mathscr{A}$ is a prime pfgc $\mathbb{V}_k$-algebra over k and let $\widetilde{C(\mathscr{A})}$ be the quotient field of $C(\mathscr{A})$. Then the central closure $\widetilde{\mathscr{A}} := \mathscr{A} \otimes_{C(\mathscr{A})} \widetilde{C(\mathscr{A})}$ of $\mathscr{A}$ is a finite dimensional central simple $\mathbb{V}_k$-algebra over $\widetilde{C(\mathscr{A})}$.*

Proof. By Proposition 3.5, $\widetilde{\mathscr{A}}$ is a prime pfgc algebra over k and hence also over $\widetilde{C(\mathscr{A})}$ (see Remark 2.10 and Lemma 3.2). Also, by Lemma 8.4, $\widetilde{\mathscr{A}}$ is a $\mathbb{V}_k$-algebra. Hence, by axiom (S), $\widetilde{\mathscr{A}}$ is the direct sum of simple algebras over $\widetilde{C(\mathscr{A})}$. Since $\widetilde{\mathscr{A}}$ is prime, there is only one summand in this sum. Thus $\widetilde{\mathscr{A}}$ is a simple algebra over $\widetilde{C(\mathscr{A})}$. Finally, by Proposition 3.5, $\widetilde{\mathscr{A}}$ is central and finite dimensional over $\widetilde{C(\mathscr{A})}$. □

Remark 8.8. Suppose that $\mathscr{A}$ is as in Proposition 8.7. Then in the terminology of [Polikarpov and Shestakov 1990, § 1], Proposition 8.7 says that $\mathscr{A}$ is a *central order* in the finite dimensional central simple algebra $\widetilde{\mathscr{A}}$.

We will also need a set $\mathbb{X}_k$ of algebras over k that play the role of the split simple Lie algebras over k.

Definition 8.9. Suppose that $\mathbb{V}_k$ is a homogeneous variety over k. A *set of archetypes* for $\mathbb{V}_k$ is a set $\mathbb{X}_k$ of finite dimensional central simple $\mathbb{V}_k$-algebras over k such that the following axioms hold:

(A1) If K/k is an algebraically closed field extension and $\mathscr{A}$ is a finite dimensional central simple $\mathbb{V}_k$-algebra over K then there exists $\mathscr{X} \in \mathbb{X}_k$ so that $\mathscr{A} \simeq_K \mathscr{X} \otimes_k K$.

(A2) If K/k is a field extension and $\mathscr{X}, \mathscr{X}' \in \mathbb{X}_k$ then

$$\mathscr{X} \otimes_k K \simeq_K \mathscr{X}' \otimes_k K \implies \mathscr{X} = \mathscr{X}'$$

In particular, the elements of $\mathbb{X}_k$ are pairwise nonisomorphic over k.

Example 8.10. In each of the cases (a)-(e) in Example 8.6 there is a natural choice for a set $\mathbb{X}_k$ of archetypes:

(a) $\operatorname{char}(k) = 0$, $\mathbb{V}_k$ is the variety of Lie algebras and $\mathbb{X}_k = \{ \mathscr{X}_\Pi \}$, where Π runs through all connected Dynkin diagrams (up to isomorphism) and $\mathscr{X}_\Pi$ denotes the split simple Lie algebra with Dynkin diagram Π [Jacobson 1962, § VII.4].

(b) $\mathbb{V}_k$ is the variety of associative algebras and $\mathbb{X}_k = \{ M_\ell(k) \mid \ell \geq 1 \}$, where $M_\ell(k)$ is the algebra of $\ell \times \ell$-matrices over k.

(c) $\mathbb{V}_k$ is the variety of commutative associative algebras and $\mathbb{X}_k = \{ k \}$.

(d) $\mathbb{V}_k$ is the variety of alternative algebras and $\mathbb{X}_k = \{ M_\ell(k) \mid \ell \geq 1 \} \cup \{ \mathbb{O} \}$, where $\mathbb{O}$ denotes the split Cayley–Dickson (octonion) algebra [Zhevlakov et al. 1982, § 2.4].

(e) $\operatorname{char}(k) \neq 2$, $\mathbb{V}_k$ is the variety of Jordan algebras and $\mathbb{X}_k$ is the set consisting of the following algebras: k; the Jordan algebra constructed from a nondegenerate symmetric bilinear form with matrix $\operatorname{diag}(1, -1, \dots, 1, -1)$ on a 2ℓ-dimensional space over k, $\ell \geq 1$; the Jordan algebra constructed from a nondegenerate symmetric bilinear form with matrix $\operatorname{diag}(1, -1, \dots, 1, -1, 1)$ on a $2\ell + 1$-dimensional space over k, $\ell \geq 1$; the algebra of $\ell \times \ell$ hermitian matrices with coordinates from the split composition algebras of dimension 1, 2 and 4, $\ell \geq 3$; and the algebra of 3×3 hermitian matrices over $\mathbb{O}$ [Jacobson 1968, § 1.4 and 4.3].

The fact that $\mathbb{X}_k$ satisfies axioms (A1) and (A2) follows from the classification of finite dimensional central simple algebras over algebraically closed fields in each case. See for example [Jacobson 1962, § IV.3, Theorem 3] in case (a), [Zhevlakov et al. 1982, § 12.2, Theorem 3] in cases (b), (c) and (d), and [Jacobson 1968, § V.6, Corollary 2] in case (e).

Remark 8.11. A homogeneous variety $\mathbb{V}_k$ over k may possess more than one set of archetypes. For example if $k = \mathbb{R}$ and $\mathbb{V}_k$ is the variety of Lie algebras over k as in Example 8.10(a) above, then an alternate choice of a set of archetypes is the set $\mathbb{X}_k = \{\mathscr{C}_\Pi\}$, where Π runs through all connected Dynkin diagrams (up to isomorphism) and $\mathscr{C}_\Pi$ denotes the compact real Lie algebra whose complexification is the simple Lie algebra with Dynkin diagram Π.

Assumption. *For the rest of this section we assume that $\mathbb{V}_k$ is a homogeneous variety over k that satisfies axiom* (S), *and that there exists a set $\mathbb{X}_k$* (*which we fix*) *of archetypes for $\mathbb{V}_k$.*

We can now prove the proposition that allows us to define the notion of type.

Proposition 8.12. *Suppose that $\mathscr{A}$ is a prime pfgc $\mathbb{V}_k$-algebra over k. If*

$$C(\mathscr{A}) \hookrightarrow K$$

is a unital k-algebra monomorphism of $C(\mathscr{A})$ into an algebraically closed field extension K of k (*such a monomorphism exists since $C(\mathscr{A})$ is an integral domain*), *then there exists a unique $\mathscr{X} \in \mathbb{X}_k$ so that*

$$\mathscr{A} \otimes_{C(\mathscr{A})} K \simeq_K \mathscr{X} \otimes_k K, \tag{8–1}$$

where on the left K is regarded as an algebra over $C(\mathscr{A})$ using the given monomorphism. Moreover, $\mathscr{X}$ is independent of the choice of the k-algebra monomorphism $C(\mathscr{A}) \hookrightarrow K$.

Proof. First let L be an algebraic closure of $\widetilde{C(\mathcal{A})}$. By Proposition 8.7, $\widetilde{\mathcal{A}}$ is a finite dimensional central simple $\mathbb{V}_k$-algebra over $\widetilde{C(\mathcal{A})}$. Therefore, by Lemma 2.4, $\widetilde{\mathcal{A}} \otimes_{\widetilde{C(\mathcal{A})}} L$ is a finite dimensional central simple algebra over L. So, by Lemma 8.4, $\widetilde{\mathcal{A}} \otimes_{\widetilde{C(\mathcal{A})}} L$ is a finite dimensional central simple $\mathbb{V}_k$-algebra over L. Thus, by axiom (A1) (see Definition 8.9), there exists $\mathcal{X} \in \mathbb{X}_k$ so that $\widetilde{\mathcal{A}} \otimes_{\widetilde{C(\mathcal{A})}} L \simeq_L \mathcal{X} \otimes_k L$. Then

$$\mathcal{A} \otimes_{C(\mathcal{A})} L \simeq_L (\mathcal{A} \otimes_{C(\mathcal{A})} \widetilde{C(\mathcal{A})}) \otimes_{\widetilde{C(\mathcal{A})}} L = \widetilde{\mathcal{A}} \otimes_{\widetilde{C(\mathcal{A})}} L \simeq_L \mathcal{X} \otimes_k L. \tag{8–2}$$

Now let $C(\mathcal{A}) \hookrightarrow K$ be an arbitrary unital k-algebra monomorphism of $C(\mathcal{A})$ into an algebraically closed extension K of k. This extends to a unital k-algebra monomorphism $\widetilde{C(\mathcal{A})} \hookrightarrow K$ which in turns extends to a unital k-algebra monomorphism $L \hookrightarrow K$. We use this latter monomorphism to identify L as a subfield of K. Then using (8–2) we have

$$\mathcal{A} \otimes_{C(\mathcal{A})} K \simeq_K (\mathcal{A} \otimes_{C(\mathcal{A})} L) \otimes_L K \simeq_K (\mathcal{X} \otimes_k L) \otimes_L K \simeq_K \mathcal{X} \otimes_k K.$$

This shows the existence of an element $\mathcal{X} \in \mathbb{X}_k$ satisfying (8–1). The uniqueness follows from Axiom (A2).

Finally if $C(\mathcal{A}) \hookrightarrow K'$ is another unital k-algebra monomorphism of $C(\mathcal{A})$ into an algebraically closed extension K' of k, then the argument just given using (8–2) shows that $\mathcal{A} \otimes_{C(\mathcal{A})} K' \simeq_{K'} \mathcal{X} \otimes_k K'$. □

Definition 8.13. Let $\mathcal{A}$ be a prime pfgc $\mathbb{V}_k$-algebra over k. The element $\mathcal{X} \in \mathbb{X}_k$ described in Proposition 8.12 is called the *type* of $\mathcal{A}$ (relative to $\mathbb{X}_k$) and denoted by $t(\mathcal{A})$. We also sometimes refer to $t(\mathcal{A})$ as the *absolute type* of $\mathcal{A}$ since it is determined by extending the base ring $C(\mathcal{A})$ to an algebraically closed field.

Example 8.14. Let $\mathrm{char}(k) = 0$, let $\mathbb{V}_k$ be the variety of Lie algebras, and let $\mathbb{X}_k = \{\mathcal{X}_\Pi\}$ be as in Example 8.10(a). If we identify $\mathcal{X}_\Pi$ with the diagram Π, then Definition 8.13 assigns to each prime pfgc $\mathbb{V}_k$-algebra $\mathcal{A}$ a connected Dynkin diagram $t(\mathcal{A})$. (If $\mathcal{A}$ is a simple pfgc algebra, this is exactly what was done in Example 8.1.)

The following result tells us that type is an isomorphism invariant for prime pfgc algebras.

Proposition 8.15. *Suppose that $\mathcal{A}$ and $\mathcal{A}'$ are prime pfgc $\mathbb{V}_k$-algebras over k. Then*

$$\mathcal{A} \simeq_k \mathcal{A}' \implies t(\mathcal{A}) = t(\mathcal{A}').$$

Proof. Let $\varphi : C(\mathcal{A}') \hookrightarrow K$ be a unital k-algebra monomorphism of $C(\mathcal{A}')$ into an algebraically closed field extension K of k. Denote the resulting action of $C(\mathcal{A}')$ on K by $(\chi', \alpha) \mapsto \chi' \cdot \alpha$.

Fix a k-algebra isomorphism $\rho : \mathcal{A} \to \mathcal{A}'$. Then ρ induces an isomorphism $C(\rho) : C(\mathcal{A}) \to C(\mathcal{A}')$ by Lemma 2.2. So the composite map $\varphi \circ C(\rho) : C(\mathcal{A}) \to K$ is a unital k-algebra monomorphism which we use to view K as an algebra over $C(\mathcal{A})$. The resulting action of $C(\mathcal{A})$ on K is given by

$$\chi \cdot \alpha = C(\rho)(\chi) \cdot \alpha.$$

for $\chi \in C(\mathcal{A})$ and $\alpha \in K$.

The biadditive map $\widetilde{\rho} : \mathcal{A} \times K \to \mathcal{A}' \otimes_{C(\mathcal{A}')} K$ satisfying $\widetilde{\rho} : (a, \alpha) \mapsto \rho(a) \otimes \alpha$ is then $C(\mathcal{A})$-balanced. Indeed if $\chi \in C(\mathcal{A})$, $a \in \mathcal{A}$ and $\alpha \in K$ we have

$$\begin{aligned}\widetilde{\rho}\big(\chi(a), \alpha\big) &= \rho\big(\chi(a)\big) \otimes \alpha = C(\rho)(\chi)\big(\rho(a)\big) \otimes \alpha \\ &= \rho(a) \otimes \big(C(\rho)(\chi) \cdot \alpha\big) = \rho(a) \otimes \chi \cdot \alpha = \widetilde{\rho}(a, \chi \cdot \alpha).\end{aligned}$$

Thus $\widetilde{\rho}$ induces a k-linear map $\mathcal{A} \otimes_{C(\mathcal{A})} K \to \mathcal{A}' \otimes_{C(\mathcal{A}')} K$ so that $a \otimes \alpha \mapsto \rho(a) \otimes \alpha$ for $a \in \mathcal{A}$ and $\alpha \in K$. This map is clearly a homomorphism of K-algebras. In a similar fashion we obtain a homomorphism of K-algebras $\mathcal{A}' \otimes_{C(\mathcal{A}')} K \to \mathcal{A} \otimes_{C(\mathcal{A})} K$ so that $a' \otimes \alpha \mapsto \rho^{-1}(a') \otimes \alpha$ for $a' \in \mathcal{A}'$ and $\alpha \in K$. These maps are inverses of each other and so we have

$$\mathcal{A} \otimes_{C(\mathcal{A})} K \simeq_K \mathcal{A}' \otimes_{C(\mathcal{A}')} K.$$

Thus, $\mathcal{X} \otimes_k K \simeq_K \mathcal{X}' \otimes_k K$, where $\mathcal{X} = t(\mathcal{A})$ and $\mathcal{X}' = t(\mathcal{A}')$, and so $t(\mathcal{A}) = t(\mathcal{A}')$. □

Our main result in this section is the following:

Theorem 8.16. (Permanence of type) *If $\mathcal{L}$ is an n-step loop algebra based on a prime pfgc $\mathbb{V}_k$-algebra $\mathcal{A}$, then $\mathcal{L}$ is a prime pfgc $\mathbb{V}_k$-algebra and*

$$t(\mathcal{L}) = t(\mathcal{A}).$$

Proof. By Theorem 5.5(iii) and (iv) and Corollary 8.5, $\mathcal{L}$ is a prime pfgc $\mathbb{V}_k$-algebra. So $t(\mathcal{A})$ and $t(\mathcal{L})$ are defined and it remains so show that these types are equal. For this note first that $C(\mathcal{A})$ is an integral domain by Lemma 3.3(i), and so the algebra $C(\mathcal{A}) \otimes_k S^{\otimes n} \simeq_k C(\mathcal{A})[z_1^{\pm 1}, \dots, z_n^{\pm 1}]$ is an integral domain. Let K be an algebraic closure of the quotient field of $C(\mathcal{A}) \otimes_k S^{\otimes n}$. Now by (7–1″) we have the isomorphism

$$\mathcal{L} \otimes_{C(\mathcal{L})} \big(C(\mathcal{A}) \otimes_k S^{\otimes n}\big) \simeq_{C(\mathcal{A}) \otimes_k S^{\otimes n}} \mathcal{A} \otimes_{C(\mathcal{A})} \big(C(\mathcal{A}) \otimes_k S^{\otimes n}\big).$$

Tensoring this over $C(\mathcal{A}) \otimes_k S^{\otimes n}$ with K yields the isomorphism

$$\mathcal{L} \otimes_{C(\mathcal{L})} K \simeq_K \mathcal{A} \otimes_{C(\mathcal{A})} K.$$

Hence we have $\mathcal{X} \otimes_k K \simeq_K \mathcal{X}' \otimes_k K$, where $\mathcal{X} = t(\mathcal{A})$ and $\mathcal{X}' = t(\mathcal{L})$, and so $\mathcal{X} = \mathcal{X}'$. □

Since finite dimensional central simple algebras over k are prime pfgc algebras, we have:

Corollary 8.17. *If $\mathcal{L}$ is an n-step loop algebra based on a finite dimensional central simple $\mathbb{V}_k$-algebra $\mathcal{A}$ over k, then $\mathcal{L}$ is a prime pfgc $\mathbb{V}_k$-algebra and $t(\mathcal{L}) = t(\mathcal{A})$.*

Corollary 8.18. *If $\mathcal{X} \in \mathbb{X}_k$ and $\mathcal{L}$ is an n-step loop algebra based on $\mathcal{X}$, then $\mathcal{L}$ is a prime pfgc $\mathbb{V}_k$-algebra of type $\mathcal{X}$. If further $\mathcal{X}' \in \mathbb{X}_k$ and $\mathcal{L}'$ is an n'-step loop algebra based on $\mathcal{X}'$, then*

$$\mathcal{L} \simeq_k \mathcal{L}' \implies \mathcal{X} = \mathcal{X}' \text{ and } n = n'.$$

Proof. The first statement follows from Corollary 8.17 since $t(\mathcal{X}) = \mathcal{X}$. For the second statement suppose that $\mathcal{L} \simeq_k \mathcal{L}'$. Then by Proposition 8.15, $t(\mathcal{L}) = t(\mathcal{L}')$ and so $\mathcal{X} = \mathcal{X}'$. Finally by Corollary 6.4, $n = n'$. □

Our primary focus in future work will be on the case when the base algebra is a finite dimensional split simple Lie algebra. For ease of reference we therefore record Corollary 8.18 explicitly in that case.

Corollary 8.19. *Suppose that $\mathcal{A}$ is a finite dimensional split simple Lie algebra over a field k of characteristic 0, and $\mathcal{L}$ is an n-step loop algebra based on $\mathcal{A}$. Then $\mathcal{L}$ is a prime pfgc Lie algebra and for any unital k-algebra monomorphism $C(\mathcal{L}) \hookrightarrow K$ into an algebraically closed extension K of k we have*

$$\mathcal{L} \otimes_{C(\mathcal{L})} K \simeq_K \mathcal{A} \otimes_k K.$$

Moreover, if $\mathcal{A}'$ is a finite dimensional split simple Lie algebra and $\mathcal{L}'$ is an n'-step loop algebra based on $\mathcal{A}'$, then

$$\mathcal{L} \simeq_k \mathcal{L}' \implies \mathcal{A} \simeq_k \mathcal{A}' \text{ and } n = n'. \tag{8–3}$$

Proof. We apply Corollary 8.18 to the case when $\mathbb{V}_k$ is the variety of Lie algebras and $\mathbb{X}_k = \{\mathcal{X}_\Pi\}$ as in Example 8.10(a). Since any finite dimensional split simple Lie algebra over k is isomorphic to exactly one algebra in $\mathbb{X}_k$ the result follows. □

If $\mathcal{L}$ is an n-step loop algebra based on a finite dimensional split simple Lie algebra $\mathcal{A}$ (in characteristic 0), then (8–3) tells us that both (the isomorphism class of) the base algebra $\mathcal{A}$ and the number of steps n are isomorphism invariants of $\mathcal{L}$. This answers a natural question that began the research described in this paper. We have now seen in Corollary 8.18 that the result is true in a much broader context. The interested reader can easily write down the results corresponding to Corollary 8.19 for the varieties of associative algebras, alternative algebras and Jordan algebras (see Example 8.10(b), (d) and (e)).

9. Two-step loop algebras

In this section we look more closely at 2-step iterated loop algebras and their centroids. We then conclude with a detailed look at two examples that illustrate several of the concepts studied in this article.

Throughout this section, we assume that m_1 and m_2 are positive integers and that k contains a primitive m_i-th root of unity ζ_{m_i}, $i = 1, 2$. We use the notation of Section 5 (for iterated loop algebras).

We start with some further notation. Let $k^\times = \{ \rho \in k \mid \rho \neq 0 \}$ be the group of units of k. If $\rho \in k^\times$, we let

$$k[u_1, u_2^{\pm 1}, w]_\rho$$

denote the unital associative commutative k-algebra presented by the generators u_1, u_2, u_2^{-1}, w subject to the relations

$$u_2 u_2^{-1} = 1 \quad \text{and} \quad w^2 = (u_1^2 - 4\rho)u_2.$$

It is clear that the set

$$\{ u_1^{i_1} u_2^{i_1} w^j : i_1 \in \mathbb{Z}_{\geq 0},\ i_2 \in \mathbb{Z},\ j = 0, 1 \}$$

is a k-basis for $k[u_1, u_2^{\pm 1}, w]_\rho$. It is also easy to verify that the group of units of $k[u_1, u_2^{\pm 1}, w]_\rho$ is given by

$$U(k[u_1, u_2^{\pm 1}, w]_\rho) = \{ \alpha u_2^{i_2} \mid \alpha \in k^\times,\ i_2 \in \mathbb{Z} \}. \tag{9–1}$$

Indeed, one way to see this is to make use of the multiplicative norm function $N : k[u_1, u_2^{\pm 1}, w]_\rho \to k[u_1, u_2^{\pm 1}]$ defined by $N(a_1 + a_2 w) = a_1^2 - a_2^2 w^2$ for $a_1, a_2 \in k[u_1, u_2^{\pm 1}]$, and use the fact that if u is a unit in $k[u_1, u_2^{\pm 1}, w]_\rho$ then $N(u)$ is a unit in $k[u_1, u_2^{\pm 1}]$. We leave the details of this to the reader.

The algebra $k[u_1, u_2^{\pm 1}, w]_\rho$ is important in the study of iterated loop algebras because of the following fact.

Lemma 9.1. *Let $\mathcal{L} = L(k, \Sigma_1, \Sigma_2)$ be a 2-step iterated loop algebra based on the algebra k, where Σ_i has modulus m_i for $i = 1, 2$. Then exactly one of the following holds*:

(a) $\mathcal{L} \simeq_k k[t_1^{\pm 1}, t_2^{\pm 2}]$ (*the algebra of Laurent polynomials in 2 variables*).

(b) $\mathcal{L} \simeq_k k[u_1, u_2^{\pm 1}, w]_\rho$ *for some* $\rho \in k^\times$.

Moreover (*a*) *holds if and only if* $z_1^{m_1} z_2^j \in \mathcal{L}$ *for some* $j \in \mathbb{Z}$.

Proof. Note that the group of units in $k[t_1^{\pm 1}, t_2^{\pm 1}]$ spans the algebra $k[t_1^{\pm 1}, t_2^{\pm 1}]$, whereas this is not true for the algebra $k[u_1, u_2^{\pm 1}, w]_\rho$ (by (9–1)). Thus (a) and (b) cannot hold simultaneously. So it remains to show that either (a) or (b) holds (the final statement will be proved along the way).

Now as noted in Remark 5.2(iii), we have

$$\mathscr{L} = L(k, \sigma_1, \sigma_2),$$

where σ_1 is an automorphism of period m_1 of $\mathscr{L}_0 = k$, and σ_2 is an automorphism of period m_2 of $\mathscr{L}_1 = L(k, \sigma_1)$. Then, since σ_1 is an algebra homomorphism, $\sigma_1 = 1$ and so

$$\mathscr{L}_1 = k[y_1^{\pm 1}], \text{ where } y_1 = z_1^{m_1}.$$

Thus σ_2 is an automorphism of period m_2 of $k[y_1^{\pm 1}]$. Hence either $\sigma_2(y_1) = \rho y_1$ for some m_2-th root of unity ρ in $k^\times$ or $\sigma_2(y_1) = \rho y_1^{-1}$ for some $\rho \in k^\times$. Moreover (for the proof of the final statement in the proposition) the first of these possibilities holds if and only if y_1 is homogeneous in the grading Σ_2 determined by σ_2 which in turn holds if and only if $y_1 z_2^j \in \mathscr{L}$ for some $j \in \mathbb{Z}$.

Case (a): Suppose that $\sigma_2(y_1) = \rho y_1$ for some m_2-th root of unity ρ in $k^\times$. Let n_2 be the order of ρ in $k^\times$. Then n_2 is a divisor of m_2,

$$\rho = \zeta_{m_2}^{p_2 r}, \text{ where } p_2 = \frac{m_2}{n_2},$$

and r is relatively prime to n_2 (take $r = 0$ if $n_2 = 1$). Choose an inverse s for r modulo n_2 (take $s = 0$ if $n_2 = 1$). Now the grading Σ_2 of $\mathscr{L}_1$ is given by $\mathscr{L}_1 = \bigoplus_{\bar{j} \in \mathbb{Z}_{m_2}} (\mathscr{L}_1)_{\bar{j}}$, where $(\mathscr{L}_1)_{\bar{j}}$ is spanned by the elements y_1^i with $i \in \mathbb{Z}$ and $\sigma_2(y_1^i) = \zeta_{m_2}^j y_1^i$. But n_2 and s are relatively prime and so any integer can be expressed in the form $an_2 + bs$, where $a, b \in \mathbb{Z}$. Also

$$\sigma_2(y_1^{an_2+bs}) = \rho^{an_2+bs} y_1^{an_2+bs} = \rho^{bs} y_1^{an_2+bs} = \zeta_{m_2}^{p_2 rbs} y_1^{an_2+bs} = \zeta_{m_2}^{p_2 b} y_1^{an_2+bs}$$

and so $y_1^{an_2+bs} \in (\mathscr{L}_1)_{\overline{p_2 b}}$. Therefore $\mathscr{L} = L(\mathscr{L}_1, \sigma_2)$ is spanned by elements of the form

$$y_1^{an_2+bs} z_2^{p_2 b}, \quad a, b \in \mathbb{Z}.$$

But $y_1^{an_2+bs} z_2^{p_2 b} = (y_1^{n_2})^a (y_1^s z_2^{p_2})^b$. Hence we obtain

$$\mathscr{L} = k[t_1^{\pm 1}, t_2^{\pm 2}], \text{ where } t_1 = y_1^{n_2} \text{ and } t_2 = y_1^s z_2^{p_2}.$$

Case (b): Suppose that $\sigma_2(y_1) = \rho y_1^{-1}$ for some $\rho \in k^\times$. Then σ_2 has order 2 and so m_2 is even. Let $p_2 = \frac{m_2}{2}$ and $y_2 = z_2^{p_2}$. Then

$$\mathscr{L} = \left(\mathscr{L}_1^+ \otimes_k k[(y_2^2)^{\pm 1}]\right) \oplus \left(\mathscr{L}_1^- \otimes_k y_2 k[(y_2^2)^{\pm 1}]\right),$$

where $\mathscr{L}_1^\pm$ is the ± 1-eigenspace for σ_2. Now it is clear that $\mathscr{L}_1^+$ has a k-basis consisting of the elements $(y_1 + \rho y_1^{-1})^a$, $a \geq 0$. Therefore $\mathscr{L}_1^+ \otimes_k k[(y_2^2)^{\pm 1}]$ has basis

$$(y_1 + \rho y_1^{-1})^a y_2^{2b}, \quad a, b \in \mathbb{Z}, \ a \geq 0.$$

Also one easily checks that $\mathcal{L}_1^- = (y_1 - \rho y_1^{-1})\mathcal{L}_1^+$, so $\mathcal{L}_1^- \otimes_k y_2 k[(y_2^2)^{\pm 1}]$ has basis

$$(y_1 - \rho y_1^{-1}) y_2 (y_1 + \rho y_1^{-1})^a y_2^{2b}, \quad a, b \in \mathbb{Z},\ a \geq 0.$$

Thus, setting

$$u_1 = y_1 + \rho y_1^{-1},\ u_2 = y_2^2 \quad \text{and} \quad w = (y_1 - \rho y_1^{-1}) y_2,$$

we see that $\mathcal{L}$ has basis $u_1^a u_2^b w^c$, $a \in \mathbb{Z}_{\geq 0}$, $b \in \mathbb{Z}$, $c = 0, 1$. Moreover, one checks directly that $w^2 = (u_1^2 - 4\rho) u_2$, and so we have identified $\mathcal{L}$ with $k[u_1, u_2^{\pm 1}, w]_\rho$. □

Remark 9.2. In Case (a) of the proof of Lemma 9.1, the conclusion is an immediate consequence of a more general "erasing theorem" that was proved in [Allison et al. 2004, Theorem 5.1]. We have included the proof above since it is short and self-contained.

Remark 9.3. If $\rho, \rho' \in k^\times$, one can show that

$$k[u_1, u_2^{\pm 1}, w]_\rho \simeq_k k[u_1, u_2^{\pm 1}, w]_{\rho'} \iff \rho' \rho^{-1} \text{ is a square in } k^\times.$$

In particular, if k is algebraically closed, the isomorphism class of $k[u_1, u_2^{\pm 1}, w]_\rho$ does not depend on ρ. In that case Lemma 9.1 tells us that, up to isomorphism, there are exactly two (one step) loop algebras based on $k[y_1^{\pm 1}]$. This fact is a special case of a more general result about (one step) loop algebras based on the algebra $\mathcal{A}$ of Laurent polynomials $k[y_1^{\pm 1}, \ldots, y_q^{\pm 1}]$ over an algebraically closed field k. Indeed, using the fact that the abstract automorphism group of $\mathcal{A}$ is $(k^\times)^q \rtimes \mathrm{GL}_q(\mathbb{Z})$ and some techniques from Galois cohomology (see Remark 4.8), one can show that there is an injective map that attaches to each R-isomorphism class of loop algebras based on $\mathcal{A}$ an invariant in the set of conjugacy classes of $GL_q(\mathbb{Z})$. (When $q = 1$, $GL_q(\mathbb{Z})$ has exactly two conjugacy classes and one can show that R-isomorphism coincides with k-isomorphism.) We omit proofs of the statements in this remark, since we will not be using these statements here and since their proofs would take us rather far afield.

Lemma 9.1 together with Theorem 6.2 implies the following more general result:

Proposition 9.4. *Let $\mathcal{L} = L(\mathcal{A}, \Sigma_1, \Sigma_2)$ be a 2-step iterated loop algebra based on a finite dimensional central simple algebra $\mathcal{A}$ over k, where Σ_i has modulus m_i for $i = 1, 2$. Then exactly one of the following holds*:

(a) $C(\mathcal{L}) \simeq_k k[t_1^{\pm 1}, t_2^{\pm 2}]$.

(b) $C(\mathcal{L}) \simeq_k k[u_1, u_2^{\pm 1}, w]_\rho$ *for some* $\rho \in k^\times$.

Moreover (a) *holds if and only if* $z_1^{m_1} z_2^j \in \bar{C}(\mathcal{L})$ *for some* $j \in \mathbb{Z}$ (see Remark 6.5).

Definition 9.5. As in Proposition 9.4, let $\mathcal{L} = L(\mathcal{A}, \Sigma_1, \Sigma_2)$ be a 2-step iterated loop algebra based on a finite dimensional central simple algebra $\mathcal{A}$ over k, where

Σ_i has modulus m_i for $i = 1, 2$. We say that $\mathscr{L}$ is of the *first kind* (resp. *second kind*) if $C(\mathscr{L})$ is isomorphic to $k[t_1^{\pm 1}, t_2^{\pm 2}]$ (resp. $k[u_1, u_2^{\pm 1}, w]_\rho$ for some $\rho \in k^\times$).

Remark 9.6. (a) It follows from Corollary 6.6 that any 2-step multiloop algebra based on a finite dimensional central simple algebra is of the first kind.

(b) Suppose k is algebraically closed of characteristic 0 and $\mathscr{L} = L(\mathscr{A}, \sigma_1, \sigma_2)$ is a 2-step iterated loop algebra based on a finite dimensional central simple Lie algebra $\mathscr{A}$ over k, where σ_i has period m_i for $i = 1, 2$. Then $L(\mathscr{A}, \sigma_1)$ is the derived algebra modulo its centre of an affine Kac–Moody Lie algebra $\mathfrak{g}$ [Kac 1990, Theorem 8.5]. Moreover one can show that the 2-step loop algebra $\mathscr{L}$ is of the first kind in the sense of Definition 9.5 if and only if the automorphism σ_2 of $L(\mathscr{A}, \sigma_1)$ is induced by an automorphism of the first kind of $\mathfrak{g}$ (as defined for example in [Levstein 1988, Part III.1]). Indeed this example is the reason for our choice of terminology.

We conclude by looking at two examples of 2-step iterated loop algebras. These examples illustrate the above proposition (Proposition 9.4) as well as a number of the concepts studied in this article.

Example 9.7. Suppose that k is of characteristic 0. In this example we consider a 2-step iterated loop algebra $\mathscr{L} = L(\mathscr{A}, \sigma_1, \sigma_2)$ based on the Lie algebra $\mathscr{A} = \mathrm{sl}_{\ell+1}(k)$ over k, where $\ell \geq 1$ and σ_1 and σ_2 have order $m_1 = m_2 = 2$.

Before beginning it will be convenient to define four commuting automorphisms η_1, η_2, κ_1 and κ_2 of $S^{\otimes 2}$ by

$$\eta_1(z_1^{i_1} z_2^{i_2}) = (-1)^{i_1} z_1^{i_1} z_2^{i_2}, \quad \eta_2(z_1^{i_1} z_2^{i_2}) = (-1)^{i_2} z_1^{i_1} z_2^{i_2},$$
$$\kappa_1(z_1^{i_1} z_2^{i_2}) = z_1^{-i_1} z_2^{i_2} \quad \text{and} \quad \kappa_2(z_1^{i_1} z_2^{i_2}) = z_1^{i_1} z_2^{-i_2}$$

for $i_1, i_2 \in \mathbb{Z}$. Each of these automorphisms restricts to an automorphism of $k[z_1^{\pm 1}]$ which we also denote by η_1, η_2, κ_1 and κ_2 respectively.

To construct $\mathscr{L}$ we first let $\mathscr{L}_0 = \mathscr{A}$. Next let $\sigma_1 \in \mathrm{Aut}(\mathscr{A})$ be defined by $\sigma_1(a) = -J a^t J$, where

$$J = \begin{bmatrix} 0 & \dots & 1 \\ \vdots & \iddots & \vdots \\ 1 & \dots & 0 \end{bmatrix}.$$

Then σ_1 has order 2 and we set

$$\mathscr{L}_1 := L(\mathscr{A}, \sigma_1, z_1),$$

using the notation of Remark 4.4. Thus $\mathscr{L}_1$ is the algebra of fixed points in $\mathscr{A} \otimes_k k[z_1^{\pm 1}]$ of the automorphism $\sigma_1 \otimes \eta_1$. (If k is an algebraically closed field of characteristic 0 and $\ell \geq 2$, then $\mathscr{L}_1$ is the derived algebra modulo its centre of the affine Kac–Moody Lie algebra of type $A_\ell^{(2)}$ [Kac 1990, Chapter 8].)

Next the automorphisms $1_{\mathscr{A}} \otimes \kappa_1$ and $\sigma_1 \otimes \eta_1$ of $\mathscr{A} \otimes_k k[z_1^{\pm 1}]$ commute, so $1_{\mathscr{A}} \otimes \kappa_1$ stabilizes $\mathscr{L}_1$. We set $\sigma_2 = 1_{\mathscr{A}} \otimes \kappa_1 |_{\mathscr{L}_1} \in \mathrm{Aut}_k(\mathscr{L}_1)$. Then σ_2 has order 2, and we set

$$\mathscr{L} = \mathscr{L}_2 := L(\mathscr{L}_1, \sigma_2, z_2).$$

By construction $\mathscr{L}$ is a 2-step iterated loop algebra based on $\mathscr{A}$.

It is clear from the above descriptions of $\mathscr{L}_1$ and $\mathscr{L}_2$, that $\mathscr{L}$ is the algebra of common fixed points in $\mathscr{A} \otimes_k S^{\otimes 2}$ of the automorphisms $\sigma_1 \otimes \eta_1$ and $1_{\mathscr{A}} \otimes \kappa_1 \eta_2$. From this it follows easily that

$$\mathscr{L} = \{ x \in \mathrm{sl}_{\ell+1}(K) \mid x^* = -x \}, \tag{9–2}$$

where

$$K = (S^{\otimes 2})^{\kappa_1 \eta_2}$$

is the algebra of fixed points in $S^{\otimes 2}$ of the automorphism $\kappa_1 \eta_2$, and

$$x^* = -J(\eta_1 x)^t J \tag{9–3}$$

for $x \in M_n(K)$. (Here $\eta_1 x$ denotes the matrix obtained from x by applying η_1 to the entries of x.) In more geometric language, $\mathscr{L}$ can be viewed as the Lie algebra of K-linear transformations of the free K-module $K^{\ell+1}$ that are skew relative to the hermitian form $(u, v) \mapsto (\eta_1 u)^t J v$.

Now by Remark 6.5, the centroid of $\mathscr{L}$ is isomorphic to the algebra

$$\bar{C}(\mathscr{L}) = \{ u \in S^{\otimes 2} \mid u \cdot \mathscr{L} \subset \mathscr{L} \} \tag{9–4}$$

of $S^{\otimes 2}$. This together with (9–2) implies that $\bar{C}(\mathscr{L}) \subset K$. But by (9–3), $(u \cdot x)^* = (\eta_1 u) \cdot x^*$ for $u \in K$ and $x \in \mathrm{sl}_{\ell+1}(K)$. Hence it follows from (9–2) and (9–4) that $\bar{C}(\mathscr{L}) = K^{\eta_1}$. So we have

$$\bar{C}(\mathscr{L}) = (S^{\otimes 2})^{\langle \eta_1, \kappa_1 \eta_2 \rangle}.$$

Note also that, by Theorem 7.1, $S^{\otimes 2}$ is a free $\bar{C}(\mathscr{L})$-module of rank 4 and

$$\mathscr{L} \otimes_{\bar{C}(\mathscr{L})} S^{\otimes 2} \simeq \mathrm{sl}_{\ell+1}(S^{\otimes 2}).$$

Moreover, by Corollary 8.17, $\mathscr{L}$ is a prime pfgc Lie algebra of type A_ℓ (see Example 8.14).

Finally, note that $\kappa_1 \eta_2 (z_1^2 z_2^j) = (-1)^j z_1^{-2} z_2^j \neq z_1^2 z_2^j$ and so $z_1^2 z_2^j \notin \bar{C}(\mathscr{L})$ for $j \in \mathbb{Z}$. Thus $\mathscr{L}$ is of the second kind. (In fact one can check directly that $\bar{C}(\mathscr{L})$ is isomorphic to $k[u_1, u_2^{\pm 1}, w]_\rho$ for $\rho = 1$.) So $\bar{C}(\mathscr{L})$ is not isomorphic to the algebra of Laurent polynomials in any number of variables (since $\bar{C}(\mathscr{L})$ is not spanned by its units). Hence, by Corollary 6.6, $\mathscr{L}$ is not isomorphic to a multiloop algebra in any number of steps based on a finite dimensional central simple Lie algebra.

Example 9.8. Suppose that $\ell \geq 1$ and k is a field which contains a primitive ℓ-th root of unity $\zeta = \zeta_\ell$. In this example we consider a 2-step multiloop loop algebra $\mathscr{L} = M(\mathscr{A}, \sigma_1, \sigma_2)$ based on the associative algebra $\mathscr{A} = M_\ell(k)$ of $\ell \times \ell$-matrices over k, where σ_1 and σ_2 have order $m_1 = m_2 = \ell$.

First let

$$a_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \zeta & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \zeta^{\ell-1} \end{bmatrix} \quad \text{and} \quad a_2 = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 1 & 0 & \dots & 0 \end{bmatrix}$$

in $\mathscr{A}$. Then $a_2a_1 = \zeta a_1a_2$, $a_1^\ell = a_2^\ell = 1$, and it is well known that

$$\{ a_1^{i_1} a_2^{i_2} \mid 0 \leq i_1, i_2 \leq \ell - 1 \}$$

is a basis for $\mathscr{A}$. (See for example [Draxl 1983, §11].)

Define $\sigma_i \in \mathrm{Aut}_k(\mathscr{A})$ by $\sigma_i(x) = a_i x a_i^{-1}$ for $x \in \mathscr{A}$, $i = 1, 2$. Then $\sigma_i(a_i) = a_i$, $\sigma_1(a_2) = \zeta^{-1}a_2$ and $\sigma_2(a_1) = \zeta a_1$. Hence σ_1 and σ_2 are commuting automorphisms of $\mathscr{A}$ of order ℓ. Let

$$\mathscr{L} = M(\mathscr{A}, \sigma_1, \sigma_2)$$

be the multiloop algebra of σ_1, σ_2 based on $\mathscr{A}$ (with $m_1 = m_2 = \ell$). To calculate $\mathscr{L}$ explicitly, note that

$$\sigma_1(a_2^{-i_1} a_1^{i_2}) = \zeta^{i_1} a_2^{-i_1} a_1^{i_2} \quad \text{and} \quad \sigma_2(a_2^{-i_1} a_1^{i_2}) = \zeta^{i_2} a_2^{-i_1} a_1^{i_2}$$

for $i_1, i_2 \in \mathbb{Z}$. Thus $\mathscr{A}_{\bar{\imath}_1, \bar{\imath}_2} = k a_2^{-i_1} a_1^{i_2}$ for $i_1, i_2 \in \mathbb{Z}$. Consequently

$$\mathscr{L} = \mathrm{span}_k\{ a_2^{-i_1} a_1^{i_2} \otimes z_1^{i_1} z_2^{i_2} \mid i_1, i_2 \in \mathbb{Z} \} = \mathrm{span}_k\{ x_1^{i_1} x_2^{i_2} \mid i_1, i_2 \in \mathbb{Z} \},$$

where

$$x_1 = a_2^{-1} \otimes z_1 = \begin{bmatrix} 0 & \dots & 0 & z_1 \\ z_1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & z_1 & 0 \end{bmatrix} \text{ and } x_2 = a_1 \otimes z_2 = \begin{bmatrix} z_2 & 0 & \dots & 0 \\ 0 & \zeta z_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \zeta^{\ell-1} z_2 \end{bmatrix}$$

in $\mathscr{L}$. Thus $\mathscr{L}$ is the subalgebra of $M_\ell(S^{\otimes 2})$ generated as an algebra by the matrices $x_1^{\pm 1}$, $x_2^{\pm 1}$ which satisfy the relations

(9–5) $$x_i x_i^{-1} = x_i^{-1} x_i = 1 \quad \text{and} \quad x_2 x_1 = \zeta x_1 x_2.$$

It follows that $\mathscr{L} \simeq k_{\mathbf{q}}$, where $k_{\mathbf{q}}$ is the algebra presented by the generators x_1, x_2 subject to the relations (9–5). This algebra $k_{\mathbf{q}}$, which is called the *quantum torus* determined by the matrix

$$\mathbf{q} = \begin{bmatrix} 1 & \zeta \\ \zeta^{-1} & 1 \end{bmatrix},$$

has arisen in a number of different contexts; see for example [Magid 1978; McConnell and Pettit 1988; Berman et al. 1996; Gao 2000].

Note that by Corollary 6.6, the centroid (= centre) of $\mathscr{L}$ is isomorphic to $\bar{C}(\mathscr{L}) = k[t_1^{\pm 1}, t_2^{\pm 1}]$, where $t_1 = z_1^\ell$ and $t_2 = z_2^\ell$. Moreover, by Theorem 6.1, $S^{\otimes 2}$ is a free $\bar{C}(\mathscr{L})$-module of rank ℓ^2 and $\mathscr{L} \otimes_{\bar{C}(\mathscr{L})} S^{\otimes 2} \simeq M_\ell(S^{\otimes 2})$. Consequently (see Corollary 7.4) $\mathscr{L} \simeq k_{\mathbf{q}}$ is a prime Azumaya algebra of constant rank ℓ^2 that is split by the extension $S^{\otimes 2}/k[t_1^{\pm 1}, t_2^{\pm 1}]$.

Remark 9.9. The fact that the quantum torus $k_{\mathbf{q}}$ (described in the preceding example) is an Azumaya algebra was seen by a different method some time ago in [Magid 1978, Lemma 4]. This information about the algebra $k_{\mathbf{q}}$ is important because it tells us that $k_{\mathbf{q}}$ defines an element $[k_{\mathbf{q}}]$ of the Brauer group of the ring $k[t_1^{\pm 1}, t_2^{\pm 2}]$. In fact ${}_\ell\mathrm{Br}(k[t_1^{\pm 1}, t_2^{\pm 2}])$ is cyclic of order ℓ and the element $[k_{\mathbf{q}}]$ is a generator of this group [Magid 1978, Theorem 6].

Remark 9.10. The authors wish to thank John Faulkner for conversations that led to Example 9.8. This example turns out to be a special case of a more general construction of quantum tori and their nonassociative analogs as multiloop algebras. This topic will be investigated in a article in preparation by the present authors together with John Faulkner.

References

[Allison et al. 1997a] B. N. Allison, S. Azam, S. Berman, Y. Gao, and A. Pianzola, *Extended affine Lie algebras and their root systems*, Mem. Amer. Math. Soc. **603**, Amer. Math. Soc., Providence, RI, 1997. MR 97i:17015 Zbl 0879.17012

[Allison et al. 1997b] B. N. Allison, S. Berman, Y. Gao, and A. Pianzola, "A characterization of affine Kac-Moody Lie algebras", *Comm. Math. Phys.* **185**:3 (1997), 671–688. MR 98h:17026 Zbl 0879.17013

[Allison et al. 2002] B. Allison, S. Berman, and A. Pianzola, "Covering algebras, I: Extended affine Lie algebras", *J. Algebra* **250**:2 (2002), 485–516. MR 2003d:17026 Zbl 1002.17010

[Allison et al. 2004] B. Allison, S. Berman, and A. Pianzola, "Covering algebras, II: Isomorphism of loop algebras", *J. Reine Angew. Math.* **571** (2004), 39–71. MR 2005e:17042 Zbl 1056.17018

[Allison et al. ≥ 2006] B. Allison, S. Berman, and A. Pianzola, "Covering algebras, III: The nullity 2 case". In preparation.

[Benkart and Zelmanov 1996] G. Benkart and E. Zelmanov, "Lie algebras graded by finite root systems and intersection matrix algebras", *Invent. Math.* **126**:1 (1996), 1–45. MR 97k:17044 Zbl 0871.17024

[Berman and Moody 1992] S. Berman and R. V. Moody, "Lie algebras graded by finite root systems and the intersection matrix algebras of Slodowy", *Invent. Math.* **108**:2 (1992), 323–347. MR 93e:17031 Zbl 0778.17018

[Berman et al. 1996] S. Berman, Y. Gao, and Y. S. Krylyuk, "Quantum tori and the structure of elliptic quasi-simple Lie algebras", *J. Funct. Anal.* **135**:2 (1996), 339–389. MR 97b:17007 Zbl 0847.17009

[Bourbaki 1972] N. Bourbaki, *Commutative algebra*, Hermann, Paris, 1972. Reprinted by Springer, Berlin, 1989. MR 90a:13001

[Bourbaki 1974] N. Bourbaki, *Algebra, Part I: Chapters 1–3*, Hermann, Paris, 1974. Reprinted by Springer, Berlin, 1989. MR 50 #6689

[Draxl 1983] P. K. Draxl, *Skew fields*, London Mathematical Society Lecture Note Series **81**, Cambridge University Press, Cambridge, 1983. MR 85a:16022 Zbl 0498.16015

[Erickson et al. 1975] T. S. Erickson, W. S. Martindale, and J. M. Osborn, "Prime nonassociative algebras", *Pacific J. Math.* **60**:1 (1975), 49–63. MR 52 #3264 Zbl 0355.17005

[Gao 2000] Y. Gao, "Representations of extended affine Lie algebras coordinatized by certain quantum tori", *Compositio Math.* **123**:1 (2000), 1–25. MR 2002c:17036 Zbl 0958.17008

[Helgason 1978] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Mathematics **80**, Academic Press, New York, 1978. 2nd ed., Amer. Math. Soc., Providence (RI), 2001. MR 2002b:53081 Zbl 0451.53038

[Jacobson 1962] N. Jacobson, *Lie algebras*, Interscience, New York, 1962. Reprinted by Dover, New York, 1979. MR 80k:17001 Zbl 0121.27504

[Jacobson 1968] N. Jacobson, *Structure and representations of Jordan algebras*, Colloquium Publications **39**, American Mathematical Society, Providence, R.I., 1968. MR 40 #4330 Zbl 0218.17010

[Kac 1969] V. G. Kac, "Automorphisms of finite order of semisimple Lie algebras", *Funkcional. Anal. i Priložen.* **3**:3 (1969), 94–96. In Russian; translated in *Funct. Anal. Appl.* **3** (1969), 252–254. MR 40 #4322 Zbl 0274.17002

[Kac 1990] V. G. Kac, *Infinite-dimensional Lie algebras*, 3rd ed., Cambridge University Press, Cambridge, 1990. MR 92k:17038 Zbl 0716.17022

[Knus 1991] M.-A. Knus, *Quadratic and Hermitian forms over rings*, Grundlehren der Math. Wissenschaften **294**, Springer, Berlin, 1991. MR 92i:11039 Zbl 0756.11008

[Knus and Ojanguren 1974] M.-A. Knus and M. Ojanguren, *Théorie de la descente et algèbres d'Azumaya*, Lecture Notes in Math. **389**, Springer, Berlin, 1974. MR 54 #5209 Zbl 0284.13002

[Kunz 1985] E. Kunz, *Introduction to commutative algebra and algebraic geometry*, Birkhäuser, Boston, 1985. MR 86e:14001 Zbl 0563.13001

[van de Leur 2001] J. van de Leur, "Twisted toroidal Lie algebras", preprint, 2001. math/0106119

[Levstein 1988] F. Levstein, "A classification of involutive automorphisms of an affine Kac–Moody Lie algebra", *J. Algebra* **114**:2 (1988), 489–518. MR 90g:17025 Zbl 0654.17012

[Magid 1978] A. R. Magid, "Brauer groups of linear algebraic groups with characters", *Proc. Amer. Math. Soc.* **71**:2 (1978), 164–168. MR 58 #5619 Zbl 0393.20029

[McConnell and Pettit 1988] J. C. McConnell and J. J. Pettit, "Crossed products and multiplicative analogues of Weyl algebras", *J. London Math. Soc.* (2) **38**:1 (1988), 47–55. MR 90c:16011 Zbl 0652.16007

[McCrimmon and Zel'manov 1988] K. McCrimmon and E. Zel'manov, "The structure of strongly prime quadratic Jordan algebras", *Advances in Mathematics* **69**:2 (1988), 133–222. MR 89k:17052 Zbl 0656.17015

[Pianzola 2002] A. Pianzola, "Affine Kac-Moody Lie algebras as torsors over the punctured line", *Indag. Math.* (*N.S.*) **13**:2 (2002), 249–257. MR 2004m:17035 Zbl 1029.17021

[Polikarpov and Shestakov 1990] S. V. Polikarpov and I. P. Shestakov, "Nonassociative affine algebras", *Algebra and Logic* **29**:6 (1990), 458–466. MR 93b:17009 Zbl 0786.17002

[Pollmann 1994] U. Pollmann, "Realisation der biaffinen Wurzelsysteme von Saito in Lie-Algebren", Hamburger Beiträge zur Mathematik (preprint) 29, Universität Hamburg, 1994.

[Saito and Yoshii 2000] K. Saito and D. Yoshii, "Extended affine root system, IV: Simply-laced elliptic Lie algebras", *Publ. Res. Inst. Math. Sci.* (*Kyoto*) **36**:3 (2000), 385–421. MR 2002c:17038 Zbl 0987.17012

[Wakimoto 1985] M. Wakimoto, "Extended affine Lie algebras and a certain series of Hermitian representations", preprint, 1985.

[Zhevlakov et al. 1982] K. A. Zhevlakov, A. M. Slin'ko, I. P. Shestakov, and A. I. Shirshov, *Rings that are nearly associative*, Pure and Applied Mathematics **104**, Academic Press, New York, 1982. MR 83i:17001 Zbl 0487.17001

Received December 16, 2004.

BRUCE ALLISON
DEPARTMENT OF MATHEMATICAL AND STATISTICAL SCIENCES
UNIVERSITY OF ALBERTA
EDMONTON, AB
CANADA T6G 2G1

ballison@math.ualberta.ca

STEPHEN BERMAN
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF SASKATCHEWAN
SASKATOON, SK
CANADA S7N 5E6

berman@math.usask.ca

ARTURO PIANZOLA
DEPARTMENT OF MATHEMATICAL AND STATISTICAL SCIENCES
UNIVERSITY OF ALBERTA
EDMONTON, AB
CANADA T6G 2G1

a.pianzola@ualberta.ca

PACIFIC JOURNAL OF MATHEMATICS
Vol. 227, No. 1, 2006

HAMILTONIAN-MINIMAL LAGRANGIAN SUBMANIFOLDS IN COMPLEX SPACE FORMS

ILDEFONSO CASTRO, HAIZHONG LI AND FRANCISCO URBANO

Using Legendrian immersions and, in particular, Legendre curves in odd-dimensional spheres and anti-de Sitter spaces, we construct new examples of Hamiltonian-minimal Lagrangian submanifolds in complex projective and hyperbolic spaces, including explicit one-parameter families of embeddings of quotients of certain product manifolds. We also give new examples of minimal Lagrangian submanifolds in complex projective and hyperbolic spaces. Making use of all these constructions, we get Hamiltonian-minimal and special Lagrangian cones in complex Euclidean space as well.

1. Introduction

Let $(\widetilde{M}^n, J, \langle\,,\rangle)$ be a Kähler manifold of complex dimension n, where J is the complex structure and $\langle\,,\rangle$ the Kähler metric. The Kähler 2-form is defined by $\omega(\,\cdot\,,\cdot\,) = \langle J\,\cdot\,,\cdot\,\rangle$. An immersion $\psi : M^n \to \widetilde{M}^n$ of an n-dimensional manifold M is called *Lagrangian* if $\psi^*\omega \equiv 0$. For this type of immersions, J defines a bundle isomorphism between the tangent bundle TM and the normal bundle $T^\perp M$.

A vector field X on $\widetilde{M}$ is a Hamiltonian vector field if there exists a smooth function $F : \widetilde{M} \to \mathbb{R}$ such that $X = J\widetilde{\nabla}F$, where $\widetilde{\nabla}$ is the gradient in $\widetilde{M}$. The diffeomorphisms of the flux of a Hamiltonian vector field transform Lagrangian submanifolds into Lagrangian ones.

In this setting, Oh [1990] studied the following natural variational problem. A normal vector field ξ to a Lagrangian immersion $\psi : M^n \to \widetilde{M}^n$ is called *Hamiltonian* if $\xi = J\nabla f$, where $f \in C^\infty(M)$ and ∇f is the gradient of f with respect to the induced metric. Take $f \in C_0^\infty(M)$ and let $\{\psi_t : M \to \widetilde{M}\}$ be a variation of ψ, with $\psi_0 = \psi$ and $\frac{d}{dt}\big|_{t=0}\psi_t = \xi$. The first variation of the volume functional is given by

$$\frac{d}{dt}\Big|_{t=0} \operatorname{vol}(M, \psi_t^*\langle\,,\rangle) = -\int_M f \operatorname{div} JH\, dM$$

MSC2000: primary 53C42, 53B25; secondary 53A05, 53D12.
Keywords: Hamiltonian-minimal, Lagrangian submanifolds, Legendrian immersions.
Castro's and Urbano's research is partially supported by a MEC-Feder grant MTM2004-00109.
Li is partially supported by Zhongdian grant No. 10531090 of NSFC..

(see [Oh 1990]), where H is the mean curvature vector of the immersion ψ and div denotes the divergence operator on M. Oh called the critical points of this variational problem *Hamiltonian minimal* (or *H-minimal*) Lagrangian submanifolds; they are characterized by the third-order differential equation div $JH = 0$. In particular, minimal Lagrangian submanifolds (where "minimal" means that the mean curvature vector vanishes) are trivially H-minimal; so is, more generally, any Lagrangian submanifold with parallel mean curvature vector.

Even when $\widetilde{M}$ is a simply connected complex space form, only few examples of H-minimal Lagrangian submanifolds are known outside the class of Lagrangian submanifolds with parallel mean curvature vector.

This can be a brief history of them: $\mathbb{S}^1$-invariant H-minimal Lagrangian tori in the complex Euclidean plane $\mathbb{C}^2$ were classified in [Castro and Urbano 1998]. H-minimal Lagrangian cones in $\mathbb{C}^2$ were studied in [Schoen and Wolfson 1999]. Hélein and Romon [2000; 2002a] derived a Weierstrass-type representation formula to describe all H-minimal Lagrangian tori and Klein bottles in $\mathbb{C}^2$. When the ambient space is the complex projective plane $\mathbb{CP}^2$ or the complex hyperbolic plane $\mathbb{CH}^2$, conformal parametrizations of H-minimal Lagrangian surfaces using holomorphic data were obtained in [Hélein and Romon 2002b; 2003]. Making use of this technique, Anciaux [2003] constructed H-minimal Lagrangian singly periodic cylinders and H-minimal Lagrangian surfaces with a nonconical singularity in $\mathbb{C}^2$. Only recently have examples of H-minimal Lagrangian submanifolds of arbitrary dimension in $\mathbb{C}^n$ and $\mathbb{CP}^n$ been found, in [Mironov 2004]. A classification of H-minimal Lagrangian submanifolds foliated by $(n-1)$-spheres in $\mathbb{C}^n$ is given in [Anciaux et al. 2006].

Our aim in this paper is the construction of H-minimal Lagrangian submanifolds in complex Euclidean space $\mathbb{C}^n$, complex projective space $\mathbb{CP}^n$ and complex hyperbolic space $\mathbb{CH}^n$, for arbitrary $n \geq 2$. The examples in $\mathbb{CP}^n$ are constructed by projection, via the Hopf fibration $\Pi : \mathbb{S}^{2n+1} \to \mathbb{CP}^n$, of certain family of Legendrian submanifolds of the sphere $\mathbb{S}^{2n+1}$ (Corollary 3.2). The cones with links in this family of Legendrian submanifolds provide new examples of H-minimal Lagrangian submanifolds in $\mathbb{C}^{n+1}$ (Section 5). Using the Hopf fibration $\Pi : \mathbb{H}^{2n+1}_1 \to \mathbb{CH}^n$ and a similar family of Legendrian submanifolds of the anti-de Sitter space $\mathbb{H}^{2n+1}_1$ (Corollary 6.5), we also find examples of H-minimal Lagrangian submanifolds in $\mathbb{CH}^n$. In a certain sense, our construction is reminiscent of the Smith join method (see [Eells and Ratto 1993]) for constructing harmonic maps between spheres.

In $\mathbb{CP}^n$, we emphasize two different one-parameter families of H-minimal Lagrangian immersions described in Corollaries 4.1 and 4.4; as a particular case, in Corollary 4.2 we provide explicit Lagrangian H-minimal embeddings of certain quotients of $\mathbb{S}^1 \times \mathbb{S}^{n_1} \times \mathbb{S}^{n_2}$, where $n_1 + n_2 + 1 = n$.

In $\mathbb{CH}^n$, we also point out in Corollary 6.6 a one-parameter family of H-minimal Lagrangian immersions, which (in the easiest cases) induce explicit Lagrangian H-minimal embeddings of certain quotients of $\mathbb{S}^1 \times \mathbb{S}^{n_1} \times \mathbb{RH}^{n_2}$, for $n_1 + n_2 + 1 = n$ (see Corollary 6.7). Here $\mathbb{RH}^{n_2}$ denotes real hyperbolic space.

As a byproduct, using our method of construction, we also obtain new examples of minimal Lagrangian submanifolds in $\mathbb{CP}^n$ (Corollary 4.1, Remark 4.3 and Corollary 4.4) and $\mathbb{CH}^n$ (Corollaries 6.5 and 6.9), as well as special Lagrangian cones in $\mathbb{C}^{n+1}$ (see Section 5).

2. Lagrangian submanifolds versus Legendrian submanifolds

Let $\mathbb{C}^{n+1}$ be complex Euclidean space endowed with the Euclidean metric $\langle\, ,\rangle$ and standard complex structure J. The Liouville 1-form is given by $\Lambda_z(v) = \langle v, Jz\rangle$ for all $z \in \mathbb{C}^{n+1}$ and all $v \in T_z\mathbb{C}^{n+1}$, and the Kähler 2-form is $\omega = d\Lambda/2$. We denote the $(2n+1)$-dimensional unit sphere in $\mathbb{C}^{n+1}$ by $\mathbb{S}^{2n+1}$ and by $\Pi : \mathbb{S}^{2n+1} \to \mathbb{CP}^n$, $\Pi(z) = [z]$, the Hopf fibration of $\mathbb{S}^{2n+1}$ on the complex projective space $\mathbb{CP}^n$. We denote the Fubini–Study metric, the complex structure and the Kähler two-form in $\mathbb{CP}^n$ by $\langle\, ,\rangle$, J and ω. This metric has constant holomorphic sectional curvature 4.

We will also denote by Λ the restriction to $\mathbb{S}^{2n+1}$ of the Liouville 1-form of $\mathbb{C}^{n+1}$. So Λ is the contact 1-form of the canonical Sasakian structure on the sphere $\mathbb{S}^{2n+1}$. An immersion $\phi : M^n \to \mathbb{S}^{2n+1}$ of an n-dimensional manifold M is said to be *Legendrian* if $\phi^*\Lambda \equiv 0$. In this case ϕ is isotropic in $\mathbb{C}^{n+1}$, that is, $\phi^*\omega \equiv 0$; in particular, the normal bundle $T^\perp M$ splits as $J(TM) \oplus \operatorname{span}\{J\phi\}$. This means that ϕ is horizontal with respect to the Hopf fibration $\Pi : \mathbb{S}^{2n+1} \to \mathbb{CP}^n$, and hence $\Phi = \Pi \circ \phi : M^n \to \mathbb{CP}^n$ is a Lagrangian immersion and the metrics induced on M^n by ϕ and Φ are the same. It is easy to check that $J\phi$ is a totally geodesic normal vector field, so the second fundamental forms of ϕ and Φ are related by

$$\Pi_*(\sigma_\phi(v, w)) = \sigma_\Phi(\Pi_* v, \Pi_* w) \quad \text{for all } v, w \in TM.$$

Thus the mean curvature vector H of ϕ satisfies $\langle H, J\phi\rangle = 0$. In particular, $\phi : M^n \to \mathbb{S}^{2n+1}$ *is minimal if and only if* $\Phi = \Pi \circ \phi : M^n \to \mathbb{CP}^n$ *is minimal.*

In this way, we can construct (minimal) Lagrangian submanifolds in $\mathbb{CP}^n$ by projecting (minimal) Legendrian manifolds in $\mathbb{S}^{2n+1}$ via the Hopf fibration Π.

Conversely, any Lagrangian immersion $\Phi : M^n \to \mathbb{CP}^n$ has a *local* horizontal lift to $\mathbb{S}^{2n+1}$ with respect to the Hopf fibration Π; this local lift is unique up to rotations. Only Lagrangian immersions in $\mathbb{CP}^n$ have such lifts.

In this article we construct examples of Lagrangian submanifolds of $\mathbb{CP}^n$ by constructing examples of Legendrian submanifolds of $\mathbb{S}^{2n+1}$. We start with some geometric properties of Legendrian submanifolds in $\mathbb{S}^{2n+1}$.

Let Ω be the complex n-form on $\mathbb{S}^{2n+1}$ given by

$$\Omega_z(v_1, \dots, v_n) = \det_{\mathbb{C}}\{z, v_1, \dots, v_n\}.$$

If $\phi : M^n \to \mathbb{S}^{2n+1}$ is a Legendrian immersion of a manifold M, then $\phi^*\Omega$ is a complex n-form on M. In the next result we analyze this n-form $\phi^*\Omega$.

Lemma 2.1. *If $\phi : M^n \to \mathbb{S}^{2n+1}$ is a Legendrian immersion of a manifold M, then*

$$\nabla(\phi^*\Omega) = \alpha_H \otimes \phi^*\Omega, \tag{1}$$

where α_H is the one-form on M defined by $\alpha_H(v) = n\,i\langle H, Jv\rangle$ and H is the mean curvature vector of ϕ. Consequently, M is orientable if ϕ is minimal.

Proof. Let $\{E_1, \dots, E_n\}$ be an orthonormal frame on an open subset $U \subset M$ containing p, such that $\nabla_v E_i = 0$ for all $v \in T_pM$ and $i = 1, \dots, n$. We define $A : U \to U(n+1)$ by $A = \{\phi, \phi_*(E_1), \dots, \phi_*(E_n)\}$. Then

$$(\nabla_v \phi^*\Omega)(E_1, \dots, E_n) = v(\det_{\mathbb{C}} A) = \det_{\mathbb{C}} A \ \operatorname{Trace}(v(A)\bar{A}^t),$$

where $\bar{A}^t$ denotes the transpose conjugate matrix of A. We easily see that

$$v(A) = \big\{\phi_*(v),\ \sigma_\phi(v, E_1(p)) - \langle v, E_1(p)\rangle\phi,\ \dots,\ \sigma_\phi(v, E_n(p)) - \langle v, E_n(p)\rangle\phi\big\},$$

and so we deduce that

$$(\nabla_v \phi^*\Omega)\big(E_1(p), \dots, E_n(p)\big) = n\,i\langle H(p), Jv\rangle(\phi^*\Omega)(E_1, \dots, E_n)(p).$$

Using this in the preceding expression we get the result. □

Suppose that our Legendrian submanifold M is oriented. Then we can consider the well defined map $\beta : M^n \to \mathbb{R}/2\pi\mathbb{Z}$ given by

$$e^{i\beta(p)} = (\phi^*\Omega)_p(e_1, \dots, e_n),$$

where $\{e_1, \dots, e_n\}$ is an oriented orthonormal frame in T_pM. We will call β the *Legendrian angle* map of ϕ. As a consequence of (1) we obtain

$$J\nabla\beta = nH, \tag{2}$$

and so we deduce:

Proposition 2.2. *A Legendrian immersion $\phi : M^n \to \mathbb{S}^{2n+1}$ of an oriented manifold M is minimal if and only if the Legendrian angle map β of ϕ is constant.*

A vector field X on $\mathbb{S}^{2n+1}$ is a *contact vector field* if $\mathcal{L}_X\Lambda = g\Lambda$, for some function $g \in C^\infty(\mathbb{S}^{2n+1})$, where $\mathcal{L}$ is the Lie derivative in $\mathbb{S}^{2n+1}$. As shown in [McDuff and Salamon 1998], for instance, X is a contact vector field if and only if there exists $F \in C^\infty(\mathbb{S}^{2n+1})$ such that

$$X_z = J(\overline{\nabla} F)_z + 2FJz, \quad z \in \mathbb{S}^{2n+1},$$

where $\overline{\nabla} F$ is the gradient of F. The diffeomorphisms of the flux $\{\varphi_t\}$ of X are contactomorphisms of $\mathbb{S}^{2n+1}$, that is, $\varphi_t^* \Lambda = e^{h_t} \Lambda$, and so they transform Legendrian submanifolds into same. The Lie algebra of the group of contactomorphisms of $\mathbb{S}^{2n+1}$ is the space of contact vector fields. In this setting, it is natural to study the following variational problem.

Let $\phi : M^n \to \mathbb{S}^{2n+1}$ a Legendrian immersion with mean curvature vector H. A normal vector field ξ_f to ϕ is called a *contact field* if

$$\xi_f = J\nabla f + 2 f J\phi,$$

where $f \in C^\infty(M)$ and ∇f is the gradient of f with respect to the induced metric. If $f \in C_0^\infty(M)$ and $\{\phi_t : M \to \mathbb{S}^{2n+1}\}$ is a variation of ϕ with $\phi_0 = \phi$ and $\frac{d}{dt}\big|_{t=0} \phi_t = \xi_f$, the first variation of the volume functional is given by

$$\frac{d}{dt}\bigg|_{t=0} \operatorname{vol}(M, \phi_t^* \langle\, ,\rangle) = -\int_M \langle H, \xi_f \rangle \, dM.$$

But using Stokes' Theorem,

$$\begin{aligned} \int_M \langle H, \xi_f \rangle \, dM &= \int_M \langle H, J\nabla f + 2 f J\phi \rangle \, dM \\ &= -\int_M \langle JH, \nabla f \rangle \, dM = \int_M f \operatorname{div} JH \, dM. \end{aligned}$$

This means that the critical points of the above variational problem are Legendrian submanifolds such that

$$\operatorname{div} JH = 0.$$

Definition 2.3. A Legendrian immersion $\phi : M^n \to \mathbb{S}^{2n+1}$ is said to be *contact minimal* (or briefly *C-minimal*) if it is a critical point of the preceding variational problem, that is, if $\operatorname{div} JH = 0$.

Clearly, minimal Legendrian submanifolds and Legendrian submanifolds with parallel mean curvature vector are C-minimal. As a consequence of (2) and the geometric relationship between Legendrian and Lagrangian submanifolds mentioned at the beginning of this section, we get:

Proposition 2.4. *Let $\phi : M^n \to \mathbb{S}^{2n+1}$ be a Legendrian immersion of a Riemannian manifold M.*

(1) *If M is oriented, ϕ is C-minimal if and only if the Legendrian angle β of ϕ is a harmonic map.*

(2) *ϕ is C-minimal if and only if $\Phi = \Pi \circ \phi : M^n \to \mathbb{CP}^n$ is H-minimal.*

3. A new construction of C-minimal Legendrian immersions

After Proposition 2.4, it is clear that constructing C-minimal Legendrian immersions in odd-dimensional spheres is a good way to find H-minimal Lagrangian submanifolds in $\mathbb{CP}^n$. This is the purpose of this section.

Let $n_1, n_2 \geq 0$ be integers with $n = n_1 + n_2 + 1$. The product $\mathrm{SO}(n_1+1) \times \mathrm{SO}(n_2+1)$ of special orthogonal groups acts on $\mathbb{S}^{2n+1} \subset \mathbb{C}^{n+1}$ as a subgroup of isometries:

$$(A_1, A_2) \in \mathrm{SO}(n_1+1) \times \mathrm{SO}(n_2+1) \mapsto \left(\begin{array}{c|c} A_1 & \\ \hline & A_2 \end{array}\right) \in \mathrm{SO}(n+1). \tag{3}$$

Theorem 3.1. *Let n, n_1, n_2 be nonnegative integers with $n = 1 + n_1 + n_2$. For $i = 1, 2$, let $\psi_i : N_i \to \mathbb{S}^{2n_i+1} \subset \mathbb{C}^{n_i+1}$ be Legendrian isometric immersions of n_i-dimensional oriented Riemannian manifolds (N_i, g_i). Suppose $\gamma = (\gamma_1, \gamma_2) : I \to \mathbb{S}^3 \subset \mathbb{C}^2$ is a Legendre curve, where I is an interval in $\mathbb{R}$. The map*

$$\phi : I \times N_1 \times N_2 \to \mathbb{S}^{2n+1} \subset \mathbb{C}^{n+1} = \mathbb{C}^{n_1+1} \times \mathbb{C}^{n_2+1}$$

defined by

$$\phi(s, p, q) = \big(\gamma_1(s)\psi_1(p), \gamma_2(s)\psi_2(q)\big) \tag{4}$$

is a Legendrian immersion in $\mathbb{S}^{2n+1}$ whose induced metric is

$$\langle\, , \rangle = |\gamma'|^2 ds^2 + |\gamma_1|^2 g_1 + |\gamma_2|^2 g_2 \tag{5}$$

and whose Legendrian angle map is

$$\beta_\phi \equiv n_1\pi + \beta_\gamma + n_1 \arg\gamma_1 + n_2 \arg\gamma_2 + \beta_{\psi_1} + \beta_{\psi_2} \mod 2\pi, \tag{6}$$

where β_γ, β_{ψ_1} and β_{ψ_2} are the Legendre angle maps of γ, ψ_1 and ψ_2.

If $n_1, n_2 \geq 2$, a Legendrian immersion $M^n \to \mathbb{S}^{2n+1}$ is invariant under the action (3) of $\mathrm{SO}(n_1{+}1) \times \mathrm{SO}(n_2{+}1)$ if and only if it is locally of the form (4), where ψ_i $(i = 1, 2)$ is the totally geodesic Legendrian embedding of $\mathbb{S}^{n_i}$ in $\mathbb{S}^{2n_i+1}$ and γ is some Legendre curve in $\mathbb{S}^3$. That is, such immersions are locally congruent to $(s, x_1, x_2) \mapsto \big(\gamma_1(s)x_1,\ \gamma_2(s)x_2\big)$, where $x_i \in \mathbb{S}^{n_i}$.

Note that Legendrian immersions of the form (4) have singularities at the points $(s, p, q) \in I \times N_1 \times N_2$ where either $\gamma_1(s) = 0$ or $\gamma_2(s) = 0$.

Proof. If $'$ denotes differentiation with respect to s, and v and w are arbitrary tangent vectors to N_1 and N_2 respectively, it is clear that

$$\begin{aligned} \phi_s &= \phi_*(\partial_s, 0, 0) = (\gamma_1'\,\psi_1, \gamma_2'\,\psi_2), \\ \phi_*(v) &:= \phi_*(0, v, 0) = (\gamma_1\,\psi_{1*}(v), 0), \\ \phi_*(w) &:= \phi_*(0, 0, w) = (0, \gamma_2\,\psi_{2*}(w)). \end{aligned}$$

(Recall that g_1, g_2 are the metrics on N_1, N_2 induced by ψ_1, ψ_2.) Because ψ_1 and ψ_2 are Legendrian immersions, we deduce from these equalities that the induced metric on $I \times N_1 \times N_2$ by ϕ is $|\gamma'|^2 ds^2 + |\gamma_1|^2 g_1 + |\gamma_2|^2 g_2$. It follows that, γ, ψ_1 and ψ_2 being Legendrian, so is the immersion ϕ.

To compute the Legendrian angle map β_ϕ, let $\{e_1, \dots, e_{n_1}\}$ and $\{e'_1, \dots, e'_{n_2}\}$ be oriented local orthonormal frames on N_1 and N_2. Then the frame

$$\{u_1, v_1, \dots, v_{n_1}, w_1, \dots, w_{n_2}\} \tag{7}$$

defined by

$$u_1 = \Big(\frac{\partial_s}{|\gamma'|}, 0, 0\Big), \quad v_j = \Big(0, \frac{e_j}{|\gamma_1|}, 0\Big), \quad w_k = \Big(0, 0, \frac{e'_k}{|\gamma_2|}\Big)$$

(with $1 \le j \le n_1,\ 1 \le k \le n_2$) is a local oriented orthonormal frame on $I \times N_1 \times N_2$. Putting

$$\phi = \gamma_1(\psi_1, 0) + \gamma_2(0, \psi_2),$$
$$\phi_*(u_1) = \frac{\gamma'_1}{|\gamma'|}(\psi_1, 0) + \frac{\gamma'_2}{|\gamma'|}(0, \psi_2),$$

we have

$$\begin{aligned} e^{i\beta_\phi} &= \det{}_{\mathbb{C}}\{\phi, \phi_*(u_1), \dots, \phi_*(v_j), \dots, \phi_*(w_k), \dots\} \\ &= \frac{\gamma_1^{n_1}\gamma_2^{n_2}(\gamma_1\gamma'_2 - \gamma'_1\gamma_2)}{|\gamma'||\gamma_1|^{n_1}|\gamma_2|^{n_2}} \\ &\qquad \times \det{}_{\mathbb{C}}\{(\psi_1, 0), (0, \psi_2), \dots, (\psi_{1*}(e_j), 0), \dots, (0, \psi_{2*}(e'_k)), \dots\}. \end{aligned}$$

In this way we obtain

$$e^{i\beta_\phi(s,p,q)} = (-1)^{n_1} e^{i(n_1 \arg\gamma_1 + n_2 \arg\gamma_2)(s)} \frac{(\gamma_1\gamma'_2 - \gamma'_1\gamma_2)(s)}{|\gamma'(s)|} \det{}_{\mathbb{C}} A_1(p) \det{}_{\mathbb{C}} A_2(q),$$

where A_1 and A_2 are the matrices

$$A_1 = \{\psi_1, \psi_{1*}(e_1), \dots, \psi_{1*}(e_{n_1})\},$$
$$A_2 = \{\psi_2, \psi_{2*}(e'_1), \dots, \psi_{2*}(e'_{n_2})\}.$$

Taking into account the definition of the Legendrian angle map given in Section 2, we finally arrive at

$$e^{i\beta_\phi(s,p,q)} = (-1)^{n_1} e^{i(\beta_\gamma + n_1 \arg\gamma_1 + n_2 \arg\gamma_2)(s)} e^{i\beta_{\psi_1}(p)} e^{i\beta_{\psi_2}(q)}.$$

This proves the first part of the result.

Conversely, let $\psi : M^n \to \mathbb{S}^{2n+1} \subset \mathbb{C}^{n+1}$ be a Legendrian immersion that is invariant under the action (3) of $\mathrm{SO}(n_1+1)\times\mathrm{SO}(n_2+1)$. Let p be any point of M and set $z = (z_1, \dots, z_{n+1}) = \psi(p)$. By the invariance assumption, for any matrix $X = (X_1, X_2)$ in the Lie algebra of $\mathrm{SO}(n_1+1)\times\mathrm{SO}(n_2+1)$, the curve $t \mapsto ze^{t\hat{X}}$ given by

$$\hat{X} = \left(\begin{array}{c|c} X_1 & \\ \hline & X_2 \end{array}\right)$$

lies in the submanifold. Thus its tangent vector at $t = 0$ satisfies

$$z\hat{X} \in \psi_*(T_pM).$$

Since ψ is a Legendrian immersion, this implies that

$$\mathrm{Im}(z\hat{X}\hat{Y}\bar{z}^t) = 0$$

for any matrices $X = (X_1, X_2)$, $Y = (Y_1, Y_2)$ in the Lie algebra of $\mathrm{SO}(n_1+1)\times\mathrm{SO}(n_2+1)$. As $n_1 + 1 \geq 3$ and $n_2 + 1 \geq 3$, it is easy to see from the last equation that $\mathrm{Re}(z_1, \dots, z_{n_1+1})$ and $\mathrm{Im}(z_1, \dots, z_{n_1+1})$ are linearly dependent, and so are $\mathrm{Re}(z_{n_1+2}, \dots, z_{n+1})$ and $\mathrm{Im}(z_{n_1+2}, \dots, z_{n+1})$. But $\mathrm{SO}(n_1+1)$ acts transitively on $\mathbb{S}^{n_1}$ and $\mathrm{SO}(n_2+1)$ acts transitively on $\mathbb{S}^{n_2}$; hence z is in the orbit (under the action of $\mathrm{SO}(n_1+1)\times\mathrm{SO}(n_2+1)$ described above) of a point of the form

$$(z_1^0, 0, \dots, 0, z_{n_1+2}^0, 0, \dots, 0),$$

with

$$|z_1^0|^2 = \sum_{i=1}^{n_1+1} |z_i|^2 \quad \text{and} \quad |z_{n_1+2}^0|^2 = \sum_{j=n_1+2}^{n+1} |z_j|^2.$$

This implies that locally ψ is the orbit under the action of $\mathrm{SO}(n_1+1)\times\mathrm{SO}(n_2+1)$ of a curve γ in $\mathbb{C}^2 \equiv \mathbb{C}^n \cap \{z_2 = \dots = z_{n_1+1} = z_{n_1+3} = \dots = z_{n+1} = 0\}$. Therefore M is locally $I \times \mathbb{S}^{n_1} \times \mathbb{S}^{n_2}$, with I an interval in $\mathbb{R}$. Moreover, ψ is given by

$$\psi(s, x, y) = (\gamma_1(s)\, x, \gamma_2(s)\, y),$$

where $\gamma = (\gamma_1, \gamma_2)$ must be a Legendre curve in $\mathbb{S}^3 \subset \mathbb{C}^2$. Finally, as ψ is a Legendrian submanifold, the result follows using the first part of this theorem. □

In the next result we make use of the method described in Theorem 3.1 to obtain new minimal and C-minimal Legendrian immersions, which will provide (projecting via the Hopf fibration) new nontrivial minimal and H-minimal immersions in $\mathbb{CP}^n$.

Corollary 3.2. *Let $\psi_i : N_i \to \mathbb{S}^{2n_i+1}$, $i = 1, 2$, be C-minimal Legendrian immersions of n_i-dimensional oriented Riemannian manifolds N_i, $i = 1, 2$, and let $\gamma = (\gamma_1, \gamma_2) : I \to \mathbb{S}^3 \subset \mathbb{C}^2$ be a Legendre curve. As before, set $n = n_1 + n_2 + 1$. Then the Legendrian immersion $\phi : I \times N_1 \times N_2 \to \mathbb{S}^{2n+1}$ of Theorem 3.1, given by*

$$\phi(t, p, q) = \big(\gamma_1(t)\psi_1(p), \gamma_2(t)\psi_2(q)\big),$$

is C-minimal if and only if there exist real constants λ, μ such that (γ_1, γ_2) is a solution of the system of ordinary differential equations

$$(\gamma_1'\overline{\gamma}_1)(t) = -(\gamma_2'\overline{\gamma}_2)(t) = -\, e^{i(\lambda+\mu t)}\, \overline{\gamma}_1(t)^{n_1+1}\, \overline{\gamma}_2(t)^{n_2+1}. \tag{8}$$

This Legendrian immersion ϕ is minimal if and only if ψ_1 and ψ_2 are minimal and there exists some λ such that (γ_1, γ_2) is a solution of the system (8) with $\mu = 0$.

Remark. If we apply a rotation through θ to a Legendre curve γ that is a solution of (8) with parameters (λ, μ), the new Legendre curve is a solution of the same equation with parameters $(\lambda - (n+1)\theta, \mu)$. The corresponding immersions given in Corollary 3.2 are related by $\tilde{\phi} = e^{i\theta}\phi$ and are therefore congruent. By choosing θ appropriately, then, we can assume that $\lambda = \pi$; that is, it suffices (up to congruence) to consider solutions of the one-parameter family of equations

$$(\gamma_j'\overline{\gamma}_j)(t) = (-1)^{j-1} i\, e^{i\mu t}\, \overline{\gamma}_1(t)^{n_1+1}\, \overline{\gamma}_2(t)^{n_2+1}, \quad \text{with } \mu \in \mathbb{R},\ j = 1, 2. \tag{9}$$

Proof of Corollary 3.2. We know from Proposition 2.4 that ϕ is C-minimal if and only if $\Delta\beta_\phi = 0$, where β_ϕ is given by (6). So we must compute the Laplacian of β_ϕ. We use the orthonormal frame (7) and after a long but direct computation we obtain

$$\Delta\beta_\phi = \frac{1}{|\gamma'|^2}\left(\frac{\partial^2\beta_\phi}{\partial s^2} + \frac{\mathrm{d}}{\mathrm{d}s}\left(\log\frac{|\gamma_1|^{n_1}|\gamma_2|^{n_2}}{|\gamma'|}\right)\frac{\partial\beta_\phi}{\partial s}\right) + \frac{\Delta_1\beta_{\psi_1}}{|\gamma_1|^2} + \frac{\Delta_2\beta_{\psi_2}}{|\gamma_2|^2}, \tag{10}$$

where the Δ_i are the Laplace operators in (N_i, g_i).

The assumptions of the Corollary 3.2 imply that $\Delta_1\beta_{\psi_1} = \Delta_2\beta_{\psi_2} = 0$ again by Proposition 2.4. So ϕ is C-minimal if and only if

$$\frac{\partial^2\beta_\phi}{\partial s^2} + \frac{\mathrm{d}}{\mathrm{d}s}\left(\log\frac{|\gamma_1|^{n_1}|\gamma_2|^{n_2}}{|\gamma'|}\right)\frac{\partial\beta_\phi}{\partial s} = 0. \tag{11}$$

Since we want ϕ to be regular, we impose that $\gamma_1(0)$ and $\gamma_2(0)$ not vanish (see after statement of Theorem 3.1). Up to a reparametrization, we can assume that γ satisfies $|\gamma'(t)| = |\gamma_1(t)|^{n_1}|\gamma_2(t)|^{n_2}$. Thus (11) becomes

$$\frac{\partial^2\beta_\phi}{\partial t^2} = 0.$$

This means that $\beta_\phi(t,p,q)=f(p,q)+t\,g(p,q)$, for certain functions f, g defined on $N_1\times N_2$. Using (6), we obtain that $g(p,q)$ is constant and that

$$(\beta_\gamma+n_1\arg\gamma_1+n_2\arg\gamma_2)(t)=\lambda+\mu t,\quad \text{with } \lambda,\mu\in\mathbb{R}. \tag{12}$$

The definition of the Legendrian angle β_γ of γ is given, in particular, by

$$e^{i\beta_\gamma}=\frac{1}{|\gamma'|}(\gamma_1\gamma_2'-\gamma_2\gamma_1').$$

Using this, it is easy to rewrite (12) as

$$\gamma_1'\bar\gamma_1=-\gamma_2'\bar\gamma_2=-\,e^{i(\lambda+\mu t)}\,\bar\gamma_1^{\,n_1+1}\,\bar\gamma_2^{\,n_2+1},$$

which is exactly (8).

Finally, by Proposition 2.2, ϕ is minimal if and only if β_ϕ is constant. This is equivalent to β_{ψ_1}, β_{ψ_2} being constant (i.e., the ψ_i are minimal, again by the same proposition) and $\beta_\gamma+n_1\arg\gamma_1+n_2\arg\gamma_2$ is constant. But this corresponds to the case $\mu=0$ in (12) and so to the case $\mu=0$ in (8). □

It is difficult to describe the general solution of (9). However it is an exercise to check that for any $\delta\in(0,\pi/2)$ the Legendre curve

$$\gamma_\delta(t)=\big(c_\delta\,\exp(is_\delta^{n_1+1}c_\delta^{n_2-1}t),\,s_\delta\,\exp(-is_\delta^{n_1-1}c_\delta^{n_2+1}t)\big), \tag{13}$$

satisfies (9) for $\mu=s_\delta^{n_1-1}c_\delta^{n_2-1}\big((n_1+1)s_\delta^2-(n_2+1)c_\delta^2\big)$, where $c_\delta=\cos\delta$ and $s_\delta=\sin\delta$. This value of μ vanishes if and only if $\tan^2\delta=(n_2+1)/(n_1+1)$. In this way we are able to obtain an explicit family of examples:

Corollary 3.3. *Let $\psi_i:N_i\to\mathbb{S}^{2n_i+1}$, for $i=1,2$, be C-minimal Legendrian immersions of n_i-dimensional Riemannian manifolds N_i, and let $n=n_1+n_2+1$. Given $\delta\in(0,\pi/2)$, set $c_\delta=\cos\delta$ and $s_\delta=\sin\delta$. Then the map $\phi_\delta:\mathbb{R}\times N_1\times N_2\to\mathbb{S}^{2n+1}$ defined by*

$$\phi_\delta(t,p,q)=\big(c_\delta\,\exp(is_\delta^{n_1+1}c_\delta^{n_2-1}t)\,\psi_1(p)\,,\,s_\delta\,\exp(-is_\delta^{n_1-1}c_\delta^{n_2+1}t)\,\psi_2(q)\big)$$

is a C-minimal Legendrian immersion.

In particular, using minimal Legendrian immersions ψ_1, ψ_2 and the value $\delta=\delta_0:=\arctan\sqrt{(n_2+1)/(n_1+1)}$, we obtain a minimal Legendrian immersion $\phi_{\delta_0}:\mathbb{R}\times N_1\times N_2\to\mathbb{S}^{2n+1}$.

Proof. We simply remark that we do not need the orientability assumption because, in the case at hand, the Legendrian immersions ϕ_δ are easily seen to satisfy $\operatorname{div}JH=0$ and thus are C-minimal (see Definition 2.3). □

To finish this section, we turn our attention to Equation (9) with $\mu=0$. We observe that this is exactly equation (6) in [Castro and Urbano 2004, Lemma 2] (in the notation of that paper, put $p=n_1$ and $q=n_2$). If we choose the initial conditions $\gamma(0) = (\cos\theta, \sin\theta)$, with $\theta \in (0, \pi/2)$, we can make use of the study made in that reference.

Lemma 3.4. *Let $\gamma_\theta = (\gamma_1, \gamma_2) : I \subset \mathbb{R} \to \mathbb{S}^3$ be the unique curve solution of*

$$\gamma_j' \bar{\gamma}_j = (-1)^{j-1} i\, \bar{\gamma}_1^{n_1+1} \bar{\gamma}_2^{n_2+1}, \ \ j = 1, 2, \tag{14}$$

satisfying the real initial conditions $\gamma_\theta(0) = (\cos\theta, \sin\theta)$, $\theta \in (0, \pi/2)$.

(1) $\operatorname{Re}(\gamma_1^{n_1+1} \gamma_2^{n_2+1}) = \cos^{n_1+1}\theta\ \sin^{n_2+1}\theta$.

(2) *For $j = 1, 2$ and any $t \in I$, we have $\bar{\gamma}_j(t) = \gamma_j(-t)$.*

(3) *The functions $|\gamma_1|$ and $|\gamma_2|$ are periodic with the same period $T = T(\theta)$, and γ_θ is a closed curve if and only if*

$$\theta \in (0, \pi/2) \quad \textit{and} \quad \frac{\cos^{n_1+1}\theta \sin^{n_2+1}\theta}{2\pi} \left(\int_0^T \frac{dt}{|\gamma_1|^2(t)}, \int_0^T \frac{dt}{|\gamma_2|^2(t)} \right) \in \mathbb{Q}^2.$$

(4) *If θ takes the value $\delta_0 = \arctan\sqrt{(n_2{+}1)/(n_1{+}1)}$ from Corollary 3.3, we recover the curve of Equation (13), with $\delta = \delta_0$.*

Proof. (1) and (2) follow directly from parts 2 and 3 of [Castro and Urbano 2004, Lemma 2]. To prove (3) we set $f(\theta) = \cos^{2(n_1+1)}\theta \sin^{2(n_2+1)}\theta$, for $\theta \in (0, \pi/2)$. It is easy to prove that $f(\theta) \le (n_1+1)^{n_1+1}(n_2+1)^{n_2+1}/(n+1)^{n+1}$ and the equality holds if and only if $\theta = \delta_0$. Using this in parts 4 and 5 of [Castro and Urbano 2004, Lemma 2] completes the proof. □

4. H-minimal Lagrangian submanifolds in complex projective space

In Section 2 we explained that we can construct (minimal, H-minimal) Lagrangian submanifolds in $\mathbb{CP}^n$ by projecting (minimal, C-minimal) Legendrian submanifolds in $\mathbb{S}^{2n+1}$ by the Hopf fibration $\Pi : \mathbb{S}^{2n+1} \to \mathbb{CP}^n$ (Proposition 2.4). The aim of this section is to analyze the Lagrangian immersions in $\mathbb{CP}^n$ that we obtain just by projecting the Legendrian ones deduced in Section 3.

First we mention that if $n_2 = 0$ in Theorem 3.1, projection by the Hopf fibration Π yields Examples 1 of [Castro et al. 2001]. In this sense, the construction given in Theorem 3.1 can be considered as a generalization of the family introduced in that reference. Some applications of our construction of Theorem 3.1 when $n = 3$ have been used recently in [Montealegre and Vrancken 2006] to the study of minimal Lagrangian submanifolds in $\mathbb{CP}^3$.

The Legendrian immersions described in Corollary 3.2 provide new examples of Lagrangian H-minimal immersions in $\mathbb{CP}^n$ when we project them via Π. If we consider the case $n_2 = 0$ (so $n_1 = n - 1$) in the minimal case of Corollary 3.2, we recover (by projecting via the Hopf fibration Π) the minimal Lagrangian submanifolds of $\mathbb{CP}^n$ described in [Castro et al. 2002, Proposition 6], although we used there a unit speed parametrization for γ.

We write in more detail what we obtain with this procedure if we consider the special case coming from Corollary 3.3.

Corollary 4.1. *Let $\psi_i : N_i \to \mathbb{S}^{2n_i+1}$, for $i = 1, 2$, be C-minimal Legendrian immersions of n_i-dimensional Riemannian manifolds N_i, and let $n = n_1 + n_2 + 1$. Suppose $\delta \in (0, \pi/2)$. Then the map $\Phi_\delta : \mathbb{S}^1 \times N_1 \times N_2 \to \mathbb{CP}^n$ given by*

$$\Phi_\delta(e^{is}, p, q) = \big[\big(\cos\delta \exp(is\sin^2\delta)\psi_1(p),\ \sin\delta \exp(-is\cos^2\delta)\psi_2(q)\big)\big]$$

is an H-minimal Lagrangian immersion. Φ_δ is minimal if and only if ψ_1 and ψ_2 are minimal and $\tan^2\delta = (n_2+1)/(n_1+1)$. (Recall that the brackets denote the image under Π.)

Proof. We consider the C-minimal Legendrian immersions

$$\phi_\delta : \mathbb{R} \times N_1 \times N_2 \to \mathbb{S}^{2n+1}$$

given in Corollary 3.3. Projecting via the Hopf fibration $\Pi : \mathbb{S}^{2n+1} \to \mathbb{CP}^n$ and using Proposition 2.4 we conclude that

$$\Pi \circ \phi_\delta : \mathbb{R} \times N_1 \times N_2 \to \mathbb{CP}^n$$

is a one-parameter family of H-minimal Lagrangian immersions. We study when $\Pi \circ \phi_\delta$ is periodic in its first variable. It is easy to see that there exists $A > 0$ such that $(\Pi \circ \phi_\delta)(t + A, p, q) = (\Pi \circ \phi_\delta)(t, p, q)$, $\forall (t, p, q) \in \mathbb{R} \times N_1 \times N_2$ if and only if there exists $\theta \in \mathbb{R}$ satisfying

$$\exp(is_\delta^{n_1+1} c_\delta^{n_2-1} A) = e^{i\theta} = \exp(-is_\delta^{n_1-1} c_\delta^{n_2+1} A).$$

We deduce that the smallest period A must equal $A = 2\pi/(s_\delta^{n_1-1} c_\delta^{n_2-1})$. Applying the change of variables

$$s \mapsto t = s/(s_\delta^{n_1-1} c_\delta^{n_2-1})$$

for $s \in [0, 2\pi]$, the equation for the Legendre curve γ_δ of (13) becomes

$$\gamma_\delta(s) = \big(c_\delta \exp(is_\delta^2 s), s_\delta \exp(-ic_\delta^2 s)\big),\ \ s \in [0, 2\pi],$$

which leads to the expression of Φ_δ.

We conclude the proof by observing that $\Pi \circ \phi_\delta$ is minimal if and only if ϕ_δ is minimal (see Section 2) and using Corollary 3.3 again. □

We get H-minimal Lagrangian embeddings as a particular case:

Corollary 4.2. *Let $\delta \in (0, \pi/2)$ and $n = n_1+n_2+1$. The immersion Φ_δ of Corollary 3.3, where ψ_i, for $i = 1, 2$, is the totally geodesic Legendrian embedding of $\mathbb{S}^{n_i}$ into $\mathbb{S}^{2n_i+1}$, gives rise to an H-minimal Lagrangian embedding*

$$\overline{(e^{is}, x, y)} \mapsto \big[\big(\cos\delta\ \exp(is\sin^2\delta)x, \sin\delta\ \exp(-is\cos^2\delta)y\big)\big]$$

of the quotient $(\mathbb{S}^1\times\mathbb{S}^{n_1}\times\mathbb{S}^{n_2})/(\mathbb{Z}_2\times\mathbb{Z}_2)$ into $\mathbb{CP}^n$, the action of $\mathbb{Z}_2\times\mathbb{Z}_2$ being generated by the involutions $(e^{is},x,y)\mapsto(-e^{is},-x,y)$, $(e^{is}, x, y)\mapsto(-e^{is}, x, -y)$.

Proof. Consider the H-minimal Lagrangian immersion $\Phi_\delta : \mathbb{S}^1\times\mathbb{S}^{n_1}\times\mathbb{S}^{n_2}\to\mathbb{CP}^n$ defined by

$$\Phi_\delta(e^{is}, x, y) = \big[\big(\cos\delta\ \exp(i\sin^2\delta\, s)\, x, \sin\delta\ \exp(-i\cos^2\delta\, s)\, y\big)\big].$$

Take $(e^{is}, x, y), (e^{i\hat{s}}, \hat{x}, \hat{y}) \in \mathbb{S}^1\times\mathbb{S}^{n_1}\times\mathbb{S}^{n_2}$. Then $\Phi_\delta(e^{is}, x, y) = \Phi_\delta(e^{i\hat{s}}, \hat{x}, \hat{y})$ if and only if there exists $\theta\in\mathbb{R}$ such that

$$\hat{x} = \exp\big(i(\theta+\sin^2\delta(s-\hat{s}))\big)\, x, \quad \hat{y} = \exp\big(i(\theta-\cos^2\delta(s-\hat{s}))\big)\, y. \tag{15}$$

Since some coordinate of $x\in\mathbb{S}^{n_1}$ and $y\in\mathbb{S}^{n_2}$ is nonzero, we deduce that

$$\begin{aligned}\epsilon_1 &:= \exp\big(i(\theta+\sin^2\delta(s-\hat{s}))\big) = \pm 1,\\ \epsilon_2 &:= \exp\big(i(\theta-\cos^2\delta(s-\hat{s}))\big) = \pm 1.\end{aligned} \tag{16}$$

We distinguish two cases:

(i) $\epsilon_1 = \epsilon_2$: From (16) we get $e^{i\hat{s}} = e^{is}$; using (15) we obtain $\hat{x} = x$, $\hat{y} = y$ if $\epsilon_1 = \epsilon_2 = 1$ or $\hat{x} = -x$, $\hat{y} = -y$ if $\epsilon_1 = \epsilon_2 = -1$. In either case $(e^{i\hat{s}}, \hat{x}, \hat{y})$ and (e^{is}, x, y) are equivalent under the $\mathbb{Z}_2\times\mathbb{Z}_2$ action.

(ii) $\epsilon_1 = -\epsilon_2$: From (16) we get $e^{i\hat{s}} = -e^{is}$ and using (15) we obtain that either $\hat{x} = x$ and $\hat{y} = -y$ or $\hat{x} = -x$ and $\hat{y} = y$. Again we see that $(e^{i\hat{s}}, \hat{x}, \hat{y})$ and (e^{is}, x, y) are equivalent under the $\mathbb{Z}_2\times\mathbb{Z}_2$ action. □

If $\tan^2\delta = (n_2+1)/(n_1+1)$ the minimal Lagrangian embedding of Corollary 4.2 admits as a special case ($n_2 = 0$) the example $(\mathbb{S}^1\times\mathbb{S}^{n-1})/\mathbb{Z}_2\to\mathbb{CP}^n$ studied in [Naitoh 1981].

Remark 4.3. As can easily be checked, the action of $\mathbb{Z}_2\times\mathbb{Z}_2$ on $\mathbb{S}^1\times\mathbb{S}^{n_1}\times\mathbb{S}^{n_2}$ preserves orientation (and hence the quotient is an orientable manifold) if and only if both n_1 and n_2 are odd.

To conclude this section, we use the information given by Lemma 3.4 on the solutions of equation (9) with $\mu = 0$.

Assume $\theta\in(0,\pi/2)$ and let γ_θ be the only solution of (14) satisfying $\gamma_\theta(0) = (\cos\theta, \sin\theta)$. Consider the C-minimal Legendrian immersions

$$\phi_\theta : I\times N_1\times N_2\to\mathbb{S}^{2n+1}$$

constructed with γ_θ. Projecting by the Hopf fibration $\Pi : \mathbb{S}^{2n+1} \to \mathbb{CP}^n$ and using Proposition 2.4 we obtain a one-parameter family

$$\Pi \circ \phi_\theta : I \times N_1 \times N_2 \to \mathbb{CP}^n$$

of H-minimal Lagrangian immersions.

Lemma 3.4(3) tells us when γ_θ is a closed curve, but now we want to find when $\Pi \circ \phi_\theta$ is periodic of period T, say, in its first variable. Write $\gamma_\theta = (\rho_1 e^{i\nu_1}, \rho_2 e^{i\nu_2})$; then $\rho_i(t+T) = \rho_i(t)$ for $i = 1, 2$. It is not hard to deduce that there exists $A > 0$ such that $(\Pi \circ \phi_\theta)(t + A, p, q) = (\Pi \circ \phi_\theta)(t, p, q)$ if and only if there exist $\nu \in \mathbb{R}$ and $m \in \mathbb{Z}$ satisfying

$$e^{i\nu_j(t+mT)} = e^{i\nu} e^{i\nu_j(t)}, \quad j = 1, 2 \tag{17}$$

(and then $A = mT$). From (14) we can deduce that

$$\rho_j^2 \nu_j' = (-1)^{j-1} c_\theta^{n_1+1} s_\theta^{n_2+1}, \quad j = 1, 2. \tag{18}$$

Then it is easy to check that $\nu_j(t + mT) = \nu_j(t) + m\nu_j(T)$, $j = 1, 2$, and (17) is equivalent to $e^{im\nu_j(T)} = e^{i\nu}$, $j = 1, 2$. This means that $(\nu_2(T) - \nu_1(T))/2\pi$ must be a rational number. In view of (18), this implies that θ lies in

$$\Gamma := \left\{ \alpha \in \left(0, \frac{\pi}{2}\right) : \frac{\cos^{n_1+1}\alpha \sin^{n_2+1}\alpha}{2\pi} \int_0^T \frac{dt}{|\gamma_1|^2(t)\,|\gamma_2|^2(t)} \in \mathbb{Q} \right\}.$$

Hence:

Corollary 4.4. *For $\theta \in \Gamma$ and fixed C-minimal Legendrian immersions $\psi_i : N_i \to \mathbb{S}^{2n_i+1}$, $i = 1, 2$, we obtain from ϕ_θ a one-parameter family of H-minimal Lagrangian immersions*

$$\Phi_\theta : \mathbb{S}^1 \times N_1 \times N_2 \to \mathbb{CP}^n, \quad n = n_1 + n_2 + 1, \;\; \theta \in \Gamma.$$

In particular, Φ_θ is minimal if and only if ψ_1 and ψ_2 are.

5. H-minimal Lagrangian cones in complex Euclidean space

Given a Legendrian immersion $\phi : M^n \to \mathbb{S}^{2n+1}$, the *cone* with link ϕ in $\mathbb{C}^{n+1}$ is the map $C(\phi) : \mathbb{R} \times M^n \to \mathbb{C}^{n+1}$ given by

$$(s, p) \mapsto s\,\phi(p).$$

$C(\phi)$ is a Lagrangian immersion with singularities at $s = 0$.

M. Haskins [2004b; 2004a] has studied in depth special Lagrangian cones using the fact that ϕ is minimal if and only if $C(\phi)$ is minimal. Following a reasoning similar to Haskin's, a straightforward computation leads to the next result.

Proposition 5.1. *Let $\phi : M^n \to \mathbb{S}^{2n+1}$ be a Legendrian immersion of an oriented manifold M and $C(\phi) : \mathbb{R} \times M \to \mathbb{C}^{n+1}$ the cone with link ϕ. Then ϕ is C-minimal if and only if $C(\phi)$ is H-minimal.*

Thanks to Proposition 5.1 we have a fruitful and simple construction method for examples of H-minimal Lagrangian cones in $\mathbb{C}^{n+1}$ using the C-minimal Legendrian immersions described in Section 3.

6. The complex hyperbolic case

In this section we summarize the analogous results when the ambient space is complex hyperbolic space. We omit proofs.

Let $\mathbb{C}_1^{n+1}$ be complex Euclidean space $\mathbb{C}^{n+1}$ endowed with the indefinite metric $\langle\,,\rangle = \operatorname{Re}(\,,)$, where

$$(z, w) = \sum_{i=1}^{n} z_i \overline{w}_i - z_{n+1}\overline{w}_{n+1}$$

for $z, w \in \mathbb{C}^{n+1}$, here $\bar{z}$ stands for the conjugate of z. The Liouville 1-form is given by $\Lambda_z(v) = \langle v, Jz\rangle$, for all $z \in \mathbb{C}^{n+1}$ and all $v \in T_z\mathbb{C}^{n+1}$, and the Kähler 2-form is $\omega = d\Lambda/2$. We denote by $\mathbb{H}_1^{2n+1}$ the anti-de Sitter space, defined as the hypersurface of $\mathbb{C}_1^{n+1}$ given by

$$\mathbb{H}_1^{2n+1} = \{z \in \mathbb{C}^{n+1} \,/\, (z, z) = -1\},$$

and by $\Pi : \mathbb{H}_1^{2n+1} \to \mathbb{CH}^n$, $\Pi(z) = [z]$, the Hopf fibration of $\mathbb{H}_1^{2n+1}$ onto complex hyperbolic space $\mathbb{CH}^n$. The metric, complex structure and Kähler two-form in $\mathbb{CH}^n$ are written $\langle\,,\rangle$, J and ω. This metric has constant holomorphic sectional curvature -4. We also denote by Λ the restriction to $\mathbb{H}_1^{2n+1}$ of the Liouville 1-form of $\mathbb{C}_1^{n+1}$. Thus Λ is the contact 1-form of the canonical (indefinite) Sasakian structure on the anti-de Sitter space $\mathbb{H}_1^{2n+1}$. An immersion $\phi : M^n \to \mathbb{H}_1^{2n+1}$ of an n-dimensional manifold M is said to be *Legendrian* if $\phi^*\Lambda \equiv 0$. So ϕ is isotropic in $\mathbb{C}_1^{n+1}$, that is, $\phi^*\omega \equiv 0$. In particular, the normal bundle $T^\perp M$ has the decomposition $J(TM) \oplus \operatorname{span}\{J\phi\}$. This means that ϕ is horizontal with respect to the Hopf fibration $\Pi : \mathbb{H}_1^{2n+1} \to \mathbb{CH}^n$, and hence $\Phi = \Pi \circ \phi : M^n \to \mathbb{CH}^n$ is a Lagrangian immersion and the induced metrics on M^n by ϕ and Φ are the same.

It is easy to check that $J\phi$ is a totally geodesic normal vector field, so the second fundamental forms of ϕ and Φ are related by

$$\Pi_*(\sigma_\phi(v, w)) = \sigma_\Phi(\Pi_* v, \Pi_* w) \quad \text{for all } v, w \in TM.$$

Thus the mean curvature vector H of ϕ satisfies $\langle H, J\phi\rangle = 0$. In particular, $\phi : M^n \to \mathbb{H}_1^{2n+1}$ *is minimal if and only if* $\Phi = \Pi \circ \phi : M^n \to \mathbb{CH}^n$ *is minimal.*

In this way, we can construct (minimal) Lagrangian submanifolds in $\mathbb{CH}^n$ by projecting (minimal) Legendrian manifolds in $\mathbb{H}_1^{2n+1}$ via the Hopf fibration Π.

Let Ω be the complex n-form on $\mathbb{H}_1^{2n+1}$ given by

$$\Omega_z(v_1, \dots, v_n) = \det_{\mathbb{C}}\{z, v_1, \dots, v_n\}.$$

If $\phi : M^n \to \mathbb{H}_1^{2n+1}$ is a Legendrian immersion of a manifold M, then $\phi^*\Omega$ is a complex n-form on M. In the following result we analyze this n-form $\phi^*\Omega$.

Lemma 6.1. *If $\phi : M^n \to \mathbb{H}_1^{2n+1}$ is a Legendrian immersion of a manifold M, then*

$$\nabla(\phi^*\Omega) = \alpha_H \otimes \phi^*\Omega, \tag{19}$$

where α_H is the one-form on M defined by $\alpha_H(v) = n i \langle H, Jv\rangle$ and H is the mean curvature vector of ϕ. Consequently, M is orientable if ϕ is minimal.

Suppose that our Legendrian submanifold M is oriented. Consider the well defined map $\beta : M^n \to \mathbb{R}/2\pi\mathbb{Z}$ given by

$$e^{i\beta(p)} = (\phi^*\Omega)_p(e_1, \dots, e_n)$$

where $\{e_1, \dots, e_n\}$ is an oriented orthonormal frame in T_pM. We will call β the *Legendrian angle* map of ϕ. As a consequence of (19) we obtain

$$J\nabla\beta = nH,$$

and so we deduce:

Proposition 6.2. *Let $\phi : M^n \to \mathbb{H}_1^{2n+1}$ be a Legendrian immersion of an oriented manifold M. Then ϕ is minimal if and only if the Legendrian angle map β of ϕ is constant.*

In this context we can also consider *contact minimal* (or briefly *C-minimal*) Legendrian submanifolds of $\mathbb{H}_1^{2n+1}$ as critical points of the volume functional for compactly supported variations with variational vector field a (normal) contact field $\xi_f = J\nabla f - 2fJ\phi$, where f lies in $C_0^\infty(M)$ and ∇f is the gradient of f respect to the induced metric. Such fields are also characterized by the equation $\operatorname{div} JH = 0$, and we have a counterpart to Proposition 2.4:

Proposition 6.3. *Let $\phi : M^n \to \mathbb{H}_1^{2n+1}$ be a Legendrian immersion of a Riemannian manifold M.*

(1) *If M is oriented, ϕ is C-minimal if and only if the Legendrian angle β of ϕ is a harmonic map.*

(2) *ϕ is C-minimal if and only if $\Phi = \Pi \circ \phi : M^n \to \mathbb{CH}^n$ is H-minimal.*

The identity component of the indefinite special orthogonal group will be denoted by $SO_0^1(m)$. So $SO(n_1+1)\times SO_0^1(n_2+1)$ acts on $\mathbb{H}_1^{2n+1}\subset\mathbb{C}^{n+1}$, where $n=n_1+n_2+1$, as a subgroup of isometries:

$$(20)\qquad (A_1, A_2)\in SO(n_1+1)\times SO_0^1(n_2+1)\mapsto \left(\begin{array}{c|c} A_1 & \\ \hline & A_2 \end{array}\right)\in SO_0^1(n+1).$$

We now state the main results of Section 3 adapted to this context. We denote by $\mathbb{RH}^n=\left\{(y_1,\dots,y_{n+1})\in\mathbb{R}^{n+1} : \sum_{i=1}^n y_i^2-y_{n+1}^2=-1,\ y_{n+1}>0\right\}$ the real hyperbolic space of dimension n.

Theorem 6.4. *Let n, n_1, n_2 be nonnegative integers with $n=1+n_1+n_2$. Let $\psi_1 : N_1\to\mathbb{S}^{2n_1+1}\subset\mathbb{C}^{n_1+1}$ and $\psi_2 : N_2\to\mathbb{H}_1^{2n_2+1}\subset\mathbb{C}^{n_2+1}$ be Legendrian immersions of n_i-dimensional oriented Riemannian manifolds (N_i, g_i). Suppose $\gamma=(\gamma_1,\gamma_2) : I\to\mathbb{H}_1^3\subset\mathbb{C}^2$ is a Legendre curve. The map*

$$\phi : I\times N_1\times N_2\to\mathbb{H}_1^{2n+1}\subset\mathbb{C}^{n+1}=\mathbb{C}^{n_1+1}\times\mathbb{C}^{n_2+1}$$

defined by

$$(21)\qquad \phi(s,p,q)=\big(\gamma_1(s)\psi_1(p),\ \gamma_2(s)\psi_2(q)\big)$$

is a Legendrian immersion in $\mathbb{H}_1^{2n+1}$ whose induced metric is

$$(22)\qquad \langle\,,\rangle=|\gamma'|^2ds^2+|\gamma_1|^2g_1+|\gamma_2|^2g_2$$

and whose Legendrian angle map is

$$(23)\qquad \beta_\phi\equiv n_1\pi+\beta_\gamma+n_1\arg\gamma_1+n_2\arg\gamma_2+\beta_{\psi_1}+\beta_{\psi_2}\mod 2\pi,$$

where β_γ, β_{ψ_1} and β_{ψ_2} are the Legendre angle maps of γ, ψ_1 and ψ_2.

If $n_1, n_2\geq 2$, a Legendrian immersion $M^n\to\mathbb{H}_1^{2n+1}$ is invariant under the action (20) of $SO(n_1+1)\times SO_0^1(n_2+1)$ if and only if it is locally of the form (21), where ψ_1 is the totally geodesic Legendrian embedding of $\mathbb{S}^{n_1}$ in $\mathbb{S}^{2n_1+1}$ and ψ_2 is the totally geodesic Legendrian embedding of $\mathbb{RH}^{n_2}$ in $\mathbb{H}_1^{2n_2+1}$ That is, such immersions are locally congruent to $\phi(s,x,y)=(\gamma_1(s)x,\gamma_2(s)y)$, where $x\in\mathbb{S}^{n_1}$, $y\in\mathbb{RH}^{n_2}$.

Remark. If $n_2=0$ in the theorem, we recover Examples 2 of [Castro et al. 2001] by projection via the Hopf fibration $\Pi : \mathbb{H}_1^{2n+1}\to\mathbb{CH}^n$. When $n_1=0$ we obtain Examples 3.

Corollary 6.5. *Let $\psi_1 : N_1\to\mathbb{S}^{2n_1+1}\subset\mathbb{C}^{n_1+1}$ and $\psi_2 : N_2\to\mathbb{H}_1^{2n_2+1}\subset\mathbb{C}^{n_2+1}$ be C-minimal Legendrian immersions of n_i-dimensional oriented Riemannian manifolds N_i, $i=1,2$, and let $\gamma=(\gamma_1,\gamma_2) : I\to\mathbb{H}_1^3\subset\mathbb{C}^2$ be a Legendre curve. As before, set $n=n_1+n_2+1$. Then the Legendrian immersion $\phi : I\times N_1\times N_2\to\mathbb{H}_1^{2n+1}$ of*

Theorem 6.4, given by

$$\phi(t,p,q) = \big(\gamma_1(t)\psi_1(p), \gamma_2(t)\psi_2(q)\big),$$

is C-minimal if and only if, up to congruences, there exists a real constant μ such that (γ_1,γ_2) is a solution of the system of ordinary differential equations

$$(\gamma_1'\bar\gamma_1)(t) = (\gamma_2'\bar\gamma_2)(t) = i\, e^{i\mu t}\,\bar\gamma_1(t)^{n_1+1}\,\bar\gamma_2(t)^{n_2+1}. \tag{24}$$

This Legendrian immersion ϕ is minimal if and only if ψ_1 and ψ_2 are minimal and (γ_1,γ_2) is a solution of (24) with $\mu = 0$.

If we consider the particular cases $n_2 = 0$ and $n_1 = 0$ in the minimal case of Corollary 6.5, we recover (projecting via the Hopf fibration Π) the minimal Lagrangian submanifolds of $\mathbb{CH}^n$ described in [Castro et al. 2002, Propositions 3 and 5], although we used there a unit speed parametrization for γ.

From these two last results we can get similar examples to the ones given in Section 4 in the projective case. Concretely, it is easy to check that for any $\rho > 0$ the Legendre curve

$$\gamma_\rho(t) = (s_\rho\,\exp(i\,s_\rho^{n_1-1}c_\rho^{n_2+1}t), c_\rho\,\exp(i\,s_\rho^{n_1+1}c_\rho^{n_2-1}t)), \tag{25}$$

satisfies (24) for $\mu = s_\rho^{n_1-1}c_\rho^{n_2-1}\big((n_1+1)c_\rho^2 + (n_2+1)s_\rho^2\big)$, where $c_\rho = \cosh\rho$, $s_\rho = \sinh\rho$.

Hence an analogous reasoning to that in Corollary 4.1 yields following explicit family of examples.

Corollary 6.6. *Let $\psi_1 : N_1 \to \mathbb{S}^{2n_1+1} \subset \mathbb{C}^{n_1+1}$ and $\psi_2 : N_2 \to \mathbb{H}_1^{2n_2+1} \subset \mathbb{C}^{n_2+1}$ be C-minimal Legendrian immersions of n_i-dimensional Riemannian manifolds N_i, $i = 1, 2$, and let $n = n_1 + n_2 + 1$. Given $\rho > 0$, set $c_\rho = \cosh\rho$ and $s_\rho = \sinh\rho$. Then the map $\Phi_\rho : \mathbb{S}^1 \times N_1 \times N_2 \to \mathbb{CH}^n$ given by*

$$\Phi_\rho(e^{it},p,q) = \big[(s_\rho\,\exp(it\,c_\rho^2)\,\psi_1(p)\,,\ c_\rho\,\exp(it\,s_\rho^2)\,\psi_2(q))\big]$$

is a H-minimal Lagrangian immersion.

A particular case of Corollary 6.6 gives a one-parameter family of H-minimal Lagrangian embeddings.

Corollary 6.7. *Let $\rho > 0$ and $n = n_1+n_2+1$. The immersion Φ_ρ of Corollary 6.6, where ψ_1 (resp. ψ_2) is the totally geodesic Legendrian embedding of $\mathbb{S}^{n_1}$ into $\mathbb{S}^{2n_1+1}$ (resp. of $\mathbb{RH}^{n_2}$ into $\mathbb{H}_1^{2n_2+1}$), provides a H-minimal Lagrangian embedding*

$$\overline{(e^{it},x,y)} \mapsto \big[(s_\rho\,\exp(it\,c_\rho^2)\,x\,,\ c_\rho\,\exp(it\,s_\rho^2)\,y)\big]$$

of the quotient of $\mathbb{S}^1 \times \mathbb{S}^{n_1} \times \mathbb{RH}^{n_2}$ by the action of the group $\mathbb{Z}_2$ into $\mathbb{CH}^n$, the action of $\mathbb{Z}_2$ being generated by the involution $(e^{is},x,y) \mapsto (-e^{is},-x,y)$.

We finally turn our attention to (24) with $\mu = 0$. We observe that this is exactly equation (3) in [Castro and Urbano 2004, Lemma 2] (with $p = n_1$ and $q = n_2$). If we choose the initial conditions $\gamma(0) = (\sinh \varrho, \cosh \varrho)$, $\varrho > 0$, we can make use of the study made in that paper.

Lemma 6.8. *Let $\gamma_\varrho = (\gamma_1, \gamma_2) : I \subset \mathbb{R} \to \mathbb{H}^3_1$ be the unique curve solution of*

$$\gamma_j' \bar{\gamma}_j = i\, \bar{\gamma}_1^{n_1+1} \bar{\gamma}_2^{n_2+1}, \;\; j = 1, 2,$$

satisfying the real initial conditions $\gamma_\varrho(0) = (\sinh \varrho, \cosh \varrho)$, $\varrho > 0$.

(1) $\mathrm{Re}(\gamma_1^{n_1+1} \gamma_2^{n_2+1}) = \sinh^{n_1+1} \varrho \, \cosh^{n_2+1} \varrho$.

(2) *For $j = 1, 2$ and any $t \in I$, we have $\bar{\gamma}_j(t) = \gamma_j(-t)$.*

(3) *The curves γ_1 and γ_2 are embedded and can be parametrized by $\gamma_j(t) = \rho_j(t) e^{i\theta_j(t)}$, where we have set (with $c_\varrho = \cosh \varrho$, $s_\varrho = \sinh \varrho$)*

$$\rho_1(t) = \sqrt{t^2 + s_\varrho^2},$$

$$\theta_1(t) = \int_0^t \frac{s_\varrho^{n_1+1} c_\varrho^{n_2+1} x \, dx}{(x^2 + s_\varrho^2)\sqrt{(x^2 + s_\varrho^2)^{n_1+1}(x^2 + c_\varrho^2)^{n_2+1} - s_\varrho^{2(n_1+1)} c_\varrho^{2(n_2+1)}}},$$

$$\rho_2(t) = \sqrt{t^2 + c_\varrho^2},$$

$$\theta_2(t) = \int_0^t \frac{s_\varrho^{n_1+1} c_\varrho^{n_2+1} x \, dx}{(x^2 + c_\varrho^2)\sqrt{(x^2 + s_\varrho^2)^{n_1+1}(x^2 + c_\varrho^2)^{n_2+1} - s_\varrho^{2(n_1+1)} c_\varrho^{2(n_2+1)}}}.$$

In this way, the immersions ϕ_ϱ constructed with the curves γ_ϱ of Lemma 6.8 induce a one-parameter family of H-minimal Lagrangian immersions

$$\Phi_\varrho : \mathbb{R} \times N_1 \times N_2 \to \mathbb{CH}^n, \;\; n = n_1 + n_2 + 1, \;\; \varrho > 0.$$

In particular, Φ_ϱ is minimal if and only if ψ_1 and ψ_2 are minimal. We conclude with the following particular case, which leads to a one-parameter family of minimal Lagrangian embeddings.

Corollary 6.9. *Let $\varrho > 0$ and set $c_\varrho = \cosh \varrho$, $s_\varrho = \sinh \varrho$. Then*

$$\mathbb{R} \times \mathbb{S}^{n_1} \times \mathbb{RH}^{n_2} \to \mathbb{CH}^n, \;\; n = n_1 + n_2 + 1,$$

$$(s, x, y) \mapsto [(\sqrt{s^2 + s_\varrho^2}\, \exp(i\, \theta_1(s)) x \, , \, \sqrt{s^2 + c_\varrho^2}\, \exp(i\, \theta_2(s)) y)],$$

where the $\theta_i(s)$ are given in Lemma 6.8(3), is a minimal Lagrangian embedding.

Acknowledgements. The authors thank the referee for helpful comments and suggestions.

References

[Anciaux 2003] H. Anciaux, "Construction of many Hamiltonian stationary Lagrangian surfaces in Euclidean four-space", *Calc. Var. Partial Diff. Equations* **17**:2 (2003), 105–120. MR 2004i:53118 Zbl 1042.53004

[Anciaux et al. 2006] H. Anciaux, I. Castro, and P. Romon, "Lagrangian submanifolds foliated by $(n-1)$-spheres in $\mathbb{R}^{2n}$", *Acta Math. Sinica* **22**:4 (2006), 1197–1214.

[Castro and Urbano 1998] I. Castro and F. Urbano, "Examples of unstable Hamiltonian-minimal Lagrangian tori in $\mathbb{C}^2$", *Compositio Math.* **111**:1 (1998), 1–14. MR 98m:53075 Zbl 0896.53039

[Castro and Urbano 2004] I. Castro and F. Urbano, "On a new construction of special Lagrangian immersions in complex Euclidean space", *Q. J. Math.* **55**:3 (2004), 253–265. MR 2005g:53087 Zbl 02138381

[Castro et al. 2001] I. Castro, C. R. Montealegre, and F. Urbano, "Closed conformal vector fields and Lagrangian submanifolds in complex space forms", *Pacific J. Math.* **199**:2 (2001), 269–302. MR 2002g:53137 Zbl 1057.53063

[Castro et al. 2002] I. Castro, C. R. Montealegre, and F. Urbano, "Minimal Lagrangian submanifolds in the complex hyperbolic space", *Illinois J. Math.* **46**:3 (2002), 695–721. MR 2004e:53089 Zbl 1032.53052

[Eells and Ratto 1993] J. Eells and A. Ratto, *Harmonic maps and minimal immersions with symmetries*, Annals of Mathematics Studies **130**, Princeton University Press, Princeton, NJ, 1993. MR 94k:58033 Zbl 0783.58003

[Haskins 2004a] M. Haskins, "The geometric complexity of special Lagrangian T^2-cones", *Invent. Math.* **157**:1 (2004), 11–70. MR 2005m:53085 Zbl 1064.53032

[Haskins 2004b] M. Haskins, "Special Lagrangian cones", *Amer. J. Math.* **126**:4 (2004), 845–871. MR 2005e:53074 Zbl 1074.53067

[Hélein and Romon 2000] F. Hélein and P. Romon, "Weierstrass representation of Lagrangian surfaces in four-dimensional space using spinors and quaternions", *Comment. Math. Helv.* **75**:4 (2000), 668–680. MR 2001m:53102 Zbl 0973.53065

[Hélein and Romon 2002a] F. Hélein and P. Romon, "Hamiltonian stationary Lagrangian surfaces in $\mathbb{C}^2$", *Comm. Anal. Geom.* **10**:1 (2002), 79–126. MR 2003e:53108 Zbl 1007.53060

[Hélein and Romon 2002b] F. Hélein and P. Romon, "Hamiltonian stationary Lagrangian surfaces in Hermitian symmetric spaces", pp. 161–178 in *Differential geometry and integrable systems* (Tokyo, 2000), edited by M. Guest et al., Contemp. Math. **308**, Amer. Math. Soc., Providence, RI, 2002. MR 2004j:53103 Zbl 1036.53058

[Hélein and Romon 2003] F. Hélein and P. Romon, "Hamiltonian stationary tori in complex projective plane", preprint, 2003. math.DG/0310095

[McDuff and Salamon 1998] D. McDuff and D. Salamon, *Introduction to symplectic topology*, Oxford University Press, New York, 1998. MR 2000g:53098 Zbl 1066.53137

[Mironov 2004] A. E. Mironov, "New examples of Hamilton-minimal and minimal Lagrangian submanifolds in $\mathbb{C}^n$ and $\mathbb{C}\mathrm{P}^n$", *Mat. Sb.* **195**:1 (2004), 89–102. In Russian; translated in *Sb. Math.* **195**:1 (2004), 85–96. MR 2005e:53127 Zbl 1078.53079

[Montealegre and Vrancken 2006] C. R. Montealegre and L. Vrancken, "Pinching properties of Lagrangian submanifolds of complex projective space", preprint, 2006.

[Naitoh 1981] H. Naitoh, "Isotropic submanifolds with parallel second fundamental form in $P^m(c)$", *Osaka J. Math.* **18**:2 (1981), 427–464. MR 83b:53051 Zbl 0471.53036

[Oh 1990] Y.-G. Oh, "Second variation and stabilities of minimal Lagrangian submanifolds in Kähler manifolds", *Invent. Math.* **101**:2 (1990), 501–519. MR 91f:58022 Zbl 0721.53060

[Schoen and Wolfson 1999] R. Schoen and J. Wolfson, "Minimizing volume among Lagrangian submanifolds", pp. 181–199 in *Differential equations: La Pietra 1996* (Florence, 1996), edited by M. Giaquinta et al., Proc. Sympos. Pure Math. **65**, Amer. Math. Soc., Providence, RI, 1999. MR 99k:53130 Zbl 1031.53112

Received December 1, 2004. Revised June 29, 2005.

ILDEFONSO CASTRO
DEPARTAMENTO DE MATEMÁTICAS
ESCUELA POLITÉCNICA SUPERIOR
UNIVERSIDAD DE JAÉN
23071 JAÉN
SPAIN
icastro@ujaen.es

HAIZHONG LI
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
100084 BEIJING
P. R. CHINA
hli@math.tsinghua.edu.cn

FRANCISCO URBANO
DEPARTAMENTO DE GEOMETRÍA Y TOPOLOGÍA
UNIVERSIDAD DE GRANADA
18071 GRANADA
SPAIN
furbano@ugr.es

PACIFIC JOURNAL OF MATHEMATICS
Vol. 227, No. 1, 2006

A SPECTRAL SEQUENCE DETERMINING THE HOMOLOGY OF Out(F_n) IN TERMS OF ITS MAPPING CLASS SUBGROUPS

MATTHEW HORAK

We construct a covering of the spine of the Culler–Vogtmann outer space Out(F_n) by complexes of ribbon graphs. By considering the equivariant homology for the action of Out(F_n) on this covering, we construct a spectral sequence converging to the homology of Out(F_n) that has its E^1 terms given by the homology of mapping class groups and their subgroups. This spectral sequence can be seen as encoding all of the information of how the homology of Out(F_n) is related to the homology of mapping class groups and their subgroups

1. Introduction

Much is known about the cohomology of mapping class groups of surfaces. (*All surfaces considered in this work are assumed orientable.*) Let Σ be a surface with boundary, and let $P\Gamma(\Sigma)$ be the group of isotopy classes, relative to the boundary, of homeomorphisms of Σ that fix the boundary pointwise. We call $P\Gamma(\Sigma)$ the pure mapping class group of Σ. Harer [1985] proved that the k-th integral homology group of $P\Gamma(\Sigma)$ is independent of the genus and number of boundary components of Σ if the genus of Σ is at least $3k$. Later, Ivanov [1989] and Harer [1993] improved these bounds, and Harer was able to find the exact location at which the rational homology stabilizes. He also computed in [Harer 1986] the virtual cohomological dimension (VCD) of $P\Gamma(\Sigma)$ and showed that this group has no rational homology at its VCD. Madsen and Weiss [2002] have determined the entire stable integral cohomology algebra of pure mapping class groups. In particular, their result verifies the conjecture of Mumford that the stable rational cohomology algebra is a polynomial algebra with a single generator in each even dimension.

For outer automorphism groups of free groups, much less is known. Culler and Vogtmann [1986] have compute the VCD of Out(F_n) by considering the action of this group on a contractible simplicial complex known as the spine of outer space. Recently, Hatcher and Vogtmann [2004] have shown that the k-th integral homology of Out(F_n) is independent of n if $n \geq 2k+5$, but the exact stability

MSC2000: primary 20F65; secondary 57M07.
Keywords: free group automorphisms, mapping class groups, ribbon graphs.

range remains unknown. Indeed, there are no nontrivial stable rational homology or cohomology classes known for $\mathrm{Out}(F_n)$. For a good survey of current knowledge about $\mathrm{Out}(F_n)$ and $\mathrm{Aut}(F_n)$, see [Vogtmann 2002].

Since the mapping class groups of surfaces appear as subgroups of $\mathrm{Out}(F_n)$, it is natural to try to understand the homology of $\mathrm{Out}(F_n)$ in terms of the homology of mapping class groups. This paper represents an attempt to clarify this relationship. For a punctured surface, the mapping class group is simply the group of isotopy classes of orientation preserving homeomorphisms of the surface. The group of all isotopy classes of homeomorphisms of a punctured surface will be called the extended mapping class group of that surface, so that extended mapping class groups contain orientation reversing homeomorphisms.

We construct a first quadrant spectral sequence that converges to $H_*(\mathrm{Out}(F_n))$, many of whose terms consist of the homology of mapping class groups. The spectral sequence arises from a covering of the spine of outer space by a collection of subcomplexes called ribbon graph subcomplexes. We prove that the nerve of this covering is contractible. The spectral sequence mentioned above is the equivariant homology spectral sequence of the action of $\mathrm{Out}(F_n)$ on this nerve.

All of the terms on the E^1 page of this spectral sequence are given by the homology simplex stabilizers. For a 0-simplex, the stabilizer is simply the extended mapping class group of a punctured surface Σ, or equivalently the stabilizer of the set conjugacy classes in $F_n \cong \pi_1 \Sigma$ that correspond to positively and negatively oriented curves about the punctures of Σ. For higher-dimensional simplices, stabilizers are given by the generalized stabilizers $\mathcal{A}_{U,G}$ of m-tuples of conjugacy classes, which are studied in [McCool 1975]. (These groups are finite-index subgroups of the ordinary stabilizers of certain sets of conjugacy classes in F_n.) We prove:

Theorem. *For any* $\mathrm{Out}(F_n)$*-module* M *there is a spectral sequence of the form*

$$E^1_{pq} = \bigoplus_{\sigma \in \Delta_p} H_q(G_\sigma; M_\sigma) \Rightarrow H_{p+q}(\mathrm{Out}(F_n); M),$$

where Δ_0 *is the set of homeomorphism classes of punctured orientable surfaces with fundamental group* F_n *and where for a vertex* $v \in \Delta_0$ *corresponding to surface* Σ, *the stabilizer* G_v *is the extended mapping class group* $\mathrm{MCG}^{\pm}(\Sigma)$. *Moreover, for* $p > 0$, *each* G_σ *is a generalized stabilizer of the form* $\mathcal{A}_{U_\sigma, H_\sigma}$.

The rest of this paper is organized as follows. In Sections 2 and 3 we review the definitions of outer space, the spine of outer space, ribbon graphs and some related objects. In Section 4, we construct a covering of the spine of outer space by subcomplexes of ribbon graphs. Section 5 is devoted to the proof of the fact that the nerve of this covering is contractible. In Section 6 we determine simplex stabilizers for the action of $\mathrm{Out}(F_n)$ on the nerve. The analysis of the equivariant homology spectral sequence for this action appears in final two sections where we prove the above theorem and use Harer's stability theorems to find rough upper bounds on the dimensions of some portions of the E^∞ page of the spectral sequence. These bounds limit the possible contribution that the mapping class subgroups of $\mathrm{Out}(F_n)$ can make to the homology of $\mathrm{Out}(F_n)$.

2. Outer space

For convenience and to set notation, we briefly review the construction in [Culler and Vogtmann 1986] of outer space and its spine. A *graph* is a connected, one-dimensional CW-complex. We will consider only finite graphs with all vertices having valence at least 3. A *subforest* of a graph Γ is a subgraph of Γ that contains no circuits; a forest is a disjoint union of trees.

Fix an integer $n \geq 2$. Denote by R_0 the standard n-petal rose; R_0 has one vertex and n edges. Fix an identification $\pi_1(R_0) = F_n$. A *marking* on a graph is a homotopy equivalence, $g : R_0 \to \Gamma$. We define an equivalence relation on the set of markings by setting $(\Gamma_1, g_1) \sim (\Gamma_2, g_2)$ if

there is a graph isomorphism $h : \Gamma_1 \to \Gamma_2$ such that $g_2 \simeq h \circ g_1$, that is, such that the diagram

(1)
$$\begin{array}{ccc} & & \Gamma_1 \\ & \overset{g_1}{\nearrow} & \\ R_0 & & \downarrow h \\ & \underset{g_2}{\searrow} & \\ & & \Gamma_2 \end{array}$$

commutes up to free homotopy. An equivalence class of markings is called a *marked graph* and can be denoted by (Γ, g). The marking g identifies $\pi_1(\Gamma)$ with F_n up to composition with an inner automorphism.

The marked graph (Γ, g) is usually represented by a labeled graph as follows. Fix an identification of $\pi_1(R_0)$ with F_n. Choose a spanning tree T in Γ and a homotopy inverse to g that collapses T to the vertex of R_0 and maps each edge of $\Gamma - T$ to a reduced edge path in R_0. A directed edge $\vec{e}$ in the complement of T corresponds, via this homotopy equivalence, to an of element in F_n. Label e with a direction and the corresponding element of F_n. Note that the same marked graph can be represented by many different labeled graphs, depending on the choice of T and the particular representative of (Γ, g). For a marked rose, the spanning tree must consist of the single vertex, so we get a label for each directed edge. The set of labels on a marked rose is a basis of F_n, which is determined up to conjugacy. Two labeled roses correspond to equivalent marked roses if and only if their edges are labeled by conjugate bases of F_n.

If Φ is a forest in the marked graph (Γ, g), then collapsing each component of Φ to a point produces another marked graph, denoted by $(\Gamma/\Phi, q \circ g)$, where q is the quotient map collapsing each component of Φ to a point. Passing from (Γ, g) to $(\Gamma/\Phi, q \circ g)$ is called a *forest collapse*. There is a partial order on the set of marked graphs with fundamental group F_n defined by $(\Gamma_1, g_1) \leq (\Gamma_2, g_2)$ if there is a forest collapse taking (Γ_2, g_2) to (Γ_1, g_1). The geometric realization of the poset of marked graphs is the spine of outer space and is denoted by K_n.

The group $\mathrm{Out}(F_n)$ acts on K_n by changing the markings of the underlying graphs. Explicitly, for $\psi \in \mathrm{Out}(F_n)$,

$$(\Gamma, g) \cdot \psi := (\Gamma, g \circ |\psi|), \tag{2}$$

where $|\psi| : R_0 \to R_0$ is a homotopy equivalence inducing an automorphism of $F_n = \pi_1(R_0)$ that represents the outer automorphism class ψ. Culler and Vogtmann observe that this action is cocompact and that vertex stabilizers are finite.

Culler and Vogtmann also define a larger space, called *outer space*, consisting of metric marked graphs. This space has the disadvantages of not being a simplicial complex and the $\mathrm{Out}(F_n)$ action not being cocompact. The complex K_n can be constructed as a simplicial spine onto which of outer space deformation retracts.

3. Ribbon graphs

There are similar constructions for mapping class groups that use marked ribbon graphs rather than ordinary marked graphs. A *ribbon graph* is a graph Γ together with, at each vertex v, a cyclic ordering of the set $h(v)$ of half-edges incident to v. The collection of cyclic orderings at the vertices is called a *ribbon structure* for Γ, and is denoted by $\mathbb{O}$. The term "ribbon graph" is used because one can construct a bounded surface from a ribbon graph $(\Gamma, \mathbb{O})$ by fattening its edges to ribbons. We give a formal construction of this surface after Definition 3.1, but informally, the surface is constructed from $(\Gamma, \mathbb{O})$ by replacing each edge by a ribbon and gluing the ribbons together at their ends according to the cyclic order of the corresponding half-edges. The gluing is done in such a way as to produce an oriented surface. Figure 1 shows this process for two different ribbon structures on a rose with 2 edges. In these figures, ribbon structures are specified by the given embeddings of a neighborhood of the vertices into the plane. The ribbon graph $(\Gamma, \mathbb{O}_1)$ produces a pair of pants while $(\Gamma, \mathbb{O}_2)$ produces a torus with one boundary component.

The boundary curves of the surface produced from $(\Gamma, \mathbb{O})$ correspond to reduced edge paths in Γ that follow the cyclic ordering at the vertices in the sense of the following definition. Following [Mulase and Penkava 1998], we view a directed edge $\vec{e}$ as an ordering (e^+, e^-) of the half-edges e^+ and e^- comprising e.

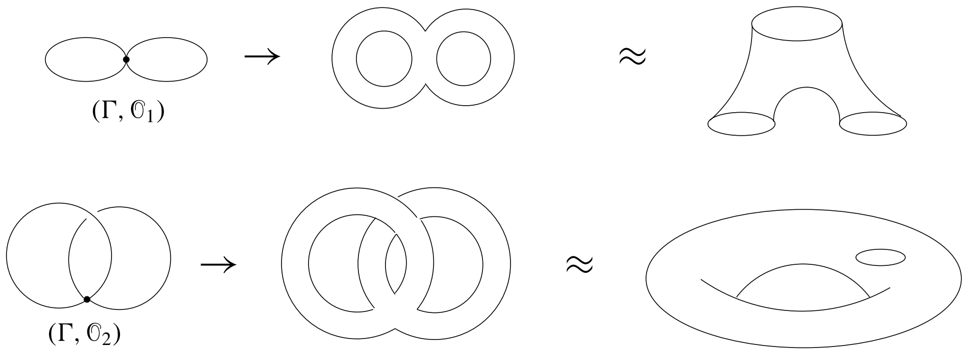


Figure 1. Fattenings of ribbon graphs.

Definition 3.1. A *boundary cycle* in the ribbon graph $(\Gamma, \mathbb{O})$ is a directed reduced edge cycle,

$$(\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_{l-1}, \vec{e}_l = \vec{e}_1)$$

such that for each i the half-edges e_i^+ and e_{i+1}^- are incident to the same vertex, and in the cyclic ordering at that vertex, e_{i+1}^- directly follows e_i^+.

For our purposes, it will be more convenient to work with punctured surfaces, so we now give a precise construction of a punctured surface $|\Gamma, \mathbb{O}|$ from a ribbon graph $(\Gamma, \mathbb{O})$. First note that each edge of Γ is traversed exactly once in each direction by the set of boundary cycles of $(\Gamma, \mathbb{O})$. Construct a space $|\Gamma, \mathbb{O}|$ by gluing a once punctured disk to Γ along each boundary cycle γ of $(\Gamma, \mathbb{O})$. By verifying that a small neighborhood of each vertex in $|\Gamma, \mathbb{O}|$ is indeed a disk, one can verify that $|\Gamma, \mathbb{O}|$ is a surface that deformation retracts onto Γ. One can also verify that $|\Gamma, \mathbb{O}|$ is orientable and we orient it such that a small positively oriented simple closed curve around a vertex v of Γ intersects the half-edges in $h(v)$ in the cyclic order determined by $\mathbb{O}$.

If Γ is marked by the homotopy equivalence $g : R_0 \to \Gamma$, then the composition of g with the inclusion $i : \Gamma \hookrightarrow |(\Gamma, \mathbb{O})|$ is a homotopy equivalence that identifies $\pi_1(\Sigma)$ with F_n up to inner automorphism, just as in the case of marked graphs. This gives the notion of a homotopy marked surface.

Definition 3.2. A *homotopy marked surface* is an equivalence class of pairs (Σ, s), where Σ is a punctured, orientable surface with $\pi_1(\Sigma) \cong F_n$ and $s : R_0 \to \Sigma$ is a homotopy equivalence. The equivalence relation on pairs is given by $(\Sigma_1, s_1) \sim (\Sigma_2, s_2)$ if there is an orientation preserving homeomorphism $h : \Sigma_1 \to \Sigma_2$ with $h \circ s_1 \simeq s_2$.

Recall that we have fixed an integer $n \geq 2$. Often we drop the word "homotopy" and simply use "marked surface" for a homotopy marked surface. Unless otherwise stated, marked surfaces will always be punctured surfaces without boundary and with fundamental group F_n. In Section 4, the equivalence relation defined by homeomorphisms that do not necessarily preserve orientation will be useful. We will denote this equivalence relation by $\sim_\pm$, and use brackets to denote its equivalence classes: $[\Sigma, s]$. We say that the marked graph (Γ, g) can be *drawn in* the marked surface (Σ, s) is there is a ribbon structure $\mathbb{O}$ on Γ such that $|(\Gamma, g, \mathbb{O})| \sim (\Sigma, s)$. In this case, there is an embedding $i : \Gamma \hookrightarrow \Sigma$ such that $s \simeq i \circ g$.

Definition 3.3. The *ribbon graph complex* for the marked surface (Σ, s) is the subcomplex of K_n spanned by graphs that can be drawn in (Σ, s). This complex is denoted by $\mathfrak{R}_{(\Sigma,s)}$.

We will often identify a marked graph or ribbon graph with the corresponding vertex of K_n or $\mathfrak{R}_{(\Sigma,s)}$. Thus for example, if ρ is a marked rose in $\mathfrak{R}_{(\Sigma,s)}$ then $\operatorname{lnk}_{\mathfrak{R}_{(\Sigma,s)}}(\rho)$ will be the link in $\mathfrak{R}_{(\Sigma,s)}$ of the vertex corresponding to ρ.

The ribbon graph complex $\mathfrak{R}_{(\Sigma,s)}$ and related complexes have been important tools in the study of mapping class groups surfaces. In particular, $\mathfrak{R}_{(\Sigma,s)}$ is a subcomplex of the first barycentric subdivision of the arc complex that Harer uses compute the VCD of the pure mapping class group of a surface with boundary [Harer 1986]. Also, for a punctured surface Σ, Bowditch and Epstein [1988] and Penner [1987] use arc systems on Σ to give an open cell decomposition of a space they call the decorated Teichmüller space of Σ. By taking the dual graph of an arc system in Σ, this decomposition may be interpreted in terms of metric ribbon graphs. In the same way that K_n is a simplicial spine of outer space, $\mathfrak{R}_\Sigma$ is a simplicial spine of the decorated Teichmüller space of Σ.

4. The ribbon cover of K_n

The ribbon subcomplex of K_n associated to a marked surface does not depend on the surface's orientation. This is because if the ribbon structure $\mathbb{O}$ draws (Γ, g) in (Σ, s), then $\mathbb{O}^{op}$ draws (Γ, g) in $(\Sigma, s)^{op}$, where $\mathbb{O}^{op}$ is the ribbon structure obtained by reversing all cyclic ordering of $\mathbb{O}$ and $(\Sigma, s)^{op}$ the marked surface obtained by reversing the orientation of (Σ, s). Therefore there is a well-defined subcomplex $\mathfrak{R}_{[\Sigma,s]}$ of K_n.

Proposition 4.1. *K_n is covered by its ribbon graph subcomplexes.*

Proof. Recall that K_n is the geometric realization of the poset of marked graphs with fundamental group F_n, so the vertices of K_n are partially ordered. For a vertex v of K_n, let $\operatorname{st}(v)$ be the star of v and let $\operatorname{st}^+(v)$ be the subcomplex of $\operatorname{st}(v)$ spanned by v together with vertices of $\operatorname{st}(v)$ that are greater than v in the partial order. Similarly let $\operatorname{st}^-(v)$ be the subcomplex of $\operatorname{st}(v)$ spanned by v and vertices less than v. Thus if v corresponds to the marked graph (Γ, g), then $\operatorname{st}^+(v)$ consists of the vertices of K_n corresponding to graphs that may be collapsed to (Γ, g) and $\mathrm{st}^-(v)$ consists of vertices corresponding to graphs to which (Γ, g) collapses.

Suppose that $v \in \mathfrak{R}_{[\Sigma,s]}$ corresponds to the marked graph (Γ, g). Then Γ has a ribbon structure $\mathbb{O}$ that draws (Γ, g) in (Σ, s). If e is any edge in Γ that is not a loop, the marked graph $(\Gamma/e, q \circ g)$ inherits a ribbon structure $\mathbb{O}/e$ from $\mathbb{O}$ that draws $(\Gamma/e, q \circ g)$ in (Σ, s). Therefore $\mathrm{st}^-(v) \subset \mathfrak{R}_{[\Sigma,s]}$.

To see that every simplex of K_n belongs to some ribbon graph subcomplex, let σ be a simplex of K_n. If w is the vertex of σ that is the greatest in the partial ordering of the vertices, then σ is contained in the complex $\mathrm{st}^-(w)$. Suppose that w corresponds to the marked graph (Γ_0, g_0) Choose any ribbon structure $\mathbb{O}_0$ on Γ_0 and set $(\Sigma, s) := |(\Gamma_0, g_0, \mathbb{O}_0)|$. Then $w \in \mathfrak{R}_{[\Sigma,s]}$ so that $\mathrm{st}^-(w)$ is contained in $\mathfrak{R}_{[\Sigma,s]}$. Since σ has w as its greatest vertex, σ is a simplex of $\mathrm{st}^-(w) \subseteq \mathfrak{R}_{[\Sigma,s]}$. $\square$

We begin our study of the nerve of this cover with definitions and lemmas.

Definition 4.2. Suppose that the homotopy-marked, oriented surface (Σ, s) has k punctures $p_1, \dots, p_k$. Let γ_j be a simple closed curve in Σ that disconnects Σ by cutting off a disk punctured at p_j. By virtue of the marking and orientation of Σ, the curve γ_j corresponds to a conjugacy class in F_n. The set of such conjugacy classes is called the set of *boundary classes* of Σ and is denoted by $W_{(\Sigma,s)}$ or simply W_Σ. Similarly, the set of conjugacy classes in F_n represented by the boundary cycles of the marked ribbon graph $(\Gamma, g, \mathbb{O})$ is called the set of *boundary classes* of $(\Gamma, g, \mathbb{O})$.

The boundary classes carry a lot of information about the surface. For example, if $(\Sigma, s) = |(\Gamma, g, \mathbb{O})|$ then the boundary classes of (Σ, s) and the boundary classes of $(\Gamma, g, \mathbb{O})$ are the same. Another important observation about the boundary classes is that if two (necessarily homeomorphic) marked surfaces have the same boundary classes, they are equivalent marked surfaces. This is proved by using a theorem of Zieschang [1980, Theorem 5.15.3] that states that an element of $\mathrm{Out}(\pi_1\Sigma)$ is induced by a mapping class of Σ if and only if it stabilizes the boundary classes of Σ. The boundary classes are also used to prove the following:

Lemma 4.3. *If the marked graph (Γ, g) can be drawn in (Σ, s), then (Γ, g) has exactly one ribbon structure giving (Σ, s).*

Proof. Since Γ can be drawn in (Σ, s), there is a ribbon structure $\mathbb{O}$ on (Γ, g) with $|(\Gamma, g, \mathbb{O})| = (\Sigma, s)$. Suppose that $\mathbb{O}'$ is a different ribbon structure on Γ. We may choose a vertex v and half-edges e^+, e_1^- and e_2^- of Γ with $e_1^- \neq e_2^-$, with e_1^- following e^+ in the cyclic ordering $\mathbb{O}$ but with e_2^- following e^+ in the cyclic ordering $\mathbb{O}'$. This means that the sequence $\vec{e}\vec{e}_1$ appears in the boundary cycles of $(\Gamma, \mathbb{O})$ while the sequence $\vec{e}\vec{e}_2$ appears in the boundary classes of $(\Gamma, \mathbb{O}')$. Since each directed edge of Γ appears exactly once in the set of boundary cycles for any given ribbon structure, the set of boundary cycles of $(\Gamma, g, \mathbb{O}')$ must differ from those of $(\Gamma, g, \mathbb{O})$. Therefore the set of boundary classes of $(\Gamma, g, \mathbb{O}')$ differ from those of $(\Gamma, g, \mathbb{O})$. Hence $|(\Gamma, g, \mathbb{O}')|$ is different from (Σ, s), because equivalent marked surfaces have the same boundary classes. □

Each of the two orientations of Σ gives a unique ribbon structure to (Γ, g). These two ribbon structures are opposite of each other. We now prove that (Σ, s) and (Σ_1, s_1) give the same ribbon graph subcomplexes of K_n if and only if $[\Sigma, s] = [\Sigma_1, s_1]$. The main step is this:

Lemma 4.4. *Let (Σ, s) be a marked surface and let $\mathfrak{R} = \mathfrak{R}_{[\Sigma,s]}$ be the corresponding ribbon graph subcomplex of K_n. The ribbon structure given by (Σ, s) to a marked rose ρ in $\mathfrak{R}$ can be reconstructed, up to reversal of the cyclic order, by the (nonribbon) marked graphs in $\mathrm{lnk}_{\mathfrak{R}}(\rho)$.*

Proof. Choose a direction for each edge in ρ. The marking of ρ determines a labeling of the directed edges by a basis $X = \{a_1, a_2, \ldots, a_n\}$ of F_n. The basis X and this labeling are determined up to composition with an inner automorphism of F_n.

Consider the marked graphs in $\mathrm{lnk}_{K_n}(\rho)$ with exactly two vertices, one of which is trivalent. These graphs are constructed from ρ as follows. Let e^+ and f^+ be any two half-edges of ρ. Construct a new marked graph $\rho(e^+, f^+)$ by deleting the vertex of ρ and replacing it with two new vertices v_0 and v_1 joined by a new edge $\tilde{e}$. Attach the half-edges e^+ and f^+ to v_0 and attach the rest of the half-edges of ρ to v_1. Mark $\rho(e^+, f^+)$ so that collapsing $\tilde{e}$ to a point gives the original marking on ρ.

By Lemma 4.3, ρ has exactly one ribbon structure giving (Σ, s). This means that if we allow for orientation reversing homeomorphisms of the surface, ρ has two ribbon structures giving $[\Sigma, s]$, and these ribbon structures are opposite of each other. A marked graph of the form $\rho(e^+, f^+)$ lies in $\mathrm{lnk}_{\mathfrak{R}}(\rho)$ if and only if e^+ and f^+ are adjacent in these ribbon structures. Thus, given a half-edge a_i^+ in ρ, exactly two graphs of the form $\rho(a_i^+, a_j^{\epsilon_j})$ and $\rho(a_i^+, a_k^{\epsilon_k})$ will lie in $\mathrm{lnk}_{\mathfrak{R}}(\rho)$ and they will be the graphs for which a_i^+ is adjacent to the half-edges $a_j^{\epsilon_j}$ and $a_k^{\epsilon_k}$ in the ribbon structure on ρ. Therefore, for each i, the nonribbon graphs in $\mathrm{lnk}_{\mathfrak{R}}(\rho)$ determine the half-edges adjacent to a_i^+ and a_i^- in the ribbon structure on ρ. There are only two cyclic orderings of the half-edges that satisfy these adjacency data, and they are opposites of each other. Example 4.5 works this out for a ribbon rose with $n = 3$. □

Example 4.5. Figure 2 shows the graphs of the form $\rho(e^+, f^+)$ in the link of a marked ribbon rose in $\mathfrak{R}_\Sigma$. The fact that graphs (1) and (6) have the half-edges a^+ and c^+, respectively, adjacent to the edge a^- implies that the half-edges in

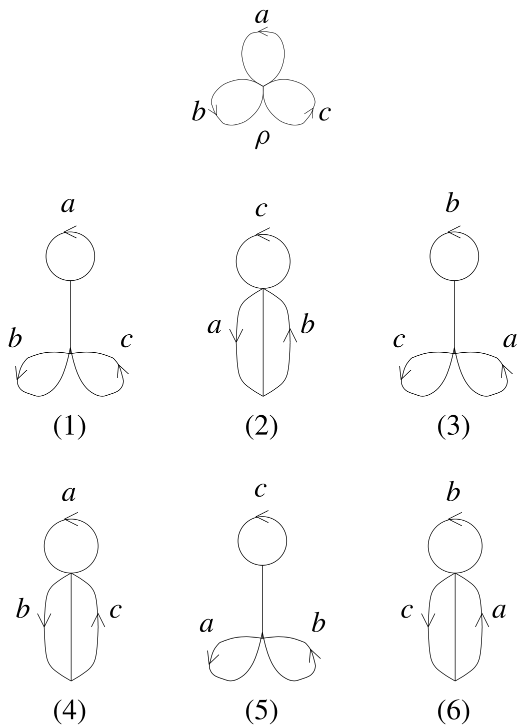


Figure 2. Some graphs in $\mathrm{lnk}^+_{\mathfrak{R}_\Sigma}(\rho)$.

ρ adjacent to a^- are a^+ and c^+. Similarly, the half-edges adjacent to any other half-edge can be determined by some pair of the graphs in Figure 2.

Proposition 4.6. *$\mathfrak{R}_{[\Sigma_1,s_1]} = \mathfrak{R}_{[\Sigma_2,s_2]}$ if and only if $[\Sigma_1, s_1] = [\Sigma_2, s_2]$.*

Proof. First suppose that (Σ_1, s_1) and (Σ_2, s_2) are equivalent via the (possibly orientation-reversing) homeomorphism $h : \Sigma_1 \to \Sigma_2$. Then h can be used to draw in (Σ_2, s_2) any graph that can be drawn in (Σ_1, s_1) and h^{-1} can be used to draw in (Σ_1, s_1) any graph that can be drawn in (Σ_2, s_2), so $\mathfrak{R}_{[\Sigma_1,s_1]} = \mathfrak{R}_{[\Sigma_2,s_2]}$.

Now suppose that $\mathfrak{R}_{[\Sigma_1,s_1]} = \mathfrak{R}_{[\Sigma_2,s_2]}$. Set $\mathfrak{R} := \mathfrak{R}_{[\Sigma_1,s_1]} = \mathfrak{R}_{[\Sigma_2,s_2]}$. Fix a marked rose $\rho \in \mathfrak{R}$; it inherits a ribbon structure from $\mathfrak{R}_{[\Sigma_1,s_1]}$ giving (Σ_1, s_1), and a ribbon structure from $\mathfrak{R}_{[\Sigma_2,s_2]}$ giving (Σ_2, s_2). By Lemma 4.4, these structures are determined up to reversal by the nonribbon graphs in $\mathrm{lnk}_{\mathfrak{R}_{[\Sigma_1,s_1]}}(\rho)$ and $\mathrm{lnk}_{\mathfrak{R}_{[\Sigma_2,s_2]}}(\rho)$, respectively. But $\mathfrak{R}_{[\Sigma_1,s_1]} = \mathfrak{R}_{[\Sigma_2,s_2]}$,so $\mathrm{lnk}_{\mathfrak{R}_{[\Sigma_1,s_1]}}(\rho) = \mathrm{lnk}_{\mathfrak{R}_{[\Sigma_2,s_2]}}(\rho)$. Therefore the ribbon structures must coincide or be opposites of each other. In the first case, $(\Sigma_1, s_1) \sim (\Sigma_2, s_2)$, and in the second, $(\Sigma_1, s_1) \sim (\Sigma_2, s_2)^{op}$. Thus $[\Sigma_1, s_1] = [\Sigma_2, s_2]$. □

This proposition gives a convenient description of the covering of K_n by its ribbon graph subcomplexes. The covering is locally finite because each different homotopy marked surface that contains a specific graph endows that graph with a different ribbon structure. A graph has only finitely many different ribbon structures so a given marked graph can be drawn in only finitely many marked surfaces and hence lies in only finitely many different ribbon graph subcomplexes.

Let $\mathcal{N}_n$ denote the nerve of the ribbon cover of K_n. That is, $\mathcal{N}_n$ is the simplicial complex containing a k-simplex $\langle \mathfrak{R}_{[\Sigma_0,s_0]}, \ldots, \mathfrak{R}_{[\Sigma_k,s_k]} \rangle$ for every collection $\{\mathfrak{R}_{[\Sigma_0,s_0]}, \ldots, \mathfrak{R}_{[\Sigma_k,s_k]}\}$ of ribbon graph complexes such that the intersection $\bigcap_{i=0}^{k} \mathfrak{R}_{[\Sigma_i,s_i]}$ is nonempty. By Proposition 4.6, the vertex set of $\mathcal{N}_n$ is the set of unoriented equivalence classes, $[\Sigma, s]$.

The action of $\mathrm{Out}(F_n)$ on K_n permutes the ribbon graph subcomplexes because if (Γ, g) can be drawn in (Σ, s), then $(\Gamma, g) \cdot \psi = (\Gamma, g \circ |\psi|)$ can be drawn in $(\Sigma, s \circ |\psi|)$. Therefore $\mathrm{Out}(F_n)$ maps intersections of ribbon graph subcomplexes to intersections of ribbon graph subcomplexes, so it acts on $\mathcal{N}_n$. The equivariant homology of this action provides the spectral sequence, which we will study, that relates the homology of $\mathrm{Out}(F_n)$ to that of mapping class groups.

Although it will not be necessary for the development here, we remark briefly on the compactness properties of $\mathcal{N}_n$ and the $\mathrm{Out}(F_n)$ action. In general ($n \geq 3$), all vertices of $\mathcal{N}_n$ have infinite valence:

Proposition 4.7. *For $n \geq 3$, the ribbon complex for any homotopy marked surface intersects the ribbon complexes of infinitely many other homotopy marked surfaces.*

Proof. We first show that the ribbon graph subcomplex of a marked surface $[\Sigma, s]$ with fundamental group of rank at least 3, contains infinitely many different marked roses. Choose a marked rose $(\rho, r) \in \mathfrak{R}_{[\Sigma,s]}$ and an automorphism $\psi \in \mathrm{Out}(F_n)$ representing a Dehn twist about a nonboundary curve in Σ. Since vertex stabilizers in the spine of outer space are finite and ψ has infinite order, there are infinitely many different equivalence classes of marked roses of the form $\psi^n \cdot (\rho, r)$. All of these marked roses lie in $\mathfrak{R}_{[\Sigma,s]}$.

Recall from Section 2 that a marking of a rose is equivalent to a choice of conjugacy class of basis labeling its directed edges. If ρ is a marked rose with edges labeled by the basis $X = \{a_1, \ldots, a_n\}$, then ρ can be drawn in a marked $(n+1)$-times punctures sphere Σ_1 with boundary classes

$$W_{\Sigma_1} = \{a_1, \ldots, a_n, a_n^{-1}, \cdots, a_1^{-1}\}.$$

By the discussion following Definition 4.2, two marked spheres with different boundary classes cannot be equivalent. Therefore the infinitely many different marked roses in $\mathfrak{R}_{[\Sigma,s]}$ give rise to infinitely many different marked spheres all of whose ribbon complexes intersect $\mathfrak{R}_{[\Sigma,s]}$. $\square$

Proposition 4.8. Out(F_n) *acts cocompactly on* $\mathcal{N}_n$.

Proof. Fix a marked rose ρ. For each p-simplex $\langle \Sigma_0, \ldots, \Sigma_p \rangle$, the subcomplex $\bigcap_{i=0}^{p} \mathfrak{R}_{\Sigma_i}$ contains a rose. This rose may be taken to ρ by an element of Out(F_n), so each orbit of p-simplex has a representative all of whose surfaces contain ρ. Since a marked rose can be drawn in only finitely many different marked surfaces, there are only finitely many orbits of p-simplices. □

5. Contractibility of $\mathcal{N}_n$

To show that $\mathcal{N}_n$ is contractible, we will need the following result from Čech theory.

Lemma 5.1 [Hatcher 2002, Section 4.G]. *Let* $\mathfrak{U}$ *be a cover of the CW-complex* X *by a family of subcomplexes. If every nonempty intersection of finitely many complexes in* $\mathfrak{U}$ *is contractible, then the nerve of the cover is homotopy equivalent to* X.

We will apply Lemma 5.1 to the covering of K_n by ribbon graph complexes. Thus the remainder of this section is devoted to the proof of,

Proposition 5.2. *For any finite collection* $\{\mathfrak{R}_{\Sigma_0}, \ldots, \mathfrak{R}_{\Sigma_k}\}$ *of ribbon graph subcomplexes of* K_n, *the subcomplex*

$$\bigcap_{i=0}^{k} \mathfrak{R}_{\Sigma_i}$$

of K_n *is either empty or contractible.*

The main tool in analyzing these intersections is the $K_{\min}$ subcomplexes of K_n, which are used by Culler and Vogtmann [1986] to show that K_n is contractible. The definition of the $K_{\min}$ complexes involves the following norm defined for each finite set of conjugacy classes of F_n. Let $\mathscr{C}$ denote the set of all conjugacy classes of F_n. For a marked rose ρ, an element $w \in \mathscr{C}$ can be represented by a unique reduced edge path in ρ.

Definition 5.3. Let W be a finite set of conjugacy classes of F_n and ρ a marked rose in K_n. The *norm* $\|\rho\|_W$ of ρ with respect to W is the sum of the number of edges in each reduced edge path in ρ that corresponds to an element of W.

If X is a basis labeling the edges of ρ, then $\|\rho\|_W$ is sometimes written $\|X\|_W$. The $K_{\min}$ subcomplex for W is defined as the union of the stars of the roses ρ for which $\|\rho\|_W$ is minimal over all marked roses. In [Vogtmann 2002] these complexes are denoted K_W, and we will follow that notation here. To prove that the entire complex K_n is contractible, Culler and Vogtmann prove that $K_W \simeq K_n$ for any finite set W. They then find a set of conjugacy classes such that K_W is the star of a single marked rose and therefore contractible. Putting these two facts together we have:

Lemma 5.4 [Culler and Vogtmann 1986]. *K_W is contractible for any finite set $W \subseteq \mathscr{C}$.*

Proposition 5.2 is proved by finding a deformation retraction from a suitable K_W to $\bigcap \mathfrak{R}_{\Sigma_i}$. We begin by studying of the behavior of the norm with respect to Whitehead automorphisms. For us, the traditional Whitehead automorphisms are less convenient to work with than a slightly modified version, given in [Hoare 1979]. This is because the effect of an automorphism on the star graph of a set of conjugacy classes (defined below), is easier to describe using this definition rather than the classical definition of Whitehead automorphism.

Definition 5.5. For a basis X and a subset $A \subseteq X \cup X^{-1}$ for which there is a letter $a \in X \cup X^{-1}$ such that $a \in A$ but $a^{-1} \notin A$, the automorphism mapping a to a^{-1} whose action on $X \cup X^{-1} - \{a, a^{-1}\}$ is given by

$$\begin{cases} x \mapsto axa^{-1} & \text{if } x \in A \text{ and } x^{-1} \in A; \\ x \mapsto xa^{-1} & \text{if } x \in A \text{ and } x^{-1} \notin A; \\ x \mapsto ax & \text{if } x \notin A \text{ and } x^{-1} \in A; \\ x \mapsto x & \text{if } x \notin A \text{ and } x^{-1} \notin A \end{cases} \tag{3}$$

will be called a *Whitehead automorphism* and will be denoted by (A, a).

Warning. This definition differs from the classical Whitehead automorphism in that the latter fix a. This is the only difference, but it allows us to prove the next result.

Lemma 5.6. *The totality of Whitehead automorphisms obtained as in the preceding definition generate the group* $\mathrm{Aut}(F_n)$.

Proof. The Neilson automorphisms generate $\mathrm{Aut}(F_n)$ [Magnus et al. 1966, Theorem 3.2]. It is straightforward to write any Neilson automorphism as a product of the Whitehead automorphisms of Definition 5.5. □

Another important fact about Whitehead automorphisms of this type is the peak reduction lemma (for Whitehead automorphisms as defined here).

Peak reduction lemma [Hoare 1979, Lemma 3]. *Fix a basis X of F_n and finite set $W \subseteq \mathscr{C}$. If there is an automorphism $\psi \in \mathrm{Aut}(F_n)$ such that $\|X\|_W \geq \|X\|_{\psi W}$ then ψ can be written as a product $\psi = \tau_1 \tau_2 \cdots \tau_k$ of Whitehead automorphisms such that*

$$\begin{aligned} \|X\|_W > \|X\|_{\tau_k W} > \|X\|_{\tau_{k-1}\tau_k W} > \cdots > \|X\|_{\tau_l \tau_{l+1} \cdots \tau_k W} \\ = \|X\|_{\tau_{l-1}\tau_l \tau_{l+1} \cdots \tau_k W} = \cdots = \|X\|_{\psi W}. \end{aligned} \tag{4}$$

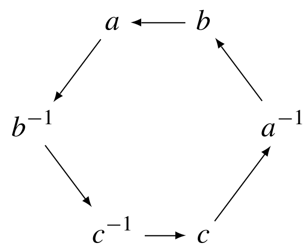


Figure 3. $S_W(X)$ for $X = \{a,\ b,\ c\}$, $W = \{aba^{-1}b^{-1}c,\ c^{-1}\}$.

The star graph of W with respect to X will allow us to study the behavior of $\|\cdot\|_W$ with respect to Whitehead automorphisms. Recall that star graph of $W \subseteq \mathscr{C}$ with respect to the basis X is the graph with vertex set $X \cup X^{-1}$ and with a directed edge from x to y^{-1} for every time the subword xy appears among the conjugacy classes in W, viewed as cyclic words in the alphabet $X \cup X^{-1}$; see Figure 3. The star graph of W with respect to X will be denoted by $S_W(X)$, or by $S_W(\rho)$ if we are thinking of X as a set of labels on the marked rose ρ.

To prove the peak reduction lemma, Hoare describes a three-step process for constructing $S_{\tau W}(X)$ from $S_W(X)$ for a Whitehead automorphism τ. If $\tau = (A, a)$, the steps are:

(1) Add two new vertices $\alpha, \bar{\alpha}$. Replace every edge going from a vertex in A to a vertex in A' (the complement of A) by a pair of edges, one from the vertex in A to α and another from $\bar{\alpha}$ to the vertex in A'. Replace every edge going from a vertex in A' to a vertex in A by a pair of edges, one from the vertex in A' to $\bar{\alpha}$ and another from α to the vertex in A.

(2) Switch the letter a with α, and a^{-1} with $\bar{\alpha}$.

(3) Do the reverse of (1), reconnecting edges incident to α and $\bar{\alpha}$ according to the cyclic words in W that produced them.

Because $\|X\|_W$ is the number of edges in $S_W(X)$, this process gives the following procedure for calculating the effect of a Whitehead automorphism on the norm. Consider the Whitehead automorphism $\tau = (A, a)$. Draw a circle C in the plane and immerse $S_W(X)$ in the plane in such a way that each vertex of A lies inside the circle, each vertex of A' lies outside the circle, no pair of edges of $S_W(X)$ intersect each other at a point of C, and $\#(S_W(X) \cap C)$ is minimal over all such immersions. Then

$$\|X\|_W - \|X\|_{\tau W} = \mathrm{val}(a) - \#(S_W(X) \cap C), \tag{5}$$

where val(a) is the valence of the vertex of $S_W(X)$ corresponding to a.

When $W = W_\Sigma$ is the set of boundary classes of a marked surface, we will be more concerned with $\|X\|_W - \|\tau^{-1}X\|_W$ than with $\|X\|_W - \|X\|_{\tau W}$, because if X is the set of labels on the edges of a marked rose ρ, then $\tau^{-1}X$ is the set of labels on $\rho \cdot \tau$. (5) will suffice because

$$\|\tau^{-1}X\|_W = \|X\|_{\tau W};$$

the latter equality can be seen from the observation that if $\tau(a) = x_1 x_2 \cdots x_k$ is an expression for $\tau(a)$ in terms of the basis X, then $a = \tau^{-1}(x_1)\tau^{-1}(x_2)\cdots\tau^{-1}(x_k)$ is an expression for a in terms of the basis $\tau^{-1}X$. The interpretation of this observation in terms of star graphs is

$$S_W(\tau^{-1}X) \approx S_{\tau W}(X). \tag{6}$$

Lemma 5.7. *Let W be a finite set of conjugacy classes of F_n. Suppose that for some basis X, the graph $S_W(X)$ is a cycle. Then $\|X\|_W$ is minimal over all bases of F_n, and if Y is another basis with $\|Y\|_W = \|X\|_W$, then $S_W(Y)$ is also a cycle.*

Proof. Since $S_W(X)$ is a cycle, any circle separating some generator from its inverse must intersect at least two edges of the graph. Since all vertices have valence 2, Equations (5) and (6) imply that no Whitehead automorphism can take X to a basis that reduces the sum of the lengths of the minimal representatives for the classes in W. But if there is any automorphism reducing the sum of the lengths of the conjugacy classes in W, the peak reduction lemma implies that a Whitehead automorphism reduces the length. Thus $\|X\|_W$ must be minimal over all bases for F_n.

Now suppose that Y is another basis with $\|Y\|_W = \|X\|_W$. Since $\mathrm{Aut}(F_n)$ acts transitively on bases of F_n, we may choose $\psi \in \mathrm{Aut}(F_n)$ with $Y = \psi^{-1}X$. By the peak reduction lemma and Equation (6), there is a sequence $\tau_1, \ldots, \tau_l$ of Whitehead automorphisms such that $\psi = \tau_1, \ldots, \tau_l$ and $\|\tau_l^{-1}\tau_{l-1}^{-1}\cdots\tau_i^{-1}X\|_W = \|X\|_W$ for $i = 1, \ldots, l$. Thus without loss of generality, we may assume that $\psi = \tau$ is a Whitehead automorphism.

We now use Hoare's method to construct the star graph $S_{\tau W}(X) \approx S_W(\tau^{-1}X)$. Since $\|\tau^{-1}X\|_W = \|X\|_W$, the circle separating A from A' in the star graph must intersect only two edges of the graph; otherwise the norm would increase. Therefore, the subgraph spanned by the vertices of A is a simple path, and the same is true for the subgraph spanned by A'. Step (1) of Hoare's procedure produces a graph consisting of two disjoint cycles, one containing α and the other containing $\bar{\alpha}$. Step (2) keeps α and $\bar{\alpha}$ in separate cycles, but they may switch cycles. Step (3) breaks these two cycles at α and $\bar{\alpha}$, and reconnects the ends of the resulting line segments to form a cycle, which is $S_{\tau W}(X)$. Since $S_W(Y) = S_W(\tau^{-1}X) \approx S_{\tau W}(X)$, $S_W(Y)$ is a cycle. □

We will use Lemma 5.7 to analyze which marked graphs can be drawn in a particular surface (Σ, s). If Σ has genus g and s punctures, and if

$$X = \{a_1, b_1, a_2, b_2, \ldots, a_{2g}, b_{2g}, c_1, \ldots, c_{s-1}\}$$

is a standard, geometric basis of $F_n = \pi_1(\Sigma)$, we have

$$W_{(\Sigma,s)} = \{[a_1, b_1][a_2, b_2] \cdots [a_{2g}, b_{2g}]c_1 \cdots c_{s-1},\ c_1^{-1},\ c_2^{-1},\ \ldots,\ c_{s-1}^{-1}\}.$$

Thus $S_W(X)$ is a cycle, and as a consequence of Lemma 5.7 we have:

Corollary 5.8. *If W is the set of boundary classes of a surface Σ, then*

$$\min_{\rho} \|\rho\|_W = 2n$$

and $S_W(\rho')$ is a cycle for any rose ρ' minimizing $\|\cdot\|_W$.

The next two lemmas characterize the marked graphs that lie in $\mathfrak{R}_{[\Sigma,s]}$.

Lemma 5.9. *The marked graph $\Gamma = (\Gamma, g)$ can be drawn in (Σ, s) if and only if the set of reduced edge cycles of Γ representing W_Σ traverses each edge of Γ exactly once in each direction.*

Proof. Suppose that (Γ, g) can be drawn in (Σ, s). Cutting Σ along Γ produces a collection of punctured disks, one for each puncture. The oriented boundaries of these disks correspond to the conjugacy classes of the boundary of Σ. Together they traverse each edge of Γ once in each direction. Thus, if (Γ, g) can be drawn in (Σ, s), the set of reduced edge cycles of Γ representing the boundary classes of (Σ, s) traverses each edge exactly once in each direction.

For the converse, we first construct another marked surface Σ' whose set of oriented boundary classes is also W_Σ, by showing that the boundary cycles in Γ induce a ribbon structure. We then use this surface to draw Γ in Σ. Let v be a vertex of Γ and define a polycyclic order on the half-edges at v by declaring that b^- follows a^+ if ab appears in the reduced boundary cycles in W_Σ. This definition may give more than one cycle of half-edges at some vertices, so it may not provide a cyclic order at each vertex. To show that it is indeed a cyclic order, we work by induction on the number of vertices in Γ.

If Γ has one vertex, then Γ is a rose and $\|\Gamma\|_{W_\Sigma} = 2n$. Since W_Σ is the set of boundary classes for a surface, Corollary 5.8 implies that the star graph $S_{W_\Sigma}(\Gamma)$ is a single connected cycle, which means that there is only one cycle of half-edges at the vertex of Γ, and our definition gives a cyclic ordering. Now, suppose that Γ has k vertices and assume by induction that any graph with fewer than k vertices and with the boundary classes traversing each edge exactly once in each direction has a single cycle at each vertex. Choose an edge e of Γ that is not a loop, and collapse it to obtain the marked graph $\Gamma' = \Gamma/e$. The reduced edge cycles of Γ'

representing the elements of W_Σ traverse each edge exactly once in each direction, and Γ' has $k-1$ vertices. Therefore Γ' has one cycle at each vertex. If v_1 and v_2 are the two vertices coalesced to the vertex $v \in \Gamma'$ during the collapse of edge e, then there is a single cycle at every vertex of Γ other than v_1 and v_2. A priori, the cycles at v can be formed by taking the cycles v_1 and v_2 and combining the one containing e^+ with the one containing e^-. Since there is only one cycle at v, each of v_1 and v_2 must possess only one cycle, which finishes the induction step that we have a cyclic ordering of the half-edges at each vertex.

Let $\mathbb{O}$ be the ribbon structure just constructed. Both Σ and the marked surface $\Sigma' = |(\Gamma, g, \mathbb{O})|$ are orientable surfaces with the same number of punctures and the same fundamental group; hence they are homeomorphic. By the construction of $\mathbb{O}$, the set of boundary classes of Σ' is W_Σ. Label each puncture of Σ and Σ' with the corresponding conjugacy class of W_Σ and choose a homeomorphism $f : \Sigma' \to \Sigma$ that preserves the labels of the punctures. Now, $f \circ i$ embeds Γ into Σ as a strong deformation retract, but this embedding may not induce the same marking as g. That is to say, the diagram

$$\begin{array}{ccc} \Gamma & \overset{i}{\hookrightarrow} & \Sigma' \\ {\scriptstyle g}\uparrow & & \downarrow{\scriptstyle f} \\ R_0 & \underset{s}{\longrightarrow} & \Sigma \end{array}$$

may not commute up to homotopy. However, the outer automorphism given by $f_* \circ i_* \circ g_* \circ (s_*)^{-1}$ stabilizes W_Σ. By [Zieschang et al. 1980, Theorem 5.15.3], it is induced by an element θ in the orientation-preserving mapping class group of Σ. Embedding Γ into Σ by $\theta^{-1} \circ f \circ i$ gives the same marking as g. Thus Γ can be drawn in Σ. □

Lemma 5.10. *Let (Σ, s) be a homotopy marked surface and ρ a marked rose. Then $\rho \in K_{W_\Sigma}$ if and only if $\rho \in \mathfrak{R}_\Sigma$.*

Proof. By Corollary 5.8, the minimal value of $\|\cdot\|_{W_\Sigma}$ is $2n$. Any $\rho \in \mathfrak{R}_\Sigma$ can be drawn in Σ. By cutting Σ along ρ, we see that $\|\rho\|_{W_\Sigma} = 2n$, so $\rho \in K_{W_\Sigma}$.

Conversely, suppose that $\rho \in K_{W_\Sigma}$. Since ρ minimizes $\|\cdot\|_{W_\Sigma}$, Corollary 5.8 implies that the star graph $S_{W_\Sigma}(\rho)$ must be a cycle. Therefore each label in ρ appears exactly once with exponent $+1$ and once with exponent -1 in the minimal expressions for conjugacy classes of W_Σ in terms of a set of labels of ρ. This means that the set of reduced edge cycles in ρ that represents W_Σ traverses each edge of ρ exactly once in each direction. By Lemma 5.9, ρ can be drawn in (Σ, s). □

This lemma implies that the roses in $\mathfrak{R}_\Sigma$ coincide with those in K_{W_Σ}. Since K_{W_Σ} is the union of the stars of its roses, $\mathfrak{R}_\Sigma \subseteq K_{W_\Sigma}$. To find a graph in $\mathfrak{R}_\Sigma$ lying near a particular graph in $K_{W_\Sigma} - \mathfrak{R}_\Sigma$, we use the following lemma.

Lemma 5.11. *If* $\Gamma = (\Gamma, g) \in K_{W_\Sigma}$, *there exists a* (*possibly empty*) *forest* $\Phi_\Sigma(\Gamma)$ *such that for any forest* $\Phi \subseteq \Gamma$,

$$\Gamma/\Phi \in \mathfrak{R}_\Sigma \iff \Phi \supseteq \Phi_\Sigma(\Gamma).$$

Proof. Let $\Phi_\Sigma(\Gamma)$ be the subgraph of Γ consisting of all the edges of Γ that are not traversed exactly once in each direction by the set of reduced edge paths representing the boundary classes of Σ. Since $\Gamma \in K_{W_\Sigma}$, and K_{W_Σ} is the union of the stars of its roses, there is a maximal tree T in Γ such that Γ/T is a rose in K_{W_Σ}. By Lemma 5.10 this rose is in $\mathfrak{R}_\Sigma$, so it can be drawn in Σ. Therefore every edge of $\Gamma - T$ is traversed exactly once in each direction by the set of conjugacy classes in W_Σ. This means that $\Phi_\Sigma(\Gamma) \subseteq T$, so that $\Phi_\Sigma(\Gamma)$ is a forest.

Given any forest Φ in Γ, Lemma 5.9 implies that Γ/Φ can be drawn in Σ exactly when the boundary cycles traverse each edge of Γ/Φ once in each direction. This happens exactly when $\Phi_\Sigma(\Gamma) \subseteq \Phi$. □

These lemmas would allow us, at this time, to define a retraction from K_{W_Σ} to $\mathfrak{R}_\Sigma$ by taking a graph $\Gamma \in K_W$ to $\Gamma/\Phi_\Sigma(\Gamma)$ proving the following well-known proposition without having to appeal to the contractibility of Teichmüller space or the identification of the ribbon graph complex with the decorated Teichmüller space.

Proposition 5.12. *For any marked surface* $[\Sigma, s]$, *the ribbon graph complex* $\mathfrak{R}_{[\Sigma,s]}$ *is contractible.*

We postpone this proof until it is covered by the proof of contractibility for arbitrary simplices of the nerve. For higher-dimensional simplices, we need a set of conjugacy classes that captures the properties of a graph that can be drawn in several different surfaces. This set emphasizes a conjugacy class according to the number of the surfaces in question of which it is a boundary. We start by describing some general properties of collections of finite sets of conjugacy classes of F_n. For the proof of Proposition 5.2, we will specialize to the case that the sets of conjugacy classes are actually the boundary classes of marked surfaces.

Definition 5.13. For a collection $\sigma = \{W_0, \dots, W_k\}$ of finite sets of conjugacy classes of F_n, define

$$W_\sigma := \left\{[\alpha_1]^{n_1}, \dots, [\alpha_l]^{n_l}\right\},$$

where $\bigcup_{i=0}^k W_i = \{[\alpha_1], \dots, [\alpha_l]\}$, and n_j is the number of times that the conjugacy class $[\alpha_j]$ appears in the W_i.

Note that $[\alpha]$ and $[\alpha^{-1}]$ both may appear in W_σ. We use the letter σ for the set $\{W_0, \dots, W_k\}$ because this definition will be applied to a simplex $\sigma = \langle \Sigma_0, \dots, \Sigma_k \rangle$ of $\mathcal{N}$, with $W_i = W_{\Sigma_i}$. We will use the notation W_σ in this situation as well.

Lemma 5.14. *For σ and W_σ as above, let*

$$A = \min_\rho \|\rho\|_{W_\sigma} \quad \textit{and} \quad A_i = \min_\rho \|\rho\|_{W_i}.$$

Then $A = A_0 + \cdots + A_k$ if and only if $\bigcap_{i=0}^k K_{W_i} \neq \varnothing$.

Proof. Suppose that $W_\sigma = \{w_1^{n_1}, \ldots, w_l^{n_l}\}$. For any marked rose ρ,

$$\|\rho\|_{W_\sigma} = \sum_{i=0}^{l} n_i \|\rho\|_{\{w_i\}} = \sum_{j=0}^{k} \|\rho\|_{W_j}. \tag{7}$$

Choose any marked rose ρ_1 with $\|\rho_1\|_{W_\sigma} = A$. Now, ρ_1 may not minimize every $\|\cdot\|_{W_i}$, so

$$A_0 + \cdots + A_k \leq \sum_{j=0}^{k} \|\rho_1\|_{W_j} = \|\rho_1\|_{W_\sigma} = A, \tag{8}$$

where the first equality comes from Equation (7).

If $\bigcap_{i=0}^k K_{W_i}$ is nonempty, there is a single marked rose ρ_2 with $\|\rho_2\|_{W_i} = A_i$ for all i. Thus

$$A \leq \|\rho_2\|_{W_\sigma} = \sum_{i=0}^{k} \|\rho_2\|_{W_i} = A_0 + \cdots + A_k.$$

Together with (8) this implies that $A = A_0 + \cdots + A_k$.

Conversely, if $A = A_0 + \cdots + A_k$, then using the ρ_1 from above we have

$$A_0 + \cdots + A_k = A = \|\rho_1\|_{W_\sigma} = \sum_{j=0}^{k} \|\rho_1\|_{W_j}. \tag{9}$$

Again the last equality comes from (7). Now, $\|\rho_1\|_{W_i} \geq A_i$, so by (9) we have $\|\rho_1\|_{W_i} = A_i$ for each i. Hence, $\rho_1 \in K_{W_i}$ for each i, and $\bigcap K_{W_i} \neq \varnothing$. □

Changing the viewpoint slightly we get:

Corollary 5.15. *For any finite collection of finite sets of conjugacy classes, $\sigma = \{W_0, \ldots, W_k\}$, $K_{W_\sigma} = \bigcap_{i=0}^k K_{W_i}$ if the right-hand side is nonempty.*

Setting $k = 0$ here provides a proof of Proposition 5.12. The final lemma we need for the proof of Proposition 5.2 is this:

Poset lemma [Quillen 1973]. *Let $f : P \to P$ be a poset map from the poset P to itself such that $p \leq f(p)$ for all $p \in P$. Then f induces a deformation retraction from the geometric realization of P to the geometric realization of its image $f(P)$.*

Proof of Proposition 5.2. Let $\sigma = \langle \Sigma_0, \ldots, \Sigma_k \rangle$ be a simplex of $\mathcal{N}_n$. Denote by $\mathfrak{R}_\sigma$ the intersection

$$\mathfrak{R}_\sigma := \bigcap_{i=0}^{k} \mathfrak{R}_{[\Sigma_i, s_i]},$$

and let W_σ be the set of conjugacy classes given by Definition 5.13. Now $\mathfrak{R}_\sigma$ contains a rose, since $\langle \Sigma_0 \ldots \Sigma_k \rangle$ is a simplex of $\mathcal{N}_n$. To simplify the notation, set $W_i := W_{\Sigma_i}$. Since $\mathfrak{R}_{\Sigma_i} \subseteq K_{W_i}$, we have $\mathfrak{R}_\sigma \subseteq \bigcap K_{W_i}$. Hence, $\bigcap K_{W_i} \neq \varnothing$, and by Corollary 5.15, $\bigcap K_{W_i} = K_{W_\sigma}$. We will define a deformation retraction $K_{W_\sigma} \to \mathfrak{R}_\sigma$ by collapsing in each graph the minimal forest that takes that graph to a graph in $\mathfrak{R}_\sigma$.

Let $\Gamma \in K_{W_\sigma}$. Since $\bigcap K_{W_i} = K_{W_\sigma}$, Lemma 5.11 implies that there exists a minimal forest $\Phi_{\Sigma_i} \subseteq \Gamma$ collapsing Γ to a graph in $\mathfrak{R}_{\Sigma_i}$. Set

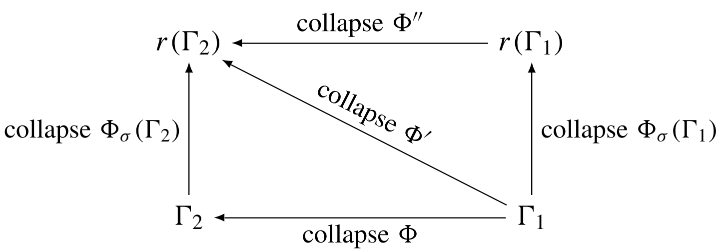
$$\Phi_\sigma(\Gamma) := \Phi_{\Sigma_0}(\Gamma) \cup \cdots \cup \Phi_{\Sigma_k}(\Gamma).$$

Since $\Gamma \in K_{W_\sigma}$, there is a spanning tree T collapsing Γ to a rose: $\Gamma/T \in K_{W_\sigma}$. By Corollary 5.15, $\Gamma/T \in K_{W_i}$ for each i. So, by Lemma 5.10, $\Gamma/T \in \mathfrak{R}_{\Sigma_i}$ for each i and therefore $\Phi_{\Sigma_i}(\Gamma) \subseteq T$ for each i. Hence $\Phi_\sigma(\Gamma) \subseteq T$. Since T is a tree, $\Phi_\sigma(\Gamma)$ is a forest.

Now we define a map r from the vertex set of K_{W_σ} to the vertex set of $\mathfrak{R}_\sigma$ by $r(\Gamma) = \Gamma/\Phi_\sigma(\Gamma)$. We claim that r induces a simplicial map

$$r : K_{W_\sigma} \to \mathfrak{R}_\sigma.$$

It will suffice to show that r takes adjacent vertices to the same vertex or adjacent vertices because both K_{W_σ} and $\mathfrak{R}_\sigma$ are determined by their 1-skeletons. To do this, suppose that Γ_1 and Γ_2 represent adjacent vertices in K_{W_σ}. By possibly switching the names of the graphs, we can write $\Gamma_2 = \Gamma_1/\Phi$ for some forest Φ. If $\Phi \subseteq \Phi_\sigma(\Gamma_1)$, then $r(\Gamma_1) = r(\Gamma_2)$. If $\Phi \not\subseteq \Phi_\sigma(\Gamma)$ then $r(\Gamma_1) \neq r(\Gamma_2)$. The diagram below represents a small portion of K_n in this case, with edges represented by arrows.



To show that r takes Γ_1 and Γ_2 to adjacent vertices, as the diagram suggests, we need justify that there is such a forest Φ'', as indicated in the diagram. The forest Φ'' is constructed as follows. Let $\Phi' = \Phi_\sigma(\Gamma_1) \cup \Phi$. Then Φ' is the subforest of Γ_1

such that $r(\Gamma_2) = \Gamma_1/\Phi'$. If Φ'' is the subgraph of $r(\Gamma_1)$ consisting of the images of the edges in $\Phi' - \Phi_\sigma(\Gamma_1)$ then Φ'' is a forest and $r(\Gamma_1)/\Phi'' = r(\Gamma_2)$. Therefore $r(\Gamma_1)$ and $r(\Gamma_2)$ are adjacent, proving that r induces a simplicial map.

That r is a retraction follows from the implication $\Gamma \in \mathfrak{N}_\sigma \Rightarrow \Phi_\sigma(\Gamma) = \varnothing$ and the fact that, by Lemma 5.9, the image of r is contained in $\mathfrak{N}_{\Sigma_i}$ for each i. Therefore $r(K_{W_\sigma}) = \mathfrak{N}_\sigma$. To see that r is a deformation retraction, we will use the poset lemma. Partially order the vertices of K_{W_σ} by setting $\Gamma_1 < \Gamma_2$ if Γ_1 can be collapsed to Γ_2. Then K_{W_σ} is the geometric realization of the poset of its vertices under this partial order. With respect to this partial order, r has the property that $\Gamma \le r(\Gamma)$. The poset lemma implies that r is a deformation retraction. Since K_{W_σ} is contractible, this finishes the proof. □

By Lemma 5.1, Proposition 5.2 proves that $\mathcal{N}_n \simeq K_n$. Since K_n is contractible by [Culler and Vogtmann 1986], so is $\mathcal{N}_n$. We record this:

Theorem 5.16. *$\mathcal{N}_n$ is contractible for all n.*

6. Simplex stabilizers

As mentioned before, the action of $\mathrm{Out}(F_n)$ on K_n gives an action of $\mathrm{Out}(F_n)$ on $\mathcal{N}_n$. To describe stabilizers of this action, we fix some notation. Let Σ be a surface with boundary and/or punctures and with free fundamental group. As in the introduction, the pure mapping class group of Σ is the group of homeomorphisms of Σ that are the identity on the boundary and fix the punctures, up to isotopy relative to the boundary. The extended mapping class group of Σ is the group of isotopy classes of homeomorphisms of Σ. Thus the extended mapping class group contains orientation reversing homeomorphisms, while the pure mapping class group does not, provided that Σ has boundary. We will use $P\Gamma(\Sigma)$ and $\Gamma(\Sigma)$ to represent the pure and extended mapping class groups of Σ.

If (Σ, s) is a homotopy marked surface then the identification of $\pi_1(\Sigma)$ with F_n given by the marking s induces a homomorphism from $\Gamma(\Sigma)$ to $\mathrm{Out}(F_n)$. This homomorphism is defined by sending a homeomorphism of Σ to the outer automorphism of $\pi_1(\Sigma)$ that it represents. By [Zieschang et al. 1980, Theorem 5.15.3], this homomorphism is injective, and its image is the subgroup of $\mathrm{Out}(F_n)$ consisting of outer automorphisms that take W_Σ to W_Σ or $(W_\Sigma)^{-1}$. Denote this subgroup by $\mathrm{MCG}^\pm(\Sigma, s)$. Denote the image of $P\Gamma(\Sigma)$ by $\mathrm{PMCG}(\Sigma, s)$ and the image of the orientation-preserving subgroup of $\Gamma(\Sigma)$ by $\mathrm{MCG}(\Sigma, s)$. Note that these subgroups depend on the marking s. Thus the difference between $\Gamma(\Sigma)$ and $\mathrm{MCG}^\pm((\Sigma, s))$ is that $\mathrm{MCG}^\pm(\Sigma)$ is viewed as a subgroup of $\mathrm{Out}(F_n)$, and this subgroup depends on the marking s. The same is true for the pure mapping class groups.

Finally, let $\mathrm{Stab}(\mathfrak{R}_{[\Sigma,s]})$ be the subgroup of $\mathrm{Out}(F_n)$ stabilizing $\mathfrak{R}_{[\Sigma,s]}$ setwise, so $\mathrm{Stab}(\mathfrak{R}_{[\Sigma,s]})$ is the stabilizer of the vertex of $\mathcal{N}_n$ that corresponds to $[\Sigma, s]$.

Theorem 6.1. $\mathrm{Stab}(\mathfrak{R}_{[\Sigma,s]}) = \mathrm{MCG}^{\pm}(\Sigma, s)$.

Proof. Since $\psi \in \mathrm{MCG}^{\pm}(\Sigma, s)$ implies $(\Sigma, s) \cdot \psi = (\Sigma, s \circ |\psi|) \sim_{\pm} (\Sigma, s)$ (and hence $\mathfrak{R}_{[\Sigma,s]} \cdot \psi = \mathfrak{R}_{[\Sigma,s]}$), we have $\mathrm{Stab}(\mathfrak{R}_{[\Sigma,s]}) \supset \mathrm{MCG}^{\pm}(\Sigma, s)$. For the other inclusion, suppose that $\psi \in \mathrm{Stab}(\mathfrak{R}_{[\Sigma,s]})$. Then

$$\mathfrak{R}_{[\Sigma,s]} = (\mathfrak{R}_{[\Sigma,s]}) \cdot \psi = \mathfrak{R}_{[\Sigma,s\circ|\psi|]}.$$

By Proposition 4.6, $[\Sigma, s] = [\Sigma, s \circ |\psi|]$, so there is a homeomorphism $h : \Sigma \to \Sigma$ that makes the diagram

$$\begin{array}{ccc} R_0 & \xrightarrow{\;s\;} & \Sigma \\ {\scriptstyle |\psi|}\big\downarrow & & \big\downarrow{\scriptstyle h} \\ R_0 & \xrightarrow[\;s\;]{} & \Sigma \end{array}$$

commute up to homotopy. Now, h takes the boundary classes W_Σ to W_Σ or $(W_\Sigma)^{-1}$. Thus ψ does also. By Zieschang's theorem, $\psi \in \mathrm{MCG}^{\pm}(\Sigma)$. □

To describe the stabilizer of a higher-dimensional simplex, we use a certain kind of stabilizer of a set of conjugacy classes of F_n, as studied in [McCool 1975]. Following the definitions there, we consider ordered m-tuples

$$(w_1, \ldots, w_m)$$

of conjugacy classes in F_n. The symmetric group S_m acts on the set of m-tuples by permuting the coordinates. The *inverting operations* $\tau_1, \ldots, \tau_m$ act on the set of m-tuples by

$$\tau_i(w_1, \ldots, w_i, \ldots, w_m) = (w_1, \ldots, w_i^{-1}, \ldots, w_m).$$

The group S_m, together with the τ_i's, generates the subgroup $\Omega_m \cong S_m \wr \mathbb{Z}_2$ of permutations of the set of m-tuples of conjugacy classes of F_n known as the extended symmetric group. The group $\mathrm{Out}(F_n)$ also acts on the set of m-tuples of conjugacy classes by acting individually on the coordinates. McCool [1975] makes the following definition in the setting of $\mathrm{Aut}(F_n)$, but we will use it also in the setting of $\mathrm{Out}(F_n)$.

Definition 6.2. For U an m-tuple of conjugacy classes and subgroup $G \leq \Omega_m$, define the subgroup $\mathcal{A}_{U,G}$ of $\mathrm{Out}(F_n)$ by

$$\mathcal{A}_{U,G} := \{\theta \in \mathrm{Out}(F_n) : \theta U \in GU\},$$

where $GU = \{gU : g \in G\}$.

For a simplex σ of $\mathcal{N}_n$, let G_σ denote the stabilizer of the simplex σ of $\mathcal{N}_n$. If $\sigma = v$ is the vertex corresponding to the marked surface (Σ, s), then $G_v = \mathrm{MCG}^{\pm}(\Sigma, s) = \mathcal{A}_{U,G}$ where U is the m-tuple of boundary classes (in any order) of a marked surface (Σ, s), and $G \leq \Omega_m$ is the subgroup generated by S_m together with the extended permutation $\tau_1\tau_2\cdots\tau_m$. To describe G_σ for a higher-dimensional simplex, we introduce some terminology.

Definition 6.3. Let $\sigma = \langle \Sigma_0, \dots, \Sigma_k \rangle$ be a simplex of $\mathcal{N}_n$, and let U_i be the m_i-tuple of boundary classes of Σ_i (again in any order). Set $m = m_0 + \cdots + m_k$ and denote by $U_\sigma = (U_0, \dots, U_k)$ the m-tuple constructed by listing the conjugacy classes from the U_i one after another, starting with those of U_0. Define $H_\sigma \leq \Omega_m$ as the subgroup generated by extended permutations of the following types:

(1) $\alpha \in S_m$ such that there is a permutation $\lambda \in S_k$ such that, for each i, α takes U_i to $U_{\lambda(i)}$, possibly with the entries of $U_{\lambda(i)}$ permuted;

(2) $\tau \in \Omega_m$ such that $\tau U_i = U_i$ or $\tau U_i = U_i^{-1}$ for each i.

Proposition 6.4. *For any simplex σ of $\mathcal{N}$, G_σ has the form $\mathcal{A}_{U_\sigma, H_\sigma}$.*

Proof. Formally, (1) can be written as $\alpha U_i \in S_{m_{\lambda(i)}} U_i$. An element $\theta \in G_\sigma$ permutes the equivalence classes of the marked surfaces Σ_i. This means that θ takes W_{Σ_i} to W_{Σ_j} or $W_{\Sigma_j}^{-1}$ for some j. Thus $\theta U_i \in S_{m_j} U_j$ or $S_{m_j} U_j^{-1}$ for some j. Since no two surfaces are taken to the same surface by θ, this means precisely that $\theta \in \mathcal{A}_{U_\sigma, G_\sigma}$ as defined above. □

7. Equivariant homology of the action of Out(F_n) on $\mathcal{N}_n$

For a cellular action of a group G on a contractible cell complex X, the equivariant spectral sequence for the action of G on X is a well-known spectral sequence that converges to a grading of the homology of G; see [Brown 1982, Chapter VII.7]. To describe this spectral sequence, let M be any G-module. Consider, for each p-cell σ of X, the G_σ-module $\mathbb{Z}_\sigma$. As an additive group, $\mathbb{Z}_\sigma$ is isomorphic to $\mathbb{Z}$. The module structure of $\mathbb{Z}_\sigma$ is given by having $g \in G_\sigma$ act as multiplication by $+1$ or -1, depending on whether g preserves or reverses the orientation of σ. The module $\mathbb{Z}_\sigma$ is called the *orientation module* of σ. Let

$$M_\sigma := \mathbb{Z}_\sigma \otimes_{\mathbb{Z}} M.$$

Fix a set Δ_p of representatives for the orbits of the p-cells of X under the action of G. The equivariant spectral sequence for the action of G on X takes the form

$$E^1_{pq} = \bigoplus_{\sigma \in \Delta_p} H_q(G_\sigma; M_\sigma) \Rightarrow H_{p+q}(G; M). \tag{10}$$

Applying this to the action of $\mathrm{Out}(F_n)$ on $\mathcal{N}_n$ and any $\mathrm{Out}(F_n)$-module M, we get a spectral sequence converging to $H_*(\mathrm{Out}(F_n); M)$. Since vertex stabilizers are extended mapping class groups and there is one orbit of vertex for each homeomorphism type of surface, the $p = 0$ column of the spectral sequence consists of direct sums of the homology groups of the extended mapping class groups. For $p > 0$, the simplex stabilizers are given by Proposition 6.4 and we have:

Theorem 7.1. *For any* $\mathrm{Out}(F_n)$*-module* M*, there is a spectral sequence of the form*

$$E^1_{pq} = \bigoplus_{\sigma \in \Delta_p} H_q(G_\sigma; M_\sigma) \Rightarrow H_{p+q}(\mathrm{Out}(F_n); M), \tag{11}$$

where Δ_0 *is the set of homeomorphism classes of punctured orientable surfaces with fundamental group* F_n*, and for a vertex* $v \in \Delta_0$ *corresponding to surface* Σ *the stabilizer* G_v *is the extended mapping class group* $\mathrm{MCG}^\pm(\Sigma)$*. For* $p > 0$*, each* G_σ *is a generalized stabilizer of the form* $\mathcal{A}_{U_\sigma, H_\sigma}$*.*

The map induced on homology by the inclusion $\mathrm{MCG}^\pm(\Sigma) \hookrightarrow \mathrm{Out}(F_n)$ appears in the spectral sequence as the left-hand edge map, which is defined for the general spectral sequence (10) as follows. Since there is nothing but zeroes to the left of the $p = 0$ column in the spectral sequence (10), E^∞_{pq} is a quotient of E^1_{pq}. The spectral sequence converges to a grading of $H_*(G; M)$, and the composition

$$\bigoplus_{v \in \Delta_0} H_q(G_v; M) = E^1_{0q} \twoheadrightarrow E^\infty_{0q} = Gr_0 H_q(G; M) \hookrightarrow H_q(G; M) \tag{12}$$

is the left-hand edge map of this spectral sequence. The left hand edge map is equal to the map induced on homology by the inclusion of G_v into G.

For sequence (11), if the vertex v corresponds to marked surface Σ, we have $G_v = \mathrm{MCG}^\pm(\Sigma)$ and the restriction of the left-hand edge map to the subspace $H_q(G_v; M)$ is the map

$$H_q(\mathrm{MCG}^\pm(\Sigma); M) \to H_q(\mathrm{Out}(F_n); M)$$

induced by the inclusion $\mathrm{MCG}^\pm(\Sigma) \hookrightarrow \mathrm{Out}(F_n)$. Thus finding a bound on the rank of the left-hand edge map gives an upper bound on the contribution that the mapping class subgroups of $\mathrm{Out}(F_n)$ can make to the homology of $\mathrm{Out}(F_n)$. This will be the subject of the next section.

8. Analysis of E^∞

In this section, we specialize to rational coefficients and give a method for using Harer's homology stability theorems [1985] for mapping class groups to analyze the E^∞ page of spectral sequence (11). We continue to use $\mathrm{MCG}^\pm(\Sigma, s)$ for the image of $\Gamma(\Sigma, s)$ in $\mathrm{Out}(F_n)$, but we extend this notation to surfaces with boundary.

Hence, if Σ is a surface with boundary and s is a homotopy equivalence from the standard rose R_0 to Σ, we use $\mathrm{MCG}^{\pm}(\Sigma, s)$, $\mathrm{PMCG}(\Sigma, s)$, and $\mathrm{MCG}(\Sigma, s)$ to denote the images in $\mathrm{Out}(F_n)$ of the extended, pure and orientation-preserving mapping class groups of Σ in $\mathrm{Out}(F_n)$. As usual, these images depend on the marking s. We remark that there is a natural inclusion $\mathrm{P}\Gamma(\Sigma) \hookrightarrow \Gamma(\Sigma)$, which agrees with the inclusion of $\mathrm{PMCG}(\Sigma, s)$ into $\mathrm{MCG}^{\pm}(\Sigma, s)$.

More generally, if Σ_0 is a subsurface with boundary of the surface Σ, the inclusion $\Sigma_0 \hookrightarrow \Sigma$ induces a map $\alpha : \mathrm{P}\Gamma(\Sigma_0) \to \mathrm{P}\Gamma(\Sigma)$ defined by extending a homeomorphism of Σ_0 to all of Σ by the identity. Harer's stability theorem, quoted below, implies that α induces an isomorphism on homology in sufficiently high dimensions.

Theorem 8.1 [Harer 1985, Theorem 0.1]. *Let Σ_0 be a subsurface of Σ such that $\Sigma - \Sigma_0$ is connected, contains no punctures but is not simply connected. If the genus of Σ_0 is at least $3k - 1$, then $\alpha_* : H_k(\mathrm{P}\Gamma(\Sigma_0); \mathbb{Q}) \to H_k(\mathrm{P}\Gamma(\Sigma); \mathbb{Q})$ is an isomorphism.*

We will also need to analyze the effect of plugging a boundary component of Σ with a punctured disk. To this end, consider the maps Θ, Θ', Υ and Φ (between the appropriate surfaces) defined, respectively, by plugging a boundary component with a disk, plugging a boundary component with a punctured disk, plugging a puncture, and sewing a pair of pants to a boundary component. In the stable range, Theorem 8.1 applies to Φ. By making the appropriate identifications, $\Theta \circ \Phi$ induces the identity on homology, so in the stable range for Φ, the map Θ must induce an isomorphism. Since $\Theta = \Upsilon \circ \Theta'$, we have:

Lemma 8.2. *For $g \geq 3k - 2$, $(\Theta')_*$ is injective and Υ_* is surjective on the k-th homology.*

In order to relate these stability maps to the d^1 terms in the spectral sequence of Theorem 7.1, consider two marked surfaces (Σ, s) and (Σ', s'). Let ρ and ρ' be the marked images in Σ and Σ' of the marking rose. Suppose that there are separating simple closed curves $\gamma \subset \Sigma$ and $\gamma' \subset \Sigma'$ cutting off subsurfaces $\widetilde{\Sigma} \subset \Sigma$ and $\widetilde{\Sigma}' \subset \Sigma'$ with the following properties, illustrated in Figure 4.

(1) The basepoints of the marking roses lie on γ and γ'.

(2) No edge of the marking roses meets γ or γ' anywhere but at the basepoints of the roses.

(3) $\rho \cap \widetilde{\Sigma} \simeq \widetilde{\Sigma}$ and $\rho' \cap \widetilde{\Sigma}' \simeq \widetilde{\Sigma}'$.

(4) $\widetilde{\Sigma}$ and $\widetilde{\Sigma}'$ are homeomorphic by a homeomorphism taking each directed edge of $\rho \cap \widetilde{\Sigma}$ to a directed edge of $\rho' \cap \widetilde{\Sigma}'$ with the same labeling.

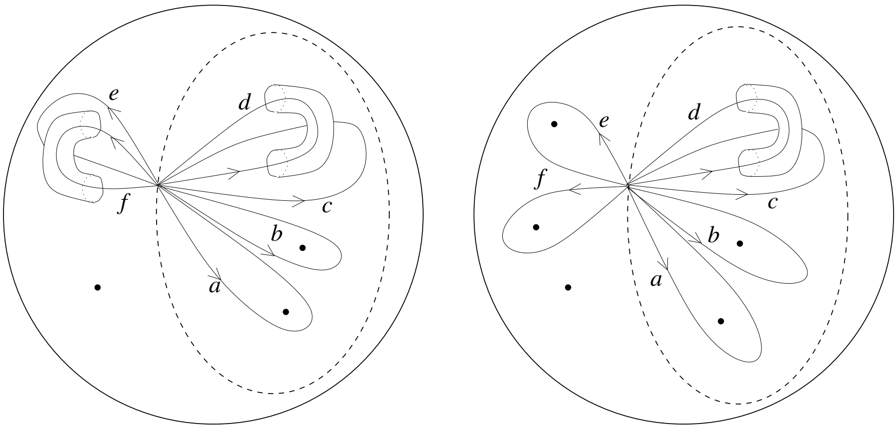


Figure 4. Markings that agree on a subsurface.

Definition 8.3. If the marked surfaces (Σ, s) and (Σ', s') satisfy conditions (1)–(4) above, the markings are said to *agree on the subsurfaces* $\widetilde{\Sigma}$ *and* $\widetilde{\Sigma}'$.

If the marked surfaces Σ and Σ' agree on the subsurface $\widetilde{\Sigma} = \widetilde{\Sigma}'$ then the corresponding vertices v_Σ and $v_{\Sigma'}$ in $\mathcal{N}$ span an edge e. The group $\mathrm{PMCG}(\widetilde{\Sigma})$ can be identified with a subgroup of the stabilizer of v_e. If Σ and Σ' are not homeomorphic surfaces, then the component of the d^1 map in spectral sequence (11) from $\mathrm{Stab}(e)$ to $\mathrm{Stab}(v_\Sigma)$ is the map induced by inclusion of $\mathrm{Stab}(e)$ into $\mathrm{Stab}(v_\Sigma)$. The same is true for Σ'. Therefore the following lemma will play the key role in determining bounds on how much homology in the E^1 page can survive until E^∞.

Lemma 8.4. *Let* $[\Sigma, s]$ *and* $[\Sigma', s']$ *be nonhomeomorphic marked surfaces such that*

(1) *the markings* s *and* s' *agree on subsurfaces* $\widetilde{\Sigma}$ *and* $\widetilde{\Sigma}'$ *of genus* $3k-1$,

(2) $\widetilde{\Sigma}$ *contains all but one of the punctures of* Σ, *and*

(3) $\mathfrak{R}_{[\Sigma,s]} \cap \mathfrak{R}_{[\Sigma',s']} \neq \varnothing$.

Let v *and* v' *be the vertices of* $\mathcal{N}_n$ *corresponding to* Σ *and* Σ' *and let* e *be the edge of* $\mathcal{N}_n$ *between* v *and* v'. *Then* $G_e = G_v \cap G_{v'}$ *and* $i_* : H_k(G_e; \mathbb{Q}) \to H_k(G_v; \mathbb{Q})$ *has rank at least*

$$\dim H_k(G_v; \mathbb{Q}) - \big(\dim H_k(\mathrm{P}\Gamma(\Sigma); \mathbb{Q}) - \dim H_k(\mathrm{P}\Gamma(\widetilde{\Sigma}); \mathbb{Q})\big).$$

Proof. First, $G_e = G_v \cap G_{v'}$ because no outer automorphism of F_n can switch $[\Sigma, s]$ with $[\Sigma', s']$.

Since the Dehn twists generate $\mathrm{P}\Gamma(\widetilde{\Sigma})$, the map $\alpha : \mathrm{P}\Gamma(\widetilde{\Sigma}) \to \mathrm{P}\Gamma(\Sigma)$ induced by $\widetilde{\Sigma} \hookrightarrow \Sigma$ is determined by its effect Dehn twists. Now, suppose that c is a simple closed curve in $\widetilde{\Sigma}$ with corresponding simple closed curve c' in $\widetilde{\Sigma}'$. When we use the maps i_s and $i_{s'}$ of Zieschang's theorem to identify $\mathrm{P}\Gamma(\Sigma)$ and $\mathrm{P}\Gamma(\Sigma')$ with the subgroups $\mathrm{PMCG}(\Sigma)$ and $\mathrm{PMCG}(\Sigma') \subseteq \mathrm{Out}(F_n)$, the Dehn twists $\delta_c \in \mathrm{P}\Gamma(\widetilde{\Sigma})$ and $\delta_{c'} \in \mathrm{P}\Gamma(\widetilde{\Sigma}')$ about c and c' correspond to the same outer automorphism of F_n. Therefore the image of the map $\alpha_1 : \mathrm{P}\Gamma(\widetilde{\Sigma}) \to \mathrm{P}\Gamma(\Sigma) = \mathrm{PMCG}(\Sigma)$ lies in the intersection $\mathrm{PMCG}(\Sigma) \cap \mathrm{PMCG}(\Sigma') \subseteq G_v \cap G_{v'}$, and we have a commutative diagram

$$\begin{array}{ccc} \mathrm{P}\Gamma(\widetilde{\Sigma}) & \xrightarrow{\alpha_1} & \mathrm{PMCG}(\Sigma) \\ \downarrow{\scriptstyle \beta_1} & & \cap\downarrow{\scriptstyle \beta_2} \\ G_e & \underset{i}{\hookrightarrow} & \mathrm{MCG}^{\pm}(\Sigma) = G_v \end{array} \tag{13}$$

where β_1 is the map α_1 viewed with a different range. We claim that the composition $\beta_2 \circ \alpha_1$ induces a map on homology that has rank at least

$$\dim H_k(G_v; \mathbb{Q}) - \big(\dim H_k(\mathrm{P}\Gamma(\Sigma); \mathbb{Q}) - \dim H_k(\mathrm{P}\Gamma(\widetilde{\Sigma}); \mathbb{Q})\big).$$

To see this, note that β_2 is the inclusion of a finite index subgroup into a supergroup. Therefore it induces a surjection on homology with rational coefficients. By Theorem 8.1 and Lemma 8.2, the map α_1 is injective, and the claim follows by a dimension counting argument. Now, the rank of the map induced by i must be at least the rank of the map induced by $i \circ \beta_1 = \beta_2 \circ \alpha_1$, finishing the proof. □

Our last proposition gives a bound on the rank of the restriction of the left-hand edge map in spectral sequence (11) to surfaces of large rank. It bounds the contribution that the homology of these surfaces' mapping class groups can make to the homology of $\mathrm{Out}(F_n)$. To simplify notation, let $\Gamma^s_{g,0}$ denote the extended mapping class of the surface of genus g with s punctures and no boundary, and let $\mathrm{P}\Gamma^s_{g,0}$ be the pure, orientation preserving mapping class group of this surface.

Proposition 8.5. *Let $k \geq 0$ and $n \geq 6k-2$. For $g \geq 3k-1$, let Σ^s_g be the punctured surface of genus g with s punctures and with $2g+s-1 = n$ (so that $\pi_1(\Sigma^s_g) \cong F_n$). By choosing particular markings of the Σ^s_g, we may identify the vector space*

$$A := \bigoplus_{g \geq 3k-1} H_k(\mathrm{MCG}(\Sigma^s_g); \mathbb{Q})$$

with a subspace of the E^1_{0k} term of spectral sequence (11) using trivial $\mathbb{Q}$ coefficients. The image of A in E^∞_{0k} has dimension no larger than

$$\dim H_k(\Gamma^t_{3k-1,0}; \mathbb{Q}) + \sum_{\substack{g \geq 3k \\ 2g+s-1=n}} \big(\dim H_k(\mathrm{P}\Gamma^s_{g,0}; \mathbb{Q}) - \dim H_k(\mathrm{P}\Gamma^{s-1}_{g,0}; \mathbb{Q})\big),$$

where $2(3k-1)+t-1 = n$.

Proof. Fix a marking $[\Sigma^t_{3k-1}, s_{3k-1}]$ of Σ^t_{3k-1}. For each $g > 3k-1$ we may choose a marking $[\Sigma^s_g, s_g]$ of Σ^s_g satisfying the conditions of Lemma 8.4 with $\Sigma = \Sigma^s_g$ and $\Sigma' = \Sigma^t_{3k-1}$. For $g \geq 3k-1$ let v_g denote the vertex of $\mathcal{N}_n$ corresponding to $[\Sigma^s_g, s_g]$ and for $g \geq 3k$, let e_g denote the edge between v_{3k-1} and v_g. By choosing the v_g and e_g as representatives for their $\mathrm{Out}(F_n)$ orbits, the vector spaces

$$(14)\qquad A := \bigoplus_{g\geq 3k-1} H_k(G_{v_g}; \mathbb{Q}_{v_g}) \quad \text{and} \quad B := \bigoplus_{g\geq 3k} H_k(G_{e_g}; \mathbb{Q}_{e_g})$$

can be identified with subspaces of E^1_{0k} and E^1_{1k} respectively. Note that $d^1(B) \subseteq A$. Now, G_{e_g} fixes e_g pointwise since no outer automorphism can switch v_g and v_{3k}. The same is true of G_{v_g}, so the modules $\mathbb{Q}_{v_g}$ and $\mathbb{Q}_{e_g}$ are actually trivial modules; $\mathbb{Q}_{v_g} = \mathbb{Q}$, $\mathbb{Q}_{e_g} = \mathbb{Q}$. Since $G_{v_g} = \mathrm{MCG}(\Sigma^s_g, s_g)$, the above definition of A agrees with the definition in the statement of the proposition, $A = \bigoplus_{g\geq 3k-1} H_k(\mathrm{MCG}(\Sigma^s_g); \mathbb{Q})$. Now, $G_{e_g} = G_{v_g} \cap G_{v_{3k-1}}$ as in Lemma 8.4. With these definitions, the (e_g, v_g)-component of d^1 is simply the map induced by the inclusion $G_{e_g} \hookrightarrow G_{v_g}$. By Lemma 8.4, this map has rank at least

$$(15)\qquad R_g = \dim H_k(G_{v_g}; \mathbb{Q}) - \big(\dim H_k(\mathrm{P}\Gamma^s_{g,0}; \mathbb{Q}) - \dim H_k(\mathrm{P}\Gamma^{s-1}_{3k-1,1}; \mathbb{Q})\big),$$

where $2g+s-1=n$. Note that R_g depends on g because, even though $H_k(\mathrm{P}\Gamma^s_{g,0}; \mathbb{Q})$ is independent of g, it depends on s, and s depends on g.

By (15), for each $g \geq 3k$ we may choose R_g vectors $\{w^g_1, \dots, w^g_{R_g}\}$ in $(\mathrm{im}\, d^1 \cap A)$ such that their projections onto $H_k(G_{v_g}; \mathbb{Q})$ are linearly independent. Let A_g be the subspace of A spanned by $\{w^g_1, \dots, w^g_{R_g}\}$. By the first direct sum decomposition in (14) and the choice of the vectors w^g_i, the subspaces $A_{3k}, \dots, A_{\lfloor \frac{n}{2} \rfloor}$ are linearly independent. Let A^∞ denote the image of A in E^∞. Since A^∞ is a quotient of $A/d^1(B)$, this means that

$$(16)\qquad \dim(A^\infty) \leq \dim(A) - \sum_{g\geq 3k} R_g.$$

By Harer's stability theorems, $H_k(\mathrm{P}\Gamma^{s-1}_{3k-1,1}) \cong H_k(\mathrm{P}\Gamma^{s-1}_{g,0})$. Now, substituting

$$\dim A = \sum_{g\geq 3k-1} \dim H_k(G_{v_g}; \mathbb{Q})$$

and (15) with $\dim H_k(\mathrm{P}\Gamma^{s-1}_{3k-1,1}) = \dim H_k(\mathrm{P}\Gamma^{s-1}_{g,0})$ into (16) gives

$$\dim(A^\infty) \leq$$

$$\dim H_k(\mathrm{MCG}(\Sigma_{3k-1}); \mathbb{Q}) + \sum_{\substack{g\geq 3k \\ 2g+s-1=n}} \big(\dim H_k(\mathrm{P}\Gamma^s_{g,0}; \mathbb{Q}) - \dim H_k(\mathrm{P}\Gamma^{s-1}_{g,0}; \mathbb{Q})\big),$$

as required. ∎

Acknowledgments

This paper presents a portion of the results of my thesis, completed at Cornell University in 2003. I thank my advisor Karen Vogtmann for all of the time and support she gave me during my time at Cornell. These ideas would not have been possible without the many useful conversations we had about this material. I also thank the referee for suggestions on filling gaps in several proofs and for advice in improving the exposition of the paper.

References

[Bowditch and Epstein 1988] B. H. Bowditch and D. B. A. Epstein, "Natural triangulations associated to a surface", *Topology* **27**:1 (1988), 91–117. MR 89e:57004 Zbl 0649.32017

[Brown 1982] K. S. Brown, *Cohomology of groups*, Graduate Texts in Mathematics **87**, Springer, New York, 1982. MR 83k:20002 Zbl 0584.20036

[Culler and Vogtmann 1986] M. Culler and K. Vogtmann, "Moduli of graphs and automorphisms of free groups", *Invent. Math.* **84**:1 (1986), 91–119. MR 87f:20048 Zbl 0589.20022

[Harer 1985] J. L. Harer, "Stability of the homology of the mapping class groups of orientable surfaces", *Ann. of Math.* (2) **121**:2 (1985), 215–249. MR 87f:57009 Zbl 0579.57005

[Harer 1986] J. L. Harer, "The virtual cohomological dimension of the mapping class group of an orientable surface", *Invent. Math.* **84**:1 (1986), 157–176. MR 87c:32030 Zbl 0592.57009

[Harer 1993] J. L. Harer, "Improved homology stability for the homology of the mapping class groups of surfaces", preprint M-93-09, Duke University, 1993, online at http://www.math.duke.edu/preprints/93-09.dvi.

[Hatcher 2002] A. Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR 2002k:55001 Zbl 1033.55001

[Hatcher and Vogtmann 2004] A. Hatcher and K. Vogtmann, "Homology stability for outer automorphism groups of free groups", preprint, 2004. math.GT/0406377

[Hoare 1979] A. H. M. Hoare, "Coinitial graphs and Whitehead automorphisms", *Canad. J. Math.* **31**:1 (1979), 112–123. MR 80e:20050 Zbl 0412.20030

[Ivanov 1989] N. V. Ivanov, "Stabilization of the homology of Teichmüller modular groups", *Algebra i Analiz* **1**:3 (1989), 110–126. Translated in *Leningrad Mathematical J.* **1**:3 (1990), 675–691. MR 91g:57010 Zbl 0727.30036

[Madsen and Weiss 2002] I. Madsen and M. S. Weiss, "The stable moduli space of Riemann surfaces: Mumford's conjecture", preprint, 2002. math.AT/0212321

[Magnus et al. 1966] W. Magnus, A. Karrass, and D. Solitar, *Combinatorial group theory*, Pure and applied mathematics **13**, Interscience Publishers, New York, 1966. Reprinted by Dover, New York, 1976. MR 54 #10423 Zbl 0138.25604

[McCool 1975] J. McCool, "Some finitely presented subgroups of the automorphism group of a free group", *J. Algebra* **35** (1975), 205–213. MR 53 #624 Zbl 0325.20025

[Mulase and Penkava 1998] M. Mulase and M. Penkava, "Ribbon graphs, quadratic differentials on Riemann surfaces, and algebraic curves defined over $\overline{\mathbf{Q}}$", *Asian J. Math.* **2**:4 (1998), 875–919. MR 2001g:30028 Zbl 0964.30023

[Penner 1987] R. C. Penner, "The decorated Teichmüller space of punctured surfaces", *Comm. Math. Phys.* **113**:2 (1987), 299–339. MR 89h:32044 Zbl 0642.32012

[Quillen 1973] D. Quillen, “Higher algebraic K-theory. I”, pp. 85–147 in *Algebraic K-theory, vol. I, Higher K-theories* (Seattle, WA, 1972), edited by H. Bass, Lecture Notes in Math. **341**, Springer, Berlin, 1973. MR 49 #2895 Zbl 0292.18004

[Vogtmann 2002] K. Vogtmann, “Automorphisms of free groups and outer space”, *Geom. Dedicata* **94** (2002), 1–31. MR 2004b:20060 Zbl 1017.20035

[Zieschang et al. 1980] H. Zieschang, E. Vogt, and H.-D. Coldewey, *Surfaces and planar discontinuous groups*, Lecture Notes in Mathematics **835**, Springer, Berlin, 1980. MR 82h:57002 Zbl 0438.57001

Received July 2, 2004.

MATTHEW HORAK
DEPARTMENT OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE
HARVEY HALL 237E
UNIVERSITY OF WISCONSIN STOUT
MENOMONIE, WI 54751
UNITED STATES

horakm@uwstout.edu
http://faculty.uwstout.edu/horakm

PACIFIC JOURNAL OF MATHEMATICS
Vol. 227, No. 1, 2006

AN OPTIMAL SYSTOLIC INEQUALITY FOR CAT(0) METRICS IN GENUS TWO

MIKHAIL G. KATZ AND STÉPHANE SABOURAU

We prove an optimal systolic inequality for CAT(0) metrics on a genus 2 surface. We use a Voronoi cell technique, introduced by C. Bavard in the hyperbolic context. The equality is saturated by a flat singular metric in the conformal class defined by the smooth completion of the curve $y^2 = x^5 - x$. Thus, among all CAT(0) metrics, the one with the best systolic ratio is composed of six flat regular octagons centered at the Weierstrass points of the Bolza surface.

1. Hyperelliptic surfaces of nonpositive curvature

Over half a century ago, a student of C. Loewner's named P. Pu [1952] presented in this journal the first two optimal systolic inequalities, which came to be known as the Loewner inequality for the torus and Pu's inequality for the real projective plane. (See (5–2) on page 104 for the latter.)

The last couple of years have seen the discovery of a number of new systolic inequalities [Ammann 2004; Bangert and Katz 2003; 2004; Bangert et al. 2005; 2006a; 2006b; Ivanov and Katz 2004; Katz 2006; Katz and Lescop 2005; Katz and Sabourau 2006; Katz et al. 2006; Sabourau 2004], as well as near-optimal asymptotic bounds [Hamilton 2005; Katz 2003; Katz and Sabourau 2005; Katz et al. 2005; Rudyak and Sabourau ≥ 2006; Sabourau 2006; ≥ 2006]. A number of questions posed in [Croke and Katz 2003] have thus been answered. A general framework for systolic geometry in a topological context is proposed in [Katz and Rudyak 2005; 2006]. See [Katz ≥ 2006] for an overview of systolic problems. The homotopy 1-systole, denoted $\mathrm{sys}\pi_1(X)$, of a compact metric space X is the least length of a noncontractible loop of X.

Given a metric $\mathcal{G}$ on a surface, let $\mathrm{SR}(\mathcal{G})$ denote its systolic ratio

$$\mathrm{SR}(\mathcal{G}) = \frac{\mathrm{sys}\pi_1(\mathcal{G})^2}{\mathrm{area}(\mathcal{G})}.$$

MSC2000: 53C20, 53C23.

Keywords: Bolza surface, CAT(0) space, hyperelliptic surface, Voronoi cell, Weierstrass point, systole.

Katz was supported by the Israel Science Foundation (grants no. 620/00-10.0 and 84/03).

The *optimal systolic ratio* of a compact Riemann surface Σ is defined as $\mathrm{SR}(\Sigma) = \sup_{\mathcal{G}} \mathrm{SR}(\mathcal{G})$, where the supremum is over all metrics in the conformal type of Σ. Finally, given a smooth compact surface M, its optimal systolic ratio is defined by setting $\mathrm{SR}(M) = \sup_{\Sigma} \mathrm{SR}(\Sigma)$, where the supremum is over all conformal structures Σ on M. The latter ratio is known for the Klein bottle — see the bound (5–1) on page 104 — in addition to the torus and real projective plane already mentioned.

In the class of all metrics without any curvature restrictions, no singular flat metric on a surface of genus 2 can give the optimal systolic ratio in this genus [Sabourau 2004]. The best available upper bound for the systolic ratio of an arbitrary genus 2 surface is $\gamma_2 \simeq 1.1547$ [Katz and Sabourau 2006].

The precise value of SR for the genus 2 surface has so far eluded researchers [Calabi 1996; Bryant 1996]. We propose an answer in the framework of negatively curved, or more generally, CAT(0) metrics.

The term "CAT(0) space" evokes an extension of the notion of a manifold of nonpositive curvature to encompass singular spaces. We will use the term to refer to surfaces with metrics with only mild quotient singularities, defined below. Here the condition of nonpositive curvature translates into a lower bound of 2π for the total angle at the singularity. We need such an extension so as to encompass the metric that saturates our optimal inequality (1–1).

A mild quotient singularity is defined as follows. Consider a smooth metric on $\mathbb{R}^2$. Let $q \geq 1$ be an integer. Consider the q-fold cover X_q of $\mathbb{R}^2 \setminus \{0\}$ with the induced metric. We compactify X_q in the neighborhood of the origin to obtain a complete metric space $X_q^c = X_q \cup \{0\}$.

Definition 1.1. Suppose X_q^c admits an isometric action of $\mathbb{Z}_p$ fixing the origin. Then we can form the orbit space $Y_{p,q} = X_q^c/\mathbb{Z}_p$. The space $Y_{p,q}$ is then called mildly singular at the origin.

The total angle at the singularity is then $2\pi q/p$, and the CAT(0) condition is $q/p \geq 1$.

Remark 1.2. Alternatively, a point is singular of total angle $2\pi(1+\beta)$ if the metric is of the form $e^{h(z)}|z|^{2\beta}|dz|^2$ in its neighborhood, where $|dz|^2 = dx^2 + dy^2$. See [Troyanov 1990, p. 915].

Theorem 1.3. *Every* CAT(0) *metric* $\mathcal{G}$ *on a surface* Σ_2 *of genus* 2 *satisfies the optimal inequality*

$$\mathrm{SR}(\Sigma_2, \mathcal{G}) \leq \frac{1}{3}\cot\frac{\pi}{8} = \frac{1}{3}\bigl(\sqrt{2}+1\bigr) = 0.8047\ldots \tag{1–1}$$

The inequality is saturated by a singular flat metric, *with* 16 *conical singularities*, *in the conformal class of the Bolza surface.*

The Bolza surface is described in Section 2. The optimal metric is described in more detail in Section 3. Theorem 1.3 is proved in Section 4 based on the octahedral triangulation of S^2.

Remark 1.4. A similar optimal inequality can be proved for hyperelliptic surfaces of genus 5 based on the icosahedral triangulation [Bavard 1986].

2. Distinguishing 16 points on the Bolza surface

The Bolza surface $\mathcal{B}$ is the smooth completion of the affine algebraic curve

$$y^2 = x^5 - x. \tag{2–1}$$

It is the unique Riemann surface of genus 2 with a group of holomorphic automorphisms of order 48. (A way of passing from an affine hyperelliptic surface to its smooth completion is described in [Miranda 1995, p. 60–61].)

Definition 2.1. A conformal involution J of a compact Riemann surface Σ of genus g is called *hyperelliptic* if J has precisely $2g+2$ fixed points. The fixed points of J are called the Weierstrass points of Σ.

The quotient Riemann surface Σ/J is then necessarily the Riemann sphere, denoted henceforth S^2. Let $Q : \Sigma \to S^2$ be the conformal ramified double cover, with $2g+2$ branch points. Thus, J acts on Σ by sheet interchange. Recall that every surface of genus 2 is hyperelliptic, that is, admits a hyperelliptic involution [Farkas and Kra 1992, Proposition III.7.2].

We make note of 16 special points on $\mathcal{B}$. We call a point *special* if it is a fixed point of an order 3 automorphism of $\mathcal{B}$.

Consider the regular octahedral triangulation of $S^2 = \mathbb{C} \cup \infty$. Its set of vertices is conformal to the set of roots of the polynomial $x^5 - x$ of formula (2–1), together with the unique point at infinity. Thus the six points in question can be thought of as the ramification points of the ramified conformal double cover $Q : \mathcal{B} \to S^2$, while the 16 special points of $\mathcal{B}$ project to the eight vertices of the cubic subdivision dual to the octahedral triangulation.

In other words, the x-coordinates of the ramification points are

$$\{0, \infty, 1, -1, i, -i\},$$

which stereographically correspond to the vertices of a regular inscribed octahedron. The conformal type therefore admits the symmetries of the cube. If one includes both the hyperelliptic involution and the real (antiholomorphic) involution of $\mathcal{B}$ corresponding to the complex conjugation $(x, y) \to (\bar{x}, \bar{y})$ of $\mathbb{C}^2$, one obtains the full symmetry group $\mathrm{Aut}(\mathcal{B})$, of order

$$|\mathrm{Aut}(\mathcal{B})| = 96; \tag{2–2}$$

see [Kuusalo and Näätänen 1995, p. 404] for more details.

Lemma 2.2. *The hyperbolic metric of* $\mathcal{B}$ *admits* 12 *systolic loops. The* 12 *loops are in one-to-one correspondence with the edges of the octahedral decomposition of* S^2*. The correspondence is given by taking the inverse image under* Q *of an edge. The* 12 *systolic loops cut the surface into* 16 *hyperbolic triangles. The centers of the triangles are the* 16 *special points.*

See [Schmutz 1993, §5] for further details. The Bolza surface is extremal for two distinct problems:

- systole of hyperbolic surfaces [Bavard 1992; Schmutz 1993, Theorem 5.2];
- conformal systole of Riemann surfaces [Buser and Sarnak 1994].

The square of the conformal systole of a Riemann surface is also known as its Seshadri constant [Kong 2003]. The Bolza surface is also conjectured to be extremal for the first eigenvalue of the Laplacian. Such extremality has been verified numerically [Jakobson et al. 2005]. The evidence above suggests that the systolically extremal surface may lie in the conformal class of $\mathcal{B}$, as well. Meanwhile, we have the following result, proved in Section 5.

Theorem 2.3. *The Bolza surface* $\mathcal{B}$ *satisfies* $\mathrm{SR}(\mathcal{B}) \le \frac{\pi}{3}$.

Note that Theorems 2.3 and 1.3 imply that $\mathrm{SR}(\mathcal{B}) \in [0.8, 1.05]$.

3. A flat singular metric in genus two

The optimal systolic ratio of a genus 2 surface $(\Sigma_2, \mathcal{G})$ is unknown, but it satisfies the Loewner inequality [Katz and Sabourau 2006]. Here we discuss a *lower* bound for the optimal systolic ratio in genus 2, briefly described in [Croke and Katz 2003].

The example of M. Berger (see [Gromov 1983, Example 5.6.B′]) in genus 2 is a singular flat metric with conical singularities. It has systolic ratio $\mathrm{SR} = 0.6666$. This ratio was improved by F. Jenni [Jenni 1984], who identified the hyperbolic genus 2 surface with the optimal systolic ratio among all hyperbolic genus 2 surfaces (see also C. Bavard [Bavard 1992] and P. Schmutz [Schmutz 1993, Theorem 5.2]). The surface in question is a (2,3,8) triangle surface. Its conformal class is that of the Bolza surface (Section 2). It admits a regular hyperbolic octagon as a fundamental domain, and has 12 systolic loops of length $2x$, where $x = \cosh^{-1}(1+\sqrt{2})$. It has

$$\mathrm{sys}\pi_1 = 2\log\bigl(1+\sqrt{2}+\sqrt{2+2\sqrt{2}}\,\bigr),$$

area 4π, and systolic ratio $\mathrm{SR} \simeq 0.7437$. This ratio can be improved to 0.8047, as we shall see. The history for genus 2 so far can be summarized as follows:

$$\mathrm{SR}(\mathcal{G}) = \frac{\operatorname{sys}\pi_1(\mathcal{G})^2}{\operatorname{area}(\mathcal{G})} = \begin{cases} 0.6666 & \text{(Berger)} \\ 0.7437 & \text{(Jenni)} \\ 0.8047 & \text{(our metric } \mathcal{G}_{\mathbb{O}} \text{ on Bolza surface)} \end{cases}$$

Proposition 3.1. *The conformal class of the Bolza surface $\mathcal{B}$ admits a metric, denoted $\mathcal{G}_{\mathbb{O}}$, with the following properties*:

(1) *the metric is singular flat, with conical singularities precisely at the* 16 *special points of* Section 2;

(2) *each singularity is of total angle $\frac{9}{4}\pi$, so that the metric $\mathcal{G}_{\mathbb{O}}$ is* CAT(0);

(3) *the metric is glued from six flat regular octagons, centered on the Weierstrass points, while the* 1*-skeleton projects under $Q : \mathcal{B} \to S^2$ to that of the dual cube in S^2*;

(4) *the systolic ratio equals* $\mathrm{SR}(\mathcal{G}_{\mathbb{O}}) = \frac{1}{3}\left(\sqrt{2}+1\right) > \frac{4}{5}$.

Proof. The octahedral triangulation of the sphere, discussed in Section 2, lifts to a triangulation of $\mathcal{B}$ consisting of 16 triangles, which we think of as being "equilateral". Here eight equilateral triangles are connected cyclically around each of the six Weierstrass vertices of the triangulation of $\mathcal{B}$.

We further subdivide each equilateral triangle into three isosceles triangles, with a common vertex at the center of the equilateral triangle. We equip each of the 48 isosceles triangles with a flat metric with obtuse angle $\frac{3}{4}\pi$.

Each of the six Weierstrass vertices of the original triangulation is a smooth point, since the total angle is eight times $\pi/4$. Each equilateral triangle possesses a singularity at the center with total angle $\frac{9}{4}\pi > 2\pi$. Alternatively, we can apply the Gauss–Bonnet formula $\sum_\sigma \alpha(\sigma) = 2g-2$ in genus 2, with 16 isometric singularities. Here the sum is over all singularities σ of a singular flat metric on a surface of genus g, where the cone angle at singularity σ is $2\pi(1+\alpha(\sigma))$. Since the metric $\mathcal{G}_{\mathbb{O}}$ is smooth at a Weierstrass point of $\mathcal{B}$, the metric has only 16 singularities, precisely at the special points of Section 2, proving items 1 and 2 of the proposition.

Let x denote the side length of the equilateral triangle. The barycentric subdivision of each equilateral triangle consists of six copies of a flat right angle triangle, denoted $\mathcal{R}$, with side $x/2$ and adjacent angle $\pi/8$. We thus obtain a decomposition of the metric $\mathcal{G}_{\mathbb{O}}$ into 96 copies of the triangle $\mathcal{R}$, which can be thought of as a fundamental domain for the action of $\mathrm{Aut}(\mathcal{B})$; see (2–2).

We have $\operatorname{sys}\pi_1(\mathcal{G}_{\mathbb{O}}) = 2x$ by Lemma 3.2, proving item 4 of the proposition. To prove item 3, note that the union of the 16 triangles $\mathcal{R}$ with a common Weierstrass

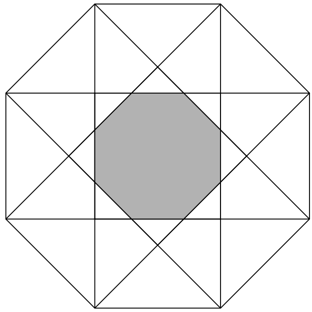


Figure 1. Flat regular octagon obtained as the union of 16 right triangles $\mathcal{R}$ with side $x/2$ and adjacent angle $\pi/8$. The shaded interior octagon represents the region with four geodesic loops through every point.

vertex is a flat regular octagon. The latter is represented in Figure 1, together with the systolic loops passing through it. □

Lemma 3.2. *The systole of the singular flat* CAT(0) *metric on the Bolza surface equals twice the distance between a pair of adjacent Weierstrass points.*

Proof. Consider the smooth closed geodesic $\gamma \subset \mathcal{B}$ that is the inverse image under the map $Q : \mathcal{B} \to S^2$ of an edge of the octahedron; see Lemma 2.2. Let x be the distance between a pair of opposite sides of the regular flat octagon in $\mathcal{B}$, or equivalently, the distance between a pair of adjacent Weierstrass points. Thus, $\text{length}(\gamma) = 2x$. Consider a loop $\alpha \subset \mathcal{B}$ whose length satisfies

$$\text{length}(\alpha) < 2x. \tag{3–1}$$

We will prove that there are two possibilities for α: it is either contractible, or freely homotopic to one of the 12 geodesics γ of the type described above. On the other hand, γ is necessarily length minimizing in its free homotopy class, by the CAT(0) property of the metric [Bridson and Haefliger 1999, Theorem 6.8]. This will rule out the second possibility, and prove the lemma.

Denote by $\mathcal{B}^{(1)}$ the graph on $\mathcal{B}$ given by the inverse image under Q of the 1-skeleton of the cubic subdivision of S^2. The graph $\mathcal{B}^{(1)}$ partitions the surface into six regular octagons, denoted Ω_k:

$$\mathcal{B} = \bigcup_{k=1}^{6} \Omega_k. \tag{3–2}$$

We will deform α to a loop $\beta \subset \mathcal{B}^{(1)}$ as follows. The partition (3–2) induces a partition of the loop α into arcs α_i, each lying in its respective octagon Ω_{k_i}. We deform each α_i, without increasing length, to the line segment $[p_i, q_i] \subset \Omega_{k_i}$. The boundary points p_i, q_i of α_i split the boundary closed curve $\partial\Omega_{k_i} \subset \mathcal{B}^{(1)}$ into a pair of paths. Let $\beta_i \subset \partial\Omega_{k_i}$ be the shorter of the two paths. Denote by y the distance between adjacent vertices of the octagon. Then clearly

$$\text{length}(\beta_i) \le \frac{4y}{x}\,\text{length}(\alpha_i). \tag{3–3}$$

We first deform the loop α into the graph $\mathcal{B}^{(1)}$. The deformation fixes the intersection points $\alpha \cap \mathcal{B}^{(1)}$. Inside Ω_{k_i} we deform the arc α_i to the path β_i. The length of the resulting loop is at most

$$\frac{4y}{x}\,\text{length}(\alpha) < 8y$$

by (3–1) and (3–3). Therefore, its homotopy class in $\mathcal{B}^{(1)}$ can be represented by an imbedded loop $\beta \subset \mathcal{B}^{(1)}$ of length at most $8y$. Thus, β contains fewer than eight edges of $\mathcal{B}^{(1)}$. Since the number of edges must be even, its image under Q must retract to a circuit with at most six edges in the 1-skeleton of the cubical subdivision of S^2. If the number is four, then the circuit lies in the boundary of a square face of the cube in S^2. But the boundary of a face does not lift to $\mathcal{B}$, since it surrounds a single ramification point, namely the center of the square face.

Hence there must be six edges in the circuit. There are two types of circuits with six edges in the 1-skeleton of the cubical subdivision of S^2:

(a) the boundary of the union of a pair of adjacent squares;

(b) a path consisting of the edges meeting a suitable great circle.

However, a path of type (b) surrounds an odd number, namely 3, of ramification points, and hence does not lift to the genus 2 surface. Meanwhile, a path of type (a) surrounds two ramification points, and hence does lift to the surface. Such a path is freely homotopic in $\mathcal{B}$ to one of the 12 geodesics of type γ (Lemma 2.2), completing the proof. □

4. Voronoi cells and Euler characteristic

The following proposition provides a preliminary lower bound on the area of hyperelliptic surfaces with nonpositive curvature.

Proposition 4.1. *Every J-invariant* CAT(0) *metric $\mathcal{G}$ on a closed hyperelliptic surface Σ_g of genus g satisfies the bound*

$$\text{SR}(\Sigma_g, \mathcal{G}) \le 8((g+1)\pi)^{-1}.$$

Proof. To prove this scale-invariant inequality, we normalize the metric on $\Sigma = \Sigma_g$ to unit systole, that is, $\text{sys}\pi_1(\Sigma, \mathcal{G}) = 1$. The preimage by $Q : \Sigma \to S^2$ of an arc of S^2 joining two distinct branch points forms a noncontractible loop on Σ. Therefore the distance between two Weierstrass points is at least $\frac{1}{2}\operatorname{sys}\pi_1(\Sigma, \mathcal{G}) = \frac{1}{2}$. Thus we obtain $2g+2$ disjoint disks of radius $\frac{1}{4}$, centered at the Weierstrass points. Since the metric is CAT(0), the area of each disk is at least $\frac{\pi}{16}$. Thus,

$$\operatorname{area}(\Sigma, \mathcal{G}) \geq \frac{g+1}{8}\pi. \qquad \square$$

An *optimal* lower bound requires a more precise estimate on the area of the Voronoi cells. The idea is to replace area of balls by area of polygons, where control over the number of sides is provided by the Euler characteristic [Bavard 1992].

Denote by $u : \tilde{\Sigma} \to \Sigma$ its universal cover. Let $\{x_i \mid i \in \mathbb{N}\}$ be an enumeration of the lifts of Weierstrass points on $\tilde{\Sigma}$. The Voronoi cell $V_i \subset \tilde{\Sigma}$ centered at x_i is defined as the set of points closer to x_i than to any other lift of a Weierstrass point. In formulas,

$$V_i = \left\{ x \in \tilde{\Sigma} \;\middle|\; d(x, x_i) \leq d(x, x_j) \text{ for every } j \neq i \right\}.$$

The Voronoi cells on $\tilde{\Sigma}$ are polygons whose edges are arcs of the equidistant curves between a pair of lifts of Weierstrass points. Note that these edges are not necessarily geodesics. The Voronoi cells on $\tilde{\Sigma}$ are topological disks, while their projections $u(V_i) \subset \Sigma$ may have more complicated topology. Thus, the surface Σ decomposes into $2g+2$ images of Voronoi cells, centered at the $2g+2$ Weierstrass points. By the number of sides of $u(V)$ we will mean the number of sides of the polygon V.

Lemma 4.2. *Let* $\mathcal{G}_1$ *and* $\mathcal{G}_2$ *be two* CAT(0) *metrics lying in the same conformal class. Then, the averaged metric* $\mathcal{G} = \frac{1}{2}(\mathcal{G}_1 + \mathcal{G}_2)$ *is* CAT(0), *as well.*

Proof. Choose a point $x \in \Sigma$ and a metric $\mathcal{G}_0$ in its conformal class, which is flat in a neighborhood of x. Every metric $\mathcal{G} = H\,\mathcal{G}_0$ conformal to it satisfies

$$K_{\mathcal{G}} H = K_{\mathcal{G}_0} - \tfrac{1}{2}\Delta \log H$$

(see [Gallot et al. 1990, p. 265]), where $K_{\mathcal{G}}$ and $K_{\mathcal{G}_0}$ are the Gaussian curvatures of $\mathcal{G}$ and $\mathcal{G}_0$, and Δ is the Laplacian of $\mathcal{G}_0$ with $\Delta f = \operatorname{div}\nabla f$. Thus, the metrics $\mathcal{G}_i$ can be written as $\mathcal{G}_i = e^{h_i}\,\mathcal{G}_0$, where h_i is subharmonic in the neighborhood of x, that is, $\Delta h_i \geq 0$. A simple computation shows that

$$\Delta \log H \geq \frac{e^{h_1}\Delta h_1 + e^{h_2}\Delta h_2}{2H} \geq 0,$$

where $H = \frac{1}{2}(e^{h_1} + e^{h_2})$, proving the lemma if both points are regular. For singular points with positive angle excess, the CAT(0) property for the averaged metric is immediate from Remark 1.2. $\square$

Proof of Theorem 1.3. Since averaging by J can only improve the systolic ratio, we may assume without loss of generality that our metric is already J-invariant.

There exists an extension of the notions of tangent plane and exponential map to surfaces with singularities. Namely, let $A \in \Sigma$. There exists a CAT(0) piecewise flat plane T_A with conical singularities and a covering $\exp_A : T_A \to \Sigma$ with the following properties:

(1) $\exp_A$ sends the origin O of T_A to A;

(2) $\exp_A$ takes the conical singularities of T_A to the singularities of Σ;

(3) $\exp_A$ sends every pair of geodesic arcs issuing from the origin $O \in T_A$ to a pair of geodesic arcs of the same lengths and forming the same angle at their basepoint A.

By the Rauch Comparison Theorem, the exponential map $\exp_A$ does not decrease distances.

Now assume $A \in \Sigma$ is a Weierstrass point, and let $B \in \Sigma$ be another Weierstrass point. Fix a lift B_0 of B to the tangent plane T_A, along a minimizing arc. Consider the equidistant line

$$L_{O,B_0} \subset T_A$$

between the origin $O \in T_A$ and the point B_0. Consider a point $X_0 \in L_{O,B_0}$. Let $X = \exp_A(X_0)$. Since the exponential map does not decrease distances, we have

$$\mathrm{dist}_\Sigma(A, X) = \mathrm{dist}_{T_A}(O, X_0) = \mathrm{dist}_{T_A}(B_0, X_0) \leq \mathrm{dist}_\Sigma(B, X).$$

Now consider the polygon in the tangent plane T_A, obtained as the intersection of the half-spaces containing the origin, defined by the lines L_{O,B_0}, as B runs over all Weierstrass points. It follows from the preceding equality that the exponential image of this polygon is contained in the Voronoi cell of A. Since the exponential map does not decrease distances, the area of the polygon is a lower bound for the area of the Voronoi cell. If k is the number of sides of V, then V is partitioned into k triangles with angle θ_i at O, whose area is bounded below by

$$\left(\frac{\mathrm{sys}\pi_1}{4}\right)^2 \tan\frac{\theta_i}{2} \tag{4–1}$$

since $\mathrm{dist}(A, B) \geq \frac{1}{2}\,\mathrm{sys}\pi_1$.

Consider the graph on S^2 defined by the projections of the Voronoi cells to the sphere. Thus we have $f = 6$ faces. Applying the formula $v - e + f = 2$ and the well-known fact that $3v \leq 2e$, we obtain

$$e \leq 3f - 6 = 12.$$

Hence the spherical graph has at most 12 edges. Note that the maximum is attained by the 1-skeleton of the cubical subdivision.

The area of a flat isosceles triangle with third angle θ and with unit altitude from the third vertex is $\tan(\theta/2)$. This formula provides a lower bound for the area of the Voronoi cells as in (4–1). The proof is completed by Jensen's inequality applied to the convex function $\tan(x/2)$ when $0 < x < \pi$. In the boundary case of equality, we have $e = 12$, all angles θ_i as in (4–1) must be equal, curvature must be zero because of equality in the Rauch Comparison Theorem, and we easily deduce that each Voronoi cell is a regular octagon. To minimize the area of the octagon, we must choose θ as small as possible. The CAT(0) hypothesis at the center of the octagon imposes a lower bound $\theta \geq \frac{\pi}{4}$. Hence the optimal systolic ratio is achieved for the regular flat octagon with a smooth point at the center. □

5. Arbitrary metrics on the Bolza surface

The conformal class of the Bolza surface $\mathcal{B}$ of Section 2 is likely to contain a systolically optimal surface in genus 2, as discussed in Section 1.

Theorem 5.1. *Every metric $\mathcal{G}$ in the conformal class of the Bolza surface satisfies the bound*

$$\mathrm{SR}(\mathcal{G}) \leq \frac{\pi}{3} = 1.0471\ldots.$$

Remark 5.2. In particular, every metric $\mathcal{G}$ in the conformal class of the Bolza surface satisfies Bavard's inequality

$$\mathrm{SR}(\mathcal{G}) \leq \frac{\pi}{2^{3/2}} \simeq 1.1107 \tag{5–1}$$

for the Klein bottle [Bavard 1986]. This suggests a possible monotonicity of $\chi(\Sigma)$ as a function of $\mathrm{SR}(\Sigma)$.

Lemma 5.3. *Let $\mathcal{G}$ be an* $\mathrm{Aut}(\mathcal{B})$*-invariant metric on $\mathcal{B}$. Let*

$$\delta(J) = \min_{x \in \mathcal{B}^{(1)}} \mathrm{dist}(x, J(x))$$

be the displacement on the 1-skeleton $\mathcal{B}^{(1)}$ of the Voronoi subdivision of $\mathcal{B}$. Then

$$\mathrm{area}(\mathcal{B}, \mathcal{G}) \geq 6\left(\tfrac{2}{\pi}\right)\delta(J)^2.$$

Proof. Consider the Voronoi subdivision with respect to the set of six Weierstrass points on $\mathcal{B}$. Since each Voronoi cell $\Omega \subset \mathcal{B}$ is J-invariant, we can identify all pairs of opposite points of the boundary $\partial\Omega$, to obtain a projective plane

$$\mathbb{RP}^2 = \Omega/\sim.$$

We now apply Pu's inequality [Pu 1952] to each of the six Voronoi cells, to obtain

$$\mathrm{area}(\mathbb{RP}^2) \geq \frac{2}{\pi}\,\mathrm{sys}\pi_1(\mathbb{RP}^2)^2. \tag{5–2}$$

The lemma now follows from the bound $\operatorname{sys}\pi_1(\mathbb{RP}^2) \geq \delta(J)$. □

Lemma 5.4. *Every* $\operatorname{Aut}(\mathcal{B})$*-invariant metric* $\mathcal{G}$ *on* $\mathcal{B}$ *satisfies the bound*

$$2\delta(J) \geq \operatorname{sys}\pi_1(\mathcal{B}, \mathcal{G}).$$

Proof. Let $p, J(p) \in \partial\Omega$ satisfy $\operatorname{dist}(p, J(p)) = \delta(J)$. Let $\alpha \subset \Omega$ be a minimizing path joining p to $J(p)$. Let $\Omega' \subset \mathcal{B}$ be the adjacent Voronoi cell containing this pair of boundary points, and $r : \mathcal{B} \to \mathcal{B}$ the anticonformal involution that switches Ω and Ω', and fixes their common boundary. The loop $\alpha \cup r(\alpha)$ belongs to the free homotopy class of the noncontractible loop $\gamma \subset \mathcal{B}$ obtained as the inverse image under $Q : \mathcal{B} \to S^2$ of the edge of the octahedral decomposition of S^2 joining the images of the centers of Ω and Ω' (Lemma 2.2). Since the length of $\alpha \cup r(\alpha)$ is $2\delta(J)$, the lemma follows. □

Proof of Theorem 5.1. We may assume that the metric on $\mathcal{B}$ is $\operatorname{Aut}(\mathcal{B})$-invariant, since averaging the metric by a finite group of holomorphic and antiholomorphic diffeomorphisms can only improve the systolic ratio. We combine the inequalities of Lemmas 5.4 and 5.3 to prove the theorem. □

References

[Ammann 2004] B. Ammann, "Dirac eigenvalue estimates on two-tori", *J. Geom. Phys.* **51**:3 (2004), 372–386. MR 2005f:53068 Zbl 1075.53041

[Bangert and Katz 2003] V. Bangert and M. Katz, "Stable systolic inequalities and cohomology products", *Comm. Pure Appl. Math.* **56**:7 (2003), 979–997. MR 2004g:53047 Zbl 1038.53031

[Bangert and Katz 2004] V. Bangert and M. Katz, "An optimal Loewner-type systolic inequality and harmonic one-forms of constant norm", *Comm. Anal. Geom.* **12**:3 (2004), 703–732. MR 2005k:53035 Zbl 1068.53027

[Bangert et al. 2005] V. Bangert, C. Croke, S. Ivanov, and M. Katz, "Filling area conjecture and ovalless real hyperelliptic surfaces", *Geom. Funct. Anal.* **15**:3 (2005), 577–597. MR 2221144 Zbl 1082.53033

[Bangert et al. 2006a] V. Bangert, C. Croke, S. Ivanov, and M. Katz, "Boundary case of equality in optimal Loewner-type inequalities", *Trans. Amer. Math. Soc.* **358** (2006).

[Bangert et al. 2006b] V. Bangert, M. Katz, S. Shnider, and S. Weinberger, "E_7, Wirtinger inequalities, Cayley 4-form, and homotopy", preprint, 2006. math.DG/0608006

[Bavard 1986] C. Bavard, "Inégalité isosystolique pour la bouteille de Klein", *Math. Ann.* **274**:3 (1986), 439–441. MR 87i:53059 Zbl 0578.53032

[Bavard 1992] C. Bavard, "La systole des surfaces hyperelliptiques", preprint 71, Ec. Norm. Sup. Lyon, 1992.

[Bridson and Haefliger 1999] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Math. Wissenschaften **319**, Springer, Berlin, 1999. MR 2000k:53038 Zbl 0988.53001

[Bryant 1996] R. L. Bryant, "On extremals with prescribed Lagrangian densities", pp. 86–111 in *Manifolds and geometry* (Pisa, 1993), edited by P. de Bartolomeis et al., Sympos. Math. **36**, Cambridge Univ. Press, Cambridge, 1996. MR 99a:58043 Zbl 0866.53004

[Buser and Sarnak 1994] P. Buser and P. Sarnak, "On the period matrix of a Riemann surface of large genus", *Invent. Math.* **117**:1 (1994), 27–56. MR 95i:22018 Zbl 0814.14033

[Calabi 1996] E. Calabi, "Extremal isosystolic metrics for compact surfaces", pp. 167–204 in *Actes de la Table Ronde de Géométrie Différentielle* (Luminy, 1992), Soc. Math. France, Paris, 1996. MR 97k:53037 Zbl 0884.58029

[Croke and Katz 2003] C. B. Croke and M. Katz, "Universal volume bounds in Riemannian manifolds", pp. 109–137 in *Surveys in differential geometry* (Boston, 2002), vol. 8, edited by S. T. Yau, Int. Press, Somerville, MA, 2003. MR 2005d:53061 Zbl 1051.53026

[Farkas and Kra 1992] H. M. Farkas and I. Kra, *Riemann surfaces*, 2nd ed., Graduate Texts in Mathematics **71**, Springer, New York, 1992. MR 93a:30047 Zbl 0764.30001

[Gallot et al. 1990] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian geometry*, Universitext, Springer, Berlin, 1990. MR 91j:53001 Zbl 0716.53001

[Gromov 1983] M. Gromov, "Filling Riemannian manifolds", *J. Differential Geom.* **18**:1 (1983), 1–147. MR 85h:53029 Zbl 0515.53037

[Hamilton 2005] M. Hamilton, "On the conformal systoles of four-manifolds", preprint, 2005. math.DG/0512127

[Ivanov and Katz 2004] S. V. Ivanov and M. G. Katz, "Generalized degree and optimal Loewner-type inequalities", *Israel J. Math.* **141** (2004), 221–233. MR 2005k:53047 Zbl 1067.53031

[Jakobson et al. 2005] D. Jakobson, M. Levitin, N. Nadirashvili, N. Nigam, and I. Polterovich, "How large can the first eigenvalue be on a surface of genus two?", *Int. Math. Res. Not.* 63 (2005), 3967–3985. MR 2202582 Zbl 05017528

[Jenni 1984] F. Jenni, "Über den ersten Eigenwert des Laplace-Operators auf ausgewählten Beispielen kompakter Riemannscher Flächen", *Comment. Math. Helv.* **59**:2 (1984), 193–203. MR 85i:58118 Zbl 0541.30034

[Katz 2003] M. Katz, "Four-manifold systoles and surjectivity of period map", *Comment. Math. Helv.* **78**:4 (2003), 772–786. MR 2005d:53062 Zbl 1048.53031

[Katz 2006] M. Katz, "Systolic inequalities and Massey products in simply-connected manifolds", preprint, 2006. math.DG/0604012

[Katz $\geq$ 2006] M. Katz, *Systolic geometry and topology*, Mathematical Surveys and Monographs, Amer. Math. Soc., Providence, RI. To appear.

[Katz and Lescop 2005] M. G. Katz and C. Lescop, "Filling area conjecture, optimal systolic inequalities, and the fiber class in abelian covers", pp. 181–200 in *Geometry, spectral theory, groups, and dynamics* (Haifa, Israel, 2003/4), edited by M. Entov et al., Contemp. Math. **387**, Amer. Math. Soc., Providence, RI, 2005. MR 2006h:53030 Zbl 02236598

[Katz and Rudyak 2005] M. Katz and Y. Rudyak, "Bounding volume by systoles of 3-manifolds", preprint, 2005. math.DG/0504008

[Katz and Rudyak 2006] M. Katz and Y. Rudyak, "Lusternik–Schnirelmann category and systolic category of low dimensional manifolds", *Comm. Pure Appl. Math.* **59**:10 (2006), 1433–1456.

[Katz and Sabourau 2005] M. G. Katz and S. Sabourau, "Entropy of systolically extremal surfaces and asymptotic bounds", *Ergodic Theory Dynam. Systems* **25**:4 (2005), 1209–1220. MR 2006d:53038 Zbl 02234476

[Katz and Sabourau 2006] M. G. Katz and S. Sabourau, "Hyperelliptic surfaces are Loewner", *Proc. Amer. Math. Soc.* **134**:4 (2006), 1189–1195. MR MR2196056 Zbl 05007253

[Katz et al. 2005] M. Katz, M. Schaps, and U. Vishne, "Logarithmic growth of systole of arithmetic Riemann surfaces along congruence subgroups", preprint, 2005. math.DG/0505007

[Katz et al. 2006] M. Katz, Y. Rudyak, and S. Sabourau, "Systoles of 2-complexes, Reeb graph, and Grushko decomposition", *Int. Math. Res. Not.* (2006). To appear. math.DG/0602009

[Kong 2003] J. Kong, "Seshadri constants on Jacobian of curves", *Trans. Amer. Math. Soc.* **355**:8 (2003), 3175–3180. MR 2004f:14041 Zbl 1016.14014

[Kuusalo and Näätänen 1995] T. Kuusalo and M. Näätänen, "Geometric uniformization in genus 2", *Ann. Acad. Sci. Fenn. Ser. A I Math.* **20**:2 (1995), 401–418. MR 96h:30083 Zbl 0856.30031

[Miranda 1995] R. Miranda, *Algebraic curves and Riemann surfaces*, Graduate Studies in Mathematics **5**, American Mathematical Society, Providence, RI, 1995. MR 96f:14029 Zbl 0820.14022

[Pu 1952] P. M. Pu, "Some inequalities in certain nonorientable Riemannian manifolds", *Pacific J. Math.* **2** (1952), 55–71. MR 14,87e Zbl 0046.39902

[Rudyak and Sabourau ≥ 2006] Y. Rudyak and S. Sabourau, "Systolic invariants of groups and 2-complexes via Grushko decomposition".

[Sabourau 2004] S. Sabourau, "Systoles des surfaces plates singulières de genre deux", *Math. Z.* **247**:4 (2004), 693–709. MR 2005i:53049 Zbl 1074.53034

[Sabourau 2006] S. Sabourau, "Systolic volume and minimal entropy of aspherical manifolds", preprint, 2006. To appear in *J. Differential Geom.* math.DG/0603695

[Sabourau ≥ 2006] S. Sabourau, "Asymptotic bounds for separating systoles on surfaces". In preparation.

[Schmutz 1993] P. Schmutz, "Riemann surfaces with shortest geodesic of maximal length", *Geom. Funct. Anal.* **3**:6 (1993), 564–631. MR 95f:30060 Zbl 0810.53034

[Troyanov 1990] M. Troyanov, "Coordonnées polaires sur les surfaces riemanniennes singulières", *Ann. Inst. Fourier (Grenoble)* **40**:4 (1990), 913–937 (1991). MR 92d:30029 Zbl 0711.53017

Received October 13, 2004.

MIKHAIL G. KATZ
DEPARTMENT OF MATHEMATICS
BAR ILAN UNIVERSITY
RAMAT GAN 52900
ISRAEL

katzmik@math.biu.ac.il

STÉPHANE SABOURAU
LABORATOIRE DE MATHÉMATIQUES ET PHYSIQUE THÉORIQUE
UNIVERSITÉ DE TOURS
PARC DE GRANDMONT
37400 TOURS
FRANCE

sabourau@lmpt.univ-tours.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 227, No. 1, 2006

OPERATOR MULTIPLIERS

EDWARD KISSIN AND VICTOR S. SHULMAN

We introduce a noncommutative version of Schur multipliers relative to an operator ideal. In this setting the functions of two variables are replaced by elements from a tensor product of C*-algebras, and the measures (or spectral measures) by representations. For commutative C*-algebras this approach agrees with Birman and Solomyak's theory of double operator integrals. We study the dependence of the spaces of multipliers on the choice of representations and find that the question is closely related to Voiculescu and Arveson's theory of approximately equivalent representations. The space of multipliers universal with respect to the chosen measures is related to the Haagerup tensor product of the algebras.

1. Introduction

Let H and K be Hilbert spaces, let $B(H, K)$ be the Banach space of all bounded linear operators from H into K, and let $\mathfrak{S}_2(H, K)$ be the Hilbert space of Hilbert–Schmidt operators. Each symmetrically normed ideal I induces the norm $|\cdot|_I$ on $\mathscr{X}_I = I(H, K) \cap \mathfrak{S}_2(H, K)$. Let $\Phi : \varphi \to \Phi_\varphi$ be a map from a set G into the algebra $B(\mathfrak{S}_2(H, K))$ of all bounded operators on $\mathfrak{S}_2(H, K)$. If for some $\varphi \in G$, the operator Φ_φ preserves $\mathscr{X}_I$ and is bounded in $|\cdot|_I$, so that $|\Phi_\varphi(R)|_I \le C|R|_I$, for all $R \in \mathscr{X}_I$, then φ is called a *(Φ, I)-multiplier*. Below we consider some examples of (Φ, I)-multipliers with increasing generality.

Let X, Y be arbitrary sets, let $H = l_2(X)$, $K = l_2(Y)$, and let $B(X \times Y)$ be the set of all bounded complex-valued functions on $X \times Y$. Identify each T in $\mathfrak{S}_2(H, K)$ with the corresponding matrix $(t(x, y))$. For $\varphi \in B(X \times Y)$, set $S_\varphi(T) = (\varphi(x, y)t(x, y))$. Then $S : \varphi \mapsto S_\varphi$ is a map from $B(X \times Y)$ into $B(\mathfrak{S}_2(H, K))$, and we call (S, I)-multipliers *Schur I-multipliers*. It is not difficult to check that, at least for separable ideals I, they coincide with Schur I-multipliers as defined in [Bennett 1977] and, for $I = B(H, K)$, in [Pisier 2001].

More generally, for arbitrary measures μ on X and ν on Y, let $H = L_2(X, \mu)$ and $K = L_2(Y, \nu)$. Then $\mathfrak{S}_2(H, K)$ consists of integral operators R with kernels

MSC2000: primary 46L06, 47B49; secondary 47B10, 47L20.

Keywords: C*-algebra, representation, approximate equivalence, tensor product, Schur multiplier, operator ideal, ω-continuity.

$r(x, y)$ on $X \times Y$ such that

$$|R|_2 = \left(\int_{X \times Y} |r(x, y)|^2 \, d\mu(x) \, d\nu(y) \right)^{1/2} < \infty. \tag{1–1}$$

Each $\varphi \in L_\infty(X \times Y, \mu \times \nu)$ defines a bounded linear map Φ_φ on $\mathfrak{S}_2(H, K)$, where $\Phi_\varphi(R)$ is the integral operator with kernel $\varphi(x, y)r(x, y)$. The (Φ, I)-multipliers in this case are called *(μ, ν, I)-multipliers*.

Birman and Solomyak [1967; 1973; 1989] developed a powerful machinery of *double operator integrals* (DOI) in their study of multipliers related to various problems arising in mathematical physics. Starting with two spectral measures $\mathscr{E}$ and $\mathscr{F}$ on sets X and Y, respectively, they define, for each bounded measurable function φ, a map Φ_φ on $\mathfrak{S}_2(H, K)$ by

$$\Phi_\varphi(R) = \int_X \int_Y \varphi(x, y) \, d\mathscr{F}(y) R \, d\mathscr{E}(x).$$

The corresponding (Φ, I)-multipliers are called “functions that define bounded DOI on I”; we will call them $(\mathscr{E}, \mathscr{F}, I)$-multipliers or DOI I-multipliers. For multiplicity-free spectral measures, they coincide with (μ, ν, I)-multipliers.

We consider now a noncommutative version of the example above. In this setting the functions of two variables are replaced by elements of the tensor product $\mathscr{A} \otimes \mathscr{B}$ of C*-algebras, and the spectral measures by representations π, ρ of these algebras. For $\varphi \in \mathscr{A} \otimes \mathscr{B}$, the operator $(\pi \otimes \rho)(\varphi)$ acts on the tensor product $\mathscr{H} = H_\pi \otimes H_\rho$. Identifying $\mathscr{H}$ with $\mathfrak{S}_2(H_\pi^d, H_\rho)$, where H_π^d is the dual of H_π, we may consider this operator as an operator Φ_φ on $\mathfrak{S}_2(H_\pi^d, H_\rho)$ and, in the sense above, speak about I-multipliers. We call them $(\pi \otimes \rho, I)$-multipliers. For commutative C*-algebras $\mathscr{A} = C_0(X)$ and $\mathscr{B} = C_0(Y)$, the $(\pi \otimes \rho, I)$-multipliers coincide with $(\mathscr{E}, \mathscr{F}, I)$-multipliers, where $\mathscr{E}$ and $\mathscr{F}$ are the spectral measures corresponding to the representations π and ρ. Even for commutative algebras the precise description of the spaces of multipliers is known only for $I = B(H)$; for Schur multipliers it was obtained in [Grothendieck 1953], for DOI $B(H)$-multipliers in [Peller 1985].

In this paper we mainly study the dependence of the spaces of multipliers on the choice of the representations and, in the commutative case, on the choice of spectral or scalar measures. Our initial aim was to prove that a continuous function $\varphi(x, y)$ is a $(\mu, \nu, B(H))$-multiplier if and only if it is a Schur multiplier on the product of the supports of μ and ν. In other words, we were going to prove that the space of continuous $(\mu, \nu, B(H))$-multipliers depends only on the supports of the measures. This was conjectured by B. E. Johnson in a discussion with the second author and previously proved in [Kissin and Shulman 1996] for functions of the form $(f(x) - f(y))/(x - y)$. Here this result will be deduced from a result of much more general nature: the space of all (π, ρ, I)-multipliers does not change if the representations π and ρ are replaced by approximately equivalent representations. For its proof we use Voiculescu's noncommutative Weyl–von Neumann theorem. As far as we know, this is the first application of a deep result of the theory of C*-algebras to multipliers and, in particular, to Schur multipliers. In fact, the desire to understand the relation between these branches of the operator theory was our main motivation during this work.

The restriction to the C*-tensor products of C*-algebras reflects our interest in continuous multipliers. However, in the last sections we go further and study "non-continuous" multipliers. More precisely, we consider (μ, ν, I)-multipliers continuous in a pseudotopology instead of a topology. It was shown in [Erdos et al. 1998] that each pair μ and ν of standard measures on X and Y defines a pseudotopology ω on $X \times Y$, and we study (μ, ν, I)-multipliers that are ω-continuous functions. It should be noted that the space of such multipliers is much wider than $C_0(X \times Y)$ and, moreover, all $(\mu, \nu, B(H))$-multipliers are necessarily ω-continuous. The main result here states that an ω-continuous function is a (μ, ν, I)-multiplier if and only if it becomes a Schur multiplier after deleting from X and Y suitable null subsets. As a consequence we show how one can deduce Peller's theorem on DOI $B(H)$-multipliers from Grothendieck's description of Schur multipliers. We also prove that the space of all ω-continuous (μ, ν, I)-multipliers does not change if the measures μ and ν are replaced by equivalent measures.

2. Preliminaries

We need some notions and results from the theory of symmetrically normed (s.n.) ideals. The general reference for this topic is Gohberg and Kreĭn [1965]. We denote the dual space of a Banach space X by X^d and the conjugate operator of $A \in B(X)$ by A^d. Let $\mathcal{F}$ and $C(H)$ be the ideals of finite rank operators and of compact operators in the Banach algebra $(B(H), \|\cdot\|)$ of all bounded operators on a Hilbert space H. A two-sided ideal I of $B(H)$ is *symmetrically normed* if it is a Banach space with respect to a norm $|\cdot|_I$, and

$$|AXB|_I \leq \|A\| |X|_I \|B\|, \qquad \text{for } A, B \in B(H) \text{ and } X \in I.$$

Such an ideal I is selfadjoint and, by the Calkin theorem, $\mathcal{F} \subseteq I \subseteq C(H)$.

There is a one-to-one correspondence between the set of symmetrically normed functions (see [Gohberg and Kreĭn 1965]) on the space c_0 of all sequences of real numbers converging to 0 and the set $\mathfrak{J}$ of all pairs (J_0, J) of s.n. ideals, where J_0 is a separable ideal that coincides with the closure of $\mathcal{F}$ in $|\cdot|_{J_0}$, and J is the largest s.n. ideal such that $J_0 \subseteq J$ and the norms $|\cdot|_{J_0}$ and $|\cdot|_J$ coincide on J_0. We call J *coseparable* because there is another, "dual" pair $(\widehat{J_0}, \widehat{J})$ in $\mathfrak{J}$ such that J is isometrically isomorphic to the dual space of $\widehat{J_0}$ via the following correspondence:

every bounded linear functional on $\widehat{J_0}$ has the form

$$F_T(X) = (X, T)_2 = \mathrm{Tr}(T^* X), \qquad \text{with } T \in J \text{ and } \|F_T\| = |T|_J. \tag{2–1}$$

In turn, the ideal $\widehat{J}$ is isometrically isomorphic to the dual space of J_0.

Many ideals are separable and coseparable simultaneously. An important class of such ideals consists of the Schatten ideals $\mathfrak{S}_p$, with $1 \le p < \infty$. We will denote $C(H)$ by $\mathfrak{S}_\infty$ and $B(H)$ by $\mathfrak{S}_b$. The dual ideal $\widehat{\mathfrak{S}}_p$ of the Schatten ideal $\mathfrak{S}_p$ is $\mathfrak{S}_{p'}$, where

$$\frac{1}{p} + \frac{1}{p'} = 1 \text{ if } 1 < p < \infty; \qquad p' = 1 \text{ if } p = \infty; \qquad p' = b \text{ if } p = 1.$$

For each s.n. ideal I, there is a unique pair (J_0, J) in $\mathfrak{J}$ such that $J_0 \subseteq I \subseteq J$ and the norms $|\cdot|_{J_0}$, $|\cdot|_I$ and $|\cdot|_J$ coincide on J_0.

If I, J are s.n. ideals and $J \subseteq I$, Proposition 2.1 of [Kissin and Shulman 2005b] tells us that there is $c > 0$ such that

$$|A|_I \le c|A|_J, \qquad \text{for } A \in J. \tag{2–2}$$

For their dual spaces I^d, J^d, we have $I^d \subseteq J^d$ and

$$\|F\|_{J^d} \le c\|F\|_{I^d}, \qquad \text{for } F \in I^d. \tag{2–3}$$

Lemma 2.1. (i) *For coseparable ideals I, J, the following conditions are equivalent*:

$$1)\ I \subseteq J; \qquad 2)\ I_0 \subseteq J; \qquad 3)\ I_0 \subseteq J_0; \qquad 4)\ \widehat{J} \subseteq \widehat{I}.$$

In particular, the following conditions are equivalent:

$$1)\ \mathfrak{S}_2 \subseteq J_0; \qquad 2)\ \mathfrak{S}_2 \subseteq J; \qquad 3)\ \widehat{J} \subseteq \mathfrak{S}_2; \qquad 4)\ \widehat{J_0} \subseteq \mathfrak{S}_2.$$

(ii) *Let $J \subseteq I$ be s.n. ideals. If a bounded map $M\colon I \to I$ preserves J, its restriction to J is bounded.*

Proof. Part (i) follows from (2–1)–(2–3).

Let $A_n \to A$ and $M(A_n) \to B$ in $(J, |\cdot|_J)$. By (2–2), $|A_n - A|_I \to 0$, so $|M(A_n) - M(A)|_I \to 0$. Therefore,

$$\begin{aligned}|M(A) - B|_I &\le |M(A) - M(A_n)|_I + |M(A_n) - B|_I \\ &\le |M(A) - M(A_n)|_I + c|M(A_n) - B|_J \longrightarrow 0.\end{aligned}$$

Thus $M(A) = B$. Hence M is closed on J and, therefore, bounded. □

Let H, K be Hilbert spaces and I be an s.n. ideal of $B(H)$. Then

$$I(H, K) = \{A \in B(H, K) : (A^* A)^{1/2} \in I\}$$

is a closed left $B(K)$- and right $B(H)$-module supplied with norm

$$|A|_I = \left|(A^*A)^{1/2}\right|_I.$$

If U is an isometry from H onto K, then $I(H, K) = UI$. For $R \in B(H_1, H)$, $S \in B(K, K_1)$, and $A \in I(H, K)$,

$$SAR \in I(H_1, K_1) \quad \text{and} \quad |SAR|_I \leq \|S\|\,|A|_I\,\|R\|.$$

If S and R are isometries, then $|SAR|_I = |A|_I$.

The dual space H^d of H is a Hilbert space; there is an antiisometry map ∂ from H onto H^d, where $x^d := \partial(x)$ is given by $x^d(y) = (y, x) = (x^d, y^d)$ [Wegge-Olsen 1993]. The space $\mathfrak{S}_2(H^d, K)$, being a Hilbert space with respect to the scalar product $(T, R) = Tr(R^*T)$, is isometrically isomorphic to the tensor product space $H \otimes K$. More precisely, the linear map θ from the algebraic tensor product $H \odot K$ into the set of all finite rank operators in $B(H^d, K)$ defined by

$$\theta(h \otimes k)x^d = x^d(h)k = (h, x)k \quad \text{for } x \in H,$$

extends to an isometric isomorphism from $H \otimes K$ to $\mathfrak{S}_2(H^d, K)$:

$$(\theta(\xi), \theta(\eta))_2 = \mathrm{Tr}(\theta(\eta)^*\theta(\xi)) = (\xi, \eta) \quad \text{for } \xi, \eta \in H \otimes K.$$

Let θ_1 be the isomorphism from $H_1 \otimes K_1$ on $\mathfrak{S}_2(H_1^d, K_1)$. For $R \in B(H, H_1)$, denote by R^* its adjoint, acting from H_1 to H, and by R^d its conjugate from H_1^d to H^d. Then

$$\|R^d\| = \|R\|, \quad R^d x^d = (R^*x)^d \quad \text{for } x \in H_1,$$

and

$$(RT)^d = T^d R^d, \quad (R^*)^d = (R^d)^*, \quad (\lambda R)^d = \lambda R^d \ \text{ for } \lambda \in C. \tag{2–4}$$

The second of these equalities can be written in the form

$$R^d = \partial R^* \partial_1^{-1}. \tag{2–5}$$

Let $S \in B(K, K_1)$. We have $\theta_1\big((R \otimes S)(h \otimes k)\big) = S\theta(h \otimes k)R^d$ for $h \in H$ and $k \in K$, so

$$\theta_1((R \otimes S)\xi) = S\theta(\xi)R^d, \quad \text{for } \xi \in H \otimes K. \tag{2–6}$$

3. Multipliers and approximate equivalence

A *normed subspace* X of a Hilbert space H is a linear subspace supplied with its own norm $\|\cdot\|_X$. By $b_1(X)$ we denote the closed unit ball of $(X, \|\cdot\|_X)$. As important classes of normed subspaces we mention full and normal subspaces:

(1) X is *full* if $X = H$, while $\|\cdot\|_X$ need not coincide with the norm of H;

(2) X is *normal* if $b_1(X)$ is closed in H.

We define the dual normed subspace $X^\natural$ of a normed subspace X in H by setting

$$X^\natural = \{y \in H \colon \|y\|_{X^\natural} < \infty\}, \qquad \text{where } \|y\|_{X^\natural} = \sup_{x \in X} \frac{|(x, y)|}{\|x\|_X}. \tag{3–1}$$

Then $X \subseteq X^{\natural\natural}$ and $\|x\|_{X^{\natural\natural}} \le \|x\|_X$, for $x \in X$. Thus $b_1(X) \subseteq b_1(X^{\natural\natural})$.

Proposition 3.1. (i) *For a normed subspace X of H the following conditions are equivalent*:

(1) *X is normal*;

(2) *X is a dual of some normed subspace*;

(3) $b_1(X) = b_1(X^{\natural\natural})$.

(ii) *Let X be full. It is normal if and only if*

$$\|x\|_X \le C\|x\|, \qquad \textit{for all } x \in H \textit{ and some } C > 0. \tag{3–2}$$

Proof. (3) $\Rightarrow$ (2). Set $Y = X^\natural$. 2) $\Rightarrow$ 1) follows from (3–1).

(1) $\Rightarrow$ (3). Let $z \in b_1(X^{\natural\natural}) \setminus b_1(X)$. Since $b_1(X)$ is closed in H, then, by the Hahn–Banach theorem, there is $y \in H$ such that $|(z, y)| > 1$ and $|(x, y)| \le 1$, for all $x \in b_1(X)$. Therefore, by (3–1), $y \in X^\natural$ and $\|y\|_{X^\natural} \le 1$, so $\|z\|_{X^{\natural\natural}} > 1$. This contradiction shows that $b_1(X) = b_1(X^{\natural\natural})$. Part (i) is proved.

Let X be full: $H = X = \bigcup_{n=1}^{\infty} n\, b_1(X)$. If X is normal, then $b_1(X)$ is closed. By Baire's theorem, $b_1(X)$ contains an open subset of H and this implies (3–2). The converse is evident. □

Let X be a normed subspace of H. An operator $M \in B(H)$ is called *bounded on the pair* (X, H), if it preserves X and is bounded on X in $\|\cdot\|_X$. The proof of the following result is straightforward.

Lemma 3.2. *Let X be a normed subspace of H, and let M be bounded on* (X, H).

(i) *M^* is bounded on $(X^\natural, H)$ and* $\|M^*\|_{B(X^\natural)} \le \|M\|_{B(X)}$.

(ii) *M is bounded on* $(X^{\natural\natural}, H)$.

(iii) *If X is normal, then* $\|M^*\|_{B(X^\natural)} = \|M\|_{B(X)}$.

Let π be a representation of a C*-algebra $\mathscr{A}$ on H and X be a normed subspace of H. An element $a \in \mathscr{A}$ is a (π, X)-*multiplier* if $\pi(a)$ is bounded on (X, H). Set

$$|a|_X^\pi = \|\pi(a)\|_{B(X)} = \sup_{x \in X} \frac{\|\pi(a)x\|_X}{\|x\|_X}. \tag{3–3}$$

We now recall the notions of approximate equivalence and approximate subordination for representations of C*-algebras, introduced in [Voiculescu 1976] (see also [Arveson 1974]) and [Hadwin 1981], respectively. Among the various possible definitions, we use the one given in this last reference.

Definition 3.3. (i) Let π and π' be *-representations of a C*-algebra $\mathcal{A}$ on Hilbert spaces H and H'. The representation π' is approximately subordinate to π (we write $\pi' \ll_a \pi$) if there is a net $\{U_\lambda\}$ of isometries from H' into H such that

$$\|\pi(a)U_\lambda - U_\lambda\pi'(a)\| \to 0, \qquad \text{for all } a \in \mathcal{A}. \tag{3–4}$$

(ii) If the operators U_λ are unitary, then π and π' are approximately equivalent, and we write $\pi \sim_a \pi'$.

Let X' and X be normed subspaces of H' and H, respectively. We say that an approximate subordination $\pi' \ll_a \pi$ or approximate equivalence $\pi' \sim_a \pi$) is (X', X)-consistent if the operators U_λ in Equation (3–3) can be chosen in such a way that

$$U_\lambda X' \subseteq X, \quad U_\lambda^* X \subseteq X', \quad \|U_\lambda x'\|_X \le C\|x'\|_{X'}, \quad \|U_\lambda^* x\|_{X'} \le C\|x\|_X, \tag{3–5}$$

for some $C > 0$ and all $x' \in X'$, $x \in X$.

Proposition 3.4. *Let π' and π be *-representations of $\mathcal{A}$ on H' and H, let X' and X be normed subspaces of H' and H, and let there exist an (X', X)-consistent approximate subordination $\pi' \ll_a \pi$. Suppose that X' is normal. Then any (π, X)-multiplier a in $\mathcal{A}$ is also a (π', X')-multiplier, and*

$$|a|_{X'}^{\pi'} \le C^2 |a|_X^\pi . \tag{3–6}$$

Proof. Set $F_\lambda = \pi(a)U_\lambda - U_\lambda\pi'(a)$. Given $x' \in X'$, we have $U_\lambda^*\pi(a)U_\lambda x' \in X'$,

$$\|U_\lambda^*\pi(a)U_\lambda x'\|_{X'} \le C\|\pi(a)U_\lambda x'\|_X \le C|a|_X^\pi \|U_\lambda x'\|_X \le C^2|a|_X^\pi \|x'\|_{X'},$$

and $\pi'(a)x' = U_\lambda^* U_\lambda \pi'(a)x' = U_\lambda^*\pi(a)U_\lambda x' - U_\lambda^* F_\lambda x'$.

Set $C_1 = C^2|a|_X^\pi \|x'\|_{X'}$. Then all $U_\lambda^*\pi(a)U_\lambda x'$ belong to $C_1 b_1(X')$. Since X' is normal, the ball $C_1 b_1(X')$ is closed in H. Since $\|U_\lambda^* F_\lambda x'\|_{H'} \to 0$, we obtain that $\pi'(a)x' \in C_1 b_1(X')$. Thus $\pi'(a)$ preserves X', and

$$\|\pi'(a)x'\|_{X'} \le C_1 = C^2|a|_X^\pi \|x'\|_{X'},$$

which gives (3–6). □

4. Operators bounded on normed subspaces of $\mathfrak{S}_2$

For an s.n. ideal I, define a normed subspace $\mathcal{X}_I$ of $\mathfrak{S}_2(H, K)$ by setting

$$\mathcal{X}_I = I(H, K) \cap \mathfrak{S}_2(H, K), \qquad \text{with } \|X\|_{\mathcal{X}_I} = |X|_I \text{ for } X \in \mathcal{X}_I.$$

As in general (see Section 2), set

$$\mathscr{X}_I^{\sim} = \{T \in \mathfrak{S}_2(H, K)\colon \text{the map } X \to (X, T)_2 = \mathrm{Tr}(T^* X) \text{ is bounded on } \mathscr{X}_I\}.$$

Let a pair (J_0, J) in $\mathfrak{J}$ be such that $J_0 \subseteq I \subseteq J$ and the norms $|\cdot|_{J_0}$ and $|\cdot|_I$ coincide on J_0. If $\mathfrak{S}_2 \subseteq I$, then, by Lemma 2.1(i), $\mathfrak{S}_2 \subseteq J_0$, so $\mathscr{X}_I = \mathscr{X}_{J_0}$. Let $(\widehat{J_0}, \widehat{J})$ be the corresponding "dual" pair.

Lemma 4.1. (i) $(\mathscr{X}_{J_0})^\natural = \mathscr{X}_{\widehat{J}}$.

(ii) *If* $\mathfrak{S}_2 \subseteq I$ *or if* I *is coseparable* $(I = J)$, *then the normed space* $\mathscr{X}_I$ *is normal.*

(iii) *If* $J \subseteq \mathfrak{S}_2$, *then* $(\mathscr{X}_J)^\natural = (\mathscr{X}_I)^\natural = (\mathscr{X}_{J_0})^\natural = \mathscr{X}_{\widehat{J}} = \mathscr{X}_{\widehat{J_0}}$ *and* $(\mathscr{X}_I)^{\natural\natural} = \mathscr{X}_J$.

Proof. Since J_0 is separable and the space $\mathscr{F}(H, K)$ of all finite rank operators from H into K lies in $\mathscr{X}_{J_0}$, we see that $\mathscr{X}_{J_0}$ is dense in $J_0(H, K)$. From this, from (2–1) and (3–1) we obtain $(\mathscr{X}_{J_0})^\natural = \mathscr{X}_{\widehat{J}}$. Part (i) is proved.

If $\mathfrak{S}^2 \subseteq I$, then $\mathscr{X}_I$ is full. By (2–2) and Proposition 3.1, $\mathscr{X}_I$ is normal.

Let $I = J$. By (i), $(\mathscr{X}_{\widehat{J_0}})^\natural = \mathscr{X}_J$. Thus, by Proposition 3.1, $\mathscr{X}_I$ is normal. Part (ii) is proved.

Let $I \subseteq \mathfrak{S}_2$. By Lemma 2.1, $J \subseteq \mathfrak{S}_2$, so that $\mathscr{X}_{J_0} \subseteq \mathscr{X}_I \subseteq \mathscr{X}_J$. It follows from (2–2) that

$$(\mathscr{X}_J)^\natural \subseteq (\mathscr{X}_I)^\natural \subseteq (\mathscr{X}_{J_0})^\natural.$$

By (i), $(\mathscr{X}_{J_0})^\natural = \mathscr{X}_{\widehat{J}}$. By Lemma 2.1, $\mathfrak{S}_2 \subseteq \widehat{J_0} \subseteq \widehat{J}$, so $\mathscr{X}_{\widehat{J}} = \mathscr{X}_{\widehat{J_0}}$. From (2–1) we have $\mathscr{X}_{\widehat{J_0}} \subseteq (\mathscr{X}_J)^\natural$. Combining all, we obtain

$$(\mathscr{X}_J)^\natural = (\mathscr{X}_I)^\natural = (\mathscr{X}_{J_0})^\natural = \mathscr{X}_{\widehat{J}} = \mathscr{X}_{\widehat{J_0}}.$$

Applying (i) again, we complete the proof. □

Denote by $\mathscr{L}(I)$ the algebra of all operators bounded on $(\mathscr{X}_I, \mathfrak{S}_2(H, K))$. Recall that this means that they are bounded operators on $\mathfrak{S}_2(H, K)$, preserve $\mathscr{X}_I$ and are bounded on $\mathscr{X}_I$ in $\|\cdot\|_{\mathscr{X}_I}$. Set

(4–1) $$\mathscr{L}(I)^* = \{M^*\colon M \in \mathscr{L}(I)\} \quad \text{and} \quad \|M\|_I = \|M\|_{B(\mathscr{X}_I)}.$$

If there is an s.n. ideal J such that $\mathscr{X}_J = (\mathscr{X}_I)^\natural$, then it follows from Lemma 3.2 that

(4–2) $$\mathscr{L}(I)^* \subseteq \mathscr{L}(J) \quad \text{and} \quad \|M^*\|_J \le \|M\|_I, \quad \text{for } M \in \mathscr{L}(I).$$

If $\mathscr{X}_I$ is normal, then

(4–3) $$\|M^*\|_J = \|M\|_I.$$

Let $(J_0, J) \in \mathfrak{J}$ and let $(\widehat{J_0}, \widehat{J})$ be the corresponding "dual" pair.

Proposition 4.2. (i) $\mathscr{L}(J_0)^* \subseteq \mathscr{L}(\widehat{J})$ *and* $\|M^*\|_{\widehat{J}} \le \|M\|_{J_0}$, *for all* $M \in \mathscr{L}(J_0)$.

(ii) *If* $J_0 = J$, *then* $\|M^*\|_{\widehat{J}} = \|M\|_J$, *for all* $M \in \mathscr{L}(J)$. *If* J *is reflexive* (*that is,* $J_0 = J$ *and* $\widehat{J_0} = \widehat{J}$), *then also* $\mathscr{L}(J)^* = \mathscr{L}(\widehat{J})$.

(iii) *Let* $J \subseteq \mathfrak{S}_2$ *and let* I *be an s.n. ideal such that* $\widehat{J_0} \subseteq I \subseteq \widehat{J}$ *and the norms* $|\cdot|_{\widehat{J_0}}, |\cdot|_I$, *coincide on* $\widehat{J_0}$. *Then*

$$\mathscr{L}(J_0) \subseteq \mathscr{L}(\widehat{J_0})^* = \mathscr{L}(I)^* = \mathscr{L}(\widehat{J})^* = \mathscr{L}(J),$$

and the inclusion and the equalities are isometric.

In particular, $\mathscr{L}(\mathfrak{S}_p)^* = \mathscr{L}(\mathfrak{S}_{p'})$, *if* $1 < p < \infty$, *where* $p' = p/(p-1)$; $\mathscr{L}(\mathfrak{S}_1)^* = \mathscr{L}(\mathfrak{S}_\infty) = \mathscr{L}(\mathfrak{S}_b)$ *and the norms coincide.*

Proof. Part (i) follows from Lemma 3.2(i) and (4–2).

If $J_0 = J$, then, by Lemma 4.1(ii), the space $\mathscr{X}_{J_0} = \mathscr{X}_J$ is normal, and part (ii) follows from (4–3) and (i).

By Lemma 2.1, $\mathfrak{S}_2 \subseteq \widehat{J_0}$, so that $\mathscr{X}_{\widehat{J_0}} = \mathscr{X}_I = \mathscr{X}_{\widehat{J}}$ $(= \mathfrak{S}_2)$ and the norms coincide. Hence $\mathscr{L}(\widehat{J_0}) = \mathscr{L}(I) = \mathscr{L}(\widehat{J})$ and the norms coincide. By Lemma 4.1, $\mathscr{X}_{\widehat{J}}$ is normal and $(\mathscr{X}_{\widehat{J}})^\natural = (\mathscr{X}_{\widehat{J_0}})^\natural = \mathscr{X}_J$. It follows from (4–2) and (4–3) that

$$\mathscr{L}(\widehat{J})^* \subseteq \mathscr{L}(J) \quad \text{and} \quad \|M\|_{\widehat{J}} = \|M^*\|_J \ \text{ for } M \in \mathscr{L}(\widehat{J}). \tag{4–4}$$

Combining this with (i), we have

$$\mathscr{L}(J_0) \subseteq \mathscr{L}(\widehat{J_0})^* \subseteq \mathscr{L}(J), \tag{4–5}$$

$$\|M\|_J = \|M^*\|_{\widehat{J_0}} = \|M^*\|_{\widehat{J}} \le \|M\|_{J_0} \quad \text{for } M \in \mathscr{L}(J_0). \tag{4–6}$$

By Lemma 4.1(iii), $(\mathscr{X}_J)^\natural = \mathscr{X}_{\widehat{J_0}}$. Hence, by (4–2),

$$\mathscr{L}(J)^* \subseteq \mathscr{L}(\widehat{J_0}) \quad \text{and} \quad \|M^*\|_{\widehat{J_0}} \le \|M\|_J \ \text{ for } M \in \mathscr{L}(J).$$

Since $\mathscr{X}_{J_0} \subseteq \mathscr{X}_J$ and the norms $|\cdot|_J$ and $|\cdot|_{J_0}$ coincide on J_0, we have $\|M\|_{J_0} \le \|M\|_J$, for $M \in \mathscr{L}(J_0)$. Combining this with (4–4)–(4–6), we conclude the proof of (iii). □

Let $I \subset R \subset J$ be s.n. ideals. The ideal R is called an *interpolation ideal* for the pair (I, J), if every bounded operator T on J preserving I also preserves R. It follows from Lemma 2.1 that $T|I$ and $T|R$ are bounded operators. All coseparable ideals are interpolation ideals for the pair $(\mathfrak{S}_1, \mathfrak{S}_\infty)$ (see [Mitjagin 1965]).

Using the results of [Boyd 1969], Arazy [1978] associated the Boyd indices (p_{J_0}, q_{J_0}), where $1 \le p_{J_0} \le q_{J_0} \le \infty$, with each separable ideal J_0 and proved that J_0 is an interpolation ideal for a pair $(\mathfrak{S}_p, \mathfrak{S}_q)$ if $p < p_{J_0}$ and $q_{J_0} < q$. For $J_0 = \mathfrak{S}_p$, one has $p_{J_0} = q_{J_0} = p$. In particular, $\mathfrak{S}_r$ is an interpolation ideal for $(\mathfrak{S}_p, \mathfrak{S}_q)$ if $p < r < q$.

Corollary 4.3. *If R is an interpolation ideal for a pair (I, I_1) of separable ideals, then $\mathscr{L}(I) \cap \mathscr{L}(I_1) \subseteq \mathscr{L}(R)$. In particular,*

$$\mathscr{L}(\mathfrak{S}_\infty)^* \cap \mathscr{L}(\mathfrak{S}_\infty) \subseteq \mathscr{L}(J) \quad \text{for each coseparable ideal } J,$$
$$\mathscr{L}(\mathfrak{S}_r) \subseteq \mathscr{L}(\mathfrak{S}_p) \quad \text{if } 2 \le p < r \text{ or } 1 \le r < p \le 2,$$
$$\mathscr{L}(\mathfrak{S}_p) \cap \mathscr{L}(\mathfrak{S}_q) \subseteq \mathscr{L}(J_0) \quad \text{if } p < p_{J_0} \text{ and } q_{J_0} < q.$$

5. Multipliers for tensor products of representations

Let $\mathscr{A} \otimes \mathscr{B}$ be the minimal tensor product of C*-algebras $\mathscr{A}$ and $\mathscr{B}$ — the completion of the algebraic tensor product $\mathscr{A} \odot \mathscr{B}$ in the minimal C*-norm $\|\cdot\|_{\min}$. If π and ρ are *-representations of $\mathscr{A}$ and $\mathscr{B}$ on Hilbert spaces H and K, we denote by $\pi \otimes \rho$ their tensor product; it is a *-representation of $\mathscr{A} \otimes \mathscr{B}$ on $H \otimes K$.

Let $\xi \in H \otimes K$. Then $\theta(\xi) \in \mathfrak{S}_2(H^d, K)$. It follows from (2–6) that, for $a \odot b \in \mathscr{A} \odot \mathscr{B}$,

$$\theta((\pi \otimes \rho)(a \odot b)\xi) = \theta((\pi(a) \otimes \rho(b))\xi) = \rho(b)\theta(\xi)\pi(a)^d,$$

where $\pi(a)^d$ is the conjugate of $\pi(a)$ on H^d. Thus the representation $\pi \otimes \rho$ is equivalent to the representation $\sigma_{\pi,\rho}$ of $\mathscr{A} \otimes \mathscr{B}$ on $\mathfrak{S}_2(H^d, K)$ such that

$$\text{(5–1)} \qquad \sigma_{\pi,\rho}(a \odot b)T = \rho(b)T\pi(a)^d, \qquad \text{for } a \in \mathscr{A},\ b \in \mathscr{B},\ T \in \mathfrak{S}_2(H^d, K).$$

We say that $\varphi \in \mathscr{A} \otimes \mathscr{B}$ is a $(\pi \otimes \rho, I)$-multiplier if it is a $(\sigma_{\pi,\rho}, I)$-multiplier, that is, $\sigma_{\pi,\rho}(\varphi) \in \mathscr{L}(I)$. Recall that it means that $\sigma_{\pi,\rho}(\varphi)$ preserves $\mathscr{X}_I = I(H^d, K) \cap \mathfrak{S}_2(H^d, K)$, and its restriction to $\mathscr{X}_I$ is bounded in $|\cdot|_I$.

Denote by $\mathbf{M}_I^{\pi,\rho}(\mathscr{A} \otimes \mathscr{B})$ (or just $\mathbf{M}_I^{\pi,\rho}$) the algebra of all $(\pi \otimes \rho, I)$-multipliers, and by $\|\varphi\|_I^{\pi,\rho}$ the norm of $\sigma_{\pi,\rho}(\varphi)$ on $\mathscr{X}_I$ (see (4–1)):

$$\|\varphi\|_I^{\pi,\rho} = \|\sigma_{\pi,\rho}(\varphi)\|_I.$$

Then $\mathbf{M}_{\mathfrak{S}_2}^{\pi,\rho} = \mathscr{A} \otimes \mathscr{B}$. We have $\mathbf{M}_{\mathfrak{S}_\infty}^{\pi,\rho} = \mathbf{M}_{\mathfrak{S}_b}^{\pi,\rho}$ and, omitting the subscript, write $\mathbf{M}^{\pi,\rho}$ and $\|\varphi\|^{\pi,\rho}$.

Remark. It follows immediately from our definitions that all results of Proposition 4.2 and Corollary 4.3 hold if $\mathscr{L}(I)$ is replaced by $\mathbf{M}_I^{\pi,\rho}$.

Clearly, all algebras $\mathbf{M}_I^{\pi,\rho}$ contain $\mathscr{A} \odot \mathscr{B}$, so they are dense in $\mathscr{A} \otimes \mathscr{B}$. We will see now (and use later on) that the unit ball of $\mathbf{M}_I^{\pi,\rho}$ is norm closed in $\mathscr{A} \otimes \mathscr{B}$. In fact, it is closed in a much stronger sense — with respect to a weaker convergence, which can be considered as the analog of the point convergence in the case of usual Schur multipliers.

Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Banach spaces. Probably the weakest condition for an operator T from $\mathfrak{X}$ into $\mathfrak{Y}$ to be considered as a "limit point" for a set W of operators is the condition that $Tx \in \overline{Wx}$, for any vector $x \in \mathfrak{X}$. In this case one says that T belongs to the *reflexive hull* $\mathrm{Ref}(W)$ of W.

Let τ be a *-representation of a C*-algebra $\mathfrak{D}$ on a Hilbert space $\mathcal{H}$. We say that $\varphi \in \mathfrak{D}$ is a τ*-cluster point* of a convex subset W of $\mathfrak{D}$, if $\tau(\varphi) \in \mathrm{Ref}(\tau(W))$. For $\mathcal{H} = \mathfrak{S}_2(H_1, H_2)$, this means that there are $\varphi_n \in W$ such that

$$(5\text{–}2)\qquad \lim_{n\to\infty} \left|\tau(\varphi)(T) - \tau(\varphi_n)(T)\right|_2 = 0 \quad \text{for all } T \in \mathfrak{S}_2(H_1, H_2).$$

We say that φ is a *weak* τ*-cluster point of* W if, for each $T \in \mathfrak{S}_2(H_1, H_2)$, the operator $\tau(\varphi)(T) \in \mathfrak{S}_2(H_1, H_2)$ belongs to $\mathrm{Ref}\{\tau(w)(T)\colon w \in W\}$. By the Hahn–Banach theorem, this means that, for each $x \in H_1$ and $y \in H_2$, there are $\varphi_n \in W$ such that

$$(5\text{–}3)\qquad (\tau(\varphi_n)(T)x, y) \to (\tau(\varphi)(T)x, y).$$

Recall that, for any normed space $(\mathfrak{X}, \|\cdot\|)$, we denote by $b_r(\mathfrak{X})$ the closed ball $\{x \in \mathfrak{X}\colon \|x\| \le r\}$.

Proposition 5.1. (i) *If* $\mathfrak{S}_2 \subseteq I$, *then* $b_r(\mathbf{M}_I^{\pi,\rho})$ *contains all its* $\sigma_{\pi,\rho}$*-cluster points.*

(ii) $b_r(\mathbf{M}^{\pi,\rho})$ *contains all its weak* $\sigma_{\pi,\rho}$*-cluster points.*

Proof. Let φ be a $\sigma_{\pi,\rho}$-cluster point of $b_r(\mathbf{M}_I^{\pi,\rho})$. Then, for $T \in \mathfrak{S}_2(H^d, K)$, there are $\varphi_n \in b_r(\mathbf{M}_I^{\pi,\rho})$ such that (5–2) holds. Hence, by (2–2),

$$|\sigma_{\pi,\rho}(\varphi)(T)|_I \le |\sigma_{\pi,\rho}(\varphi-\varphi_n)(T)|_I + |\sigma_{\pi,\rho}(\varphi_n)(T)|_I$$

$$\le c|\sigma_{\pi,\rho}(\varphi-\varphi_n)(T)|_2 + \|\varphi_n\|_I^{\pi,\rho}|T|_I \le c|\sigma_{\pi,\rho}(\varphi-\varphi_n)(T)|_2 + r|T|_I,$$

for some $c > 0$. Thus $|\sigma_{\pi,\rho}(\varphi)(T)|_I \le r|T|_I$, so $\varphi \in b_r(\mathbf{M}_I^{\pi,\rho})$. Part (i) is proved.

Let $I = \mathfrak{S}_\infty$ and let φ be a weak $\sigma_{\pi,\rho}$-cluster of $b_r(\mathbf{M}^{\pi,\rho})$. For $T \in \mathfrak{S}_2(H^d, K)$, $x \in H^d$, $y \in K$, choose $\varphi_n \in b_r(\mathbf{M}^{\pi,\rho})$ satisfying (5–3). Then a similar argument gives

$$|(\sigma_{\pi,\rho}(\varphi)(T)x, y)| \le r\|T\|\|x\|\|y\|.$$

Hence $\varphi \in b(\mathbf{M}^{\pi,\rho})$. □

We consider now how the space of multipliers depends on the choice of representations. The next theorem establishes that $\mathbf{M}_I^{\pi,\rho}$ does not change if π and ρ are replaced by approximately equivalent representations.

Theorem 5.2. *Let* $\pi' \ll_a \pi$ *and* $\rho' \ll_a \rho$. *If* I *is either a coseparable ideal or contains* $\mathfrak{S}_2$, *then*

$$\mathbf{M}_I^{\pi,\rho} \subseteq \mathbf{M}_I^{\pi',\rho'} \quad \textit{and} \quad \|\varphi\|_I^{\pi',\rho'} \le \|\varphi\|_I^{\pi,\rho} \textit{ for } \varphi \in \mathbf{M}_I^{\pi,\rho}.$$

As a consequence, if $\pi' \sim_a \pi$ *and* $\rho' \sim_a \rho$, *then*

$$\mathbf{M}_I^{\pi,\rho} = \mathbf{M}_I^{\pi',\rho'} \quad \textit{and} \quad \|\varphi\|_I^{\pi',\rho'} = \|\varphi\|_I^{\pi,\rho} \textit{ for } \varphi \in \mathbf{M}_I^{\pi,\rho}.$$

Proof. Let isometries $U_\lambda : H' \to H$ and $V_\mu : K' \to K$ satisfy (3–4). Then, for $a \in \mathcal{A}$ and $b \in \mathcal{B}$, $\|\pi(a)U_\lambda - U_\lambda\pi'(a)\| \to 0$ and $\|\rho(b)V_\mu - V_\mu\rho'(b)\| \to 0$. The operators $W_{\lambda\mu} = U_\lambda \otimes V_\mu$ are isometries from $H' \otimes K'$ into $H \otimes K$, and

$$\begin{aligned}\big\|(\pi\otimes\rho)(a\otimes b)U_\lambda\otimes V_\mu - U_\lambda\otimes V_\mu(\pi'\otimes\rho')(a\otimes b)\big\|\\ \le \big\|(\pi(a)U_\lambda - U_\lambda\pi'(a))\otimes\rho(b)V_\mu\big\| + \big\|U_\lambda\pi'(a)\otimes(\rho(b)V_\mu - V_\mu\rho'(b))\big\|,\end{aligned}$$

which tends to 0. By linearity, $\|(\pi\otimes\rho)(x)W_{\lambda\mu} - W_{\lambda\mu}(\pi'\otimes\rho')(x)\| \to 0$, for all $x \in \mathcal{A}\odot\mathcal{B}$. Since $\|W_{\lambda\mu}\| = 1$, it also holds for all $x \in \mathcal{A}\otimes\mathcal{B}$. Thus

$$\pi'\otimes\rho' \ll_a \pi\otimes\rho.$$

By Lemma 4.1(ii), the normed space $\mathcal{X}_I((H')^d, K')$ is normal. Identifying $H\otimes K$ with $\mathfrak{S}_2(H^d, K)$ and $H'\otimes K'$ with $\mathfrak{S}_2((H')^d, K')$, we have from (2–6) that,

$$W_{\lambda\mu}T = (U_\lambda\otimes V_\mu)T = V_\mu T U_\lambda^d \quad\text{and}\quad W_{\lambda\mu}^*R = (U_\lambda^*\otimes V_\mu^*)R = V_\mu^* R (U_\lambda^*)^d,$$

for $T \in \mathfrak{S}_2((H')^d, K')$ and $R \in \mathfrak{S}_2(H^d, K)$. Since I is an ideal,

$$W_{\lambda\mu}\mathcal{X}_I((H')^d, K') \subseteq \mathcal{X}_I(H^d, K);$$

see (4–6). By (2–4),

$$|W_{\lambda\mu}T|_I \le \|V_\mu\||T|_I\|U_\lambda^d\| \le |T|_I \quad\text{and}\quad |W_{\lambda\mu}^*R|_I \le \|V_\mu^*\||R|_I\|(U_\lambda^*)^d\| \le |R|_I.$$

Hence the approximate subordination $\pi'\otimes\rho' \ll_a \pi\otimes\rho$ satisfies (3–5). Applying Proposition 3.4, we complete the proof. □

Remark. We do not know whether Theorem 5.2 extends to all separable ideals contained in $\mathfrak{S}_2$. Proposition 4.2(i) only gives that $(\mathbf{M}_{J_0}^{\pi,\rho})^* \subseteq \mathbf{M}_J^{\pi',\rho'}$, if $J_0 \subseteq \mathfrak{S}_2$.

Recall that for $T \in B(H)$, $\operatorname{rank}(T) = \dim\overline{(TH)}$. Let π and π' be representations of a C*-algebra $\mathcal{A}$. It was proved in Theorem 5.1 of [Hadwin 1981] that

$$\text{(5–4)}\qquad \pi' \ll_a \pi \iff \operatorname{rank}(\pi'(a)) \le \operatorname{rank}(\pi(a)) \quad \text{for each } a \in \mathcal{A}.$$

Thus it follows from Theorem 5.2 and (5–4) that, if

$$\operatorname{rank}(\pi'(a)) = \operatorname{rank}(\pi(a)) \quad\text{and}\quad \operatorname{rank}(\rho'(b)) = \operatorname{rank}(\rho(b))$$

for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$, then $\mathbf{M}_I^{\pi,\rho} = \mathbf{M}_I^{\pi',\rho'}$, and the corresponding norms are equal.

For some applications (see Section 6) it is important that, for representations of separable algebras on the spaces of arbitrary dimension, one need not distinguish infinite values of the rank.

Corollary 5.3. *Let an s.n. ideal I be either coseparable or contain $\mathfrak{S}_2$. Let π, π' and, respectively, ρ, ρ' be representations of separable C^*-algebras $\mathscr{A}$ and $\mathscr{B}$ on Hilbert spaces H, H' and K, K'. If*

$$\min\{\aleph_0, \operatorname{rank}(\pi'(a))\} \le \min\{\aleph_0, \operatorname{rank}(\pi(a))\}$$

and

$$\min\{\aleph_0, \operatorname{rank}(\rho'(b))\} \le \min\{\aleph_0, \operatorname{rank}(\rho(b))\},$$

for all $a \in \mathscr{A}$, $b \in \mathscr{B}$, then $\mathbf{M}_I^{\pi,\rho} \subseteq \mathbf{M}_I^{\pi',\rho'}$ and $\|\varphi\|_I^{\pi',\rho'} \le \|\varphi\|_I^{\pi,\rho}$, for $\varphi \in \mathbf{M}_I^{\pi,\rho}$.

Proof. Let $\varphi \in \mathbf{M}_I^{\pi,\rho}$ and $\|\varphi\|_I^{\pi,\rho} = C$. We have to prove that for every $T \in I(H'^d, K')$,

$$\|\sigma_{\pi',\rho'}(\varphi)(T)\|_I \le C\|T\|_I.$$

Since T is compact, there are separable subspaces $K_0 \subset K'$ and $G_0 \subset H'^d$ such that $T = P_{K_0} T P_{G_0}$, where P_{K_0}, P_{G_0} are the corresponding projections. The subspaces $K_1 = \overline{\rho'(\mathscr{B})K_0}$ of K' and $G_1 = \overline{\pi'^d(\mathscr{A})G_0}$ of H'^d are also separable because $\mathscr{A}$ and $\mathscr{B}$ are separable. We denote by H_1 the orthogonal complement in H' of the annihilator of G_1.

Since H_1 and K_1 are invariant for π' and ρ', respectively, define new representations π_1' and ρ_1' of $\mathscr{A}$ and $\mathscr{B}$ by

$$\pi_1'(a) = \pi'(a)|_{H_1} \oplus \mathbf{0} \quad \text{and} \quad \rho_1'(b) = \rho'(b)|_{K_1} \oplus \mathbf{0} \quad \text{for } a \in \mathscr{A},\ b \in \mathscr{B}.$$

Since H_1 and K_1 are separable, it follows from our assumptions that $\operatorname{rank}(\pi_1'(a)) \le \operatorname{rank}(\pi(a))$ and $\operatorname{rank}(\rho_1'(b)) \le \operatorname{rank}(\rho(b))$, for all $a \in \mathscr{A}$, $b \in \mathscr{B}$. By (5–4),

$$\pi_1' \ll_a \pi \quad \text{and} \quad \rho_1' \ll_a \rho.$$

Hence, by Theorem 5.2,

$$\varphi \in \mathbf{M}_I^{\pi_1',\rho_1'} \quad \text{and} \quad \|\varphi\|_I^{\pi_1',\rho_1'} \le \|\varphi\|_I^{\pi,\rho}.$$

Thus $\|\sigma_{\pi_1',\rho_1'}(\varphi)(T)\|_I \le C\|T\|_I$. But by the construction of π_1' and ρ_1', we have

$$\sigma_{\pi_1',\rho_1'}(\varphi)(T) = \sigma_{\pi',\rho'}(\varphi)(T),$$

whence $\|\sigma_{\pi',\rho'}(\varphi)(T)\|_I \le C\|T\|_I$. □

Corollary 5.4. *Let an ideal I be either coseparable or contain $\mathfrak{S}_2$, let π, π' be representations of $\mathscr{A}$, and ρ, ρ' be representations of $\mathscr{B}$. Suppose that $\pi(\mathscr{A})$ and $\rho(\mathscr{B})$ contain no nonzero finite rank operators, and that π' and ρ' are separable and satisfy the condition*

(5–5) $$\operatorname{Ker}(\pi) \subseteq \operatorname{Ker}(\pi') \quad \textit{and} \quad \operatorname{Ker}(\rho) \subseteq \operatorname{Ker}(\rho').$$

Then $\mathbf{M}_I^{\pi,\rho} \subseteq \mathbf{M}_I^{\pi',\rho'}$ and $\|\varphi\|_I^{\pi',\rho'} \le \|\varphi\|_I^{\pi,\rho}$, for $\varphi \in \mathbf{M}_I^{\pi,\rho}$.

Proof. It follows from (5–5) that $\operatorname{rank}(\pi'(a)) \le \operatorname{rank}(\pi(a))$, for $a \in \mathcal{A}$, and $\operatorname{rank}(\rho'(b)) \le \operatorname{rank}(\rho(b))$, for $b \in \mathcal{B}$. Hence, by (5–4),

$$\pi' \ll_a \pi \quad \text{and} \quad \rho' \ll_a \rho,$$

and it remains now only to apply Theorem 5.2. □

Remark 5.5. (1) The first condition in Corollary 5.4 can be replaced by the conditions

$$\operatorname{rank}(\pi'(a)) \le \operatorname{rank}(\pi(a)) \quad \text{and} \quad \operatorname{rank}(\rho'(b)) \le \operatorname{rank}(\rho(b)),$$

whenever $\pi(a)$ and $\rho(b)$ are nonzero finite rank operators.

(2) If $\mathcal{A}$ and $\mathcal{B}$ are separable, the condition in Corollary 5.4 that π' and ρ' are separable can be omitted.

Applying Corollary 5.4 to simple C*-algebras we get the following result.

Corollary 5.6. *Let I be either a coseparable ideal, or $\mathfrak{S}_2 \subseteq I$. If $\mathcal{A}$ and $\mathcal{B}$ are simple C^*-algebras different from $\mathfrak{S}_\infty$, then $\mathbf{M}_I^{\pi,\rho}$ is the same for all separable representations π of $\mathcal{A}$ and ρ of $\mathcal{B}$.*

For $I = \mathfrak{S}_\infty$ or $\mathfrak{S}_b$, the conditions in Corollary 5.4 can be further simplified if the representations π, ρ have separating vectors. This simplification is based on the results of Smith [1991].

Recall that a vector $x \in H$ is *separating* for a representation π of $\mathcal{A}$ if the map $T \to Tx$ is injective on the second commutant $\pi(\mathcal{A})''$. This is equivalent to the existence of a *cyclic vector* for the commutant $\pi(\mathcal{A})'$. In the lemma below, $\mathbf{1}$ is the identity operator on a fixed Hilbert space $\mathcal{H}$. The representations $\pi \otimes \mathbf{1}$ and $\rho \otimes \mathbf{1}$ act on $H \otimes \mathcal{H}$ and $K \otimes \mathcal{H}$, respectively.

Lemma 5.7. *Let *-representations π and ρ of C^*-algebras $\mathcal{A}$ and $\mathcal{B}$ on H and K have separating vectors. Then $\mathbf{M}^{\pi,\rho} = \mathbf{M}^{\pi\otimes\mathbf{1},\rho\otimes\mathbf{1}}$ and the norms coincide.*

The proof of the lemma follows along the lines of the proof of in [Smith 1991, Theorem 2.1] and we omit it.

Corollary 5.8. *Let representations π of $\mathcal{A}$ and ρ of $\mathcal{B}$ have separating vectors. Then $\mathbf{M}^{\pi,\rho} \subseteq \mathbf{M}^{\pi',\rho'}$ and $\|\varphi\|^{\pi',\rho'} \le \|\varphi\|^{\pi,\rho}$, for $\varphi \in \mathbf{M}^{\pi,\rho}$, if representations π' and ρ' satisfy (5–5).*

Proof. If $\dim \mathcal{H}$ is sufficiently large, then condition (5–5) implies

$$\operatorname{rank}(\pi'(a)) \le \operatorname{rank}(\pi(a)\otimes\mathbf{1}) \quad \text{and} \quad \operatorname{rank}(\rho'(b)) \le \operatorname{rank}(\rho(b)\otimes\mathbf{1}),$$

for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$. Hence, by (5–4), $\pi' \ll_a \pi \otimes \mathbf{1}$ and $\rho' \ll_a \rho \otimes \mathbf{1}$, and to complete the proof it remains only to apply Theorem 5.2 and Lemma 5.7. □

6. Universal multipliers

In this section we consider only the case $I = \mathfrak{S}_b$. We saw in Corollary 5.8 that in this case all multipliers for a pair of faithful representations with separating vectors are multipliers for all pairs of representations. Let us denote by $\mathbf{M}(\mathcal{A} \otimes \mathcal{B})$ the algebra of "universal" multipliers; it consists of all elements of $\mathcal{A} \otimes \mathcal{B}$ that are $(\pi \otimes \rho, \mathfrak{S}_b)$-multipliers for all pairs (π, ρ). Clearly $\mathcal{A} \odot \mathcal{B} \subseteq \mathbf{M}(\mathcal{A} \otimes \mathcal{B}) \subseteq \mathcal{A} \otimes \mathcal{B}$. For $\varphi \in \mathbf{M}(\mathcal{A} \otimes \mathcal{B})$, set

$$\|\varphi\|_r = \sup_{\pi,\rho} \|\varphi\|^{\pi,\rho}. \tag{6–1}$$

It is not difficult to see that $\|\varphi\|_r < \infty$. Indeed, if $\|\varphi\|^{\pi_n,\rho_n} \to \infty$, consider the representations $\pi = \bigoplus \pi_n$ and $\rho = \bigoplus \rho_n$. Then $\|\varphi\|^{\pi_n,\rho_n} \le \|\varphi\|^{\pi,\rho}$ for all n, a contradiction.

As usual, we denote by $\mathcal{A}^{\text{op}}$ the C*-algebra that consists of all elements of $\mathcal{A}$ and has the same norm and involution, but the reverse multiplication: $a \circ b = ba$. If π is a *-representation of $\mathcal{A}$ on H, the map $\pi^{\text{op}} : a \to \pi(a)^d$ is a *-representation of $\mathcal{A}^{\text{op}}$ on H^d.

Recall that the Haagerup norm on $\mathcal{A} \odot \mathcal{B}$ is defined by

$$\|w\|_h = \inf\Big\{ \big\| \sum a_i a_i^* \big\|^{1/2} \big\| \sum b_i^* b_i \big\|^{1/2} : w = \sum a_i \otimes b_i \Big\}.$$

It is known that

$$\|w\|_{\min} \le \|w\|_h, \qquad \text{for } w \in \mathcal{A} \odot \mathcal{B}.$$

Define a "pseudo-Haagerup" norm on $\mathcal{A} \odot \mathcal{B}$ by setting

$$\|w\|_{ph} = \inf\Big\{ \big\| \sum a_i a_i^* \big\|^{1/2} \big\| \sum b_i b_i^*) \big\|^{1/2} : w = \sum a_i \otimes b_i \Big\}. \tag{6–2}$$

It is a norm, because $\|w\|_{ph} = \|\Gamma w\|_h$, where $\Gamma\colon \mathcal{A} \odot \mathcal{B} \to \mathcal{B} \odot \mathcal{A}^{\text{op}}$ is a linear bijection defined by $\Gamma(a \otimes b) = b \otimes a$.

Theorem 6.1. *The Haagerup and pseudo-Haagerup norms satisfy $\|w\|_r = \|w\|_{ph}$ for $w \in \mathcal{A} \odot \mathcal{B}$.*

Proof. Let π and ρ be representations of $\mathcal{A}$ and $\mathcal{B}$ on H and K. For $w = \sum a_i \otimes b_i$ in $\mathcal{A} \odot \mathcal{B}$, set $A_i = \pi(a_i)$, $B_i = \rho(b_i)$. Take $T \in \mathfrak{S}_2(H^d, K)$, $x \in H$, $y \in K$. By (5–1),

$$\begin{aligned} |(\sigma_{\pi,\rho}(w)Tx, y)| &= \Big|\sum_i (B_i T A_i^d x, y)\Big| \le \sum_i \big|(T A_i^d x, B_i^* y)\big| \\ &\le \sum_i \|T A_i^d x\| \|B_i^* y\| \le \|T\| \Big(\sum_i \|A_i^d x\|^2\Big)^{1/2} \Big(\sum_i \|B_i^* y\|^2\Big)^{1/2}. \end{aligned}$$

We have

$$\sum_i \|B_i^* y\|^2 = \sum_i (B_i^* y, B_i^* y) = \Big(y, \Big(\sum_i B_i B_i^*\Big) y\Big)$$
$$\leq \Big\|\rho\Big(\sum_i b_i b_i^*\Big)\Big\| \, \|y\|^2 \leq \Big\|\sum_i b_i b_i^*\Big\| \, \|y\|^2.$$

We see from (2–4) that $(\pi(a)^d)^* \pi(a)^d = \pi(aa^*)^d$ for $a \in \mathscr{A}$. From this, and using (2–4) again, we get $\sum_i \|A_i^d x\|^2 \leq \big\|\sum_i a_i a_i^*\big\| \, \|x\|^2$. Therefore

$$\big|(\sigma_{\pi,\rho}(w) T x, y)\big| \leq \|T\| \, \Big\|\sum_i a_i a_i^*\Big\|^{1/2} \Big\|\sum_i b_i b_i^*\Big\|^{1/2} \|x\| \, \|y\|.$$

Hence

$$\|\sigma_{\pi,\rho}(w) T\| \leq \|w\|_{ph} \|T\|,$$

so $\|w\|^{\pi,\rho} \leq \|w\|_{ph}$. Thus

$$\|w\|_r \leq \|w\|_{ph} \quad \text{for } w \in \mathscr{A} \odot \mathscr{B}. \tag{6–3}$$

To prove the converse inequality, denote by $\mathscr{G}$ the space of all linear functionals g on $\mathscr{A} \odot \mathscr{B}$ such that

$$|g(w)| \leq \|w\|_{ph}, \quad \text{for } w \in \mathscr{A} \odot \mathscr{B}.$$

For $g \in \mathscr{G}$, let $\hat{g}$ be the linear functional on $\mathscr{B} \odot \mathscr{A}^{\mathrm{op}}$ acting by the rule $\hat{g}(w) = g(\Gamma w)$ for $w \in \mathscr{B} \odot \mathscr{A}^{\mathrm{op}}$. By (6–2), $|\hat{g}(w)| \leq \|w\|_h$. Hence $\hat{g}$ extends to a bounded functional on the Haagerup tensor product $\mathscr{B} \otimes_h \mathscr{A}^{\mathrm{op}}$ and $\|\hat{g}\| \leq 1$. Consider now the bilinear map on $\mathscr{B} \times \mathscr{A}^{\mathrm{op}}$ defined by the formula: $G(b, a) = \hat{g}(b \otimes a)$, for $b \in \mathscr{B}$ and $a \in \mathscr{A}^{\mathrm{op}}$. It follows from Theorems 1.5.2 and 1.5.4 of [Sinclair and Smith 1995] that there exist *-representations ρ of $\mathscr{B}$ on K and τ of $\mathscr{A}^{\mathrm{op}}$ on L, a bounded operator $T \colon L \to K$, and elements $x \in L$ and $y \in K$ with $\|x\| = \|y\| = 1$, such that

$$G(b, a) = (\rho(b) T \tau(a) x, y), \quad \text{for } b \in \mathscr{B},\ a \in \mathscr{A}^{\mathrm{op}},$$

and $\|G\|_{cb} = \|\hat{g}\|_h = \|T\| \leq 1$.

Set $H = L^d$ and $\pi(a) = \tau(a)^d$. Then π is a *-representation of $\mathscr{A}$ on H and

$$g(a \otimes b) = \hat{g}(b \otimes a) = G(b, a) = (\rho(b) T \pi(a)^d x, y), \tag{6–4}$$

for $a \in \mathscr{A}$ and $b \in \mathscr{B}$. For $w = \sum_i a_i \otimes b_i$, denote by $\sigma_{\pi,\rho}^{\infty}(w)$ the extension of $\sigma_{\pi,\rho}(w)$ from $\mathfrak{S}_2(H^d, K)$ to $\mathfrak{S}_\infty(H^d, K)$. Let $\lambda_{\pi,\rho}(w)$ be the second adjoint of $\sigma_{\pi,\rho}^{\infty}(w)$ acting on the second dual space $B(H^d, K)$. By (6–1),

$$\|\lambda_{\pi,\rho}(w)\| = \|\sigma_{\pi,\rho}^{\infty}(w)\| = \|w\|^{\pi,\rho} \leq \|w\|_r. \tag{6–5}$$

For $T \in \mathfrak{S}_\infty(H^d, K)$, we have $\sigma_{\pi,\rho}^{\infty}(w) T = \sum \rho(b_i) T \pi(a_i)^d$. This implies that $\lambda_{\pi,\rho}(w) T = \sum \rho(b_i) T \pi(a_i)^d$ for all $T \in B(H^d, K)$. Hence, by (6–4), $g(w) =$

$(\lambda_{\pi,\rho}(w)Tx, y)$. Using (6–5), we obtain $|g(w)| \le \|\lambda_{\pi,\rho}(w)T\| \le \|w\|_r$. Thus

$$|g(w)| \le \|w\|_r \quad \text{for } g \in \mathcal{G} \text{ and } w \in \mathcal{A} \odot \mathcal{B}.$$

Making use of the Hahn–Banach theorem, we have $\|w\|_{ph} = \sup_{g\in\mathcal{G}} |g(w)| \le \|w\|_r$, for $w \in \mathcal{A} \odot \mathcal{B}$. Combining this with (6–3), we complete the proof. □

We say that a net $\{d_\nu\}$ of elements of a C*-algebra $\mathcal{D}$ *point-weakly converges* to $d \in \mathcal{D}$, and write

$$d_\nu \overset{\text{pw}}{\to} d,$$

if for each irreducible representation τ of $\mathcal{D}$, $\tau(d_\nu) \to \tau(d)$ in the weak operator topology. Denote by $(\mathcal{A} \odot \mathcal{B})^\sim$ the linear space of all $\varphi \in \mathcal{A} \otimes \mathcal{B}$ for which there is a net $\{w_\nu\}$ in $\mathcal{A} \odot \mathcal{B}$ point-weakly converging to φ such that $\sup \|w_\nu\|_{ph} < \infty$.

Theorem 6.2. $(\mathcal{A} \odot \mathcal{B})^\sim \subseteq \mathbf{M}(\mathcal{A} \otimes \mathcal{B})$.

Proof. Let $w_\nu \in \mathcal{A} \odot \mathcal{B}$, $w_\nu \overset{\text{pw}}{\to} \varphi \in \mathcal{A} \otimes \mathcal{B}$ and $D = \sup \|w\|_{ph} < \infty$. To prove that $\varphi \in \mathbf{M}(\mathcal{A} \otimes \mathcal{B})$, we have to check that $\|\varphi\|^{\pi,\rho} \le D$ for all representations π, ρ.

Let firstly π and ρ be direct sums of irreducible representations: $\pi = \bigoplus_{\lambda\in\Lambda} \pi_\lambda$ and $\rho = \bigoplus_{\gamma\in\Gamma} \rho_\gamma$ act on Hilbert spaces $H = \bigoplus H_\lambda$ and $K = \bigoplus K_\lambda$, respectively. By Theorem 6.1, $\|w_\nu\|^{\pi,\rho} \le D$, for each ν, so $\|\sigma_{\pi,\rho}(w_\nu)(T)\| \le \|w_\nu\|^{\pi,\rho}\|T\| \le D\|T\|$, for any operator $T \in \mathfrak{S}_2(H^d, K)$. To prove that $\|\sigma_{\pi,\rho}(\varphi)(T)\| \le D\|T\|$, it suffices to show that the operators $\sigma_{\pi,\rho}(w_\nu)(T)$ tend to $\sigma_{\pi,\rho}(\varphi)(T)$ in the weak operator topology. Moreover, the standard boundedness arguments show that it suffices to prove that

$$(6\text{–}6) \qquad (\sigma_{\pi,\rho}(w_\nu)(T)x, y) \to (\sigma_{\pi,\rho}(\varphi)(T)x, y),$$

for each $x \in U = \bigcup H^d_\lambda$ and $y \in V = \bigcup K_\gamma$, since U, V are generating subsets in H^d and K, respectively.

For $x \in H^d_\lambda$ and $y \in K_\gamma$, set $R = x \otimes y$. We have from (5–1) that, for each $\psi \in \mathcal{A} \otimes \mathcal{B}$, $\sigma_{\pi,\rho}(\psi)(R) = \sigma_{\pi_\lambda,\rho_\gamma}(\psi)(R)$. Hence, we obtain from (2–1) that

$$\begin{aligned}(\sigma_{\pi,\rho}(\psi)(T)x, y) &= \operatorname{Tr}(y \otimes \sigma_{\pi,\rho}(\psi)(T)x) = \operatorname{Tr}(\sigma_{\pi,\rho}(\psi)(T)(x \otimes y)^*) \\ &= \overline{(x \otimes y, \sigma_{\pi,\rho}(\psi)(T))_2} = \overline{(\sigma_{\pi,\rho}(\psi^*)(R), T)_2} \\ &= \overline{(\sigma_{\pi_\lambda,\rho_\gamma}(\psi^*)(R), T)_2}\,.\end{aligned}$$

Since $\sigma_{\pi_\lambda,\rho_\gamma}$ is an irreducible representation $\mathcal{A} \otimes \mathcal{B}$ and $w_\nu^* \overset{\text{pw}}{\to} \varphi^*$, it follows that (6–6) holds.

Now let π and ρ be arbitrary. Consider the representation τ of $\mathcal{A}$, which is the direct sum of all irreducible representations of $\mathcal{A}$ repeated $\dim(\mathcal{H}_\pi)$ times. Then, for each $a \in \mathcal{A}$, we have $\operatorname{rank}(\pi(a)) \le \operatorname{rank}(\tau(a))$, whence $\pi \ll_a \tau$; see (5–4). Similarly, there is a representation χ of $\mathcal{B}$, which is a direct sum of irreducible representations, such that $\rho \ll_a \chi$. By Theorem 5.2, $\mathbf{M}^{\tau,\chi} \subseteq \mathbf{M}^{\pi,\rho}$, so $\varphi \in \mathbf{M}^{\pi,\rho}$. Thus $\varphi \in \mathbf{M}(\mathcal{A} \otimes \mathcal{B})$ and $(\mathcal{A} \odot \mathcal{B})^{\sim} \subseteq \mathbf{M}(\mathcal{A} \otimes \mathcal{B})$. □

Problem 6.3. Does $(\mathcal{A} \odot \mathcal{B})^{\sim}$ coincide with $\mathbf{M}(\mathcal{A} \otimes \mathcal{B})$?

We will see further that for commutative C*-algebras the answer is positive.

7. Multipliers of commutative algebras; (μ, ν)-multipliers

The commutativity of the C*-algebras $\mathcal{A}$ and $\mathcal{B}$ implies significant simplifications to the previous results and constructions. To begin with, each representation of a commutative algebra has a separating vector. Hence, by Corollary 5.8, the algebra $\mathbf{M}^{\pi,\rho}(\mathcal{A} \otimes \mathcal{B})$ depends only on the kernels of the representations π, ρ. In particular, if π, ρ are faithful, then $\mathbf{M}^{\pi,\rho}(\mathcal{A} \otimes \mathcal{B}) = \mathbf{M}(\mathcal{A} \otimes \mathcal{B})$ and $\|\varphi\|^{\pi,\rho} = \|\varphi\|_r$. Since $\|\varphi\|_{ph} = \|\varphi\|_h$ for commutative algebras, Theorem 6.1 shows that, for faithful π and ρ,

$$\|\varphi\|^{\pi,\rho} = \|\varphi\|_h, \qquad \text{for } \varphi \in \mathcal{A} \odot \mathcal{B}.$$

It was proved in [1953] that the norm $\|\cdot\|_h$ on $\mathcal{A} \otimes \mathcal{B}$ is equivalent to the projective tensor norm $\|\cdot\|_\gamma$. Thus in the case of commutative algebras our Theorem 6.2 implies that the Varopoulos tensor algebra $V(X, Y) = C(X) \hat{\otimes} C(Y)$ and its "tilde-algebra" (see [Graham and McGehee 1979]) are topologically included into $\mathbf{M}(C(X) \otimes C(Y))$. In fact, this theorem deals with a wider "tilde-extension" consisting of pointwise limits of $\|\cdot\|_\gamma$-bounded nets. We will return to this topic later.

Let $\mathcal{U}$ be a commutative operator C*-algebra on H with the space Λ of all maximal ideals. Then

$$H = \bigoplus_{\gamma \in \Gamma} H_\gamma,$$

where all $H_\gamma \approx L_2(\Lambda, \mu_\gamma)$ are invariant for $\mathcal{U}$, and each $f \in \mathcal{U}$ acts on H_γ as a multiplication operator. The antiisometric involution $j : \{g_\gamma(\lambda)\} \mapsto \{\overline{g_\gamma(\lambda)}\}$ on H induces an involution on $\mathcal{U}$ given by $jAj = A^*$, for $A \in \mathcal{U}$. Taking into account (2–5), which here becomes $\partial A^* \partial^{-1} = A^d$, we see that the unitary operator $V = \partial j$ from H to H^d establishes a unitary equivalence of A and A^d: $A^d = VAV^{-1}$. For representations π of $\mathcal{A}$ on H and ρ of $\mathcal{B}$ on K, we identify $\mathfrak{S}_2(H^d, K)$ with $\mathfrak{S}_2(H, K)$ by the formula $U(T) = TV$, for $T \in \mathfrak{S}_2(H^d, K)$. Using this and (5–1), we will assume that $\sigma_{\pi,\rho}$ acts on $\mathfrak{S}_2(H, K)$ by the formula

$$\sigma_{\pi,\rho}(a \otimes b)R = \rho(b)R\pi(a).$$

We now prove that, for commutative $\mathcal{A}$, $\mathcal{B}$, the subalgebras $M_I^{\pi,\rho}$ in $\mathcal{A} \otimes \mathcal{B}$ are involutive.

Proposition 7.1. *If $\mathcal{A}$ and $\mathcal{B}$ are commutative, then $\mathbf{M}_I^{\pi,\rho}(\mathcal{A}\otimes\mathcal{B})$ is a *-subalgebra of $\mathcal{A}\otimes\mathcal{B}$ for each pair of representations π, ρ and each s.n. ideal I. Moreover,*

$$\|\varphi^*\|_I^{\pi,\rho} = \|\varphi\|_I^{\pi,\rho}, \quad \textit{for } \varphi \in \mathbf{M}_I^{\pi,\rho}(\mathcal{A}\otimes\mathcal{B}).$$

Proof. Consider the antiisometric involutions j on H_π and i on K_ρ such that $\pi(a^*) = j\pi(a)j$ and $\rho(b^*) = i\rho(b)i$, for all $a \in \mathcal{A}$, $b \in \mathcal{B}$. Then, for $T \in \mathfrak{S}_2(H, K)$,

$$\sigma_{\pi,\rho}(a^*\otimes b^*)(T) = \rho(b^*)T\pi(a^*) = i\{\sigma_{\pi,\rho}(a\otimes b)(iTj)\}j.$$

Hence $\sigma_{\pi,\rho}(\varphi^*)(T) = i\{\sigma_{\pi,\rho}(\varphi)(iTj)\}j$, for all $\varphi \in \mathcal{A}\otimes\mathcal{B}$. For any s.n. ideal I and any operator $T \in I(H, K)$, $iTj \in I(H, K)$ and $|iTj|_I = |T|_I$. Thus it follows that $\varphi \in \mathbf{M}_I^{\pi,\rho}(\mathcal{A}\otimes\mathcal{B})$ implies $\varphi^* \in \mathbf{M}_I^{\pi,\rho}(\mathcal{A}\otimes\mathcal{B})$, and $\|\varphi^*\|_I^{\pi,\rho} = \|\sigma_{\pi,\rho}(\varphi^*)\|_I = \|\sigma_{\pi,\rho}(\varphi)\|_I = \|\varphi\|_I^{\pi,\rho}$. □

Let X be the space of all maximal ideals of a commutative C*-algebra $\mathcal{A}$. Then $\mathcal{A} = C_0(X)$ and each representation π of $\mathcal{A}$ corresponds to a spectral measure $\mathcal{E}_\pi$ on X, that is, a σ-additive map from the σ-algebra of all Borel subsets of X to the lattice of projections in H_π. An isolated point x in the support, supp($\mathcal{E}_\pi$), of $\mathcal{E}_\pi$ must be an atom: $\mathcal{E}_\pi(\{x\}) \neq 0$. To apply the results of the previous sections we need to express rank($\pi(f)$) in terms of the spectral measure. Set

$$\mathcal{S}(f, \mathcal{E}_\pi) = \{x \in \operatorname{supp}(\mathcal{E}_\pi) : f(x) \neq 0\}.$$

Lemma 7.2. *For $f \in C_0(X)$,* rank($\pi(f)$) $< \infty$ *if and only if $\mathcal{S}(f, \mathcal{E}_\pi)$ consists of a finite number of points of finite multiplicity* (dim($\mathcal{E}_\pi(\{x\})$) $< \infty$). *In this case*

$$\operatorname{rank}(\pi(f)) = \sum_{x\in S(f,\mathcal{E}_\pi)} \dim(\mathcal{E}_\pi(\{x\})).$$

Proof. If $\mathcal{S}(f, \mathcal{E}_\pi)$ is infinite, it contains a countable set of points with disjoint neighbourhoods. Hence rank($\pi(f)$) is infinite. Let $\mathcal{S}(f, \mathcal{E}_\pi) = \{x_1, \dots, x_n\}$. Since f is continuous, $\mathcal{E}_\pi(\{x_i\}) \neq 0$ and $\pi(f) = \sum_i f(x_i)\mathcal{E}_\pi(\{x_i\})$. □

It follows from Lemma 7.2 that the kernel of a representation depends only on the support of the corresponding spectral measure.

Corollary 7.3. *Let π, π' and ρ, ρ' be, respectively, representations of commutative C*-algebras $\mathcal{A} = C_0(X)$ and $\mathcal{B} = C_0(Y)$. Let*

(7–1) $$\operatorname{supp}(\mathcal{E}_{\pi'}) \subset \operatorname{supp}(\mathcal{E}_\pi) \quad \textit{and} \quad \operatorname{supp}(\mathcal{E}_{\rho'}) \subset \operatorname{supp}(\mathcal{E}_\rho).$$

Then

(i) $\mathbf{M}^{\pi,\rho}(\mathcal{A}\otimes\mathcal{B}) \subseteq \mathbf{M}^{\pi',\rho'}(\mathcal{A}\otimes\mathcal{B})$, *and the inclusion is contractive.*

(ii) *Suppose that I is either a coseparable s.n. ideal or $\mathfrak{S}_2 \subseteq I$. Let the representations π' and ρ' be separable, and let, for all isolated points $x \in \operatorname{supp}(\mathscr{E}_\pi)$ and $y \in \operatorname{supp}(\mathscr{E}_\rho)$,*

$$\dim(\mathscr{E}_{\pi'}(\{x\})) \le \dim(\mathscr{E}_\pi(\{x\})) \quad \textit{and} \quad \dim(\mathscr{E}_{\rho'}(\{y\})) \le \dim(\mathscr{E}_\rho(\{y\})).$$

Then $\mathbf{M}_I^{\pi,\rho}(\mathscr{A} \otimes \mathscr{B}) \subseteq \mathbf{M}_I^{\pi',\rho'}(\mathscr{A} \otimes \mathscr{B})$, and the inclusion is contractive.

Proof. The inclusions (7–1) imply (5–4). Since all representations of commutative C*-algebras have separating vectors, (i) follows from Corollary 5.8.

Applying Lemma 7.2, Corollary 5.4 and Remark 5.5, we get (ii). □

Let μ and ν be measures on X and Y, let $H = L_2(X, \mu)$ and $K = L_2(Y, \nu)$. Then $\mathfrak{S}_2(H, K)$ consists of integral Hilbert–Schmidt operators R with kernels $r(x, y)$ on $X \times Y$ satisfying (1–1). Each $\varphi \in L_\infty(X \times Y, \mu \times \nu)$ defines a bounded linear map Φ_φ on $\mathfrak{S}_2(H, K)$: $\Phi_\varphi(R)$ is the integral operator with kernel $\varphi(x, y)r(x, y)$. Recall from the Introduction that if Φ_φ preserves $\mathscr{X}_I$ and is bounded in $|\cdot|_I$, then φ is called a (μ, ν, I)-multiplier. We denote by $\|\Phi_\varphi\|_I$ the norm of the operator Φ_φ acting on $\mathscr{X}_I$, and by $\mathfrak{M}_{\mu,\nu}(I)$ the set of all (μ, ν, I)-multipliers.

Every multiplicity-free representation of $C_0(X)$ is defined by a regular σ-finite Borel measure μ on X, and acts on $L_2(X, \mu)$ by multiplication operators:

$$\pi_\mu(f)h(x) = f(x)h(x).$$

Let ρ_ν be a multiplicity-free representation of $C_0(Y)$ defined by a regular σ-finite Borel measure ν on Y. Then $C_0(X) \otimes C_0(Y) = C_0(X \times Y)$ and, for each φ in $C_0(X \times Y)$, $\sigma_{\pi_\mu,\rho_\nu}(\varphi)$, acts on $\mathfrak{S}_2(H, K)$ by multiplying the integral kernels of operators $R \in \mathfrak{S}_2(H, K)$ by φ. Thus

$$\sigma_{\pi_\mu,\rho_\nu}(\varphi) = \Phi_\varphi \quad \text{for } \varphi \in C_0(X \times Y).$$

Therefore $(\pi_\mu \otimes \pi_\nu, I)$-multipliers are *continuous* (μ, ν, I)-multipliers, and

$$\mathbf{M}_I^{\pi_\mu,\rho_\nu}(C_0(X) \otimes C_0(Y)) = \mathfrak{M}_{\mu,\nu}(I) \cap C_0(X \times Y).$$

We will use the simplified notations and write

$$\mathbf{M}_I^{\mu,\nu} \text{ instead of } \mathbf{M}_I^{\pi_\mu,\rho_\nu} \quad \text{and} \quad \|\varphi\|_I^{\mu,\nu} \text{ instead of } \|\varphi\|_I^{\pi_\mu,\rho_\nu}.$$

Thus

$$\|\varphi\|_I^{\mu,\nu} \stackrel{\text{def}}{=} \|\varphi\|_I^{\pi_\mu,\rho_\nu} \stackrel{\text{def}}{=} \|\sigma_{\pi_\mu,\rho_\nu}(\varphi)\|_I = \|\Phi_\varphi\|_I \quad \text{for } \varphi \in C_0(X \times Y).$$

Corollary 7.4. *Let X, Y be locally compact spaces with countable bases. Let μ, μ' and ν, ν' be σ-finite Borel measure on X and Y, respectively. Let I be either a coseparable s.n. ideal or $\mathfrak{S}_2 \subseteq I$. If*

$$\operatorname{supp}(\mu') \subseteq \operatorname{supp}(\mu) \quad \textit{and} \quad \operatorname{supp}(\nu') \subseteq \operatorname{supp}(\nu),$$

then $\mathbf{M}_I^{\mu,\nu}(C_0(X)\otimes C_0(Y)) \subseteq \mathbf{M}_I^{\mu',\nu'}(C_0(X)\otimes C_0(Y))$, *and the inclusion is contractive.*

Proof. Since X, Y have countable bases and all measures are σ-finite, the corresponding $L_2(\cdot\,,\cdot)$ spaces are separable. For any $A \subset X$, $\mathscr{E}_{\pi_\mu}(A)$ is the multiplication operator by the characteristic function of A. Hence $\operatorname{supp}(\mu)$ coincides with $\operatorname{supp}(\mathscr{E}_{\pi_\mu})$. Since $\dim(\mathscr{E}_{\pi_\mu}(\{x\})) = 1$ for each isolated point $x \in \operatorname{supp}(\mu)$, our result follows from Corollary 7.3. □

Our next aim is to relate continuous (μ, ν, I)-multipliers to Schur I-multipliers. Let $H = l_2(X)$ be the Hilbert space of all complex-valued functions g on X such that $\sum_{x\in X}|g(x)|^2 < \infty$. Denote by τ_X the representation of $C_0(X)$ on $l_2(X)$ by diagonal operators

$$(\tau_X(h)g)(x) = h(x)g(x) \quad \text{for } h \in C_0(X),\, g \in l_2(X).$$

Let $K = l_2(Y)$. Each $T \in \mathfrak{S}_2(H, K)$ corresponds to a matrix $(t(x, y))$ with $\sum |t(x, y)|^2 < \infty$. For a bounded complex-valued function φ on $X \times Y$, the operator $S_\varphi(T) = (\varphi(x, y)t(x, y))$ is bounded on $\mathfrak{S}_2(H, K)$. If S_φ preserves $\mathscr{X}_I = I(H, K) \cap \mathfrak{S}_2(H, K)$ and is bounded in $|\cdot|_I$, then φ is called a Schur I-multiplier and $\|S_\varphi\|_I$ denotes the norm of the operator S_φ acting on $\mathscr{X}_I$. Clearly, Schur I-multipliers on $X \times Y$ are exactly (τ_X, τ_Y, I)-multipliers.

Theorem 7.5. *Let X, Y be locally compact spaces with countable bases and let μ, ν be Borel σ-finite measures on X and Y, with* $\operatorname{supp}(\mu) = X$, $\operatorname{supp}(\nu) = Y$. *Suppose that an s.n. ideal I is either coseparable or $\mathfrak{S}_2 \subseteq I$. A function $\varphi \in C_0(X \times Y)$ is a (μ, ν, I)-multiplier on $X \times Y$ if and only if it is a Schur I-multiplier on $X \times Y$. In this case $\|S_\varphi\|_I = \|\varphi\|_I^{\mu,\nu}$.*

Proof. Since $L_2(X, \mu)$ is a separable space, $\operatorname{rank}(\pi_\mu(f)) \le \aleph_0$, for $f \in C_0(X)$. Let us show that

$$\operatorname{rank}(\pi_\mu(f)) = \min\{\aleph_0, \operatorname{rank}(\tau_X(f))\}. \tag{7–2}$$

If $\operatorname{rank}(\tau_X(f)) \ge \aleph_0$, then, by Lemma 7.2, $\operatorname{rank}(\pi_\mu(f))$ can not be finite, so (7–2) holds. If $\operatorname{rank} \tau_X(f) = n < \infty$, then the set $\mathscr{S}(f, \mathscr{E}_\tau)$ consists of n points. By the continuity of f, these points must be isolated in X. Hence, by Lemma 7.2, $\operatorname{rank}(\pi_\mu(f)) = n$, and (7–2) holds. Since the same equality holds for π_ν and τ_Y, and the C*-algebras $C_0(X)$, $C_0(Y)$ are separable, our result follows from Corollary 5.3. □

Problem 7.6. Let X and Y be locally compact spaces with countable bases and $I = \mathfrak{S}_p$. Is each Schur I-multiplier $\varphi \in C_0(X \times Y)$ a $(\pi \otimes \rho, I)$-multiplier for all separable representations π of $C_0(X)$ and ρ of $C_0(Y)$?

The positive answer to this problem follows from the previous results in two cases: if X and Y have no isolated points, and if $I = \mathfrak{S}_\infty$.

(1) *Assume X and Y have no isolated points.* Let μ, ν be Borel σ-finite measures without atoms, with $\mathrm{supp}(\mu) = X$, $\mathrm{supp}(\nu) = Y$. Then $\pi_\mu(C_0(X))$ and $\rho_\nu(C_0(Y))$ have no nonzero finite rank operators, and $\mathrm{Ker}(\pi_\mu) = \mathrm{Ker}(\rho_\nu) = \{0\}$. By Theorem 7.5, φ is a (μ, ν, I)-multiplier on $X \times Y$. By Corollary 5.4, it is a $(\pi \otimes \rho, I)$-multiplier for all separable representations π of $C_0(X)$ and ρ of $C_0(Y)$.

(2) *Assume $I = \mathfrak{S}_\infty$.* Let $I = \mathfrak{S}_\infty$. Every cyclic representation of $C_0(X)$ is of the form π_μ. Each separable representation π of $C_0(X)$ is equivalent to a subrepresentation of $\pi_\mu \otimes \mathbf{1}_{\mathcal{H}}$, for separable $\mathcal{H}$ and some cyclic representation π_μ. By Theorem 7.5, φ is a (μ, ν, I)-multiplier on $X \times Y$. By Lemma 5.7, it is a $((\pi_\mu \otimes \mathbf{1}_{\mathcal{H}}) \otimes (\rho_\nu \otimes \mathbf{1}_{\mathcal{H}}), I)$-multiplier. Hence it is a $(\pi \otimes \rho, I)$-multiplier.

For $I = \mathfrak{S}_\infty$ (or equivalently $\mathfrak{S}_b$, $\mathfrak{S}_1$), Schur I-multipliers were described by Grothendieck in [1953] (see also Theorems 5.1 and 5.5 in [Pisier 2001]): φ is a Schur $\mathfrak{S}_\infty$-multiplier if and only if there are bounded families $\{u_\lambda\}$ and $\{v_\lambda\}$ of functions on X and Y, such that φ belongs to the pointwise closure of the convex hull of $\{u_\lambda(x)v_\lambda(y)\}$. It can be easily seen from the proof in [Pisier 2001] that if $\varphi \in C_0(X, Y)$, one can choose u_λ, v_λ among Borel functions. Since each Borel function $u(x)$ with $|u(x)| \le 1$ can be pointwise approximated by functions from $b_1(C_0(X))$, the inclusion of Theorem 6.2 is, in fact, an equality for commutative $\mathcal{A}$ and $\mathcal{B}$.

Corollary 7.7. *If $\mathcal{A}$ and $\mathcal{B}$ are commutative, then* $(\mathcal{A} \odot \mathcal{B})^\sim = M(\mathcal{A} \otimes \mathcal{B})$.

Recall one of the equivalent definitions (cf. [Birman and Solomyak 1967]) of a double operator integral (DOI). Let $\mathcal{E}$, $\mathcal{F}$ be spectral measures on X and Y with values in the sets $\mathcal{P}(H)$ and $\mathcal{P}(K)$ of all projections in $B(H)$ and $B(K)$, respectively. One defines their direct product $\mathcal{G}$ as a spectral measure on $X \times Y$ with values in $\mathcal{P}(\mathfrak{S}_2(H, K))$ by $\mathcal{G}(A \times B)(T) = \mathcal{F}(B)T\mathcal{E}(A)$, and further extends it from rectangulars to all Borel sets. For a bounded Borel function φ on $X \times Y$, one defines the operator $\mathfrak{I}_\varphi$ on $\mathfrak{S}_2(H, K)$ by

$$\mathfrak{I}_\varphi = \int \varphi(x, y)\, d\mathcal{G}.$$

If $\mathfrak{I}_\varphi$ is bounded on $I \cap \mathfrak{S}_2$ in $|\cdot|_I$, then one says that φ defines DOI on I.

Let now $\varphi \in C_0(X \times Y)$, and let spectral measures $\mathcal{E}_\pi$, $\mathcal{F}_\rho$ correspond to representations π and ρ of $C_0(X)$ and $C_0(Y)$, respectively. Then φ defines DOI on I if and only if $\varphi \in \mathbf{M}_I^{\pi,\rho}(C_0(X) \otimes C_0(Y))$. Thus the DOI theory, restricted to continuous functions, can be considered as a part of the general operator multipliers theory for tensor products of representations of C*-algebras. In particular, Corollary 7.3 states that the space of continuous functions that define bounded DOI depends only on the supports of the spectral measures and (if $I \neq \mathfrak{S}_1, \mathfrak{S}_\infty, \mathfrak{S}_b$) on the multiplicity of their atoms.

Some results in this section are known or can be deduced from the DOI theory. Proposition 7.1 was, in fact, proved by Birman and Solomyak [1967; 1973]. We presented the proof here because it is short and "coordinate free". For functions that define DOI on $\mathfrak{S}_\infty$ (or, equivalently, on $\mathfrak{S}_1$, $\mathfrak{S}_b$), a precise description was obtained by Peller [1985], completing previous results of Birman and Solomyak [1967; 1973] (a transparent proof of Peller's theorem can be found in the recent book [Hiai and Kosaki 2003]). Without stating this directly, Peller's theorem shows that only supports of the spectral measures are essential in the description of $\mathbf{M}^{\pi,\rho}$. No definitive description of $\mathbf{M}_I^{\pi,\rho}$ is known for other I. We will discuss Peller's theorem at the end of Section 8.

8. The notion of ω-continuity and an analog of Luzin's theorem

Our goal now is to remove the restriction of continuity on (μ, ν, I)-multiplier in the main results of Section 6. Moreover, we are going to extend these results to functions on the product of measure spaces (X, μ) and (Y, ν) without distinguished topologies. On the other hand, even in this case, in order to be a (μ, ν, I)-multiplier (at least if $I = \mathfrak{S}_\infty$; see Proposition 9.1), a function still has to be "continuous" in some natural *pseudotopology*, called ω-pseudotopology, associated in [Erdos et al. 1998] with the product of measure spaces. In this section we establish some auxiliary results on ω-continuous functions.

Recall that a pseudotopology on a set is defined by a family of its subsets (called *pseudoopen*), which is closed under finite intersections and countable unions. The complements of pseudoopen sets are called *pseudoclosed*. A complex-valued function is *pseudocontinuous* if the preimages of open sets are pseudoopen.

The ω-pseudotopology on the product of measure spaces is defined as follows. A subset N of $X \times Y$ is called *marginally null* if there are subsets $F \subseteq X$ and $S \subseteq Y$ of zero measure such that $N \subseteq (F \times Y) \cup (X \times S)$. A set E is *ω-open* if there is a countable family of measurable rectangles $A_n \times B_n$ such that the symmetric difference of $\bigcup (A_n \times B_n)$ and E is marginally null. The space of all ω-continuous functions on $X \times Y$ is denoted by $C_{\mu,\nu}(X \times Y)$.

A measure space (X, μ) is called *standard* if there is a topology on X (called *admissible*) with respect to which μ is a σ-finite Radon measure, that is, for each measurable set A of finite measure and each $\varepsilon > 0$, there is a compact set F such that $F \subseteq A$ and $\mu(A \setminus F) < \varepsilon$. A standard space (X, μ) is *separable* if there is an admissible topology in which X has a countable base.

Lemma 8.1. *Let $Z \times W \subseteq \bigcup_{i=1}^n (A_i \times B_i)$ for $A_i \subseteq X$, $B_i \subseteq Y$ and $n < \infty$. Then there are finite families of disjoint sets $\{X_p\}_{p=1}^m$ in Z and $\{Y_j\}_{j=1}^k$ in W such that* *each* $X_p \times Y_j$ *is contained in at least one of* $A_i \times B_i$ *and*

$$Z = \bigcup_{p=1}^{m} X_p, \quad W = \bigcup_{j=1}^{k} Y_j.$$

Proof. When z spans Z, there is only a finite number of different sets

$$A_z = Z \cap \Big(\bigcap_{z \in A_i} A_i \Big) \cap \Big(\bigcap_{z \notin A_i} (Z \setminus A_i) \Big).$$

Denote them by $X_1, \ldots, X_m$. Choosing, similarly, sets $Y_1, \ldots, Y_k$ in W, we obtain the sets $\{X_p\}$, $\{Y_j\}$ satisfying all conditions of the lemma. □

Denote by χ_E the characteristic function of a set E. We say that a function g on $X \times Y$ is *simple* if there are measurable, disjoint sets $\{X_i\}_{i=1}^n$, $\{Y_j\}_{j=1}^m$, with $n, m < \infty$, such that

$$X = \bigcup_{i=1}^{n} X_i, \quad Y = \bigcup_{j=1}^{m} Y_j, \quad \text{and} \quad g = \sum_{i,j} \alpha_{ij} \chi_{X_i} \chi_{Y_j} \ \text{ with } \ \alpha_{ij} \in \mathbb{C}.$$

Let φ be a function on $X \times Y$ and let $Z \subseteq X$, $W \subseteq Y$ be measurable. Set

$$\lambda(\varphi, Z \times W) = \sup\big\{ |\varphi(x, y) - \varphi(x', y')| : x, x' \in Z, \ y, y' \in W \big\}.$$

For $\varepsilon > 0$, a function φ is called ε-*decomposable* on $Z \times W$ if there are measurable sets $\{X_i\}_{i=1}^n$, $\{Y_j\}_{j=1}^m$, with $n, m < \infty$, such that

$$Z \subseteq \bigcup_{i=1}^{n} X_i, \quad W \subseteq \bigcup_{j=1}^{m} Y_j, \quad \text{and} \quad \lambda(\varphi, X_i \times Y_j) < \varepsilon \ \text{ for all } i, j.$$

Theorem 8.2. *Let* (X, μ) *and* (Y, ν) *be standard finite measure spaces. For a function* φ *on* $X \times Y$, *the following conditions are equivalent.*

(i) φ *is* ω*-continuous.*

(ii) *For each* $\varepsilon > 0$, *there are measurable sets* X_ε *and* Y_ε *such that* $\mu(X \setminus X_\varepsilon) < \varepsilon$, $\nu(Y \setminus Y_\varepsilon) < \varepsilon$ *and* φ *is* ε*-decomposable on* $X_\varepsilon \times Y_\varepsilon$.

(iii) *For each* $\varepsilon > 0$, *there are measurable sets* X_ε *and* Y_ε *such that* $\mu(X \setminus X_\varepsilon) < \varepsilon$, $\nu(Y \setminus Y_\varepsilon) < \varepsilon$ *and* $\varphi | X_\varepsilon \times Y_\varepsilon$ *is a uniform limit of simple functions.*

Proof. (i) $\Longrightarrow$ (ii). Choosing admissible topologies on X, Y and compacts $Q \subset X$ and $K \subset Y$ such that $\mu(X \setminus Q) < \varepsilon/2$ and $\nu(Y \setminus K) < \varepsilon/2$, we only need to prove the implication for Q, K and $\varepsilon/2$. Thus we may assume that X and Y are compacts in these topologies.

Cover the range of φ by open disks D^k of radius $\varepsilon/2$. Since φ is continuous, the sets $\varphi^{-1}(D^k)$ are ω-open. Hence, for each k, there are marginally null sets N_k

in $X \times Y$ and measurable sets A_i^k in X and B_i^k in Y such that

$$\varphi^{-1}(D^k) = N_k \cup \bigcup_{i=1}^{\infty} (A_i^k \times B_i^k)$$

Thus

$$X \times Y = N \cup \bigcup_{i=k=1}^{\infty} (A_i^k \times B_i^k),$$

where N is a marginally null set, and $\lambda(\varphi, A_i^k \times B_i^k) < \varepsilon$, for all i, k.

Choose $U \subset X$, $V \subset Y$ such that

$$N \subseteq (X \times V) \cup (U \times Y), \quad \mu(U) \leq \varepsilon/2, \quad \nu(V) \leq \varepsilon/2.$$

The set $\kappa = (X \setminus U) \times (Y \setminus V)$ is ω-closed and

$$\kappa \subseteq \bigcup_{i=k=1}^{\infty} (A_i^k \times B_i^k).$$

By [Erdos et al. 1998, Lemma 3.4], there are sets $R_\varepsilon \subset X$ and $T_\varepsilon \subset Y$, with $\mu(X \setminus R_\varepsilon) < \varepsilon/2$ and $\nu(Y \setminus T_\varepsilon) < \varepsilon/2$, such that the set

$$\kappa \cap (R_\varepsilon \times T_\varepsilon) = (R_\varepsilon \setminus U) \times (T_\varepsilon \setminus V)$$

is covered by a finite number of the rectangles $A_i^k \times B_i^k$. Setting $X_\varepsilon = R_\varepsilon \setminus U$, $Y_\varepsilon = T_\varepsilon \setminus V$, we obtain what we need.

(ii) $\Longrightarrow$ (iii). For $\varepsilon_n = 2^{-n}\varepsilon$, choose X_{ε_n}, Y_{ε_n} as in (ii): $\mu(X \setminus X_{\varepsilon_n}) < \varepsilon_n$, $\nu(Y \setminus Y_{\varepsilon_n}) < \varepsilon_n$ and $X_{\varepsilon_n} \times Y_{\varepsilon_n}$ is covered by a finite family of rectangles

$$\{A_j^n \times B_j^n\}_{j=1}^{p(n)}, \quad \text{with } \lambda(\varphi, A_j^n \times B_j^n) < \varepsilon_n.$$

Set

$$X_\varepsilon = \bigcap_{n=1}^{\infty} X_{\varepsilon_n}, \quad Y_\varepsilon = \bigcap_{n=1}^{\infty} Y_{\varepsilon_n}.$$

Then $\mu(X \setminus X_\varepsilon) < \varepsilon$, $\nu(Y \setminus Y_\varepsilon) < \varepsilon$, and, for each n,

$$X_\varepsilon \times Y_\varepsilon \subseteq X_{\varepsilon_n} \times Y_{\varepsilon_n} \subseteq \bigcup_{j=1}^{p(n)} (A_j^n \times B_j^n).$$

It follows from Lemma 8.1 that there is a simple function φ_n on $X \times Y$ such that $\sup\{|\varphi(x, y) - \varphi_n(x, y)| : (x, y) \in X_\varepsilon \times Y_\varepsilon\} < \varepsilon_n$.

(iii) $\Longrightarrow$ (i). Every simple function is ω-continuous. By [Erdos et al. 1998, Lemma 3.3], the uniform limit of ω-continuous functions is ω-continuous. Hence, for each $\varepsilon > 0$, the function $\varphi|X_\varepsilon \times Y_\varepsilon$ is ω-continuous. The set

$$N = (X \times Y) \setminus \bigcup_{n=1}^{\infty} \big(X_{1/n} \times Y_{1/n}\big)$$

is marginally null. Hence, for any open set $G \subset \mathbb{C}$,

$$\varphi^{-1}(G) \setminus \bigcup_{n=1}^{\infty} (\varphi^{-1}(G) \cap (X_{1/n} \times Y_{1/n})) = \varphi^{-1}(G) \cap N$$

is a marginally null set. Since all $\varphi^{-1}(G) \bigcap (X_{1/n} \times Y_{1/n})$ are ω-open, $\varphi^{-1}(G)$ is ω-open and φ is ω-continuous. □

A sequence $\{X_n\}$ of measurable sets in (X, μ) is *exhaustive* if

$$X_n \subseteq X_{n+1} \quad \text{and} \quad \mu\Big(X - \bigcup_{n=1}^{\infty} X_n\Big) = 0.$$

Fix an admissible topology on a standard measure space (X, μ). Then X has an exhaustive sequence $\{X_n\}$ of compact sets. For each n, there are disjoint compacts $\{K_i(n)\}$ in X_{n+1} such that $K_1(n) = X_n$ and $\mu(X_{n+1} - \bigcup_i K_i(n)) = 0$. Hence there are disjoint compact sets $\{K_n\}$ in X such that $\mu(X - \bigcup_n K_n) = 0$.

The following result can be considered as an ω-version of Luzin's theorem.

Theorem 8.3. *Let μ and ν be Radon σ-finite measures on topological spaces X and Y. For a function φ on $X \times Y$ the following conditions are equivalent.*

(i) *φ is ω-continuous.*

(ii) *For any $\varepsilon > 0$, there are measurable sets $X_\varepsilon \subseteq X$ and $Y_\varepsilon \subseteq Y$ such that $\mu(X - X_\varepsilon) < \varepsilon$, $\nu(Y - Y_\varepsilon) < \varepsilon$ and φ is continuous on $X \times Y$.*

(iii) *There are exhaustive sequences $\{X_n\}$ and $\{Y_n\}$ of compacts in X and Y such that φ is continuous on each $X_n \times Y_n$.*

Proof. Step 1. First we will prove the theorem for compact X and Y.

(i) $\Rightarrow$ (ii). Let E be a measurable subset of X. By Luzin's theorem, for $\delta > 0$, there is a compact subset K of X such that $\mu(X \setminus K) < \delta$ and χ_E is continuous on K. Hence if g is a simple function on $X \times Y$, there are compacts $K \subseteq X$, $R \subseteq Y$ such that $\mu(X \setminus K) < \delta$, $\nu(Y \setminus R) < \delta$ and g is continuous on $K \times R$.

Let φ be ω-continuous. For $\varepsilon > 0$, let sets X_ε and Y_ε be chosen as in Theorem 8.2 (iii) and let simple functions φ_n uniformly converge to $\varphi|X_\varepsilon \times Y_\varepsilon$. Set $\varepsilon_n = 2^{-n}\varepsilon$. By the argument above, there are compacts $K_n \subseteq X$, $R_n \subseteq Y$ such that $\mu(X \setminus K_n) < \varepsilon_n$, $\nu(Y \setminus R_n) < \varepsilon_n$ and the functions φ_n are continuous on $K_n \times R_n$. Set

$$L(\varepsilon) = X_\varepsilon \cap \bigcap_{n=1}^{\infty} K_n \quad \text{and} \quad M(\varepsilon) = Y_\varepsilon \cap \bigcap_{n=1}^{\infty} R_n.$$

Then $\mu(X \setminus L(\varepsilon)) \leq 2\varepsilon$ and $\nu(Y \setminus M(\varepsilon)) \leq 2\varepsilon$. All φ_n are continuous on $L(\varepsilon) \times M(\varepsilon)$ and uniformly converge to $\varphi|L(\varepsilon) \times M(\varepsilon)$. Hence φ is continuous on $L(\varepsilon) \times M(\varepsilon)$.

(ii) $\Rightarrow$ (iii). Set

$$L_n = \bigcap_{k=n}^{\infty} L(\varepsilon_k) \quad \text{and} \quad M_n = \bigcap_{k=n}^{\infty} M(\varepsilon_k).$$

Then $L_n \subseteq L_{n+1}$, $M_n \subseteq M_{n+1}$, and φ is continuous on $L_n \times M_n$. Furthermore,

$$\mu(X \setminus L_n) \leq \sum_{k=n}^{\infty} \mu(X \setminus L(\varepsilon_k)) < \varepsilon 2^{2-n},$$

$$\nu(Y \setminus M_n) \leq \sum_{k=n}^{\infty} \nu(Y \setminus M(\varepsilon_k)) < \varepsilon 2^{2-n}.$$

Thus $\{L_n\}$, $\{M_n\}$ are exhaustive sequences. Since μ, ν are Radon measures, there are compacts $E_n \subseteq L_n$ and $F_n \subseteq M_n$ such that $\mu(L_n \setminus E_n) < 1/n$, $\nu(M_n \setminus F_n) < 1/n$. Hence

$$X_n = \bigcup_{k=1}^{n} E_k \quad \text{and} \quad Y_n = \bigcup_{k=1}^{n} F_k$$

form exhaustive sequences of compacts in X and Y.

Step II. Now assume that X and Y are not compact spaces. Let $\{F_n\}$ and $\{G_n\}$ be disjoint compact sets in X and Y such that $\mu(X \setminus \bigcup_n F_n) = \nu(Y \setminus \bigcup_n G_n) = 0$. For $\varepsilon > 0$, set $\varepsilon_n = \varepsilon 2^{-n}$.

(i) $\Rightarrow$ (ii). It follows from step I that, for each pair (n, m), there are sets $R_{n,m}(\varepsilon) \subset F_n$ and $T_{n,m}(\varepsilon) \subset G_m$ such that φ is continuous on $R_{n,m}(\varepsilon) \times T_{n,m}(\varepsilon)$, $\mu(F_n \setminus R_{n,m}(\varepsilon)) \leq \varepsilon_n \varepsilon_m$, and $\nu(G_m \setminus T_{n,m}(\varepsilon)) \leq \varepsilon_n \varepsilon_m$. Set

(8–1) $$R_n(\varepsilon) = \bigcap_{m=1}^{\infty} R_{n,m}(\varepsilon) \quad \text{and} \quad T_m(\varepsilon) = \bigcap_{n=1}^{\infty} T_{n,m}(\varepsilon).$$

Then $\mu(F_n \setminus R_n(\varepsilon)) \leq \varepsilon_n \varepsilon$, $\mu(G_m \setminus T_m(\varepsilon)) \leq \varepsilon_m \varepsilon$, and the map φ is continuous on $R_n(\varepsilon) \times T_m(\varepsilon)$ for each pair (n, m). Set

(8–2) $$X_\varepsilon = \bigcup_{n=1}^{\infty} R_n(\varepsilon) \quad \text{and} \quad Y_\epsilon = \bigcup_{m=1}^{\infty} T_m(\varepsilon).$$

These are the sets we need.

(ii) $\Rightarrow$ (iii). We preserve the notations above. Let $\varepsilon_k = 2^{-k}$. It follows from step I that, for each pair (n, m), there are increasing sequences of compact sets $\{R_{n,m}(\varepsilon_k)\}_{k=1}^{\infty}$ in F_n and $\{T_{n,m}(\varepsilon_k)\}_{k=1}^{\infty}$ in G_m such that

$$\mu(F_n \setminus R_{n,m}(\varepsilon_k)) \leq \varepsilon_n \varepsilon_m \varepsilon_k, \qquad \nu(G_m \setminus T_{n,m}(\varepsilon_k)) \leq \varepsilon_n \varepsilon_m \varepsilon_k,$$

and φ is continuous on $R_{n,m}(\varepsilon_k) \times T_{n,m}(\varepsilon_k)$. The compact sets $R_n(\varepsilon_k) \subseteq F_n$ and $T_m(\varepsilon_k) \subseteq G_m$ (see (8–1)) increase with k, $\mu(F_n \setminus R_n(\varepsilon_k)) \leq \varepsilon_n \varepsilon_k$, $\mu(G_m \setminus T_m(\varepsilon_k)) \leq \varepsilon_m \varepsilon_k$, and φ is continuous on $R_n(\varepsilon_k) \times T_m(\varepsilon_k)$. The compact sets $X_k = X_{\varepsilon_k}$ and

$Y_k = Y_{\varepsilon_k}$ (see (8–2)) form exhaustive sequences in X and Y, and φ is continuous on each $X_k \times Y_k$.

The proof of (iii) $\Rightarrow$ (i) is the same as in Theorem 8.2. □

9. Schur multipliers and discontinuous (μ, ν)-multipliers

In contrast to Schur multipliers, (μ, ν, I)-multipliers are not sensitive to changes on null sets. Therefore one cannot expect that the classes of noncontinuous (μ, ν, I)-multipliers and of Schur I-multipliers coincide. In this section we will show that for any I, ω-continuous (μ, ν, I)-multipliers coincide marginally a.e. with Schur I-multipliers. More precisely, an ω-continuous function is a (μ, ν, I)-multiplier if and only if it becomes a Schur I-multiplier after deleting a marginally null subset.

Remark. Two ω-continuous functions φ, φ' coincide a.e. if and only if they coincide marginally a.e. Indeed, set $\psi = \varphi - \varphi'$. If $\psi \equiv 0$ marginally a.e., then $\psi \equiv 0$ a.e. Suppose that ψ vanishes a.e. The set $L = \{z \in \mathbb{C} : \psi(z) \neq 0\}$ is ω-open and $(\mu \otimes \nu)(L) = 0$. Therefore it coincides with some union of rectangles $A_n \times B_n$ up to a marginally null set. Hence $\mu(A_n)\nu(B_n) = 0$, so all $A_n \times B_n$ are marginally null. Thus L is a marginally null set.

Our restriction to ω-continuous functions is strongly motivated by the following result.

Proposition 9.1. *If* (X, μ) *and* (Y, ν) *are standard measure spaces, then*

$$\mathfrak{M}_{\mu,\nu}(\mathfrak{S}_\infty) \subseteq C_{\mu,\nu}(X \times Y).$$

Proof. Choose admissible topologies on X and Y, so that $X = \bigcup_n X_n$ and $Y = \bigcup_n Y_n$, with $\mu(X_n) < \infty$ and $\nu(Y_n) < \infty$. Let $\varphi \in \mathfrak{M}_{\mu,\nu}(\mathfrak{S}_\infty)$ and let G be an open set in $\mathbb{C}$. Since

$$\varphi^{-1}(G) = \bigcup_n \left(\varphi^{-1}(G) \cap (X_n \times Y_n)\right),$$

we only need to show that each set $\varphi^{-1}(G) \cap (X_n \times Y_n)$ is ω-open. Hence we may assume that $\mu(X) < \infty$ and $\nu(Y) < \infty$.

Set $H = L_2(X, \mu)$ and $K = L_2(Y, \nu)$. All results of Proposition 4.2 hold if $\mathcal{L}(I)$ is replaced by $\mathfrak{M}_{\mu,\nu}(I)$. Hence $\varphi \in \mathfrak{M}_{\mu,\nu}(\mathfrak{S}_1)$. The operator A with kernel $a(x, y) \equiv 1$ is a rank one operator. Hence the operator $\Phi_\varphi(A)$ with kernel $\varphi(x, y)$ belongs to $\mathfrak{S}_1(H, K)$. Hence $\varphi(x, y)$ belongs to the projective tensor product $H \hat{\otimes} K$ and is ω-continuous by Theorem 6.5 of [Erdos et al. 1998]. □

Let $\{X_n\}$, $\{Y_n\}$ be exhaustive sequences of measurable subsets of (X, μ) and (Y, ν), and let χ_n and χ'_n be the characteristic functions of X_n and Y_n. Let μ_n and

ν_n be the restrictions of μ and ν to X_n and Y_n. For $\varphi \in L_\infty(X \times Y, \mu \otimes \nu)$, set

$$\varphi_n = \chi_n \chi_n' \varphi \quad \text{and} \quad \widehat{\varphi}_n = \varphi | X_n \times Y_n.$$

Lemma 9.2. (i) *Let an s.n. ideal I be either coseparable, or contain $\mathfrak{S}_2$. Then φ is a (μ, ν, I)-multiplier if and only if all φ_n are (μ, ν, I)- multipliers and* $\sup_n \|\Phi_{\varphi_n}\|_I < \infty$. *In this case*

$$\|\Phi_\varphi\|_I = \sup_n \|\Phi_{\varphi_n}\|_I.$$

(ii) *Let I be either a separable or coseparable s.n. ideal, or $\mathfrak{S}_b$. Let $X = \bigcup_n X_n$ and $Y = \bigcup_n Y_n$. Then φ is a Schur I-multiplier if and only if all φ_n are Schur I-multipliers and* $\sup_n \|S_{\varphi_n}\|_I < \infty$. *In this case* $\|S_\varphi\|_I = \sup_n \|S_{\varphi_n}\|_I$.

(iii) *φ_n is a (μ, ν, I)-multiplier on $X \times Y$ if and only if $\widehat{\varphi}_n$ is a (μ_n, ν_n, I)-multiplier on $X_n \times Y_n$. Moreover, $\|\Phi_{\varphi_n}\|_I = \|\Phi_{\widehat{\varphi}_n}\|_I$.*

(iv) *φ_n is a Schur I-multiplier on $X \times Y$ if and only if $\widehat{\varphi}_n$ is a Schur I-multiplier on $X_n \times Y_n$. Moreover, $\|S_{\varphi_n}\|_I = \|S_{\widehat{\varphi}_n}\|_I$.*

Proof. Set $\Phi = \Phi_\varphi$ and $\Phi_n = \Phi_{\varphi_n}$. The operators P_n on $H = L_2(X, \mu)$ (identified with H^d as usual) and Q_n on $K = L_2(Y, \nu)$, acting by the multiplication by χ_n and χ_n', respectively, are projections, and

$$\text{(9–1)} \qquad \Phi_n(R) = Q_n \Phi(R) P_n, \qquad \text{for } R \in \mathfrak{S}_2(H, K).$$

Since P_n and Q_n strongly converge to the identity operators, $\Phi_n(R)$ strongly converge to $\Phi(R)$.

Let I be coseparable. If φ_n are (μ, ν, I)-multipliers, then $\Phi_n(R) \in I$, for $R \in \mathscr{X}_I$. If $\sup_n \|\Phi_n\|_I < \infty$, then $\sup_n |\Phi_n(R)|_I < \infty$, and it follows from Theorem III.5.1 of [Gohberg and Kreĭn 1965] that $\Phi(R) \in I$ and $|\Phi(R)|_I \le \sup_n |\Phi_n(R)|_I$. On the other hand, by (9–1), all

$$|\Phi_n(R)|_I \le \|Q_n\| |\Phi(R)|_I \|P_n\| \le |\Phi(R)|_I.$$

Hence $\|\Phi\|_I = \sup_n \|\Phi_n\|_I$. The proof of the converse statement follows from (9–1) immediately.

If $\mathfrak{S}_2 \subseteq I$, then $\mathfrak{S}_2 \subseteq I \subseteq J$, for some coseparable ideal J. Since all results of Proposition 4.2 hold, if $\mathscr{L}(I)$ is replaced by $\mathfrak{M}_{\mu,\nu}(I)$, the sets of (μ, ν, I)- and (μ, ν, J)-multipliers coincide. Part (i) is proved.

Let I be a coseparable ideal, let I_0 be the corresponding separable ideal and $(\widehat{I_0}, \widehat{I})$ be the pair of the dual ideals. It is well known and follows from duality (see, for example, [Kissin and Shulman 2005a, Lemma 5.1]) that the sets of Schur I-, I_0-, $\widehat{I}$- and $\widehat{I_0}$-multipliers (in particular, $\mathfrak{S}_b$-, $\mathfrak{S}_1$- and $\mathfrak{S}_\infty$-multipliers) coincide and the norms of the multipliers are equal. Hence we only need to prove (ii) for coseparable ideals. This proof is identical to the proof of (i) with Φ replaced by S, $L_2(X,\mu)$ by $l_2(X)$ and $L_2(Y,\nu)$ by $l_2(Y)$.

Set $H_n = L_2(X_n,\mu_n)$, $K_n = L_2(Y_n,\nu_n)$. For $R \in \mathfrak{S}_2(H,K)$ with kernel r, let $\delta_n(R)$ be the integral operator from H_n into K_n with kernel $\widehat{r} = r|X_n \times Y_n$. Then δ_n maps $\mathscr{X}_I(H,K)$ onto $\mathscr{X}_I(H_n,K_n)$, and it is an isometry from $Q_n\mathscr{X}_I(H,K)P_n$ onto $\mathscr{X}_I(H_n,K_n)$. Conversely, for $\widehat{R} \in \mathfrak{S}_2(H_n,K_n)$ with kernel $\widehat{r}$, let $\Delta_n(\widehat{R})$ be the integral operator from H into K with kernel r that vanishes outside $X_n \times Y_n$ and $r|X_n \times Y_n = \widehat{r}$. Then Δ_n is an isometry from $\mathscr{X}_I(H_n,K_n)$ onto $Q_n\mathscr{X}_I(H,K)P_n$, and

$$\delta_n(\Delta_n(\widehat{R})) = \widehat{R} \quad \text{for } \widehat{R} \in \mathscr{X}_I(H_n,K_n),$$
$$\Delta_n(\delta_n(R)) = R \quad \text{for } R \in Q_n\mathscr{X}_I(H,K)P_n.$$

We also have $\Phi_{\varphi_n}(R) \in Q_n\mathscr{X}_I(H,K)P_n$ for $R \in \mathscr{X}_I(H,K)$, $\delta_n\Phi_{\varphi_n} = \Phi_{\widehat{\varphi}_n}\delta_n$, $\Phi_{\varphi_n}\Delta_n = \Delta_n\Phi_{\widehat{\varphi}_n}$, and

$$\|\Phi_{\varphi_n}\|_I = \sup\{\left|\Phi_{\varphi_n}(R)\right|_I : R \in Q_n\mathscr{X}_I(H,K)P_n,\ |R|_I = 1\}.$$

Making use of these formulae, one obtains a proof of (iii). Part (iv) is evident. □

We will prove now an analogue of Theorem 7.5 for ω-continuous functions.

Theorem 9.3. *Let I be either a coseparable ideal, or a separable ideal containing $\mathfrak{S}_2$. Let (X,μ) and (Y,ν) be standard measure spaces with countable bases. An ω-continuous function φ on $X \times Y$ is a (μ,ν,I)-multiplier if and only if there are null sets $X_0 \subset X$, $Y_0 \subset Y$ such that φ is a Schur I-multiplier on $(X \setminus X_0) \times (Y \setminus Y_0)$. In this case the sets X_0, Y_0 can be chosen in such a way that $\|\varphi\|_I^{\mu,\nu} = \|S_{\tilde{\varphi}}\|_I$, where $\tilde{\varphi} = \varphi|(X \setminus X_0) \times (Y \setminus Y_0)$.*

Proof. Choose admissible topologies on X and Y. By Theorem 8.3, there are exhaustive sequences $\{X_n\}$ and $\{Y_n\}$ of compact sets in X and Y such that φ is continuous on each $X_n \times Y_n$. Let μ_n and ν_n be the restrictions of μ and ν to X_n and Y_n. One can assume that $\mathrm{supp}(\mu_n) = X_n$ and $\mathrm{supp}(\nu_n) = Y_n$. Indeed, set $K_n = \mathrm{supp}(\mu_n)$. If $K_n \neq X_n$, replace X_n by K_n. If $x \in K_n$ then, for each neighbourhood U_x of x in X, we have $\mu(X_n \cap U_x) \neq 0$. Hence $\mu(X_{n+1} \cap U_x) \neq 0$, so $x \in K_{n+1}$. Thus $K_n \subseteq K_{n+1}$. Since $X_n = K_n \cup N_n$ and $\mu(N_n) = 0$, the sequence $\{K_n\}$ is exhaustive and $\mathrm{supp}(\mu|K_n) = K_n$.

By Lemma 9.2(i), φ is a (μ,ν,I)-multiplier if and only if all its restrictions φ_n to $X_n \times Y_n$ are (μ,ν,I)-multipliers and the norms are bounded. Moreover, $\|\Phi_\varphi\|_I = \sup_n \|\Phi_{\varphi_n}\|_I$. Since $\widehat{\varphi}_n$ is continuous on $X_n \times Y_n$, it follows from Theorem 7.5 that $\widehat{\varphi}_n$ is a (μ_n,ν_n,I)-multiplier if and only $\widehat{\varphi}_n$ is a Schur I-multiplier on $X_n \times Y_n$; in this case $\|S_{\widehat{\varphi}_n}\|_I = \|\widehat{\varphi}_n\|_I^{\mu_n,\nu_n}$. By Lemma 9.2(ii), the restriction $\tilde{\varphi}$ of φ to $\left(\bigcup_n X_n\right) \times \left(\bigcup_n Y_n\right)$ is a Schur I-multiplier if and only if all φ_n are Schur I-multipliers and the norms are bounded. In this case, $\|S_{\widetilde{\varphi}}\|_I = \sup_n \|S_{\varphi_n}\|$. Taking now into account Lemma 9.2(iii) and (iv), we complete the proof. □

Remark. Theorem 9.3 allows us to deduce Peller's [1985] description of double operator integrable functions from Grothendieck's description of Schur $\mathfrak{S}_\infty$-multipliers. Indeed, let $\mathscr{E}$, $\mathscr{F}$ be spectral measures on locally compact spaces X and Y. Denote by π and ρ the representations of $C_0(X)$ and $C_0(Y)$ corresponding to $\mathscr{E}$ and $\mathscr{F}$. Then (see the discussion at the end of Section 6) the set of double operator integrable functions with respect to $\mathscr{E}$, $\mathscr{F}$ coincides with $\mathbf{M}^{\pi,\rho}$. Let μ, ν be scalar measures such that $\operatorname{supp}(\mathscr{E}) = \operatorname{supp}(\mu)$ and $\operatorname{supp}(\mathscr{F}) = \operatorname{supp}(\nu)$. Then it follows from Corollary 7.3 that $\mathbf{M}^{\pi,\rho} = \mathbf{M}^{\mu,\nu}$. By Proposition 9.1, every function $\varphi \in \mathbf{M}^{\mu,\nu}$ is ω-continuous. Hence, by Theorem 9.3, φ becomes a Schur $\mathfrak{S}_\infty$-multiplier after deleting some null subsets from X and Y. Applying [Pisier 2001, Theorem 5.1], we get $\varphi(x, y) = (a(x), b(y))$, where a, b are bounded Hilbert space valued functions. This is the first part of Peller's theorem. Furthermore, by the proof of Theorem 5.5 of [Pisier 2001], there are a probability space (T, τ) and bounded functions $a(x, t)$, $b(y, t)$ on $X \times T$ and $Y \times T$ such that $\varphi(x, y) = \int_T a(x, t)b(y, t)\, d\tau$. This is the second (much stronger) statement in Peller's result.

We now relate ω-continuous (μ, ν, I)-multipliers for different pairs of measures, just as we did for continuous (μ, ν, I)-multipliers in Corollary 7.4. Let μ be a σ-finite Radon measure on a topological space X and let Σ be the σ-algebra of all μ-measurable sets in (X, μ). Let a measure μ' on Σ be *absolutely continuous with respect to* μ, that is, $\mu(E) = 0$ implies $\mu'(E) = 0$ for $E \in \Sigma$. Then, for every μ-measurable subset Z of X, $\operatorname{supp}(\mu'|Z) \subseteq \operatorname{supp}(\mu|Z)$.

Theorem 9.4. *Let I be either a coseparable ideal, or $\mathfrak{S}_2 \subseteq I$. Let μ and ν be σ-finite Radon measures on topological spaces X and Y with countable bases. Let σ-finite measures μ' and ν' on X and Y be absolutely continuous with respect to μ and ν, respectively. Then every (μ, ν, I)-multiplier φ is also a (μ', ν', I)-multiplier and $\|\Phi_{\varphi,\mu',\nu'}\|_I \leq \|\Phi_{\varphi,\mu,\nu}\|_I$.*

Proof. By Theorem 8.3, there are exhaustive (with respect to μ and ν) sequences $\{X_n\}$ and $\{Y_n\}$ of compact sets in X and Y such that the functions

$$\widehat{\varphi}_n = \varphi | X_n \times Y_n$$

are continuous. Then $\{X_n\}$ and $\{Y_n\}$ are also exhaustive sequences with respect to μ' and ν'. Let μ_n and μ'_n be the restrictions of μ and μ' to X_n, and let ν_n and ν'_n be the restrictions of ν and ν' to Y_n. By Lemma 9.2(i) and (iii), the functions $\widehat{\varphi}_n$ are (μ_n, ν_n, I)-multipliers and $\|\Phi_{\varphi,\mu,\nu}\|_I = \sup_n \|\Phi_{\widehat{\varphi}_n,\mu_n,\nu_n}\|_I$.

Since $\operatorname{supp}(\mu'_n) \subseteq \operatorname{supp}(\mu_n)$ and $\operatorname{supp}(\nu'_n) \subseteq \operatorname{supp}(\nu_n)$, it follows from Corollary 7.4 that $\widehat{\varphi}_n$ are also (μ'_n, ν'_n, I)-multipliers and $\|\Phi_{\widehat{\varphi}_n,\mu'_n,\nu'_n}\|_I \leq \|\Phi_{\widehat{\varphi}_n,\mu_n,\nu_n}\|_I$. Applying again Lemma 9.2(i) and (iii), we conclude that φ is a (μ', ν', I)-multiplier and $\|\Phi_{\varphi,\mu',\nu'}\|_I \leq \|\Phi_{\varphi,\mu',\nu'}\|_I$. □

Acknowledgment

The authors are grateful to Charles Batty, Aristides Katavolos, Vladimir Peller, Ivan Todorov and, especially, to the late Barry Johnson for helpful discussions.

References

[Arazy 1978] J. Arazy, "Some remarks on interpolation theorems and the boundness of the triangular projection in unitary matrix spaces", *Integral Equations Operator Theory* **1**:4 (1978), 453–495. MR 81k:47056a Zbl 0395.47030

[Arveson 1974] W. Arveson, "Operator algebras and invariant subspaces", *Ann. of Math.* (2) **100** (1974), 433–532. MR 51 #1420 Zbl 0334.46070

[Bennett 1977] G. Bennett, "Schur multipliers", *Duke Math. J.* **44**:3 (1977), 603–639. MR 58 #12490 Zbl 0389.47015

[Birman and Solomyak 1967] M. S. Birman and M. Z. Solomyak, "Double Stieltjes operator integrals, II", pp. 26–60 in Проблемы математическоӥ физики, 2: спектральная теория, задачи дифракциа, edited by M. S. Birman, Izdat. Leningrad. Univ., Leningrad, 1967. MR 38 #2621 Zbl 0182.46202

[Birman and Solomyak 1973] M. S. Birman and M. Z. Solomyak, "Double Stieltjes operator integrals, III", pp. 27–53 in Проблемы математическоӥ физики, 6: теория функциӥ, спектральная теория, распространение волн, edited by M. S. Birman, Izdat. Leningrad. Univ., Leningrad, 1973. MR 50 #992 Zbl 0281.47013

[Birman and Solomyak 1989] M. S. Birman and M. Z. Solomyak, "Operator integration, perturbations and commutators", *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov.* **170** (1989), 34–66. Translation in *J. Sov. Math.* **63**:2, (1993), 129–148. MR 91b:47086 Zbl 0784.47033

[Boyd 1969] D. W. Boyd, "Indices of function spaces and their relationship to interpolation", *Canad. J. Math.* **21** (1969), 1245–1254. MR 54 #909 Zbl 0184.34802

[Erdos et al. 1998] J. A. Erdos, A. Katavolos, and V. S. Shulman, "Rank one subspaces of bimodules over maximal abelian selfadjoint algebras", *J. Funct. Anal.* **157**:2 (1998), 554–587. MR 99f:47054 Zbl 0977.47062

[Gohberg and Kreĭn 1965] I. C. Gohberg and M. G. Kreĭn, Введение в теорию линеиных несамосопряженных операторов в гилбертовом пространстве, Nauka, Moscow, 1965. MR 36 #3137 Zbl 0138.07803

[Graham and McGehee 1979] C. C. Graham and O. C. McGehee, *Essays in commutative harmonic analysis*, Grundlehren der Mathematischen Wissenschaften **238**, Springer, New York, 1979. MR 81d:43001 Zbl 0439.43001

[Grothendieck 1953] A. Grothendieck, "Résumé de la théorie métrique des produits tensoriels topologiques", *Bol. Soc. Mat. São Paulo* **8** (1953), 1–79. MR 20 #1194 Zbl 0074.32303

[Hadwin 1981] D. W. Hadwin, "Nonseparable approximate equivalence", *Trans. Amer. Math. Soc.* **266**:1 (1981), 203–231. MR 82e:46078 Zbl 0462.46039

[Hiai and Kosaki 2003] F. Hiai and H. Kosaki, *Means of Hilbert space operators*, Lecture Notes in Mathematics **1820**, Springer, Berlin, 2003. MR 2004h:47029 Zbl 1048.47001

[Kissin and Shulman 1996] E. V. Kissin and V. S. Shulman, "Operator-differentiable functions and derivations of operator algebras", *Funktsional. Anal. i Prilozhen.* **30**:4 (1996), 75–77. Translation in *Funct. Anal. Appl.* **30**:4 (1996), 280–282. MR 98c:46101 Zbl 0888.46020

[Kissin and Shulman 2005a] E. Kissin and V. S. Shulman, "Classes of operator-smooth functions, I: operator-Lipschitz functions", *Proc. Edinb. Math. Soc.* (2) **48**:1 (2005), 151–173. MR 2005i:47020 Zbl 1065.47010

[Kissin and Shulman 2005b] E. Kissin and V. S. Shulman, "Classes of operator-smooth functions, III: stable functions and Fuglede ideals", *Proc. Edinb. Math. Soc.* (2) **48**:1 (2005), 175–197. MR 2005i:47021 Zbl 1065.47011

[Mitjagin 1965] B. S. Mitjagin, "An interpolation theorem for modular spaces", *Mat. Sb.* (*N.S.*) **66 (108)** (1965), 473–482. MR 31 #1562 Zbl 0142.10701

[Peller 1985] V. V. Peller, "Hankel operators in the theory of perturbations of unitary and selfadjoint operators", *Funktsional. Anal. i Prilozhen.* **19**:2 (1985), 37–51, 96. Translation in *Funct. Anal. Appl.* **19**:2 (1985), 111–123. MR 87e:47029 Zbl 0587.47016

[Pisier 2001] G. Pisier, *Similarity problems and completely bounded maps*, Lecture Notes in Mathematics **1618**, Springer, Berlin, 2001. MR 2001m:47002 Zbl 0971.47016

[Sinclair and Smith 1995] A. M. Sinclair and R. R. Smith, *Hochschild cohomology of von Neumann algebras*, London Mathematical Society Lecture Note Series **203**, Cambridge University Press, Cambridge, 1995. MR 96d:46094 Zbl 0826.46050

[Smith 1991] R. R. Smith, "Completely bounded module maps and the Haagerup tensor product", *J. Funct. Anal.* **102**:1 (1991), 156–175. MR 93a:46115 Zbl 0745.46060

[Voiculescu 1976] D. Voiculescu, "A non-commutative Weyl–von Neumann theorem", *Rev. Roumaine Math. Pures Appl.* **21**:1 (1976), 97–113. MR 54 #3427 Zbl 0335.46039

[Wegge-Olsen 1993] N. E. Wegge-Olsen, *K-theory and C^*-algebras*, Oxford University Press, New York, 1993. MR 95c:46116 Zbl 0780.46038

Received December 11, 2004.

EDWARD KISSIN
DEPARTMENT OF COMPUTING, COMMUNICATIONS TECHNOLOGY AND MATHEMATICS
LONDON METROPOLITAN UNIVERSITY
HOLLOWAY ROAD
LONDON N7 8DB
GREAT BRITAIN
e.kissin@londonmet.ac.uk

VICTOR S. SHULMAN
DEPARTMENT OF MATHEMATICS
VOLOGDA STATE TECHNICAL UNIVERSITY
VOLOGDA
RUSSIA
shulman_v@yahoo.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 227, No. 1, 2006

AN EICHLER–ZAGIER MAP FOR JACOBI FORMS OF HALF-INTEGRAL WEIGHT

M. MANICKAM AND B. RAMAKRISHNAN

We construct an Eichler–Zagier map for Jacobi cusp forms of half-integral weight. As an application, we show there exists no Hecke-equivariant map from index 1 to index p (p prime), when the weight is half-integral.

The aim of this paper is to generalize the Eichler–Zagier map for Jacobi forms of half-integral weight, which is formally defined as

$$\mathfrak{Z}_m : \sum_{\substack{0>D,r\in\mathbb{Z}\\ D\equiv r^2\ (4m)}} c(D,r)\, e\left(\frac{r^2-D}{4m}\tau + rz\right) \mapsto \sum_{0>D\in\mathbb{Z}} \left(\sum_{\substack{r\ (2m)\\ r^2\equiv D\ (4m)}} c(D,r) \right) e(|D|\tau).$$

We prove that it is a Hecke-equivariant map from Jacobi cusp forms of weight $k+\frac{1}{2}$ on $\Gamma_0(4M)$, index m and character χ (k and χ are even) into a certain subspace of cusp forms of weight k on $\Gamma_1(16m^2M)$. First we derive this assertion for $m=1$ by proving that $\mathfrak{Z}_1$ maps respective Poincaré series. For the general index m, we apply certain operator I_m (see (2) for the definition) which changes the index m into index 1 and then apply $\mathfrak{Z}_1$ to obtain the required mapping property.

In order to give a Maass relation for each prime p for Siegel modular forms of half-integral weight and degree two, Y. Tanigawa [1986] obtained a Hecke-equivariant map from the space of index 1 Jacobi forms of half-integral weight into certain modular forms of integral weight and he constructed the map V_{p^2} from the space of Jacobi forms of index 1 into index p^2. As a natural question, he asked the existence of a connection between Jacobi forms of index 1 and index p (p is a prime) in the case of half-integral weight. We show that there is no such Hecke-equivariant map as an application of the nature of the map $\mathfrak{Z}_m$.

Notation and background. Throughout this paper, unless otherwise specified, the letters k, m, M, N will stand for natural numbers and τ for an element of $\mathcal{H}$, the complex upper half-plane.

MSC2000: primary 11F11, 11F50; secondary 11F37.
Keywords: Modular forms, Siegel modular forms, Jacobi forms.

For a complex number z, we write $\sqrt{z}$ for the square root with argument in $(-\pi/2, \pi/2]$, and we set $z^{k/2} = (\sqrt{z})^k$ for any $k \in \mathbb{Z}$.

For integers a, b, let $\left(\frac{a}{b}\right)$ denote the generalized quadratic residue symbol. Let $S_k(N, \psi)$ denote the space of cusp forms of weight k and level N with character ψ. We write the Fourier expansion of a modular form f as

$$f(\tau) = \sum_{n\geq 1} a_f(n) e^{2\pi i n\tau}.$$

For $z \in \mathbb{C}$ and $c, d \in \mathbb{Z}$, we put $e_d^c(z) = e^{2\pi i c z/d}$. We also write $e_1^c(z) = e^c(z)$, $e_c^1(z) = e_c(z)$, and $e_1^1(z) = e(z)$. The symbol $a \equiv \square\ (b)$ means that a is a square modulo b. For two forms f and g (either in the space of modular forms of integral weight or in the space of Jacobi forms of half-integral weight), $\langle f, g\rangle$ denotes the Petersson inner product of f and g. For a Dirichlet character ψ modulo $4m$, the twisting operator on modular forms of integral weight is given by

$$R_\psi = \frac{1}{W_\psi} \sum_{u \bmod 4m} \overline{\psi}(u) \begin{pmatrix} 4m & u \\ 0 & 4m \end{pmatrix}, \tag{1}$$

where $W_\psi = \sum_{u \bmod (4m)} \psi(u) e(u/4m)$. It follows that $\langle f \mid R_\psi, g\rangle = \langle f, g \mid R_{\overline{\psi}}\rangle$, where $f, g \in S_k(\Gamma_1(16mM))$ and

$$R_\psi : \sum_{n\geq 1} a_f(n) e(n\tau) \mapsto \sum_{n\geq 1} \psi(n) a_f(n) e(n\tau).$$

For a natural number d, the operators $U(d)$ and $B(d)$ are defined on formal power series by

$$U(d) : \sum_{n\geq 1} a(n) e(n\tau) \mapsto \sum_{n\geq 1} a(nd) e(n\tau),$$

$$B(d) : \sum_{n\geq 1} a(n) e(n\tau) \mapsto \sum_{n\geq 1} a(n) e(nd\tau).$$

For $n \geq 1$, let P_n denote the n-th Poincaré series in $S_k(N, \psi)$ whose ℓ-th Fourier coefficient is given by

$$g_n(\ell) = \delta(\ell, n) + 2\pi i^{-k} (\ell/n)^{(k-1)/2} \sum_{c\geq 1,\, N\mid c} K_{N,\chi}(n, \ell; c) J_{k-1}\left(\frac{4\pi\sqrt{n\ell}}{c}\right),$$

where $\delta(\ell, n)$ is the Kronecker-delta function, $J_{k-1}(x)$ is the Bessel function of order $k-1$ and $K_{N,\chi}(n, \ell; c)$ is the Kloosterman sum defined by

$$K_{N,\chi}(n, \ell; c) = \frac{1}{c} \sum_{\substack{d(c)^* \\ dd^{-1}\equiv 1\ (c)}} \overline{\psi}(d) e_c(nd^{-1} + \ell d).$$

1. A certain space of cusp forms of integral weight

For $m, M \in \mathbb{N}$, let $\chi \bmod M$ be a Dirichlet character and $\chi_m(n) = \left(\frac{m}{n}\right)$ be the quadratic character modulo m or $4m$ according as $m \equiv 1$ or $m \equiv 3 \pmod 4$.

Let

$$S = \left\{\ell \in \mathbb{N} : 1 \le \ell \le 4m,\ \ell \equiv \square\ (4m)\right\},$$
$$S^* = \left\{\ell \in S : p^2 \mid 4mM \text{ implies } p \nmid \ell, \text{ with } p \text{ prime}\right\}.$$

If $\ell \in S$, define

$$S_k^{\square,\ell}(16mM, \chi\chi_m) := S_k(16mM, \chi\chi_m) \,\Big|\, R_\ell,$$

where

$$R_\ell : \sum_{n\ge 1} a(n)e(n\tau) \mapsto \sum_{\substack{n\ge 1\\ -n\equiv \ell\ (4m)}} a(n)e(n\tau).$$

For $\ell \in S$, let $t = (\ell, 4m)$. A formal computation shows that

$$R_\ell = U(t)R(\ell)B(t),$$

with

$$R(\ell) = \frac{1}{\varphi(4m/t)} \sum_{\psi \bmod 4m/t} \overline{\psi}(-\ell/t)R_\psi,$$

where $\varphi(n)$ is the Euler totient function. Using the mapping properties of $U(t)$, R_ψ and $B(t)$ in the said order, we verify that $S_k^{\square,\ell}(16mM, \chi\chi_m)$ is a subspace of $S_k(\Gamma_1(16m^2M))$. Finally we define

$$S_k^{\square}(16mM, \chi\chi_m) = \sum_{\ell\in S} S_k^{\square,\ell}(16mM, \chi\chi_m).$$

2. Jacobi forms of half-integral weight

For $\alpha = \left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$, let $\tilde\alpha = (\alpha, \phi(\tau))$, where $\phi(\tau)$ is a holomorphic function on $\mathcal{H}$ such that $\phi^2(\tau) = t(c\tau + d)$, with $t \in \{1, -1\}$. Then the set $G = \{\tilde\alpha : \alpha \in \mathrm{SL}_2(\mathbb{R})\}$ is a group with group law

$$\left(\alpha_1, \phi_1(\tau)\right)\left(\alpha_2, \phi_2(\tau)\right) = \left(\alpha_1\alpha_2, \phi_1(\alpha_2\tau)\phi_2(\tau)\right).$$

If $\alpha \in \Gamma_0(4)$, set

$$j(\alpha, \tau) = \left(\frac{c}{d}\right)\left(\frac{-4}{d}\right)^{-1/2}(c\tau + d)^{1/2}.$$

We set $\alpha^* = \left(\alpha, j(\alpha, \tau)\right)$; the association $\alpha \mapsto \alpha^*$ is an injective map from $\Gamma_0(4)$ into G. Let G^J be the set of all triplets $[\tilde\alpha, X, s]$, $\alpha \in \mathrm{SL}_2(\mathbb{R})$, $X \in \mathbb{R}^2$, $s \in \mathbb{C}$, $|s| = 1$. Then G^J is a group, with group law given by

$$[\tilde\alpha_1, X_1, s_1]\,[\tilde\alpha_2, X_2, s_2] = \left[\tilde\alpha_1\tilde\alpha_2,\ X_1\alpha_2 + X_2,\ s_1s_2 \cdot \left(\det\begin{pmatrix} X_1\alpha_2\\ X_2\end{pmatrix}\right)\right].$$

The stroke operator $\big|_{k+1/2,m}$ is defined on functions $\phi : \mathcal{H} \times \mathbb{C} \to \mathbb{C}$ by

$$\phi\,\big|_{k+1/2,m}[\tilde{\alpha}, X, s] = \\ s^m \phi(\tau)^{-2k-1} e^m\left(\frac{-c(z+\lambda\tau+\mu)^2}{c\tau+d} + 2\lambda^2\tau + 2\lambda z + \lambda\mu\right)\phi\left(\frac{a\tau+b}{c\tau+d}, \frac{z+\lambda\tau+\mu}{c\tau+d}\right),$$

where $[\tilde{\alpha}, X, s] \in G^J$.

The Jacobi group for $\Gamma_0(4N)$ is a subgroup $\Gamma_0^J(4N)^*$ of G^J, given by

$$\Gamma_0^J(4N)^* = \left\{[\alpha^*, X] : \alpha \in \Gamma_0(4N),\, X \in \mathbb{Z}^2\right\}.$$

A *Jacobi form* $\phi(\tau, z)$ of weight $k + \frac{1}{2}$ and index m for the group $\Gamma_0(4M)$, with character χ, is a holomorphic function $\phi : \mathcal{H} \times \mathbb{C} \to \mathbb{C}$ satisfying the following conditions:

(i) $\phi\,\big|_{k+1/2,m}[\gamma^*, X](\tau, z) = \chi(d)\phi(\tau, z)$, where χ is a Dirichlet character mod $4M$ and $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(4M)$.

(ii) For every $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, the image $\phi\,\big|_{k+1/2,m}[\tilde{\alpha}, (0,0)](\tau, z)$ has a Fourier development of the form

$$\sum_{\substack{n,r\in\mathbb{Q} \\ r^2\le 4nm}} c_\alpha(n, r)\, e\,(n\tau + rz)\,,$$

where the sum ranges over rational numbers n, r with bounded denominators subject to the condition $r^2 \le 4nm$.

Further, if $r^2 < 4nm$ whenever $c_\alpha(n, r) \neq 0$, then ϕ is called a *Jacobi cusp form*. We denote by $J_{k+1/2,m}(4M, \chi)$ the space of Jacobi forms of weight $k + \frac{1}{2}$, index m for $\Gamma_0(4M)$ with character χ, and by $J^{\text{cusp}}_{k+1/2,m}(4M, \chi)$ the subspace of $J_{k+1/2,m}(4M, \chi)$ consisting of Jacobi cusp forms. A Jacobi form ϕ has a Fourier expansion of the form

$$\phi(\tau, z) = \sum_{\substack{n,r\in\mathbb{Z} \\ r^2\le 4nm}} c(n, r)\, e\,(n\tau + rz)\,.$$

Since $c(n, r) = c(n', r')$ if $r'^2 - 4n'm = r^2 - 4nm$ and $r' \equiv r \pmod{2m}$, we write the Fourier expansion of ϕ as

$$\phi(\tau, z) = \sum_{\substack{0\ge D,\, r\in\mathbb{Z} \\ D\equiv r^2\ (4m)}} c_\phi(D, r)\, e\left(\frac{r^2 - D}{4m}\tau + rz\right).$$

Let $D < 0$ be a discriminant and r an integer modulo $2m$ with $D \equiv r^2\ (4m)$. Then the (D, r)-th Poincaré series, denoted by $P_{(D,r)}$, is defined by

$$P_{(D,r)}(\tau, z) = \sum_{\gamma \in \Gamma_0(4M)^J_\infty \backslash \Gamma_0(4M)^J} \overline{\chi}(\gamma) e(n\tau + rz)\,\big|_{k+1/2,m}\gamma.$$

We state the following proposition without proof.

Proposition 2.1. *The Poincaré series $P_{(D,r)}$ lies in $J^{\mathrm{cusp}}_{k+1/2,m}(4M, \chi)$ and satisfies*

$$\langle \phi, P_{(D,r)} \rangle = \alpha_{k,m} |D|^{-k+1}\, c_\phi(D, r),$$

for each $\phi \in J^{\mathrm{cusp}}_{k+1/2,m}(4M, \chi)$, where $\alpha_{k,m} = \Gamma(k-1)m^{k-3/2}/(2\pi^{k-1})$. It has a Fourier development of the form

$$P_{(D,r)}(\tau, z) = \sum_{\substack{0 > D',\, r' \in \mathbb{Z} \\ D' \equiv r'^2\ (4m)}} \big(g_{D,r}(D', r') + \chi(-1) g_{D,r}(D', -r')\big)\, e\left(\frac{r'^2 - D'}{4m}\tau + r'z\right),$$

where $D = r^2 - 4mn$, $D' = r'^2 - 4mn'$, and $g_{D,r}(D', r')$ is given by

$$\delta_m(D, r, D', r') + i^{-k-3/2}\pi \sqrt{\frac{2}{m}} \left(\frac{D'}{D}\right)^{k/2} \sum_{\substack{c \geq 1 \\ 4M|c}} H_{m,c,\chi}(D, r, D', r')\, J_k\left(\frac{\pi\sqrt{DD'}}{mc}\right),$$

with

$$\delta_m(D, r, D', r') = \begin{cases} 1 & \text{if } D' = D \text{ and } r' \equiv r \ (\text{mod } 2m), \\ 0 & \text{otherwise.} \end{cases}$$

and

$$\begin{aligned} H_{m,c,\chi}(D, r, D', r') &= c^{-3/2} e^{-rr'/(2mc)} \\ &\quad \times \sum_{\substack{d,\, \lambda (c) \\ dd^{-1} \equiv 1\ (c)}} \overline{\chi}(d) \left(\frac{c}{d}\right)\left(\frac{-4}{d}\right)^{1/2} e_c\left(d^{-1}(m\lambda^2 + r\lambda + n) + dn' - \lambda r'\right). \end{aligned}$$

3. The Eichler–Zagier map

First we consider the space $J^{\mathrm{cusp}}_{k+1/2,1}(4M, \chi)$. Put $D = D_0\ell^2$, $r = r_0\ell$ in Proposition 2.1. In the Fourier coefficient of $P_{(D_0\ell^2, r_0\ell)}$, the Kloosterman-type sum is periodic as a function of ℓ of period $2c$. Hence, for any h (mod $2c$), its Fourier transform (after replacing ℓ by ℓd and λ by λd) becomes

$$\begin{aligned} \frac{1}{2c^{5/2}} \sum_{\substack{\ell (2c),\, d(c)^* \\ \lambda (c)}} &\overline{\chi}(d) \left(\frac{c}{d}\right)\left(\frac{-4}{d}\right)^{1/2} \\ &\times e_{2c}\big(d(2\lambda^2 + 2r_0\ell\lambda + 2n_0\ell^2 + 2n - 2r\lambda - r_0\ell r - h\ell)\big). \end{aligned}$$

Since $4|c$, the sum over λ is nonzero only if $r_0\ell \equiv r \pmod 2$. Hence, the sum over λ becomes

$$\sum_{\lambda(c)} e_c(d\lambda^2)\, e_c\Bigl(-d\Bigl(\frac{r_0\ell - r}{2}\Bigr)^2\Bigr).$$

Again, the fact that $4\,|\,c$ and $\gcd(c, d) = 1$ gives the identity

$$\frac{1}{\sqrt{2ic}} \sum_{\lambda(c)} e_c(d\lambda^2) = \Bigl(\frac{c}{d}\Bigr)\Bigl(\frac{-4}{d}\Bigr)^{-1/2}.$$

Thus, the Fourier transform simplifies to

$$\frac{\sqrt{i}}{\sqrt{2}c^2} \sum_{\ell(2c),d(c)^*} \overline{\chi}(d) e_{4c}\bigl(d(D_0\ell^2 + D - 2h\ell)\bigr) = \frac{\sqrt{i}}{4\sqrt{2}c^2} \sum_{\ell(2c),d(4c)^*} \overline{\chi}(d)\, e_{4c}\bigl(d(D_0\ell^2 + D - 2h\ell)\bigr),$$

which is the Fourier transform of the corresponding Kloosterman sum of integral weight.

More precisely:

Theorem 3.1. *The Eichler–Zagier map $\mathcal{Z}_1$ maps $J^{\text{cusp}}_{k+1/2,1}(4M,\chi)$ into $S^{\Box}_k(16M,\chi)$.*

Proof. We shall prove that the (D, r)-th Fourier coefficient of $P_{(D_0\ell^2,r_0\ell)}$ is equal (up to constant) $|D|$-th Fourier coefficient of $P_{|D_0|\ell^2}$. It is easy to see that

$$\delta_1(D_0\ell^2, r_0\ell, D, r) = \delta_{|D_0|\ell^2,|D|}.$$

We consider both the Kloosterman sums as periodic functions of period $2c$. The arguments put forth above shows that for each $c \geq 1$, with $4M|c$, the Fourier transform of $H_{1,c,\chi}(D_0\ell^2, r_0\ell, D, r)$ is equal to (up to the required constants) the Fourier transform of the Kloosterman sum (corresponding to integral weight) $K_{16M,\chi}(|D_0|\ell^2, |D|; 4c)$. This proves the theorem. □

The index-changing operator I_m. If $\phi \in J^{\text{cusp}}_{k+1/2,m}(4M, \chi)$, define I_m by

$$\phi \,\big|\, I_m(\tau, z) = \sum_{\lambda \pmod m} e(\lambda^2\tau + 2\lambda z)\,\phi(m\tau, z + \lambda\tau). \tag{2}$$

Proposition 3.2. *I_m maps $J^{\text{cusp}}_{k+1/2,m}(4M, \chi)$ into $J^{\text{cusp}}_{k+1/2,1}(4mM, \chi\chi_m)$. The Fourier development of $\phi \,\big|\, I_m$ is of the form*

$$\phi \,\big|\, I_m(\tau, z) = \sum_{\substack{0<D,\, r\in\mathbb{Z} \\ D\equiv r^2 \pmod 4}} \Bigl(\sum_{\substack{s \pmod{2m} \\ s\equiv r \pmod 2}} c_\phi(D, s)\Bigr)\, e\Bigl(\frac{r^2 - D}{4}\tau + rz\Bigr).$$

Proof. It is easy to see that

$$\phi \,\big|\, I_m(\tau, z) = m^{-k/2-1/4} \sum_{\lambda \pmod m} \phi_{1/\sqrt{m}} \,\big|_{k,1} \big[\, \widetilde{\Delta}_m,\ (\lambda, 0)\big]\,(\tau, z),$$

where $\phi_{1/\sqrt{m}}(\tau, z) = \phi(\tau, z/\sqrt{m})$ and Δ_m is the diagonal matrix $\operatorname{diag}\big(\sqrt{m}, 1/\sqrt{m}\big)$. The proposition now follows directly from the preceding expression. □

Using the equality $\mathfrak{X}_m = I_m \mathfrak{X}_1$, together with Theorem 3.1 and Proposition 3.2, we have:

Theorem 3.3. *The map $\mathfrak{X}_m$ takes $J^{\text{cusp}}_{k+1/2,m}(4M, \chi)$ into $S^{\square}_k(16mM, \chi\chi_m)$.*

4. Half-integral weight Jacobi forms of index 1 and index p

In the case of integral weight Jacobi forms, the well-known map V_p is a Hecke-equivariant map from $J_{k,1}$ into $J_{k,p}$ (p is a prime). If we replace k by $k+\frac{1}{2}$, then we have a Hecke-equivariant map V_{p^2} from $J_{k+1/2,1}(4M)$ into $J_{k+1/2,p^2}(4M)$, which was given by Tanigawa [1986]. Therefore, existence of a Hecke-equivariant map from index 1 into p in the case of half-integral weight Jacobi forms seems to be a natural question.

As an application of the map $\mathfrak{X}_m$, we show that there does not exist a Hecke-equivariant map from $J^{\text{cusp}}_{k+1/2,1}(4)$ into $J^{\text{cusp}}_{k+1/2,p}(4)$.

Let

$$N = \begin{cases} p & \text{if } p \equiv 1 \pmod 4, \\ p^2 & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

Let $\psi \pmod N$ be a primitive Dirichlet character such that $\psi^2 = \chi_p$. Let R_ψ be the twisting operator defined as in (1). Then, R_ψ maps $S_k(16N^2, \chi_p)$ into $S_k(16N^2)$ and commutes with Hecke operators T_n, $(n, p) = 1$. Further, if $f \in S_k(16N^2, \chi_p)$, we have

$$\big(f \,\big|\, R_\psi\big) \,\big|\, W_p = f \,\big|\, R_\psi,$$

where W_p is the W-operator on $S_k(16N^2)$ for the prime p.

***Case 1:* $p \equiv 3 \pmod 4$.** Let $f \in S_k(4p, \chi_p)$ be a normalized Hecke eigenform. Since $f \,\big|\, R_\psi \in S_k(4p^4)$ and it is an eigenform for all the Hecke operators and the W operators, it is a newform in $S^{\text{new}}_k(4p^4)$. Hence, by the theory of newforms, it is not equivalent to a level-1 Hecke eigenform.

***Case 2:* $p \equiv 1 \pmod 4$.** Let $f \in S^{\text{new}}_k(4p, \chi_p)$ be a normalized Hecke eigenform. Then, $f \,\big|\, R_\psi \in S^{\text{new}}_k(4p^2)$. Since $f \,\big|\, R_{\overline{\psi}} \in S^{\text{new}}_k(4p^2)$, and $\psi^3 = \overline{\psi}$ (as ψ^2 is quadratic), we get $f \,\big|\, R_\psi$ and $f \,\big|\, R_\psi \,\big|\, R_{\chi_p}$ are newforms in $S^{\text{new}}_k(4p^2)$. Thus, the form f is not equivalent to a level-1 Hecke eigenform. Now, we let $f \in S_k(p, \chi_p)$.

Arguments as above again show that f is not equivalent to a level-1 Hecke eigenform.

Thus, we conclude that a normalized Hecke eigenform in $S_k(4p, \chi_p)$ is not equivalent to a normalized Hecke eigenform in $S_k(4)$. In view of the mapping property proved in Theorem 3.3, we have proved:

Theorem 4.1. *There is no Hecke-equivariant map from the space* $J^{\text{cusp}}_{k+1/2,1}(4)$ *into the space* $J^{\text{cusp}}_{k+1/2,p}(4)$.

In this connection the following question seems natural.

What contribution do half-integral weight Jacobi forms of square-free index make to the construction of a "Maass space" (*if one exists*) *for degree*-2 *Siegel modular forms of half-integral weight*?

References

[Tanigawa 1986] Y. Tanigawa, "Modular descent of Siegel modular forms of half integral weight and an analogy of the Maass relation", *Nagoya Math. J.* **102** (1986), 51–77. MR 87i:11066 Zbl 0564.10031

Received November 1, 2004.

M. MANICKAM
mmmanickam@yahoo.com
DEPARTMENT OF MATHEMATICS
RKM VIVEKANANDA COLLEGE
MYLAPORE, CHENNAI 600 004
INDIA

B. RAMAKRISHNAN
ramki@mri.ernet.in
HARISH-CHANDRA RESEARCH INSTITUTE
CHHATNAG ROAD
JHUNSI, ALLAHABAD 211 019
INDIA

PACIFIC JOURNAL OF MATHEMATICS
Vol. 227, No. 1, 2006

THE TANGENT GRUPOID OF A HEISENBERG MANIFOLD

RAPHAËL PONGE

As a step towards proving an index theorem for hypoelliptic operators on Heisenberg manifolds, including for those on CR and contact manifolds, we construct an analogue for Heisenberg manifolds of Connes' tangent groupoid of a manifold. As is well known for a Heisenberg manifold (M, H) the relevant notion of tangent bundle is rather that of a Lie group bundle of graded 2-step nilpotent Lie groups GM. We define the tangent groupoid of (M, H) as a differentiable groupoid $\mathcal{G}_H M$ encoding the smooth deformation of $M \times M$ to GM. In particular, this construction makes a crucial use of a refined notion of privileged coordinates and of a tangent-approximation result for Heisenberg diffeomorphisms.

1. Introduction

A somewhat long standing open question is the existence of an index theorem for geometric operators on contact and CR manifolds. In this context the operators are not elliptic, so we cannot apply the classical index theorem of Atiyah–Singer [1968a; 1968b]. The natural pseudodifferential tool to deal with hypoelliptic operators on contact and CR manifolds is provided by the Heisenberg calculus of Beals–Greiner [1988] and Taylor [1984]. The latter holds in full generality for Heisenberg manifolds, that is, manifolds M together with a distinguished hyperplane bundle $H \subset TM$. This definition includes that of CR and contact manifolds, as well as that of codimension one foliations and confoliations. Therefore, what we would like to have is an analogue of the Atiyah–Singer theorem for hypoelliptic operators on Heisenberg manifolds.

There are various proofs of the Atiyah–Singer index theorem. A simple and fairly general proof is that of Connes [1994, Sect. II.5]. A salient feature in Connes' proof is the use of the tangent groupoid of a manifold, that is, the differentiable groupoid encoding the smooth deformation of $M \times M$ to TM (see [Connes 1994; Hilsum and Skandalis 1987]).

In this paper, as a step towards proving an index theorem for hypoelliptic operators on Heisenberg manifolds, we construct an analogue for Heisenberg manifolds

MSC2000: primary 58H05; secondary 53C10, 53D10, 32V05.
Keywords: differentiable groupoid, Heisenberg group, foliation, contact structure, CR structure.

of Connes' tangent groupoid. The existence of such a tangent groupoid was conjectured by Bellaïche [1996, p. 74] for Carnot–Carathéodory manifolds and by Ponge [2000, p. 37] for Heisenberg manifolds. We also refer to Van Erp [2005] for an alternative description of this groupoid.

Our construction of the tangent groupoid of a Heisenberg manifold is carried out in two steps. The first step consists in giving a suitable description of the tangent Lie group bundle GM of a Heisenberg manifold (M, H). The latter is a bundle of graded 2-step nilpotent Lie groups and provides us with a more accurate tangent structure for Heisenberg manifolds than the classical tangent tangent space TM. There are various descriptions of GM in the literature (see, e.g., [Bellaïche 1996; Beals and Greiner 1988; Epstein et al. 1991; Folland and Stein 1974; Gromov 1996; Rockland 1987]).

Our description of GM stems from the existence of an intrinsic real-valued Levi form,

$$\mathscr{L} : H \times H \longrightarrow TM/H. \tag{1-1}$$

This is a 2-form on H with values in the normal bundle TM/H (see Lemma 2.3). It allows us to define the tangent Lie group bundle GM as the bundle $(TM/H) \oplus H$ together with the grading and Lie group law, such that for sections X_0, Y_0 of TM/H and sections X', Y' of H, we have

$$t.(X_0 + X') = t^2 X_0 + tX', \qquad t \in \mathbb{R}, \tag{1-2}$$

$$(X_0 + X').(Y_0 + Y') = X_0 + Y_0 + \tfrac{1}{2}\mathscr{L}(X', Y') + X' + Y'. \tag{1-3}$$

This description of GM is simple and is completely intrinsic. What is crucial, and more difficult, in the construction of the tangent groupoid is to relate the above description to the extrinsic tangent nilpotent approximations of some previous approaches (see, e.g., [Bellaïche 1996; Beals and Greiner 1988; Epstein et al. 1991; Folland and Stein 1974; Gromov 1996; Rockland 1987]). More precisely, given a point $x \in M$ the tangent Lie group $G_x M$ in these approaches is obtained as the Lie group associated to a Lie algebra of model vector fields in some privileged coordinates centered at x. We point out that by using a refined notion of privileged coordinates, which we call Heisenberg coordinates (see Definition 2.18), this approach coincides with ours (Proposition 2.20).

An important consequence of the equivalence between these two descriptions of GM is a tangent approximation result for Heisenberg diffeomorphisms (Proposition 2.21). Namely, in Heisenberg coordinates a Heisenberg diffeomorphism is well approximated by a Lie group isomorphism between the tangent groups at the points. We really do need to use Heisenberg coordinates, because in general privileged coordinates we only get a Lie algebra isomorphism between the Lie algebras of the tangent group, and the corresponding Lie group isomorphism need not approximate the Heisenberg diffeomorphism (compare [Bellaïche 1996, Prop. 5.20]).

The second step is the construction of the tangent groupoid $\mathcal{G}_H M$ of a Heisenberg manifold (M, H) as the b-differentiable groupoid $\mathcal{G}_H M$ that encodes the smooth deformation of $M \times M$ to GM (Theorem 3.7). As an abstract groupoid the definition of $\mathcal{G}_H M$ is similar to that of Connes' tangent groupoid. In particular, at the set-theoretic level we have

$$\mathcal{G}_H M = GM \sqcup (M \times M \times (0, \infty)). \tag{1-4}$$

In order to endow $\mathcal{G}_H M$ with a consistent topology, with a differentiable structure, as well as with a smooth composition map, we make crucial uses of the Heisenberg coordinates and of the tangent approximation of Heisenberg diffeomorphisms alluded to above. In this sense our construction differs from the usual construction of Connes' tangent groupoid. In addition, this construction is functorial with respect to Heisenberg diffeomorphisms (see Proposition 3.8 for the precise statement).

Beside potential applications towards an index theorem for hypoelliptic operators on Heisenberg manifolds, the construction of the tangent groupoid $\mathcal{G}_H M$ is also interesting from the point of view of Carnot–Carathéodory geometry. Namely, Gromov [1996] and Bellaïche [1996] proved that the tangent group at a point of a Carnot–Carathéodory manifold is tangent to the manifold in a topological sense (i.e. in terms of Gromov–Hausdorff limits). However, our tangent groupoid construction shows that, in the special case of Heisenberg manifolds, this tangency occurs in a differentiable sense.

More generally, it would be interesting to construct a tangent groupoid for more general Carnot–Carathéodory manifolds. As mentionned earlier this has been conjectured by Bellaïche, but it is believed that the approach of this paper could be extended to deal with such a construction. Notice that in this setting the tangent Lie group bundle GM should rather be an orbifold-bundle of Lie groups, but it should be an actual Lie group bundle when the Carathéodory distribution is equiregular in the sense of [Gromov 1996]. We hope to address these issues in a subsequent paper.

The remainder of the paper is organized as follows. In Section 2, after recalling the main facts about Heisenberg manifolds, we describe the tangent group bundle GM of a Heisenberg manifold (M, H) and prove our approximation result for Heisenberg diffeomorphisms. In Section 3 we construct the tangent groupoid of (M, H) as a the differentiable groupoid that encodes the smooth deformation of $M \times M$ to GM.

Acknowledgements. I am grateful to Alain Connes, Pierre Julg, Henri Moscovici, Jean Renault and Erik Van Erp for discussions related to the subject matter of this paper. I also thank the hospitality of the IHÉS, where part of this paper was written.

2. The tangent Lie group bundle of a Heisenberg manifold

In this section, after recalling the main definitions and examples of Heisenberg manifolds, we describe the tangent Lie group bundle of a Heisenberg manifold in terms of an intrinsic Levi form. We then relate this approach to the nilpotent approximation of vector fields from previous approaches by using Heisenberg coordinates, which refine the privileged coordinates of [Beals and Greiner 1988] and [Bellaïche 1996]. As a consequence we get a tangent-approximation result for Heisenberg diffeomorphism which will be crucial later in the construction of the tangent groupoid of a Heisenberg manifold.

2.1. *Heisenberg manifolds.*

Definition 2.1. (1) A *Heisenberg manifold* is a smooth manifold M equipped with a distinguished hyperplane bundle $H \subset TM$.

(2) A *Heisenberg diffeomorphism* ϕ from a Heisenberg manifold (M, H) onto another Heisenberg manifold (M, H') is a diffeomorphism $\phi : M \to M'$ such that $\phi^* H = H'$.

Definition 2.2. Let (M^{d+1}, H) be a Heisenberg manifold. Then:

(1) A (*local*) *H-frame* for TM is a (local) frame $X_0, X_1, \ldots, X_d$ such that X_1, $\ldots$, X_d span H.

(2) A *local Heisenberg chart* is a local chart with a local H-frame of TM over its domain.

Following are the main examples of Heisenberg manifolds.

Heisenberg group. The $(2n+1)$-dimensional Heisenberg group $\mathbb{H}^{2n+1}$ consists in $\mathbb{R}^{2n+1} = \mathbb{R} \times \mathbb{R}^{2n}$ equipped with the group law,

$$x \cdot y = \Big(x_0 + y_0 + \sum_{1 \le j \le n} (x_{n+j} y_j - x_j y_{n+j}),\ x_1 + y_1, \ldots, x_{2n} + y_{2n}\Big).$$

A left-invariant basis for its Lie algebra $\mathfrak{h}^{2n+1}$ is provided by the vector fields,

$$X_0 = \partial_{x_0}, \qquad X_j = \partial_{x_j} + x_{n+j}\partial_{x_0}, \qquad X_{n+j} = \partial_{x_{n+j}} - x_j \partial_{x_0},$$

where j ranges over $1, \ldots, n$. In particular, for $j, k = 1, \ldots, n$ and $k \neq j$ we have the Heisenberg relations,

$$[X_j, X_{n+k}] = -2\delta_{jk} X_0, \qquad [X_0, X_j] = [X_j, X_k] = [X_{n+j}, X_{n+k}] = 0. \tag{2-1}$$

In particular, the subbundle spanned by the vector fields $X_1, \ldots, X_{2n}$ gives rise to a left-invariant Heisenberg structure on $\mathbb{H}^{2n+1}$.

Foliations. A (smooth) *foliation* is a manifold M together with a subbundle $\mathcal{F} \subset TM$ which is integrable in the Frobenius sense, that is, the space of sections of $\mathcal{F}$ is closed under the Lie bracket of vector fields. Thus any codimension-1 foliation is a Heisenberg manifold.

Contact manifolds. Opposite to foliations are *contact manifolds*. A contact manifold is a Heisenberg manifold (M^{2n+1}, H) such that H can be locally realized as the kernel of a contact form, that is, a 1-form θ such that $d\theta_{|H}$ is nondegenerate. When M is orientable it is equivalent to require H to be globally the kernel of a contact form. Furthermore, by Darboux's theorem any contact manifold is locally Heisenberg-diffeomorphic to the Heisenberg group $\mathbb{H}^{2n+1}$ equipped with the standard contact form $\theta^0 = dx_0 + \sum_{j=1}^{n}(x_j dx_{n+j} - x_{n+j} dx_j)$.

Confoliations. According to Eliashberg and Thurston [1998], a *confoliation* on an oriented manifold M^{2n+1} is given by a global nonvanishing 1-form θ on M such that $(d\theta)^n \wedge \theta \geq 0$. In particular, if we let $H = \ker\theta$ then (M, H) is a Heisenberg manifold which turns to be a foliation when $d\theta \wedge \theta = 0$ and a contact manifold when $(d\theta)^n \wedge \theta > 0$.

CR manifolds. A CR structure on an orientable manifold M^{2n+1} is given by a rank-n complex subbundle $T_{1,0} \subset T_{\mathbb{C}}M$ such that $T_{1,0}$ is integrable in Frobenius' sense and we have$T_{1,0} \cap T_{0,1} = \{0\}$, where we have let $T_{0,1} = \overline{T_{1,0}}$. Equivalently, the subbundle $H = \Re(T_{1,0} \oplus T_{0,1})$ has the structure of a complex bundle of (real) dimension $2n$. In particular, (M, H) is a Heisenberg manifold.

The main example of a CR manifold is that of a (smooth) boundary $M = \partial D$ of a complex domain $D \subset \mathbb{C}^n$. In particular, when D is strongly pseudoconvex (or strongly pseudoconcave) with defining function ρ then $\theta = i(\partial - \bar{\partial})\rho$ is a contact form on M.

2.2. ***The tangent Lie group bundle.*** The tangent Lie group bundle of a Heisenberg manifold (M^{d+1}, H) can be described as follows.

First, we have:

Lemma 2.3. *The Lie bracket on vector fields induces a TM/H-valued 2-form on H*

$$\mathcal{L} : H \times H \longrightarrow TM/H,$$

such that, for any sections X and Y of H near a point $m \in M$, we have

$$\mathcal{L}_m\big(X(m), Y(m)\big) = [X, Y](m) \mod H_m.$$

Proof. We only need to check that given two sections X and Y of H near $m \in M$ the value of $[X, Y](m)$ modulo H_m depends only on $X(m)$ and $Y(m)$. Indeed, if

f and g are smooth functions near m then we have

$$\begin{aligned}[fX, gY](m) &= f(m)g(m)[X,Y](m) - Y(f)(m)X(m) + X(g)(m)Y(m) \\ &= f(m)g(m)[X,Y](m) \mod H_m.\end{aligned}$$

This shows that if $X(m)$ or $Y(m)$ vanish then so does the class of $[X,Y](m)$ modulo H_m. Therefore, the latter only depends on the values of $X(m)$ and $Y(m)$. Hence the result. ☐

Definition 2.4. The 2-form $\mathscr{L}$ is called the *Levi form* of (M, H).

The Levi form $\mathscr{L}$ allows us to define a bundle $\mathfrak{g}M$ of graded Lie algebras by endowing $(TM/H)\oplus H$ with the smooth fields of Lie Brackets $[\,.\,,\,.\,]_{\mathfrak{g}M}$ and gradings $X \to t.X$, $t\in\mathbb{R}$, such that, for $m\in M$ and X_0, Y_0 in T_mM/H_m and X', Y' in H_m, we have

$$\begin{aligned}[X_0+X', Y_0+Y']_{\mathfrak{g}_mM} &= \mathscr{L}_m(X',Y'), \\ t.(X_0+X') &= t^2X_0+tX'.\end{aligned}$$

Definition 2.5. The bundle $\mathfrak{g}M$ is called the *tangent Lie algebra bundle* of M.

Proposition 2.6. *The tangent Lie algebra bundle is* 2*-step nilpotent and contains the normal bundle TM/H in its center.*

Proof. It follows from 2.2 that TM/H is contained in the center of $\mathfrak{g}M$ and that the Lie bracket $[\,.\,,\,.\,]_{\mathfrak{g}M}$ maps to TM/H, so $\mathfrak{g}M$ is 2-step nilpotent. ☐

Since $\mathfrak{g}M$ is nilpotent its associated graded Lie group bundle GM can be described as follows. As a bundle GM is $(TM/H)\oplus H$ and the exponential map is merely the identity. In particular the grading of GM is as in 2.2. Moreover, as $\mathfrak{g}M$ is 2-step nilpotent the Campbell–Hausdorff formula gives

$$(\exp X)(\exp Y) = \exp\big(X+Y+\tfrac{1}{2}[X,Y]\big) \qquad \text{for sections } X, Y \text{ of } \mathfrak{g}M.$$

We thus deduce that the product on GM is such that, for $m\in M$, and X_0, Y_0 in T_mM/H_m and X', Y' in H_m, we have

$$(X_0+X')\cdot(Y_0+X') = X_0+Y_0+\tfrac{1}{2}\mathscr{L}(X',Y')+X'+Y'. \tag{2-2}$$

Definition 2.7. The bundle GM is called the *tangent Lie group bundle* of M.

The fibers of GM are classified by the Levi form $\mathscr{L}$ as follows:

Proposition 2.8. (1) *The form $\mathscr{L}_m$ at $m\in M$ has rank $2n$ if, and only if, G_mM is isomorphic to $\mathbb{H}^{2n+1}\times\mathbb{R}^{d-2n}$ as a graded Lie group.*

(2) *The Levi form $\mathscr{L}$ has constant rank $2n$ if, and only if, GM is a fiber bundle with typical fiber $\mathbb{H}^{2n+1}\times\mathbb{R}^{d-2n}$.*

Proof. Let g be a Riemannian metric on H. Notice that (1) is a pointwise statement and that (2) is a local statement, since GM is a Lie group bundle already. Therefore, without any loss of generality, we may suppose that TM/H is orientable, so that it admits a global nonvanishing section X_0. Then let A denote the smooth section of End H such that

$$\mathscr{L}(X, Y) = g(X, AY)X_0 \qquad \text{for sections } X, Y \text{ of } H. \tag{2-3}$$

Let $m \in M$. Since $\mathscr{L}_m$ is real-antisymmetric its rank has to be an even integer, say $\operatorname{rk} \mathscr{L}_m = 2n$. Let us first assume that $\mathscr{L}_m$ is nondegenerate, i.e., A_m is invertible. Let $A_m = J_m|A_m|$ be the polar decomposition of A_m and on H_m define the positive definite scalar product,

$$h_m(X, Y) = \tfrac{1}{2} g_m(X, |A_m|Y) \qquad X, Y \in H_m. \tag{2-4}$$

Notice that J_m is antisymmetric and unitary with respect to h_m, so we have $J_m^2 = -J_m^t J_m = -1$, that is, J_m is a unitary complex structure on H_m. Therefore, we can construct a basis $X_1, \dots, X_{2n}$ of H_m which is orthonormal with respect to h_m and such that $X_{n+j} = J_m X_j$ for $j = 1, \dots, n$.

On the other hand, for X and Y in $H_m \subset \mathfrak{g}_m M$ we have

$$[X, Y]_{\mathfrak{g}_m M} = \mathscr{L}_m(X, Y) = g_m(X, A_m Y)X_0 = 2h_m(X, J_m Y)X_0. \tag{2-5}$$

Thus, for $j = 1, \dots, n$ and $k = 1, \dots, n+j-1, n+j+1, \dots, 2n$ we get

$$[X_j, X_{n+j}]_{\mathfrak{g}_m M} = 2h_m(X_j, J_m^2 X_j)X_0 = -2h_m(X_j, X_j)X_0 = -2X_0, \tag{2-6}$$

$$[X_j, X_k]_{\mathfrak{g}_m M} = h_m(X_j, J_m X_k)X_0 = -h_m(X_{n+j}, X_k)X_0 = 0. \tag{2-7}$$

These relations are the same as those in (2-1) for the Lie algebra of $\mathbb{H}^{2n+1}$. Thus $G_m M$ is isomorphic to $\mathbb{H}^{2n+1}$ as a graded Lie group.

Next, assume that A_m has a nontrivial kernel. Then as A_m is real antisymmetric with respect to g_m we have an orthogonal direct sum $H_m = \operatorname{im} A_m \oplus \ker A_m$. In fact, it follows from (2-3) that if $X \in \ker A_m$ and $Y \in H_m$ then

$$[X, Y]_{\mathfrak{g}_m M} = \mathscr{L}_m(X, Y) = g_m(X, A_m Y)X_0 = 0. \tag{2-8}$$

Thus $\ker A_m$ is contained in the center of $\mathfrak{g}_m M$. Moreover, as A_m is invertible on $\operatorname{im} A_m$ the same reasoning as above shows that the Lie subalgebra $(T_m M/H_m) \oplus \operatorname{im} A_m$ is isomorphic to the (graded) Lie algebra $\mathfrak{h}^{2n+1}$ of $\mathbb{H}^{2n+1}$. Therefore, $\mathfrak{g}_m M = (T_m M/H_m) \oplus \operatorname{im} A_m \oplus \ker A_m$ is isomorphic to $\mathfrak{h}^{2n+1} \times \mathbb{R}^{d-2n}$, and so $G_m M$ is isomorphic to $\mathbb{H}^{2n+1} \times \mathbb{R}^{d-2n}$.

Conversely, suppose that $G_m M$ is isomorphic to $\mathfrak{h}^{2n+1} \times \mathbb{R}^{d-2n}$. Then $\mathfrak{g}_m M$ is isomorphic to $\mathfrak{h}^{2n+1} \times \mathbb{R}^{d-2n}$, so admits a basis $X_0, \dots, X_d$ such that

$$[X_j, X_{n+j}]_{\mathfrak{g}_m M} = -2X_0 \quad \text{and} \quad [X_j, X_k]_{\mathfrak{g}_m M} = [X_l, X_k]_{\mathfrak{g}_m M} = 0, \tag{2-9}$$

for $j = 1, \dots, n$ and $k = 1, \dots, d$ with $k \neq n + j$ and $l = 2n + 1, \dots, d$. Since $\mathcal{L}_m(X, Y) = [X, Y]_{\mathfrak{g}_m M}$ for X and Y in H_m it follows from this that $\mathcal{L}_m$ has rank $2n$. The proof of the first part of the proposition is thus complete.

Now, suppose that $\mathcal{L}$ has constant rank $2n$. We have $\operatorname{rk} A_m = 2n$ everywhere, so we get a vector bundle splitting $H = \operatorname{im} A \oplus \ker A$. Furthermore, the polar decomposition of A_m is smooth with respect to m, i.e., J and $|A|$ are smooth sections of $\operatorname{End} H$. Therefore, the above process for constructing the basis $X_0, X_1, \dots, X_d$ can be carried out near every point $m \in M$ in such way to yield a smooth H-frame satisfying the relations (2-6)–(2-7). Thus, near every point of M we get a Lie group bundle trivialization of GM as a trivial fiber bundle with fiber $\mathbb{H}^{2n+1} \times \mathbb{R}^{d-2n}$. Consequently, GM is fiber bundle with typical fiber $\mathbb{H}^{2n+1} \times \mathbb{R}^{d-2n}$.

Conversely, assume that GM is a fiber bundle with typical fiber $\mathbb{H}^{2n+1} \times \mathbb{R}^{d-2n}$. Then at every point $m \in M$ the Lie group $G_m M$ is isomorphic to $\mathbb{H}^{2n+1} \times \mathbb{R}^{d-2n}$, so it follows from the first part of the proposition that $\mathcal{L}$ has constant rank $2n$. □

In presence of a foliation or contact structure we have more precise results.

Proposition 2.9. *Let (M, H) be a Heisenberg manifold. Then the following are equivalent*:

(1) *(M, H) is a foliation.*

(2) *(M, H) is Levi flat, i.e., $\mathcal{L}$ vanishes.*

(3) *As a Lie group bundle GM agrees with $(TM/H) \oplus H$.*

Proof. It follows from its definition that $\mathcal{L}$ vanishes if, and only if, for any sections X and Y of H the Lie bracket vector field $[X, Y]$ is again a section of H, that is, if, and only if, H defines a foliation.

On the other hand, in view of the definition of the group law of GM the Levi form $\mathcal{L}$ vanishes if, and only if, the group law is $X.Y = X + Y$, that is, if, and only if, GM is the Abelian Lie group bundle $(TM/H) \oplus H$. □

Proposition 2.10. *Suppose that (M^{2n+1}, H) is a Heisenberg manifold. Then the following are equivalent*:

(1) *(M, H) is a contact manifold.*

(2) *The Levi form $\mathcal{L}$ is (everywhere) nondegenerate.*

(3) *The Lie group tangent bundle GM is a fiber bundle with typical fiber $\mathbb{H}^{2n+1}$.*

Proof. Since the equivalence of (2) and (3) follows from Proposition 2.8, we only have to prove that (1) and (2) are equivalent. Since these are local statements we may assume that TM/H is orientable, i.e., there exists a global nonzero 1-form θ such that $H = \ker \theta$. As any nonzero 1-form annihiliting H is a nonzero multiple of θ we see that (M, H) is a contact manifold if, and only if, θ is a contact form.

Now, for any sections X and Y of H we have

$$\mathcal{L}(X,Y)=\theta([X,Y])X_0=-d\theta(X,Y)X_0. \tag{2-10}$$

This shows that $\mathcal{L}$ and $d\theta_{|_H}$ have same rank, so θ is a contact form if, and only if, $\mathcal{L}$ is everywhere nondegenerate. This proves the equivalence of (1) and (2) and thus completes the proof of the proposition. □

Finally, let $\phi:(M,H)\to(M',H')$ be a Heisenberg diffeomorphism from (M,H) onto another Heisenberg manifold (M',H'). Since we have $\phi_*H=H'$, we see that ϕ' induces a smooth vector bundle isomorphism $\bar{\phi}$ from TM/H onto TM'/H'.

Definition 2.11. We let $\phi'_H:(TM/H)\oplus H\to(TM'/H')\oplus H'$ denote the vector bundle isomorphism such that

$$\phi'_H(m)(X_0+X')=\bar{\phi}'(m)X_0+\phi'(m)X', \tag{2-11}$$

for any $m\in M$, $X_0\in T_m/H_m$, and $X'\in H_m$.

Proposition 2.12. *The vector bundle isomorphism ϕ'_H is an isomorphism of graded Lie group bundles from GM onto GM'.*

Proof. If X and Y are sections of H then we have

$$\mathcal{L}(\phi'_H(X),\phi'_H(Y))=[\phi_*X,\phi_*Y]=\phi'_*[X,Y]=\phi'_H(\mathcal{L}_m(X,Y))\mod H'. \tag{2-12}$$

In view of (2-2) this implies that ϕ'_H is a Lie group bundle isomorphism from GM onto GM'. Furthemore, it follows from (2-11) that, for any $t\in\mathbb{R}$ and any section X of GM, we have $\phi'_H(t.X)=t.\phi'_H(X)$, i.e., ϕ'_H is graded. □

Corollary 2.13. *The Lie group bundle isomorphism class of GM depends only on the Heisenberg diffeomorphism class of (M,H).*

2.3. *Heisenberg coordinates and nilpotent approximations of vector fields.* In the sequel it will be useful to combine the above intrinsic description of GM with a more extrinsic description of the tangent Lie group at a point in terms of the Lie group associated to a nilpotent Lie algebra of model vector field. Incidentally, this will show that our approach is equivalent to previous ones [Beals and Greiner 1988; Bellaïche 1996; Epstein et al. 1991; Folland and Stein 1974; Gromov 1996; Rockland 1987].

First, pick $m\in M$ and let us describe $\mathfrak{g}_mM$ as the graded Lie algebra of left-invariant vector field on G_mM by identifying any $X\in\mathfrak{g}_mM$ with the left-invariant vector field L_X on G_mM given by

$$L_Xf(x)=\frac{d}{dt}f\big(t\exp(X)\cdot x\big)\Big|_{t=0}=\frac{d}{dt}f(tX\cdot x)\Big|_{t=0},\qquad f\in C^\infty(G_mM).$$

This allows us to associate, to any vector field X near m, a unique left-invariant vector field X^m on G_mM such that

$$X^m = \begin{cases} L_{X_0(m)} & \text{if } X(m) \notin H_m, \\ L_{X(m)} & \text{otherwise,} \end{cases} \tag{2-13}$$

where $X_0(m)$ denotes the class of $X(m)$ modulo H_m.

Definition 2.14. The left-invariant vector field X^m is called the *model vector field* of X at m.

Let us look at this construction in terms of an H-frame $X_0, \ldots, X_d$ near m, i.e., of a local trivialization of the vector bundle $(TM/H) \oplus H$. For $j, k = 1, \ldots, d$ set

$$\mathcal{L}(X_j, X_k) = [X_j, X_k] = L_{jk} X_0 \mod H.$$

With respect to the coordinate system $(x_0, \ldots, x_d)$ corresponding to $X_0(m), \ldots, X_d(m)$ we can write the product law of G_mM as

$$x \cdot y = \big(x_0 + \tfrac{1}{2} \sum_{j,k=1}^{d} L_{jk} x_j y_k,\ x_1 + y_1, \ldots, x_d + y_d\big).$$

The vector fields X_j^m, $j = 1, \ldots, d$, in (2-13) are just the left-invariant vector fields corresponding to the vectors of the canonical basis $e_1, \ldots, e_d$, i.e., we have

$$X_0^m = \partial_{x_0} \quad \text{and} \quad X_j^m = \partial_{x_j} - \big(\tfrac{1}{2} \sum_{k=1}^{d} L_{jk} x_k\big) \partial_{x_0}, \quad 1 \leq j \leq d. \tag{2-14}$$

In particular, for $j, k = 1, \ldots, d$, we have the relations

$$[X_j^m, X_k^m] = L_{jk}(m) X_0^m \quad \text{and} \quad [X_j^m, X_0^m] = 0. \tag{2-15}$$

Let X be a vector field near m. Then X is of the form $X = a_0(x) X_0 + \cdots + a_d(x) X_d$ near m, and its model vector field X^m is thus given by the formula

$$X^m = \begin{cases} a_0(m) X_0^m & \text{if } a_0(m) \neq 0, \\ a_1(m) X_1^m + \cdots + a_d(m) X_d^m & \text{otherwise.} \end{cases} \tag{2-16}$$

Now, let $\kappa : \operatorname{dom} \kappa \to U$ be a Heisenberg chart near $m = \kappa^{-1}(u)$ and let $X_0, \ldots, X_d$ be the associated H-frame of TU. There exists a unique affine coordinate change $v \to \psi_u(v)$ such that $\psi_u(u) = 0$ and $\psi_{u*} X_j(0) = \partial_{x_j}$ for $j = 0, 1, \ldots, d$. Indeed, if for $j = 1, \ldots, d$ we set $X_j(x) = \sum_{k=0}^{d} B_{jk}(x) \partial_{x_k}$ then we have

$$\psi_u(x) = A(u)(x - u), \qquad \text{where } A(u) = \big(B(u)^t\big)^{-1}.$$

Definition 2.15 [Beals and Greiner 1988].

(1) The coordinates provided by ψ_u are called the *privileged coordinates* at u with respect to the H-frame $X_0, \dots, X_d$.

(2) The map ψ_u is called the *privileged coordinate map* at u with respect to the H-frame $X_0, \dots, X_d$.

Remark 2.16. In [Beals and Greiner 1988] the privileged coordinates at u are called u-coordinates, but in the special case of a Heisenberg manifold they correspond to the privileged coordinates of [Bellaïche 1996] and [Gromov 1996].

Notice that in the privileged coordinates at u we can write

$$X_j = \partial_{x_j} + \sum_{k=0}^{d} a_{jk}(x)\,\partial_{x_k}, \qquad j = 0, 1, \dots d,$$

where the a_{jk}'s are smooth functions such that $a_{jk}(0) = 0$.

Next, on $\mathbb{R}^{d+1}$ we consider the dilations

$$\delta_t(x) = t \cdot x = (t^2 x_0, t x_1, \dots, t x_d), \qquad t \in \mathbb{R}, \tag{2-17}$$

with respect to which ∂_{x_0} is homogeneous of degree -2, while $\partial_{x_1}, \dots, \partial_{x_d}$ are homogeneous of degree -1. Therefore, we may let

$$X_0^{(u)} = \lim_{t\to 0} t^2 \delta_t^* X_0 \;=\; \partial_{x_0}, \tag{2-18}$$

$$X_j^{(u)} = \lim_{t\to 0} t \delta_t^* X_j \;=\; \partial_{x_j} + \sum_{k=1}^{d} b_{jk} x_k\,\partial_{x_0}, \qquad j = 1, \dots, d, \tag{2-19}$$

where $b_{jk} = \partial_{x_k} a_{j0}(0)$ for $j, k = 1, \dots, d$. In fact, for any vector field $X = a_0(x)X_0 + \dots + a_d(x)X_d$ we have

$$\lim_{t\to 0} t^2 \delta_t^* X = a_0(0)\, X_0^{(u)},$$

$$\lim_{t\to 0} t^{-1} \delta_t^* X = a_1(0)\, X_1^{(u)} + \dots + a_d(0)\, X_d^{(u)} \quad \text{when } a_0(0) = 0. \tag{2-20}$$

Observe that $X_0^{(u)}$ is homogeneous of degree -2 and $X_1^{(u)}, \dots, X_d^{(u)}$ are homogeneous of degree -1. Moreover, for $j, k = 1, \dots, d$, we have

$$[X_j^{(u)}, X_0^{(u)}] = 0 \qquad \text{and} \qquad [X_j^{(u)}, X_0^{(u)}] = (b_{kj} - b_{jk}) X_0^{(u)}. \tag{2-21}$$

Thus, the linear space spanned by $X_0^{(u)}, X_1^{(u)}, \dots, X_d^{(u)}$ is a graded 2-step nilpotent Lie algebra $\mathfrak{g}^{(u)}$. In particular, $\mathfrak{g}^{(u)}$ is the Lie algebra of left-invariant vector fields over the graded Lie group $G^{(u)}$, consisting of $\mathbb{R}^{d+1}$ equipped with the grading (2-17) and the group law

$$x \cdot y = \Big(x_0 + \sum_{j,k=1}^{d} b_{kj} x_j y_k,\; x_1 + y_1, \dots, x_d + y_d \Big).$$

Now, if near m we set $\mathcal{L}(X_j, X_k) = [X_j, X_k] = L_{jk}X_0 \bmod H$, then we have

$$[X_j^{(u)}, X_k^{(u)}] = \lim_{t\to 0}\big[t\delta_t^* X_j, t\delta_t^* X_k\big] = \lim_{t\to 0} t^2\delta_t^*(L_{jk}X_0) = L_{jk}(m)X_0^{(u)}. \tag{2-22}$$

Comparing this with (2-15) and (2-21) shows that $\mathfrak{g}^{(u)}$ has the same constants of structure as $\mathfrak{g}_m M$, and is therefore isomorphic to it. Consequently, the Lie groups $G^{(u)}$ and $G_m M$ are isomorphic. An explicit isomorphism can be obtained as follows.

Lemma 2.17. *Consider a diffeomorphism $\phi : \mathbb{R}^{d+1} \to \mathbb{R}^{d+1}$ of the form*

$$\phi(x_0, \dots, x_d) = (x_0 + \tfrac{1}{2}c_{jk}x_j x_k, x_1, \dots, x_d), \tag{2-23}$$

where $c = (c_{jk})$, $c^t = c$, is a symmetric matrix in $M_d(\mathbb{R})$. Then ϕ is a graded isomorphism from $G^{(u)}$ onto the Lie group G consisting of $\mathbb{R}^{d+1}$ equipped with the group law,

$$x.y = (x_0 + y_0 + \sum_{j,k=1}^{d}(b_{kj} + c_{kj})x_j y_k, x_1 + y_1, \dots, x_d + y_d). \tag{2-24}$$

Moreover, under ϕ the vector fields $X_0^{(u)}, \dots, X_d^{(u)}$ transform into

$$\phi_* X_0^{(u)} = \frac{\partial}{\partial x_0}, \tag{2-25}$$

$$\phi_* X_j^{(u)} = \partial_{x_j} + \sum_{k=1}^{d}(b_{jk} + c_{jk})x_k\partial_{x_0}, \qquad j = 1, \dots, d. \tag{2-26}$$

Proof. First, since $\phi(t\cdot x) = t\cdot\phi(x)$ for any $t \in \mathbb{R}$, we see that ϕ is graded. Second, for x and y in $\mathbb{R}^{d+1}$ the product $\phi(x)\cdot\phi(y)$ is equal to

$$\begin{aligned}
&\phi\big(x_0 + y_0 + \sum_{j,k=1}^{d} b_{kj}x_j y_k,\ x_1 + y_1, \dots, x_d + y_d\big)\\
&\ = \big(x_0 + y_0 + \sum_{j,k=1}^{d} b_{kj}x_j y_k + \tfrac{1}{2}\sum_{j,k=1}^{d} c_{jk}(x_j + y_j)(x_k + y_k),\ x_1 + y_1, \dots, x_d + y_d\big)\\
&\ = \big(x_0 + \tfrac{1}{2}\sum_{j,k=1}^{d} c_{jk}x_j x_k + y_0 + \tfrac{1}{2}\sum_{j,k=1}^{d}\big(c_{jk}y_j y_k + (b_{kj} + c_{kj})x_j y_k\big),\ x_1 + y_1, \dots, x_d + y_d\big).
\end{aligned}$$

Thus, in view of the law group of G we have $\phi(x\cdot y) = \phi(x)\cdot\phi(y)$ and ϕ is a Lie group isomorphism. Consequently, for each $j = 0, \dots, d$, the vector field

$$\phi_* X_j^{(u)} = \phi'\big(\phi^{-1}(x)\big)\big(X_j(\phi^{-1}(x))\big)$$

is left-invariant on G. In fact, as $\phi'(0) = \mathrm{id}$ and $X_j^{(u)}(0) = \partial_{x_j}$ we see that $\phi_* X_j^{(u)}$ is the left-invariant vector field on G that coincides with ∂_{x_j} at $x = 0$. Therefore, by substituting $b_{jk}+c_{jk}$ for b_{jk} in (2-18)–(2-19), we get the formulas (2-25)–(2-26) for $\phi_* X_j^{(u)}$. The lemma is thus proved. □

Since by (2-21) and (2-22) we have $L_{jk}=b_{kj}-b_{jk}$ for $j,k=1,\ldots,d$, we deduce from Lemma 2.17 that an isomorphism of graded Lie groups from $G^{(u)}$ onto $G_m M$ is given by

$$\phi_u(x_0,\ldots,x_d)=\Big(x_0-\tfrac{1}{4}\sum_{j,k=1}^{d}(b_{jk}+b_{kj})x_j x_k,\ x_1,\ldots,x_d\Big). \tag{2-27}$$

Definition 2.18. Let $\varepsilon_u=\phi_u\circ\psi_u$.

(1) The new coordinates provided by ε_u are called *Heisenberg coordinates* at u with respect to the H-frame $X_0,\ldots,X_d$.

(2) The map ε_u is called the *u-Heisenberg coordinate map*.

Remark 2.19. The Heisenberg coordinates were first introduced in [Beals and Greiner 1988], where they were called "antisymmetric u-coordinates" and used as a technical tool for inverting the principal symbol of a hypoelliptic sublaplacian.

Next, Lemma 2.17 also tells us that

$$\phi_* X_0^{(u)}=\partial_{x_0}=X_0^m,$$
$$\phi_* X_j^{(u)}=\partial_{x_j}-\tfrac{1}{2}\sum_{k=1}^{d}L_{jk}x_k\partial_{x_0}=X_j^m,\qquad j=1,\ldots,d.$$

Since ϕ_u commutes with the Heisenberg dilations (2-17), by using (2-18) and (2-19) we get

$$\lim_{t\to 0} t^2\delta_t^*\phi_{u*}X_0^{(u)}=X_0^m \quad\text{and}\quad \lim_{t\to 0} t\delta_t^*\phi_{u*}X_j^{(u)}=X_j^m,\quad j=1,\ldots,d.$$

Combining this with (2-16) and (2-20) shows that, for any vector field X near m, in Heisenberg coordinates at m we have, as $t\to 0$,

$$\delta_t^* X=\begin{cases} t^{-2}X^m+O(t^{-1}) & \text{if } X(m)\in H_m,\\ t^{-1}X^m+O(1) & \text{otherwise.}\end{cases} \tag{2-28}$$

Therefore, we obtain:

Proposition 2.20. *In Heisenberg coordinates centered at $m=\kappa^{-1}(u)$, the tangent Lie group $G_m M$ coincides with $G^{(u)}$.*

2.4. ***Tangent approximation of Heisenberg diffeomorphisms.*** If $\phi: M\to M'$ is a smooth map between (standard) smooth manifolds, then for any $m\in M$ the derivative $\phi'(m)$ yields a tangent linear approximation for ϕ in local coordinates around m. We shall now prove an analogous result in the Heisenberg setting. To this end, it will be useful to endow $\mathbb{R}^{d+1}$ with the pseudonorm,

$$\|x\| = \left(x_0^2 + (x_1^2 + \ldots + x_d^2)^2\right)^{1/4}, \qquad x \in \mathbb{R}^{d+1},$$

so that, for any $x \in \mathbb{R}^{d+1}$ and any $t \in \mathbb{R}$, we have

$$\|t \cdot x\| = |t|\,\|x\|. \tag{2-29}$$

From now on we let $\phi : (M, H) \to (M', H')$ be a Heisenberg diffeomorphism from (M, H) to another Heisenberg manifold (M', H').

Proposition 2.21. *Let $m \in M$ and set $m' = \phi(m)$. Then, in Heisenberg coordinates at m and at m', the diffeomorphism $\phi(x)$ has a behavior near $x = 0$ of the form*

$$\phi(x) = \phi_H'(0)x + \left(O(\|x\|^3), O(\|x\|^2), \ldots, O(\|x\|^2)\right), \tag{2-30}$$

where ϕ_H is as in Definition 2.11. In particular, there is no term of the form $x_j x_k$, $1 \le j, k \le d$, in the Taylor expansion of $\phi_0(x)$ at $x = 0$.

Proof. Let $X_0, \ldots, X_d$ be an H-frame of TM over a Heisenberg chart κ near m and let $Y_0, \ldots, Y_d$ be an H'-frame of TM' over a Heisenberg chart κ_1 near m'. Set $u = \kappa(m)$, so that in privileged coordinates at u we have $X_j(0) = \partial_{x_j}$ for $j = 0, \ldots, d$. As the change of variables ϕ_u from privileged coordinates to Heisenberg coordinates at u is such that $\phi_u(0) = 0$ and $\phi_u'(0) = \mathrm{id}$, we see that in Heisenberg coordinates at m we also have $X_j(0) = \partial_{x_j}$ for $j = 0, \ldots, d$. Similarly, in Heisenberg coordinates at m' we have $Y_j(0) = \partial_{x_j}$ for $j = 0, \ldots, d$. As $\phi'(0)$ maps H_0 to H_0' it then follows that, with respect to the basis $\partial_{x_0}, \ldots, \partial_{x_d}$, the matrices of $\phi'(0)$ and $\phi_H'(0)$ take the forms

$$\phi'(0) = \begin{pmatrix} a_{00} & 0 \\ b & A_\| \end{pmatrix} \qquad \text{and} \qquad \phi_H'(0) = \begin{pmatrix} a_{00} & 0 \\ 0 & A_\| \end{pmatrix}, \tag{2-31}$$

for some scalar $a_{00} \neq 0$ and some matrices $b \in M_{d1}(\mathbb{R})$ and $A_\| \in GL_d(\mathbb{R})$. In particular, we have $\phi'(0)x = \phi_H'(0)x + x_0(0, b_1, \ldots, b_d)$. Thus, the Taylor expansion of $\phi(x)$ at $x = 0$ takes the form

$$\phi(x) = \hat{\phi}(x) + \theta(x), \qquad \hat{\phi}(x) = \left(x_0 + \tfrac{1}{2}\sum_{j,k=1}^{d} c_{jk} x_j x_k,\ x_1, \ldots, x_d\right), \tag{2-32}$$

where $c_{jk} = \partial^2_{x_j,x_k}\phi_0(0)$ and $\theta(x) = \left(\theta_0(x), \ldots, \theta_d(x)\right)$ is such that

$$\theta_0(x) = O\left(|x_0||x| + |x|^3\right) = O(\|x\|^3), \tag{2-33}$$

$$\theta_j(x) = O\left(|x_0| + |x|^2\right) = O(\|x\|^2), \qquad j = 1, \ldots, d. \tag{2-34}$$

To complete the proof we need only to show that $c_{jk} = 0$ for $j, k = 1, \ldots, d$. Possibly after replacing ϕ by $\phi_H'(0)^{-1} \circ \phi$ we may assume that $\phi_H'(0) = \mathrm{id}$. Since by Proposition 2.12 $\phi'_H(0)$ is a Lie group isomorphism from $G = G_0M$ onto $G' = G_0M'$, this implies that G and G' have the same group law, namely,

$$x \cdot y = \Big(x_0 + y_0 + \tfrac{1}{2}\sum_{j,k=1}^{d} L_{jk}x_jx_k,\ x_1 + y_1, \ldots, x_d + y_d\Big),$$

where the structure constants L_{jk} are such that

$$\mathcal{L}(X_j, X_k)(0) = \mathcal{L}(Y_j, Y_k)(0) = L_{jk}X_0(0).$$

Therefore, using (2-14) we see that, at the level of the model vector fields (2-13), we have

$$\text{(2-35)}\qquad \begin{aligned} X_0^m &= Y_0^{m'} = \partial_{x_0},\\ X_j^m &= Y_j^{m'} = \partial_{x_j} - \tfrac{1}{2}\sum_{k=1}^{d} L_{jk}x_k\partial_{x_0}, \qquad j = 1, \ldots, d. \end{aligned}$$

As in (2-31) $\phi'_H(0)$ is the diagonal part of $\phi'(0)$ we have $\phi_*X_0(0) = Y_j(0) \bmod H'_0$ and $\phi_*X_0(0) = Y_j(0)$ for $j = 1, \ldots, d$. Therefore, using (2-13) we obtain

$$\text{(2-36)}\qquad (\phi_*X_j)^{m'} = Y_j^{m'} = X_j^m \qquad \text{for } j = 0, \ldots, d.$$

On the other hand, as we are using Heisenberg coordinates both at m and m', from (2-28) we get

$$X_j^m = \lim_{t\to 0} t\delta_t^* X_j \qquad \text{and} \qquad (\phi_*X_j)^{m'} = \lim_{t\to 0} t\delta_t^*\phi_*X_j = \lim_{t\to 0}(\delta_t^{-1}\circ\phi\circ\delta_t)_*(t\delta_t^*X_j).$$

Since (2-32)–(2-34) imply that $\lim_{t\to 0}\delta_t^{-1}\circ\phi\circ\delta_t = \hat{\phi}$, we see that

$$(\phi_*X_j)^{m'} = \lim_{t\to 0}(\delta_t^{-1}\circ\phi\circ\delta_t)_* \lim_{t\to 0}(t\delta_t^*X_j) = \hat{\phi}_*X_j^m.$$

Combining this with (2-36) we then get

$$\text{(2-37)}\qquad \hat{\phi}_*X_j^m = (\phi_*X_j)^{m'} = X_j^m \qquad \text{for } j = 1, \ldots, d.$$

Now, the form of $\hat{\phi}$ in (2-32) allows us to apply Lemma 2.17 to get

$$\hat{\phi}_*X_j^m = \partial_{x_j} + \sum_{k=1}^{d}\big(-\tfrac{1}{2}L_{jk} + c_{jk}\big)x_k\partial_{x_0}.$$

Combining with (2-35) and (2-37) then gives $L_{jk} = L_{jk} - 2c_{jk}$, from which we deduce that $c_{jk} = 0$ for $j, k = 1, \ldots, d$. The proof is thus complete. □

Remark 2.22. An asymptotics similar to (2-30) is given in [Bellaïche 1996, Proposition 5.20] by using privileged coordinates at u and $u' = \kappa_1(m')$, but the leading term there is only a Lie algebra isomorphism from $\mathfrak{g}^{(u)}$ onto $\mathfrak{g}^{(u')}$. It is only in Heisenberg coordinates that we recover the Lie group isomorphism $\phi'_H(m)$ as the leading term of the asymptotics.

Finally, for future use we mention the following version of Proposition 2.21.

Proposition 2.23. *In local coordinates and as $t \to 0$ we have*

$$t^{-1} \cdot \varepsilon_{\phi(u)} \circ \phi \circ \varepsilon_u^{-1}(t \cdot x) = (\varepsilon_{\phi(u)} \circ \phi \circ \varepsilon_u^{-1})'_H(0)x + O(t),$$

locally uniformly with respect to u and x.

Proof. By combining Proposition 2.21 and (2-29) we get

$$t^{-1} \cdot \varepsilon_{\phi(u)} \circ \phi \circ \varepsilon_u^{-1}(t \cdot x) = (\varepsilon_{\phi(u)} \circ \phi \circ \varepsilon_u^{-1})'_H(0)x + O(t). \tag{2-38}$$

A priori this holds only pointwise with respect to u and x. However, the asymptotic bound above comes from remainder terms in Taylor formulas at $t = 0$ for components of $\Psi(u, x, t) := \varepsilon_{\phi(u)} \circ \phi \circ \varepsilon_u^{-1}(t \cdot x)$. Since Ψ is smooth with respect to u and x, it follows that the bounds in (2-38) are locally uniform with respect to u and x. Hence the result. □

3. The tangent groupoid of a Heisenberg manifold

In this section we construct the tangent groupoid of a Heisenberg manifold (M, H) as a groupoid encoding the smooth deformation of $M \times M$ to GM. In this construction a crucial use is made of Heisenberg coordinates and of the tangent approximation of Heisenberg diffeomorphisms provided by Proposition 2.21.

3.1. ***Differentiable groupoids.*** Here we recall the main definitions on groupoids and illustrate them with the example of Connes' tangent groupoid.

Definition 3.1. A *groupoid* consists of a set $\mathcal{G}$, a distinguished subset $\mathcal{G}^{(0)} \subset \mathcal{G}$, two maps r and s from $\mathcal{G}$ to $\mathcal{G}^{(0)}$ (called the *range* and *source* maps) and a composition map,

$$\circ : \mathcal{G}^{(2)} = \left\{(\gamma_1, \gamma_2) \in \mathcal{G} \times \mathcal{G} \,\middle|\, s(\gamma_1) = r(\gamma_2)\right\} \longrightarrow \mathcal{G},$$

such that the following properties are satisfied:

(1) $s(\gamma_1 \circ \gamma_2) = s(\gamma_2)$ and $r(\gamma_1 \circ \gamma_2) = r(\gamma_1)$, for any $(\gamma_1, \gamma_2) \in \mathcal{G}^{(2)}$;

(2) $s(x) = r(x) = x$ for any $x \in \mathcal{G}^{(0)}$;

(3) $\gamma \circ s(\gamma) = r(\gamma) \circ \gamma = \gamma$ for any $\gamma \in \mathcal{G}$;

(4) $(\gamma_1 \circ \gamma_2) \circ \gamma_3 = \gamma_1 \circ (\gamma_2 \circ \gamma_3)$;

(5) each element $\gamma \in \mathcal{G}$ has a two-sided inverse γ^{-1} so that $\gamma \circ \gamma^{-1} = r(\gamma)$ and $\gamma^{-1} \circ \gamma = s(\gamma)$.

The groupoids interpolate between spaces and groups. This aspect especially pertains in the construction by Connes [1994, Section II.5] (see also [Hilsum and Skandalis 1987]) of the tangent groupoid $\mathcal{G} = \mathcal{G}M$ of a smooth manifold M.

At the set-theoretic level we let

$$\mathcal{G} = TM \sqcup \big(M \times M \times (0, \infty)\big) \qquad \text{and} \qquad \mathcal{G}^{(0)} = M \times [0, \infty),$$

where TM denotes the (total space) of the tangent bundle of M. The inclusion ι of $\mathcal{G}^{(0)}$ into $\mathcal{G}$ is given by

$$\iota(m,t)=\begin{cases}(m,m,t) & \text{for } t>0 \text{ and } m\in M,\\ (m,0)\in TM & \text{for } t=0 \text{ and } m\in M.\end{cases} \tag{3-1}$$

The range and source maps of $\mathcal{G}$ are such that

$$r(p,q,t)=(p,t) \quad \text{and} \quad s(p,q,t)=(q,t) \quad \text{for } t>0 \text{ and } p,q\in M,$$
$$r(p,X)=s(p,X)=(p,0) \quad \text{for } t=0 \text{ and } (p,X)\in TM,$$

while the composition law is given by

$$(p,m,t)\circ(m,q,t)=(p,q,t) \quad \text{for } t>0 \text{ and } m,p,q\in M, \tag{3-2}$$
$$(p,X)\circ(p,Y)=(p,X+Y) \quad \text{for } t=0 \text{ and } (p,X),(p,Y)\in TM. \tag{3-3}$$

In fact, $\mathcal{G}M$ is a b-differentiable groupoid in the sense of the definition below.

Definition 3.2. A *b-differentiable groupoid* is a groupoid $\mathcal{G}$ so that $\mathcal{G}$ and $\mathcal{G}^{(0)}$ are smooth manifolds with boundary and the following properties hold:

(1) the inclusion of $\mathcal{G}^{(0)}$ into $\mathcal{G}$ is smooth;

(2) the source and range maps are smooth submersions, so that $\mathcal{G}^{(2)}$ is a submanifold (with boundary) of $\mathcal{G}\times\mathcal{G}$;

(3) the composition map $\circ:\mathcal{G}^{(2)}\to\mathcal{G}$ is smooth.

In the case of the tangent groupoid $\mathcal{G}=\mathcal{G}M$ the topology is such that:

- the inclusions of $\mathcal{G}^{(0)}$ and $\mathcal{G}^{(1)}:=M\times M\times(0,\infty)$ into $\mathcal{G}$ are continuous and $\mathcal{G}^{(1)}$ is an open subset of $\mathcal{G}$;
- a sequence (p_n,q_n,t_n) from $\mathcal{G}^{(1)}$ converges to $(p,X)\in TM$ if, and only if, $\lim(p_n,q_n,t_n)=(p,p,0)$ and for any local chart κ near p we have

$$\lim_{n\to\infty} t_n^{-1}\big(\kappa(q_n)-\kappa(p_n)\big)=\kappa'(p)X.$$

One can check that this condition does not depend on the choice of a particular chart near p.

The differentiable structure of $\mathcal{G}M$ is obtained by gluing those of TM and of $\mathcal{G}^{(1)}=M\times M\times(0,\infty)$ by means of a chart of the form,

$$\gamma(p,X,t)=\begin{cases}\big(p,\exp_p(-tX),t\big) & \text{if } t>0 \text{ and } (p,tX)\in\operatorname{dom}\exp,\\ (p,X) & \text{if } t=0 \text{ and } (p,X)\in\operatorname{dom}\exp.\end{cases} \tag{3-4}$$

Here $\exp:\operatorname{dom}\exp\to M\times M$ denotes the exponential map associated to an arbitrary Riemannian metric on M, so that γ maps an open subset of $TM\times[0,\infty)$

onto an open neighborhood in $\mathcal{G}$ of the boundary TM (see [Connes 1994], [Hilsum and Skandalis 1987], [Cariñena et al. 1999]).

3.2. ***The tangent groupoid of a Heisenberg manifold.*** We now construct the tangent groupoid $\mathcal{G} = \mathcal{G}_H M$ of a Heisenberg manifold (M^{d+1}, H).

As an abstract groupoid $\mathcal{G}_H M$ is defined as follows. First, we set

$$\mathcal{G} = GM \sqcup \big(M \times M \times (0, \infty)\big) \qquad \text{and} \qquad \mathcal{G}^{(0)} = M \times [0, \infty),$$

where GM denotes the (total space) of the tangent Lie group bundle of M. We have an inclusion $\iota : \mathcal{G}^{(0)} \to \mathcal{G}$ as in (3-1), namely,

$$\iota(m, t) = \begin{cases} (m, m, t) & \text{for } t > 0 \text{ and } m \in M, \\ (m, 0) \in GM & \text{for } t = 0 \text{ and } m \in M. \end{cases}$$

The range and source maps are defined similarly to (3-2)–(3-3) by letting

$$r(p, q, t) = (p, t) \quad \text{and} \quad s(p, q, t) = (q, t) \quad \text{for } t > 0 \text{ and } p, q \in M,$$
$$r(p, X) = s(p, X) = (p, 0) \quad \text{for } t = 0 \text{ and } (p, X) \in GM.$$

In addition, we endow $\mathcal{G}$ with the composition law

$$(p, m, t) \circ (m, q, t) = (p, q, t) \quad \text{for } t > 0 \text{ and } m, p, q \in M, \tag{3-5}$$
$$(p, X) \circ (p, Y) = (p, X.Y) \quad \text{for } t = 0 \text{ and } (p, X), (p, Y) \in GM. \tag{3-6}$$

It is immediate to check the properties (1)–(5) of Definition 3.1, noticing that the inverse map is here given by

$$(p, q, t)^{-1} = (q, p, t) \quad \text{for } t > 0 \text{ and } p, q \in M,$$
$$(p, X)^{-1} = (p, X^{-1}) = (p, -X) \quad \text{for } t = 0 \text{ and } (p, X) \in GM.$$

Therefore, $\mathcal{G} = \mathcal{G}_H M$ is a groupoid.

Definition 3.3. The groupoid $\mathcal{G}_H M$ is called the *tangent groupoid* of (M, H).

We now turn the groupoid $\mathcal{G} = \mathcal{G}_H M$ into a b-differentiable groupoid. First, we endow $\mathcal{G}$ with the topology such that:

- the inclusions of $\mathcal{G}^{(0)}$ and $\mathcal{G}^{(1)} := M \times M \times (0, \infty)$ into $\mathcal{G}$ are continuous and make $\mathcal{G}^{(1)}$ an open subset of $\mathcal{G}$;
- a sequence (p_n, q_n, t_n) from $\mathcal{G}^{(1)}$ converges to $(p, X) \in GM$ if, and only if, $\lim(p_n, q_n, t_n) = (p, p, 0)$ and for any local Heisenberg chart $\kappa : \operatorname{dom} \kappa \to U$ near p we have

$$\lim_{n\to\infty} t_n^{-1} \cdot \varepsilon_{\kappa(p_n)}\big(\kappa(q_n)\big) = (\varepsilon_{\kappa(p)} \circ \kappa)'_H(p) X, \tag{3-7}$$

where $t\cdot x$ is the Heisenberg dilation (2-17) and ε_u denotes the coordinate change to the Heisenberg coordinates at $u \in U$ with respect to the H-frame of the Heisenberg chart κ (see Definition 2.18).

Lemma 3.4. *The condition (3-7) is independent of the choice of Heisenberg chart.*

Proof. Assume that (3-7) holds for κ. Let κ_1 be another Heisenberg chart near p, and let $\phi = \kappa_1 \circ \kappa^{-1}$. Letting $x_n = \kappa(p_n)$ and $y_n = \kappa(q_n)$, we have

$$t_n^{-1}\cdot \varepsilon_{\kappa_1(p_n)}(\kappa_1(q_n)) = t_n^{-1}\cdot \varepsilon_{\phi(x_n)}(\phi(y_n)) = \delta_{t_n}^{-1}\circ \varepsilon_{\phi(x_n)} \circ \phi \circ \varepsilon_{x_n}^{-1} \circ \delta_{t_n}\big(t_n \cdot \varepsilon_{x_n}(y_n)\big). \tag{3-8}$$

On the other hand, since ϕ is a Heisenberg diffeomorphism it follows from Proposition 2.23 that as t goes to zero we have

$$\delta_t^{-1}\circ \varepsilon_{\phi(x)} \circ \phi \circ \varepsilon_x^{-1} \circ \delta_t(y) - \partial_y(\varepsilon_{\phi(x)} \circ \phi \circ \varepsilon_x^{-1})_H(0)\, y \longrightarrow 0,$$

locally uniformly with respect to x and y. Since $(x_n, y_n, t_n) \to \big(\kappa(p), \kappa(p), 0\big)$ and $t_n^{-1}\cdot \varepsilon_{\kappa(p_n)}(\kappa(q_n)) \to (\varepsilon_{\kappa(p)} \circ \kappa)'_H(p)\, X$, by combining this with (3-8) we get

$$\lim_{n\to\infty} t_n^{-1}\cdot \varepsilon_{\kappa_1(p_n)}(\kappa_1(q_n)) = \big(\varepsilon_{\phi(\kappa(p))} \circ \phi \circ \varepsilon_{\kappa(p)}^{-1}\big)'_H(0)\,\big((\varepsilon_{\kappa(p)} \circ \kappa)'_H(p)\, X\big) = (\varepsilon_{\kappa_1(p)} \circ \kappa_1)'_H(p)\, X.$$

Hence the lemma. □

Next, in order to endow $\mathcal{G}_H M$ with a manifold structure we cannot make use of an exponential chart as in (3-4) because, unless GM is a fiber bundle, the Lie-algebraic structures of its fibers vary from point to point. Instead we can proceed as follows.

Let $\kappa : \operatorname{dom}\kappa \to U$ be a local Heisenberg chart near $m \in M$. We get a local coordinate system near $GM_{|\operatorname{dom}\kappa} \subset \mathcal{G}$ by letting

$$\gamma_\kappa(x, X, t) = \begin{cases} \big(\kappa^{-1}(x),\ \kappa^{-1}\circ \varepsilon_x^{-1}(t\cdot X),\ t\big) & \text{if } t > 0 \text{ and } x,\ \varepsilon_x^{-1}(t\cdot X) \in U, \\ \big(\kappa^{-1}(x),\ (\kappa^{-1}\circ \varepsilon_x^{-1})'_H(0)\, X\big) & \text{if } t = 0 \text{ and } (x, X) \in U \times \mathbb{R}^{d+1}. \end{cases}$$

The map γ_κ is one-to-one from an open neighborhood of the boundary $U\times\mathbb{R}^{d+1}\times 0$ in $U \times \mathbb{R}^{d+1} \times [0,\infty)$. Moreover, γ_κ is continuous off the boundary. It is also continuous near any boundary point $(x, X, 0)$ because if a sequence $(x_n, X_n, t_n) \in \operatorname{dom}\gamma_\kappa$ with $t_n > 0$ converges to $(x, X, 0)$ then $(p_n, q_n, t_n) = \gamma_\kappa(x_n, X_n, t_n)$ has limit $\big(\kappa^{-1}(x),\ (\kappa^{-1})'_H(x)\, X\big) = \gamma_\kappa(x, X, 0)$, for we have

$$t_n^{-1}\cdot \varepsilon_{\kappa(p_n)}(\kappa(q_n)) = X_n \longrightarrow X = \kappa'_H(\kappa(x))\,(\kappa^{-1})'_H(x)\, X.$$

The inverse γ_κ^{-1} is given by

$$\gamma_\kappa^{-1}(p,q,t) = \big(\kappa(p),\ t^{-1}\cdot \varepsilon_{\kappa(p)} \circ \kappa(q),\ t\big) \quad \text{for } t>0, \tag{3-9}$$

$$\gamma_{\kappa_1}^{-1}(p,X) = \big(\kappa(p),\ \kappa'_H(p)X\big) \quad \text{for } (p,X)\in GM \text{ in the range of } \gamma_{\kappa_1}. \tag{3-10}$$

Therefore, if κ_1 is another local Heisenberg chart near m then, in terms of $\phi = \kappa_1^{-1}\circ\kappa$, the transition map $\gamma_\kappa^{-1}\circ\gamma_{\kappa_1}$ is

$$\gamma_\kappa^{-1}\circ\gamma_{\kappa_1}(x,X,t) = \begin{cases} \big(\phi(x),\ t^{-1}\cdot \varepsilon_{\phi(x)}\circ\phi\circ\varepsilon_x^{-1}(t\cdot X),\ t\big) & \text{for } t>0,\\ \big(\phi(x),\ \phi'_H(x)X,\ 0\big) & \text{for } t=0.\end{cases}$$

This shows that $\gamma_\kappa^{-1}\circ\gamma_{\kappa_1}(x,X,t)$ is smooth with respect to x and X and is meromorphic with respect to t with at worst a possible singularity at $t=0$ only. However, by Proposition 2.23 we have

$$\lim_{t\to 0} t^{-1}\cdot \varepsilon_{\phi(x)}\circ\phi\circ\varepsilon_x^{-1}(t\cdot X) = \phi'_H(x)X,$$

so there is no singularity at $t=0$. Hence $\gamma_\kappa^{-1}\circ\gamma_{\kappa_1}$ is a smooth diffeomorphism between open subsets of $\mathbb{R}^{d+1}\times[0,\infty)$. Therefore the coordinates system γ_κ allows us to glue together the differentiable structures of GM and $\mathcal{G}^{(1)} = M\times M\times(0,\infty)$ to turn $\mathcal{G}$ into a smooth manifold with boundary.

Next, $\mathcal{G}^{(0)} = M\times[0,\infty)$ is a manifold with boundary and the inclusion $\iota:\mathcal{G}^{(0)}\to\mathcal{G}$ is smooth. In addition, the range map r and the source maps s are submersions off the boundary. Moreover, in a coordinate system γ_κ near the boundary of $\mathcal{G}$ the maps r and s are given by

$$r(x,X,t) = (x,t) \qquad \text{and} \qquad s(x,X,t) = \big(\varepsilon_x^{-1}(t\cdot X),t\big), \tag{3-11}$$

which shows that $\partial_{x,t}r$ and $\partial_{X,t}s$ are invertible near the boundary. Hence r and s are submersions on all $\mathcal{G}$.

Let us now look at the smoothness of the composition map.

Proposition 3.5. *The composition map $\circ:\mathcal{G}^2\to\mathcal{G}$ is smooth.*

Proof. Since $\circ$ is clearly smooth off the boundary we only need to understand what happens near the boundary. Using (3-11) we see that in a local coordinate system γ_κ near the boundary two elements (x,X,t) and (y,Y,t) can be composed if, and only if, we have $y=\varepsilon_x(t\cdot X)$. Then for $t>0$ using (3-5) and (3-9) we see that $(x,X,t)\circ\big(\varepsilon_x^{-1}(t\cdot X),Y,t\big)$ is equal to

$$\begin{aligned}\gamma_\kappa^{-1}\Big(\big(\kappa^{-1}(x),\ \kappa^{-1}\varepsilon_x^{-1}(t\cdot X),\ t\big)\circ\big(\kappa^{-1}\varepsilon_x^{-1}(t\cdot X),\ \kappa^{-1}\circ\varepsilon^{-1}_{\varepsilon_x^{-1}(t\cdot X)}(t\cdot Y),\ t\big)\Big)&\\ = \gamma_\kappa^{-1}\Big(\big(\kappa^{-1}(x),\ \kappa^{-1}\circ\varepsilon^{-1}_{\varepsilon_x^{-1}(t\cdot X)}(t\cdot Y),\ t\big)\Big)&\\ = \big(x,\ t^{-1}\cdot\varepsilon_x\circ\varepsilon^{-1}_{\varepsilon_x^{-1}(t\cdot X)}(t\cdot Y),\ t\big).&\end{aligned}$$

On the other hand, for $t=0$ from (3-6) and (3-10) we see that $(x, X, 0)\circ(x, Y, 0)$ is equal to

$$\begin{aligned}\gamma_\kappa^{-1}\Big(\big(\kappa^{-1}, (\kappa^{-1}\circ\varepsilon_x^{-1})'_H(0)\,X\big)\circ\big(\kappa^{-1}, (\kappa^{-1}\circ\varepsilon_x^{-1})'_H(0)\,Y\big)\Big)&\\ =\gamma_\kappa^{-1}\Big(\big(\kappa^{-1}(x),\ \big((\kappa^{-1}\circ\varepsilon_x^{-1})'_H(0)\,X\big)\cdot\big((\kappa^{-1}\circ\varepsilon_x^{-1})'_H(0)\,Y\big)\Big)&\\ =\gamma_\kappa^{-1}\big(\kappa^{-1}(x),\ (\kappa^{-1}\circ\varepsilon_x^{-1})'_H(0)\,(X\cdot Y)\big)&\\ =(x, X\cdot Y, 0),&\end{aligned}$$

where we have used the fact that $(\kappa^{-1}\circ\varepsilon_x^{-1})'_H(0)$ is a morphism of Lie groups (cf. Proposition 2.12). Therefore, we get

$$(x, X, t)\circ\big(\varepsilon_x^{-1}(t\cdot X),\ Y, t\big)=\begin{cases}\big(x,\ t^{-1}\cdot\varepsilon_x\circ\varepsilon^{-1}_{\varepsilon_x^{-1}(t\cdot X)}(t\cdot Y),\ t\big) & \text{for } t>0,\\ (x, X\cdot Y, 0) & \text{for } t=0.\end{cases}$$

This shows that $\circ$ is smooth with respect to x, X, and Y and is meromorphic with respect to t with at worst a singularity at $t=0$. Therefore, in order to show the smoothness of $\circ$ at $t=0$ it is enough to prove that

$$\lim_{t\to 0^+} t^{-1}\cdot\varepsilon_x\circ\varepsilon^{-1}_{\varepsilon_x^{-1}(t\cdot X)}(t\cdot Y)=X\cdot Y. \tag{3-12}$$

Lemma 3.6. *Let ψ_u denote the affine change to the privileged coordinates at u as in Definition 2.15. Then with respect to the law group of the u-group $G^{(u)}$ we have*

$$\lim_{t\to 0} t^{-1}.\psi_u\circ\psi^{-1}_{\psi_u^{-1}(t.v)}(t.w)=v.w, \tag{3-13}$$

locally uniformly with respect to w.

Proof. Let $\lambda_v(w)=v\cdot w$ and $\mu_t(w)=t^{-1}\cdot\psi_u\circ\psi^{-1}_{\psi_u^{-1}(t\cdot v)}(t\cdot w)$. For $w=0$ we have

$$\mu_t(0)=t^{-1}\cdot\psi_u\circ\psi^{-1}_{\psi_u^{-1}(t\cdot v)}(0)=t^{-1}\cdot\psi_u\big(\psi_u^{-1}(t\cdot v)\big)=v=\lambda_v(0). \tag{3-14}$$

Remark also that μ_t and λ_v are both affine maps and we have

$$\mu'_t=\delta_t^{-1}\circ\psi'_u\circ\big(\psi^{-1}_{\psi_u^{-1}(t\cdot v)}\big)'\circ\delta_t. \tag{3-15}$$

Let $X_0, \dots, X_d$ be the H-frame associated to the Heisenberg chart κ (seen as an H-frame on $U=\operatorname{range}\kappa$) and set $w_0=2$ and $w_1=\cdots=w_d=1$. By (2-18) and (2-19) we have $X_j(u)=(\psi_u^{-1})'(\partial_{x_j})$ for $j=0, \dots, d$. Therefore, we get

$$(\delta_t^*\psi_{u*}X_j)(v)=\delta_t^{-1}\circ\psi'_u\big(X_j(\psi_u^{-1}\circ\delta_t(v))\big)=\delta_t^{-1}\circ\psi'_u\circ\big(\psi^{-1}_{\psi_u^{-1}(t\cdot v)}\big)'(\partial_{x_j}).$$

Combining with (3-15) we thus obtain

$$t^{w_j}(\delta_t^*\psi_{u*}X_j)(v) = \delta_t^{-1}\circ\psi_u'\circ\big(\psi^{-1}_{\psi_u^{-1}(t\cdot v)}\big)'(t^{w_j}\partial_{x_j})$$
$$= \delta_t^{-1}\circ\psi_u'\circ\big(\psi^{-1}_{\psi_u^{-1}(t\cdot v)}\big)'\circ\delta_t(\partial_{x_j}) = \mu_t'(\partial_{x_j}).$$

Now, for $j=1,\ldots,d$, let $X_j^{(u)}$ be the left-invariant field on $G^{(u)}$ with $X_j^{(u)}=\partial_{x_j}$. Recall that, by the very definition of $G^{(u)}$ we have $X_j^{(u)}=\lim_{t\to 0} t^{w_j}(\delta_t^*\psi_{u*}X_j)$. Thus,

$$X_j^{(u)}(v) = \lim_{t\to 0}\mu_t'(\partial_{x_j}).$$

In fact, as $X_j^{(u)}$ is left-invariant, we have

$$X_j^{(u)}(v) = \big(\lambda_{v*}X_j^{(u)}\big)(v) = \lambda_v'\big(X_j^{(u)}(0)\big) = \lambda_v'(\partial_{x_j}).$$

Therefore, we have $\lim_{t\to 0}\mu_t'(\partial_{x_j}) = \lambda_v'(\partial_{x_j})$ for $j=0,\ldots,d$, which yields

$$\lim_{t\to 0}\mu_t' = \lambda_v'.$$

Since by (3-14) we have $\mu_t(0)=\lambda_v(0)$ and since μ_t and λ_v are affine maps, it follows that as t goes to zero $\mu_t(w)=t^{-1}\cdot\psi_u\circ\psi^{-1}_{\psi_u^{-1}(t\cdot v)}(t\cdot w)$ converges to $\lambda_v(w)=v\cdot w$ locally uniformly with respect to w. Hence the lemma. □

Next, let ϕ_x be the map (2-27), that is, the transition map from x-coordinates to Heisenberg coordinates centered at x. Recall that ϕ_x is an isomorphism of graded Lie groups from $G^{(x)}$ to the tangent group $G_x=(\kappa_* GM)_x$. Therefore, as $\varepsilon_x = \phi_x\circ\psi_x$ we get

$$t^{-1}\cdot\varepsilon_x\circ\varepsilon^{-1}_{\varepsilon_x^{-1}(t\cdot X)}(t\cdot Y) = \delta_t^{-1}\circ\phi_x\circ\psi_x\circ\psi^{-1}_{\psi_x^{-1}\circ\phi_x(t\cdot X)}\circ\phi_{\varepsilon_x^{-1}(t\cdot X)}\circ\delta_t(Y)$$
$$= \phi_x\big(\delta_t^{-1}\circ\psi_x\circ\psi^{-1}_{\psi_x^{-1}(t\cdot v)}\circ\delta_t(w_t)\big),$$

where we have let $v=\phi_x^{-1}(X)$ and $w_t=\phi_{\varepsilon_x^{-1}(t\cdot X)}(Y)$. Combining this with (3-13) we get

$$\lim_{t\to 0} t^{-1}\cdot\varepsilon_x\circ\varepsilon^{-1}_{\varepsilon_x^{-1}(t\cdot X)}(t\cdot Y) = \phi_x(v\cdot\lim_{t\to 0} w_t) = \phi_x\big(\phi_x^{-1}(X)\cdot\phi_x^{-1}(Y)\big) = X\cdot Y.$$

This proves (3-12) and thus completes the proof of Proposition 3.5. □

Summarizing all this we have proved:

Theorem 3.7. *The groupoid $\mathcal{G}_H M$ is a b-differentiable groupoid.*

Finally, let ϕ be a Heisenberg diffeomorphism from (M, H) onto a Heisenberg manifold (M', H') and let us compare the tangent groupoids $\mathcal{G}_H M$ and $\mathcal{G}_{H'} M'$. To this end consider the map $\Phi_H : \mathcal{G}_H M \to \mathcal{G}_{H'} M'$ given by

$$\Phi_H(p, q, t) = (\phi(p), \phi(q), t) \qquad \text{for } t > 0 \text{ and } p, q \in M, \tag{3-16}$$

$$\Phi_H(p, X) = (\phi(p), \phi'_H(p)X) \qquad \text{for } (p, X) \in GM. \tag{3-17}$$

For $t > 0$ and $p, q \in M$ we have

$$r_{M'} \circ \Phi_H(p, q, t) = (\phi(q), t) = \Phi_H \circ r_M(p, q, t),$$
$$s_{M'} \circ \Phi_H(p, q, t) = (\phi(p), t) = \Phi_H \circ s_M(p, q, t),$$

while for $(p, X) \in GM$ we have

$$\begin{aligned} s_{M'} \circ \Phi_H(p, X) = r_{M'} \circ \Phi_H(p, X) &= (\phi(p), 0) \\ &= \Phi_H \circ r_M(p, X) = \Phi_H \circ s_M(p, X). \end{aligned}$$

Hence $r_{M'} \circ \Phi_H = \Phi_H \circ r_M$ and $s_{M'} \circ \Phi_H = \Phi_H \circ s_M$. Incidentally $\Phi_H(\mathcal{G}_H^{(2)} M)$ agrees with $\mathcal{G}_{H'}^{(2)} M'$.

Moreover, for $t > 0$ and $m, p, q \in M$ we have

$$\begin{aligned} \Phi_H(m, p, t) \circ_{M'} \Phi_H(p, q, t) &= \big(\phi(m), \phi(q), t\big) \\ &= \Phi_H\big((m, p, t) \circ_M (p, q, t)\big), \end{aligned}$$

and for $p \in M$ and $X, Y \in G_p M$ we get

$$\begin{aligned} \Phi_H(p, X) \circ_{M'} \Phi_H(p, Y) &= \big(\phi(p), \phi'_H(p)(X \cdot Y)\big) \\ &= \Phi_H\big((p, X) \circ_M \Phi_H(p, Y)\big). \end{aligned}$$

All this shows that Φ_H is a morphism of groupoids. In fact, the map defined by replacing ϕ with ϕ^{-1} in (3-16) and (3-17) is an inverse for Φ_H, so Φ_H is in fact a groupoid isomorphism from $\mathcal{G}_H M$ onto $\mathcal{G}_{H'} M'$.

Next, it follows from (3-16) that Φ_H is continuous off the boundary. To see what happens at the boundary consider a sequence (p_n, q_n, t_n) converging to $(p, X) \in GM$ and let κ be a local Heisenberg chart for M' near $p' = \phi(p)$. By pulling back the H'-frame of κ by ϕ we turn $\kappa \circ \phi$ into a Heisenberg chart so setting $(p'_n, q'_n, t_n) = \Phi_H(p_n, q_n, t_n)$ we get

$$t_n^{-1} \cdot \varepsilon_{\kappa(p'_n)}(\kappa(q'_n)) = t_n \cdot \varepsilon_{\kappa \circ \phi(p_n)}(\kappa \circ \phi(q_n)) \longrightarrow (\kappa \circ \phi)'_H(p)X = \kappa'_H(p)(\phi'_H(p)X).$$

Thus, Φ_H is continuous from $\mathcal{G}_H M$ to $\mathcal{G}_{H'} M'$.

It also follows from (3-16) that Φ_H is smooth off the boundary. Moreover, if κ is a local Heisenberg chart for M' then $\Phi_H \circ \gamma_{\kappa \circ \phi}(p, X, t)$ coincides for $t > 0$ with

$$\begin{aligned} \Big(\phi\big(\phi^{-1} \circ \kappa^{-1}(x)\big), \phi\big(\phi^{-1} \circ \kappa^{-1} \circ \varepsilon_x^{-1}(t \cdot X)\big), t\Big) &= \big(\kappa^{-1}(x), \kappa^{-1} \circ \varepsilon_x^{-1}(t \cdot X), t\big) \\ &= \gamma_\kappa(x, X, t), \end{aligned}$$

while for $t = 0$ it is equal to

$$\Big(\phi\big(\phi^{-1}\circ\kappa^{-1}(x)\big),\ \phi'_H\big(\phi^{-1}\circ\kappa^{-1}(x)\big)\big((\kappa^{-1}\circ\varepsilon_x^{-1})'_H(0)\,X\big),\ 0\Big)$$
$$= \big(\kappa^{-1}(x),\ (\kappa^{-1}\circ\varepsilon_x^{-1})'_H(0)\,X,\ t\big) = \gamma_\kappa(x, X, 0).$$

Hence $\gamma_\kappa \circ \Phi \circ \gamma_{\kappa\circ\phi} = \mathrm{id}$, which shows that Φ_H is smooth map. Since similar arguments show that Φ_H^{-1} is smooth, it follows that Φ_H is a diffeomorphism. We have thus proved:

Proposition 3.8. *The map $\Phi_H : \mathcal{G}_H M \to \mathcal{G}_{H'} M'$ given by (3-16)–(3-17) is an isomorphism of b-differentiable groupoids. Hence the isomorphism class of b-groupoids of $\mathcal{G}_H M$ depends only on the Heisenberg diffeomorphism class of (M, H).*

References

[Atiyah and Singer 1968a] M. F. Atiyah and I. M. Singer, "The index of elliptic operators, I", *Ann. of Math.* (2) **87** (1968), 484–530. MR 38 #5243 Zbl 0164.24001

[Atiyah and Singer 1968b] M. F. Atiyah and I. M. Singer, "The index of elliptic operators, III", *Ann. of Math.* (2) **87** (1968), 546–604. MR 38 #5245 Zbl 0164.24301

[Beals and Greiner 1988] R. Beals and P. Greiner, *Calculus on Heisenberg manifolds*, Ann. of Math. Studies **119**, Princeton University Press, Princeton, NJ, 1988. MR 89m:35223 Zbl 0654.58033

[Bellaïche 1996] A. Bellaïche, "The tangent space in sub-Riemannian geometry", pp. 1–78 in *Sub-Riemannian geometry*, edited by A. Bellaïche and J.-J. Risler, Progr. Math. **144**, Birkhäuser, Basel, 1996. MR 98a:53108 Zbl 0862.53031

[Cariñena et al. 1999] J. F. Cariñena, J. Clemente-Gallardo, E. Follana, J. M. Gracia-Bondía, A. Rivero, and J. C. Várilly, "Connes' tangent groupoid and strict quantization", *J. Geom. Phys.* **32**:2 (1999), 79–96. MR 2000i:53127 Zbl 0961.53047

[Connes 1994] A. Connes, *Noncommutative geometry*, Academic Press, San Diego, CA, 1994. MR 95j:46063 Zbl 0818.46076

[Eliashberg and Thurston 1998] Y. M. Eliashberg and W. P. Thurston, *Confoliations*, University Lecture Series **13**, American Math. Society, Providence, RI, 1998. MR 98m:53042 Zbl 0893.53001

[Epstein et al. 1991] C. L. Epstein, R. B. Melrose, and G. A. Mendoza, "Resolvent of the Laplacian on strictly pseudoconvex domains", *Acta Math.* **167** (1991), 1–106. MR 92i:32016 Zbl 0758.32010

[Folland and Stein 1974] G. B. Folland and E. M. Stein, "Estimates for the $\bar\partial_b$ complex and analysis on the Heisenberg group", *Comm. Pure Appl. Math.* **27** (1974), 429–522. MR 51 #3719 Zbl 0293.35012

[Gromov 1996] M. Gromov, "Carnot-Carathéodory spaces seen from within", pp. 79–323 in *Sub-Riemannian geometry*, edited by A. Bellaïche and J.-J. Risler, Progr. Math. **144**, Birkhäuser, Basel, 1996. MR 2000f:53034 Zbl 0864.53025

[Hilsum and Skandalis 1987] M. Hilsum and G. Skandalis, "Morphismes K-orientés d'espaces de feuilles et fonctorialité en théorie de Kasparov (d'après une conjecture d'A. Connes)", *Ann. Sci. École Norm. Sup.* (4) **20**:3 (1987), 325–390. MR 90a:58169 Zbl 0656.57015

[Ponge 2000] R. Ponge, *Calcul hypoelliptique sur les variétés de Heisenberg, résidu non commutatif et géométrie pseudo-hermitienne*, Ph.D. thesis, University of Paris-Sud (Orsay), 2000.

[Rockland 1987] C. Rockland, "Intrinsic nilpotent approximation", *Acta Appl. Math.* **8**:3 (1987), 213–270. MR 88i:35032 Zbl 0621.22010

[Taylor 1984] M. E. Taylor, "Noncommutative microlocal analysis, I", *Mem. Amer. Math. Soc.* **52**:313 (1984), iv+182. MR 86f:58156 Zbl 0554.35025

[Van Erp 2005] E. Van Erp, *The Atiyah–Singer index formula for subelliptic operators on contact manifolds*, Ph.D. thesis, Pennsylvania State University, 2005.

Received August 30, 2004. Revised May 9, 2005.

RAPHAËL PONGE
DEPARTMENT OF MATHEMATICS
OHIO STATE UNIVERSITY
COLUMBUS, OH 43210-1174
UNITED STATES

ponge@math.osu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 227, No. 1, 2006

MOSER–TRUDINGER TRACE INEQUALITIES ON A COMPACT RIEMANNIAN SURFACE WITH BOUNDARY

YUNYAN YANG

Let (M, g) be a compact smooth Riemannian surface with boundary. In this paper, we use blowing-up analysis to prove that some Moser–Trudinger trace inequalities hold on certain function spaces, and that the extremal functions exist in those function spaces without any additional hypothesis on (M, g).

1. Introduction and main results

Let (M, g) be a compact smooth Riemannian surface, and $H^{1,2}(M)$ the completion of $C^\infty(M)$ under the norm

$$\|u\|_{H^{1,2}(M)} = \left(\int_M (|\nabla u|^2 + |u|^2)\, dV_g \right)^{1/2}.$$

A result of N. Trudinger [1967] implies that there exists a constant α such that

$$\sup_{\|u\|_{H^{1,2}(M)}=1} \int_M e^{\alpha u^2} dV_g < +\infty.$$

J. Moser proved the following theorems:

Theorem A [Moser 1970/71]. *Let Ω be an open domain in $\mathbb{R}^n$, $n \geq 2$. There exists a constant C which depends only on n such that if u is smooth, has compact support contained in Ω and its gradient ∇u satisfies $\int_M |\nabla u|^n dx \leq 1$, then*

$$\int_\Omega e^{\alpha_n |u|^{n/(n-1)}} dx \leq C|\Omega|,$$

where $\alpha_n = n(\omega_{n-1})^{1/n-1}$ and ω_{n-1} is the surface measure of the unit sphere in $\mathbb{R}^n$. If α_n is replaced by any $\alpha > \alpha_n$, the integral on the left-hand is still finite, but can be made arbitrarily large by an appropriate choice of u.

MSC2000: 58J05, 58J32.

Keywords: Moser–Trudinger inequality, trace imbedding theorem, blowing-up analysis, extremal functions.

Theorem B [Moser 1970/71]. *There exists an absolute constant c_0 such that if u is a smooth function on S^2 with $\int_{S^2} |\nabla u|^2\, dS = 1$ and $\int_{S^2} u\, dV_g = 0$, then*

$$\int_{S^2} e^{4\pi u^2} dS \le c_0.$$

The constant 4π is the best possible in the same sense as α_n in Theorem A.

Recall that Sobolev's theorems, see e.g. [10], assert existence of imbedding $W_0^{1,p}(\Omega) \to L^q(\Omega)$ for $1 < p < n$ and $W_0^{1,p}(\Omega) \to C^0(\Omega)$ for $p > n$, where $1/q = 1/p - 1/n$. Thus Theorem A represents a sharp way to fill in the gap at the critical exponent $p = n$. Theorem B plays the same role for the Sobolev theorems on S^2.

Moser's work was extended in [Adams 1988; Fontana 1993; Nolasco and Tarantello 1998; Chang and Yang 1988; Ding et al. 1997]. Generally, the inequalities obtained by those mathematicians are also called Moser–Trudinger inequalities.

It is well known that Moses–Trudinger inequalities play an important role in the study of partial differential equations, especially those that arise in geometry and physics. There has been much work on such inequalities and their applications; see, for example, [Trudinger 1967; Cohn and Lu 2002; Carleson and Chang 1986; Chang 1996; Flucher 1992; Lin 1996; Jost and Wang 2001] and the references therein.

Li and Zhu [1997] established some sharp Sobolev trace inequalities on n-dimensional compact Riemannian manifolds with smooth boundaries. Recently, Liu generalized a result of Osgood, Phillips and Sarnak [Osgood et al. 1988]:

Theorem C [Liu 2002]. *Let (M, g) be a compact Riemannian surface with boundary ∂M, then there exists a constant C, which depends only on the geometry of M, such that for all $u \in H^{1,2}(M)$*

$$\log \int_{\partial M} e^u ds_g \le \frac{1}{4\pi} \int_M |\nabla u|^2 dV_g + \int_{\partial M} u\, ds_g + C. \tag{1–1}$$

The value $\dfrac{1}{4\pi}$ is sharp.

A strong version of (1–1) has also been obtained:

Theorem D [Li and Liu 2005]. *Let (M, g) be a compact Riemannian surface with boundary ∂M. Then*

$$\sup_{\substack{\int_M |\nabla u|^2\, dV_g = 1 \\ \int_{\partial M} u\, dS_g = 0}} \int_{\partial M} e^{\pi u^2} dS_g < +\infty, \tag{1–2}$$

and

$$\sup_{\substack{\int_M |\nabla u|^2 dV_g = 1 \\ \int_{\partial M} u\, dS_g = 0}} \int_{\partial M} e^{\alpha u^2} dS_g = +\infty$$

for any $\alpha > \pi$. *Moreover, there is a function* $u \in C^\infty(\bar{M})$ *which satisfies that* $\int_M |\nabla u|^2 dV_g = 1$, $\int_{\partial M} u = 0$, *and*

$$\int_{\partial M} e^{\pi u^2} dS_g = \sup_{\substack{\int_M |\nabla v|^2 dV_g = 1 \\ \int_{\partial M} v\, dS_g = 0}} \int_{\partial M} e^{\pi v^2} dS_g.$$

Theorems C and D are proved by blowing-up analysis, a method closely related to those used by Schoen [1984] in his solution of the Yamabe problem, Escobar and Schoen [1986] for finding conformal metrics with prescribed curvatures in higher dimensions, and Ding, Jost, Li and Wang [Ding et al. 1997] in their solution of the differential equation $\Delta u = 8\pi - 8\pi h e^u$ on a compact Riemannian surface.

In this paper we study some trace inequalities similar to (1–2). Let

$$\mathcal{H}_1 = \left\{u \in H^{1,2}(M) : \textstyle\int_M |\nabla u|^2 dV_g = 1,\ \int_M u\, dV_g = 0\right\},$$
$$\mathcal{H}_2 = \left\{u \in H^{1,2}(M) : \textstyle\int_M (|\nabla u|^2 + u^2)\, dV_g = 1\right\}.$$

Theorem 1.1. *Let* (M, g) *be a compact Riemannian surface with boundary* ∂M. *Then*

$$\sup_{u \in \mathcal{H}_1} \int_{\partial M} e^{\pi u^2} dS_g < +\infty$$

and $\sup_{u \in \mathcal{H}_1} \int_{\partial M} e^{\alpha u^2} dS_g = +\infty$ *for any* $\alpha > \pi$. *Moreover, there is a function* $u \in C^\infty(\bar{M}) \cap \mathcal{H}_1$ *such that*

$$\int_{\partial M} e^{\pi u^2} dS_g = \sup_{v \in \mathcal{H}_1} \int_{\partial M} e^{\pi v^2} dS_g. \tag{1–3}$$

Our method to prove Theorem 1.1 is similar to that of [Li and Liu 2005]. Precisely speaking, we divide the proof into two steps. Firstly, for any $\varepsilon > 0$, let $u_\varepsilon \in \mathcal{H}_1$ be a maximizer of the functional

$$J_{\pi-\varepsilon}(u) = \int_{\partial M} e^{(\pi-\varepsilon)u^2} dS_g$$

on the space $\mathcal{H}_1$. Let G be a Green's function on M. Then G takes the form

$$G(x, p) = -\frac{1}{\pi} \log r(x) + A_p + O(r)$$

in a normal coordinate system around p, where $r(x) = \mathrm{dist}(x, p)$ and A_p is a constant. If the sequence $\{u_\varepsilon\}$ blows up, i.e.,

$$|u_\varepsilon|(x_\varepsilon) = \sup_{x \in M} |u_\varepsilon|(x) \to +\infty \quad \text{as} \quad \varepsilon \to 0,$$

we obtain

$$\sup_{u\in\mathcal{H}_1}\int_{\partial M} e^{\pi u^2}dS_g \le \operatorname{Vol}\partial M + 2\pi e^{\pi A_p}. \tag{1–4}$$

In the second step, we construct a blowing up sequence $\phi_\varepsilon \in \mathcal{H}_1$ such that

$$J_\pi(\phi_\varepsilon) = \int_{\partial M} e^{\pi\phi_\varepsilon^2}dS_g > \operatorname{Vol}\partial M + 2\pi e^{\pi A_p}$$

for sufficiently small ε. This contradicts step 1, and implies that blowing up cannot occur. The weak compactness of $L^p(M)$ $(p>1)$ gives the existence of the extremal function, i.e., (1–3) holds.

It should be mentioned that x_ε lies on ∂M naturally in [Li and Liu 2005] because u_ε is a harmonic function there. But in our case, passing to any subsequence, we cannot assume $x_\varepsilon \in \partial M$ and $u_\varepsilon(x_\varepsilon) \to +\infty$ simultaneously. Also, in the second step, the blowing up sequence we constructed (see Section 5) is different from that of [Li and Liu 2005].

Using the same idea described above, we also obtain:

Theorem 1.2. *Let* (M, g) *be as in Theorem 1.1. Then*

$$\sup_{u\in\mathcal{H}_2}\int_{\partial M} e^{\pi u^2}dS_g < +\infty$$

and $\sup_{u\in\mathcal{H}_2}\int_{\partial M} e^{\alpha u^2}dS_g = +\infty$ *for any* $\alpha > \pi$. *Moreover, there is a function* $u \in C^\infty(\overline{M})\cap\mathcal{H}_2$ *such that*

$$\int_{\partial M} e^{\pi u^2}dS_g = \sup_{u\in\mathcal{H}_2}\int_{\partial M} e^{\pi v^2}dS_g.$$

Clearly, Theorem C is a corollary of Theorem D. Similar results can also be derived from Theorems 1.1 and 1.2; for instance, we can substitute $\int_M u\,dV_g$ for $\int_{\partial M} u\,ds_g$, or $(1/4\pi)\,\|u\|_{H^{1,2}(M)}$ for $(1/4\pi)\int_M |\nabla u|^2 dV_g + \int_{\partial M} u\,ds_g$ in the right side of inequality (1–1). Theorems 1.1 and 1.2 are independent of Theorems C and D. They are more interesting than Theorem C because we obtain boundary estimates without direct boundary conditions.

For simplicity, we often omit the volume elements dV_g and dS_g when we write the integrals on M and ∂M respectively, and sometimes denote different constants by the same c. The reader can distinguish them easily from the context.

Most of the remainder of this paper is devoted to the proof of Theorem 1.1. In Section 2, we establish two regularity lemmas for use later. In Section 3, we prove that π is the best constant. And we derive an upper bound of $J_\pi(u)$ under the assumption that u_ε blows up in Section 4. A blowing up sequence ϕ_ε is constructed to reach a contradiction in Section 5, and this completes the proof of Theorem 1.1. In Section 6 we outline the proof of Theorem 1.2.

2. Regularity lemmas

Lemma 2.1. *Suppose $f \in L^q(M)$, $h \in H^{1,q}(M)$, $1 < q < 2$, and $2 < p < 2q/(2-q)$. let $u \in H^{1,2}(M)$ be a solution of the equation*

$$\begin{cases} \Delta u = f & \text{in } \mathring{M} \\ \dfrac{\partial u}{\partial n} = h & \text{on } \partial M, \end{cases}$$

where $\mathring{M}$ denotes the interior of M. Then u lies in $L^\infty(M)$ and we have

$$\|u\|_{L^\infty(M)} \le c\big(\|f\|_{L^q(M)} + \|h\|_{L^p(M)} + \|\nabla h\|_{L^q(M)} + \|u\|_{L^2(M)}\big),$$

where c is a constant depending only on M.

Proof. We use De Giorgi iteration. Choose a C^∞ vector field ζ whose restriction on ∂M is the outward unit normal vector field. By Stokes' theorem we have, for any $\varphi \in C^\infty(M)$,

$$\text{(2–1)} \qquad \begin{aligned} -\int_M \nabla u \nabla \varphi &= \int_M f\varphi - \int_{\partial M} \varphi \frac{\partial u}{\partial n} = \int_M f\varphi - \int_M \operatorname{div}(\varphi h \zeta) \\ &= \int_M \big(f - h \operatorname{div}\zeta - \langle \zeta, \nabla h\rangle_g\big)\varphi - \int_M h\langle \zeta, \nabla\varphi\rangle_g \\ &\equiv \int_M f^0 \varphi - \int_M \langle \vec{h}, \nabla \varphi\rangle_g, \end{aligned}$$

where $\langle\,\cdot\,,\cdot\,\rangle$ denotes the Riemannian inner product, $f^0 = f - h \operatorname{div}\zeta - \langle \zeta, \nabla h\rangle_g$, and $\vec{h} = h\zeta$. Clearly, $f^0 \in L^q(M)$ and $\vec{h} \in L^p(M)$.

For $0 < k < +\infty$, define $v_k = (u-k)^+$, $M_k = \{x \in M : v_k(x) > 0\}$. By Hölder's inequality,

$$\text{(2–2)} \qquad |M_k| \le \frac{\|u\|_{L^1(M)}}{k} \le \frac{|M|^{1/2}\|u\|_{L^2(M)}}{k},$$

where $|M_k|$ and $|M|$ represent the 2-dimensional measure of M_k and M respectively. Inserting $\varphi = v_k$ into (2–1), one has

(2–3)

$$\begin{aligned} \int_M |\nabla v_k|^2 &= \int_M \nabla u \nabla v_k = -\int_M f^0 v_k + \int_M \langle \vec{h}, \nabla v_k\rangle_g \\ &\le \Big(\int_M (f^0)^q\Big)^{1/q}\Big(\int_M v_k^{q'}\Big)^{1/q'} + \Big(\int_{M_k} |\vec{h}|^2\Big)^{1/2}\Big(\int_M |\nabla v_k|^2\Big)^{1/2}, \end{aligned}$$

where $|\cdot| = \sqrt{\langle\cdot,\cdot\rangle}$ and $1/q'+1/q=1$. Since $1<q<2$, we have $q'>2$. Choose r sufficiently large that $1/q'-1/r>1/2-1/p$ and $r(1/2-1/p)>1$. By the Sobolev imbedding theorem,

$$\|v_k\|_{L^r(M)} \le c(\|v_k\|_{L^2(M)} + \|\nabla v_k\|_{L^2(M)}) \le c(\|v_k\|_{L^r(M)}|M_k|^{1/2-1/r} + \|\nabla v_k\|_{L^2(M)}),$$

where c is a constant depending only on M.

Without loss of generality we assume that $\|u\|_{L^2(M)}=1$. According to (2–2), there exists a large integer number k_0 such that $c|M_k|^{1/2-1/r}<1$ for $k>k_0$. Hence

$$\|v_k\|_{L^r(M)} \le c\|\nabla v_k\|_{L^2(M)} \quad \text{for} \quad k \ge k_0. \tag{2–4}$$

By (2–3), we have

$$\begin{aligned}\|\nabla v_k\|^2_{L^2(M)} &\le \|f^0\|_{L^q(M)}\|v_k\|_{L^r(M)}|M_k|^{1/q'-1/r} + \|\vec{h}\|_{L^2(M_k)}\|\nabla v_k\|_{L^2(M)} \\ &\le c\|f^0\|_{L^q(M)}|M_k|^{1/q'-1/r}\|\nabla v_k\|_{L^2(M)} + \|\vec{h}\|_{L^2(M_k)}\|\nabla v_k\|_{L^2(M)},\end{aligned}$$

which gives

$$\|\nabla v_k\|_{L^2(M)} \le c\|f^0\|_{L^q(M)}|M_k|^{1/q'-1/r} + \|\vec{h}\|_{L^p(M)}|M_k|^{1/2-1/p}.$$

Note that $1/q'-1/r>1/2-1/p$. We have

$$\|\nabla v_k\|_{L^2(M)} \le c\tau|M_k|^{1/2-1/p}, \tag{2–5}$$

where $\tau = \|f^0\|_{L^q(M)} + \|\vec{h}\|_{L^p(M)}$.

On the other hand, for $h>k$, we have

$$\int_M v_k^r \ge \int_{M_h} (u-k)^r \ge |M_h|(h-k)^r.$$

Combining this with (2–4) and (2–5), we get $|M_h| \le K(h-k)^{-r}|M_k|^\beta$, with $K \equiv \tilde{c}\tau^r$ for some constant $\tilde{c}$, $\beta \equiv (1/2-1/p)r>1$, and $k_0<k<h<h_1<+\infty$ for any sufficiently large h_1. By [Troianiello 1987, Lemma 2.9], $|M_{k_0+\hat{k}}|=0$ for some $\hat{k}>0$; that is, $u \le k_0+\hat{k}$ in M. With the same argument, one can deduce that $-u \le k_0+\hat{k}$ in M. □

Theorem 3.17 of [Troianiello 1987] yields an immediate consequence:

Lemma 2.2. *Suppose that $f \in L^p(M)$ and $h \in H^{1,p}(M)$ for some $p \ge 2$, and that $u \in H^{1,2}(M)$ is a solution of*

$$\begin{cases} \Delta u = f & \text{in } \mathring{M} \\ \dfrac{\partial u}{\partial n} = h & \text{on } \partial M. \end{cases}$$

Then $u \in H^{2,p}(M)$.

3. The best constants

We now prove that the best constant in Theorem 1.1 is π. Here best means that

$$\sup_{u\in\mathscr{H}_1}\int_{\partial M} e^{\alpha u^2} < +\infty \quad \text{for } \alpha < \pi,$$
$$\sup_{u\in\mathscr{H}_1}\int_{\partial M} e^{\alpha u^2} = +\infty \quad \text{for } \alpha > \pi.$$

The following lemma is well known:

Lemma 3.1. *Let M be a compact Riemannian surface with boundary. Then there exists a positive number α such that* $\sup_{u\in\mathscr{H}_1}\int_M e^{\alpha u^2} < \infty$.

Lemma 3.2. *Set* $\alpha_2 = \sup\{\alpha : \sup_{u\in\mathscr{H}_1}\int_M e^{\alpha u^2} < +\infty\}$. *Then* $\alpha_2 = 2\pi$.

Proof. Step 1. We first prove that $\alpha_2 \geq 2\pi$.

Suppose $\alpha_2 < 2\pi$. There exists a sequence $u_\varepsilon \in \mathscr{H}_1$ such that $\int_M e^{(\alpha_2+\varepsilon)u_\varepsilon^2} \to +\infty$ as $\varepsilon \to 0$. One can see that there exists a $p \in M$ such that for any $r > 0$,

$$\int_{B_r(p)} e^{(\alpha_2+\varepsilon)u_\varepsilon^2} \to +\infty \quad \text{as} \quad \varepsilon \to 0, \tag{3–1}$$

where $B_r(p)$ is a geodesic ball centered at p with radius r. For otherwise, using a covering argument, one has $\int_M e^{(\alpha_2+\varepsilon)u_\varepsilon^2} \leq c$ for ε small enough, which contradicts the definition of u_ε. By the Poincaré inequality, $\{u_\varepsilon\}$ is bounded in $H^{1,2}(M)$, and so is $\{|u_\varepsilon|\}$. Hence there is $u \in H^{1,2}(M)$ such that $|u_\varepsilon| \rightharpoonup u$ (weak convergence) in $H^{1,2}(M)$ and $|u_\varepsilon| \to u$ (strong convergence) in $L^2(M)$ as $\varepsilon \to 0$. For any $\eta > 0$, we claim that

$$\lim_{\varepsilon\to 0}\int_M |\nabla(|u_\varepsilon| - \eta)^+|^2 = 1, \tag{3–2}$$

where $(|u_\varepsilon| - \eta)^+$ is the positive part of $|u_\varepsilon| - \eta$. Suppose (3–2) does not hold. Clearly, $\liminf_{\varepsilon\to 0}\int_M |\nabla(|u_\varepsilon| - \eta)^+|^2 < 1$. By the definition of α_2, passing to a subsequence, we can choose $\alpha' > \alpha_2$ such that

$$\int_M \exp\biggl(\alpha'\Bigl((|u_\varepsilon| - \eta)^+ - \frac{1}{\operatorname{Vol} M}\int_M (|u_\varepsilon| - \eta)^+\Bigr)^2\biggr) \leq c$$

for sufficiently small ε. Using the Poincaré inequality and the inequality $ab \leq \delta a^2 + b^2/(4\delta)$ for any $\delta > 0$, we can choose some $\varepsilon' > 0$ such that $\alpha'/(1+\varepsilon') > \alpha_2$ and $\int_M e^{\alpha' u_\varepsilon^2/(1+\varepsilon')} \leq c$, which contradicts (3–1) for ε small enough, and implies (3–2).

Let $v_\varepsilon = \min\{|u_\varepsilon|, \eta\}$. Then v_ε is bounded in $H^{1,2}(M)$. So there exists $v \in H^{1,2}(M)$ such that $v_\varepsilon \rightharpoonup v$ weakly in $H^{1,2}(M)$ and $v_\varepsilon \to v$ strongly in $L^2(M)$. Obviously,

$$|u_\varepsilon| = v_\varepsilon + (|u_\varepsilon| - \eta)^+. \tag{3–3}$$

Note that

$$1=\int_M |\nabla u_\varepsilon|^2 \ge \int_M \big|\nabla|u_\varepsilon|\big|^2 = \int_M |\nabla v_\varepsilon|^2 + \int_M \big|\nabla(|u_\varepsilon|-\eta)^+\big|^2.$$

By (3–2), we have $\int_M |\nabla v_\varepsilon|^2 \to 0$ as $\varepsilon \to 0$. By the Poincaré lemma, $\int_M |v_\varepsilon - \bar{v}_\varepsilon|^2 \to 0$ as $\varepsilon \to 0$, where $\bar{v}_\varepsilon = (\mathrm{Vol}\, M)^{-1} \int_M v_\varepsilon$. Note that $v_\varepsilon \to v$ strongly in $L^2(M)$. One has $v = \bar{v}$ almost everywhere in M. From (3–3), we know that

$$u = v + (u-\eta)^+ \quad \text{a.e. in } M.$$

By an appropriate choice of η, one easily derives that $v = 0$ and $u = 0$ a.e. in M.

Recall that $|u_\varepsilon| \to u$ strongly in $L^2(M)$. One has

(3–4) $$|u_\varepsilon| \to 0 \quad \text{strongly in } L^2(M) \quad \text{as } \varepsilon \to 0.$$

Now we turn to (3–1). Take $p \in \partial M$. Choose an isothermal coordinate system (U, ψ) around p such that $\psi : U \to \mathbb{B}^+_{2r}$. Choose a cut-off function $\varphi \in C^\infty(M)$ such that $\varphi \equiv 1$ on $B_r(p)$ and $\varphi \equiv 0$ outside $B_{4r/3}(p)$. By (3–4), we have

$$\int_{B_r(p)} |\nabla(\eta u_\varepsilon)|^2 \le \int_M |\nabla(\eta u_\varepsilon)|^2 \le 1+\varepsilon''$$

for some $\varepsilon'' > 0$ with $2\pi/(1+\varepsilon'') > \alpha_2$, provided that ε is sufficiently small. Define

$$\tilde{u}_\varepsilon(s,t) = \begin{cases} (\eta u_\varepsilon)(s,t) & \text{for } t \ge 0, \\ (\eta u_\varepsilon)(s,-t) & \text{for } t < 0. \end{cases}$$

Then $\int_{\mathbb{B}_{2r}} |\nabla \tilde{u}_\varepsilon|^2\, ds\, dt \le 2+2\varepsilon''$. By Moser's inequality, we then obtain the bound $\int_{\mathbb{B}_{2r}} e^{4\pi \tilde{u}_\varepsilon^2/(2+2\varepsilon')}\, ds\, dt \le c$. Hence

(3–5) $$\int_{B_r(p)} e^{2\pi u_\varepsilon^2/(1+\varepsilon')} \le 2\int_{\mathbb{B}_{2r}} e^{4\pi \tilde{u}_\varepsilon^2/(2+2\varepsilon')}\, ds\, dt \le 2c$$

for sufficiently small r. This contradicts (3–1).

When p is an interior point in M, one can get a contradiction as above without any difficulty. In this case, $\tilde{u}_\varepsilon$ is not needed any more; one need only consider u_ε itself. This completes the proof of step 1.

Step 2. To prove the opposite inequality, $\alpha_2 \le 2\pi$, take any $p \in \partial M$ and choose an isothermal coordinate system around p. Set

(3–6) $$u_\varepsilon = \begin{cases} -\sqrt{\dfrac{1}{2\pi}\log\dfrac{1}{\varepsilon}} & \text{in } B_{\delta\sqrt{\varepsilon}}(p), \\ \dfrac{\sqrt{2}}{\sqrt{-\pi\log\varepsilon}}\log\dfrac{r}{\delta} & \text{in } B_\delta(p)\setminus B_{\delta\sqrt{\varepsilon}}(p), \\ C_\varepsilon \varphi & \text{in } M\setminus B_\delta(p), \end{cases}$$

where $\varphi \in C_0^\infty(M \backslash B_\delta(p))$, $0 \le \varphi \le 1$, and C_ε is chosen to satisfy $\int_M u_\varepsilon = 0$. It is easy to check that

$$\int_M |\nabla u_\varepsilon|^2 \to 1 \ \text{ as } \ \varepsilon \to 0$$

and that

$$\int_M \exp\left(\alpha\left(\frac{u_\varepsilon}{\|u_\varepsilon\|_{L^2}}\right)^2\right) \ge \exp\left(\frac{\alpha}{2\pi \|\nabla u_\varepsilon\|_{L^2}^2} \log \frac{1}{\varepsilon}\right) \operatorname{Vol} B_{\delta\sqrt{\varepsilon}} \ge C\varepsilon^{1-\alpha/(2\pi\|\nabla u_\varepsilon\|_{L^2}^2)}$$

for any $\alpha > 2\pi$; the latter lower bound approaches $+\infty$ as $\varepsilon \to 0$. Therefore $\alpha_2 \le 2\pi$. □

Lemma 3.3. *Set* $J_\alpha(u) = \int_{\partial M} e^{\alpha u^2}$. *Then*

$$\sup_{u\in\mathcal{H}_1} J_\alpha(u) < +\infty \quad \textit{for } \alpha < \pi \quad \textit{and} \quad \sup_{u\in\mathcal{H}_1} J_\alpha(u) = +\infty \quad \textit{for } \alpha > \pi.$$

Proof. Take a smooth vector field ζ whose restriction on ∂M is the outward unit normal vector field. Using the divergence theorem and Lemma 3.2, one has

$$\begin{aligned}\int_{\partial M} e^{(\pi-\varepsilon)u^2} &= \int_M \operatorname{div}(\zeta e^{(\pi-\varepsilon)u^2}) = \int_M \big(\operatorname{div}(\zeta) + 2(\pi-\varepsilon)u\langle\zeta, \nabla u\rangle_g\big)e^{(\pi-\varepsilon)u^2} \\ &\le C\left(1 + \int_M |\nabla u||u|e^{(\pi-\varepsilon)u^2}\right) \\ &\le C\left(1 + \|\nabla u\|_{L^2(M)} \|u\|_{L^p(M)} \|e^{(\pi-\varepsilon)u^2}\|_{L^{(2\pi-\varepsilon)/(\pi-\varepsilon)}(M)}\right)\end{aligned}$$

for all $u \in \mathcal{H}_1$, where $1/p + 1/2 + (\pi-\varepsilon)/(2\pi-\varepsilon) = 1$. Combining this estimate with the Sobolev imbedding theorem, one has $\sup_{u\in\mathcal{H}_1} J_{\pi-\varepsilon}(u) < +\infty$ for any $\varepsilon > 0$, which implies that $\sup_{u\in\mathcal{H}_1} J_\alpha(u) < +\infty$ for any $\alpha < \pi$.

To complete the proof of the lemma, we employ (3–6) to check that for any $\alpha > \pi$, $J_\alpha(u_\varepsilon)$ diverges to $+\infty$ as $\varepsilon \to 0$. □

4. Blowing up analysis

We now use the method of blowing up to prove (1–4). The same method has also been used in [Li 2001; Li 2005].

The proof consists of several lemmas.

Lemma 4.1. *The functional* $J_{\pi-\varepsilon}(u)$ *defined in the space* $\mathcal{H}_1$ *admits a smooth maximizer* $u_\varepsilon \in \mathcal{H}_1$.

Proof. It is obvious that there exists $u_\varepsilon \in \mathcal{H}_1$ such that

$$J_{\pi-\varepsilon}(u_\varepsilon) = \sup_{u\in\mathcal{H}_1} J_{\pi-\varepsilon}(u).$$

The function u_ε satisfies the Euler–Lagrange equation

$$\begin{cases} \Delta u_\varepsilon = \dfrac{\mu_\varepsilon}{2\lambda_\varepsilon} & \text{in } \mathring{M}, \\ \dfrac{\partial u_\varepsilon}{\partial n} = \dfrac{\pi-\varepsilon}{\lambda_\varepsilon} u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} & \text{on } \partial M, \end{cases} \tag{4–1}$$

where

$$\lambda_\varepsilon = (\pi-\varepsilon)\int_{\partial M} u_\varepsilon^2 e^{(\pi-\varepsilon)u_\varepsilon^2} \quad \text{and} \quad \mu_\varepsilon = \frac{2(\pi-\varepsilon)}{\operatorname{Vol} M}\int_{\partial M} u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2}. \tag{4–2}$$

Write $h(u_\varepsilon) = (\pi-\varepsilon)/\lambda_\varepsilon u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2}$. By the Orlicz space imbedding (see [Struwe 1988]), $e^{u_\varepsilon^2} \in L^p(M)$ for any $p > 0$. Hence $h(u_\varepsilon) \in H^{1,q}(M)$ for any $1 < q < 2$. By Lemma 2.1 we have $u_\varepsilon \in L^\infty(M)$, hence $h(u_\varepsilon) \in H^{1,2}(M)$. By Lemma 2.2, $u_\varepsilon \in H^{2,2}(M)$. The Sobolev imbedding theorem then implies that $h(u_\varepsilon) \in H^{1,p}(M)$ for some $p > 2$. Again, by Lemma 2.2, $u_\varepsilon \in H^{2,p}(M)$. The Sobolev imbedding theorem gives $u_\varepsilon \in C^1(M)$. Using Lemma 2.2 repeatedly, we conclude that $u_\varepsilon \in C^\infty(M)$. □

Lemma 4.2. $\liminf_{\varepsilon\to 0} \lambda_\varepsilon > 0$.

Proof. The following estimate is elementary

$$\operatorname{Vol}\partial M < \sup_{u\in\mathcal{H}_1}\int_{\partial M} e^{\pi u^2} = \lim_{\varepsilon\to 0}\int_{\partial M} e^{(\pi-\varepsilon)u_\varepsilon^2} \le \operatorname{Vol}\partial M + \liminf_{\varepsilon\to 0}\lambda_\varepsilon,$$

which gives $\liminf_{\varepsilon\to 0} \lambda_\varepsilon > 0$. □

Lemma 4.3. *$\mu_\varepsilon/\lambda_\varepsilon$ is bounded with respect to ε.*

Proof. By (4–2) and Lemma 4.2, we have

$$\frac{|\mu_\varepsilon|}{\lambda_\varepsilon} \le \frac{2(\pi-\varepsilon)}{\operatorname{Vol} M}\int_{\partial M}\frac{|u_\varepsilon|}{\lambda_\varepsilon} e^{(\pi-\varepsilon)u_\varepsilon^2} \le \frac{2(\pi-\varepsilon)}{\operatorname{Vol} M}\left(\frac{e^{\pi-\varepsilon}}{\lambda_\varepsilon} + \frac{1}{\pi-\varepsilon}\right) \le C. \qquad \square$$

Write $c_\varepsilon = |u_\varepsilon|(x_\varepsilon) = \max_{x\in M}(x)$. If $\{c_\varepsilon\}$ is bounded, then by the standard elliptic estimate with respect to Equation (4–1), there exists $u \in \mathcal{H}_1 \cap C^\infty(M)$ such that $u_\varepsilon \to u$ in $C^\infty(M)$ as $\varepsilon \to 0$, and Theorem 1.1 follows immediately. Henceforth we assume $c_\varepsilon \to +\infty$ as $\varepsilon \to 0$.

Passing to a subsequence, we may assume that $\mu_\varepsilon \ge 0$ for all $\varepsilon > 0$, for otherwise we consider $-u_\varepsilon$ instead of u_ε in (4–1)–(4–2). We consider separately the possibilities that $\{u_\varepsilon(x_\varepsilon)\}$ approaches $+\infty$ or $-\infty$ or as $\varepsilon \to 0$.

Take first the case $u_\varepsilon(x_\varepsilon) \to +\infty$. Applying the maximum principle to (4–1), we see that $x_\varepsilon \in \partial M$. Passing to a subsequence, we may assume $x_\varepsilon \to p$ for some $p \in \partial M$.

Lemma 4.4. *Define*

$$r_\varepsilon = \frac{1}{\pi-\varepsilon}\frac{\lambda_\varepsilon}{c_\varepsilon^2} e^{-(\pi-\varepsilon)c_\varepsilon^2}. \tag{4–3}$$

Then $r_\varepsilon c_\varepsilon \to 0$ *as* $\varepsilon \to 0$.

Proof. By the first equality in (4–2), we have

$$1 = \frac{\pi-\varepsilon}{\lambda_\varepsilon}\int_{\partial M} u_\varepsilon^2 e^{(\pi-\varepsilon)u_\varepsilon^2} \le \frac{\pi-\varepsilon}{\lambda_\varepsilon} e^{\pi c_\varepsilon^2}\int_{\partial M} u_\varepsilon^2 \le c\frac{\pi-\varepsilon}{\lambda_\varepsilon}e^{\pi c_\varepsilon^2}$$

for some constant c, where we have used the Sobolev trace imbedding theorem. This implies that $r_\varepsilon c_\varepsilon \to 0$ as $\varepsilon \to 0$. □

Choose an isothermal coordinate system (U, ϕ) near p such that $\phi(p) = 0$, ϕ maps U to $\mathbb{R}^2_+ := \{x = (x_1, x_2) \in \mathbb{R}^2 : x_2 > 0\}$ and $\phi(U \cap \partial M) \subset \partial\mathbb{R}^2_+$.

Set

$$\psi_\varepsilon(x) = u_\varepsilon(x_\varepsilon + r_\varepsilon x)/c_\varepsilon, \qquad \varphi_\varepsilon(x) = c_\varepsilon\big(u_\varepsilon(x_\varepsilon + r_\varepsilon x) - c_\varepsilon\big). \tag{4–4}$$

Lemma 4.5. $\psi_\varepsilon \to 1$ *in* $C^2_{\text{loc}}(\overline{\mathbb{R}^2_+})$ *as* $\varepsilon \to 0$.

Proof. By (4–1), for ε is sufficiently small we have

$$\begin{cases} \Delta\psi_\varepsilon = \dfrac{r_\varepsilon^2}{c_\varepsilon}\dfrac{\mu_\varepsilon}{2\lambda_\varepsilon} & \text{in } B_R^+(0), \\ \dfrac{\partial\psi_\varepsilon}{\partial n} = \dfrac{r_\varepsilon}{c_\varepsilon}\dfrac{\pi-\varepsilon}{\lambda_\varepsilon}u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} & \text{on } B_R(0)\cap\partial\mathbb{R}^2_+, \end{cases}$$

for any $R > 0$. As in the proof of Lemma 4.1, it is not hard to see that $\psi_\varepsilon \to 1$ in $C^2(\overline{B^+_{R/2}(0)})$ as $\varepsilon \to 0$. □

Lemma 4.6. *The functions* φ_ε *converge in* $C^2_{\text{loc}}(\overline{\mathbb{R}^2_+})$ *as* $\varepsilon \to 0$ *to some* φ *satisfying*

$$\begin{cases} -\Delta_{\mathbb{R}^2}\varphi = 0 & \text{in } \mathbb{R}^2_+ \\ \dfrac{\partial\varphi}{\partial n} = e^{2\pi\varphi} & \text{on } \partial\mathbb{R}^2_+, \\ \varphi(0) = \sup\varphi = 0. \end{cases}$$

Proof. By (4–1), we have

$$\begin{cases} \Delta\varphi_\varepsilon(x) = c_\varepsilon r_\varepsilon^2\dfrac{\mu_\varepsilon}{2\lambda_\varepsilon} & \text{in } B_R^+(0), \\ \dfrac{\partial\varphi_\varepsilon}{\partial n} = \dfrac{u_\varepsilon}{c_\varepsilon}\exp\Big((\pi-\varepsilon)\varphi_\varepsilon\Big(1+\dfrac{u_\varepsilon}{c_\varepsilon}\Big)\Big) & \text{on } \partial\mathbb{R}^2_+\cap B_R(0) \end{cases}$$

for any $R > 0$. Using Lemma 2.2, we have $\varphi_\varepsilon \to \varphi$ in $C^2(\overline{B^+_{R/2}(0)})$ as $\varepsilon \to 0$ for some $u \in C^2(B^+_{R/2}(0))$. Clearly u satisfies the required conditions. □

It is not difficult to see that

$$\int_{B_R(0)\cap\partial\mathbb{R}^2_+} e^{2\pi\varphi} \le \liminf_{\varepsilon\to 0}\int_{B_{Rr_\varepsilon}(x_\varepsilon)\cap\partial M}\frac{\pi-\varepsilon}{\lambda_\varepsilon}u_\varepsilon^2 e^{(\pi-\varepsilon)u_\varepsilon^2}\le 1,$$

which gives

$$\int_{\partial\mathbb{R}^2_+} e^{2\pi\varphi}\le 1.$$

By a result in [Li and Zhu 1995], we have

$$\varphi(x) = -\frac{1}{2\pi}\log\bigl(\pi^2x_1^2+(1+\pi x_2)^2\bigr).$$

A direct calculation gives

$$\int_{\partial\mathbb{R}^2_+} e^{2\pi\varphi} = 1.$$

Following [Li and Liu 2005], we define $u_\varepsilon^c = \min\{\frac{c_\varepsilon}{c}, u_\varepsilon\}$.

Lemma 4.7. *For any* $c>1$, *we have* $\lim_{\varepsilon\to 0}\int_M |\nabla u_\varepsilon^c|^2 = \frac{1}{c}$.

Proof. Using Stokes' formula, (4–1) and Lemma 4.5, we have

$$\begin{aligned}\int_M\Bigl|\nabla\Bigl(u_\varepsilon-\frac{c_\varepsilon}{c}\Bigr)^+\Bigr|^2 &= \int_M \nabla u_\varepsilon\nabla\Bigl(\Bigl(u_\varepsilon-\frac{c_\varepsilon}{c}\Bigr)^+\Bigr)\\ &= \int_{\partial M}\Bigl(u_\varepsilon-\frac{c_\varepsilon}{c}\Bigr)^+\frac{\partial u_\varepsilon}{\partial n} - \int_M\Bigl(u_\varepsilon-\frac{c_\varepsilon}{c}\Bigr)^+\Delta u_\varepsilon\\ &= \int_{\partial M}\Bigl(u_\varepsilon-\frac{c_\varepsilon}{c}\Bigr)^+\frac{\pi-\varepsilon}{\lambda_\varepsilon}u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} - \int_M\Bigl(u_\varepsilon-\frac{c_\varepsilon}{c}\Bigr)^+\frac{\mu_\varepsilon}{2\lambda_\varepsilon}\\ &\ge \int_{\partial M\cap B_{Rr_\varepsilon(x_\varepsilon)}}\bigl(u_\varepsilon-c_\varepsilon/c\bigr)^+\frac{\pi-\varepsilon}{\lambda_\varepsilon}u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2}+o_\varepsilon(1)\\ &= \frac{c-1}{c}\int_{\partial\mathbb{R}^{2+}\cap B_R(0)} e^{2\pi\varphi}+o_\varepsilon(R)+o_\varepsilon(1),\end{aligned}$$

where $o_\varepsilon(1)\to 0$ as $\varepsilon\to 0$, and $o_\varepsilon(R)\to 0$ for any fixed R as $\varepsilon\to 0$. Letting $\varepsilon\to 0$ first, and then $R\to+\infty$, we obtain

$$\liminf_{\varepsilon\to 0}\int_M\Bigl|\nabla\Bigl(u_\varepsilon-\frac{c_\varepsilon}{c}\Bigr)^+\Bigr|^2\ge\frac{c-1}{c}.$$

With the same argument, we get

$$\liminf_{\varepsilon\to 0}\int_M|\nabla u_\varepsilon^c|^2\ge\frac{1}{c}.$$

Note that since
$$\int_M \left|\nabla\left(u_\varepsilon - \frac{c_\varepsilon}{c}\right)^+\right|^2 + \int_M |\nabla u_\varepsilon^c|^2 = 1,$$
we have $\liminf_{\varepsilon\to 0} \int_M |\nabla u_\varepsilon^c|^2 = c^{-1}$. □

Lemma 4.8. *Under the assumption that $c_\varepsilon \to +\infty$ as $\varepsilon \to 0$, we have the estimate*
$$\sup_{u\in\mathcal{H}_1} J_\pi(u) \le \operatorname{Vol}\partial M + \frac{1}{\pi}\limsup_{\varepsilon\to 0}\frac{\lambda_\varepsilon}{c_\varepsilon^2}.$$

Proof. For any $c > 1$, we have
$$\begin{aligned}\int_{\partial M} e^{(\pi-\varepsilon)u_\varepsilon^2} &= \int_{\partial M\cap\{u_\varepsilon\le c_\varepsilon/c\}} e^{(\pi-\varepsilon)u_\varepsilon^2} + \int_{\partial M\cap\{u_\varepsilon> c_\varepsilon/c\}} e^{(\pi-\varepsilon)u_\varepsilon^2}\\ &\le \int_{\partial M} e^{(\pi-\varepsilon)(u_\varepsilon^c)^2} + c^2\frac{\lambda_\varepsilon}{c_\varepsilon^2}\int_{\partial M}\frac{u_\varepsilon^2}{\lambda_\varepsilon}e^{(\pi-\varepsilon)u_\varepsilon^2}.\end{aligned}$$

By Lemma 4.7, according to step 1 in the proof of Lemma 3.2, one can see that $u_\varepsilon^c \to 0$ a.e. in M as $\varepsilon \to 0$. Substituting u_ε^c for u in (3–5), one immediately has
$$\int_{\partial M} e^{(\pi-\varepsilon)(u_\varepsilon^c)^2} \to \operatorname{Vol}\partial M \quad \text{as} \quad \varepsilon\to 0.$$
Hence
$$\sup_{u\in\mathcal{H}_1} J_\pi(u) = \lim_{\varepsilon\to 0}\int_{\partial M} e^{(\pi-\varepsilon)u_\varepsilon^2} \le \operatorname{Vol}\partial M + \frac{c^2}{\pi}\limsup_{\varepsilon\to 0}\frac{\lambda_\varepsilon}{c_\varepsilon}.$$
Letting $c \to 1$, the conclusion of the lemma follows. □

The next result is an immediate consequence of Lemma 4.8:

Corollary 4.9. $\lambda_\varepsilon/c_\varepsilon \to +\infty$ *as* $\varepsilon \to 0$.

Lemma 4.10. *For any $\phi \in C^\infty(\partial M)$, we have*

$$\lim_{\varepsilon\to 0}\int_{\partial M}\phi\,\frac{\pi-\varepsilon}{\lambda_\varepsilon}c_\varepsilon u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} = \phi(p). \tag{4–5}$$

Proof. For any fixed $c > 1$, we partition ∂M into its intersections with
$$D_1 = \left(\left\{u_\varepsilon > \frac{c_\varepsilon}{c}\right\}\setminus B_{Rr_\varepsilon}(x_\varepsilon)\right),\quad D_2 = \left(\left\{u_\varepsilon \le \frac{c_\varepsilon}{c}\right\}\setminus B_{Rr_\varepsilon}(x_\varepsilon)\right),\quad D_3 = B_{Rr_\varepsilon}(x_\varepsilon).$$
Denote by I_1, I_2, I_3 the partial integrals in (4–5) taken over D_1, D_2, D_3. Then
$$\begin{aligned}|I_1| &\le c\sup_{\partial M}|\varphi|\int_{\partial M\cap(\{u_\varepsilon>\frac{c_\varepsilon}{c}\}\setminus B_{Rr_\varepsilon}(x_\varepsilon))}\frac{\pi-\varepsilon}{\lambda_\varepsilon}u_\varepsilon^2 e^{(\pi-\varepsilon)u_\varepsilon^2}\\ &\le c\sup_{\partial M}|\varphi|\left(1-\int_{\partial M\cap B_{Rr_\varepsilon}(x_\varepsilon)}\frac{\pi-\varepsilon}{\lambda_\varepsilon}u_\varepsilon^2 e^{(\pi-\varepsilon)u_\varepsilon^2}\right)\\ &\le c\sup_{\partial M}|\varphi|\left(1-\int_{\partial B_R^+(0)\cap\partial\mathbb{R}_+^2} e^{2\pi\varphi} + o_\varepsilon(R)\right),\end{aligned}$$

where $o_\varepsilon(R) \to 0$ as $\varepsilon \to 0$ for any fixed R. Letting $\varepsilon \to 0$ first, and then $R \to +\infty$, one has $I_1 \to 0$. Next,

$$\begin{aligned}|I_2| &\leq (\pi-\varepsilon)\sup_{\partial M}|\varphi|\frac{c_\varepsilon}{\lambda_\varepsilon}\int_{\partial M}|u_\varepsilon|e^{(\pi-\varepsilon)(u_\varepsilon^c)^2}\\ &\leq \pi\sup_{\partial M}|\varphi|\frac{c_\varepsilon}{\lambda_\varepsilon}\|u_\varepsilon\|_{L^{(c+1)/(c-1)}(\partial M)}\|e^{(\pi-\varepsilon)(u_\varepsilon^c)^2}\|_{L^{(c+1)/2}(\partial M)}\\ &\leq \tilde{C}\sup_{\partial M}|\varphi|\frac{c_\varepsilon}{\lambda_\varepsilon},\end{aligned}$$

where $\tilde{C}$ is a constant depending on M and c, here we have used Hölder's inequality and Sobolev imbedding theorem. By Corollary 4.9, we get $I_2 \to 0$ as $\varepsilon \to 0$. Finally,

$$\begin{aligned}I_3 &= \int_{\partial M\cap B_{Rr_\varepsilon}(x_\varepsilon)}\varphi\,\frac{\pi-\varepsilon}{\lambda_\varepsilon}\,c_\varepsilon u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} = \int_{\partial\mathbb{R}^2_+\cap\partial B_R^+(0)}\varphi\,\frac{u_\varepsilon}{c_\varepsilon}\,e^{(\pi-\varepsilon)\varphi_\varepsilon(1+u_\varepsilon/c_\varepsilon)}\\ &= \varphi(p)\bigg(\int_{\partial B_R^+(0)\cap\partial\mathbb{R}^2_+}e^{2\pi\varphi}+o_\varepsilon(R)\bigg).\end{aligned}$$

As before, letting $\varepsilon \to 0$ first, then $R \to +\infty$, we get $I_3 \to \varphi(p)$. Combining all three estimates, we get the conclusion of the lemma. □

Lemma 4.11. $|\nabla u_\varepsilon|^2 \rightharpoonup \delta_p$ *weakly in the sense of measure.*

Proof. Set

$$A = \Big\{q \in M : \lim_{r\to 0}\liminf_{\varepsilon\to 0}\int_{B_r(q)}|\nabla u_\varepsilon|^2 > 0\Big\}.$$

We claim that A contains only one point.

Suppose not. Then, for any $q \in M$, we have $\lim_{r\to 0}\liminf_{\varepsilon\to 0}\int_{B_r(q)}|\nabla u_\varepsilon|^2 < 1$. There exist positive numbers r and δ such that

$$\int_{B_r(q)}|\nabla u_\varepsilon|^2 \leq \delta(q) < 1.$$

With the assumption $c_\varepsilon \to +\infty$ as $\varepsilon \to 0$, step 1 in the proof of Lemma 3.2 implies that $u_\varepsilon \to 0$ in $L^2(M)$, and hence $\int_{B_r(q)} u_\varepsilon \to 0$ as $\varepsilon \to 0$. It is not difficult to see that there exists a constant $\alpha(q) > \pi$ such that

$$\int_{\partial M\cap\partial B_r(q)}e^{\alpha(q)u_\varepsilon^2} \leq C_q$$

for some constant C_q depending on q. By a covering argument, there exists an $\alpha > \pi$ such that

$$\int_{\partial M}e^{\alpha u_\varepsilon^2} \leq C$$

for some constant C. This contradicts the choice of u_ε, and our claim follows.

Next we claim that $A=\{p\}$. Let q be the unique point in A, and suppose $q\neq p$. Choose a smooth function ψ such that $\psi(p)\neq\psi(q)$. By Stokes' theorem and Equation (4–1), we have

$$\int_M \psi|\nabla u_\varepsilon|^2=\int_{\partial M}\psi\frac{\pi-\varepsilon}{\lambda_\varepsilon}u_\varepsilon^2 e^{(\pi-\varepsilon)u_\varepsilon^2}-\int_M u_\varepsilon\frac{\mu_\varepsilon}{2\lambda_\varepsilon}-\int_M u_\varepsilon\nabla\psi\nabla u_\varepsilon.$$

Clearly the last two terms here tend to 0 as $\varepsilon\to 0$. As in the proof of Lemma 4.10, we can show that

$$\lim_{\varepsilon\to 0}\int_{\partial M}\psi\frac{\pi-\varepsilon}{\lambda_\varepsilon}u_\varepsilon^2 e^{(\pi-\varepsilon)u_\varepsilon^2}=\psi(p).$$

On the other hand, $\lim_{\varepsilon\to 0}\int_M\psi|\nabla u_\varepsilon|^2=\psi(q)$. Hence $\psi(p)=\psi(q)$, which contradicts the choice of ψ. This completes the proof of the lemma. □

Lemma 4.12. *$c_\varepsilon u_\varepsilon\rightharpoonup G$ weakly in $H^{1,q}(M)$ for any $q:1<q<2$. For any $\Omega\Subset M\setminus\{p\}$, we have $c_\varepsilon u_\varepsilon\to G$ in $C^\infty(\overline{\Omega})$, where G satisfies*

$$\begin{cases}-\Delta G=\delta_p-\dfrac{1}{\operatorname{Vol} M} & \text{in } M,\\ \displaystyle\int_M G=0, \quad \frac{\partial G}{\partial n}\Big|_{\partial M\setminus\{p\}}=0. \end{cases}\tag{4–6}$$

Proof. By Equation (4–1), we have

$$\begin{cases}\Delta(c_\varepsilon u_\varepsilon)=c_\varepsilon\dfrac{\mu_\varepsilon}{2\lambda_\varepsilon} & \text{in } M,\\ \dfrac{\partial(c_\varepsilon u_\varepsilon)}{\partial n}=\dfrac{\pi-\varepsilon}{\lambda_\varepsilon}c_\varepsilon u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} & \text{on } \partial M.\end{cases}$$

Integrating both sides on M, one has

$$\int_M c_\varepsilon\frac{\mu_\varepsilon}{2\lambda_\varepsilon}=\int_M\Delta(c_\varepsilon u_\varepsilon)=\int_{\partial M}\frac{\pi-\varepsilon}{\lambda_\varepsilon}c_\varepsilon u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2}.$$

By Lemma 4.10, we immediately get $c_\varepsilon\mu_\varepsilon/(2\lambda_\varepsilon)\to 1/\operatorname{Vol} M$ as $\varepsilon\to 0$.

For any q in the range $1<q<2$, denote its conjugate by q', so $1/q+1/q'=1$. It is well known that

$$\int_M|\nabla(c_\varepsilon u_\varepsilon)|^q\le\sup\left\{\int_M\nabla\phi\nabla(c_\varepsilon u_\varepsilon)\,dV_g:\|\phi\|_{H^{1,q'}}=1\right\}.$$

The Sobolev embedding theorem yields $\|\phi\|_{C^0(M)} \le C$, where C is a constant depending only on M. Using the divergence theorem and (4–1), we have

$$\begin{aligned}\int_M \nabla\phi\,\nabla(c_\varepsilon u_\varepsilon) &= \int_{\partial M} \phi\,\frac{\partial(c_\varepsilon u_\varepsilon)}{\partial n} - \int_M \phi\Delta(c_\varepsilon u_\varepsilon)\\ &= \int_{\partial M} \phi\frac{\pi-\varepsilon}{\lambda_\varepsilon} c_\varepsilon u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} - c_\varepsilon\frac{\mu_\varepsilon}{2\lambda_\varepsilon}\int_M \phi.\end{aligned}$$

By Lemma 4.10 again, we obtain

$$\int_M |\nabla(c_\varepsilon u_\varepsilon)|^q \le C\|\phi\|_{C^0(M)} \le C.$$

This, together with Poincaré's inequality, implies that $c_\varepsilon u_\varepsilon$ is bounded in $H^{1,q}(M)$. Hence there exists $G \in H^{1,q}(M)$ such that $c_\varepsilon u_\varepsilon \rightharpoonup G$ weakly in $H^{1,q}(M)$ as $\varepsilon \to 0$. For any $\phi \in C^\infty(M)$, we have

$$\begin{aligned}\int_M \nabla\phi\nabla(c_\varepsilon u_\varepsilon) &= \int_{\partial M} \phi\frac{\partial(c_\varepsilon u_\varepsilon)}{\partial n} - \int_M \phi\Delta(c_\varepsilon u_\varepsilon)\\ &= \int_{\partial M} \phi\frac{\pi-\varepsilon}{\lambda_\varepsilon} c_\varepsilon u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} - c_\varepsilon\frac{\mu_\varepsilon}{2\lambda_\varepsilon}\int_M \phi\\ &\longrightarrow \phi(p) - \frac{1}{\operatorname{Vol} M}\int_M \phi \quad \text{as } \varepsilon \to 0.\end{aligned}$$

Hence

$$\int_M \nabla G\,\nabla\phi = \phi(p) - \frac{1}{\operatorname{Vol} M}\int_M \phi,$$

and Equation (4–6) holds.

For any $\Omega \Subset M \setminus \{p\}$, we choose a smooth function η on M such that $\eta \equiv 1$ on Ω, and $\eta \equiv 0$ near p. By Lemma 4.11, $\eta u_\varepsilon \to 0$ in $L^2(M)$ as $\varepsilon \to 0$. This, together with the convergence $u_\varepsilon \to 0$ in $L^2(M)$ as $\varepsilon \to 0$, implies that $e^{(\pi-\varepsilon)u_\varepsilon^2}$ is uniformly bounded in $L^r(\overline{\Omega})$ with respect to ε for any $r > 1$. Standard elliptic estimates imply that $c_\varepsilon u_\varepsilon \to G$ in $C^k(\overline{\Omega})$ for any positive integer k. This completes the proof of the lemma. □

In the following, we use the capacity technique to derive the upper bound of $J_\pi(u)$. Take an isothermal coordinate system (U, ϕ) near p such that $\phi(p) = 0$ and ϕ maps U inside $\mathbb{R}^2_+$ and $U \cap \partial M$ inside $\partial\mathbb{R}^2_+$. In this coordinate system we can write $g = e^{2f}(dx_1^2 + dx_2^2)$, with $f(0) = 0$. Set $\phi(x_\varepsilon) = (x_\varepsilon^1, 0)$. Let $\mathbb{B}_r = \mathbb{B}_r(x_\varepsilon^1, 0) \subset \mathbb{R}^2$ be the standard ball centered at $(x_\varepsilon^1, 0)$ with radius r. Define

$$i_\varepsilon = \inf_{\partial\mathbb{B}^+_{Rr_\varepsilon}\setminus\partial\mathbb{R}^2_+} u_\varepsilon\circ\phi^{-1}, \quad s_\varepsilon = \sup_{\partial\mathbb{B}^+_{\delta}\setminus\partial\mathbb{R}^2_+} u_\varepsilon\circ\phi^{-1}, \quad \tilde{u}_\varepsilon = \max\{s_\varepsilon, \min\{u_\varepsilon\circ\phi^{-1}, i_\varepsilon\}\}.$$

Clearly,

$$\int_{\mathbb{B}_\delta^+\setminus\mathbb{B}_{Rr_\varepsilon}^+}|\nabla\tilde{u}_\varepsilon|^2\le\int_{\phi^{-1}(B_\delta^+)\setminus\phi^{-1}(B_{Rr_\varepsilon}^+)}|\nabla u_\varepsilon|^2 \tag{4–7}$$

$$\le 1-\int_{\phi^{-1}(B_\delta^+)}|\nabla u_\varepsilon|^2-\int_{\phi^{-1}(B_{Rr_\varepsilon}^+)}|\nabla u_\varepsilon|^2.$$

Define a function space

$$\Lambda_\varepsilon=\Big\{u\in H^{1,2}(\mathbb{B}_\delta^+\setminus\mathbb{B}_{Rr_\varepsilon}^+)\;:\;u|_{\partial\mathbb{B}_\delta^+\setminus\partial\mathbb{R}_+^2}=s_\varepsilon,\;u|_{\partial\mathbb{B}_{Rr_\varepsilon}^+\setminus\partial\mathbb{R}_+^2}=i_\varepsilon,\;\frac{\partial u}{\partial n}\Big|_{\partial\mathbb{R}_+^2}=0\Big\}.$$

It is easy to see that $\inf_{u\in\Lambda_\varepsilon}\int_{\mathbb{B}_\delta^+\setminus\mathbb{B}_{Rr_\varepsilon}^+}|\nabla_{\mathbb{R}^2}u|^2$ is attained by the unique solution of the equation

$$\begin{cases}\Delta\Phi=0 \quad \text{in} \quad \mathbb{B}_\delta^+\setminus\mathbb{B}_{Rr_\varepsilon}^+\\ \Phi\in\Lambda_\varepsilon.\end{cases}$$

One can check that

$$\Phi=\frac{s_\varepsilon(\log r-\log(Rr_\varepsilon))+i_\varepsilon(\log\delta-\log r)}{\log\delta-\log(Rr_\varepsilon)},$$

whence

$$\int_{\mathbb{B}_\delta^+\setminus\mathbb{B}_{Rr_\varepsilon}^+}|\nabla\Phi|^2=\pi\frac{(s_\varepsilon-i_\varepsilon)^2}{\log\delta-\log Rr_\varepsilon}. \tag{4–8}$$

By Lemma 4.7, we have

$$\int_{\phi^{-1}(\mathbb{B}_{Rr_\varepsilon}^+)}|\nabla u_\varepsilon|^2=\frac{1}{c_\varepsilon^2}\Big(\frac{1}{\pi}\log R+\frac{1}{\pi}\log\frac{\pi}{2}+O\Big(\frac{\log R}{R}\Big)+o_\varepsilon(1)\Big).$$

Lemma 4.12 then yields

$$\int_{M\setminus\phi^{-1}(B_\delta^+)}|\nabla u_\varepsilon|^2=\frac{1}{c_\varepsilon^2}\Big(-\frac{1}{\pi}\log\delta+A_p+O(\delta\log\delta)+o_\varepsilon(1)\Big). \tag{4–9}$$

By (4–7) and (4–8), we have

$$\frac{\pi s_\varepsilon^2-2\pi s_\varepsilon i_\varepsilon+\pi i_\varepsilon^2}{\log\delta-\log(Rr_\varepsilon)}<1-\frac{1}{c_\varepsilon^2}\Big(-\frac{1}{\pi}\log\delta+A_p+O(\delta\log\delta)+o_\varepsilon(1)\Big) \tag{4–10}$$

$$-\frac{1}{c_\varepsilon^2}\Big(\frac{1}{\pi}\log R+\frac{1}{\pi}\log\frac{\pi}{2}+O\Big(\frac{\log R}{R}\Big)+o_\varepsilon(1)\Big).$$

From Lemma 4.7 and Lemma 4.12, one can see that

$$i_\varepsilon=c_\varepsilon-\frac{\log(1+\pi^2R^2)+o_\varepsilon(R)}{2\pi c_\varepsilon},\qquad s_\varepsilon=\frac{-\log\delta+\pi A_p+O(\delta)+o_\varepsilon(R)}{\pi c_\varepsilon}.$$

Adding this and (4–3) to (4–10), we have

$$\log\frac{\lambda_\varepsilon}{c_\varepsilon^2}\le -\varepsilon c_\varepsilon^2+\log(2\pi^2)+\pi A_p+o_\varepsilon(\delta)+o_\varepsilon(R)+o_\varepsilon(1)+o_\delta(1)+o_R(1).$$

Letting $\varepsilon\to 0$ first, then $\delta\to 0$ and $R\to+\infty$, we obtain

$$\limsup_{\varepsilon\to 0}\frac{\lambda_\varepsilon}{c_\varepsilon^2}\le 2\pi^2 e^{\pi A_p}.$$

Together with Lemma 4.8, this estimate yields $\sup_{u\in\mathscr{H}_1} J_\pi(u)\le \operatorname{Vol}\partial M+2\pi e^{\pi A_p}$. In fact, we have proved the following:

Proposition 4.13. *Under the assumption that $\mu_\varepsilon\ge 0$ and $u_\varepsilon(x_\varepsilon)\to+\infty$ as $\varepsilon\to 0$, we obtain*

$$\sup_{u\in\mathscr{H}_1} J_\pi(u)\le \operatorname{Vol}\partial M+2\pi e^{\pi\max_{p\in\partial M}A_p}.$$

For the other case, $\mu_\varepsilon\ge 0$ and $u_\varepsilon(x_\varepsilon)\to-\infty$, we only need to replace (4–4) by $\varphi_\varepsilon(x)=-c_\varepsilon\big(u_\varepsilon(x_\varepsilon+r_\varepsilon x)+c_\varepsilon\big)$. Using the same arguments we have used from Lemma 4.5 to Proposition 4.13, we also get:

Proposition 4.14. *Under the assumption that $\mu_\varepsilon\ge 0$ and $u_\varepsilon(x_\varepsilon)\to-\infty$ as $\varepsilon\to 0$, we obtain*

$$\sup_{u\in\mathscr{H}_1} J_\pi(u)\le \operatorname{Vol}\partial M+2\pi e^{\pi\max_{p\in\partial M}A_p}.$$

5. Existence results

Assume $A_p=\max_{p\in\partial M}A_p$ for some $p\in\partial M$. In this section, we will construct a blowing up sequence ϕ_ε with $\int_M|\nabla\phi_\varepsilon|^2=1$, and

$$\int_{\partial M} e^{\pi(\phi_\varepsilon-\bar\phi_\varepsilon)^2}>\operatorname{Vol}\partial M+2\pi e^{\pi A_p},\quad \text{where } \bar\phi_\varepsilon=\frac{1}{\operatorname{Vol}M}\int_{\partial M}\phi_\varepsilon.$$

Take an isothermal coordinate system (U,ψ) around p such that $\psi(p)=(0,0)$, ψ maps $\partial M\cap U)$ inside $\partial\mathbb{R}^2_+$, and $g=e^{2f}(ds^2+dt^2)$ with $f(0)=0$. Let R be a function of ε such that $R\to+\infty$ and $R\varepsilon\to 0$ as $\varepsilon\to 0$. For sufficiently small $r>0$, write $\mathbb{B}_r^+=\mathbb{B}_r^+(0,-\varepsilon/\pi)=\mathbb{B}_r(0,-\varepsilon/\pi)\cap\overline{\mathbb{R}^2_+}$, $B_r^+=\psi^{-1}(\mathbb{B}_r^+)$ and

$$\tilde\phi_\varepsilon(s,t)=c+\frac{-(1/2\pi)\,\log(\pi^2s^2/\varepsilon^2+(\pi t/\varepsilon+1)^2)+B}{c}\quad \text{in } \mathbb{B}^+_{R\varepsilon},$$

for some constants B, c.

Set

$$\phi_\varepsilon=\begin{cases}\tilde\phi_\varepsilon\circ\psi(x) & \text{if } x\in B^+_{R\varepsilon},\\ (G-\eta\beta)/c & \text{if } x\in B^+_{2R\varepsilon}\setminus B^+_{R\varepsilon},\\ G/c & \text{if } x\in M\setminus B^+_{2R\varepsilon},\end{cases}$$

where B, c are constants to be defined later, $\beta = G + (1/\pi)\log r - A_p = O(r)$, $\eta \in C_0^\infty(B_{2R\varepsilon})$ with $\eta \equiv 1$ on $B_{R\varepsilon}$, and $\max |\nabla\eta| = O(1/(R\varepsilon))$.

To ensure that $\phi_\varepsilon \in H^{1,2}(M)$, we assume

$$c + \frac{-(1/2\pi)\log(\pi^2 R^2) + B}{c} = \frac{-(1/\pi)\log(R\varepsilon) + A_p}{c},$$

which gives

$$2\pi c^2 = 2\log\pi - 2\pi B - 2\log\varepsilon + 2\pi A_p. \tag{5–1}$$

By (4–9), we have

$$\int_{B^+_{R\varepsilon}} |\nabla\phi_\varepsilon|^2 = \frac{1}{\pi c^2}\log\frac{\pi}{2} + \frac{1}{\pi c^2}\log R + O\Big(\frac{\log R}{R}\Big),$$

$$\int_{M\setminus B^+_{R\varepsilon}} |\nabla\phi_\varepsilon|^2 = \int_{M\setminus B^+_{R\varepsilon}} \frac{|\nabla G|^2}{c^2} + \int_{B^+_{2R\varepsilon}\setminus B^+_{R\varepsilon}} \frac{|\nabla(\eta\beta)|^2}{c^2} - \frac{2}{c^2}\int_{B^+_{2R\varepsilon}\setminus B^+_{R\varepsilon}} \nabla G\nabla(\eta\beta).$$

Let I_1, I_2, I_3 be the three summands on the right-hand side of the last equation. Clearly, $I_2 = c^{-2}O(R\varepsilon)$ and $I_3 = c^{-2}O(R\varepsilon)$. Next,

$$\begin{aligned} I_1 &= \frac{1}{c^2}\int_{\partial(M\setminus B^+_{R\varepsilon})} G\frac{\partial G}{\partial n} - \frac{1}{c^2}\int_{M\setminus B^+_{R\varepsilon}} G\Delta G \\ &= \frac{1}{c^2}\int_{\partial M\setminus\partial B^+_{R\varepsilon}} G\frac{\partial G}{\partial n} - \frac{1}{c^2}\int_{\partial B^+_{R\varepsilon}\setminus\partial M} G\frac{\partial G}{\partial n} + \frac{1}{c^2}\frac{1}{\mathrm{Vol}\, M}\Big(\int_{B^+_{R\varepsilon}} G\Big) + \frac{1}{c^2}O(\frac{1}{R}) \\ &= \frac{1}{c^2}\Big(-\frac{1}{\pi}\log(R\varepsilon) + A_p + O(R\varepsilon\log(R_\varepsilon)) + O\Big(\frac{1}{R}\Big)\Big), \end{aligned}$$

whence

$$\int_{M\setminus B^+_{R\varepsilon}} |\nabla\phi_\varepsilon|^2 = \frac{1}{c^2}\Big(-\frac{1}{\pi}\log(R\varepsilon) + A_p + O(R\varepsilon\log(R_\varepsilon)) + O\Big(\frac{1}{R}\Big)\Big).$$

Combining the two estimates above, one has

$$\int_M |\nabla\phi_\varepsilon|^2 = \frac{1}{c^2}\Big(-\frac{1}{\pi}\log\varepsilon + \frac{1}{\pi}\log\frac{\pi}{2} + A_p + O(R\varepsilon\log(R\varepsilon)) + O\Big(\frac{\log R}{R}\Big)\Big).$$

To ensure that $\int_M |\nabla\phi_\varepsilon|^2 = 1$, we set

$$c^2 = -\frac{1}{\pi}\log\varepsilon + \frac{1}{\pi}\log\frac{\pi}{2} + A_p + O(R\varepsilon\log(R\varepsilon)) + O\Big(\frac{\log R}{R}\Big).$$

By (5–1), one can determine B as

$$B = \frac{1}{\pi}\log 2 + O(R\varepsilon\log(R\varepsilon)) + O\Big(\frac{\log R}{R}\Big).$$

A straightforward computation gives

$$\bar{\phi}_\varepsilon = \frac{1}{\operatorname{Vol} M}\int_M \phi_\varepsilon = \frac{1}{c}\Big(O\big((R\varepsilon)^2\log R\big) + O\big((R\varepsilon)^2\log\varepsilon\big) + O\big((R\varepsilon)^2\log(R\varepsilon)\big)\Big).$$

Then

$$\begin{aligned}
&\int_{\partial B^+_{R\varepsilon}\cap\partial M} \exp\big(\pi(\phi_\varepsilon - \bar{\phi}_\varepsilon)^2\big)\\
&= \int_{\partial\mathbb{B}^+_{R\varepsilon}\cap\partial\mathbb{R}^2_+} \exp\left(\pi\left(c - \frac{\log\left(\pi^2 s^2/\varepsilon^2 + (1+\pi t/\varepsilon)^2\right) + c\bar{\phi}_\varepsilon - 2\pi B}{2\pi c}\right)^2 + O(R\varepsilon)\right)ds\\
&\geq \int_{\partial\mathbb{B}^+_{R\varepsilon}\cap\partial\mathbb{R}^2_+} \exp\left(\pi c^2 - \log\left(\pi^2\frac{s^2}{\varepsilon^2}+1\right) + 2\pi B - c\bar{\phi}_\varepsilon\right)e^{O(R\varepsilon)}ds\\
&= 2\pi e^{\pi A_p}\Big(\frac{2}{\pi}\arctan(\pi R)\Big)\exp\left(O(R\varepsilon\log R\varepsilon) + O\Big(\frac{\log R}{R}\Big)\right)\\
&= 2\pi e^{\pi A_p}\left(1 + O(R\varepsilon\log(R\varepsilon)) + O\Big(\frac{\log R}{R}\Big)\right).
\end{aligned}$$

Moreover,

$$\begin{aligned}
\int_{\partial M\setminus\partial B^+_{R\varepsilon}} \exp\big(\pi(\phi_\varepsilon - \bar{\phi}_\varepsilon)^2\big) &\geq \int_{\partial M\setminus\partial B^+_{R\varepsilon}} \big(1 + \pi^2(\phi_\varepsilon - \bar{\phi}_\varepsilon)^2\big)\\
&\geq \operatorname{Vol}\partial M - \operatorname{Vol}(\partial M\cap\partial B^+_{R\varepsilon}) + \pi^2\int_{\partial M\setminus B^+_{2R\varepsilon}} \frac{(G - c\bar{\phi}_\varepsilon)^2}{c^2}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\int_{\partial M} e^{\pi(\phi_\varepsilon - \bar{\phi}_\varepsilon)^2} \geq \operatorname{Vol}\partial M + \frac{\pi^2}{c^2}\int_{\partial M\setminus B^+_{2R\varepsilon}} &\big(G - O(R\varepsilon\log R\varepsilon)\big)^2\\
&+ O\big(R\varepsilon\log(R\varepsilon)\big) + O\Big(\frac{\log R}{R}\Big).
\end{aligned}$$

Set $R = \log^2\varepsilon$. Then $R\to+\infty$, $R\varepsilon\to 0$, $c^2(\log R)/R\to 0$, $c^2R\varepsilon\log R\varepsilon\to 0$. Hence

$$\int_{\partial M} e^{\pi(\phi_\varepsilon - \bar{\phi}_\varepsilon)^2} > \operatorname{Vol}\partial M + 2\pi e^{\pi A_p}$$

when ε is sufficiently small. This contradicts Proposition 4.13 or Proposition 4.14. Hence Theorem 1.1 holds. □

6. Proof of Theorem 1.2

Lemma 6.1. *Set* $\tilde{\alpha}_2 = \sup\left\{\alpha;\ \sup_{u\in\mathcal{H}_2}\int_M e^{\alpha u^2} < +\infty\right\}$. *Then* $\tilde{\alpha}_2 = 2\pi$.

Proof. We first show that $\tilde{\alpha}_2 \geq 2\pi$. For any $\alpha < 2\pi$ and $u \in \mathcal{H}_2$, we set

$$\tilde{u} = u - \frac{1}{\operatorname{Vol} M} \int_M u.$$

Then $\tilde{u} \in \mathcal{H}_1$. So, by Lemma 3.2,

$$\int_M e^{\alpha \tilde{u}^2} \leq \sup_{v \in \mathcal{H}_1} \int_M e^{\alpha v^2} < +\infty.$$

For any $u \in \mathcal{H}_2$, we have

$$\int_M e^{\alpha u^2} \leq e^{c(\varepsilon')} \int_M e^{\alpha(1+\varepsilon')\tilde{u}^2}$$

for some $\varepsilon' > 0$. One can choose ε' such that $\alpha(1+\varepsilon') < 2\pi$, which gives

$$\int_M e^{\alpha u^2} \leq \sup_{v \in \mathcal{H}_1} \int_M e^{\alpha(1+\varepsilon')v^2} < +\infty.$$

Hence

$$\sup_{u \in \mathcal{H}_2} \int_M e^{\alpha u^2} < +\infty.$$

Next we prove that $\tilde{\alpha}_2$ cannot be greater than 2π. To do this, the example in the proof of Lemma 3.2 still works here. For $p \in \partial M$, we set

$$u_\varepsilon = \begin{cases} -\sqrt{\dfrac{1}{2\pi} \log \dfrac{1}{\varepsilon}} & \text{in } B_{\delta\sqrt{\varepsilon}}(p), \\ \dfrac{\sqrt{2}}{\sqrt{-\pi \log \varepsilon}} \log \dfrac{r}{\delta} & \text{in } B_\delta(p) \setminus B_{\delta\sqrt{\varepsilon}}(p) \\ C_\varepsilon \varphi & \text{in } M \setminus B_\delta(p), \end{cases}$$

where $\varphi \in C_0^\infty(M \setminus B_\delta(p))$, $0 \leq \varphi \leq 1$, and C_ε is chosen to satisfy $\int_M u_\varepsilon = 0$. It is easy to check that

$$\|u_\varepsilon\|^2_{H^{1,2}(M)} = \int_M \big(|\nabla u_\varepsilon|^2 + u_\varepsilon^2\big) \to 1 \quad \text{as} \quad \varepsilon \to 0,$$

and for any $\alpha > 2\pi$

$$\begin{aligned} \int_M \exp\left(\alpha\left(\frac{u_\varepsilon}{\|u_\varepsilon\|_{H^{1,2}(M)}}\right)^2\right) &\geq \exp\left(\frac{\alpha}{2\pi\|u_\varepsilon\|^2_{H^{1,2}(M)}} \log \frac{1}{\varepsilon}\right) \operatorname{Vol} B_{\delta\sqrt{\varepsilon}} \\ &\geq C\varepsilon^{1-\frac{\alpha}{2\pi\|u_\varepsilon\|^2_{H^{1,2}(M)}}}, \end{aligned}$$

which approaches $+\infty$ as $\varepsilon \to 0$. Therefore $\tilde{\alpha}_2 \leq 2\pi$, as needed. □

Using the same argument as in the proof of Lemma 3.3, we obtain:

Lemma 6.2.

$$\sup_{u\in\mathcal{H}_2} J_\alpha(u) < +\infty \quad \textit{for } \alpha < \pi \quad \textit{and} \quad \sup_{u\in\mathcal{H}_2} J_\alpha(u) = +\infty \quad \textit{for } \alpha > \pi.$$

Similarly to Lemma 4.1, one has:

Lemma 6.3. *The functional $J_{\pi-\varepsilon}(u)$ defined in the space $\mathcal{H}_2$ admits a smooth maximizer $u_\varepsilon \in \mathcal{H}_2$.*

Proof. The proof of the existence of u_ε is the same as that of Lemma 4.1. The Euler–Lagrange equation of u_ε is

$$\begin{cases} \Delta u_\varepsilon = u_\varepsilon & \text{in } \mathring{M}, \\ \dfrac{\partial u_\varepsilon}{\partial n} = \dfrac{\pi-\varepsilon}{\lambda_\varepsilon} u_\varepsilon e^{(\pi-\varepsilon)u_\varepsilon^2} & \text{on } \partial M, \end{cases} \tag{6–1}$$

where

$$\lambda_\varepsilon = (\pi-\varepsilon)\int_M u_\varepsilon^2 e^{(\pi-\varepsilon)u_\varepsilon^2}.$$

Using Lemmas 2.1 and 2.2 repeatedly, we get $u_\varepsilon \in C^\infty(M)$. □

The rest of the proof of Theorem 1.2 is almost the same as that of Theorem 1.1; we only give its outline. Without loss of generality, we may assume $u_\varepsilon \geq 0$ in M. Set $c_\varepsilon = u_\varepsilon(x_\varepsilon) = \max_{x\in M} u_\varepsilon(x)$. If $\{c_\varepsilon\}$ is bounded, it is not difficult to see that Theorem 1.2 holds. Hence we assume that $c_\varepsilon \to +\infty$ as $\varepsilon \to 0$. This is equivalent to saying $\int_M e^{\alpha u_\varepsilon^2} \to +\infty$ for any $\alpha > 2\pi$, which implies that $u_\varepsilon \to 0$ strongly in $L^2(M)$ (see the first step in the proof of Lemma 3.2). Applying the maximum principle to (6–1), we find that $x_\varepsilon \in \partial M$. Assume that x_ε converges to p, so $p \in \partial M$. Let r_ε, $\varphi_\varepsilon(x)$ and $\psi_\varepsilon(x)$ be as in Section 4. Then

$$\varphi_\varepsilon \to \varphi = -\frac{1}{2\pi}\log\bigl(\pi^2 x_1^2 + (1+\pi x_2)^2\bigr) \quad \text{in } C^2_{\text{loc}}(\mathbb{R}^2_+).$$

Moreover $c_\varepsilon u_\varepsilon \rightharpoonup G$ weakly in $H^{1,q}(M)$ for any q such that $1 < q < 2$. The function $G \in C^\infty(M \setminus \{p\})$ satisfies

$$\begin{cases} -\Delta G + G = \delta_p & \text{in } M, \\ \int_M G = 1. \end{cases}$$

In a normal coordinate system around p, the Green's function G has the representation $G = -(1/\pi)\log r + A_p + O(r)$, where $r(x) = \operatorname{dist}(p,x)$ is the distance function, A_p is a constant depending only on p, and $A_p + O(r)$ is called the regular part. Repeating the other steps taken in Section 4, we obtain

$$\sup_{u\in\mathcal{H}_2}(u) \leq \operatorname{Vol}\partial M + 2\pi e^{\pi A_p}. \tag{6–2}$$

The blowing up sequence we constructed in Section 5 still works here; one can check that

$$J_\pi\left(\frac{\phi_\varepsilon}{\|\phi_\varepsilon\|_{H^{1,2}(M)}}\right) > \mathrm{Vol}\,\partial M + 2\pi e^{\pi A_p}$$

for sufficiently small ε, which contradicts (6–2) and so proves Theorem 1.2. □

Acknowledgements. The author thanks Professor Jiayu Li for his constant support, Yuxiang Li for stimulating conversations and the referee for his helpful suggestions.

References

[Adams 1988] D. R. Adams, "A sharp inequality of J. Moser for higher order derivatives", *Ann. of Math.* (2) **128**:2 (1988), 385–398. MR 89i:46034 Zbl 0672.31008

[Carleson and Chang 1986] L. Carleson and S.-Y. A. Chang, "On the existence of an extremal function for an inequality of J. Moser", *Bull. Sci. Math.* (2) **110**:2 (1986), 113–127. MR 88f:46070 Zbl 0619.58013

[Chang 1996] S.-Y. A. Chang, "The Moser–Trudinger inequality and applications to some problems in conformal geometry", pp. 65–125 in *Nonlinear partial differential equations in differential geometry* (Park City, UT, 1992), edited by R. Hardt and M. Wolf, IAS/Park City Math. Ser. **2**, Amer. Math. Soc., Providence, RI, 1996. MR 96j:58173 Zbl 0924.58106

[Chang and Yang 1988] S.-Y. A. Chang and P. C. Yang, "Conformal deformation of metrics on S^2", *J. Differential Geom.* **27**:2 (1988), 259–296. MR 89b:53079

[Cohn and Lu 2002] W. S. Cohn and G. Z. Lu, "Best constants for Moser-Trudinger inequalities, fundamental solutions and one-parameter representation formulas on groups of Heisenberg type", *Acta Math. Sin.* (*Engl. Ser.*) **18**:2 (2002), 375–390. MR 2003e:46046 Zbl 1011.22004

[Ding et al. 1997] W. Ding, J. Jost, J. Li, and G. Wang, "The differential equation $\Delta u = 8\pi - 8\pi h e^u$ on a compact Riemann surface", *Asian J. Math.* **1**:2 (1997), 230–248. MR 98k:58062 Zbl 0955.58010

[Escobar and Schoen 1986] J. F. Escobar and R. M. Schoen, "Conformal metrics with prescribed scalar curvature", *Invent. Math.* **86**:2 (1986), 243–254. MR 88b:58030 Zbl 1056.53026

[Flucher 1992] M. Flucher, "Extremal functions for the Trudinger–Moser inequality in 2 dimensions", *Comment. Math. Helv.* **67**:3 (1992), 471–497. MR 93k:58073 Zbl 0763.58008

[Fontana 1993] L. Fontana, "Sharp borderline Sobolev inequalities on compact Riemannian manifolds", *Comment. Math. Helv.* **68**:3 (1993), 415–454. MR 94h:46048 Zbl 0844.58082

[Jost and Wang 2001] J. Jost and G. Wang, "Analytic aspects of the Toda system, I: a Moser–Trudinger inequality", *Comm. Pure Appl. Math.* **54**:11 (2001), 1289–1319. MR 2002m:35050 Zbl 01860529

[Li 2001] Y. Li, "Moser-Trudinger inequality on compact Riemannian manifolds of dimension two", *J. Partial Differential Equations* **14**:2 (2001), 163–192. MR 2002h:58033 Zbl 0995.58021

[Li 2005] Y. Li, "Extremal functions for the Moser–Trudinger inequalities on compact Riemannian manifolds", *Sci. China Ser. A* **48**:5 (2005), 618–648. MR 2158479 Zbl 02234372

[Li and Liu 2005] Y. Li and P. Liu, "A Moser–Trudinger inequality on the boundary of a compact Riemann surface", *Math. Z.* **250**:2 (2005), 363–386. MR MR2178789 Zbl 1070.26017

[Li and Zhu 1995] Y. Li and M. Zhu, "Uniqueness theorems through the method of moving spheres", *Duke Math. J.* **80**:2 (1995), 383–417. MR 96k:35061 Zbl 0846.35050

[Li and Zhu 1997] Y. Li and M. Zhu, "Sharp Sobolev trace inequalities on Riemannian manifolds with boundaries", *Comm. Pure Appl. Math.* **50**:5 (1997), 449–487. MR 98c:46065 Zbl 0869.58054

[Lin 1996] K.-C. Lin, "Extremal functions for Moser's inequality", *Trans. Amer. Math. Soc.* **348**:7 (1996), 2663–2671. MR 96i:58043 Zbl 0861.49001

[Liu 2002] P. Liu, *A Moser–Trudinger type inequality and blow-up analysis on Riemannian surfaces*, Ph.D. thesis, Univ. Leipzig, 2002.

[Moser 1970/71] J. Moser, "A sharp form of an inequality by N. Trudinger", *Indiana Univ. Math. J.* **20** (1970/71), 1077–1092. MR 46 #662 Zbl 0213.13001

[Nolasco and Tarantello 1998] M. Nolasco and G. Tarantello, "On a sharp Sobolev-type inequality on two-dimensional compact manifolds", *Arch. Ration. Mech. Anal.* **145**:2 (1998), 161–195. MR 2000j:58021 Zbl 0980.46022

[Osgood et al. 1988] B. Osgood, R. Phillips, and P. Sarnak, "Extremals of determinants of Laplacians", *J. Funct. Anal.* **80**:1 (1988), 148–211. MR 90d:58159 Zbl 0653.53022

[Schoen 1984] R. Schoen, "Conformal deformation of a Riemannian metric to constant scalar curvature", *J. Differential Geom.* **20**:2 (1984), 479–495. MR 86i:58137 Zbl 0576.53028

[Struwe 1988] M. Struwe, "Critical points of embeddings of $H_0^{1,n}$ into Orlicz spaces", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **5**:5 (1988), 425–464. MR 90c:35084 Zbl 0664.35022

[Troianiello 1987] G. M. Troianiello, *Elliptic differential equations and obstacle problems*, Plenum Press, New York, 1987. MR 92b:35004 Zbl 0655.35002

[Trudinger 1967] N. S. Trudinger, "On imbeddings into Orlicz spaces and some applications", *J. Math. Mech.* **17** (1967), 473–483. MR 35 #7121 Zbl 0163.36402

Received November 26, 2004. Revised January 27, 2005.

YUNYAN YANG
DEPARTMENT OF MATHEMATICS
INFORMATION SCHOOL
RENMIN UNIVERSITY OF CHINA
BEIJING 100872
P. R. CHINA

yunyan_yang2002@yahoo.com.cn

Guidelines for Authors

Authors may submit manuscripts at pjm.math.berkeley.edu/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 Web site 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 227 No. 1 September 2006