Pacific Journal of Mathematics

Volume 264 No. 1 July 2013

PACIFIC JOURNAL OF MATHEMATICS

msp.org/pjm

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

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

The subscription price for 2013 is US $400/year for the electronic version, and $485/year for print and electronic.
Subscriptions, requests for back issues and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published 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.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.1

ON THE CENTER OF FUSION CATEGORIES

ALAIN BRUGUIÈRES AND ALEXIS VIRELIZIER

Müger proved in 2003 that the center of a spherical fusion category $\mathscr{C}$ of nonzero dimension over an algebraically closed field is a modular fusion category whose dimension is the square of that of $\mathscr{C}$. We generalize this theorem to a pivotal fusion category $\mathscr{C}$ over an arbitrary commutative ring $\Bbbk$, without any condition on the dimension of the category. (In this generalized setting, modularity is understood as 2-modularity in the sense of Lyubashenko.) Our proof is based on an explicit description of the Hopf algebra structure of the coend of the center of $\mathscr{C}$. Moreover we show that the dimension of $\mathscr{C}$ is invertible in $\Bbbk$ if and only if any object of the center of $\mathscr{C}$ is a retract of a "free" half-braiding. As a consequence, if $\Bbbk$ is a field, then the center of $\mathscr{C}$ is semisimple (as an abelian category) if and only if the dimension of $\mathscr{C}$ is nonzero. If in addition $\Bbbk$ is algebraically closed, then this condition implies that the center is a fusion category, so that we recover Müger's result.

Introduction

Given a monoidal category $\mathscr{C}$, Joyal and Street [1991], Drinfeld (unpublished), and Majid [1991] defined a braided category $\mathscr{Z}(\mathscr{C})$, called the center of $\mathscr{C}$, whose objects are half-braidings of $\mathscr{C}$. Müger [2003] showed that the center $\mathscr{Z}(\mathscr{C})$ of a spherical fusion category $\mathscr{C}$ of nonzero dimension over an algebraically closed field $\Bbbk$ is a modular fusion category, and that the dimension of $\mathscr{Z}(\mathscr{C})$ is the square of that of $\mathscr{C}$. Müger's proof of this remarkable result relies on algebraic constructions due to Ocneanu (such as the "tube" algebra) and involves the construction of a

MSC2010: 18D10, 16T05, 18C20.
Keywords: categorical center, fusion categories, Hopf monads, modularity.

weak monoidal Morita equivalence between $\mathscr{Z}(\mathscr{C})$ and $\mathscr{C} \otimes \mathscr{C}^{\mathrm{op}}$. The modularity of the center is of special interest in three-dimensional quantum topology, since spherical fusion categories and modular categories are respectively the algebraic input for the construction of the Turaev–Viro/Barrett–Westbury invariant and of the Reshetikhin–Turaev invariant. Indeed it has been shown recently in [Turaev and Virelizier 2010] (see also [Balsam 2010]) that, under the hypotheses of Müger's theorem, the Barrett–Westbury generalization of the Turaev–Viro invariant for $\mathscr{C}$ is equal to the Reshetikhin–Turaev invariant for $\mathscr{Z}(\mathscr{C})$.

In this paper, we generalize Müger's theorem to pivotal fusion categories over an arbitrary commutative ring. More precisely, given a pivotal fusion category $\mathscr{C}$ over a commutative ring $\Bbbk$, we prove the following:

(i) the center $\mathscr{Z}(\mathscr{C})$ of $\mathscr{C}$ is always modular (but not necessarily semisimple) and has dimension $\dim(\mathscr{C})^2$;

(ii) the scalar $\dim(\mathscr{C})$ is invertible in $\Bbbk$ if and only if every half braiding is a retract of a so-called *free half braiding*;

(iii) if $\Bbbk$ is a field, then $\mathscr{Z}(\mathscr{C})$ is abelian semisimple if and only if $\dim(\mathscr{C}) \neq 0$;

(iv) if $\Bbbk$ is an algebraically closed field, then $\mathscr{Z}(\mathscr{C})$ is fusion if and only if $\dim(\mathscr{C}) \neq 0$.

Our proof is different from that of Müger. It relies on the principle that if a braided category $\mathscr{B}$ has a coend, then all the relevant information about $\mathscr{B}$ is encoded in its coend, which is a universal Hopf algebra sitting in $\mathscr{B}$ and endowed with a canonical Hopf algebra pairing. For instance, modularity means that the canonical pairing is nondegenerate, and the dimension of $\mathscr{B}$ is that of its coend. In particular we do not need to introduce an auxiliary category.

The center $\mathscr{Z}(\mathscr{C})$ of a pivotal fusion category $\mathscr{C}$ always has a coend. We provide a complete and explicit description of the Hopf algebra structure of this coend, which enables us to exhibit an integral for the coend and an "inverse" to the pairing. Our proofs are based on a "handleslide" property for pivotal fusion categories.

A general description of the coend of the center of a rigid category $\mathscr{C}$, together with its structural morphisms, was given in [Bruguières and Virelizier 2012]. It is an application of the theory of Hopf monads, and in particular, of the notion of double of a Hopf monad, which generalizes the Drinfeld double of a Hopf algebra. It is based on the fact that $\mathscr{Z}(\mathscr{C})$ is the category of modules over a certain quasitriangular Hopf monad Z on $\mathscr{C}$ (generalizing the braided equivalence $\mathscr{Z}(\mathrm{mod}_H) \simeq \mathrm{mod}_{D(H)}$ between the center of the category of modules over a finite-dimensional Hopf algebra H and the category of modules over the Drinfeld double $D(H)$ of H). It turns out that, when $\mathscr{C}$ is a fusion category, we can make this description very explicit and in particular, we can depict the structural morphisms of the coend by means of a graphical formalism for fusion categories.

Part of the results of this paper were announced (without proofs) in [Bruguières and Virelizier 2008], where they were used to define and compute a 3-manifolds invariant of Reshetikhin–Turaev type associated with the center of $\mathscr{C}$, even when the dimension of $\mathscr{C}$ is not invertible.

Organization of the text. In Section 1, we recall definitions, notations and basic results concerning pivotal and fusion categories over a commutative ring. A graphical formalism for representing morphisms in fusion categories is provided. In Section 2, we state the main results of this paper, that is, the description of the coend of the center of a pivotal fusion category and its structural morphisms, the modularity of the center of such a category, its dimension, and a semisimplicity criterion. Section 3 is devoted to coends, Hopf algebras in braided categories, and modular categories. Section 4 contains the proofs of the main results.

1. Pivotal and fusion categories

Monoidal categories are assumed to be strict. This does not lead to any loss of generality, since, in view of Mac Lane's coherence theorem for monoidal categories (see [Mac Lane 1998]), all definitions and statements remain valid for nonstrict monoidal categories after insertion of the suitable canonical isomorphisms.

1A. ***Rigid categories.*** Let $\mathscr{C} = (\mathscr{C}, \otimes, \mathbb{1})$ be a monoidal category. A *left dual* of an object X of $\mathscr{C}$ is an object ${}^\vee X$ of $\mathscr{C}$ together with morphisms $\mathrm{ev}_X : {}^\vee X \otimes X \to \mathbb{1}$ and $\mathrm{coev}_X : \mathbb{1} \to X \otimes {}^\vee X$ such that

$$(\mathrm{id}_X \otimes \mathrm{ev}_X)(\mathrm{coev}_X \otimes \mathrm{id}_X) = \mathrm{id}_X \quad \text{and} \quad (\mathrm{ev}_X \otimes \mathrm{id}_{{}^\vee X})(\mathrm{id}_{{}^\vee X} \otimes \mathrm{coev}_X) = \mathrm{id}_{{}^\vee X}.$$

Similarly a *right dual* of X is an object $X^\vee$ with morphisms $\widetilde{\mathrm{ev}}_X : X \otimes X^\vee \to \mathbb{1}$ and $\widetilde{\mathrm{coev}}_X : \mathbb{1} \to X^\vee \otimes X$ such that

$$(\widetilde{\mathrm{ev}}_X \otimes \mathrm{id}_X)(\mathrm{id}_X \otimes \widetilde{\mathrm{coev}}_X) = \mathrm{id}_X \quad \text{and} \quad (\mathrm{id}_{X^\vee} \otimes \widetilde{\mathrm{ev}}_X)(\widetilde{\mathrm{coev}}_X \otimes \mathrm{id}_{X^\vee}) = \mathrm{id}_{X^\vee}.$$

The left and right duals of an object, if they exist, are unique up to an isomorphism (preserving the (co)evaluation morphisms).

A monoidal category $\mathscr{C}$ is *rigid* (or *autonomous*) if every object of $\mathscr{C}$ admits a left and a right dual. The choice of left and right duals for each object of a rigid $\mathscr{C}$ defines a left dual functor ${}^\vee ? : \mathscr{C}^{\mathrm{op}} \to \mathscr{C}$ and a right dual functor $?^\vee : \mathscr{C}^{\mathrm{op}} \to \mathscr{C}$, where $\mathscr{C}^{\mathrm{op}}$ is the opposite category to $\mathscr{C}$ with opposite monoidal structure. The left and right dual functors are strong monoidal. Note that the actual choice of left and right duals is innocuous in the sense that different choices of left (respectively, right) duals define canonically monoidally isomorphic left (respectively, right) dual functors.

There are canonical natural monoidal isomorphisms ${}^\vee(X^\vee) \simeq X \simeq ({}^\vee X)^\vee$, but in general the left and right dual functors are not monoidally isomorphic.

1B. ***Pivotal categories.*** A rigid category $\mathscr{C}$ is *pivotal* (or *sovereign*) if it is endowed with a monoidal isomorphism between the left and the right dual functors. We may assume that this isomorphism is the identity without loss of generality. In other words, for each object X of $\mathscr{C}$, we have a *dual object* X^* and four morphisms

$$\begin{aligned} &\mathrm{ev}_X : X^* \otimes X \to \mathbb{1}, \quad &&\mathrm{coev}_X : \mathbb{1} \to X \otimes X^*, \\ &\widetilde{\mathrm{ev}}_X : X \otimes X^* \to \mathbb{1}, \quad &&\widetilde{\mathrm{coev}}_X : \mathbb{1} \to X^* \otimes X, \end{aligned}$$

such that $(X^*, \mathrm{ev}_X, \mathrm{coev}_X)$ is a left dual for X, $(X^*, \widetilde{\mathrm{ev}}_X, \widetilde{\mathrm{coev}}_X)$ is a right dual for X, and the induced left and right dual functors coincide as monoidal functors. In particular, the *dual* $f^* : Y^* \to X^*$ of any morphism $f : X \to Y$ in $\mathscr{C}$ is

$$\begin{aligned} f^* &= (\mathrm{ev}_Y \otimes \mathrm{id}_{X^*})(\mathrm{id}_{Y^*} \otimes f \otimes \mathrm{id}_{X^*})(\mathrm{id}_{Y^*} \otimes \mathrm{coev}_X) \\ &= (\mathrm{id}_{X^*} \otimes \widetilde{\mathrm{ev}}_Y)(\mathrm{id}_{X^*} \otimes f \otimes \mathrm{id}_{Y^*})(\widetilde{\mathrm{coev}}_X \otimes \mathrm{id}_{Y^*}). \end{aligned}$$

In what follows, for a pivotal category $\mathscr{C}$, we will suppress the duality constraints $\mathbb{1}^* \cong \mathbb{1}$ and $X^* \otimes Y^* \cong (Y \otimes X)^*$. For example, we will write $(f \otimes g)^* = g^* \otimes f^*$ for morphisms f, g in $\mathscr{C}$.

1C. ***Traces and dimensions.*** For an endomorphism f of an object X of a pivotal category $\mathscr{C}$, one defines the *left* and *right traces* $\mathrm{tr}_l(f), \mathrm{tr}_r(f) \in \mathrm{End}_{\mathscr{C}}(\mathbb{1})$ by

$$\mathrm{tr}_l(f) = \mathrm{ev}_X(\mathrm{id}_{X^*} \otimes f)\widetilde{\mathrm{coev}}_X \quad \text{and} \quad \mathrm{tr}_r(f) = \widetilde{\mathrm{ev}}_X(f \otimes \mathrm{id}_{X^*})\mathrm{coev}_X.$$

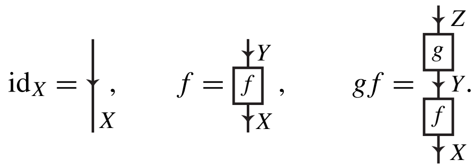
They satisfy $\mathrm{tr}_l(gh) = \mathrm{tr}_l(hg)$ and $\mathrm{tr}_r(gh) = \mathrm{tr}_r(hg)$ for any morphisms $g : X \to Y$ and $h : Y \to X$ in $\mathscr{C}$. Also we have $\mathrm{tr}_l(f) = \mathrm{tr}_r(f^*) = \mathrm{tr}_l(f^{**})$ for any endomorphism f in $\mathscr{C}$. If

$$\alpha \otimes \mathrm{id}_X = \mathrm{id}_X \otimes \alpha \quad \text{for all } \alpha \in \mathrm{End}_{\mathscr{C}}(\mathbb{1}) \text{ and } X \text{ in } \mathscr{C}, \tag{1}$$

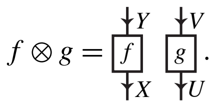
then tr_l, tr_r are $\otimes$-multiplicative; that is, $\mathrm{tr}_l(f \otimes g) = \mathrm{tr}_l(f)\,\mathrm{tr}_l(g)$ and $\mathrm{tr}_r(f \otimes g) = \mathrm{tr}_r(f)\,\mathrm{tr}_r(g)$ for all endomorphisms f, g in $\mathscr{C}$.

The *left* and the *right dimensions* of an object X of $\mathscr{C}$ are defined by $\dim_l(X) = \mathrm{tr}_l(\mathrm{id}_X)$ and $\dim_r(X) = \mathrm{tr}_r(\mathrm{id}_X)$. Isomorphic objects have the same dimensions, $\dim_l(X) = \dim_r(X^*) = \dim_l(X^{**})$, and $\dim_l(\mathbb{1}) = \dim_r(\mathbb{1}) = \mathrm{id}_{\mathbb{1}}$. If $\mathscr{C}$ satisfies (1), then left and right dimensions are $\otimes$-multiplicative: $\dim_l(X \otimes Y) = \dim_l(X)\dim_l(Y)$ and $\dim_r(X \otimes Y) = \dim_r(X)\dim_r(Y)$ for any X, Y in $\mathscr{C}$.

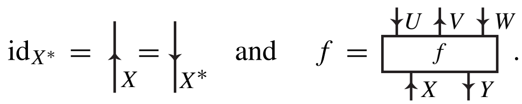
1D. ***Penrose graphical calculus.*** We represent morphisms in a category $\mathscr{C}$ by plane diagrams to be read from the bottom to the top. In a pivotal category $\mathscr{C}$, the diagrams are made of oriented arcs colored by objects of $\mathscr{C}$ and of boxes colored by morphisms of $\mathscr{C}$. The arcs connect the boxes and have no mutual intersections or self-intersections. The identity id_X of an object X of $\mathscr{C}$, a morphism $f : X \to Y$, and the composition of two morphisms $f : X \to Y$ and $g : Y \to Z$ are represented respectively as



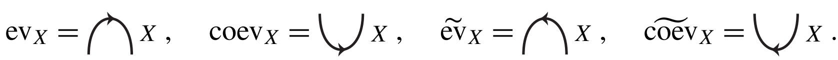
The monoidal product of two morphisms $f : X \to Y$ and $g : U \to V$ is represented by juxtaposition:



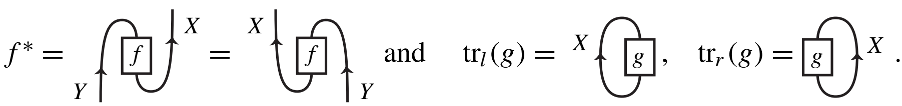
If an arc colored by X is oriented upwards, then the corresponding object in the source/target of morphisms is X^*. For example, id_{X^*} and a morphism $f : X^* \otimes Y \to U \otimes V^* \otimes W$ may be depicted as



The duality morphisms are depicted as follows:



The dual of a morphism $f : X \to Y$ and the traces of a morphism $g : X \to X$ can be depicted as follows:



In a pivotal category, the morphisms represented by the diagrams are invariant under isotopies of the diagrams in the plane keeping fixed the bottom and top endpoints.

1E. *Spherical categories.* A *spherical category* is a pivotal category whose left and right traces are equal; i.e., $\mathrm{tr}_l(g) = \mathrm{tr}_r(g)$ for every endomorphism g of an object. Then $\mathrm{tr}_l(g)$ and $\mathrm{tr}_r(g)$ are denoted $\mathrm{tr}(g)$ and called the *trace of* g. Similarly, the left and right dimensions of an object X are denoted $\dim(X)$ and called the *dimension of* X.

Note that sphericity can be interpreted in graphical terms: it means that the morphisms represented by closed diagrams are invariant under isotopies of diagrams in the 2-sphere $S^2 = \mathbb{R}^2 \cup \{\infty\}$, i.e., are preserved under isotopies pushing arcs of the diagrams across ∞.

1F. *Additive categories.* Let $\Bbbk$ be a commutative ring. A $\Bbbk$-*additive category* is a category where Hom-sets are $\Bbbk$-modules, the composition of morphisms is $\Bbbk$-bilinear, and any finite family of objects has a direct sum. In particular, such a category has a zero object.

An object X of a $\Bbbk$-additive category $\mathscr{C}$ is *scalar* if the map $\Bbbk \to \mathrm{End}_{\mathscr{C}}(X)$, $\alpha \mapsto \alpha\, \mathrm{id}_X$ is bijective.

A $\Bbbk$-*additive monoidal category* is a monoidal category which is $\Bbbk$-additive in such a way that the monoidal product is $\Bbbk$-bilinear. Note that a $\Bbbk$-additive monoidal category whose unit object $\mathbb{1}$ is scalar satisfies (1) and so its traces tr_l, tr_r are $\Bbbk$-linear and $\otimes$-multiplicative.

1G. *Fusion categories.* A *fusion category* over a commutative ring $\Bbbk$ is a $\Bbbk$-additive rigid category $\mathscr{C}$ such that

(a) each object of $\mathscr{C}$ is a finite direct sum of scalar objects;

(b) for any nonisomorphic scalar objects i, j of $\mathscr{C}$, we have $\mathrm{Hom}_{\mathscr{C}}(i, j) = 0$;

(c) the set of isomorphism classes of scalar objects of $\mathscr{C}$ is finite;

(d) the unit object $\mathbb{1}$ is scalar.

Let $\mathscr{C}$ be a fusion category. The Hom spaces in $\mathscr{C}$ are free $\Bbbk$-modules of finite rank. We identify $\mathrm{End}_{\mathscr{C}}(\mathbb{1})$ with $\Bbbk$ via the canonical isomorphism. Given a scalar object i of $\mathscr{C}$, the i-*isotypical component* $X^{(i)}$ of an object X is the largest direct factor of X isomorphic to a direct sum of copies of i. The actual number of copies of i is

$$\nu_i(X) = \mathrm{rank}_{\Bbbk} \mathrm{Hom}_{\mathscr{C}}(i, X) = \mathrm{rank}_{\Bbbk} \mathrm{Hom}_{\mathscr{C}}(X, i).$$

An i-*decomposition* of X is an explicit direct sum decomposition of $X^{(i)}$ into copies of i, that is, a family $(p_\alpha : X \to i, q_\alpha : i \to X)_{\alpha \in A}$ of pairs of morphisms in $\mathscr{C}$ such that

(a) $p_\alpha\, q_\beta = \delta_{\alpha,\beta}\, \mathrm{id}_i$ for all $\alpha, \beta \in A$,

(b) the set A has $\nu_i(X)$ elements,

where $\delta_{\alpha,\beta}$ is the Kronecker symbol.

A *representative set of scalar objects* of $\mathscr{C}$ is a set I of scalar objects such that $\mathbb{1} \in I$ and every scalar object of $\mathscr{C}$ is isomorphic to exactly one element of I.

Note that if $\Bbbk$ is a field, a fusion category over $\Bbbk$ is abelian and semisimple. Recall that an abelian category is *semisimple* if its objects are direct sums of simple[1] objects.

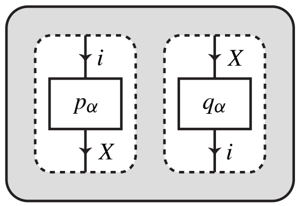
A pivotal fusion category is spherical (see Section 1E) if and only if the left and right dimension of any of its scalar objects coincide.

[1] An object of an abelian category is *simple* if it is nonzero and has no other subobject than the zero object and itself.

1H. ***Graphical calculus in pivotal fusion categories.*** Let $\mathscr{C}$ be a pivotal fusion category. Let X be an object of $\mathscr{C}$ and i be a scalar object of $\mathscr{C}$. Then the tensor

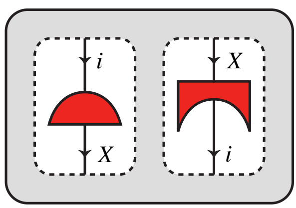
$$\sum_{\alpha\in A} p_\alpha \otimes_{\Bbbk} q_\alpha \in \mathrm{Hom}_{\mathscr{C}}(X, i) \otimes_{\Bbbk} \mathrm{Hom}_{\mathscr{C}}(i, X),$$

where $(p_\alpha, q_\alpha)_{\alpha\in A}$ is an i-decomposition of X, does not depend on the choice of the i-decomposition $(p_\alpha, q_\alpha)_{\alpha\in A}$ of X. Consequently, a sum of the type

$\sum_{\alpha\in A}$ 

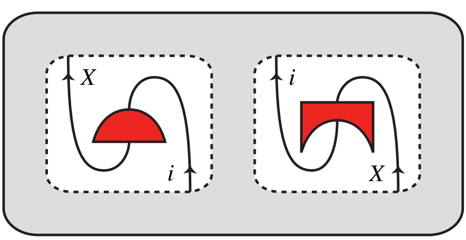
,

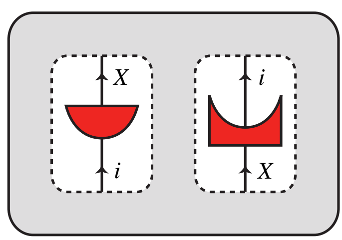
where $(p_\alpha, q_\alpha)_{\alpha\in A}$ is an i-decomposition of an object X and the gray area does not involve α, represents a morphism in $\mathscr{C}$ which is independent of the choice of the i-decomposition. We depict it as

(2) 

,

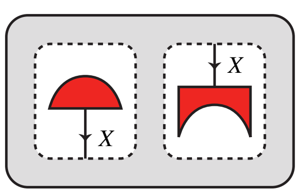
where the two curvilinear boxes should be shaded with the same color. If several such pairs of boxes appear in a picture, they must have different colors. We will also depict



as


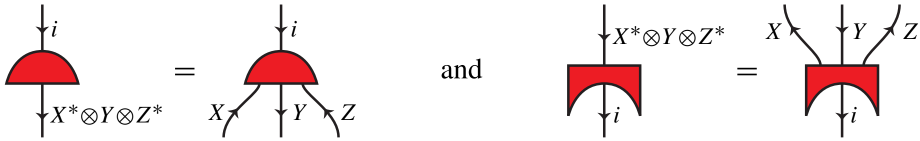
.

As usual, the edges labeled with $i = \mathbb{1}$ may be erased and then (2) becomes



.

Note also that tensor products of objects may be depicted as bunches of strands. For example,



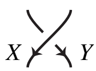
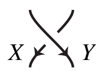
where the equality sign means that the pictures represent the same morphism of $\mathscr{C}$.

1I. *Braided and ribbon categories.* A *braiding* in a monoidal category $\mathscr{B}$ is a natural isomorphism $\tau = \{\tau_{X,Y} : X \otimes Y \to Y \otimes X\}_{X,Y\in\mathscr{B}}$ such that

$$\tau_{X,Y\otimes Z} = (\mathrm{id}_Y \otimes \tau_{X,Z})(\tau_{X,Y} \otimes \mathrm{id}_Z) \quad \text{and} \quad \tau_{X\otimes Y,Z} = (\tau_{X,Z} \otimes \mathrm{id}_Y)(\mathrm{id}_X \otimes \tau_{Y,Z})$$

for all X, Y, Z objects of $\mathscr{C}$. These conditions imply that $\tau_{X,\mathbb{1}} = \tau_{\mathbb{1},X} = \mathrm{id}_X$.

A monoidal category endowed with a braiding is said to be *braided*. The braiding and its inverse are depicted as

$$\tau_{X,Y} =$$  and $$\tau^{-1}_{Y,X} =$$  .

Note that any braided category satisfies the condition (1) of Section 1C.

For any object X of a braided pivotal category $\mathscr{B}$, the morphism

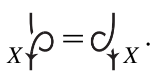
$$\theta_X =$$  $$= (\mathrm{id}_X \otimes \widetilde{\mathrm{ev}}_X)(\tau_{X,X} \otimes \mathrm{id}_{X^*})(\mathrm{id}_X \otimes \mathrm{coev}_X) : X \to X$$

is called the *twist*. The twist is natural in X and invertible, with inverse

$$\theta_X^{-1} =$$  $$= (\mathrm{ev}_X \otimes \mathrm{id}_X)(\mathrm{id}_{X^*} \otimes \tau^{-1}_{X,X})(\widetilde{\mathrm{coev}}_X \otimes \mathrm{id}_X) : X \to X.$$

It satisfies $\theta_{X\otimes Y} = (\theta_X \otimes \theta_Y)\tau_{Y,X}\tau_{X,Y}$ for all objects X, Y of $\mathscr{B}$ and $\theta_{\mathbb{1}} = \mathrm{id}_{\mathbb{1}}$.

A *ribbon category* is a braided pivotal category $\mathscr{B}$ whose twist θ is self-dual; i.e., $(\theta_X)^* = \theta_{X^*}$ for any object X of $\mathscr{B}$. This is equivalent to the equality



A ribbon category is spherical.

1J. *The center of a monoidal category.* Let $\mathscr{C}$ be a monoidal category. A *half braiding* of $\mathscr{C}$ is a pair (A, σ), where A is an object of $\mathscr{C}$ and

$$\sigma = \{\sigma_X : A \otimes X \to X \otimes A\}_{X\in\mathscr{C}}$$

is a natural isomorphism such that

$$\sigma_{X\otimes Y} = (\mathrm{id}_X \otimes \sigma_Y)(\sigma_X \otimes \mathrm{id}_Y) \tag{3}$$

for all X, Y objects of $\mathscr{C}$. This implies that $\sigma_{\mathbb{1}} = \mathrm{id}_A$.

The *center of* $\mathscr{C}$ is the braided category $\mathscr{Z}(\mathscr{C})$ defined as follows. The objects of $\mathscr{Z}(\mathscr{C})$ are half braidings of $\mathscr{C}$. A morphism $(A, \sigma) \to (A', \sigma')$ in $\mathscr{Z}(\mathscr{C})$ is a morphism $f : A \to A'$ in $\mathscr{C}$ such that $(\mathrm{id}_X \otimes f)\sigma_X = \sigma'_X(f \otimes \mathrm{id}_X)$ for any object X of $\mathscr{C}$. The unit object of $\mathscr{Z}(\mathscr{C})$ is $\mathbb{1}_{\mathscr{Z}(\mathscr{C})} = (\mathbb{1}, \{\mathrm{id}_X\}_{X \in \mathscr{C}})$ and the monoidal product is

$$(A, \sigma) \otimes (B, \rho) = \big(A \otimes B, (\sigma \otimes \mathrm{id}_B)(\mathrm{id}_A \otimes \rho)\big).$$

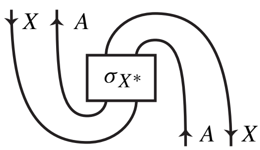
The braiding τ in $\mathscr{Z}(\mathscr{C})$ is defined by

$$\tau_{(A,\sigma),(B,\rho)} = \sigma_B : (A, \sigma) \otimes (B, \rho) \to (B, \rho) \otimes (A, \sigma).$$

There is a *forgetful functor* $\mathscr{U} : \mathscr{Z}(\mathscr{C}) \to \mathscr{C}$ assigning to every half braiding (A, σ) the underlying object A and acting in the obvious way on the morphisms. This is a strict monoidal functor.

If $\mathscr{C}$ satisfies (1), then $\mathrm{End}_{\mathscr{Z}(\mathscr{C})}(\mathbb{1}_{\mathscr{Z}(\mathscr{C})}) = \mathrm{End}_{\mathscr{C}}(\mathbb{1})$.

If $\mathscr{C}$ is rigid, then so is $\mathscr{Z}(\mathscr{C})$. If $\mathscr{C}$ is pivotal, then so is $\mathscr{Z}(\mathscr{C})$ with $(A, \sigma)^* = (A^*, \sigma^\natural)$, where

$\sigma^\natural_X =$  $: A^* \otimes X \to X \otimes A^*,$

and $\mathrm{ev}_{(A,\sigma)} = \mathrm{ev}_A$, $\mathrm{coev}_{(A,\sigma)} = \mathrm{coev}_A$, $\widetilde{\mathrm{ev}}_{(A,\sigma)} = \widetilde{\mathrm{ev}}_A$, $\widetilde{\mathrm{coev}}_{(A,\sigma)} = \widetilde{\mathrm{coev}}_A$. In that case the forgetful functor $\mathscr{U}$ preserves (left and right) traces of morphisms and dimensions of objects.

If $\mathscr{C}$ is a $\Bbbk$-additive monoidal category, then so is $\mathscr{Z}(\mathscr{C})$ and the forgetful functor is $\Bbbk$-linear. If $\mathscr{C}$ is an abelian rigid category, then so is $\mathscr{Z}(\mathscr{C})$, and the forgetful functor is exact.

If $\mathscr{C}$ is a fusion category over the ring $\Bbbk$, then $\mathscr{Z}(\mathscr{C})$ is braided $\Bbbk$-additive rigid category whose monoidal unit is scalar. If in addition $\Bbbk$ is field, then $\mathscr{C}$ is abelian, and so is $\mathscr{Z}(\mathscr{C})$.

2. Main results

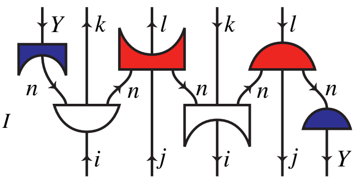
In this section, we state our main results concerning the center of a pivotal fusion category. They are proved in Section 4. Let $\mathscr{C}$ be a pivotal fusion category over a commutative ring $\Bbbk$ and I be a representative set of scalar objects of $\mathscr{C}$. Recall from Section 1J that the center $\mathscr{Z}(\mathscr{C})$ of $\mathscr{C}$ is a braided $\Bbbk$-additive pivotal category whose monoidal unit is scalar.

The coend of a rigid braided category is, if it exists, a Hopf algebra in the category which coacts universally on the objects (see Section 3C for details). The center

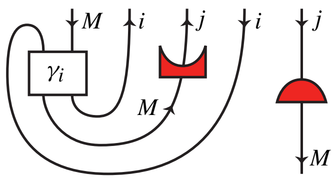
$\mathcal{Z}(\mathscr{C})$ of $\mathscr{C}$ has a coend (C,σ), where

$$C = \bigoplus_{i,j\in I} i^* \otimes j^* \otimes i \otimes j$$

and the half braiding $\sigma = \{\sigma_Y\}_{Y\in\mathscr{C}}$ is given by

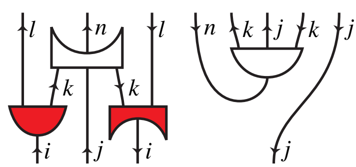
(4) $\sigma_Y = \sum_{i,j,k,\ell,n\in I}$  $: C\otimes Y \to Y\otimes C.$

The universal coaction $\delta = \{\delta_{M,\gamma}\}_{(M,\gamma)\in\mathcal{Z}(\mathscr{C})}$ of the coend (C,σ) is

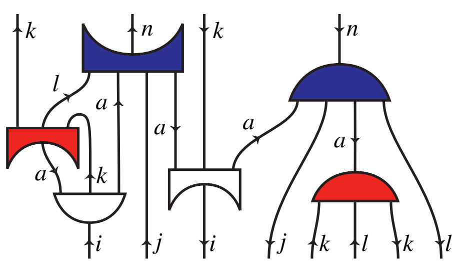
(5) $\delta_{(M,\gamma)} = \sum_{i,j\in I}$  $: (M,\gamma) \to (M,\gamma)\otimes(C,\sigma).$

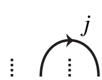
The structural morphisms and the canonical pairing of the Hopf algebra (C,σ) can be depicted as follows:

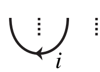
(a) The coproduct $\Delta : C \to C\otimes C$:

$\Delta = \sum_{i,j,k,\ell,n\in I}$ 

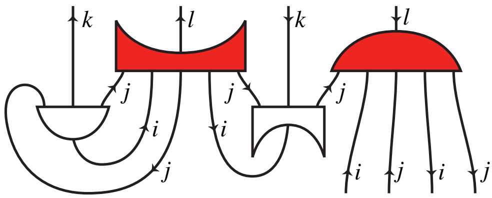
(b) The product $m : C\otimes C \to C$:

$m = \sum_{i,j,k,\ell,n,a\in I}$ 

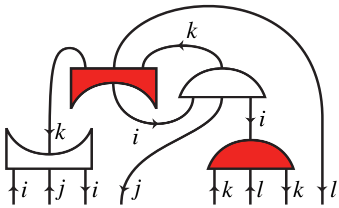
(c) The counit $\varepsilon : C \to \mathbb{1}$: $\varepsilon = \sum_{j\in I}$ 

(d) The unit $u : \mathbb{1} \to C$: $u = \sum_{i\in I}$ 

(e) The antipode $S : C \to C$:

$S = \sum_{i,j,k,\ell,n\in I}$ 

(f) The canonical pairing $\omega : C \otimes C \to \mathbb{1}$:

$\omega = \sum_{i,j,k,\ell\in I}$ 

In the pictures, the dotted lines represent $\mathrm{id}_{\mathbb{1}}$ and serve to indicate which direct factor of C is concerned. Moreover,

(6) $\Lambda = \sum_{j\in I} \dim_r(j)$  $: (\mathbb{1}, \mathrm{id}) \to (C, \sigma)$

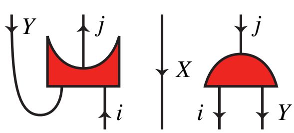
is an integral of the Hopf algebra (C, σ), which is invariant under the antipode.

By a modular category we mean a braided pivotal category admitting a coend, and whose canonical pairing is nondegenerate (see Section 3E for details). The dimension of such a category is the dimension of its coend (see Section 3D).

Theorem 2.1. *The center* $\mathcal{Z}(\mathcal{C})$ *of* $\mathcal{C}$ *is modular and has dimension* $\dim(\mathcal{C})^2$.

The forgetful functor $\mathcal{U} : \mathcal{Z}(\mathcal{C}) \to \mathcal{C}$ has a left adjoint $\mathcal{F} : \mathcal{C} \to \mathcal{Z}(\mathcal{C})$. For an object X of $\mathcal{C}$,

$$\mathcal{F}(X) = \big(Z(X), \varsigma_X = \{\varsigma_{X,Y}\}_{Y\in\mathcal{C}}\big) \quad \text{where} \quad Z(X) = \bigoplus_{i\in I} i^* \otimes X \otimes i \quad \text{and}$$

$\varsigma_{X,Y} = \sum_{i,j\in I}$  $: Z(X) \otimes Y \to Y \otimes Z(X).$

For a morphism f in $\mathcal{C}$,

$$\mathcal{F}(f) = \sum_{i\in I} \mathrm{id}_{i^*} \otimes f \otimes \mathrm{id}_i .$$

By a *free half braiding* we mean a half braiding of the form $\mathcal{F}(X)$ for some object X of $\mathcal{C}$.

Theorem 2.2. *The dimension of* $\mathcal{C}$ *is invertible in* $\Bbbk$ *if and only if every half braiding is a retract of a free half braiding.*

From Section 1J, if $\Bbbk$ is a field, then $\mathcal{Z}(\mathcal{C})$ is abelian.

Corollary 2.3. *Assume $\Bbbk$ is a field.*

(a) *The center $\mathcal{Z}(\mathcal{C})$ is semisimple (as an abelian category) if and only if* $\dim(\mathcal{C}) \neq 0$.

(b) *Assume $\Bbbk$ is algebraically closed. Then $\mathcal{Z}(\mathcal{C})$ is a fusion category if and only if* $\dim(\mathcal{C}) \neq 0$.

Since the center of a spherical fusion category is ribbon (see, for example, [Turaev and Virelizier 2010, Lemma 10.1]), we recover Müger's theorem:

Corollary 2.4 [Müger 2003, Theorem 1.2]. *If $\mathcal{C}$ is a spherical fusion category over an algebraically closed field and* $\dim(\mathcal{C}) \neq 0$, *then $\mathcal{Z}(\mathcal{C})$ is a modular ribbon fusion category* (i.e., $\mathcal{Z}(\mathcal{C})$ is modular in the sense of [Turaev 1994]).

Note that by [Etingof et al. 2005], the hypothesis $\dim(\mathcal{C}) \neq 0$ of the previous corollary is automatically fulfilled on a field of characteristic zero.

Example 2.5. Let G be a finite group and $\Bbbk$ be a commutative ring. The category $\mathcal{C}_{G,\Bbbk}$ of G-graded free $\Bbbk$-modules of finite rank is a spherical fusion category. The dimension of $\mathcal{C}_{G,\Bbbk}$ is $\dim(\mathcal{C}_{G,\Bbbk}) = |G|1_\Bbbk$, where $|G|$ is the order of G. By Theorem 2.1, the center $\mathcal{Z}(\mathcal{C}_{G,\Bbbk})$ of $\mathcal{C}_{G,\Bbbk}$ is modular of dimension $|G|^2 1_\Bbbk$. When $|G|$ is not invertible in $\Bbbk$, by Theorem 2.2, there exist half braidings of $\mathcal{C}_{G,\Bbbk}$ which are not retracts of any free half braiding. If particular, if $\Bbbk$ is a field of characteristic p which divides $|G|$, then $\mathcal{Z}(\mathcal{C}_{G,\Bbbk})$ is not semisimple.

3. Modular categories

In this section, we clarify some notions used in the previous section. More precisely, in Section 3A, we recall the definition of a Hopf algebra in a braided category and provide a criterion for the nondegeneracy of a Hopf algebra pairing. In Section 3B, we recall the definition of a coend. In Section 3C, we describe the Hopf algebra structure of the coend of a braided rigid category. Sections 3D and 3E are devoted to the definition of respectively the dimension and the modularity of a braided category admitting a coend.

3A. ***Hopf algebras, pairings, and integrals.*** Let $\mathcal{B}$ be a braided category, with braiding τ. Recall that a *bialgebra in $\mathcal{B}$* is an object A of $\mathcal{B}$ endowed with four morphisms $m : A \otimes A \to A$ (the product), $u : \mathbb{1} \to A$ (the unit), $\Delta : A \to A \otimes A$ (the coproduct), and $\varepsilon : A \to \mathbb{1}$ (the counit) such that

$$m(m \otimes \mathrm{id}_A) = m(\mathrm{id}_A \otimes m), \quad m(\mathrm{id}_A \otimes u) = \mathrm{id}_A = m(u \otimes \mathrm{id}_A),$$
$$(\Delta \otimes \mathrm{id}_A)\Delta = (\mathrm{id}_A \otimes \Delta)\Delta, \quad (\mathrm{id}_A \otimes \varepsilon)\Delta = \mathrm{id}_A = (\varepsilon \otimes \mathrm{id}_A)\Delta,$$
$$\Delta m = (m \otimes m)(\mathrm{id}_A \otimes \tau_{A,A} \otimes \mathrm{id}_A)(\Delta \otimes \Delta),$$
$$\Delta u = u \otimes u, \quad \varepsilon m = \varepsilon \otimes \varepsilon, \quad \varepsilon u = \mathrm{id}_{\mathbb{1}}.$$

An *antipode* for a bialgebra A in $\mathcal{B}$ is a morphism $S\colon A \to A$ in $\mathcal{B}$ such that

$$m(S \otimes \mathrm{id}_A)\Delta = u\varepsilon = m(\mathrm{id}_A \otimes S)\Delta.$$

If it exists, an antipode is unique. A *Hopf algebra in* $\mathcal{B}$ is a bialgebra in $\mathcal{B}$ which admits an invertible antipode.

Let A be a Hopf algebra in $\mathcal{B}$. A *Hopf pairing* for A is a morphism $\omega\colon A\otimes A \to \mathbb{1}$ such that

$$\omega(m \otimes \mathrm{id}_A) = \omega(\mathrm{id}_A \otimes \omega \otimes \mathrm{id}_A)(\mathrm{id}_{A^{\otimes 2}} \otimes \Delta), \quad \omega(u \otimes \mathrm{id}_A) = \varepsilon,$$
$$\omega(\mathrm{id}_A \otimes m) = \omega(\mathrm{id}_A \otimes \omega \otimes \mathrm{id}_A)(\Delta \otimes \mathrm{id}_{A^{\otimes 2}}), \quad \omega(\mathrm{id}_A \otimes u) = \varepsilon.$$

These axioms imply that $\omega(S \otimes \mathrm{id}_A) = \omega(\mathrm{id}_A \otimes S)$.

A Hopf pairing ω for A is *nondegenerate* if there exists a morphism $\Omega\colon \mathbb{1} \to A\otimes A$ in $\mathcal{B}$ such that

$$(\omega \otimes \mathrm{id}_A)(\mathrm{id}_A \otimes \Omega) = \mathrm{id}_A = (\mathrm{id}_A \otimes \omega)(\Omega \otimes \mathrm{id}_A).$$

If such is the case, the morphism Ω is unique and called the *inverse* of ω.

A *left* (respectively, *right*) *integral* for A is a morphism $\Lambda : \mathbb{1} \to A$ such that

$$m(\mathrm{id}_A \otimes \Lambda) = \Lambda\, \varepsilon \quad (\text{respectively, } m(\Lambda \otimes \mathrm{id}_A) = \Lambda\, \varepsilon).$$

A *left* (respectively, *right*) *cointegral* for A is a morphism $\lambda : A \to \mathbb{1}$ such that

$$(\mathrm{id}_A \otimes \lambda)\Delta = u\, \lambda \quad (\text{respectively, } (\lambda \otimes \mathrm{id}_A)\Delta = u\, \lambda).$$

A (co)integral is *two-sided* if it is both a left and a right (co)integral.

If Λ is a left (respectively, right) integral for A, then $S\Lambda$ is a right (respectively, left) integral for A. If λ is a left (respectively, right) cointegral for A, then λS is a right (respectively, left) cointegral for A.

Let ω be a Hopf pairing for A and $\Lambda : \mathbb{1} \to A$ be a morphism in $\mathcal{B}$. Assume ω is nondegenerate. Then Λ is a left integral for A if and only if $\lambda = \omega(\mathrm{id}_A \otimes \Lambda)$ is a right cointegral for A, and Λ is a right integral for A if and only if $\lambda = \omega(\Lambda \otimes \mathrm{id}_A)$ is a left cointegral for A.

Lemma 3.1. *Let ω be a Hopf pairing for a Hopf algebra A in a braided category $\mathcal{B}$. Assume there exist morphisms $\Lambda, \Lambda' : \mathbb{1} \to A$ in $\mathcal{B}$ such that*

(a) $\omega(\Lambda \otimes \mathrm{id}_A)$ *and* $\omega(\mathrm{id}_A \otimes \Lambda')$ *are left cointegrals for A;*

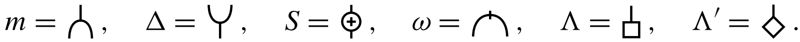
(b) $\omega(\Lambda \otimes \Lambda')$ *is invertible in* $\mathrm{End}_{\mathcal{B}}(\mathbb{1})$.

Then ω is nondegenerate, with inverse

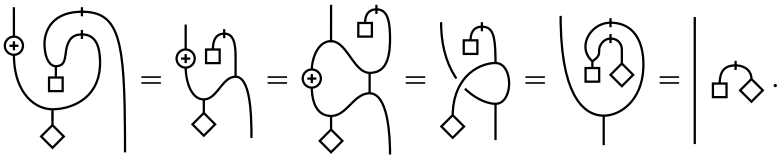
$$\Omega = \omega(\Lambda \otimes \Lambda')^{-1}\, (S \otimes \mathrm{id}_A \otimes \omega)(\mathrm{id}_A \otimes \Delta\Lambda \otimes \mathrm{id}_A)\Delta\Lambda',$$

and Λ and Λ' are right integrals for A.

Proof. Set $e = (S \otimes \mathrm{id}_A \otimes \omega)(\mathrm{id}_A \otimes \Delta\Lambda \otimes \mathrm{id}_A)\Delta\Lambda' : \mathbb{1} \to A \otimes A$. Let us depict the product m, coproduct Δ, antipode S of A, and the morphisms ω, Λ, Λ' as follows:



Then $(\mathrm{id}_A \otimes \omega)(e \otimes \mathrm{id}_A) = \omega(\Lambda \otimes \Lambda')\,\mathrm{id}_A$ since



We use the product/coproduct axioms of a Hopf pairing in the first and fourth equalities, the unit axiom and the fact that $\omega(\Lambda \otimes \mathrm{id}_A)$ is a left cointegral in the second equality, the compatibility of m and Δ and the axiom of the antipode in the third equality, and finally the fact that $\omega(\mathrm{id}_A \otimes \Lambda')$ is a left cointegral and the unit/counit axiom of a Hopf pairing in the last equality. Similarly one shows that $(\omega \otimes \mathrm{id}_A)(\mathrm{id}_A \otimes e) = \omega(\Lambda \otimes \Lambda')\,\mathrm{id}_A$. Thus $\Omega = \omega(\Lambda \otimes \Lambda')^{-1} e$ is an inverse of ω.

Finally, since ω is nondegenerate and $\omega(\Lambda \otimes A)$ and $\omega(A \otimes \Lambda')$ are left cointegrals, we conclude that Λ and Λ' are right integrals. □

3B. *Coends.* Let $\mathscr{C}$ and $\mathscr{D}$ be categories. A *dinatural transformation* from a functor $F : \mathscr{D}^{\mathrm{op}} \times \mathscr{D} \to \mathscr{C}$ to an object A of $\mathscr{C}$ is a family of morphisms in $\mathscr{C}$

$$d = \{d_Y : F(Y, Y) \to A\}_{Y \in \mathscr{D}}$$

such that for every morphism $f : X \to Y$ in $\mathscr{D}$, we have

$$d_X F(f, \mathrm{id}_X) = d_Y F(\mathrm{id}_Y, f) : F(Y, X) \to A.$$

The *composition* of such a d with a morphism $\phi : A \to B$ in $\mathscr{C}$ is the dinatural transformation $\phi \circ d = \{\phi \circ d_X : F(Y, Y) \to B\}_{Y \in \mathscr{D}}$ from F to B. A *coend* of F is a pair (C, ρ) consisting in an object C of $\mathscr{C}$ and a dinatural transformation ρ from F to C satisfying the following universality condition: every dinatural transformation d from F to an object of $\mathscr{C}$ is the composition of ρ with a morphism in $\mathscr{C}$ uniquely determined by d. If F has a coend (C, ρ), then it is unique (up to unique isomorphism). One writes $C = \int^{Y \in \mathscr{D}} F(Y, Y)$. For more on coends, see [Mac Lane 1998].

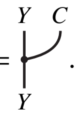
Remark 3.2. Let $F : \mathscr{D}^{\mathrm{op}} \times \mathscr{D} \to \mathscr{C}$ be a $\Bbbk$-linear functor, where $\mathscr{C}$ is a $\Bbbk$-additive category and $\mathscr{D}$ is a fusion category (over $\Bbbk$). Then F has a coend. More precisely, pick a (finite) representative set I of simple objects of $\mathscr{D}$ and set $C = \bigoplus_{i \in I} F(i, i)$. Let $\rho = \{\rho_Y : F(Y, Y) \to C\}_{Y \in \mathscr{D}}$ be defined by $\rho_Y = \sum_\alpha F(q_Y^\alpha, p_Y^\alpha)$, where $(p_Y^\alpha, q_Y^\alpha)_\alpha$ is any I-partition of Y. Then (C, ρ) is a coend of F and each dinatural transformation d from F to any object A of $\mathscr{C}$ is the composition of ρ with $\bigoplus_{i \in I} d_i : C \to A$.

3C. ***The coend of a braided rigid category.*** Let $\mathscr{B}$ be braided rigid category. The coend

$$C = \int^{Y \in \mathscr{B}} {}^{\vee}Y \otimes Y,$$

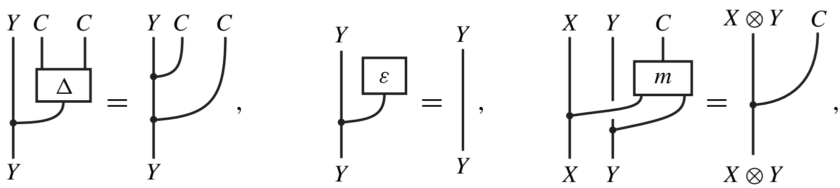
if it exists, is called the *coend of* $\mathscr{B}$.

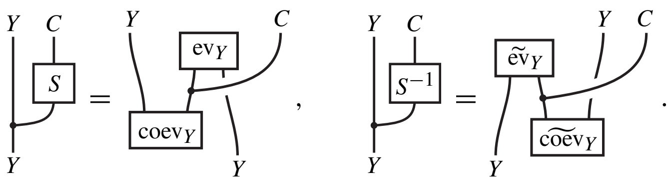
Assume $\mathscr{B}$ has a coend C and denote by $i_Y : {}^{\vee}Y \otimes Y \to C$ the corresponding universal dinatural transformation. The *universal coaction* of C on the objects of $\mathscr{B}$ is the natural transformation δ defined by

(7) $\quad \delta_Y = (\mathrm{id}_Y \otimes i_Y)(\mathrm{coev}_Y \otimes \mathrm{id}_Y) : Y \to Y \otimes C, \quad$ depicted as $\delta_Y =$ 

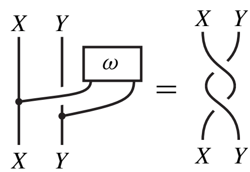
.

According to [Majid 1995], C is a Hopf algebra in $\mathscr{B}$. Its coproduct Δ, product m, counit ε, unit u, and antipode S with inverse S^{-1} are characterized by the following equalities, where $X, Y \in \mathscr{B}$:



$u = \delta_{\mathbb{1}}$, 

Furthermore, the morphism $\omega : C \otimes C \to \mathbb{1}$ defined by

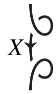


is a Hopf pairing for C, called the *canonical pairing*. Moreover this pairing satisfies the following self-duality condition: $\omega \tau_{C,C}(S \otimes S) = \omega$.

3D. ***The dimension of a braided pivotal category.*** Let $\mathscr{B}$ be a braided pivotal category admitting a coend C.

Lemma 3.3. *The left and right dimensions of C coincide.*

Proof. Let $\upsilon = \{\upsilon_X\}_{X \in \mathcal{B}}$ be the natural transformation defined by

$$\upsilon_X = $$  $$: X \to X.$$

Then υ is natural monoidal isomorphism; that is, $\upsilon_{X\otimes Y} = \upsilon_X \otimes \upsilon_Y$ and $\upsilon_{\mathbb{1}} = \mathrm{id}_{\mathbb{1}}$, which implies that $\upsilon_X{}^* = \upsilon_{X^*}^{-1}$. The full subcategory $\mathcal{B}_0$ of $\mathcal{B}$ made of the objects X of $\mathcal{B}$ satisfying $\tau_X = \mathrm{id}_X$ is a ribbon category. Let us prove that the coend C of $\mathcal{B}$ belongs to $\mathcal{B}_0$. Denote by $i = \{i_X : X^* \otimes X \to C\}_{X\in\mathcal{B}}$ the universal dinatural transformation associated with C. For any object X of $\mathcal{C}$, by naturality and monoidality of υ and dinaturality of i, the following holds:

$$\upsilon_C i_X = i_X \upsilon_{(X^*\otimes X)} = i_X(\upsilon_{X^*} \otimes \upsilon_X) = i_X(\upsilon_X^* \upsilon_{X^*} \otimes \mathrm{id}_X) = i_X.$$

So $\upsilon_C = \mathrm{id}_C$; that is, C belongs to $\mathcal{B}_0$. Hence the left and right dimensions of C coincide, since $\mathcal{B}_0$ is a ribbon category. □

We define the *dimension of* $\mathcal{B}$ as $\dim(\mathcal{B}) = \dim_l(C) = \dim_r(C)$.

This definition agrees with the standard definition of the dimension of a pivotal fusion category. Indeed, any pivotal fusion category $\mathcal{C}$ (over the ring $\Bbbk$) admits a coend $C = \bigoplus_{i\in I} i^* \otimes i$, where I is a (finite) representative set of scalar objects of $\mathcal{C}$, and so

$$\dim_l(C) = \dim_r(C) = \sum_{i\in I} \dim_l(i^*)\dim_l(i) = \sum_{i\in I} \dim_r(i)\dim_l(i).$$

3E. ***Modular categories.*** By a *modular category*, we mean a braided rigid category which admits a coend whose canonical pairing is nondegenerate. Note that when $\mathcal{B}$ is ribbon, this definition coincides with that of a *2-modular category* given in [Lyubashenko 1995].

Remark 3.4. Let $\mathcal{B}$ be a braided pivotal fusion category over $\Bbbk$. Let I be a representative set of the scalar objects of $\mathcal{B}$. Recall that $C = \bigoplus_{i\in I} i^* \otimes i$ is the coend of $\mathcal{B}$. For $i,\, j \in I$, set

$$S_{i,j} = (\mathrm{ev}_i \otimes \widetilde{\mathrm{ev}}_j)(\mathrm{id}_{i^*} \otimes \tau_{j,i}\tau_{i,j} \otimes \mathrm{id}_{j^*})(\widetilde{\mathrm{coev}}_i \otimes \mathrm{coev}_j) \in \Bbbk.$$

The matrix $S = [S_{i,j}]_{i,j\in I}$, called the *S-matrix* of $\mathcal{B}$, is invertible if and only if the canonical pairing of C is nondegenerate. In particular a modular category in the sense of [Turaev 1994] is a ribbon fusion category which is modular in the above sense.

4. Proofs

The statements of Section 2 derive directly from the theory of Hopf monads, introduced in [Bruguières and Virelizier 2007] and developed in [Bruguières and

Virelizier 2012; Bruguières et al. 2011]. Hopf monads generalize Hopf algebras in the setting of general monoidal categories. In Section 4A, we recall some basic definitions concerning Hopf monads. In Section 4B, we give a Hopf monadic description of the center $\mathcal{Z}(\mathcal{C})$ of a fusion category $\mathcal{C}$, from which is derived the explicit description of the coend of $\mathcal{Z}(\mathcal{C})$. In Section 4C, we prove a "handleslide" property for pivotal fusion categories. In Section 4D, we use the explicit description of the coend of $\mathcal{Z}(\mathcal{C})$ to prove Theorem 2.1 and prove that the morphism Λ of (6) is an integral invariant under the antipode. Sections 4E and 4F are devoted to the proofs of Theorem 2.2 and Corollary 2.3, respectively.

4A. ***Hopf monads and their modules.*** Let $\mathcal{C}$ be a category. A *monad* on $\mathcal{C}$ is a monoid in the category of endofunctors of $\mathcal{C}$, that is, a triple (T, μ, η) consisting of a functor $T : \mathcal{C} \to \mathcal{C}$ and two natural transformations

$$\mu = \{\mu_X : T^2(X) \to T(X)\}_{X\in\mathcal{C}} \quad \text{and} \quad \eta = \{\eta_X : X \to T(X)\}_{X\in\mathcal{C}},$$

called the *product* and the *unit* of T, such that, for any object X of $\mathcal{C}$,

$$\mu_X T(\mu_X) = \mu_X \mu_{T(X)} \quad \text{and} \quad \mu_X \eta_{T(X)} = \mathrm{id}_{T(X)} = \mu_X T(\eta_X).$$

Given a monad $T = (T, \mu, \eta)$ on $\mathcal{C}$, a T-module in $\mathcal{C}$ is a pair (M, r) where M is an object of $\mathcal{C}$ and $r : T(M) \to M$ is a morphism in $\mathcal{C}$ such that $rT(r) = r\mu_M$ and $r\eta_M = \mathrm{id}_M$. A morphism from a T-module (M, r) to a T-module (N, s) is a morphism $f : M \to N$ in $\mathcal{C}$ such that $fr = sT(f)$. This defines the *category* $\mathcal{C}^T$ *of* T*-modules in* $\mathcal{C}$ with composition induced by that in $\mathcal{C}$. We define a forgetful functor $U_T : \mathcal{C}^T \to \mathcal{C}$ by $U_T(M, r) = M$ and $U_T(f) = f$. The forgetful functor U_T has a left adjoint $F_T : \mathcal{C} \to \mathcal{C}^T$, called the free module functor, defined by $F_T(X) = (T(X), \mu_X)$ and $F_T(f) = T(f)$. Note that if $\mathcal{C}$ is $\Bbbk$-additive and T is $\Bbbk$-linear (that is, T induces $\Bbbk$-linear maps on Hom spaces), then the category $\mathcal{C}^T$ is $\Bbbk$-additive and the functors U_T and F_T are $\Bbbk$-linear.

Let $\mathcal{C}$ be a monoidal category. A *bimonad* on $\mathcal{C}$ is a monoid in the category of comonoidal endofunctors of $\mathcal{C}$. In other words, a bimonad on $\mathcal{C}$ is a monad (T, μ, η) on $\mathcal{C}$ such that the functor $T : \mathcal{C} \to \mathcal{C}$ and the natural transformations μ and η are comonoidal; that is, T comes equipped with a natural transformation $T_2 = \{T_2(X, Y) : T(X \otimes Y) \to T(X) \otimes T(Y)\}_{X,Y\in\mathcal{C}}$ and a morphism $T_0 : T(\mathbb{1}) \to \mathbb{1}$ such that

$$\begin{gathered}
\big(\mathrm{id}_{T(X)} \otimes T_2(Y, Z)\big)T_2(X, Y \otimes Z) = \big(T_2(X, Y) \otimes \mathrm{id}_{T(Z)}\big)T_2(X \otimes Y, Z);\\
(\mathrm{id}_{T(X)} \otimes T_0)T_2(X, \mathbb{1}) = \mathrm{id}_{T(X)} = (T_0 \otimes \mathrm{id}_{T(X)})T_2(\mathbb{1}, X);\\
T_2(X, Y)\mu_{X\otimes Y} = (\mu_X \otimes \mu_Y)T_2(T(X), T(Y))T(T_2(X, Y));\\
T_2(X, Y)\eta_{X\otimes Y} = \eta_X \otimes \eta_Y.
\end{gathered}$$

For any bimonad T on $\mathscr{C}$, the category of T-modules $\mathscr{C}^T$ has a monoidal structure with unit object $(\mathbb{1}, T_0)$ and with tensor product

$$(M, r) \otimes (N, s) = \big(M \otimes N, (r \otimes s)\, T_2(M, N)\big).$$

Note that the forgetful functor $U_T : \mathscr{C}^T \to \mathscr{C}$ is strict monoidal.

Given a bimonad (T, μ, η) on $\mathscr{C}$ and objects $X, Y \in \mathscr{C}$, one defines the *left fusion operator*

$$H^l_{X,Y} = (T(X) \otimes \mu_Y) T_2(X, T(Y)) : T(X \otimes T(Y)) \to T(X) \otimes T(Y)$$

and the *right fusion operator*

$$H^r_{X,Y} = (\mu_X \otimes T(Y)) T_2(T(X), Y) : T(T(X) \otimes Y) \to T(X) \otimes T(Y).$$

A *Hopf monad* on $\mathscr{C}$ is a bimonad on $\mathscr{C}$ whose left and right fusion operators are isomorphisms for all objects X, Y of $\mathscr{C}$. When $\mathscr{C}$ is a rigid category, a bimonad T on $\mathscr{C}$ is a Hopf monad if and only if the category $\mathscr{C}^T$ is rigid. The structure of a rigid category in $\mathscr{C}^T$ can then be encoded in terms of natural transformations

$$s^l = \{s^l_X : T({}^\vee T(X)) \to {}^\vee X\}_{X \in \mathscr{C}} \quad \text{and} \quad s^r = \{s^r_X : T(T(X)^\vee) \to X^\vee\}_{X \in \mathscr{C}},$$

called the *left and right antipodes*. They are computed from the fusion operators:

$$s^l_X = \big(T_0 T(\mathrm{ev}_{T(X)}) (H^l_{{}^\vee T(X), X})^{-1} \otimes {}^\vee \eta_X\big)\big(\mathrm{id}_{T({}^\vee T(X))} \otimes \mathrm{coev}_{T(X)}\big);$$
$$s^r_X = \big(\eta_X^\vee \otimes T_0 T(\widetilde{\mathrm{ev}}_{T(X)}) (H^r_{X, T(X)^\vee})^{-1}\big)\big(\widetilde{\mathrm{coev}}_{T(X)} \otimes \mathrm{id}_{T(T(X)^\vee)}\big).$$

The left and right duals of any T-module (M, r) are then defined by

$${}^\vee(M, r) = \big({}^\vee M, s^l_M T({}^\vee r)\big) \quad \text{and} \quad \big(M, r)^\vee = (M^\vee, s^r_M T(r^\vee)\big).$$

A *quasitriangular Hopf monad* on $\mathscr{C}$ is a Hopf monad T on $\mathscr{C}$ equipped with an R-matrix, that is, a natural transformation

$$R = \{R_{X,Y} : X \otimes Y \to T(Y) \otimes T(X)\}_{X,Y \in \mathscr{C}}$$

satisfying appropriate axioms which ensure that the natural transformation $\tau = \{\tau_{(M,r),(N,s)}\}_{(M,r),(N,s) \in \mathscr{C}^T}$ defined by

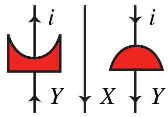
$$\tau_{(M,r),(N,s)} = (s \otimes r) R_{M,N} : (M, r) \otimes (N, s) \to (N, s) \otimes (M, r)$$

form a braiding in the category $\mathscr{C}^T$ of T-modules.

4B. ***The coend of the center of a fusion category.*** Let $\mathscr{C}$ be a pivotal fusion category (over the ring $\Bbbk$), with a representative set of scalar objects I. For each object X of $\mathscr{C}$, by Remark 3.2, the $\Bbbk$-linear functor $\mathscr{C}^{\mathrm{op}} \times \mathscr{C} \to \mathscr{C}$, defined by $(U, V) \mapsto U^* \otimes X \otimes V$, has a coend

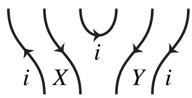
$$Z(X) = \bigoplus_{i\in I} i^* \otimes X \otimes i,$$

with dinatural transformation $\rho_X = \{\rho_{X,Y}\}_{Y\in\mathscr{C}}$ given by

$$\rho_{X,Y} = \sum_{i\in I}$$


$$: Y^* \otimes X \otimes Y \to Z(X).$$

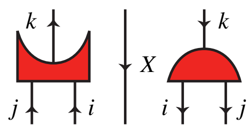
The correspondence $X \mapsto Z(X)$ extends to a functor $Z : \mathscr{C} \to \mathscr{C}$. By Theorem 6.4 and Section 9.2 of [Bruguières and Virelizier 2012], Z is a quasitriangular Hopf monad on $\mathscr{C}$, with structural morphisms as follows (the dotted lines represent $\mathrm{id}_{\mathbb{1}}$):

$$Z_2(X, Y) = \sum_{i\in I}$$


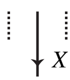
$$: Z(X \otimes Y) \to Z(X) \otimes Z(Y),$$

$$Z_0 = \sum_{i\in I}$$

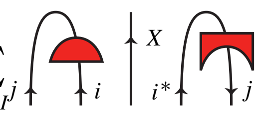

$$: Z(\mathbb{1}) \to \mathbb{1},$$

$$\mu_X = \sum_{i,j,k\in I}$$


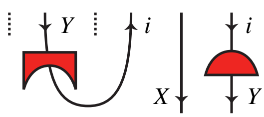
$$: Z^2(X) \to Z(X),$$

$$\eta_X =$$


$$: X \to X = \mathbb{1}^* \otimes X \otimes \mathbb{1} \hookrightarrow Z(X),$$

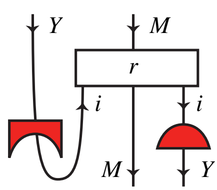
$$s^l_X = s^r_X = \sum_{i,j\in I}$$


$$: Z(Z(X)^*) \to X^*,$$

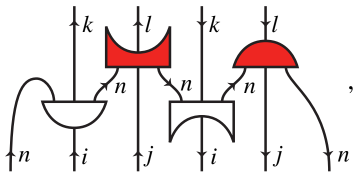
$$R_{X,Y} = \sum_{i\in I}$$


$$: X \otimes Y \to Z(Y) \otimes Z(X).$$

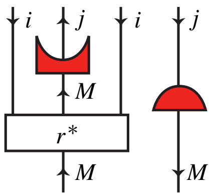
In particular, the category $\mathscr{C}^Z$ of Z-modules is a braided pivotal category. By [Bruguières and Virelizier 2012, Theorem 6.5], the functor

(8) $$\Phi : \left\{ \begin{array}{ccc} \mathscr{C}^Z & \to & \mathscr{Z}(\mathscr{C}) \\ (M, r) & \mapsto & (M, \sigma) \\ f & \mapsto & f \end{array} \right. \quad \text{where} \quad \sigma_Y = \sum_{i\in I}$$


is an isomorphism of braided pivotal categories. Note that this isomorphism is a "fusion" version of the braided isomorphism $\mathcal{Z}(\mathrm{mod}_H) \simeq \mathrm{mod}_{D(H)}$ between the center of the category of modules over a finite-dimensional Hopf algebra H and the category of modules over the Drinfeld double $D(H)$ of H. Now by [Bruguières and Virelizier 2012, Section 6.3], the coend of $\mathcal{C}^Z$ is (C, α), where

$$C = \bigoplus_{i,j\in I} i^* \otimes j^* \otimes i \otimes j \quad \text{and} \quad \alpha = \sum_{i,j,k,l,n\in I}$$  ,

with universal dinatural transformation $\iota = \{\iota_{(M,r)}\}_{(M,r)\in\mathcal{C}^Z}$ given by

$$\iota_{(M,r)} = \sum_{i,j\in I}$$  $: (M,r)^* \otimes (M,r) \to (C,\alpha).$

Thus $(C,\sigma) = \Phi(C,\alpha)$ is the coend of $\mathcal{Z}(\mathcal{C})$, with universal dinatural transformation $\{\Phi(\iota_{\Phi^{-1}(M,\gamma)})\}_{(M,\gamma)\in\mathcal{Z}(\mathcal{C})}$. Using the description of Φ and the definition of the universal coaction given in (7), we obtain that the half braiding σ is given by (4) and that the universal coaction of (C,σ) is given by (5). Finally, recall from Section 3C that (C,α) is a Hopf algebra in $\mathcal{C}^Z$ endowed with a canonical Hopf algebra pairing. By [Bruguières and Virelizier 2012, Section 9.3], the structural morphisms of (C,α) are those given on pages 10 and 11, items (a)–(f). These structural morphisms are also those of (C,σ), since Φ is the identity on morphisms.

4C. ***Slope and handleslide in pivotal fusion categories.*** Let $\mathcal{C}$ be a pivotal fusion category. Recall that the left and right dimensions of a scalar object of $\mathcal{C}$ are invertible. The *slope* of a scalar object i is the invertible scalar $\mathrm{sl}(i)$ defined by

$$\mathrm{sl}(i) = \frac{\dim_l(i)}{\dim_r(i)}.$$

The *slope* of an object X of $\mathcal{C}$ is the morphism $\mathrm{SL}_X : X \to X$ defined as

$$\mathrm{SL}_X = \sum_{\alpha\in A} \mathrm{sl}(i_\alpha)\, q_\alpha p_\alpha,$$

where $(p_\alpha : X \to i_\alpha, q_\alpha : i_\alpha \to X)_{\alpha\in A}$ is a decomposition of X as a sum of scalar objects, that is, a family of pairs of morphisms such that i_α is scalar for every $\alpha \in A$, $p_\alpha q_\beta = \delta_{\alpha,\beta}\, \mathrm{id}_{i_\alpha}$ for all $\alpha, \beta \in A$, and $\mathrm{id}_X = \sum_{\alpha\in A} q_\alpha p_\alpha$. The morphism SL_X does not depend on the choice of the decomposition of X into scalar objects. Note that

SL_X is invertible with inverse

$$\mathrm{SL}_X^{-1} = \sum_{\alpha \in A} \mathrm{sl}(i_\alpha)^{-1}\, q_\alpha\, p_\alpha .$$

The family $\mathrm{SL} = \{\mathrm{SL}_X : X \to X\}_{X \in \mathscr{C}}$ is a monoidal natural automorphism of the identity functor $1_{\mathscr{C}}$ of $\mathscr{C}$, called the *slope operator* of $\mathscr{C}$. In particular

$$\mathrm{SL}_Y f = f\,\mathrm{SL}_X \quad \text{and} \quad \mathrm{SL}_{X\otimes Y} = \mathrm{SL}_X \otimes \mathrm{SL}_Y$$

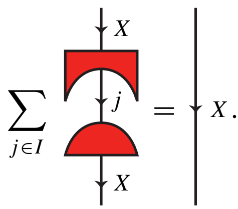
for all objects X, Y of $\mathscr{C}$ and all morphism $f : X \to Y$. The slope operator relates the left and right traces: for any endomorphism f of an object of $\mathscr{C}$,

$$\mathrm{tr}_l(f) = \mathrm{tr}_r(f\,\mathrm{SL}_X). \tag{9}$$

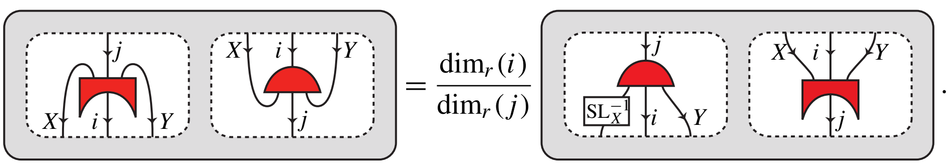
Note that $\mathscr{C}$ is spherical if and only its slope operator is the identity.

Lemma 4.1. *Let I be a representative set of scalar objects of $\mathscr{C}$.*

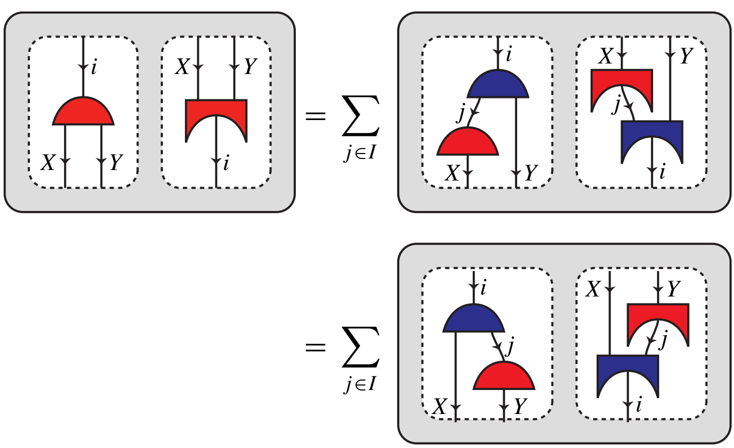
(a) *For any object X of $\mathscr{C}$,*



(b) *For $i, j \in I$ and X, Y objects of $\mathscr{C}$,*

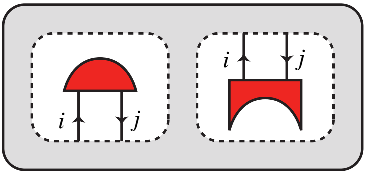


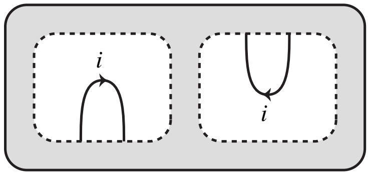
(c) *For $i \in I$ and X, Y objects of $\mathscr{C}$,*



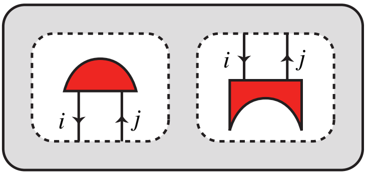
provided there are no j-colored strands in the gray area.

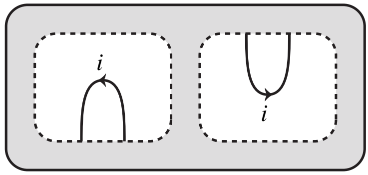
(d) *For all* $i, j \in I$,



$= \dfrac{\delta_{i,j}}{\dim_l(i)}$


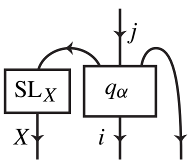
,

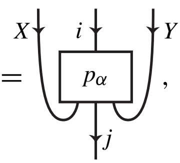


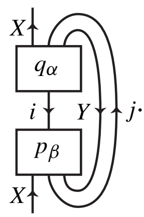
$= \dfrac{\delta_{i,j}}{\dim_r(i)}$


.

Proof. Part (a) follows directly from the definitions. We prove (b). Let $(p_\alpha, q_\alpha)_{\alpha\in A}$ be an i-decomposition of $X^* \otimes j \otimes Y^*$. For $\alpha, \beta \in A$, set

$P_\alpha = \dfrac{\dim_r(j)}{\dim_r(i)}$


, $Q_\alpha =$


, $f_{\alpha,\beta} =$


.

We need to prove that $(P_\alpha, Q_\alpha)_{\alpha\in A}$ is a j-decomposition of $X \otimes i \otimes Y$. Let $\alpha, \beta \in A$. Since $(\mathrm{SL}_X)^* = \mathrm{SL}_{X^*}^{-1}$ and using (9), we obtain

$$\begin{aligned}
P_\alpha Q_\beta &= \frac{\mathrm{tr}_r(P_\alpha Q_\beta)}{\dim_r(j)}\,\mathrm{id}_j = \frac{\mathrm{tr}_l(f_{\alpha,\beta}\mathrm{SL}_{X^*}^{-1})}{\dim_r(i)}\,\mathrm{id}_j = \frac{\mathrm{tr}_r(f_{\alpha,\beta})}{\dim_r(i)}\,\mathrm{id}_j\\
&= \frac{\mathrm{tr}_r(q_\alpha p_\beta)}{\dim_r(i)}\,\mathrm{id}_j = \frac{\mathrm{tr}_r(p_\beta q_\alpha)}{\dim_r(i)}\,\mathrm{id}_j = \frac{\mathrm{tr}_r(\delta_{\alpha,\beta}\,\mathrm{id}_i)}{\dim_r(i)}\,\mathrm{id}_j = \delta_{\alpha,\beta}\,\mathrm{id}_j.
\end{aligned}$$

We conclude using that $\mathrm{card}(A) = \nu_i(X^* \otimes j \otimes Y^*) = \nu_j(X \otimes i \otimes Y)$.

Part (c) reflects the canonical isomorphisms

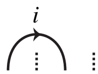
$$\begin{aligned}
\mathrm{Hom}_{\mathscr{C}}(X \otimes Y, i) &\cong \bigoplus_{j\in I} \mathrm{Hom}_{\mathscr{C}}(X, j) \otimes_{\Bbbk} \mathrm{Hom}_{\mathscr{C}}(j \otimes Y, i)\\
&\cong \bigoplus_{j\in I} \mathrm{Hom}_{\mathscr{C}}(X \otimes j, i) \otimes_{\Bbbk} \mathrm{Hom}_{\mathscr{C}}(Y, j),
\end{aligned}$$

and part (d) is a direct consequence of the duality axioms. □

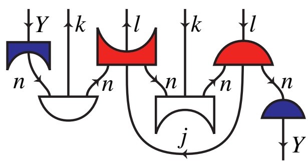
4D. ***Proof of Theorem 2.1 and of the integrality of*** $\boldsymbol{\Lambda}$**.** Recall that $\mathscr{Z}(\mathscr{C})$ is a braided pivotal category which has a coend (C, σ) with $C = \bigoplus_{i,j\in I} i^* \otimes j^* \otimes i \otimes j$. Therefore its dimension is well-defined and

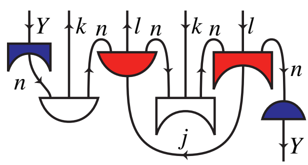
$$
\begin{aligned}
\dim \mathcal{Z}(\mathcal{C}) = \dim_l(C,\sigma) = \dim_l(C) &= \dim_l\Bigl(\sum_{i,j\in I} i^*\otimes j^* \otimes i \otimes j\Bigr)\\
&= \sum_{i,j\in I} \dim_l(i^*)\dim_l(j^*)\dim_l(i)\dim_l(j)\\
&= \Bigl(\sum_{i\in I}\dim_r(i)\dim_l(i)\Bigr)\Bigl(\sum_{j\in I}\dim_r(j)\dim_l(j)\Bigr) = \dim(\mathcal{C})^2.
\end{aligned}
$$

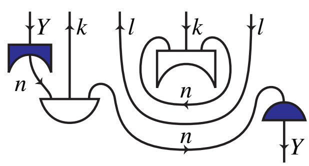
Let us prove that the canonical pairing of the coend (C,σ) is nondegenerate. Define the morphism $\lambda : C \to \mathbb{1}$ as follows and recall the definition of the morphism $\Lambda : \mathbb{1} \to C$ of (6):

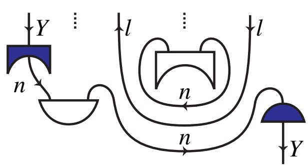
$$\lambda = \sum_{i\in I} \dim_r(i)$$  and $$\Lambda = \sum_{j\in I}\dim_r(j)$$  .

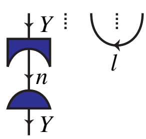
Firstly, Λ is a morphism in $\mathcal{Z}(\mathcal{C})$ from $\mathbb{1}_{\mathcal{Z}(\mathcal{C})} = (\mathbb{1}, \mathrm{id})$ to (C,σ). Indeed, using the description of the half braiding σ given in (4), we obtain that for any object Y of $\mathcal{C}$,

$$\sigma_Y(\Lambda\otimes \mathrm{id}_Y) = \sum_{j,k,\ell,n\in I} \dim_r(j)$$ 

$$= \sum_{j,k,\ell,n\in I} \frac{\dim_r(\ell)}{\mathrm{sl}(n)}$$  by Lemma 4.1(b)

$$= \sum_{k,\ell,n\in I} \frac{\dim_r(\ell)}{\mathrm{sl}(n)}$$  by Lemma 4.1(a)

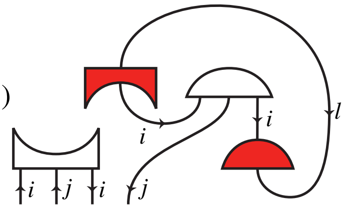
$$= \sum_{\ell,n\in I} \frac{\dim_r(\ell)}{\mathrm{sl}(n)}$$ 

$$= \sum_{\ell,n\in I} \dim_r(\ell)$$  by Lemma 4.1(d)

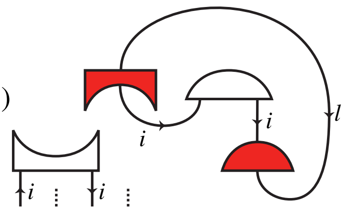
$$= \mathrm{id}_Y \otimes \Lambda$$ by Lemma 4.1(a).

Secondly, λ and Λ satisfy $\omega(\mathrm{id}_C \otimes \Lambda) = \lambda = \omega(\Lambda \otimes \mathrm{id}_C)$. Indeed, using the description of the canonical pairing ω given in item (f) on page 11, we obtain

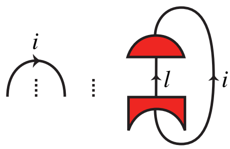
$$\omega(\mathrm{id}_C \otimes \Lambda) = \sum_{i,j,\ell \in I} \dim_r(\ell)$$



$$= \sum_{i,\ell \in I} \dim_r(\ell)$$

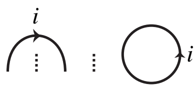


$$= \sum_{i,\ell \in I} \frac{\dim_r(\ell)}{\dim_l(i)}$$

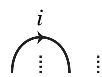


by Lemma 4.1(d)

$$= \sum_{i,\ell \in I} \frac{\dim_r(\ell)}{\dim_l(i)} \delta_{\ell,i^*}$$



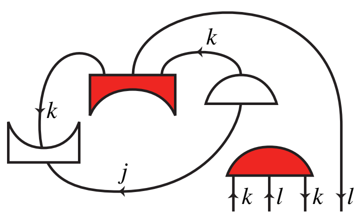
$$= \sum_{i \in I} \dim_r(i)$$



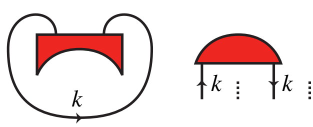
$$= \lambda,$$

and similarly

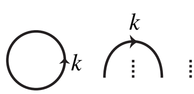
$$\omega(\Lambda \otimes \mathrm{id}_C) = \sum_{j,k,\ell \in I} \dim_r(j)$$



$$= \sum_{j,k \in I} \dim_r(j)\, \delta_{j,k^*}$$

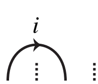


$$= \sum_{k \in I}$$



by Lemma 4.1(d)

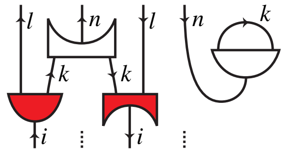
$$= \sum_{k \in I} \dim_r(k)$$

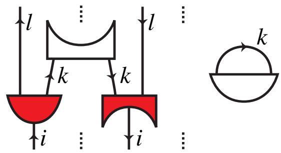


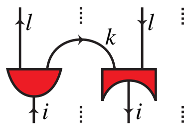
$$= \lambda.$$

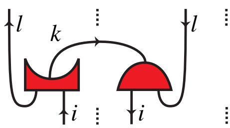
This implies in particular that λ is a morphism in $\mathfrak{Z}(\mathscr{C})$ from (C,σ) to $\mathbb{1}_{\mathfrak{Z}(\mathscr{C})}$, since ω and Λ are morphisms in $\mathfrak{Z}(\mathscr{C})$.

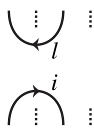
Thirdly, λ is a left cointegral for the Hopf algebra (C, σ) in $\mathcal{Z}(\mathcal{C})$. Indeed, using the description of the coproduct Δ and the unit u — items (a) and (d) on page 10 — we obtain

$$(\mathrm{id}_C \otimes \lambda)\Delta = \sum_{i,k,\ell,n\in I} \dim_r(k)$$


$$= \sum_{i,k,\ell\in I} \dim_r(k)$$


$$= \sum_{i,k,\ell\in I} \dim_r(k)$$

by Lemma 4.1(d)

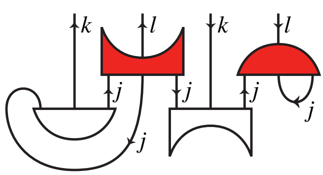
$$= \sum_{i,k,\ell\in I} \dim_r(i)$$

by Lemma 4.1(b)

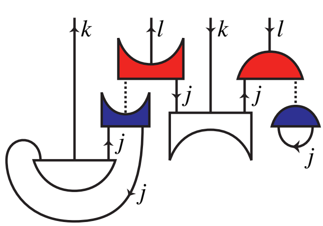
$$= \sum_{i,\ell\in I} \dim_r(i)$$

by Lemma 4.1(a)

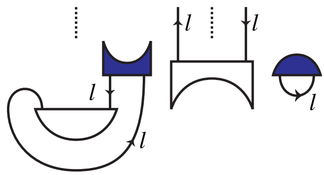
$$= u\,\lambda.$$

Since $\omega(\Lambda \otimes \Lambda) = \lambda\Lambda = \dim_r(\mathbb{1}) = 1 \in \Bbbk$ is invertible, we conclude by Lemma 3.1 that ω is nondegenerate. Hence $\mathcal{Z}(\mathcal{C})$ is modular.

Finally, let us prove that Λ is a two-sided integral of (C, σ) which is invariant under the antipode. The last part of Lemma 3.1 gives that Λ is a right integral of (C, σ). Using the description of the antipode S of (C, σ) in item (e) on page 11, we obtain

$$S\Lambda = \sum_{j,k,\ell\in I} \dim_r(j)$$


$$= \sum_{j,k,\ell\in I} \dim_r(j)$$

by Lemma 4.1(c)

$$= \sum_{\ell \in I} \dim_r(\ell^*)$$ 

$$= \sum_{\ell \in I} \dim_r(\ell)$$  by Lemma 4.1(d)

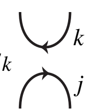
$$= \Lambda.$$

Hence Λ is S-invariant. This implies in particular that Λ, being a right integral, is also a left integral. Hence Λ is an S-invariant (two-sided) integral.

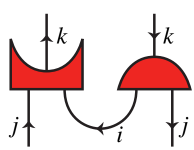
4E. ***Proof of Theorem 2.2.*** Consider the Hopf monad Z of Section 4B. Recall from [Bruguières and Virelizier 2007] that the monad Z is said to be *semisimple* if any Z-module is a Z-linear retract of a free Z-module, that is, of $(Z(X), \mu_X)$ for some object X of $\mathscr{C}$. Since the isomorphism $\Phi : \mathscr{C}^Z \to \mathscr{Z}(\mathscr{C})$ defined in (8) sends the free Z-module $(Z(X), \mu_X)$ to the free half braiding $\Phi(Z(X), \mu_X) = \mathscr{F}(X)$, we need to prove that $\dim(\mathscr{C})$ is invertible if and only if Z is semisimple. Now Theorem 6.5 of [Bruguières and Virelizier 2007] provides an analogue of Maschke's semisimplicity criterion for Hopf monads: the Hopf monad Z is semisimple if and only if there exists a morphism $\alpha : \mathbb{1} \to Z(\mathbb{1})$ in $\mathscr{C}$ such that

$$\mu_{\mathbb{1}} \alpha = \alpha Z_0 \quad \text{and} \quad Z_0 \alpha = 1. \tag{10}$$

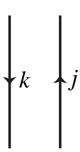
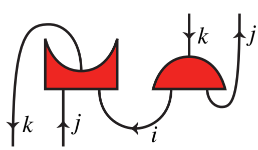
Let $\alpha : \mathbb{1} \to Z(\mathbb{1}) = \bigoplus_{i \in I} i^* \otimes i$ be a morphism in $\mathscr{C}$. Since $\mathscr{C}$ is a fusion category, α decomposes uniquely as $\alpha = \sum_{i \in I} \alpha_i \, \widetilde{\mathrm{coev}}_i$ where $\alpha_i \in \Bbbk$. From the structural morphisms of the Hopf monad Z (page 19), we obtain

$$\alpha Z_0 = \sum_{j,k \in I} \alpha_k$$ 

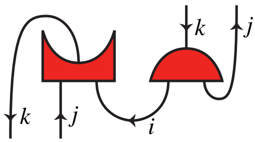
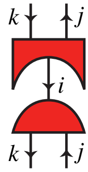
and

$$\mu_{\mathbb{1}} Z(\alpha) = \sum_{i,j,k \in I} \alpha_i \, \frac{\dim_r(k)}{\dim_r(i)}$$  .

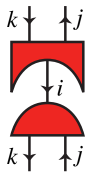
Thus, by duality, $\alpha Z_0 = \mu_{\mathbb{1}} Z(\alpha)$ if and only if

$$\sum_{j,k \in I} \alpha_k$$  $$= \sum_{i,j,k \in I} \alpha_i$$  in $\mathrm{End}_{\mathscr{C}}\Big(\bigoplus_{k,j \in I} k \otimes j^*\Big)$.

Now, for $j, k \in I$, by using Lemma 4.1(b),

$$\sum_{i\in I} \alpha_i$$  $$= \sum_{i\in I} \alpha_i \frac{\dim_r(k)}{\dim_r(i)}$$  .

Therefore $\alpha Z_0 = \mu_{\mathbb{1}} Z(\alpha)$ if and only if

(11) $$\alpha_k \, \mathrm{id}_{k\otimes j^*} = \sum_{i\in I} \alpha_i$$  $\in \mathrm{End}_{\mathscr{C}}(k \otimes j^*)$ for all $k, j \in I$.

In particular, if $\alpha Z_0 = \mu_{\mathbb{1}} Z(\alpha)$, then for any $i \in I$, setting $k = \mathbb{1}$ and $j = i^*$ we obtain $\alpha_i = \alpha_{\mathbb{1}} \dim_r(i)$. Conversely, if $\alpha_i = \alpha_{\mathbb{1}} \dim_r(i)$ for all $i \in I$, then (11) holds by Lemma 4.1(a), and so $\alpha Z_0 = \mu_{\mathbb{1}} Z(\alpha)$. In conclusion, $\alpha Z_0 = \mu_{\mathbb{1}} Z(\alpha)$ if and only if $\alpha = \alpha_{\mathbb{1}} \kappa$, where

$$\kappa = \sum_{i\in I} \dim_r(i) \, \widetilde{\mathrm{coev}}_i : \mathbb{1} \to Z(\mathbb{1}).$$

In that case,

$$Z_0\alpha = \alpha_{\mathbb{1}} Z_0 \kappa = \sum_{i\in I} \dim_r(i) Z_0 \widetilde{\mathrm{coev}}_i = \alpha_{\mathbb{1}} \sum_{i\in I} \dim_r(i) \dim_l(i) = \alpha_{\mathbb{1}} \dim(\mathscr{C}).$$

Hence there exists α satisfying (10) if and only if $\dim(\mathscr{C})$ is invertible in $\Bbbk$. This concludes the proof of Theorem 2.2. □

4F. *Proof of Corollary 2.3.* Let $\mathscr{A}$ be an abelian category. If $\mathscr{A}$ is semisimple (see Section 2), then every object of $\mathscr{A}$ is projective[2]. The converse is true if in addition we assume that all objects of $\mathscr{A}$ have finite length[3].

Assume $\Bbbk$ is a field and let $\mathscr{C}$ be a pivotal fusion category over $\Bbbk$. Then $\mathscr{C}$ is abelian semisimple and its objects have finite length. The center $\mathscr{Z}(\mathscr{C})$ of $\mathscr{C}$ is then an abelian category and the forgetful functor $\mathscr{U} : \mathscr{Z}(\mathscr{C}) \to \mathscr{C}$ is $\Bbbk$-linear, faithful, and exact. This implies that all objects of $\mathscr{Z}(\mathscr{C})$ have finite length and the Hom spaces in $\mathscr{Z}(\mathscr{C})$ are finite-dimensional. As a result, $\mathscr{Z}(\mathscr{C})$ is semisimple if and only if all of its objects are projective.

[2] An object P of $\mathscr{A}$ is *projective* if the functor $\mathrm{Hom}_{\mathscr{A}}(P, -) : \mathscr{A} \to \mathrm{Ab}$ is exact, where Ab is the category of abelian groups.

[3] An object A of $\mathscr{A}$ has *finite length* if there exists a finite sequence of subobjects $A = X_0 \supsetneq X_1 \supsetneq \cdots \supsetneq X_n = \mathbf{0}$ such that each quotient X_i / X_{i+1} is simple.

We identify $\mathcal{Z}(\mathscr{C})$ with the category $\mathscr{C}^Z$ of Z-modules via the isomorphism (8). Recall from the proof of Theorem 2.2 (see the beginning of Section 4E) that the monad Z is semisimple if and only if $\dim(\mathscr{C})$ is invertible in $\Bbbk$. The following lemma relates the notions of semisimplicity for monads and for categories.

Lemma 4.2. *Let $\mathscr{C}$ be an abelian category and T be a right exact monad on $\mathscr{C}$, so that $\mathscr{C}^T$ is abelian and the forgetful functor $U_T : \mathscr{C}^T \to \mathscr{C}$ is exact.*

(a) *If all the T-modules are projective, then T is semisimple.*
(b) *If T is semisimple and all the objects of $\mathscr{C}$ are projective, then all the T-modules are projective.*
(c) *If the objects of $\mathscr{C}$ have finite length, then the same holds in $\mathscr{C}^T$. If in addition $\mathscr{C}$ has finitely many isomorphy classes of simple objects, then so does $\mathscr{C}^T$.*

Proof. Let us prove assertion (a). Denote by $F_T : \mathscr{C} \to \mathscr{C}^T$ the free module functor (see Section 4A). Let (M, r) be a T-module. The action r defines an epimorphism $F_T(M) \to (M, r)$ in $\mathscr{C}^T$. In particular, if (M, r) is projective, it is a retract of $F_T(M)$. Therefore if all the T-modules are projective, the monad T is semisimple.

Let us prove assertion (b). Note that if X is a projective object of $\mathscr{C}$, then $F_T(X)$ is a projective T-module. Indeed,

$$\mathrm{Hom}_{\mathscr{C}^T}(F_T(X), ?) \simeq \mathrm{Hom}_{\mathscr{C}}(X, U_T)$$

by adjunction, and $\mathrm{Hom}_{\mathscr{C}}(X, U_T)$ is an exact functor when X is projective. In particular, if all objects are projective in $\mathscr{C}$ then all free T-modules are projective. If in addition T is semisimple, then any T-module, being a retract of a free T-module, is projective.

Finally, let us prove assertion (c). The first part results from the fact that U_T is faithful exact. Now if S is a simple object of $\mathscr{C}^T$ and Σ is a simple subobject of $U_T(S)$, then by adjunction the inclusion $\Sigma \subset U_T(S)$ defines a nonzero morphism $F_T(\Sigma) \to S$, which is an epimorphism because S is simple. This proves the second part of assertion (c), because under the assumptions made there are finitely many possibilities for Σ, and each $F_T(\Sigma)$ has finitely many simple quotients. □

Assertion (a) of Corollary 2.3 results immediately from the first two assertions of Lemma 4.2.

Let us prove assertion (b). A fusion category over a field is semisimple. Now assume $\Bbbk$ is algebraically closed. By assertion (a), we need to show that if $\mathcal{Z}(\mathscr{C})$ is semisimple, then it is a fusion category. Assume $\mathcal{Z}(\mathscr{C})$ is semisimple. Since $\mathscr{C}$ is fusion, by the third assertion of Lemma 4.2, the category $\mathcal{Z}(\mathscr{C})$ has finitely many classes of simple objects and its objects have finite length. So each object of $\mathcal{Z}(\mathscr{C})$ is a finite direct sum of simple objects. Since the unit object of $\mathcal{Z}(\mathscr{C})$ is scalar and any simple object S of $\mathcal{Z}(\mathscr{C})$ is scalar (because $\mathrm{End}(S)$ is a finite extension of $\Bbbk$), we obtain that $\mathcal{Z}(\mathscr{C})$ is a fusion category. This proves Corollary 2.3. □

References

[Balsam 2010] B. Balsam, "Turaev–Viro invariants as an extended TQFT II", preprint, 2010. arXiv 1010.1222

[Bruguières and Virelizier 2007] A. Bruguières and A. Virelizier, "Hopf monads", *Adv. Math.* **215**:2 (2007), 679–733. MR 2009b:18006 Zbl 1168.18002

[Bruguières and Virelizier 2008] A. Bruguières and A. Virelizier, "Categorical centers and Reshetikhin–Turaev invariants", *Acta Math. Vietnam.* **33**:3 (2008), 255–277. MR 2010h:57017 Zbl 1226.57018

[Bruguières and Virelizier 2012] A. Bruguières and A. Virelizier, "Quantum double of Hopf monads and categorical centers", *Trans. Amer. Math. Soc.* **364**:3 (2012), 1225–1279. MR 2869176 Zbl 06024127

[Bruguières et al. 2011] A. Bruguières, S. Lack, and A. Virelizier, "Hopf monads on monoidal categories", *Adv. Math.* **227**:2 (2011), 745–800. MR 2012h:18003 Zbl 1233.18002

[Etingof et al. 2005] P. Etingof, D. Nikshych, and V. Ostrik, "On fusion categories", *Ann. of Math.* (2) **162**:2 (2005), 581–642. MR 2006m:16051 Zbl 1125.16025

[Joyal and Street 1991] A. Joyal and R. Street, "Tortile Yang–Baxter operators in tensor categories", *J. Pure Appl. Algebra* **71**:1 (1991), 43–51. MR 92e:18006 Zbl 0726.18004

[Lyubashenko 1995] V. V. Lyubashenko, "Invariants of 3-manifolds and projective representations of mapping class groups via quantum groups at roots of unity", *Comm. Math. Phys.* **172**:3 (1995), 467–516. MR 97c:57018 Zbl 0844.57016

[Mac Lane 1998] S. Mac Lane, *Categories for the working mathematician*, 2nd ed., Graduate Texts in Mathematics **5**, Springer, New York, 1998. MR 2001j:18001 Zbl 0906.18001

[Majid 1991] S. Majid, "Representations, duals and quantum doubles of monoidal categories", *Rend. Circ. Mat. Palermo* (2) *Suppl.* **26** (1991), 197–206. MR 93c:18008 Zbl 0762.18005

[Majid 1995] S. Majid, *Foundations of quantum group theory*, Cambridge University Press, Cambridge, 1995. MR 97g:17016 Zbl 0857.17009

[Müger 2003] M. Müger, "From subfactors to categories and topology, II: The quantum double of tensor categories and subfactors", *J. Pure Appl. Algebra* **180**:1-2 (2003), 159–219. MR 2004f:18014 Zbl 1033.18003

[Turaev 1994] V. G. Turaev, *Quantum invariants of knots and* 3*-manifolds*, De Gruyter Studies in Mathematics **18**, de Gruyter, Berlin, 1994. MR 95k:57014 Zbl 0812.57003

[Turaev and Virelizier 2010] V. G. Turaev and A. Virelizier, "On two approaches to 3-dimensional TQFTs", preprint, 2010. arXiv 1006.3501

Received March 21, 2012. Revised August 28, 2012.

ALAIN BRUGUIÈRES
DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ MONTPELLIER II
CASE COURRIER 051
PLACE EUGÈNE BATAILLON
34095 MONTPELLIER CEDEX 5
FRANCE

bruguier@math.univ-montp2.fr

ALEXIS VIRELIZIER
DÉPARTEMENT DE MATHÉMATIQUES
UNIVERSITÉ LILLE 1
59655 VILLENEUVE D'ASCQ
FRANCE

alexis.virelizier@math.univ-lille1.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.31

CONNECTED QUANDLES ASSOCIATED WITH POINTED ABELIAN GROUPS

W. EDWIN CLARK, MOHAMED ELHAMDADI, XIANG-DONG HOU, MASAHICO SAITO AND TIMOTHY YEATMAN

A quandle is a self-distributive algebraic structure that appears in quasi-group and knot theories. For each abelian group A and $c \in A$, we define a quandle $G(A, c)$ on $\mathbb{Z}_3 \times A$. These quandles are generalizations of a class of nonmedial Latin quandles defined by V. M. Galkin, so we call them *Galkin quandles*. Each $G(A, c)$ is connected but not Latin unless A has odd order. $G(A, c)$ is nonmedial unless $3A = 0$. We classify their isomorphism classes in terms of pointed abelian groups and study their various properties. A family of symmetric connected quandles is constructed from Galkin quandles, and some aspects of knot colorings by Galkin quandles are also discussed.

1. Introduction

Sets with certain self-distributive operations called *quandles* have been studied since the 1940s in various areas. They have been studied, for example, as an algebraic system for symmetries [Takasaki 1943], as quasigroups [Galkin 1988], and in relation to modules [Nelson 2003]. The *fundamental quandle* was defined in a manner similar to the fundamental group [Joyce 1982; Matveev 1982], which made quandles an important tool in knot theory. Algebraic homology theories for quandles were defined [Carter et al. 2003b; Fenn et al. 1995] and developed and investigated ([Litherland and Nelson 2003; Mochizuki 2011; Niebrzydowski and Przytycki 2009; 2011; Nosaka 2011], for example), and extensions of quandles by cocycles have been studied [Andruskiewitsch and Graña 2003; Carter et al. 2003a; Eisermann 2007b] and applied to various properties of knots and knotted surfaces (see [Carter et al. 2004] and references therein).

Before algebraic theories of extensions were developed, Galkin [1988] defined a family of quandles that are extensions of the 3-element connected quandle R_3, and we call them *Galkin quandles*. Even though the definition of Galkin quandles is a

M. S. was supported in part by NSF grant DMS 0900671.
MSC2010: 57M25.
Keywords: quandles, pointed abelian groups, knot colorings.

special case of a cocycle extension described in [Andruskiewitsch and Graña 2003], they have curious properties such as the explicit and simple defining formula, close connections to dihedral quandles, and the fact that they appear in the list of small connected quandles.

In this paper, we generalize Galkin's definition and define a family of quandles that are extensions of R_3, characterize their isomorphism classes, and study their properties. The definition is given in Section 3 after a brief review of necessary materials in Section 2. Isomorphism classes are characterized by pointed abelian groups in Section 4. Various algebraic properties of Galkin quandles are investigated in Section 5, and their knot colorings are studied in Section 6.

2. Preliminaries

In this section we briefly review some definitions and examples of quandles. More details can be found, for example, in [Andruskiewitsch and Graña 2003; Carter et al. 2004; Fenn et al. 1995].

A *quandle* X is a set with a binary operation $(a, b) \mapsto a * b$ satisfying the following conditions.

(1)	(Idempotency)	For any $a \in X$, $a * a = a$.
(2)	(Invertibility)	For any $b, c \in X$, there is a unique $a \in X$ such that $a * b = c$.
(3)	(Right self-distributivity)	For any $a, b, c \in X$, we have $(a * b) * c = (a * c) * (b * c)$.

A *quandle homomorphism* between two quandles X, Y is a map $f : X \to Y$ such that $f(x *_X y) = f(x) *_Y f(y)$, where $*_X$ and $*_Y$ denote the quandle operations of X and Y, respectively. A *quandle isomorphism* is a bijective quandle homomorphism, and two quandles are *isomorphic* if there is a quandle isomorphism between them.

Typical examples of quandles include the following.

- Any nonempty set X with the operation $x * y = x$ for any $x, y \in X$ is a quandle called the *trivial* quandle.
- A group $X = G$ with the operation of n-fold conjugation, $a * b = b^{-n}ab^n$, is a quandle.
- Let n be a positive integer. For $a, b \in \mathbb{Z}_n$ (integers modulo n), define

  $$a * b \equiv 2b - a \pmod{n}.$$

  Then $*$ defines a quandle structure called the *dihedral quandle* R_n. This set can be identified with the set of reflections of a regular n-gon with conjugation as the quandle operation.

- Any $\mathbb{Z}[T, T^{-1}]$-module M is a quandle with $a*b = Ta + (1-T)b$ for $a, b \in M$. This is called an *Alexander quandle*. An Alexander quandle is also regarded as a pair (M, T), where M is an abelian group and $T \in \mathrm{Aut}(M)$.

Let X be a quandle. The *right translation* $\mathcal{R}_a : X \to X$ by $a \in X$ is defined by $\mathcal{R}_a(x) = x*a$ for $x \in X$. Similarly, the *left translation* $\mathcal{L}_a$ is defined by $\mathcal{L}_a(x) = a*x$. Then $\mathcal{R}_a$ is a permutation of X by Axiom (2). The subgroup of $\mathrm{Sym}(X)$ generated by the permutations $\mathcal{R}_a$, $a \in X$, is called the *inner automorphism group* of X and is denoted by $\mathrm{Inn}(X)$. We list some definitions of commonly known properties of quandles below.

- A quandle is *connected* if $\mathrm{Inn}(X)$ acts transitively on X.
- A *Latin quandle* is a quandle such that for each $a \in X$, the left translation $\mathcal{L}_a$ is a bijection. That is, the multiplication table of the quandle is a Latin square.
- A quandle is *faithful* if the mapping $a \mapsto \mathcal{R}_a$ is an injection from X to $\mathrm{Inn}(X)$.
- A quandle X is *involutory*, or a *kei*, if the right translations are involutions: $\mathcal{R}_a^2 = \mathrm{id}$ for all $a \in X$.
- The operation $\bar{*}$ on X defined by $a \mathbin{\bar{*}} b = \mathcal{R}_b^{-1}(a)$ is a quandle operation, and $(X, \bar{*})$ is called the *dual* quandle of $(X, *)$. If $(X, \bar{*})$ is isomorphic to $(X, *)$, then $(X, *)$ is called *self-dual*.
- A quandle X is *medial* if $(a*b)*(c*d) = (a*c)*(b*d)$ for all $a, b, c, d \in X$. It is also called *abelian*. It is known and easily seen that every Alexander quandle is medial.

A *coloring* of an oriented knot diagram by a quandle X is a map $\mathcal{C} : \mathcal{A} \to X$ from the set of arcs $\mathcal{A}$ of the diagram to X such that the image of the map satisfies the relation depicted in Figure 1 at each crossing. More details can be found in [Carter et al. 2004; Eisermann 2007a], for example. A coloring that assigns the same element of X for all the arcs is called trivial, and otherwise nontrivial. The number of colorings of a knot diagram by a finite quandle is known to be independent of the choice of a diagram, and hence is a knot invariant. A coloring by a dihedral quandle R_n for a positive integer $n > 1$ is called an n-coloring. If a knot is nontrivially colored by a dihedral quandle R_n for a positive integer $n > 1$, then it is called *n-colorable*. In Figure 2, a nontrivial 3-coloring of the trefoil knot (3_1 in a common notation in a knot table [Cha and Livingston 2011]) is indicated. This is presented

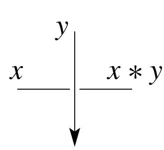


Figure 1. A coloring rule at a crossing.

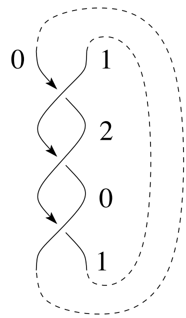


Figure 2. Trefoil as the closure of σ_1^3.

in a closed braid form. Each crossing corresponds to a standard generator σ_1 of the 2-strand braid group, and σ_1^3 represents three crossings together as in the figure. The dotted line indicates the closure; see [Rolfsen 1976] for more details of braids.

The fundamental quandle is defined in a manner similar to the fundamental group [Joyce 1982; Matveev 1982]. A *presentation* of a quandle is defined in a manner similar to groups as well, and a presentation of the fundamental quandle is obtained from a knot diagram (see, for example, [Fenn and Rourke 1992]), by assigning generators to arcs of a knot diagram, and relations corresponding to crossings. The set of colorings of a knot diagram K by a quandle X is then in one-to-one correspondence with the set of quandle homomorphisms from the fundamental quandle of K to X.

3. Definition and notation for Galkin quandles

Let A be an abelian group, also regarded naturally as a $\mathbb{Z}$-module. Let $\mu : \mathbb{Z}_3 \to \mathbb{Z}$, $\tau : \mathbb{Z}_3 \to A$ be functions. These functions μ and τ need not be homomorphisms. Define a binary operation on $\mathbb{Z}_3 \times A$ by

$$(x, a) * (y, b) = \bigl(2y - x, -a + \mu(x - y)b + \tau(x - y)\bigr), \quad x, y \in \mathbb{Z}_3, \ a, b \in A.$$

Proposition 3.1. *For any abelian group A, the operation $*$ defines a quandle structure on $\mathbb{Z}_3 \times A$ if $\mu(0) = 2$, $\mu(1) = \mu(2) = -1$, and $\tau(0) = 0$.*

Galkin [1988, p. 950] gave this definition for $A = \mathbb{Z}_p$. The proposition generalizes his result to any abelian group A. For the proof, we examine the axioms.

Lemma 3.2. (A) *The operation is idempotent — that is, it satisfies Axiom (1) — if and only if $(\mu(0) - 2)a = 0$ for any $a \in A$, and $\tau(0) = 0$.*

(B) *The operation as a right action is invertible — that is, it satisfies Axiom (2).*

Proof. Direct calculations. □

Lemma 3.3. *The operation $*$ on $\mathbb{Z}_3 \times A$ is right self-distributive — that is, it satisfies Axiom (3) — if and only if μ, τ satisfy the following conditions for any $X, Y \in \mathbb{Z}_3$ and $b, c \in A$:*

$$\mu(-X)b = \mu(X)b, \tag{4}$$

$$\big(\mu(X+Y)+\mu(X-Y)\big)c = (\mu(X)\mu(Y))c, \tag{5}$$

$$\tau(X+Y)+\tau(Y-X) = \tau(X)+\tau(-X)+\mu(X)\tau(Y). \tag{6}$$

Proof. Right self-distributivity, that is,

$$((x,a)*(y,b))*(z,c) = ((x,a)*(z,c))*((y,b)*(z,c))$$

for $x, y, z \in \mathbb{Z}_3$ and $a, b, c \in A$, is satisfied if and only if

$$\begin{aligned} \mu(x-y)b &= \mu(y-x)b, \\ \mu(2y-x-z)c &= \big(-\mu(x-z)+\mu(y-x)\mu(y-z)\big)c, \\ -\tau(x-y)+\tau(2y-x-z) &= -\tau(x-z)+\mu(y-x)\tau(y-z)+\tau(y-x). \end{aligned}$$

This is seen by equating the coefficients of b and c and the constant term. For the equivalence of the first equation with (4), set $X = x - y$. For the equivalence of the second with (5), set $X = y - x$ and $Y = z - y$. For the equivalence of the last with (6), set $X = y - x$ and $Y = y - z$. □

Proof of Proposition 3.1. Assume the conditions stated. By Lemma 3.2, Axioms (1) and (2) are satisfied under the specifications $\mu(0) = 2$, $\mu(1) = \mu(2) = -1$, and $\tau(0) = 0$.

If $X = 0$ or $Y = 0$, then (5) (together with (4)) becomes a tautology. If $X - Y = 0$ or $X + Y = 0$, then (5) reduces to $\mu(2X) + 2 = \mu(X)^2$, which is satisfied by the above specifications. For R_3, if $X + Y \neq 0$ and $X - Y \neq 0$, then either $X = 0$ or $Y = 0$. Hence (5) is satisfied. For (6), it is checked similarly, for the two cases [$X = 0$ or $Y = 0$] and [$X - Y = 0$ or $X + Y = 0$]. □

Definition 3.4. Let A be an abelian group. The quandle defined by $*$ on $\mathbb{Z}_3 \times A$ by Proposition 3.1,

$$(x,a)*(y,b) = \big(2y-x,\ -a+\mu(x-y)b+\tau(x-y)\big), \quad x, y \in \mathbb{Z}_3,\ a, b \in A,$$

with $\mu(0) = 2$, $\mu(1) = \mu(2) = -1$, and $\tau(0) = 0$, is called the *Galkin quandle* and denoted by $G(A, \tau)$.

Since τ is specified by the values $\tau(1) = c_1$ and $\tau(2) = c_2$ where $c_1, c_2 \in A$, we also denote it by $G(A, c_1, c_2)$.

Example 3.5. The Galkin quandle $G(\mathbb{Z}_2, 0, 1)$ is $\mathbb{Z}_3 \times \mathbb{Z}_2$ as a set with the quandle operation defined as above with $\mu(0) = 2$, $\mu(1) = \mu(2) = -1$, $\tau(0) = \tau(1) = 0$, and

$\tau(2)=1$. Thus, $(0,1)*(1,0)=(2,-1+\mu(2)0+\tau(2))=(2,0)$ and $(2,0)*(1,1)=(0,0+\mu(1)1+\tau(1))=(0,1)$, for example.

Lemma 3.6. *For any abelian group A and $c_1, c_2 \in A$, the quandles $G(A, c_1, c_2)$ and $G(A, 0, c_2-c_1)$ are isomorphic.*

Proof. Let $c=c_2-c_1$. Define $\eta: G(A, c_1, c_2) \to G(A, 0, c)$, as a map on $\mathbb{Z}_3 \times A$, by $\eta(x,a)=(x, a+\beta(x))$ where $\beta(0)=\beta(1)=0$ and $\beta(2)=-c_1$. This η is a bijection, and we show that it is a quandle homomorphism. We compute $\eta((x,a)*(y,b))$ and $\eta(x,a)*\eta(y,b)$ for $x, y \in \mathbb{Z}_3$ and $a, b \in A$.

If $x=y$, then $\mu(x-y)=2$ and $\tau(x-y)=0$ for both $G(A, c_1, c_2)$ and $G(A, 0, c)$, so that

$$\eta((x,a)*(x,b)) = \eta(x, 2b-a) = (x, 2b-a+\beta(x)),$$

$$\begin{aligned}\eta(x,a)*\eta(x,b) &= (x, a+\beta(x))*(x, b+\beta(x)) = \big(x, 2(b+\beta(x))-(a+\beta(x))\big)\\ &= (x, 2b-a+\beta(x)),\end{aligned}$$

as desired.

If $x-y=1 \in \mathbb{Z}_3$, then $\mu(x-y)=-1$ for both $G(A, c_1, c_2)$ and $G(A, 0, c)$ and $\tau(x-y)=c_1$ for $G(A, c_1, c_2)$ but $\tau(x-y)=0$ for $G(A, 0, c)$, so that

$$\eta((x,a)*(y,b)) = \eta(2y-x, -a-b+c_1) = (2y-x, -a-b+c_1+\beta(2y-x)),$$

$$\begin{aligned}\eta(x,a)*\eta(y,b) &= (x, a+\beta(x))*(y, b+\beta(y))\\ &= \big(2y-x,\ -(a+\beta(x))-(b+\beta(y))\big).\end{aligned}$$

The two expressions are equal if and only if $\beta(x)+\beta(y)+\beta(2y-x)=-c_1$, which is true since $x \neq y$ implies that exactly one of $x, y, 2y-x$ is $2 \in \mathbb{Z}_3$.

If $x-y=2 \in \mathbb{Z}_3$, then $\mu(x-y)=-1$ for both $G(A, c_1, c_2)$ and $G(A, 0, c)$ and $\tau(x-y)=c_2$ for $G(A, c_1, c_2)$ but $\tau(x-y)=c_2-c_1=c$ for $G(A, 0, c)$, so that

$$\eta((x,a)*(y,b)) = \eta(2y-x, -a-b+c_2) = (2y-x, -a-b+c_2+\beta(2y-x)),$$

$$\begin{aligned}\eta(x,a)*\eta(y,b) &= (x, a+\beta(x))*(y, b+\beta(y))\\ &= \big(2y-x, -(a+\beta(x))-(b+\beta(y))+c_2-c_1\big)\\ &= \big(2y-x, -a-b-\beta(x)-\beta(y)+(c_2-c_1)\big),\end{aligned}$$

and again these are equal for the same reason as above. □

Notation. Since, by Lemma 3.6, any Galkin quandle is isomorphic to $G(A, 0, c)$ for an abelian group A and $c \in A$, we denote $G(A, 0, c)$ by $G(A, c)$ for short.

Any finite abelian group is a product $\mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k}$, where the positive integers n_j satisfy $n_j \mid n_{j+1}$ for $j=1, \ldots, k-1$. In this case, any element $c \in A$ is written in a vector form $[c_1, \ldots, c_k]$, where $c_j \in \mathbb{Z}_{n_j}$. Then the corresponding Galkin quandle is denoted by $G(A, [c_1, \ldots, c_k])$.

Remark 3.7. We note that the definition of Galkin quandles induces a functor. Let $\mathbf{Ab}_0$ denote the category of pointed abelian groups; its objects are pairs (A, c), where A is an abelian group and $c \in A$, and its morphisms $f : (A, c) \to (B, d)$ are group homomorphisms $f : A \to B$ such that $f(c) = d$. Let $\mathbf{Q}$ be the category of quandles consisting of quandles as objects and quandle homomorphisms as morphisms.

Then the correspondence $(A, c) \overset{\mathcal{F}}{\mapsto} G(A, c)$ defines a functor $\mathcal{F} : \mathbf{Ab}_0 \to \mathbf{Q}$. It is easy to verify that if a morphism $f : (A, c) \to (B, d)$ is given, then the mapping $\mathcal{F}(f)(x, a) = (x, f(a))$ with $(x, a) \in G(A, c) = \mathbb{Z}_3 \times A$ is a homomorphism from $G(A, c)$ to $G(B, d)$ and satisfies $\mathcal{F}(gf) = \mathcal{F}(g)\mathcal{F}(f)$ and $\mathcal{F}(\mathrm{id}_{(A,c)}) = \mathrm{id}_{G(A,c)}$.

Remark 3.8. A reader will wonder to what extent Definition 3.4 of a Galkin quandle can be generalized. We tried several generalizations. For example, if one attempts to replace 3 by an arbitrary prime p in Definition 3.4, then Lemma 3.3 still holds. In this case for $p > 3$, we prove in Lemma 5.14 that $\mu(x) = 2$ for all $x \in \mathbb{Z}_p$, and computer experiments indicate that one almost always obtains a quandle if and only if $\tau = 0$, in which case the quandle obtained is simply a product of dihedral quandles. We have also attempted to replace $-x + 2y$ by the Alexander quandle operation $tx + (1 - t)y$ in both the left and right coordinates, but have neither been successful in finding interesting new quandles, nor been able to prove that no such generalizations exist. We note that if a generalization for $p > 3$ exists, then any such quandles will be less dense than Galkin quandles, since multiples of 3 are more numerous than multiples of p when $p > 3$.

4. Isomorphism classes

In this section we classify isomorphism classes of Galkin quandles.

Lemma 4.1. *Let A be an abelian group, and let $h : A \to A'$ be a group isomorphism. Then Galkin quandles $G(A, \tau)$ and $G(A', h\tau)$ are isomorphic as quandles.*

Proof. Define $f : G(A, \tau) \to G(A', h\tau)$, as a map from $\mathbb{Z}_3 \times A$ to $\mathbb{Z}_3 \times A'$, by $f(x, a) = (x, h(a))$. This f is a bijection, and we show that it is a quandle homomorphism by computing $f((x, a) * (y, b))$ and $f(x, a) * f(y, b)$ for $x, y \in \mathbb{Z}_3$ and $a, b \in A$:

$$\begin{aligned} f((x, a) * (y, b)) &= f(2y - x, -a + \mu(x - y)b + \tau(x - y)) \\ &= \big(2y - x, h(-a + \mu(x - y)b + \tau(x - y))\big), \\ f(x, a) * f(y, b) &= (x, h(a)) * (y, h(b)) \\ &= \big(2y - x, -h(a) + \mu(x - y)h(b) + h\tau(x - y)\big). \end{aligned}$$

The equality $f((x, a) * (y, b)) = f(x, a) * f(y, b)$ follows from the facts that h is a group homomorphism and $\mu(x - y)$ is an integer. $\square$

Lemma 4.2. *Let c, d, n be positive integers. If* $\gcd(c, n) = d$, *then $G(\mathbb{Z}_n, c)$ is isomorphic to $G(\mathbb{Z}_n, d)$.*

Proof. If $A = \mathbb{Z}_n$, then $\mathrm{Aut}(A) = \mathbb{Z}_n^* =$ units of $\mathbb{Z}_n$, and the divisors of n are representatives of the orbits of $\mathbb{Z}_n^*$ acting on $\mathbb{Z}_n$. □

Thus we may choose the divisors of n for the values of c for representing isomorphism classes of $G(\mathbb{Z}_n, c)$.

Corollary 4.3. *If A is a vector space* (*elementary p-group*), *then there are exactly two isomorphism classes of Galkin quandles $G(A, \tau)$.*

Proof. If A is a vector space containing nonzero vectors c_1 and c_2, then there is a nonsingular linear transformation h of A such that $h(c_1) = c_2$. That $G(A, 0)$ is not isomorphic to $G(A, c)$ if $c \neq 0$ follows from Lemma 4.5. □

For distinguishing isomorphism classes, cycle structures of the right action are useful, and we use the following lemmas.

Lemma 4.4. *For any abelian group A, the Galkin quandle $G(A, \tau)$ is connected.*

Proof. Recall that the operation is defined by the formula

$$(x, a) * (y, b) = \big(2y - x, -a + \mu(x - y)b + \tau(x - y)\big),$$

with $\mu(0) = 2$, $\mu(1) = \mu(2) = -1$, and $\tau(0) = 0$. If $x \neq y$, then $(x, a) * (y, b) = (2y - x, -a - b + c_i) = (z, c)$, where $i = 1$ or 2 and $x, y \in \mathbb{Z}_3$ and $a, b \in A$. Note that $\{x, y, 2y - x\} = \mathbb{Z}_3$ if $x \neq y$. In particular, for any (x, a) and (z, c) with $x \neq z$, there is (y, b) such that $(x, a) * (y, b) = (z, c)$.

For any (x, a_1) and (x, a_2) where $x \in \mathbb{Z}_3$ and $a_1, a_2 \in A$, take $(z, c) \in \mathbb{Z}_3 \times A$ such that $z \neq x$. Then there are (y, b_1), (y, b_2) such that $x \neq y \neq z$ and $(x, a_1) * (y, b_1) = (z, c)$ and $(z, c) * (y, b_2) = (x, a_2)$. Hence $G(A, \tau)$ is connected. □

Lemma 4.5. *The cycle structure of a right translation in $G(A, \tau)$, where $\tau(0) = \tau(1) = 0$ and $\tau(2) = c$, consists of* 1-*cycles*, 2-*cycles*, *and* $2k$-*cycles*, *where k is the order of c in the group A.*

Since isomorphic quandles have the same cycle structure of right translations, $G(A, c)$ and $G(A, c')$ for $c, c' \in A$ are not isomorphic unless the orders of c and c' coincide.

Proof. Let $\tau(0) = 0$, $\tau(1) = 0$, and $\tau(2) = c$. Then by Lemma 4.4, the cycle structure of each column is the same as the cycle structure of the right translation by $(0, 0)$, that is, of the permutation $f(x, a) = (x, a) * (0, 0) = (-x, -a + \tau(x))$.

We show that this permutation has cycles of length only 1, 2 and twice the order of c in A. Since $f(0, a) = (0, -a)$ for $a \in A$, $a \neq 0$, we have $f^2(0, a) = (0, a)$, so that $(0, a)$ generates a 2-cycle, or a 1-cycle if $2a = 0$. Now from $f(1, a) = (2, -a)$ and $f(2, a) = (1, -a + c)$ for $a \in A$, by induction it is easy to see that for $k > 0$,

$f^{2k}(1, a) = (1, a+kc)$ and $f^{2k}(2, a) = (2, a-kc)$. In the case of $(1, a)$, $a \neq 0$, the cycle closes when $a+kc = a$ in A. The smallest k for which this holds is the order of c, in which case the cycle is of length $2k$. A cycle beginning at $(2, a)$ similarly has this same length. □

Proposition 4.6. *Let n be a positive integer. Let $A = \mathbb{Z}_n$ and $c_i, c_i' \in \mathbb{Z}_n$ for $i = 1, 2$. Two Galkin quandles $G(A, c_1, c_2)$ and $G(A, c_1', c_2')$ are isomorphic if and only if* $\gcd(c_1 - c_2, n) = \gcd(c_1' - c_2', n)$.

Proof. If $\gcd(c_1 - c_2, n) = \gcd(c_1' - c_2', n)$, then they are isomorphic by Lemmas 3.6 and 4.2. The cycle structures are different if $\gcd(c_1 - c_2, n) \neq \gcd(c_1' - c_2', n)$ by Lemma 4.5, and hence they are not isomorphic. □

Remark 4.7. The cycle structure is not sufficient for noncyclic groups A. For example, let $A = \mathbb{Z}_2 \times \mathbb{Z}_4$. Then $G(A, [1, 0])$ and $G(A, [0, 2])$ have the same cycle structure for right translations, with cycle lengths $\{2, 2, 4, 4, 4, 4\}$ in a multiset notation, yet they are known not to be isomorphic. (In the notation of Example 4.12 below, $G(A, [1, 0]) = C[24, 29]$ and $G(A, [0, 2]) = C[24, 31]$.) We note that there is no automorphism of A carrying $[1, 0]$ to $[0, 2]$.

More generally, the isomorphism classes of Galkin quandles are characterized as follows.

Theorem 4.8. *Suppose A, A' are finite abelian groups. Two Galkin quandles $G(A, \tau)$ and $G(A', \tau')$ are isomorphic if and only if there exists a group isomorphism $h : A \to A'$ such that $h\tau = \tau'$.*

One implication in the proof of Theorem 4.8 is Lemma 4.1. For the other, first we prove the following two lemmas. We will use a well known description of the automorphisms of a finite abelian group, which can be found in [Hillar and Rhea 2007; Ranum 1907].

Lemma 4.9. *Let A be a finite abelian p-group and let $f : pA \to pA$ be an automorphism. Then f can be extended to an automorphism of A.*

Proof. Let $A = \mathbb{Z}_{p^1}^{n_1} \times \cdots \times \mathbb{Z}_{p^k}^{n_k}$. Then

$$f\left(\begin{bmatrix} px_2 \\ \vdots \\ px_k \end{bmatrix}\right) = P \begin{bmatrix} px_2 \\ \vdots \\ px_k \end{bmatrix}, \qquad \begin{bmatrix} x_2 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{Z}_{p^2}^{n_2} \times \cdots \times \mathbb{Z}_{p^k}^{n_k}, \tag{7}$$

where

$$P = \begin{bmatrix} P_{22} & P_{23} & \cdots & P_{2k} \\ pP_{32} & P_{33} & \cdots & P_{3k} \\ \vdots & \vdots & & \vdots \\ p^{k-2}P_{k2} & p^{k-3}P_{k3} & \cdots & P_{kk} \end{bmatrix}, \tag{8}$$

$P_{ij} \in \mathrm{M}_{n_i \times n_j}(\mathbb{Z})$, $\det P_{ii} \not\equiv 0 \pmod p$. Entries of the vectors are elements of finite groups as specified, and entries of the block matrices are integers. Define $g : A \to A$ by

$$g\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}\right) = \begin{bmatrix} I & \\ & P \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}, \quad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \in \mathbb{Z}_{p^1}^{n_1} \times \mathbb{Z}_{p^2}^{n_2} \times \cdots \times \mathbb{Z}_{p^k}^{n_k}.$$

Then $g \in \mathrm{Aut}(A)$ and $g|_{pA} = f$. □

Lemma 4.10. *Let A be a finite abelian p-group and let $a, b \in A \setminus pA$. If there exists an automorphism $f : pA \to pA$ such that $f(pa) = pb$, then there exists an automorphism $g : A \to A$ such that $g(a) = b$.*

Proof. Let $A = \mathbb{Z}_{p^1}^{n_1} \times \cdots \times \mathbb{Z}_{p^k}^{n_k}$ and let f be defined by (7) and (8). Write

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \quad a_i, b_i \in \mathbb{Z}_{p^i}^{n_i}.$$

Since $f(pa) = pb$, we have

$$p\left(P\begin{bmatrix} a_2 \\ \vdots \\ a_n \end{bmatrix} - \begin{bmatrix} b_2 \\ \vdots \\ b_n \end{bmatrix}\right) = 0,$$

that is,

$$P\begin{bmatrix} a_2 \\ \vdots \\ a_n \end{bmatrix} - \begin{bmatrix} b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} pc_2 \\ \vdots \\ p^{k-1}c_k \end{bmatrix}, \quad c_i \in \mathbb{Z}_{p^i}^{n_i}, \quad 2 \le i \le k. \tag{9}$$

Case 1. Assume that $\begin{bmatrix} a_2 \\ \vdots \\ a_n \end{bmatrix} \in pA$. Then by (9), $\begin{bmatrix} b_2 \\ \vdots \\ b_n \end{bmatrix} \in pA$. So $a_1 \neq 0$ and $b_1 \neq 0$. Then we have

$$\begin{bmatrix} pc_2 \\ \vdots \\ p^{k-1}c_k \end{bmatrix} = \begin{bmatrix} pQ_2 \\ \vdots \\ p^{k-1}Q_k \end{bmatrix} a_1$$

for some $Q_i \in \mathrm{M}_{n_i \times n_1}(\mathbb{Z})$ with $2 \le i \le k$. Also, there exists $P_{11} \in \mathrm{M}_{n_1 \times n_1}(\mathbb{Z})$ such that $\det P_{11} \not\equiv 0 \pmod p$ and $P_{11}a_1 = b_1$. Let $g \in \mathrm{Aut}(A)$ be defined by

$$g\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}\right) = \begin{bmatrix} P_{11} & 0 \\ -pQ_2 & \\ \vdots & P \\ -p^{k-1}Q_k & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}, \quad x_i \in \mathbb{Z}_{p^i}^{n_i}.$$

Then $g(a) = b$.

<u>Case 2.</u> Assume that $\begin{bmatrix} a_2 \\ \vdots \\ a_n \end{bmatrix} \notin pA$. Then there exists $2 \le s \le k$ such that $a_s \notin p\mathbb{Z}_{p^s}^{n_s}$.

Then we have

$$\begin{bmatrix} c_2 \\ \vdots \\ p^{k-2}c_k \end{bmatrix} = \begin{bmatrix} Q_2 \\ \vdots \\ p^{k-2}Q_k \end{bmatrix} a_s$$

for some $Q_i \in \mathrm{M}_{n_i \times n_s}(\mathbb{Z})$ with $2 \le i \le k$. Put

$$Q = \begin{bmatrix} 0 & \cdots & 0 & Q_2 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & p^{k-2}Q_k & 0 & \cdots & 0 \end{bmatrix},$$

where the (i, j) block is of size $n_i \times n_j$ and Q_2 is in the $(1, s)$ block. Then

$$Q \begin{bmatrix} a_2 \\ \vdots \\ a_k \end{bmatrix} = \begin{bmatrix} c_2 \\ \vdots \\ p^{k-2}c_k \end{bmatrix}.$$

Also, there exist $U \in \mathrm{M}_{n_1 \times (n_2 + \cdots + n_k)}(\mathbb{Z})$ such that

$$U \begin{bmatrix} a_2 \\ \vdots \\ a_k \end{bmatrix} = b_1 - a_1.$$

Now define $g \in \mathrm{Aut}(A)$ by

$$g\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}\right) = \begin{bmatrix} I & U \\ 0 & P - pQ \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}, \quad x_i \in \mathbb{Z}_{p^i}^{n_i}.$$

Then $g(a) = b$. $\square$

Proof of Theorem 4.8. We assume that $|3A'| \le |3A|$. Since $G(A', c')$ is connected, there exists an isomorphism $\phi : G(A, c) \to G(A', c')$ such that $\phi(0, 0) = (0, 0)$. Write

$$\phi(x, a) = (\alpha(x, a), \beta(x, a)), \quad (x, a) \in \mathbb{Z}_3 \times A.$$

Define $t : \mathbb{Z}_3 \to \mathbb{Z}$ by

$$t(x) = \begin{cases} 1 & \text{if } x = 2, \\ 0 & \text{if } x \neq 2, \end{cases}$$

so that for $(x, a), (y, b) \in \mathbb{Z}_3 \times A$, the operation on $G(A, c)$ is written by

$$(x, a) * (y, b) = \bigl(-x - y, -a + \mu(x - y)b + t(x - y)c\bigr).$$

Then $\phi((x, a) * (y, b)) = \phi(x, a) * \phi(y, b)$ is equivalent to

$$\alpha\bigl(-x - y, -a + \mu(x - y)b + t(x - y)c\bigr) = -\alpha(x, a) - \alpha(y, b), \tag{10}$$

$$\begin{aligned} &\beta\bigl(-x - y, -a + \mu(x - y)b + t(x - y)c\bigr) \\ &\quad = -\beta(x, a) + \mu\bigl(\alpha(x, a) - \alpha(y, b)\bigr)\beta(y, b) + t\bigl(\alpha(x, a) - \alpha(y, b)\bigr)c'. \end{aligned} \tag{11}$$

Claim 1. *The map $\alpha(0, \cdot\,) : A \to \mathbb{Z}_3$ is a homomorphism.*

Proof. Setting $x = y = 0$ in (10), we have

$$\alpha(0, -a + 2b) = -\alpha(0, a) - \alpha(0, b). \tag{12}$$

Setting $b = 0$ in (12), we have

$$\alpha(0, -a) = -\alpha(0, a). \tag{13}$$

By the symmetry of the right-hand side of (12), we also have

$$\alpha(0, -a + 2b) = \alpha(0, -b + 2a), \quad a, b \in A. \tag{14}$$

Now we have

$$\begin{aligned} \alpha(0, a + b) &= \alpha(0, a - b + 2b) \\ &= \alpha(0, -b + 2(b - a)) && \text{(by (14))} \\ &= \alpha(0, b - 2a) \\ &= -\alpha(0, -b) - \alpha(0, -a) && \text{(by (12))} \\ &= \alpha(0, a) + \alpha(0, b) && \text{(by (13))}. \end{aligned}$$

□

Claim 2. *There exists $u \in \mathbb{Z}_3$ such that*

$$\alpha(x, a) = \alpha(0, a) + ux, \quad (x, a) \in \mathbb{Z}_3 \times A. \tag{15}$$

Proof. Setting $x = 1$ and $y = 0$ in (10), we have

$$\alpha(-1, -a - b) = -\alpha(1, a) - \alpha(0, b). \tag{16}$$

Setting $b = 0$ in (16) gives

$$\alpha(-1, -a) = -\alpha(1, a). \tag{17}$$

Letting $a = 0$ in (16) and using (17), we get

$$\alpha(1, b) = \alpha(0, b) + \alpha(1, 0), \quad b \in A. \tag{18}$$

Equations (16) and (13) also imply that

$$\alpha(-1, -b) = \alpha(0, -b) - \alpha(1, 0), \quad b \in A. \tag{19}$$

Let $u = \alpha(1, 0)$. Then

$$\alpha(x, a) = \alpha(0, a) + ux, \quad (x, a) \in \mathbb{Z}_3 \times A. \qquad \square$$

Claim 3. $\alpha(0, c) = 0.$

Proof. Substituting (15) in (10), we get

$$\alpha\big(0, -a + \mu(x - y)b + t(x - y)c\big) = -\alpha(0, a) - \alpha(0, b). \tag{20}$$

Setting $x - y = 2$, we have $\alpha(0, c) = 0$. $\square$

The rest of the proof of Theorem 4.8 is divided into two cases according to whether u is zero or nonzero in (15).

Case A. Assume $u = 0$ in (15).

We have $\alpha(x, a) = \alpha(0, a)$ for all $(x, a) \in \mathbb{Z}_3 \times A$. We write $\alpha(a)$ for $\alpha(0, a)$. Then (11) becomes

$$\begin{aligned}\beta\big(-x - y, -a + \mu(x - y)b + t(x - y)c\big)& \\ = -\beta(x, a) + \mu\big(\alpha(a - b)\big)&\beta(y, b) + t\big(\alpha(a - b)\big)c'.\end{aligned} \tag{21}$$

Step A-1. We claim that $c = 0$.

Equation (21) with $x = 1$, $y = 0$, $a = b = 0$ yields

$$\beta(-1, 0) = -\beta(1, 0),$$

and with $x = -1$, $y = 0$, $a = b = 0$, it yields

$$\beta(1, c) = -\beta(-1, 0).$$

Thus $\beta(1, c) = \beta(1, 0)$. Since $\alpha(1, c) = 0 = \alpha(1, 0)$, we have $\phi(1, c) = \phi(1, 0)$. Thus $c = 0$.

Step A-2. We claim that $c' = 0$.

The homomorphism $\alpha : A \to \mathbb{Z}_3$ must be onto. (Otherwise ϕ is not onto.) Choose $d \in A$ such that $\alpha(d) = -1$. Equation (21) with $x = y = 0$, $a = d$, $b = 0$ gives

$$\beta(0, -d) = -\beta(0, d) + c',$$

and with $x = y = 0$, $a = -d$, $b = 0$, it gives

$$\beta(0, d) = -\beta(0, -d).$$

Therefore $c' = 0$.

Step A-3. Now (21) becomes

$$\beta\big(-x - y, -a + \mu(x - y)b\big) = -\beta(x, a) + \mu\big(\alpha(a - b)\big)\beta(y, b). \tag{22}$$

Setting $y = 0$ and $b = 0$ in (22), we have

$$\beta(-x, -a) = -\beta(x, a). \tag{23}$$

Step A-4. We claim that $\beta(0, \cdot\,) : 3A \to A'$ is a one-to-one homomorphism.

Note that $3A \subset \ker\alpha$. Let $a, b \in 3A$, and $x = -1$, $y = 1$ in (22). We have

$$\beta(0, -a - b) = -\beta(-1, a) + 2\beta(1, b). \tag{24}$$

Setting $b = 0$ and $a = 0$, respectively, in (24) and using (23), we have

$$\beta(0, -a) = -\beta(-1, a) + 2\beta(1, 0) = \beta(1, -a) + 2\beta(1, 0), \tag{25}$$

$$\beta(0, -b) = -\beta(-1, 0) + 2\beta(1, b) = \beta(1, 0) + 2\beta(1, b). \tag{26}$$

Setting $a = b = 0$ in (24), we have

$$3\beta(1, 0) = 0. \tag{27}$$

Combining (24)–(27), we have

$$\beta(0, -a - b) = \beta(0, -a) + \beta(0, -b).$$

If $a \in 3A$ such that $\beta(0, a) = 0$, then $\phi(0, a) = (0, 0)$, so $a = 0$. Thus

$$\beta(0, \cdot\,) : 3A \to A'$$

is one-to-one.

Step A-5. We claim that $\beta(0, 3b) \in 3A'$ for all $b \in A$.

Let $x = y = 0$ and $a = -b$ in (22). We have

$$\begin{aligned} \beta(0, 3b) &= -\beta(0, -b) + \mu(\alpha(-2b))\beta(0, b) \\ &= \beta(0, b) + \mu(\alpha(b))\beta(0, b) \\ &\equiv 0 \pmod{3A'} \qquad (\text{since } \mu(\alpha(b)) \equiv -1 \pmod 3). \end{aligned}$$

Step A-6. Now $\beta(0,\cdot): 3A \to 3A'$ is a one-to-one homomorphism. It is therefore an isomorphism, since $|3A'| \leq |3A|$. Since $|A| = |A'|$, we have $A \cong A'$. We are done in Case A.

Case B. Assume $u \neq 0$ in (15).

By the proofs of Lemma 4.5 above and Proposition 5.11 below, the map $(x', a') \mapsto (-x', a' - t(-x')c')$ is an isomorphism from $G(A', c')$ to $G(A', -c')$. Thus we may assume $u = 1$ in (15). We have $\alpha(x, a) = \alpha(0, a) + x$ for all $(x, a) \in \mathbb{Z}_3 \times A$.

Step B-1. We claim that $\beta(0,\cdot): \ker\alpha(0,\cdot) \to A'$ is a one-to-one homomorphism.

In (11) let $a, b \in \ker\alpha(0,\cdot)$ and $x = -1$, $y = 1$. We have

$$\beta(0, -a-b) = -\beta(-1, a) - \beta(1, b). \tag{28}$$

Equation (28) with $a = -b$ yields

$$\beta(-1, -b) = -\beta(1, b). \tag{29}$$

So

$$\beta(0, -a-b) = \beta(1, -a) - \beta(1, b). \tag{30}$$

Letting $b = 0$ and $a = 0$ in (30), respectively, we have

$$\begin{aligned}\beta(0, -a) &= \beta(1, -a) - \beta(1, 0),\\ \beta(0, -b) &= \beta(1, 0) - \beta(1, b).\end{aligned}$$

Thus

$$\begin{aligned}\beta(0, -a) + \beta(0, -b) &= \beta(1, -a) - \beta(1, b)\\ &= \beta(0, -a-b) \qquad \text{(by (30))}.\end{aligned}$$

If $a \in \ker\alpha(0,\cdot)$ such that $\beta(0, a) = 0$, then $\phi(0, a) = (0, 0)$, so $a = 0$. Hence $\beta(0,\cdot): \ker\alpha(0,\cdot) \to A'$ is one-to-one.

Step B-2. We claim that $\beta(0, 3a) \in 3A'$ for all $a \in A$.

Setting $x = y = 0$ in (11), we have

$$\begin{aligned}\beta(0, -a+2b) &= -\beta(0, a) + \mu\big(\alpha(0, a-b)\big)\beta(0, b) + t\big(\alpha(0, a-b)\big)c'\\ &\equiv -\beta(0, a) - \beta(0, b) + t\big(\alpha(0, a-b)\big)c' \pmod{3A'}.\end{aligned} \tag{31}$$

By (31),

$$\beta(0, 3a) = \beta\big(0, -a + 2(2a)\big) \equiv -\beta(0, a) - \beta(0, 2a) + t\big(\alpha(0, -a)\big)c' \pmod{3A'}$$

and

$$\beta(0, 2a) = \beta(0, 0 + 2a) \equiv -\beta(0, a) + t(\alpha(0, -a))c' \pmod{3A'}.$$

Thus $\beta(0, 3a) \equiv 0 \pmod{3A'}$.

Step B-3. By the argument in Step A-6, $\beta(0,\cdot):3A\to 3A'$ is an isomorphism and $A\cong A'$.

Step B-4. We claim that $\beta(0,c)=c'$.

Equation (11) with $x=1$, $y=-1$, $a=b=0$ yields

$$\begin{aligned}\beta(0,c)&=-\beta(1,0)-\beta(-1,0)+c'\\&=c'\quad\text{(by (29))}.\end{aligned}$$

Step B-5. Now we complete the proof in Case B. Write $A=A_1\oplus A_2$ and $A'=A'_1\oplus A'_2$, where neither $|A_1|$ nor $|A'_1|$ is a multiple of 3, and $|A_2|$ and $|A'_2|$ are powers of 3. Write $c=c_1+c_2$, where $c_1\in A_1$, $c_2\in A_2$. Then $c_1\in A_1\subset\ker\alpha(0,\cdot)$, so $c_2=c-c_1\in\ker\alpha(0,\cdot)$. Since $\beta(0,\cdot):\ker\alpha(0,\cdot)\to A'$ is a homomorphism, we have

$$c'=\beta(0,c_1)+\beta(0,c_2)=c'_1+c'_2,$$

where $c'_1=\beta(0,c_1)\in A'_1$ and $c'_2=\beta(0,c_2)\in A'_2$. By Step B-3, $\beta(0,\cdot):A_1\to A'_1$ is an isomorphism. So it suffices to show that there is an isomorphism $f:A_2\to A'_2$ such that $f(c_2)=c'_2$.

First assume $c_2\in 3A_2$. Then $c'_2\in 3A'_2$. By Lemma 4.9, the isomorphism $\beta(0,\cdot):3A\to 3A'$ can be extended to an isomorphism $f:A_2\to A'_2$ and we are done.

Now assume that $c_2\in A_2\setminus 3A_2$. We claim that $c_2\in A'_2\setminus 3A'_2$. Assume to the contrary that $c'_2\in 3A'_2$. By Step B-3, there exists $d\in A_2$ such that $\beta(0,3d)=c'_2=\beta(0,c_2)$. By Step B-1, $c_2=3d$, which is a contradiction.

Note that $\beta(0,\cdot):3A_2\to 3A'_2$ is an isomorphism and

$$\begin{aligned}\beta(0,3c_2)&=3\beta(0,c_2)\quad\text{(by Step B-1)}\\&=3c'_2.\end{aligned}$$

By Lemma 4.10, there exists an isomorphism $f:A_2\to A'_2$ such that $f(c_2)=c'_2$. □

Remark 4.11. The numbers of isomorphism classes of order $3n$, from $n=1$ to $n=100$, are as follows: 1, 2, 2, 5, 2, 4, 2, 10, 5, 4, 2, 10, 2, 4, 4, 20, 2, 10, 2, 10, 4, 4, 2, 20, 5, 4, 10, 10, 2, 8, 2, 36, 4, 4, 4, 25, 2, 4, 4, 20, 2, 8, 2, 10, 10, 4, 2, 40, 5, 10, 4, 10, 2, 20, 4, 20, 4, 4, 2, 20, 2, 4, 10, 65, 4, 8, 2,10, 4, 8, 2, 50, 2, 4, 10, 10, 4, 8, 2, 40, 20, 4, 2, 20, 4, 4, 4, 20, 2, 20, 4, 10, 4, 4, 4, 72, 2, 10, 10, 25.

In [Clark and Hou 2013] it is shown that the number $N(n)$ of isomorphism classes of Galkin quandles of order n is multiplicative, that is, if $\gcd(n,m)=1$, then $N(nm)=N(n)N(m)$, so it suffices to find $N(q^n)$ for all prime powers q^n. Clark and Hou established that

$$N(q^n)=\sum_{0\le m\le n}p(m)p(n-m),$$

where $p(m)$ is the number of partitions of the integer m. In particular, $N(q^n)$ is independent of the prime q. The sequence $n \mapsto N(q^n)$ appears in the *On-Line Encyclopedia of Integer Sequences* [Sloane 2011] as sequence A000712.

Example 4.12. In [Vendramin 2011], connected quandles are listed up to order 35. For a positive integer $n > 1$, let $q(n)$ be the number of isomorphism classes of connected quandles of order n. For a positive integer $n > 1$, if $q(n) \neq 0$, then we denote by $C[n, i]$ the i-th quandle of order n in their list $(1 < n \leq 35, i = 1, \ldots, q(n))$. We note that $q(n) = 0$ for $n = 2$, 14, 22, 26, and 34 (for $1 < n \leq 35$). The quandle $C[n, i]$ is denoted by $Q_{n,i}$ in [Vendramin 2012] (and they are left-distributive in that work, so the matrix of $C[n, i]$ is the transpose of the matrix of $Q_{n,i}$). Isomorphism classes of Galkin quandles are identified with those in their list in Table 1.

The 4-digit numbers to the right of each row in Table 1 indicate the numbers of knots that are colored nontrivially by these Galkin quandles, out of total 2977 knots in the table [Cha and Livingston 2011] with 12 crossings or less. See Section 6 for more on this.

5. Properties of Galkin quandles

In this section, we investigate various properties of Galkin quandles.

Lemma 5.1. *The Galkin quandle $G(A, \tau)$ is Latin if and only if $|A|$ is odd.*

Proof. To show that it is Latin if n is odd, first note that R_3 is Latin. Suppose that $(x, a) * (y, b) = (x, a) * (y', b')$. Then we have the equations

$$-x + 2y = -x + 2y', \tag{32}$$

$$-a + \mu(x - y)b + \tau(x - y) = -a + \mu(x - y')b' + \tau(x - y'). \tag{33}$$

From (32) it follows that $y = y'$, and it follows from (33) that $\mu(x-y)b = \mu(x-y)b'$. Now since $|A|$ is odd, the left module action of 2 on A is invertible, and hence $b = b'$. If $|A|$ is even, there is a nonzero element b of order 2, and hence $(0, 0) * (0, b) = (0, 0) * (0, 0)$, so the quandle is not Latin. $\square$

Lemma 5.2. *Any Galkin quandle is faithful.*

Proof. We show that if $(x, a) * (y, b) = (x, a) * (y', b')$ holds for all (x, a), then $(y, b) = (y', b')$. We have $y = y'$ immediately. From the second factor

$$-a + \mu(x - y)b + \tau(x - y) = -a + \mu(x - y)b' + \tau(x - y),$$

we have $\mu(x - y)b = \mu(x - y)b'$ for any x. Pick x such that $x \neq y$; then we have $\mu(x - y) = -1$, and hence $b = b'$. $\square$

Lemma 5.3. *If A' is a subgroup of A and c' is in A', then $G(A', c')$ is a subquandle of $G(A, c')$.*

Proof. Immediate. $\square$

Lemma 5.4. *Any Galkin quandle $G(A, \tau)$ consists of three disjoint subquandles $\{x\} \times A$ for $x \in \mathbb{Z}_3$, and each is a product of dihedral quandles.*

Proof. Immediate. □

We note the following somewhat curious quandles from Lemma 5.4: For a positive integer k, $G(\mathbb{Z}_2^k, [0, \dots, 0])$ is a connected quandle that is a disjoint union of three trivial subquandles of order 2^k.

Lemma 5.5. *The Galkin quandle $G(A, \tau)$ has R_3 as a subquandle if and only if $\tau = 0$ or 3 divides $|A|$.*

Proof. If A is any group and $\tau = 0$, then $(x, 0) * (y, 0) = (2y - x, 0)$ for any $x, y \in \mathbb{Z}_3$, so that $\mathbb{Z}_3 \times \{0\}$ is a subquandle isomorphic to R_3. If 3 divides $|A|$, then A has a subgroup B isomorphic to $\mathbb{Z}_3$. In the subquandle $\{0\} \times B$, we have $(0, a)*(0, b) = (0, -a+2b)$ for $a, b \in B$, so that $\{0\} \times B$ is a subquandle isomorphic to R_3.

Rig notation	Galkin notation	N.C.	Rig notation	Galkin notation	N.C.
$C[6, 1]$	$G(\mathbb{Z}_2, [0])$	1084	$C[24, 28]$	$G(\mathbb{Z}_8, [4])$	1084
$C[6, 2]$	$G(\mathbb{Z}_2, [1])$	1084	$C[24, 29]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_4, [1, 0], [1, 2])$	1084
$C[9, 2]$	$G(\mathbb{Z}_3, [0])$	1084	$C[24, 30]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_4, [0, 0])$	1084
$C[9, 6]$	$G(\mathbb{Z}_3, [1])$	1084	$C[24, 31]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_4, [0, 2])$	1084
$C[12, 5]$	$G(\mathbb{Z}_4, [2])$	1084	$C[24, 32]$	$G(\mathbb{Z}_8, [1])$	1051
$C[12, 6]$	$G(\mathbb{Z}_4, [0])$	1084	$C[24, 33]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_4, [0, 1], [1, 1])$	1051
$C[12, 7]$	$G(\mathbb{Z}_4, [1])$	1051	$C[24, 38]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2, [0, 0, 1])$	1084
$C[12, 8]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_2, [0, 0])$	1084	$C[24, 39]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2, [0, 0, 0])$	1084
$C[12, 9]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_2, [1, 0])$	1084	$C[27, 2]$	$G(\mathbb{Z}_3 \times \mathbb{Z}_3, [0, 0])$	1084
$C[15, 5]$	$G(\mathbb{Z}_5, [1])$	1440	$C[27, 12]$	$G(\mathbb{Z}_9, [3])$	1084
$C[15, 6]$	$G(\mathbb{Z}_5, [0])$	1512	$C[27, 13]$	$G(\mathbb{Z}_9, [0])$	1084
$C[18, 1]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_3, [0, 0])$	1084	$C[27, 23]$	$G(\mathbb{Z}_3 \times \mathbb{Z}_3, [1, 0])$	1084
$C[18, 4]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_3, [1, 0])$	1084	$C[27, 55]$	$G(\mathbb{Z}_9, [1])$	1084
$C[18, 5]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_3, [1, 1])$	1084	$C[30, 12]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_5, [0, 1])$	1440
$C[18, 8]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_3, [0, 1])$	1084	$C[30, 13]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_5, [0, 0])$	1512
$C[21, 7]$	$G(\mathbb{Z}_7, [1])$	1339	$C[30, 14]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_5, [1, 1])$	1440
$C[21, 8]$	$G(\mathbb{Z}_7, [0])$	1386	$C[30, 15]$	$G(\mathbb{Z}_2 \times \mathbb{Z}_5, [1, 0])$	1512
$C[24, 26]$	$G(\mathbb{Z}_8, [2])$	1071	$C[33, 10]$	$G(\mathbb{Z}_{11}, [0])$	1260
$C[24, 27]$	$G(\mathbb{Z}_8, [0])$	1084	$C[33, 11]$	$G(\mathbb{Z}_{11}, [1])$	1220

Table 1. Galkin quandles in the Rig table [Vendramin 2011]. The columns headed N.C. show the number of knots with at most 12 crossings that can be nontrivially colored by the quandle.

Conversely, let $S=\{(x,a),(y,b),(z,d)\}$ be a subquandle of $G(A,c)$ isomorphic to R_3. Note that the quandle operation of R_3 is commutative, and the product of any two distinct elements is equal to the third. We examine two cases.

Case 1. $x=y=z$. In this case we have

$$(x,a)*(x,b)=(x,-a+2b)=(x,d),$$
$$(x,b)*(x,a)=(x,-b+2a)=(x,d).$$

Hence we have $-a+2b=-b+2a$, so that $3(a-b)=0$. If there are no elements of order 3 in A, then we have $a-b=0$, and so $b=a$. This is a contradiction to the fact that S contains 3 elements, so there is an element of order 3 in A; hence 3 divides $|A|$.

Case 2. x, y and z are all distinct (if two are distinct then all three are). In this case consider $S=\{(0,a),(1,b),(2,d)\}$. Now we have

$$(2,d)*(0,a)=(1,-d-a+c)=(1,b),$$
$$(0,a)*(2,d)=(1,-a-d)=(1,b).$$

Hence we have $-d-a+c=-a-d$, so that $c=0$, and we have $\tau=0$. □

Lemma 5.6. *The Galkin quandle $G(A,\tau)$ is left-distributive if and only if $3A=0$, that is, every element of A has order 3.*

Proof. Let $\tau(1)=c_1$, $\tau(2)=c_2$. Let $a=(0,0)$, $b=(0,\alpha)$ and $c=(1,0)$ for $\alpha\in A$. Then we get $a*(b*c)=(1,\alpha-c_2+c_1)$ and $(a*b)*(a*c)=(1,-2\alpha-c_2+c_1)$. If these are equal, then $3\alpha=0$ for any $\alpha\in A$.

Conversely, suppose that every element of A has order 3. Then we have $\mu(x)a=2a$ for any $x\in\mathbb{Z}_3$, $a\in A$. Then one computes

$$(x,a)*[(y,b)*(z,c)]=\big(x*(y*z),-a+b+c-\tau(y-z)+\tau(x-y*z)\big), \tag{34}$$

$$\begin{aligned}&[(x,a)*(y,b)]*[(x,a)*(z,c)]\\&\quad=\big((x*y)*(x*z),-a+b+c-\tau(x-y)-\tau(x-z)+\tau(x*y-x*z)\big).\end{aligned} \tag{35}$$

If all x,y,z are distinct, then $x-y=1$ or $x-y=2$, and $x*y=z$, $x*z=y$, $y*z=x$. If $x-y=1$, then $z=x+1$ and $y-z=1$, $x-z=2$, and one computes that (34) $=(-x+y+z,-a+b+c-c_1)=$ (35). If $x-y=2$, then one computes (34) $=(-x+y+z,-a+b+c-c_2)=$ (35). The other cases for x,y,z are checked similarly. □

Proposition 5.7. *The Galkin quandle $G(A,\tau)$ is Alexander if and only if $3A=0$.*

Proof. If $G(A,\tau)$ is Alexander then it is left-distributive, and hence Lemma 5.6 implies $3A=0$. Conversely, suppose $3A=0$. Then $A=\mathbb{Z}_3^k$ for some positive integer k, and is an elementary 3-group. By Corollary 4.3, there are two isomorphism classes,

$G(\mathbb{Z}_3^k, [0, \dots, 0])$ and $G(\mathbb{Z}_3^k, [0, \dots, 0, 1])$. The quandle $G(\mathbb{Z}_3, 1) = C[9, 6]$ is isomorphic to $\mathbb{Z}_3[t]/(t+1)^2$ by a direct comparison. Hence the two classes are isomorphic to the Alexander quandles R_3^k and $R_3^{k-2} \times \mathbb{Z}_3[t]/(t+1)^2$, respectively. $\square$

Proposition 5.8. *The Galkin quandle $G(A, c)$ is medial if and only if $3A = 0$.*

Proof. We have seen that if $3A = 0$, then $G(A, c)$ is Alexander and hence is medial. Suppose $3b \neq 0$ for some $b \in A$. Then consider the products

$$X = \big((0, 0) * (1, b)\big) * \big((1, 0) * (0, 0)\big) = \big(-1, b - \tau(-1)\big),$$
$$Y = \big((0, 0) * (1, 0)\big) * \big((1, b) * (0, 0)\big) = \big(-1, -\tau(-1) - 2b\big).$$

Since $3b \neq 0$, we have $X \neq Y$ and so $G(A, c)$ is not medial. $\square$

Remark 5.9. The fact that the same condition appeared in Lemma 5.6 and Propositions 5.7 and 5.8 is explained as follows. Alexander quandles are left-distributive and medial. It is easy to check that, for a finite Alexander quandle (M, T) with $T \in \mathrm{Aut}(M)$,

$$(M, T) \text{ is connected} \iff (1-T) \text{ is an automorphism of } M \iff (M, T) \text{ is Latin.}$$

It was also proved by Toyoda [1941] that a Latin quandle is Alexander if and only if it is medial. As noted by Galkin, $G(\mathbb{Z}_5, 0)$ and $G(\mathbb{Z}_5, 1)$ are the smallest nonmedial Latin quandles and hence the smallest non-Alexander Latin quandles.

We note that medial quandles are left-distributive (by idempotency). We show in Theorem 5.10 that any left-distributive connected quandle is Latin. This implies, by Toyoda's theorem, that every medial connected quandle is Alexander and Latin. The smallest Latin quandles that are not left-distributive are the Galkin quandles of order 15.

It is known that the smallest left-distributive Latin quandle that is not Alexander is of order 81. This is due to V. D. Belousov. See, for example, [Pflugfelder 1990; Galkin 1988, Section 5].

Theorem 5.10. *Every finite left-distributive connected quandle is Latin.*

Proof. Let $(X, *)$ be a finite, connected, and left-distributive quandle. For each $a \in X$, let $X_a = \{a * x : x \in X\}$.

Step 1. We claim that $|X_a| = |X_b|$ for all $a, b \in X$. For any $a, y \in X$, we have

$$|X_a| = |X_a * y| = \big|\{(a * x) * y : x \in X\}\big| = \big|\{(a * y) * (x * y) : x \in X\}\big| = |X_{a*y}|.$$

Since X is connected, we have $|X_a| = |X_b|$ for all $a, b \in X$.

Step 2. Fix $a \in X$. If $|X_a| = |X|$, by Step 1, $X_b = X$ for all $b \in X$ and we are done. So assume $|X_a| < |X|$. Clearly, $(X_a, *)$ is a left-distributive quandle. Since

$(X, *)$ is connected and $x \mapsto a * x$ is an onto homomorphism from $(X, *)$ to $(X_a, *)$, $(X_a, *)$ is also connected. Using induction, we may assume that $(X_a, *)$ is Latin.

Step 3. For each $y \in Y$, we claim that $X_{a*y} = X_a$. In fact,

$$\begin{aligned} X_{a*y} &\supset (a * y) * X_a \\ &= X_a \quad \text{(since } X_a \text{ is Latin).} \end{aligned}$$

Since $|X_{a*y}| = |X_a|$, we must have $X_{a*y} = X_a$.

Step 4. Since $(X, *)$ is connected, by Step 3, $X_b = X_a$ for all $b \in X$. Thus $X = \bigcup_{b \in X} X_b = X_a$, which is a contradiction. □

Proposition 5.11. *Any Galkin quandle is self-dual, that is, isomorphic to its dual.*

Proof. The dual quandle structure of $G(A, \tau) = G(A, c_1, c_2)$ is written by

$$(x, a) \mathbin{\bar{*}} (y, b) = \big(x \mathbin{\bar{*}} y, -a + \mu(y - x)b + \tau(y - x)\big)$$

for $(x, a), (y, b) \in G(A, \tau)$. Note that $\mu(x - y) = \mu(y - x)$ and $\tau(y - x) = c_{-i}$ if $\tau(x - y) = c_i$ for any $x, y \in X$ and $i \in \mathbb{Z}_3$. Hence its dual is $G(A, c_2, c_1)$. The isomorphism is $f : \mathbb{Z}_3 \times A \to \mathbb{Z}_3 \times A$, defined by $f(x, a) = (-x, a)$. □

Corollary 5.12. *A Galkin quandle $G(A, c_1, c_2)$ is involutory (kei) if and only if $c_1 = c_2 \in A$.*

Proof. A quandle is a kei if and only if it is the same as its dual, that is, the identity map is an isomorphism between the dual quandle and itself. Hence this follows from Proposition 5.11. □

A *good involution* [Kamada 2007; Kamada and Oshiro 2010] ρ on a quandle $(X, *)$ is an involution $\rho : X \to X$ (a map with $\rho^2 = \mathrm{id}$) such that $x * \rho(y) = x \mathbin{\bar{*}} y$ and $\rho(x * y) = \rho(x) * y$ for any $x, y \in X$. A quandle with a good involution is called a *symmetric* quandle. A kei is a symmetric quandle with $\rho = \mathrm{id}$ (in this case ρ is said to be trivial). Symmetric quandles have been used for unoriented knots and nonorientable surface-knots.

Symmetric quandles with nontrivial good involution have been hard to find. Other than computer calculations, very few constructions have been known. In [Kamada 2007; Kamada and Oshiro 2010], nontrivial good involutions were defined on dihedral quandles of even order, which are not connected. Infinitely many symmetric connected quandles were constructed in [Carter et al. 2010] as extensions of odd order dihedral quandles: For each odd $2n+1$ ($n \in \mathbb{Z}$, $n > 0$), a symmetric connected quandle of order $(2n + 1)2^{2n+1}$ was given that is not a kei. Here we use Galkin quandles to construct more symmetric quandles.

Proposition 5.13. *For any positive integer n, there exists a symmetric connected quandle of order $6n$ that is not involutory.*

Proof. We show that if an abelian group A has an element $c \in A$ of order 2, then $G(A, c)$ is a symmetric quandle. Note that $G(A, c)$ is not involutory by Corollary 5.12.

Define the map $\rho : \mathbb{Z}_3 \times A \to \mathbb{Z}_3 \times A$ by $\rho(x, a) = (x, a + c)$, where $c \in A$ is a fixed element of order 2 and $x \in \mathbb{Z}_3$, $a \in A$. The map ρ is an involution. It satisfies the required conditions, as we show below. For $x, y \in \mathbb{Z}_3$, we have

$$\begin{aligned}(x, a) * \rho(y, b) &= (x, a) * (y, b + c) \\ &= \big(2y - x, -a + \mu(x - y)(b + c) + \tau(x - y)\big), \\ (x, a) \bar{*} (y, b) &= \big(2y - x, -a + \mu(y - x)b + \tau(y - x)\big),\end{aligned}$$

where the last equality follows from the proof of Proposition 5.11. If $x = y$, then $\mu(x - y) = 2 = \mu(y - x)$ and $\tau(x - y) = 0 = \tau(y - x)$, and the above two terms are equal. If $x \neq y$, then $\mu(x - y) = -1 = \mu(y - x)$, and exactly one of $\tau(x - y)$ and $\tau(y - x)$ is c and the other is 0, so that the equality holds.

Next we compute

$$\begin{aligned}\rho\big((x, a) * (y, b)\big) &= \rho\big(2y - x, -a + \mu(x - y)b + \tau(x - y)\big) \\ &= \big(2y - x, -a + \mu(x - y)b + \tau(x - y) + c\big), \\ \rho(x, a) * (y, b) &= (x, a + c) * (y, b) \\ &= \big(2y - x, -a - c + \mu(x - y)b + \tau(x - y)\big),\end{aligned}$$

and these are equal. □

For the equations in Lemma 3.3, we have the following for $\mathbb{Z}_p$.

Lemma 5.14. *Let $p > 3$ be a prime and let $\mu : \mathbb{Z}_p \to \mathbb{Z}$ be a function satisfying $\mu(0) = 2$ and*

$$\mu(x + y) + \mu(x - y) = \mu(x)\mu(y) \tag{36}$$

for any $x, y \in \mathbb{Z}_p$. Then $\mu(x) = 2$ for all $x \in \mathbb{Z}_p$.

Proof. Let

$$S = \sum_{x \in \mathbb{Z}_p} \mu(x).$$

Summing (36) as y runs over $\mathbb{Z}_p$, we have $2S = S\mu(x)$. So if $S \neq 0$, we have $\mu(x) = 2$ for all $x \in \mathbb{Z}_p$. Hence we only need to prove that $S \neq 0$.

Assume to the contrary that $S = 0$. Since $\mu(kx)\mu(x) = \mu((k+1)x) + \mu((k-1)x)$, it is easy to see by induction that

$$\mu(x)^k = \frac{1}{2} \sum_{0 \le i \le k} \binom{k}{i} \mu((k - 2i)x). \tag{37}$$

(Here we also use the fact that $\mu(-x)=\mu(x)$, which follows from the fact that $\mu(x-y)=\mu(x)\mu(y)-\mu(x+y)$ is symmetric in x and y.) In particular,

$$\mu(x)^{2p}=\frac{1}{2}\sum_{0\le i\le 2p}\binom{2p}{i}\mu(2(p-i)x).$$

Since $\sum_{x\in\mathbb{Z}_p}\mu(x)=0$, we have

$$\sum_{x\in\mathbb{Z}_p}\mu(x)^{2p}=\left[2+\binom{2p}{p}\right]p.$$

Since $\mu(x)=\mu\left(\frac{x}{2}\right)^2-2$, we have $\mu(x)=-2,-1,2,7,\ldots$.

<u>Case 1.</u> Assume that there exists $0\neq x\in\mathbb{Z}_p$ such that $\mu(x)\ge 7$. Then

$$\left[2+\binom{2p}{p}\right]p=\sum_{x\in\mathbb{Z}_p}\mu(x)^{2p}\ge 7^{2p},$$

which is not possible.

<u>Case 2.</u> Assume that $\mu(x)\in\{-2,-1,2\}$ for all $x\in\mathbb{Z}_p$. Let $a_i=|\mu^{-1}(i)|$. Since $\sum_{x\in\mathbb{Z}_p}\mu(x)=0$ and $\sum_{x\in\mathbb{Z}_p}\mu(x)^3=0$, where the second equation follows from (37), we have

$$\begin{cases}-2a_{-2}-a_{-1}+2a_2=0,\\ -8a_{-2}-a_{-1}+8a_2=0.\end{cases}$$

So $a_{-1}=0$, that is, $\mu(x)=\pm 2$ for all $x\in\mathbb{Z}_p$. Then

$$\sum_{x\in\mathbb{Z}_p}\mu(x)\equiv 2p\equiv 2 \pmod 4,$$

which is a contradiction. □

6. Knot colorings by Galkin quandles

In this section we investigate knot colorings by Galkin quandles. Recall from Lemma 5.4 that any Galkin quandle $G(A,\tau)$ consists of three disjoint subquandles $\{x\}\times A$ for $x\in\mathbb{Z}_3$, and each is a product of dihedral quandles. Also any Galkin quandle has R_3 as a quotient. Thus we look at relations between colorings by dihedral quandles and those by Galkin quandles. For a positive integer n, a knot is called *n-colorable* if its diagram is colored nontrivially by the dihedral quandle R_n.

First we present the numbers of n-colorable knots (for odd n) with 12 crossings or less out of 2977 knots in the knot table from [Cha and Livingston 2011], for comparison with Table 1. These are for dihedral quandles and their products that may be of interest and relevant for comparisons.

$$R_3 : 1084,\quad R_5 : 670,\quad R_7 : 479,\quad R_{11} : 285,\quad R_{15} : 1512,\quad R_{17} : 192,$$
$$R_{19} : 159,\quad R_{21} : 1386,\quad R_{23} : 128,\quad R_{29} : 97,\quad R_{31} : 87,\quad R_{33} : 1260.$$

Remark 6.1. We note that many Rig Galkin quandles in Table 1 have the same number (1084) of nontrivially colorable knots as the number of 3-colorable knots. We make a few observations on these Galkin quandles.

By Lemma 5.5, a Galkin quandle has R_3 as a subquandle if $\tau = 0$ or 3 divides $|A|$, and among Rig Galkin quandles with the number 1084, 17 of them satisfy this condition. Hence any 3-colorable knot is nontrivially colored by these Galkin quandles. The converse is not necessarily true: $G(\mathbb{Z}_5, 0)$ has $\tau = 0$ but has the number 1512. See Corollary 6.5 for more on these quandles.

The remaining 7 Rig Galkin quandles with the number 1084 have $C[6, 2]$ as a subquandle:

$$C[12, 5],\ C[12, 9],\ C[24, 28],\ C[24, 29],\ C[24, 31],\ C[24, 38].$$

It was conjectured [Carter et al. 2010] that if a knot is 3-colorable, then it is nontrivially colored by $C[6, 2]$ ($\tilde{R}_3$ in their notation). It is also seen that any nontrivial coloring by $C[6, 2]$ descends to a nontrivial 3-coloring via the surjection $C[6, 2] \to R_3$, so if the conjecture is true, then any knot is nontrivially colored by these quandles if and only if it is 3-colorable. See also Remarks 6.6 and 6.7.

The *determinant* of a knot is a well known knot invariant related to n-colorability; see [Fox 1962; Rolfsen 1976] for example, for the definition.

Proposition 6.2. *Let K be a knot with a prime determinant $p > 3$. Then K is nontrivially colored by a finite Galkin quandle $G(A, \tau)$ if and only if p divides $|A|$.*

Proof. By Fox's theorem [1962], for any prime p, a knot is p-colorable if and only if its determinant is divisible by p. Let K be a knot with the determinant that is a prime $p > 3$. Then K is p-colorable and not 3-colorable.

Let $G(A, \tau)$ be any Galkin quandle and let $\mathscr{C} : \mathscr{A} \to G(A, \tau)$ be a coloring, where $\mathscr{A}$ is the set of arcs of a knot diagram of K. By the surjection $r : G(A, \tau) \to R_3$, the coloring $\mathscr{C}$ induces a coloring $r \circ \mathscr{C} : \mathscr{A} \to R_3$. Since K is not 3-colorable, it is a trivial coloring, and therefore $\mathscr{C}(\mathscr{A}) \subset r^{-1}(x)$ for some $x \in R_3$. The subquandle $r^{-1}(x)$ for any $x \in R_3$ is an Alexander quandle $\{x\} \times A$ with the operation

$$(x, a) * (x, b) = (x, 2b - a),$$

so that it is a product of dihedral quandles $\{x\} \times A = R_{q_1} \times \cdots \times R_{q_k}$ for some positive integer k and prime powers q_j, $j = 1, \ldots, k$ (Lemma 5.4). It is known that the number of colorings by a product quandle $X_1 \times \cdots \times X_k$ is the product of numbers of colorings by X_i for $i = 1, \ldots, k$. It is also seen that a knot is nontrivially colored by R_{p^k} for a prime p if and only if it is p-colorable. Hence K is nontrivially colored by $\{x\} \times A$ if and only if one of $q_1, \dots, q_k$ is a power of p. □

A 2-*bridge knot* is a knot that can be put into a position with two maxima and two minima with respect to some height function in space (see [Rolfsen 1976], for example, for its definition and properties).

Corollary 6.3. *For any positive integer n not divisible by 3 and any finite Galkin quandle $G(A, \tau)$, all 2-bridge knots with the determinant n have the same number of colorings by $G(A, \tau)$.*

Proof. Let K be a two-bridge knot with the determinant $n = p_1^{m_1} \dots p_\ell^{m_\ell}$ (in the prime decomposition form), where $p_i \neq 3$ for $i = 1, \dots, \ell$, and let $A = R_{q_1} \times \cdots \times R_{q_k}$ be the decomposition for prime powers, as a quandle. By Fox's theorem [1962], for a prime p, K is p-colorable if and only if p divides the determinant of K. Hence K is p_i-colorable for $i = 1, \dots, \ell$ and not 3-colorable. By the proof of Proposition 6.2, the number of colorings by a Galkin quandle $G(A, \tau)$ of K is determined by the number of colorings by the dihedral quandles R_{q_j} that are factors of A.

The double branched cover $M_2(K)$ of the 3-sphere $\mathbb{S}^3$ along a 2-bridge knot K is a lens space ([Rolfsen 1976], for example), and its first homology group $H_1(M_2(K), \mathbb{Z})$ is cyclic. If the determinant of K is n, then it is isomorphic to $\mathbb{Z}_n$ ([Lickorish 1997], for example). It is known [Przytycki 1998] that the number of colorings by R_{q_j} is equal to the order of the group $\big(\mathbb{Z} \oplus H_1(M_2(K), \mathbb{Z})\big) \otimes \mathbb{Z}_{q_j}$, which is determined by n and q_j alone. □

Example 6.4. Among knots with up to 8 crossings, the following sets of knots have the same numbers of colorings by all finite Galkin quandles from Corollary 6.3: $\{4_1, 5_1\}$ (determinant 5), $\{5_2, 7_1\}$ (7), $\{6_2, 7_2\}$ (11), $\{6_3, 7_3, 8_1\}$ (13), $\{7_5, 8_2, 8_3\}$ (17), $\{7_6, 8_4\}$ (19), $\{8_6, 8_7\}$ (23), $\{8_8, 8_9\}$ (25), $\{8_{12}, 8_{13}\}$ (29). See [Cha and Livingston 2011] for notations of knots in the table. This exhausts such sets of knots up to 8 crossings.

Computer calculations show that the set of knots up to 8 crossings with determinant 9 is $\{6_1, 8_{20}\}$, and these have different numbers of colorings by some Galkin quandles. The determinant was looked up at KnotInfo [Cha and Livingston 2011].

There are two knots (7_4 and 8_{21}, up to 8 crossings) with determinant 15. They can be distinguished by the numbers of colorings by some Galkin quandles, according to computer calculations.

Corollary 6.5. *Let p be an odd prime. Then a knot K is nontrivially colored by the Galkin quandle $G(\mathbb{Z}_p, 0)$ if and only if it is $3p$-colorable.*

Proof. Suppose it is $3p$-colorable; then it is nontrivially colored by R_{3p}, which is isomorphic to $R_3 \times R_p$, so that it is either 3-colorable or p-colorable. If K is 3-colorable, then K is nontrivially colored by $G(Z_p; 0)$, since $G(Z_p; 0)$ has R_3 as a subquandle by Lemma 5.5. If K is p-colorable, then K is nontrivially colored by $G(Z_p; 0)$, since $G(Z_p; 0)$ has $\{0\} \times R_p$ as a subquandle by Lemma 5.4.

Suppose that a knot K is nontrivially colored by $G(\mathbb{Z}_p, 0)$, where p is an odd prime. If K is 3-colorable, then it is $3p$-colorable, and we are done. By the proof of Proposition 6.2, if K is not 3-colorable, then K is nontrivially colored by $\{x\} \times R_p$, where $x \in \mathbb{Z}_3$. Hence K is p-colorable, and so $3p$-colorable. □

Remark 6.6. According to computer calculations, the following sets of Galkin quandles (in the numbering of Table 1) have the same numbers of colorings for all 2977 knots with 12 crossings or less. Thus we conjecture that it is the case for all knots. If a Galkin quandle does not appear in the list, then it means that it has different numbers of colorings for some knots, compared to other Galkin quandles. The numbers of colorings are distinct for distinct sets listed below as well.

$$\begin{aligned}
&\big\{C[6, 1], C[6, 2]\big\}, \big\{C[12, 5], C[12, 6]\big\}, \big\{C[12, 8], C[12, 9]\big\},\\
&\big\{C[18, 1], C[18, 4]\big\}, \big\{C[18, 5], C[18, 8]\big\}, \big\{C[24, 27], C[24, 28]\big\},\\
&\big\{C[24, 29], C[24, 30], C[24, 31]\big\}, \big\{C[24, 38], C[24, 39]\big\},\\
&\big\{C[30, 12], C[30, 14]\big\}, \big\{C[30, 13], C[30, 15]\big\}.
\end{aligned}$$

We wish to acknowledge the use of the programs GAP [2008], Maple15 (Magma package) [Maplesoft 2011], and Prover9 and Mace4 [McCune 2009] in our computations. Computational results are posted at [Clark and Yeatman 2011].

Remark 6.7. In contrast to the preceding remark, if we relax the requirement of coloring the same number of times, and instead consider two quandles equivalent if each colors the same knots nontrivially (among these 2977 knots), then we get the following 4 equivalence classes:

$$\begin{aligned}
&\big\{C[3,1], C[6,1], C[6,2], C[9,2], C[9,6], C[12,5], C[12,6], C[12,8], C[12,9],\\
&C[18,1], C[18,4], C[18,5], C[18,8], C[24,27], C[24,28], C[24,29], C[24,30],\\
&C[24,31], C[24,38], C[24,39], C[27,2], C[27,12], C[27,13], C[27,23], C[27,55]\big\},\\
&\big\{C[12,7], C[24,32], C[24,33]\big\},\\
&\big\{C[15,5], C[30,12], C[30,14]\big\},\\
&\big\{C[15,6], C[30,13], C[30,15]\big\}.
\end{aligned}$$

Thus we conjecture that it is the case for all knots. Of these, the first family with many elements consists of quandles with $C[3, 1]$, $C[6, 1]$ or $C[6, 2]$ as a subquandle. Hence, in fact, the conjecture about this family follows from the conjecture about $\{C[6, 1], C[6, 2]\}$ in the preceding remark.

Remark 6.8. Also in contrast to Remark 6.6, there exists a virtual knot K (see, for example, [Kauffman 1999]) such that the numbers of colorings by $C[6, 1]$ and

$C[6, 2]$ are distinct. A virtual knot K with the following property was given in [Carter et al. 2010, Remark 4.6]: K is 3-colorable, but does not have a nontrivial coloring by $C[6, 2]$. Since $C[6, 1]$ has R_3 as a subquandle, this virtual knot K has a nontrivial coloring by $C[6, 1]$. Hence the numbers of colorings by $C[6, 1]$ and $C[6, 2]$ are distinct for K. Thus we might conjecture that for any pair of nonisomorphic Galkin quandles, there is a virtual knot with different numbers of colorings.

Remark 6.9. For any finite Galkin quandle $G(A, \tau)$, there is a knot K with a surjection $\pi_Q(K) \to G(A, \tau)$ from the fundamental quandle $\pi_Q(K)$. In fact, a connected sum of trefoils can be taken as K as follows (see, for example, [Rolfsen 1976] for connected sum).

First we take a set of generators of $G(A, \tau)$ as follows. Let $A = \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k}$, where $k, n_1, \ldots, n_k$ are positive integers such that n_i divides n_{i+1} for $i = 1, \ldots, k$. Let $S = \{(x, e_i) \mid x \in \mathbb{Z}_3, i = 0, \ldots, k\}$, where $e_0 = 0 \in A$ and $e_i \in A$ $(i = 1, \ldots, k)$ is an elementary vector $[0, \ldots, 0, 1, 0, \ldots, 0] \in \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k}$ with a single 1 at the i-th position. Note that R_n is generated by 0, 1 as $0 * 1 = 2$, $1 * 2 = 3$, and inductively, $i * (i + 1) = i + 2$ for $i = 0, \ldots, n - 2$. Since $\{x\} \times A$ is isomorphic to a product of dihedral quandles for each $x \in \mathbb{Z}_3$, S generates $G(A, \tau)$.

For a 2-string braid σ_1^3 whose closure is trefoil (see Figure 2), we note that if $x \neq y \in \mathbb{Z}_3$, then for any $a, b \in A$, the pair of colors $(x, a), (y, b) \in G(A, \tau)$ at top arcs extends to the bottom, that is, the bottom arcs receive the same pair. This can be computed directly.

For the copies of the trefoil, we assign pairs $[(0, e_0), (x, e_i)]$ as colors where $x = 1, 2$ and $i = 0, \ldots k$, and take connected sums on the portion of the arcs with the common color $(0, e_0)$. Further we take pairs $[(0, e_j), (1, e_0)]$ for $j = 1, \ldots, k$, for example, and take connected sums on the arcs with the common color $(1, e_0)$, to obtain a connected sum of trefoils with all elements of S used as colors, as indicated in Figure 3. Such a coloring gives rise to a quandle homomorphism $\pi_Q(K) \to G(A, \tau)$ whose image contains generators S; hence it defines a surjective homomorphism.

Acknowledgement

Special thanks to Michael Kinyon for bringing [Galkin 1988] to our attention and pointing out the construction of nonmedial, Latin quandles on page 950 of [Galkin 1988] that we call here Galkin quandles $G(\mathbb{Z}_p, c_1, c_2)$. We are also grateful to Professor Kinyon for helping us with using Mace4 for colorings of knots by quandles, and for telling us about Belousov's work on distributive quasigroups. Thanks to David Stanovsky for useful discussions on these matters. We are grateful to James McCarron for his help with the Magma package in Maple 15, especially with isomorphism testing. We also thank the referees for valuable comments.

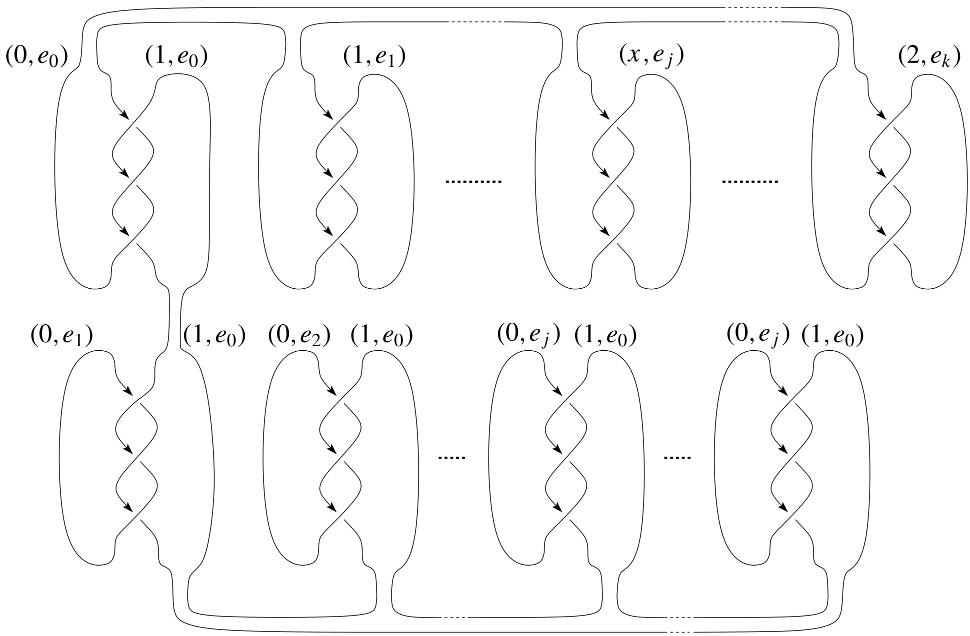


Figure 3. A coloring of a connected sum of trefoils.

References

[Andruskiewitsch and Graña 2003] N. Andruskiewitsch and M. Graña, "From racks to pointed Hopf algebras", *Adv. Math.* **178**:2 (2003), 177–243. MR 2004i:16046 Zbl 1032.16028

[Carter et al. 2003a] J. S. Carter, M. Elhamdadi, M. A. Nikiforou, and M. Saito, "Extensions of quandles and cocycle knot invariants", *J. Knot Theory Ramifications* **12**:6 (2003), 725–738. MR 2004g:57020 Zbl 1049.57008

[Carter et al. 2003b] J. S. Carter, D. Jelsovsky, S. Kamada, L. Langford, and M. Saito, "Quandle cohomology and state-sum invariants of knotted curves and surfaces", *Trans. Amer. Math. Soc.* **355**:10 (2003), 3947–3989. MR 2005b:57048 Zbl 1028.57003

[Carter et al. 2004] J. S. Carter, S. Kamada, and M. Saito, *Surfaces in* 4*-space*, Encyclopædia of Mathematical Sciences/Low-Dimensional Topology **142/III**, Springer, Berlin, 2004. MR 2005e:57065 Zbl 1078.57001

[Carter et al. 2010] J. S. Carter, K. Oshiro, and M. Saito, "Symmetric extensions of dihedral quandles and triple points of non-orientable surfaces", *Topology Appl.* **157**:5 (2010), 857–869. MR 2011c:57034 Zbl 05678844

[Cha and Livingston 2011] J. C. Cha and C. Livingston, "KnotInfo: table of knot invariants", 2011, http://www.indiana.edu/~knotinfo.

[Clark and Hou 2013] W. E. Clark and X.-D. Hou, "Galkin quandles, pointed abelian groups, and sequence *A*000712", *Electron. J. Combin.* **20**:1 (2013), P45. Zbl 06171884

[Clark and Yeatman 2011] W. E. Clark and T. Yeatman, "Small connected quandles and their knot colorings", web site, 2011, http://shell.cas.usf.edu/~saito/QuandleColor.

[Eisermann 2007a] M. Eisermann, "Knot colouring polynomials", *Pacific J. Math.* **231**:2 (2007), 305–336. MR 2008j:57014 Zbl 1152.57010

[Eisermann 2007b] M. Eisermann, "Quandle coverings and their Galois correspondence", preprint, 2007. arXiv math/0612459

[Fenn and Rourke 1992] R. Fenn and C. Rourke, "Racks and links in codimension two", *J. Knot Theory Ramifications* **1**:4 (1992), 343–406. MR 94e:57006 Zbl 0787.57003

[Fenn et al. 1995] R. Fenn, C. Rourke, and B. Sanderson, "Trunks and classifying spaces", *Appl. Categ. Structures* **3**:4 (1995), 321–356. MR 96i:57023 Zbl 0853.55021

[Fox 1962] R. H. Fox, "A quick trip through knot theory", pp. 120–167 in *Topology of* 3*-manifolds and related topics* (Athens, GA, 1961), edited by M. K. Fort, Jr., Prentice-Hall, Englewood Cliffs, NJ, 1962. MR 25 #3522 Zbl 06075327

[Galkin 1988] V. M. Galkin, "Quasigroups", *Itogi Nauki i Tekhniki. Ser. Algebra. Topol. Geom.* **26** (1988), 3–44. In Russian; translated in *J. Soviet Math.* **49**:3 (1990), 941–967. MR 89k:20103 Zbl 0675.20057

[GAP 2008] The GAP Group, "Groups, algorithms, and programming", Version 4.4.12, 2008, http://www.gap-system.org.

[Hillar and Rhea 2007] C. J. Hillar and D. L. Rhea, "Automorphisms of finite abelian groups", *Amer. Math. Monthly* **114**:10 (2007), 917–923. MR 2363058 Zbl 1156.20046

[Joyce 1982] D. Joyce, "A classifying invariant of knots, the knot quandle", *J. Pure Appl. Algebra* **23**:1 (1982), 37–65. MR 83m:57007 Zbl 0474.57003

[Kamada 2007] S. Kamada, "Quandles with good involutions, their homologies and knot invariants", pp. 101–108 in *Intelligence of low dimensional topology 2006* (Hiroshima, 2006), edited by J. S. Carter et al., Ser. Knots Everything **40**, World Scientific, Hackensack, NJ, 2007. MR 2009a:57042 Zbl 0674.15021

[Kamada and Oshiro 2010] S. Kamada and K. Oshiro, "Homology groups of symmetric quandles and cocycle invariants of links and surface-links", *Trans. Amer. Math. Soc.* **362**:10 (2010), 5501–5527. MR 2011f:57017 Zbl 1220.57016

[Kauffman 1999] L. H. Kauffman, "Virtual knot theory", *European J. Combin.* **20**:7 (1999), 663–690. MR 2000i:57011 Zbl 0938.57006

[Lickorish 1997] W. B. R. Lickorish, *An introduction to knot theory*, Graduate Texts in Mathematics **175**, Springer, New York, 1997. MR 98f:57015 Zbl 0886.57001

[Litherland and Nelson 2003] R. A. Litherland and S. Nelson, "The Betti numbers of some finite racks", *J. Pure Appl. Algebra* **178**:2 (2003), 187–202. MR 2004a:57006 Zbl 1017.55014

[Maplesoft 2011] Maplesoft, "Magma package in Maple 15", 2011, http://www.maplesoft.com/support/help/Maple/view.aspx?path=Magma.

[Matveev 1982] S. V. Matveev, "Distributive groupoids in knot theory", *Mat. Sb.* (*N.S.*) **119(161)**:1(9) (1982), 78–88. In Russian; translated in *Math. USSR Sb.* **47**:1 (1984), 73–83. MR 84e:57008 Zbl 0523.57006

[McCune 2009] W. McCune, "Prover9 and Mace4", 2009, http://www.cs.unm.edu/~mccune/prover9.

[Mochizuki 2011] T. Mochizuki, "The third cohomology groups of dihedral quandles", *J. Knot Theory Ramifications* **20**:7 (2011), 1041–1057. MR 2012e:18028 Zbl 1226.57010

[Nelson 2003] S. Nelson, "Classification of finite Alexander quandles", *Topology Proc.* **27**:1 (2003), 245–258. MR 2005b:57027 Zbl 1066.57019

[Niebrzydowski and Przytycki 2009] M. Niebrzydowski and J. H. Przytycki, "Homology of dihedral quandles", *J. Pure Appl. Algebra* **213**:5 (2009), 742–755. MR 2010a:18014 Zbl 0302.35002

[Niebrzydowski and Przytycki 2011] M. Niebrzydowski and J. H. Przytycki, "The second quandle homology of the Takasaki quandle of an odd abelian group is an exterior square of the group", *J. Knot Theory Ramifications* **20**:1 (2011), 171–177. MR 2012g:55009 Zbl 05872656

[Nosaka 2011] T. Nosaka, "On homotopy groups of quandle spaces and the quandle homotopy invariant of links", *Topology Appl.* **158**:8 (2011), 996–1011. MR 2012i:57024 Zbl 1227.57020

[Pflugfelder 1990] H. O. Pflugfelder, *Quasigroups and loops: introduction*, Sigma Series in Pure Mathematics **7**, Heldermann Verlag, Berlin, 1990. MR 93g:20132 Zbl 0715.20043

[Przytycki 1998] J. H. Przytycki, "3-coloring and other elementary invariants of knots", pp. 275–295 in *Knot theory* (Warsaw, 1995), edited by V. F. R. Jones et al., Banach Center Publ. **42**, Polish Acad. Sci., Warsaw, 1998. MR 1634462 Zbl 0904.57002 arXiv math/0608172

[Ranum 1907] A. Ranum, "The group of classes of congruent matrices with application to the group of isomorphisms of any abelian group", *Trans. Amer. Math. Soc.* **8**:1 (1907), 71–91. MR 1500775 JFM 38.0178.01

[Rolfsen 1976] D. Rolfsen, *Knots and links*, Mathematics Lecture Series **7**, Publish or Perish, Berkeley, CA, 1976. MR 58 #24236 Zbl 0339.55004

[Sloane 2011] N. J. A. Sloane, "A000712: Number of permutations of *n* into parts of 2 kinds", entry A000712 in *The On-Line Encyclopedia of Integer Sequences* (http://oeis.org), 2011.

[Takasaki 1943] M. Takasaki, "Abstraction of symmetric transformations", *Tôhoku Math. J.* **49** (1943), 145–207. In Japanese. MR 9,8c Zbl 0061.02109

[Toyoda 1941] K. Toyoda, "On axioms of linear functions", *Proc. Imp. Acad. Tokyo* **17**:7 (1941), 221–227. MR 7,241g Zbl 0061.02403

[Vendramin 2011] L. Vendramin, "Rig: a GAP package for racks and quandles", 2011, http://code.google.com/p/rig.

[Vendramin 2012] L. Vendramin, "On the classification of quandles of low order", *J. Knot Theory Ramifications* **21**:9 (2012), Art. ID #1250088. Zbl 06044378 arXiv 1105.5341

Received April 13, 2012. Revised July 13, 2012.

W. EDWIN CLARK
wclark@mail.usf.edu

MOHAMED ELHAMDADI
emohamed@usf.edu

XIANG-DONG HOU
xhou@usf.edu

MASAHICO SAITO
saito@usf.edu

TIMOTHY YEATMAN
tyeatma2@mail.usf.edu

(all authors)

DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF SOUTH FLORIDA
TAMPA, FL 33620-5700
UNITED STATES

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.61

ENTROPY AND LOWEST EIGENVALUE ON EVOLVING MANIFOLDS

HONGXIN GUO, ROBERT PHILIPOWSKI AND ANTON THALMAIER

We determine the first two derivatives of the classical Boltzmann–Shannon entropy of the conjugate heat equation on general evolving manifolds. Based on the second derivative of the Boltzmann–Shannon entropy, we construct Perelman's $\mathcal{F}$ and $\mathcal{W}$ entropy in abstract geometric flows. Monotonicity of the entropies holds when a technical condition is satisfied.

This condition is satisfied on static Riemannian manifolds with nonnegative Ricci curvature, for Hamilton's Ricci flow, List's extended Ricci flow, Müller's Ricci flow coupled with harmonic map flow and Lorentzian mean curvature flow when the ambient space has nonnegative sectional curvature.

Under the extra assumption that the lowest eigenvalue is differentiable along time, we derive an explicit formula for the evolution of the lowest eigenvalue of the Laplace–Beltrami operator with potential in the abstract setting.

1. Introduction

Geometric flows have been studied extensively. The idea is to evolve metrics in certain ways usually by heat-type equations to obtain better metrics on manifolds and thus to gain topological information of the manifolds. It is desirable to derive evolution equations in a general setting such that the formulas may be applied to various flows. For instance, very nice general approaches to get monotone quantities on evolving manifolds have been developed in [Ecker et al. 2008; Müller 2010].

We briefly introduce notation for an abstract geometric flow. Let M be an n-dimensional compact manifold. Assume that $\alpha(t, y)$ is a time-dependent symmetric two-tensor on M, and that $g(t)$ is a family of one parameter Riemannian metrics evolving along the flow equation

$$\frac{\partial g}{\partial t} = -2\alpha, \quad t \in (0, T), \tag{1-1}$$

Research supported by NSF of China (grants no. 11001203 and 11171143), Zhejiang Provincial Natural Science Foundation of China (project no. LY13A010009) and Fonds National de la Recherche, Luxembourg.
MSC2010: 53C44.
Keywords: Ricci flow, conjugate heat equation, entropy, eigenvalue.

where T is some fixed positive constant. Let $A := g^{ij}\alpha_{ij}$ be the trace of α with respect to $g(t)$.

Classical quantities on static manifolds have nice applications on evolving manifolds by certain natural modifications. The Boltzmann–Shannon entropy is such a quantity for the heat equation. Formally, the conjugate of the heat operator $\partial/\partial t - \Delta$ on space-time is $-\partial/\partial t - \Delta + A$. As Perelman [2002] showed, on evolving manifolds it is natural to work with the entropy for the conjugate heat equation. We will derive the first two derivatives of Boltzmann–Shannon entropy for the conjugate heat equation, and based on that we define Perelman's $\mathcal{F}$ and $\mathcal{W}$ entropy in the framework of abstract geometric flows.

Other classical quantities on static Riemannian manifolds are the eigenvalues of the Laplace–Beltrami operator Δ. When the metric evolves, it is natural to include a potential function. Perelman [2002] shows that the lowest eigenvalue of $-\Delta + R/4$ is monotone nondecreasing along the Ricci flow. Furthermore by deriving explicit formula of the derivative, Cao [2007; 2008] shows that the monotonicity holds for the lowest eigenvalue of $-\Delta + cR$ for any $c \geq \frac{1}{4}$; see also [Li 2007].

Reto Müller [2010] derived formulas for the reduced volume in abstract geometric flows. His formulation is very general and thus can be applied to different flows. He shows that the reduced volume is monotone when a technical assumption holds, which is satisfied for static manifolds with positive Ricci curvature, Hamilton's Ricci flow, List's extended Ricci flow, Müller's Ricci flow coupled with harmonic map flow and Lorentzian mean curvature flow when the ambient manifold has nonnegative sectional curvature. This allows him to establish new monotonicity formulas for these flows.

One of our purposes in this paper is to show that the same phenomena as for reduced volume holds for entropy and eigenvalues.

Notation and main results. Throughout the paper, M will be a compact manifold without boundary. Along the flow equation (1-1) the Riemannian volume dy of M evolves by

$$\frac{\partial}{\partial t}\,dy = -A\,dy$$

and A satisfies

$$\frac{\partial A}{\partial t} = 2|\alpha|^2 + g^{ij}\frac{\partial \alpha_{ij}}{\partial t},$$

where $|\alpha|^2 = g^{ij}g^{kl}\alpha_{ik}\alpha_{jl}$. To simplify the notation, we let $\beta_{ij} := \partial\alpha_{ij}/\partial t$ and $B := g^{ij}\beta_{ij}$, so that

$$\frac{\partial A}{\partial t} = 2|\alpha|^2 + B. \tag{1-2}$$

In particular, $A = R$ and $B = \Delta R$ under the Ricci flow.

For any time-dependent vector field V on M we define

$$\Theta_{g,\alpha}(V) := (\mathrm{Rc}-\alpha)(V,V) + \langle \nabla A - 2\,\mathrm{Div}(\alpha), V\rangle + \tfrac{1}{2}(B-\Delta A), \tag{1-3}$$

where Rc is the Ricci tensor and Div the divergence operator: $\mathrm{Div}(\alpha)_k = g^{ij}\nabla_i \alpha_{jk}$. In the rest of this paper we omit the subscripts of $\Theta_{g,\alpha}(V)$ and denote it by $\Theta(V)$.

The quantity $\Theta(V)$ appears as an error term in our main results. In the expression of $\Theta(V)$, the Rc term is caused by the Bochner's formula. This explains technically why our results are particularly useful for the Ricci flow and its various modifications. Müller [2010] introduced the quantity $\mathcal{D}$. In our notation his definition reads as

$$\mathcal{D}(V) = \partial_t A - \Delta A - 2|\alpha_{ij}|^2 + 4\nabla_i \alpha_{ij} V_j - 2\nabla_j A V_j + 2R_{ij} V_i V_j - 2\alpha_{ij} V_i V_j.$$

Note that $\mathcal{D}$ and Θ are essentially the same; indeed $\mathcal{D}(V) = 2\Theta(-V)$. Müller [2010] further explained that $\mathcal{D}$ is the difference between two differential Harnack-type quantities for the tensor α.

Let $u(t,y)$ be a nonnegative solution to the conjugate heat equation

$$\frac{\partial u(t,y)}{\partial t} = -\Delta u(t,y) + A(t,y)\,u(t,y), \quad t\in(0,T),\ y\in M, \tag{1-4}$$

where Δ is the Laplace–Beltrami operator calculated with respect to the evolving metric $g(t)$. Note that $\int_M u(t,y)\,dy$ remains constant along the flow, and without loss of generality we assume this constant to be 1.

The classical Boltzmann–Shannon entropy functional is defined by

$$\mathcal{E}(t) = \int_M u(t,y)\log u(t,y)\,dy. \tag{1-5}$$

If $\Theta(V)\ge 0$ for all V, we will show that $\mathcal{E}$ is convex. Based on this observation we construct Perelman's $\mathcal{F}$ and $\mathcal{W}$ entropy in abstract geometric flows. We then derive the explicit evolution equations of the entropies along the conjugate heat equation, and show that they are monotone if $\Theta\ge 0$. We thus present a unified formula of various $\mathcal{W}$ entropies established by various authors for different flows (including the static case); see [Feldman et al. 2005; Li 2007; List 2008; Müller 2012; Ni 2004b; 2004a; Perelman 2002].

We show indeed that the generalized entropy $\mathcal{F}_k$ $(k\ge 1)$, see Definition 4.1 below, is monotone under the additional assumption $B-\Delta A\ge 0$, which is satisfied by all previously mentioned flows. The study of the $\mathcal{F}_k$ entropy leads to a simpler argument to rule out nontrivial steady breathers.

The eigenvalues and eigenfunctions of the Laplace–Beltrami operator with potential cA where c is a constant, satisfy

$$\lambda(t) f(t,y) = -\Delta f(t,y) + cA(t,y) f(t,y). \tag{1-6}$$

Let $\lambda(t)$ be the lowest eigenvalue. We shall determine the derivative of $\lambda(t)$. A remarkable fact is that the derivative $\lambda'(t)$ does not depend on the time derivative of the corresponding eigenfunction; this allows to establish a formula for $\lambda'(t)$ not requiring knowledge of the eigenfunction evolution. We will prove eigenvalue monotonicity and apply it to rule out nontrivial steady and expanding breathers in various flows.

2. The first two derivatives of the Boltzmann–Shannon entropy

Theorem 2.1. *Suppose that $(M, g(t))$ is a solution to the abstract geometric flow (1-1), and that $u(t, y)$ is a positive solution to the conjugate heat equation (1-4), normalized by $\int_M u(t, y)\, dy = 1$. The first two derivatives of $\mathscr{E}(t)$ are given by*

$$\mathscr{E}'(t) = \int_M (|\nabla \log u|^2 + A) u\, dy, \tag{2-7}$$

$$\mathscr{E}''(t) = \int_M 2\big(|\alpha - \nabla\nabla \log u|^2 + \Theta(\nabla \log u)\big) u\, dy. \tag{2-8}$$

In particular, if Θ is nonnegative then $\mathscr{E}(t)$ is convex in time.

Proof. Since M is closed we can integrate by parts freely. Direct calculations show that

$$\begin{aligned}
\mathscr{E}'(t) &= \int_M (u_t \log u + u_t - Au \log u)\, dy \\
&= \int_M \big((-\Delta u + Au) \log u - \Delta u + Au - Au \log u\big)\, dy \\
&= \int_M (-\Delta u \log u + Au)\, dy = \int_M (|\nabla \log u|^2 + A) u\, dy,
\end{aligned}$$

and

$$\begin{aligned}
\mathscr{E}''(t) &= \int_M \bigg(\frac{\partial(|\nabla \log u|^2 + A)}{\partial t} u + (|\nabla \log u|^2 + A) \frac{\partial u}{\partial t} - (|\nabla \log u|^2 + A) u A \bigg)\, dy \\
&= \int_M \bigg(\Big(2\alpha(\nabla \log u, \nabla \log u) + 2 \Big\langle \nabla \frac{u_t}{u}, \nabla \log u \Big\rangle + 2|\alpha|^2 + B \Big) u \\
&\qquad\qquad + (|\nabla \log u|^2 + A)(-\Delta u + Au) - (|\nabla \log u|^2 + A) u A \bigg)\, dy \\
&= \int_M \bigg(\Big(2\alpha(\nabla \log u, \nabla \log u) + 2 \Big\langle \nabla \Big(-\frac{\Delta u}{u} + A \Big), \nabla \log u \Big\rangle \Big) u \\
&\qquad\qquad + (2|\alpha|^2 + B) u - (|\nabla \log u|^2 + A) \Delta u \bigg)\, dy \\
&= \int_M \bigg(2u\alpha(\nabla \log u, \nabla \log u) - 2 \Big\langle \nabla \Big(\frac{\Delta u}{u} \Big), \nabla u \Big\rangle + 2 \langle \nabla A, \nabla u \rangle \\
&\qquad\qquad + 2u|\alpha|^2 + Bu - \Delta(|\nabla \log u|^2) u - \Delta A u \bigg)\, dy.
\end{aligned}$$

Plugging in $\Delta \log u = \Delta u/u - |\nabla \log u|^2$ and

$$\Delta(|\nabla \log u|^2) = 2\,|\nabla\nabla \log u|^2 + 2\,\mathrm{Rc}(\nabla \log u,\, \nabla \log u) + 2\langle \nabla \log u,\, \nabla(\Delta \log u)\rangle,$$

we have

$$\begin{aligned}\mathscr{E}''(t) &= \int_M \big(2u(|\alpha|^2+|\nabla\nabla \log u|^2)+2u(\alpha+\mathrm{Rc})(\nabla \log u,\nabla \log u)+Bu-3\Delta A u\big)\,dy\\ &= \int_M 2\Big(u\,|\alpha-\nabla\nabla \log u|^2 + 4u\langle \alpha,\, \nabla\nabla \log u\rangle\\ &\qquad + 2u(\alpha+\mathrm{Rc})(\nabla \log u,\, \nabla \log u) + (B-\Delta A)u + 2\langle \nabla A,\, \nabla u\rangle\Big)\,dy.\end{aligned}$$

By observing that

$$\mathrm{Div}(u\alpha(\nabla \log u)) = \alpha(\nabla \log u,\, \nabla u) + u\,\mathrm{Div}(\alpha)(\nabla \log u) + u\langle \alpha,\, \nabla\nabla \log u\rangle,$$

and by the divergence theorem, we get

$$\begin{aligned}\mathscr{E}''(t) = \int_M \Big(&2u|\alpha-\nabla\nabla \log u|^2 + 2u(\mathrm{Rc}-\alpha)(\nabla \log u,\, \nabla \log u)\\ &+ (B-\Delta A)u + \langle 2\nabla A - 4\,\mathrm{Div}(\alpha),\, \nabla u\rangle\Big)\,dy,\end{aligned}$$

which is exactly (2-8). □

3. Examples where Θ and $B - \Delta A$ are nonnegative

We next list some examples where Θ and $B-\Delta A$ are nonnegative. Calculations on the Ricci flow and extended Ricci flow are carried out in detail. For other examples we list values of Θ and $B-\Delta A$, and for details we refer to [Müller 2010]. This section is organized in the same way as the corresponding section there. Recall that

$$\Theta(V) = (\mathrm{Rc}-\alpha)(V,\, V) + \langle \nabla A - 2\,\mathrm{Div}(\alpha),\, V\rangle + \tfrac{1}{2}(B-\Delta A).$$

Riemannian manifolds. In the case of a static metric we have $\alpha = 0$ and hence

(3-9) $$\Theta(V) = \mathrm{Rc}(V,\, V), \qquad B - \Delta A = 0.$$

Thus Θ is nonnegative if M has nonnegative Ricci curvature.

Hamilton's Ricci flow. In the case of Ricci flow where $\alpha = \mathrm{Rc}$, we have $A = R$. The evolution equation $\partial R/\partial t = 2|\mathrm{Rc}|^2 + \Delta R$ gives

$$B = \frac{\partial A}{\partial t} - 2\,|\alpha|^2 = \Delta R.$$

Notice that $\nabla R = 2\,\mathrm{Div}(\mathrm{Rc})$ by the second Bianchi identity; we thus get

(3-10) $$\Theta(V) = 0, \qquad B - \Delta A = 0.$$

List's extended Ricci flow. Bernhard List [2008] introduced an extended Ricci flow system, namely

$$\frac{\partial g}{\partial t} = -2\,\mathrm{Rc} + 2a_n \nabla v \otimes \nabla v, \tag{3-11}$$

where v is a solution to the heat equation $\partial v/\partial t = \Delta v$ and a_n a positive constant depending only on the dimension n of the manifold. It turns out that one can exhibit List's flow as a Ricci–DeTurck flow in one higher dimension. This connection has been observed by Jun-Fang Li according to [Akbar and Woolgar 2009]. The extended Ricci flow is interesting by itself since its stationary points are solutions to the vacuum Einstein equations, and it is desirable to work on this flow directly.

In our notation for the extended Ricci flow, $\alpha = \mathrm{Rc} - a_n\, dv \otimes dv$ and $A = R - a_n |\nabla v|^2$, which gives

$$\nabla A = \nabla R - 2a_n \nabla\nabla v(\nabla v, \cdot).$$

Since $\mathrm{Div}(dv \otimes dv)_k = g^{ij} \nabla_i (\nabla_j v \nabla_k v) = (\Delta v)\nabla_k v + g^{ij} \nabla_j v \nabla_i \nabla_k v$, we have

$$\mathrm{Div}\,\alpha = \mathrm{Div}\,\mathrm{Rc} - a_n\,\mathrm{Div}(dv \otimes dv) = \tfrac{1}{2}\nabla R - a_n(\Delta v \nabla v + \nabla\nabla v(\nabla v, \cdot)).$$

Thus we find

$$\nabla A - 2\,\mathrm{Div}(\alpha) = 2a_n \Delta v \nabla v. \tag{3-12}$$

The evolution equation of α is given by (cf. [List 2008])

$$\beta_{ij} = \frac{\partial \alpha_{ij}}{\partial t} = \Delta \alpha_{ij} - R_{ip}\alpha_{pj} - R_{jp}\alpha_{pi} + 2R_{ipqj}\alpha_{pq} + 2a_n \Delta v \nabla_i \nabla_j v.$$

(Note that in our notation $R_{ij} = g^{pq} R_{ipqj}$, while many authors, including List, write $R_{ij} = -g^{pq} R_{ipqj}$.) Hence we have $B = \Delta A + 2a_n(\Delta v)^2$ and

$$B - \Delta A = 2a_n(\Delta v)^2. \tag{3-13}$$

Plugging in our formula for Θ we arrive at

$$\Theta(V) = a_n \langle \nabla v, V\rangle^2 + 2a_n \Delta v \langle \nabla v, V\rangle + a_n(\Delta v)^2 = a_n(\langle \nabla v, V\rangle + \Delta v)^2.$$

Müller's Ricci flow coupled with harmonic map flow. The Ricci flow coupled with an harmonic map flow was introduced in [Müller 2012] as a generalization of the extended Ricci flow. Suppose that (N, γ) is a further closed static Riemannian manifold, $a(t)$ a nonnegative function depending only on time, and $\varphi(t)\colon M \to N$ a family of 1-parameter smooth maps. Then $(g(t), \varphi(t))$ is called a solution to Müller's Ricci flow coupled with harmonic map flow with coupling function $a(t)$, if it satisfies

$$\begin{cases} \partial g/\partial t = -2\mathrm{Rc} + 2a(t)\,\nabla\varphi \otimes \nabla\varphi, \\ \partial\varphi/\partial t = \tau_g \varphi, \end{cases} \tag{3-14}$$

where τ_g denotes the tension field of the map φ with respect to the evolving metric $g(t)$ and $\nabla\varphi \otimes \nabla\varphi \equiv \varphi^*\gamma$ the pullback of the metric γ on N via the map φ.

Recall that $\mathfrak{D}(V) = 2\Theta(-V)$; we have, as in [Müller 2010],

$$B - \Delta A = 2a\,|\tau_g\varphi|^2 - a'|\nabla\varphi|^2, \quad \Theta(V) = a\,|\tau_g\varphi + \nabla_V\varphi|^2 - \tfrac{1}{2}a'|\nabla\varphi|^2. \tag{3-15}$$

Thus both $B - \Delta A$ and Θ are nonnegative as long as $a(t)$ is nonincreasing in time.

Lorentzian mean curvature flow when the ambient space has nonnegative sectional curvature. Let L^{n+1} be a Lorentzian manifold, and $M(t)$ be a family of space-like hypersurfaces of L. Denote by ν the future-oriented time-like unit normal vector of M, by h_{ij} the second fundamental form, and by H its mean curvature. Let $F(t, y)$ be the position function of M in L. The Lorentzian mean curvature flow is then defined by

$$\frac{\partial F}{\partial t} = H\nu. \tag{3-16}$$

The induced metric $g(t)$ of $M(t)$ satisfies $\partial_t g = 2Hh_{ij}$. We have

$$\begin{aligned} B - \Delta A &= 2H^2|h|^2 + 2|\nabla H|^2 + 2H^2\,\overline{\mathrm{Rc}}(\nu, \nu), \\ \Theta(V) &= |\nabla H + h(V, \cdot)|^2 + \overline{\mathrm{Rc}}(H\nu + V, H\nu + V) + \overline{\mathrm{Rm}}(V, \nu, \nu, V), \end{aligned} \tag{3-17}$$

where $\overline{\mathrm{Rc}}$ and $\overline{\mathrm{Rm}}$ denote the Ricci and Riemann curvature tensors of L^{n+1}. Both $B - \Delta A$ and Θ are obviously nonnegative when the sectional curvature of L^{n+1} is nonnegative.

4. Perelman's $\mathcal{F}_k$ functional in abstract geometric flows

We proved the following. If $(M, g(t))$ is a solution to the abstract flow equation (1-1) and u a positive solution to the conjugate heat equation (1-4) then

$$\frac{d}{dt}\int_M (|\nabla \log u|^2 + A)u\,dy = \int_M 2\big(|\alpha - \nabla\nabla \log u|^2 + \Theta(\nabla \log u)\big)u\,dy. \tag{4-18}$$

We note that

$$\begin{aligned} \frac{d}{dt}\int_M Au\,dy &= \int_M \left(\frac{\partial A}{\partial t}u + A\frac{\partial u}{\partial t} - A^2u\right)dy \\ &= \int_M \big((2|\alpha|^2 + B)u + A(-\Delta u + Au) - A^2u\big)\,dy \\ &= \int_M 2\big(|\alpha|^2 + \tfrac{1}{2}(B - \Delta A)\big)u\,dy. \end{aligned} \tag{4-19}$$

Let $\phi := -\log u$; then

$$\frac{\partial \phi}{\partial t} = -\Delta\phi + |\nabla\phi|^2 - A, \tag{4-20}$$

with constraint $\int_M e^{-\phi}\,dy = 1$. We rewrite (4-18) in the more familiar form following Perelman's notation:

$$\frac{d}{dt}\int_M (|\nabla\phi|^2 + A)e^{-\phi}\,dy = \int_M 2\big(|\alpha + \nabla\nabla\phi|^2 + \Theta(-\nabla\phi)\big)e^{-\phi}\,dy. \tag{4-21}$$

Definition 4.1. For any $\phi \in C^\infty(M)$ with $\int_M e^{-\phi}\,dy = 1$ and any constant k we define Perelman's $\mathcal{F}_k$-functional for abstract geometric flows by

$$\mathcal{F}_k(g,\phi) = \int_M (|\nabla\phi|^2 + kA)e^{-\phi}\,dy. \tag{4-22}$$

When $k = 1$ we simply denote $\mathcal{F}_1$ by $\mathcal{F}$.

For Perelman's $\mathcal{F}_k$-functional in an abstract geometric flow we have:

Theorem 4.2. *If g is a solution of the abstract geometric flow equation (1-1) and ϕ a solution to (4-20) then we have*

$$\frac{d}{dt}\mathcal{F}_k = \int_M 2\Big(|\alpha+\nabla\nabla\phi|^2 + (k-1)|\alpha|^2 + \Theta(-\nabla\phi) + \frac{k-1}{2}(B - \Delta A)\Big) \cdot e^{-\phi}\,dy. \tag{4-23}$$

Thus for $k > 1$, $\mathcal{F}_k$ is monotone nondecreasing as long as $B - \Delta A$ and Θ are nonnegative. Moreover the monotonicity is strict unless

$$\alpha = 0, \quad \phi = \text{constant}, \quad B - \Delta A = 0.$$

For $k = 1$ we have

$$\frac{d}{dt}\mathcal{F} = \int_M 2\big(|\alpha + \nabla\nabla\phi|^2 + \Theta(-\nabla\phi)\big)e^{-\phi}\,dy. \tag{4-24}$$

In particular, $\mathcal{F}$ is monotone nondecreasing when $\Theta \geq 0$, and monotonicity is strict unless

$$\alpha + \nabla\nabla\phi = 0, \quad \Theta(-\nabla\phi) = 0.$$

Proof. Since

$$\mathcal{F}_k(g,\phi) = \int_M (|\nabla\phi|^2 + A)e^{-\phi}\,dy + (k-1)\int_M Ae^{-\phi}\,dy,$$

and by (4-21) and (4-19) we immediately get formula (4-23).

Furthermore for $k > 1$, the functional $\mathcal{F}_k$ is monotone nondecreasing as long as $B - \Delta A$ and Θ are nonnegative. When $d/dt\mathcal{F}_k = 0$, each term on the right side of (4-23) has to be identically zero. In particular we have

$$\alpha + \nabla\nabla\phi = 0, \quad \alpha = 0,$$

which further implies $\Delta\phi = 0$ on the closed manifold M, and thus ϕ has to be a constant. Now $\Theta(-\nabla\phi) = \Theta(0) = (B - \Delta A)/2$ and $B - \Delta A = 0$.

When $k = 1$ the statement in the theorem is obvious. $\square$

The advantage of $\mathcal{F}_k$ over $\mathcal{F}$ is that when $k > 1$, extra terms in $\mathcal{F}_k'$ can tell more about the manifold M. Li [2007] has studied $\mathcal{F}_k$ in the Ricci flow. We state an analogous application of $\mathcal{F}_k$ to rule out nontrivial steady breathers in abstract geometric flows.

Recall that a breather of a geometric flow is a periodic solution changing only by diffeomorphism and rescaling. A solution $(M, g(t))$ is called a breather if there are a diffeomorphism $\eta\colon M \to M$, a positive constant c and times $t_1 < t_2$ such that

$$g(t_2) = c\,\eta^* g(t_1), \quad \alpha(t_2) = \eta^* \alpha(t_1). \tag{4-25}$$

When $c < 1$, $c = 1$ or $c > 1$, the breather is called shrinking, steady or expanding, respectively.

We now apply monotonicity of $\mathcal{F}_k$ to rule out nontrivial steady breathers of abstract geometric flows.

Corollary 4.3. *Suppose that $(M, g(t))$ is a steady breather to an abstract geometric flow (1-1). Suppose that $\Theta \geq 0$ and $B - \Delta A \geq 0$. Then $B - \Delta A = 0$ and the steady breather is α-flat, namely $\alpha = 0$.*

Proof. The arguments are standard and follow from Perelman's proof [2002] of the no steady breather theorem for the Ricci flow. We follow [Kleiner and Lott 2008] and only sketch the proof. Define

$$\lambda(t) = \inf\left\{\mathcal{F}_k(g, \phi)\colon \int_M e^{-\phi}\, dy = 1,\ \phi \in C^\infty(M)\right\}. \tag{4-26}$$

Since we are on a steady breather we have $\lambda(t_1) = \lambda(t_2)$. Let $\bar{\phi}(t_2)$ be a minimizer of $\lambda(t_2)$. Solve the conjugate heat equation backwards with end value $e^{-\bar{\phi}(t_2)}$. Denote the solution by $u(t)$. Let $\phi(t) = -\log u(t)$ then $\phi(t)$ satisfies the constraint

$$\int_M e^{-\phi}\, dy = 1,$$

and $\mathcal{F}_k(g(t), \phi(t))$ is monotone nondecreasing as its derivative is nonnegative when $e^{-\phi(t)}$ is a solution to the conjugate heat equation. Thus we have

$$\lambda(t_1) \leq \mathcal{F}_k(g(t_1), \phi(t_1)) \leq \mathcal{F}_k(g(t_2), \bar{\phi}(t_2)) = \lambda(t_2). \tag{4-27}$$

Since on a breather $\lambda(t_1) = \lambda(t_2)$, we get

$$\mathcal{F}_k(g(t_1), \phi(t_1)) = \mathcal{F}_k(g(t_2), \phi(t_2)),$$

and in particular $\mathcal{F}_k'(g(t), \phi(t)) = 0$ when $t \in [t_1, t_2]$. Now we apply Theorem 4.2 to conclude that $\alpha = 0$ and $B - \Delta A = 0$ on M when $t \in [t_1, t_2]$. $\square$

Remark 4.4. From (4-26) we know that λ is the lowest eigenvalue of $-\Delta+(k/4)A$. Thus, by Theorem 4.2, under the assumptions that $B-\Delta A\geq 0$ and $\Theta\geq 0$, the lowest eigenvalue of $-\Delta+(k/4)A$ is monotone in t when $k\geq 1$. An explicit formula for the derivative of the lowest eigenvalue will be given in Section 7 under the assumption that λ is differentiable along time.

5. Construction of Perelman's $\mathscr{W}$ entropy

We have noted that Perelman's $\mathscr{F}$-functional is the derivative of $\mathscr{E}$, whose stationary points are steady solitons. The purpose of this section is to construct functionals corresponding to the shrinking solitons. Our construction is just completing squares of $\mathscr{E}''$ (or $\mathscr{F}'$ by Perelman's notation). Monotonicity of $\mathscr{W}$ holds in the flows mentioned in Section 3.

We rewrite the second derivative of $\mathscr{E}(t)$ in order to fit the shrinking soliton equation simply by completing squares:

$$\begin{aligned}\mathscr{E}''(t) &= \int_M 2\big(|\alpha-\nabla\nabla\log u|^2+\Theta(\nabla\log u)\big)u\,dy\\ &= \int_M\biggl(2u\left|\alpha-\nabla\nabla\log u-\frac{1}{2(T-t)}g\right|^2+\frac{2u}{T-t}(A-\Delta\log u)\\ &\qquad\qquad -\frac{2nu}{4(T-t)^2}+2u\Theta(\nabla\log u)\biggr)\,dy\\ &= \int_M 2\left(\left|\alpha-\nabla\nabla\log u-\frac{1}{2(T-t)}g\right|^2+\Theta(\nabla\log u)\right)u\,dy\\ &\qquad\qquad +\frac{2}{T-t}\mathscr{E}'(t)-\frac{n}{2(T-t)^2}.\end{aligned}$$

Hence we have

$$\begin{aligned}&\int_M 2\left(\left|\alpha-\nabla\nabla\log u-\frac{1}{2(T-t)}g\right|^2+\Theta(\nabla\log u)\right)u\,dy\\ &\qquad = \mathscr{E}''(t)-\frac{2}{T-t}\mathscr{E}'(t)+\frac{n}{2(T-t)^2}\\ &\qquad = \frac{1}{T-t}\frac{d}{dt}\left((T-t)\mathscr{E}'-\mathscr{E}-\frac{n}{2}\log(T-t)\right).\end{aligned}$$

Now in terms of

$$\mathscr{W}:=(T-t)\mathscr{E}'-\mathscr{E}-\frac{n}{2}\log(T-t)-\frac{n}{2}\log(4\pi)-n,$$

we have proved that

$$\frac{d}{dt}\mathscr{W}=(T-t)\int_M 2\left(\left|\alpha-\nabla\nabla\log u-\frac{1}{2(T-t)}g\right|^2+\Theta(\nabla\log u)\right)u\,dy. \tag{5-28}$$

Following Perelman, we let

$$\tau := T - t, \quad \phi := -\log((4\pi\tau)^{n/2}u),$$

and introduce the following definition.

Definition 5.1. For a solution (M, g) to the abstract geometric flow equation (1-1) and for $\phi \in C^\infty(M)$, let Perelman's $\mathcal{W}$-entropy be defined as

$$\mathcal{W}(g, \phi, t) = \int_M \big(\tau(|\nabla\phi|^2 + A) + \phi - n\big)(4\pi\tau\big)^{-n/2} e^{-\phi}\, dy. \tag{5-29}$$

We can rewrite (5-28) in the following way.

Theorem 5.2. *Let (M, g) be a solution to the abstract geometric flow equation (1-1). If ϕ satisfies*

$$\frac{\partial\phi}{\partial t} = -\Delta\phi + |\nabla\phi|^2 - A + \frac{n}{2\tau} \quad \textit{and} \quad \int_M (4\pi\tau)^{-n/2} e^{-\phi}\, dy = 1,$$

then

$$\frac{d}{dt}\mathcal{W} = \int_M 2\tau\left(\left|\alpha + \nabla\nabla\phi - \frac{1}{2\tau} g\right|^2 + \Theta(-\nabla\phi)\right)(4\pi\tau)^{-n/2} e^{-\phi}\, dy.$$

If $\Theta \geq 0$ then $\mathcal{W}$ is monotone nondecreasing, and the monotonicity is strict unless

$$\alpha + \nabla\nabla\phi - \frac{1}{2\tau} g = 0, \quad \Theta(-\nabla\phi) = 0.$$

The monotonicity of $\mathcal{W}$ can be applied to rule out nontrivial shrinking breathers in abstract flows with $\Theta \geq 0$. The arguments are almost identical to the Ricci flow case. We omit details.

6. Expander entropy $\mathcal{W}_+$

Feldman, Ilmanen, and Ni [2005] established expander entropy $\mathcal{W}_+$ for Ricci flow, and there has been a very nice explanation of their motivation in [Feldman et al. 2005]. We attempt to explain formally why $\mathcal{W}_+$ should be the way as they defined it. In short, the signs in $\mathcal{W}$ and $\mathcal{W}_+$ are caused by antiderivatives of $1/(t - T)$ depending on the situation whether $t > T$ or $t < T$.

We now carry out the details. Note that $t > T$ on expanders and that

$$\begin{aligned}\mathcal{E}''(t) &= \int_M 2\,\big(|\alpha - \nabla\nabla\log u|^2 + \Theta(\nabla\log u)\big)u\, dy \\ &= \int_M \left(2u\left|\alpha - \nabla\nabla\log u + \frac{1}{2(t-T)} g\right|^2 - \frac{2u}{t-T}(A - \Delta\log u) \right. \\ &\qquad\qquad \left. - \frac{2nu}{4(t-T)^2} + 2u\Theta(\nabla\log u)\right) dy\end{aligned}$$

$$= \int_M 2\left(\left|\alpha - \nabla\nabla \log u + \frac{1}{2(t-T)} g\right|^2 + \Theta(\nabla \log u)\right) u\, dy \\ - \frac{2}{t-T}\mathscr{E}'(t) - \frac{n}{2(t-T)^2};$$

moreover

$$\int_M 2\left(\left|\alpha - \nabla\nabla \log u + \frac{1}{2(t-T)} g\right|^2 + \Theta(\nabla \log u)\right) u\, dy \\ = \mathscr{E}''(t) + \frac{2}{t-T}\mathscr{E}'(t) + \frac{n}{2(t-T)^2} \\ = \frac{1}{t-T}\frac{d}{dt}\left((t-T)\mathscr{E}' + \mathscr{E} + \frac{n}{2}\log(t-T)\right).$$

The calculations suggest to define

$$\mathscr{W}_+ := (t-T)\mathscr{E}' + \mathscr{E} + \frac{n}{2}\log(t-T) + \frac{n}{2}\log(4\pi) + n$$

which is the definition of expander entropy in [Feldman et al. 2005] in the case of Ricci flow. One has

$$\frac{d\mathscr{W}_+}{dt} = (t-T)\int_M 2\left(\left|\alpha - \nabla\nabla \log u + \frac{1}{2(t-T)} g\right|^2 + \Theta(\nabla \log u)\right) u\, dy.$$

This again may be rewritten following [Feldman et al. 2005] in terms of

$$\sigma := t - T, \quad \phi_+ := -\log((4\pi\sigma)^{n/2} u).$$

Definition 6.1. For a solution (M, g) to the abstract geometric flow equation (1-1) and $\phi_+ \in C^\infty(M)$ one defines Perelman's entropy for expanders by

$$\mathscr{W}_+(g, \phi_+, t) = \int_M \big(\sigma\,(|\nabla\phi_+|^2 + A) - \phi_+ + n\big)(4\pi\sigma)^{-n/2} e^{-\phi_+}\, dy. \tag{6-30}$$

Theorem 6.2. *Let (M, g) be a solution to the abstract geometric flow equation (1-1). Assume that ϕ_+ satisfies*

$$\frac{\partial\phi_+}{\partial t} = -\Delta\phi_+ + |\nabla\phi_+|^2 - A - \frac{n}{2(t-T)} \quad \text{and} \quad \int_M (4\pi\sigma)^{-n/2} e^{-\phi_+}\, dy = 1.$$

Then

$$\frac{d\mathscr{W}_+}{dt} = \int_M 2\sigma\left(\left|\alpha + \nabla\nabla\phi_+ + \frac{1}{2(t-T)} g\right|^2 + \Theta(-\nabla\phi_+)\right)(4\pi\sigma)^{-n/2} e^{-\phi_+}\, dy.$$

Furthermore, if $\Theta \geq 0$ then $\mathscr{W}_+$ is monotone nondecreasing, and monotonicity is strict unless

$$\alpha + \nabla\nabla\phi_+ + \frac{1}{2(t-T)} g = 0, \quad \Theta(-\nabla\phi_+) = 0.$$

Remark 6.3. The constants $\pm\left(\frac{n}{2}\log(4\pi)+n\right)$ in the definition of $\mathcal{W}$ and $\mathcal{W}_+$ are for normalization purposes.

7. Evolution equation of the lowest eigenvalue

In this section, assuming that the lowest eigenvalue $\lambda(t)$ is differentiable along t, we derive an explicit formula for its derivative in terms of its normalized eigenfunction. Although monotonicity of $\mathcal{F}_k$ in Theorem 4.2 is sufficient for our geometric applications, an explicit formula which holds at points where λ is differentiable, may be of independent interest. Time derivatives of the eigenfunction do not appear in the formula.

In the literature, for instance [Kleiner and Lott 2008, Section 7], it has been stated that smooth dependence on time of the lowest eigenvalue and the corresponding eigenfunction follows from perturbation theory as presented in [Reed and Simon 1978, Chapter XII]. However it is not immediately clear how perturbation theory is applied to our context, where the operator depends only smoothly, but not analytically on t.

Lemma 7.1. *Assume that M is a closed manifold and let $\psi \in C^\infty(M)$. Let λ be the lowest eigenvalue of $-\Delta+\psi$ and f a positive eigenfunction corresponding to λ, so that $\lambda f = -\Delta f + \psi f$. Then*

$$\int_M \psi \Delta f^2 \, dy = \int_M 2\left(|\nabla\nabla \log f|^2 + \mathrm{Rc}(\nabla \log f, \nabla \log f)\right) f^2 \, dy. \tag{7-31}$$

Proof. We have $\psi f = \lambda f + \Delta f$ and

$$\begin{aligned}
\psi \Delta f^2 &= 2\psi f \Delta f + 2\psi |\nabla f|^2 \\
&= 2(\lambda f + \Delta f)\Delta f + 2(\lambda f + \Delta f)\frac{|\nabla f|^2}{f} \\
&= \lambda(2f\Delta f + 2|\nabla f|^2) + 2(\Delta f)^2 + 2\,\frac{\Delta f |\nabla f|^2}{f} \\
&= \lambda \Delta f^2 + 2(\Delta f)^2 + 2\,\frac{\Delta f |\nabla f|^2}{f}.
\end{aligned}$$

We observe that

$$\begin{aligned}
\int_M \psi \Delta f^2 \, dy &= \int_M \left(2(\Delta f)^2 + 2\,\frac{\Delta f |\nabla f|^2}{f}\right) dy \\
&= \int_M \left(-2\,\langle \nabla f, \nabla(\Delta f)\rangle - 2\left\langle \nabla f, \nabla\Big(\frac{|\nabla f|^2}{f}\Big)\right\rangle\right) dy.
\end{aligned} \tag{7-32}$$

Now we calculate the two terms on the right side of (7-32). For the first term we have by Bochner's formula

$$-2\langle \nabla f, \nabla(\Delta f)\rangle = 2|\nabla\nabla f|^2 + 2\,\mathrm{Rc}(\nabla f, \nabla f) - \Delta(|\nabla f|^2). \tag{7-33}$$

The second term can be written as

$$\begin{aligned}\Big\langle \nabla f, \nabla\Big(\frac{|\nabla f|^2}{f}\Big)\Big\rangle &= \langle \nabla f, \nabla(f|\nabla \log f|^2)\rangle \\ &= \langle \nabla f, \nabla f|\nabla \log f|^2 + 2f\nabla\nabla \log f(\nabla \log f, \cdot)\rangle \\ &= f^2|\nabla \log f|^4 + 2f^2\nabla\nabla \log f(\nabla \log f, \nabla \log f) \\ &= |\nabla\nabla f|^2 - f^2|\nabla\nabla \log f|^2,\end{aligned} \tag{7-34}$$

where in the last equality we used that

$$\nabla\nabla \log f = \frac{\nabla\nabla f}{f} - \frac{\nabla f\otimes\nabla f}{f^2} = \frac{\nabla\nabla f}{f} - \nabla \log f\otimes\nabla \log f,$$

and moreover

$$\begin{aligned}|\nabla\nabla f|^2 &= f^2|\nabla\nabla \log f + \nabla \log f\otimes\nabla \log f|^2 \\ &= f^2|\nabla\nabla \log f|^2 + 2f^2\nabla\nabla \log f(\nabla \log f, \nabla \log f) + f^2|\nabla \log f|^4.\end{aligned}$$

Plugging (7-33) and (7-34) into (7-32), we get

$$\int_M \psi\,\Delta f^2\,dy = \int_M \big(2f^2|\nabla\nabla \log f|^2 + 2\,\mathrm{Rc}(\nabla f, \nabla f)\big)\,dy. \qquad \square$$

Let $\lambda(t)$ be the lowest eigenvalue of $-\Delta + cA$, where c is a constant; indeed

$$\lambda(t) = \inf\Big\{\int_M |\nabla\phi|^2 + cA\phi^2\,dy : \int_M \phi^2\,dy = 1,\ \phi\in C^\infty(M)\Big\}. \tag{7-35}$$

Let $f(t,\cdot)$ be the corresponding positive eigenfunction normalized by

$$\int_M f^2(t,y)\,dy = 1.$$

Theorem 7.2. *At all times t_0 when the function $t\mapsto\lambda(t)$ is differentiable we have*

$$\begin{aligned}&\lambda'(t_0) \\ &= \frac{1}{2}\int_M \Big(|\alpha - 2\nabla\nabla \log f|^2 + (4c-1)|\alpha|^2 + \Theta(2\nabla \log f) + \frac{4c-1}{2}(B - \Delta A)\Big) f^2\,dy.\end{aligned} \tag{7-36}$$

In particular, for $c = \frac{1}{4}$ we have

$$\lambda' = \frac{1}{2}\int_M \big(|\alpha - 2\nabla\nabla \log f|^2 + \Theta(2\nabla \log f)\big) f^2\,dy. \tag{7-37}$$

Proof. Fix $t_0 \in (0, T)$ where the function $t \mapsto \lambda(t)$ is differentiable, and let $\varphi : (0, T) \times M \to \mathbb{R}_{>0}$ be a smooth function such that

(1) $\int_M \varphi(t, y)^2 \, dy = 1$ for all $t \in (0, T)$, and

(2) $\varphi(t_0, \cdot) = f(t_0, \cdot)$.

For instance $\varphi(t)$ may be chosen as $f(t_0)\sqrt{dy(g(t_0))/dy(g(t))}$, where $dy(g(t))$ is the volume form with respect to the metric $g(t)$. Let

$$\mu(t) := \int_M \big(|\nabla\varphi(t, y)|^2 + cA(t, y)\varphi(t, y)^2\big)\, dy. \tag{7-38}$$

Then $\mu(t)$ is a smooth function by definition. The trick to work with $\mu(t)$ rather than $\lambda(t)$ allows to bypass time derivatives of the eigenfunction $f(t, \cdot)$. Note that $\mu(t) \geq \lambda(t)$ for all $t \in (0, T)$, and $\mu(t_0) = \lambda(t_0)$, so that

$$\lambda'(t_0) = \mu'(t_0).$$

Differentiation of (7-38) gives

$$\begin{aligned}
\mu' &= \int_M \big(2\alpha(\nabla\varphi, \nabla\varphi) + 2\langle\nabla\varphi', \nabla\varphi\rangle + cA'\varphi^2 + 2cA\varphi\varphi' - (|\nabla\varphi|^2 + cA\varphi^2)A\big)\, dy \\
&= \int_M \big(2\alpha(\nabla\varphi, \nabla\varphi) - 2\varphi'\Delta\varphi + cA'\varphi^2 + 2cA\varphi\varphi' + \varphi\langle\nabla A, \nabla\varphi\rangle + A\varphi\Delta\varphi \\
&\qquad\qquad - cA^2\varphi^2\big)\, dy \\
&= \int_M \big(2\alpha(\nabla\varphi, \nabla\varphi) + cA'\varphi^2 + \varphi\langle\nabla A, \nabla\varphi\rangle\big)\, dy + \lambda \int_M (2\varphi'\varphi - A\varphi^2)\, dy \\
&= \int_M \big(2\alpha(\nabla\varphi, \nabla\varphi) + cA'\varphi^2 + \varphi\langle\nabla A, \nabla\varphi\rangle\big)\, dy \\
&= \int_M \big(2\alpha(\nabla\varphi, \nabla\varphi) + c(2|\alpha|^2 + B)\varphi^2 - \tfrac{1}{2}A\Delta\varphi^2\big)\, dy \\
&= \int_M \big(2\alpha(\nabla\varphi, \nabla\varphi) + 2c|\alpha|^2\varphi^2 + c(B - \Delta A)\varphi^2 + cA\Delta\varphi^2 - \tfrac{1}{2}A\Delta\varphi^2\big)\, dy,
\end{aligned}$$

where in the fourth equality we used that $\int_M (2\varphi\varphi' - A\varphi^2)\, dy = 0$ (which is due to the normalization of φ).

Noting that

$$\begin{aligned}
\operatorname{Div}(\varphi\alpha(\nabla\varphi, \cdot)) &= \alpha(\nabla\varphi, \nabla\varphi) + \varphi \operatorname{Div}(\alpha)(\nabla\varphi) + \varphi\langle\alpha, \nabla\nabla\varphi\rangle \\
&= 2\alpha(\nabla\varphi, \nabla\varphi) + \varphi \operatorname{Div}(\alpha)(\nabla\varphi) + \varphi^2\langle\alpha, \nabla\nabla \log\varphi\rangle,
\end{aligned}$$

and by the divergence theorem, we have

$$\begin{aligned}
\text{(7-39)} \quad \int_M 2\alpha(\nabla\varphi, \nabla\varphi)\, dy
&= \int_M 4\alpha(\nabla\varphi, \nabla\varphi) - 2\alpha(\nabla\varphi, \nabla\varphi)\, dy \\
&= \int_M -2\varphi \operatorname{Div}(\alpha)(\nabla\varphi) - 2\varphi^2\langle\alpha, \nabla\nabla \log\varphi\rangle - 2\alpha(\nabla\varphi, \nabla\varphi)\, dy.
\end{aligned}$$

In (7-31) let $\psi = cA$, we get

$$\text{(7-40)} \qquad \int_M cA\Delta\varphi^2\, dy = \int_M 2\varphi^2|\nabla\nabla\log\varphi|^2 + 2\operatorname{Rc}(\nabla\varphi, \nabla\varphi)\, dy.$$

Plugging (7-39) and (7-40) into the equation for μ' we obtain

$$\begin{aligned}
\mu' &= \int_M \big(-2\varphi \operatorname{Div}(\alpha)(\nabla\varphi) - 2\varphi^2\langle\alpha, \nabla\nabla\log\varphi\rangle - 2\alpha(\nabla\varphi, \nabla\varphi) + 2c|\alpha|^2\varphi^2 \\
&\qquad + c(B - \Delta A)\varphi^2 + 2\varphi^2|\nabla\nabla\log\varphi|^2 + 2\operatorname{Rc}(\nabla\varphi, \nabla\varphi) - \tfrac{1}{2}A\Delta\varphi^2\big)\, dy \\
&= \int_M \Big(\big(2|\nabla\nabla\log\varphi|^2 - 2\langle\alpha, \nabla\nabla\log\varphi\rangle + \tfrac{1}{2}|\alpha|^2 + \big(2c - \tfrac{1}{2}\big)|\alpha|^2\big)\varphi^2 \\
&\qquad + (2(\operatorname{Rc} - \alpha)(\nabla\log\varphi, \nabla\log\varphi) + \langle\nabla A - 2\operatorname{Div}(\alpha), \nabla\log\varphi\rangle)\varphi^2 \\
&\qquad\qquad + c(B - \Delta A)\varphi^2\Big)\, dy \\
&= \int_M \Big(\tfrac{1}{2}|\alpha - 2\nabla\nabla\log\varphi|^2 + \big(2c - \tfrac{1}{2}\big)|\alpha|^2 + \tfrac{1}{2}\Theta(2\nabla\log\varphi) + \big(c - \tfrac{1}{4}\big)(B - \Delta A)\Big) \\
&\qquad\qquad \cdot \varphi^2\, dy,
\end{aligned}$$

so that

$$\begin{aligned}
\lambda'(t_0) &= \mu'(t_0) \\
&= \frac{1}{2}\int_M \Big(|\alpha - 2\nabla\nabla\log f|^2 + (4c-1)|\alpha|^2 + \Theta(2\nabla\log f) + \frac{4c-1}{2}(B - \Delta A)\Big) \\
&\qquad\qquad \cdot f^2\, dy,
\end{aligned}$$

as claimed. □

Let us compare Theorem 4.2 with Theorem 7.2, and (4-23) with (7-36). Let $\phi = -2\log f$; then (7-36) can be rewritten as

$$\lambda' = \frac{1}{2}\int_M \Big(|\alpha + \nabla\nabla\phi|^2 + (4c-1)|\alpha|^2 + \Theta(-\nabla\phi) + \frac{4c-1}{2}(B - \Delta A)\Big)e^{-\phi}\, dy.$$

Letting $k = 4c$, we see that the two evolution equations are formally proportional. We note that in (4-23) the exponential $e^{-\phi}$ is a normalized solution to the conjugate heat equation, while $e^{-\phi/2}$ in the preceding integrand is the normalized eigenfunction of $\lambda(t)$.

8. Eigenvalue monotonicity in various flows

In this section we list explicit formulas of the eigenvalue evolution in different flows. The constant c is assumed to be no less than $\frac{1}{4}$.

Hamilton's Ricci flow. In the case of Ricci flow, monotonicity of the lowest eigenvalue of $-\Delta + cR$ for $c \geq \frac{1}{4}$ and its applications has been established by Cao [2007; 2008], as mentioned in the introduction. See also [Li 2007]. Plugging

$$\alpha = \mathrm{Rc}, \quad \Theta = 0, \quad B - \Delta A = 0$$

into (7-36) we get Cao's formula [2008] for the Ricci flow:

$$\lambda'(t) = \int_M \tfrac{1}{2}\big(|\mathrm{Rc} - 2\nabla\nabla \log f|^2 + (4c-1)|\mathrm{Rc}|^2\big) f^2\, dy. \tag{8-41}$$

This can be applied to show that every steady breather in the Ricci flow is Ricci flat.

List's extended Ricci flow. We work out the details in the extended Ricci flow.

Corollary 8.1. *Assume that $(M, g(t))$ is a solution to the extended Ricci flow equation, and that $\lambda(t)$ is the lowest eigenvalue of*

$$-\Delta + c(R - a_n|\nabla v|^2), \tag{8-42}$$

then we have

$$\lambda'(t) = \int_M \Big(\tfrac{1}{2}|\mathrm{Rc} - a_n \nabla v \otimes \nabla v - 2\nabla\nabla \log f|^2 + \big(2c - \tfrac{1}{2}\big)|\mathrm{Rc} - a_n \nabla v \otimes \nabla v|^2 \\ + \frac{a_n}{2}\big((\Delta v - 2\langle \nabla v, \nabla \log f\rangle)^2 + (4c-1)(\Delta v)^2\big)\Big) f^2\, dy.$$

In particular, a steady breather of the extended Ricci flow is trivial in the sense that

$$\mathrm{Rc} = 0, \quad v \equiv \text{constant}.$$

Proof. The formula for $\lambda'(t)$ is a direct plug-in. When $(M, g(t))$ is a steady breather, there are times $t_1 < t_2$ such that $\lambda(t_1) = \lambda(t_2)$ for any $c > \frac{1}{4}$. In particular we have $\Delta v = 0$ on the closed manifold M, thus v is constant, and moreover M is Ricci flat by $\mathrm{Rc} - a_n \nabla v \otimes \nabla v = 0$. □

Müller's Ricci flow coupled with harmonic map flow. We already used $\mathcal{F}_k$ to rule out nontrivial steady breathers. Using eigenvalue monotonicity, one does not need to solve the conjugate heat equation. The lowest eigenvalue of

$$-\Delta + c(R - a(t)|\nabla\varphi|^2)$$

is nondecreasing along the flow. The conclusions remain the same as in Corollary 4.3.

Lorentzian mean curvature flow when the ambient space has nonnegative sectional curvature. When M evolves along the Lorentzian mean curvature flow (3-16), the lowest eigenvalue of

$$-\Delta - cH^2$$

is nondecreasing provided sectional curvature of the ambient space is nonnegative.

9. Normalized eigenvalue and no expanding breathers theorem

The eigenvalue of $-\Delta + cA$ is not scale invariant. Suppose that α is invariant under scaling which is true in all of our examples. If we rescale a Riemannian metric g to εg by a positive constant ε, then

$$-\Delta_{\varepsilon g} + cA_{\varepsilon g} = \varepsilon^{-1}(-\Delta_g + cA_g),$$

and for the lowest eigenvalue we get $\lambda_{\varepsilon g} = \varepsilon^{-1}\lambda_g$. Thus the (nonnormalized) lowest eigenvalue only works in the steady case. Following [Perelman 2002] we define the scale invariant eigenvalue by

$$\bar{\lambda}_g := \lambda_g V_g^{2/n}, \tag{9-43}$$

where V denotes the volume of M.

In the following for simplicity of calculations we let $c = \frac{1}{4}$.

Proposition 9.1. *Suppose that $(M, g(t))$ is a solution to the abstract geometric flow (1-1) with α being scale invariant. Assume that Θ is nonnegative. Let $\lambda(t)$ be the lowest eigenvalue of $-\Delta + A/4$. Then whenever $\bar{\lambda}(t) \le 0$ one has $\bar{\lambda}'(t) \ge 0$.*

Proof. Recall that by (7-35) and choosing $\phi(t, y) = V^{-1/2}$ we have

$$\lambda(t) \le \frac{1}{4V}\int_M A\,dy.$$

When $\bar{\lambda}(t) \le 0$ we obtain

$$\begin{aligned}\bar{\lambda}'(t) &= \lambda'(t)V^{2/n} + \frac{2\lambda}{n}V^{n/2-1}\int_M (-A)\,dy\\ &\ge V^{n/2}\left(\lambda'(t) - \frac{8\lambda^2(t)}{n}\right)\\ &\ge \frac{V^{n/2}}{2}\left(\int_M \big(|\alpha - 2\nabla\nabla\log f|^2 + \Theta(2\nabla\log f)\big)f^2\,dy - \frac{16\lambda^2(t)}{n}\right),\end{aligned}$$

where f is the normalized positive eigenfunction corresponding to λ.

We observe that

$$|\alpha - 2\nabla\nabla\log f|^2 = \left|\alpha - 2\nabla\nabla\log f - \frac{1}{n}(A - 2\Delta\log f)g\right|^2 + \frac{1}{n}(A - 2\Delta\log f)^2.$$

Recall that f is the normalized eigenfunction and by Hölder's inequality we obtain

$$
\begin{aligned}
\int_M (A-2\Delta \log f)^2 f^2\,dy &= \int_M (A-2\Delta \log f)^2 f^2\,dy \int_M f^2\,dy \\
&\geq \left(\int_M (A-2\Delta \log f) f\cdot f\,dy\right)^2 \\
&= \left(\int_M Af^2+4|\nabla f|^2\,dy\right)^2 \\
&= 16\,\lambda^2(t).
\end{aligned}
\tag{9-44}
$$

Finally we have $\bar{\lambda}'(t)\geq 0$. ☐

If $\lambda(t)\leq 0$ we have in fact derived the inequality

(9-45)
$$
\begin{aligned}
\bar{\lambda}'(t) \geq \frac{V^{2/n}}{2}\left(\int_M \left(\left|\alpha-2\nabla\nabla \log f-\frac{1}{n}(A-2\Delta \log f)g\right|^2+\Theta(2\nabla \log f)\right)f^2\,dy\right) \\
+\frac{V^{2/n}}{2n}\left(\int_M (A-2\Delta \log f)^2 f^2\,dy-\left(\int_M (A-2\Delta \log f) f\cdot f\,dy\right)^2 dy\right).
\end{aligned}
$$

Now we may use (9-45) to rule out nontrivial expanding breathers.

Theorem 9.2. *Suppose that $(M, g(t))$ is a solution to the abstract geometric flow equation (1-1) with α being scale invariant. Assume that Θ is nonnegative. If $(M, g(t))$ is an expanding breather for $t_1 < t_2$, then it has to be a gradient soliton on (t_1, t_2) in the sense that*

$$\alpha-2\nabla\nabla \log f-\frac{4\lambda}{n}g=0$$

where f is the positive normalized eigenfunction corresponding to $\lambda(t)$. Moreover one has

$$\Theta(2\nabla \log f)=0.$$

Proof. Since $\bar{\lambda}$ is invariant under diffeomorphism and rescaling, we have $\bar{\lambda}(t_1)=\bar{\lambda}(t_2)$. Since $V(t_1)<V(t_2)$ there must be a time $t_0\in(t_1,t_2)$ such that $V'(t_0)\geq 0$. Hence

$$
\begin{aligned}
\lambda(t_0) &\leq \frac{1}{4V(t_0)}\int_M A(t_0)\,dy \\
&= -\frac{1}{4V(t_0)}V'(t_0) \\
&\leq 0.
\end{aligned}
$$

Proposition 9.1 then implies $\bar{\lambda}(t_1) \le \bar{\lambda}(t_0) \le 0$. Thus, on the whole interval $[t_1, t_2]$, the function $\bar{\lambda}(t)$ is nonpositive increasing and equals at the end points. This means that the right side of (9-45) vanishes. In particular, the second line of (9-45) being zero means that equality holds in Hölder's inequality (9-44). Thus $A - 2\Delta \log f$ must be a spatial constant which is $4\lambda(t)$ because f is a normalized eigenfunction corresponding to $\lambda(t)$. The vanishing of the first line of (9-45) means that

$$\alpha - 2\nabla\nabla \log f - \frac{4\lambda}{n} g = 0, \quad \Theta(2\nabla \log f) = 0. \qquad \square$$

References

[Akbar and Woolgar 2009] M. M. Akbar and E. Woolgar, "Ricci solitons and Einstein-scalar field theory", *Classical Quantum Gravity* **26**:5 (2009), Article ID 055015, 14. MR 2010i:53121 Zbl 1160.83351

[Cao 2007] X. Cao, "Eigenvalues of $(-\Delta + \frac{R}{2})$ on manifolds with nonnegative curvature operator", *Math. Ann.* **337**:2 (2007), 435–441. MR 2007g:53071 Zbl 1105.53051

[Cao 2008] X. Cao, "First eigenvalues of geometric operators under the Ricci flow", *Proc. Amer. Math. Soc.* **136**:11 (2008), 4075–4078. MR 2009f:53098 Zbl 1166.58007

[Ecker et al. 2008] K. Ecker, D. Knopf, L. Ni, and P. Topping, "Local monotonicity and mean value formulas for evolving Riemannian manifolds", *J. Reine Angew. Math.* **616** (2008), 89–130. MR 2008k:53137 Zbl 1170.35050

[Feldman et al. 2005] M. Feldman, T. Ilmanen, and L. Ni, "Entropy and reduced distance for Ricci expanders", *J. Geom. Anal.* **15**:1 (2005), 49–62. MR 2006b:53091 Zbl 1071.53040

[Kleiner and Lott 2008] B. Kleiner and J. Lott, "Notes on Perelman's papers", *Geom. Topol.* **12**:5 (2008), 2587–2855. MR 2010h:53098 Zbl 1204.53033

[Li 2007] J.-F. Li, "Eigenvalues and energy functionals with monotonicity formulae under Ricci flow", *Math. Ann.* **338**:4 (2007), 927–946. MR 2008c:53068 Zbl 1127.53059

[List 2008] B. List, "Evolution of an extended Ricci flow system", *Comm. Anal. Geom.* **16**:5 (2008), 1007–1048. MR 2010i:53126 Zbl 1166.53044

[Müller 2010] R. Müller, "Monotone volume formulas for geometric flows", *J. Reine Angew. Math.* **643** (2010), 39–57. MR 2011k:53086 Zbl 1204.53054

[Müller 2012] R. Müller, "Ricci flow coupled with harmonic map flow", *Ann. Sci. Éc. Norm. Supér.* (4) **45**:1 (2012), 101–142. MR 2961788 Zbl 1247.53082

[Ni 2004a] L. Ni, "The entropy formula for linear heat equation", *J. Geom. Anal.* **14**:1 (2004), 87–100. MR 2004m:53118a Zbl 1044.58030

[Ni 2004b] L. Ni, "Addenda to: The entropy formula for linear heat equation", *J. Geom. Anal.* **14**:2 (2004), 369–374. MR 2004m:53118b Zbl 1062.58028

[Perelman 2002] G. Perelman, "The entropy formula for the Ricci flow and its geometric applications", preprint, 2002. Zbl 1130.53001 arXiv math/0211159v1

[Reed and Simon 1978] M. Reed and B. Simon, *Methods of modern mathematical physics, IV: Analysis of operators*, Academic, New York, 1978. MR 58 #12429c Zbl 0401.47001

Received January 30, 2012. Revised October 23, 2012.

HONGXIN GUO
SCHOOL OF MATHEMATICS AND INFORMATION SCIENCE
WENZHOU UNIVERSITY
CHASHAN UNIVERSITY TOWN
WENZHOU, ZHEJIANG 325035
CHINA
guo@wzu.edu.cn

ROBERT PHILIPOWSKI
MATHEMATICS RESEARCH UNIT
UNIVERSITY OF LUXEMBOURG
6, RUE RICHARD COUDENHOVE-KALERGI
L-1359 LUXEMBOURG
LUXEMBOURG
robert.philipowski@uni.lu

ANTON THALMAIER
MATHEMATICS RESEARCH UNIT
UNIVERSITY OF LUXEMBOURG
6, RUE RICHARD COUDENHOVE-KALERGI
L-1359 LUXEMBOURG
LUXEMBOURG
anton.thalmaier@uni.lu
http://math.uni.lu/thalmaier/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.83

POLES OF CERTAIN RESIDUAL EISENSTEIN SERIES OF CLASSICAL GROUPS

DIHUA JIANG, BAIYING LIU AND LEI ZHANG

In memory of Steve Rallis

We study the location of possible poles of a family of residual Eisenstein series on classical groups. Special types of residues of those Eisenstein series were used as key ingredients in the automorphic descent constructions of Ginzburg, Rallis and Soudry and in the refined constructions of Ginzburg, Jiang and Soudry. We study the conditions for the existence of other possible poles of those Eisenstein series and determine the possible Arthur parameters for the residual representations if they exist. Further properties of those residual representations and their applications to automorphic constructions will be considered in our future work.

1. Introduction

Automorphic descent constructions of Ginzburg, Rallis and Soudry [Ginzburg et al. 2011] produce the inverse of the Langlands functorial transfers from classical groups to the general linear groups. More recently, the extensions of those constructions to produce endoscopy transfers for classical groups were considered in [Ginzburg 2008; Ginzburg et al. 2012; Jiang 2011; 2012]. The key ingredient in these constructions is to use certain Fourier coefficients of special types of residues of certain residual Eisenstein series as kernel functions in the corresponding integral transforms. In order to explore the possibility of more general constructions, in this paper we start to consider other possible poles and residues of these and more general residual Eisenstein series for classical groups.

1A. ***Classical groups.*** Let F be a number field and let E be a quadratic extension of F whose Galois group is denoted by $\Gamma_{E/F} = \{1, \iota\}$. Denote by $\mathbb{A} = \mathbb{A}_F$ the ring of adeles of F.

The classical groups considered in this paper, denoted by G_n, are the F-quasisplit unitary groups U_{2n} and U_{2n+1} of hermitian type, the F-split special orthogonal

The work of Dihua Jiang is supported in part by NSF DMS-1001672.
MSC2010: primary 11F70, 22E50; secondary 22E55, 11F72.
Keywords: residual representations, Arthur parameters, Eisenstein series.

group SO_{2n+1} and the symplectic group Sp_{2n}, and the F-quasisplit even special orthogonal group SO_{2n}. Define the number field F' as F if G_n is not a unitary group and as E if it is. Denote by $\mathrm{R}_{F'/F}(\mathrm{GL}_n)$ the Weil restriction of the GL_n from F' to F.

We try to follow closely the notation introduced in [Mœglin and Waldspurger 1995]. Since the groups considered in this paper are quasisplit, we fix a standard Borel subgroup $P_0 = M_0N_0$ of G_n that is realized in the upper-triangular matrices in a chosen realization of the classical group in matrices [Ginzburg et al. 2011]. Let T_0 be the maximal split torus of the center of M_0 that defines the root system $R(T_0, G_n)$ with the given positive roots $R^+(T_0, G_n)$ and the set Δ_0 of simple roots. Let $P = MN$ be a standard parabolic subgroup of G_n (containing P_0) and let T_M be the maximal split torus in the center of M. The set of restricted roots is denoted by $R(T_M, G_n)$. We define $R^+(T_M, G_n)$ and Δ_M accordingly.

Furthermore, we define $X_M = X_M^{G_n}$ to be the group of all continuous homomorphisms from $M(\mathbb{A})$ into $\mathbb{C}^\times$ that are trivial on $M(\mathbb{A})^1$. Then following page 6 of [Mœglin and Waldspurger 1995] for the explicit realization of X_M, define the real part of X_M, which is denoted by $\mathrm{Re}\, X_M$.

1B. ***Discrete spectrum of*** $\mathbf{GL}_{ab}$**.** Let τ be an irreducible unitary cuspidal automorphic representation of $\mathrm{GL}_a(\mathbb{A})$. Take the standard parabolic subgroup $Q_{a^b} = L_{a^b}U_{a^b}$ of GL_{ab}, whose Levi subgroup L_{a^b} is isomorphic to $\mathrm{GL}_a^{\times b}$. Then $\pi = \tau^{\otimes b}$ is an irreducible unitary cuspidal automorphic representation of $L_{a^b}(\mathbb{A})$. As in Section II.1.5 of [Mœglin and Waldspurger 1995], denote by $\mathfrak{P}$ the $X_{L_{a^b}}^{\mathrm{GL}_{ab}}$-orbit of the cuspidal datum (L_{a^b}, π). For an automorphic function

$$\phi_\pi \in A\big(L_{a^b}(F)U_{a^b}(\mathbb{A})\backslash \mathrm{GL}_{ab}(\mathbb{A})\big)_\pi,$$

denote by $\phi_{\pi\otimes\lambda} = \lambda\phi_\pi$ the element $\lambda \circ m_Q \phi_\pi$ for $\lambda \in X_{L_{a^b}}^{\mathrm{GL}_{ab}}$. Here the mapping m_Q from $\mathrm{GL}_{ab}(\mathbb{A})$ to $L_{a^b}(\mathbb{A})^1\backslash L_{a^b}(\mathbb{A})$ is as defined on page 7 of [Mœglin and Waldspurger 1995] by means of the Langlands decomposition with respect to $Q_{a^b}(\mathbb{A})$ and the standard maximal compact subgroup of $\mathrm{GL}_{ab}(\mathbb{A})$. An Eisenstein series attached to $\phi_{\pi\otimes\lambda}$ is defined by

$$E(\phi_{\pi\otimes\lambda}, \pi\otimes\lambda)(g) := \sum_{\gamma\in Q_{a^b}(F)\backslash \mathrm{GL}_{ab}(F)} \lambda\phi_\pi(\gamma g).$$

It converges absolutely for λ in the cone

$$\big\{\lambda \in \mathrm{Re}\, X_{L_{a^b}}^{\mathrm{GL}_{ab}} \,\big|\, \langle\lambda, \tilde{\alpha}\rangle > \langle\rho_{Q_{a^b}}, \tilde{\alpha}\rangle \text{ for all } \alpha \in R^+(T_{L_{a^b}}, \mathrm{GL}_{ab})\big\},$$

and converges uniformly for g in a compact set and λ in a neighborhood of 0 in $X_{L_{a^b}}^{\mathrm{GL}_{ab}}$. The general theory of Langlands [1976; Mœglin and Waldspurger 1995] asserts that it has meromorphic continuation to the whole parameter space $X_{L_{a^b}}^{\mathrm{GL}_{ab}}$

and satisfies the standard functional equations in terms of the relevant intertwining operators.

Take $\Lambda_b = \big((b-1)/2, (b-3)/2, \dots, (1-b)/2\big) \in \operatorname{Re} X_{L_{a^b}}^{\mathrm{GL}_{ab}}$ and define the iterated residue

$$\Delta(\tau, b)(\phi_\pi)(g) := \operatorname{Res}_{\Lambda_b}^{\mathfrak{P}_{a^b}} E(\phi_{\pi\otimes\lambda}, \pi\otimes\lambda)(g).$$

It follows from [Mœglin and Waldspurger 1989] that $\Delta(\tau, b)(\phi_\pi)(g)$ is a square-integrable automorphic function of $\mathrm{GL}_{ab}(\mathbb{A})$, or more precisely, that it defines the $\mathrm{GL}_{ab}(\mathbb{A})$-equivariant homomorphism

$$\Delta(\tau, b)\colon A\big(L_{a^b}(F)U_{a^b}(\mathbb{A})\backslash \mathrm{GL}_{ab}(\mathbb{A})\big)_\pi \to L^2_{\mathrm{disc}}\big(\mathrm{GL}_{ab}(F)\backslash \mathrm{GL}_{ab}(\mathbb{A})\big)_{\omega_\tau^b}.$$

The image is an irreducible subspace of $L^2_{\mathrm{disc}}(\mathrm{GL}_{ab}(F)\backslash \mathrm{GL}_{ab}(\mathbb{A}))_{\omega_\tau^b}$, which is denoted also by $\mathscr{E}_{(\tau,b)}$, and is usually called the *Speh residual representation*. Mœglin and Waldspurger proved that all noncuspidal automorphic representations occurring in the discrete spectrum of $\mathrm{GL}_{ab}(\mathbb{A})$ are of this type.

Theorem 1.1 [Mœglin and Waldspurger 1989]. *As b ranges over the divisors of n, with $n = ab$ and $b > 1$, and τ ranges over the irreducible unitary cuspidal automorphic representations of GL_a, with $\omega_\tau^b = \chi$, the residual representations $\mathscr{E}_{(\tau,b)}$ generated by the corresponding residues $\Delta(\tau, b)(\phi_\pi)$ span the residual spectrum $L^2_{\mathrm{res}}(\mathrm{GL}_n(F)\backslash \mathrm{GL}_n(\mathbb{A}))_\chi$, where χ is a unitary central character of $\mathrm{GL}_n(\mathbb{A})$.*

1C. ***Main results.*** We consider a family of residual Eisenstein series on $G_n(\mathbb{A})$. For a partition $n = r + m$, take the standard maximal parabolic subgroup $P_r = M_r N_r$ of G_n, whose Levi subgroup M_r is isomorphic to $\mathrm{R}_{F'/F}(\mathrm{GL}_r) \times G_m$. For any $g \in \mathrm{R}_{F'/F}(\mathrm{GL}_r)$, define $\hat{g} = w_r g^t w_r$ or $w_r \iota(g)^t w_r$ in the case of unitary groups, where w_r is the antidiagonal symmetric matrix defined inductively by

$$\begin{pmatrix} 0 & 1 \\ w_{r-1} & 0 \end{pmatrix}$$

and $\iota \in \Gamma_{E/F} = \{1, \iota\}$. Then each element $g \in M_r$ is of type $\mathrm{diag}\{t, h, \hat{t}^{-1}\}$, with $t \in \mathrm{R}_{F'/F}(\mathrm{GL}_r)$ and $h \in G_m$. Since P_r is maximal, the space of characters $X_{M_r}^{G_n}$ is one-dimensional. Using the normalization in [Shahidi 2010], it is identified with $\mathbb{C}$ by $s \mapsto \lambda_s$.

For simplicity, we state here only our results for the case of $m > 0$, and refer to Section 5 for the case of $m = 0$.

Let σ be an irreducible generic cuspidal automorphic representation of $G_m(\mathbb{A})$. Write $r = ab$. Let $\phi \in A(N_{ab}(\mathbb{A})M_{ab}(F)\backslash G_n(\mathbb{A}))_{\Delta(\tau,b)\otimes\sigma}$. Following [Langlands 1976; Mœglin and Waldspurger 1995], an Eisenstein series is defined by

$$E^n_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s) = E(\phi_{\Delta\otimes\sigma}, s) = \sum_{\gamma \in P_{ab}(F)\backslash G_n(F)} \lambda_s \phi(\gamma g).$$

It converges absolutely for the real part of s large and has meromorphic continuation to the whole complex plane $\mathbb{C}$.

The objective of this paper is to determine the location of possible poles (at $\mathrm{Re}(s) \geq 0$) of this family of residual Eisenstein series, or more precisely the normalized Eisenstein series, and basic properties of the corresponding residual representations. We take the expected normalizing factor $\beta_{b,\tau,\sigma}(s)$ of the Langlands–Shahidi type, which is given by a product of relevant automorphic L-functions:

(1-1)
$$\beta_{b,\tau,\sigma}(s) := L\Big(s + \frac{b+1}{2}, \tau \times \sigma\Big) \prod_{i=1}^{\lceil b/2 \rceil} L(e_{b,i}(s)+1, \tau, \rho) \prod_{i=1}^{\lfloor b/2 \rfloor} L(e_{b,i}(s), \tau, \rho^-),$$

where $e_{b,i}(s) := 2s + b + 1 - 2i$, and ρ and ρ^- are defined as

$$\rho := \begin{cases} \mathrm{Asai} & \text{if } G_n = \mathrm{U}_{2n}, \\ \mathrm{Asai} \otimes \delta & \text{if } G_n = \mathrm{U}_{2n+1}, \\ \mathrm{Sym}^2 & \text{if } G_n = \mathrm{SO}_{2n+1}, \\ \Lambda^2 & \text{if } G_n = \mathrm{Sp}_{2n} \text{ or } \mathrm{SO}_{2n}, \end{cases} \tag{1-2}$$

$$\rho^- := \begin{cases} \mathrm{Asai} \otimes \delta & \text{if } G_n = \mathrm{U}_{2n}, \\ \mathrm{Asai} & \text{if } G_n = \mathrm{U}_{2n+1}, \\ \Lambda^2 & \text{if } G_n = \mathrm{SO}_{2n+1}, \\ \mathrm{Sym}^2 & \text{if } G_n = \mathrm{Sp}_{2n} \text{ or } \mathrm{SO}_{2n}. \end{cases} \tag{1-3}$$

For unitary groups, "Asai" is the Asai representation of the L-group of $R_{E/F}(\mathrm{GL}_a)$ and δ is the character associated to the quadratic extension E/F via class field theory. For symplectic or orthogonal groups, Sym^2 and Λ^2 denote the symmetric and exterior second powers of the standard representation of $\mathrm{GL}_a(\mathbb{C})$, respectively. In addition, we have the following identities [Ginzburg et al. 2011, Remark (3), page 21]:

$$L(s, \tau \times \tau^*) = L(s, \tau, \rho) L(s, \tau, \rho^-),$$

where $\tau^* = \tau$ if $F' = F$ and $\tau^* = \tau^\iota$ if $F' = E$, where the involution ι is the nontrivial element in the Galois group $\Gamma_{E/F}$.

We use the function $\beta_{b,\tau,\sigma}(s)$ to normalize the Eisenstein series by

$$E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s) := \beta_{b,\tau,\sigma}(s) E^n_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s). \tag{1-4}$$

In order to determine the location of the poles of $E^*(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$, we need to consider four cases:

(1) $L(s, \tau, \rho)$ has a pole at $s = 1$, and $L\big(\frac{1}{2}, \tau \times \sigma\big) \neq 0$;

(2) $L(s, \tau, \rho)$ has a pole at $s = 1$, and $L\left(\frac{1}{2}, \tau \times \sigma\right) = 0$;

(3) $L(s, \tau, \rho^-)$ has a pole at $s = 1$, and $L(s, \tau \times \sigma)$ has a pole at $s = 1$;

(4) $L(s, \tau, \rho^-)$ has a pole at $s = 1$, and $L(s, \tau \times \sigma)$ is holomorphic at $s = 1$.

We define the sets of possible poles according to the four cases:

$$X^+_{b,\tau,\sigma} := \begin{cases} \left\{\hat{0}, \dots, \dfrac{b-2}{2}, \dfrac{b}{2}\right\} & \text{in Case (1);} \\ \left\{\hat{0}, \dots, \dfrac{b-4}{2}, \dfrac{b-2}{2}\right\} & \text{in Case (2);} \\ \left\{\hat{0}, \dots, \dfrac{b-1}{2}, \dfrac{b+1}{2}\right\} & \text{in Case (3);} \\ \left\{\hat{0}, \dots, \dfrac{b-3}{2}, \dfrac{b-1}{2}\right\} & \text{in Case (4).} \end{cases}$$

When $b = 1$ or 2, the set $X^+_{b,\tau,\sigma}$ is empty for Case (2), and when $b = 1$, the set $X^+_{b,\tau,\sigma}$ is empty for Case (4). Note that we omit 0 in the set $X^+_{b,\tau,\sigma}$, since the normalized Eisenstein series $E^*(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic at $s = 0$ (Corollary 4.3).

Theorem 1.2. *Assume that G_n is either the symplectic group or the F-quasisplit special orthogonal group, and assume that $m > 0$. Let σ be an irreducible generic cuspidal automorphic representation of $G_m(\mathbb{A})$, and let τ be an irreducible unitary self-dual cuspidal automorphic representation of $\mathrm{GL}_r(\mathbb{A})$. The normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic for $\mathrm{Re}(s) \geq 0$ except at $s = s_0 \in X^+_{b,\tau,\sigma}$, where it has possibly at most simple poles.*

This is a consequence of Proposition 4.1, Corollary 4.3, and Theorems 4.5 and 5.2.

The proof uses an induction formula (Proposition 3.2) for the constant term of $E^*(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ along the standard maximal parabolic subgroup P_a. This formula, which extends a similar one studied in [Jiang 1998], is proved in Section 3, with the unnormalized version proved in Section 2 (Proposition 2.3); it uses the Arthur classification [Arthur 2013] for the discrete spectrum of the classical groups. This yields more explicit information about the residual representations. A special case of Sp_{2n} was treated in [Brenner 2009]. We note that there are some mistakes in the arguments used there, and we have corrected them along the way in our discussion.

We remark that the calculations in both Sections 2 and 3 work also for F-quasisplit unitary groups, and the results there cover the case when G_n is either U_{2n} or U_{2n+1}.

In the proof of Theorem 1.2, the case of $m > 0$ is treated in Section 4 and the case of $m = 0$ is briefly discussed in Section 5. This makes the discussion clearer and the formulas involved easier to present. By using the induction formula (Proposition 3.2), one reduces the proof to showing that the normalized Eisenstein series $E^*(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic at $0 \le \mathrm{Re}(s) < \frac{1}{2}$, which is proved in Corollary 4.3 and Proposition 4.4. The proof of this result uses the result of Arthur [2013, Corollary 7.3.5] on behavior at $s = 0$ of the normalized intertwining operators, and on classification of the discrete spectrum. We thank James Arthur for his careful explanation of this issue. Since the results in [Arthur 2013] for the case of unitary groups are now proved in [Mok 2012], the proof of Theorem 1.2 also works for F-quasisplit unitary groups.

Another issue is to consider the possible poles of the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ at $\mathrm{Re}(s) < 0$ by the standard functional equation. This needs sufficient properties of the involved standard intertwining operator and the local Plancherel measures in this setting. We will leave this for our future consideration.

There is one more issue in extending Theorem 1.2 to cover the case when σ is tempered, but nongeneric. We need to normalize the intertwining operators involved in the calculation of the induction formula so that they are holomorphic and nonzero for $\mathrm{Re}(s) > 0$ at every local place. Following the work of Arthur [2013], one is able to define these local L-functions at all local places. According to Mœglin [2010], over p-adic local fields, for the tempered local L-packets, the normalization of these intertwining operators by the Langlands–Shahidi local factors yields the required properties of the normalized intertwining operators. It seems that at archimedean local places, this may need more work, and we decide to consider this technical issue in the future. Hence we still restrict Theorem 1.2 to the generic case in this paper, which is enough for the current applications to our work in progress on constructions of certain types of endoscopy transfers for classical groups [Jiang 2011; 2012].

In Section 4 we prove Theorem 1.2 for the case when $m > 0$, and in Section 5 we prove Theorem 1.2 for the case when $m = 0$. In the last section, we will discuss the conditions for the existence of poles of the normalized Eisenstein series $E^*(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ at $s_0 \in X^+_{b,\tau,\sigma}$ and determine the possible Arthur parameters for these residual representations of $G_n(\mathbb{A})$, which are generated by the residues at $s_0 \in X^+_{b,\tau,\sigma}$, respectively, and are square-integrable.

2. An induction formula

In this section, we take G_n to be one of the following classical groups: the F-quasisplit unitary groups U_{2n} and U_{2n+1}, the F-split odd special orthogonal group SO_{2n+1}, the symplectic group Sp_{2n}, and the F-quasisplit even special orthogonal group SO_{2n}.

Let σ be an irreducible cuspidal automorphic representation of $G_m(\mathbb{A})$, without assuming its genericity. From the Langlands theory of Eisenstein series, the possible poles of an Eisenstein series are determined by its constant terms. For the residual

Eisenstein series $E^n_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$, the general formula for constant terms along parabolic subgroups are given in [Mœglin and Waldspurger 1995, Section II.1.7], for instance. In this section, we investigate the constant term of $E^n_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ along the maximal parabolic subgroup P_a (as given in Section 1), which leads to an induction formula. This extends the formula in [Jiang 1998] to this more general setting. On the way of our calculations, we also correct some technical mistakes in [Brenner 2009], which treated a special family of residual Eisenstein series of $\mathrm{Sp}_{2n}(\mathbb{A})$.

In the explicit calculation for the induction formula, we may set P^n_r for the standard maximal parabolic subgroup P_r of $G = G_n$. We denote by Q or $Q^{ab}_{a,a(b-1)}$ a parabolic subgroup of GL_{ab} with Levi subgroup isomorphic to $\mathrm{GL}_a \times \mathrm{GL}_{a(b-1)}$.

2A. ***Constant terms of Eisenstein series.*** Here we calculate the constant term of $E^n_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ along the maximal parabolic subgroup $P^n_a = P_a$, which is defined by

$$E_{P_a}(\phi_{\Delta\otimes\sigma}, s)(g) = \int_{N_a(F)\backslash N_a(\mathbb{A})} E(\phi_{\Delta\otimes\sigma}, s)(ng)\,\mathrm{d}n.$$

Assume that $\mathrm{Re}(s)$ is large. After unfolding the Eisenstein series, we obtain

$$\text{(2-1)}\quad E_{P_a}(\phi_{\Delta\otimes\sigma}, s)(g) = \sum_{w^{-1}\in P_{ab}\backslash G/P_a}\ \sum_{\gamma\in M^w_a(F)\backslash M_a(F)} \int_{[N^w_a]}\int_{N_{a,w}(\mathbb{A})} \lambda_s\phi(w^{-1}\gamma n'n''g)\,\mathrm{d}n'\,\mathrm{d}n'',$$

where we define $M^w_a := wP_{ab}w^{-1}\cap M_a$ and $N^w_a := wP_{ab}w^{-1}\cap N_a$ and $[N^w_a] := N^w_a(F)\backslash N^w_a(\mathbb{A})$. Note that the unipotent radical N_a can be decomposed as a product $N_{a,w}N^w_a$, where $N_{a,w}$ satisfies $N_{a,w}\cap N^w_a = \{1\}$ and $N_a = N_{a,w}N^w_a = N^w_a N_{a,w}$.

For the first summation in (2-1), we consider the generalized Bruhat decomposition $P_{ab}\backslash G/P_a$. As in [Shahidi 2010, Lemma 4.2.1], the representative w^{-1} of the double coset $P_{ab}w^{-1}P_a$ is chosen to have the *minimal length*. Following the explicit calculations done in [Ginzburg et al. 2011, Chapter 4], it is not hard to figure out that by the cuspidal support of the Eisenstein series, all terms vanish except the two double cosets, whose representatives are given by $w = \mathrm{Id}$ and

$$w = \omega = (-1)^a \begin{pmatrix} 0 & 0 & 0 & I_a & 0 \\ I_{a(b-1)} & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{a(b-1)} \\ 0 & \pm I_a & 0 & 0 & 0 \end{pmatrix},$$

with ω^{-1} being in the open cell. Here we use $(-1)^a$ and $\pm$ to make sure that ω belongs to G_n. Define, for $w = \mathrm{Id}$ or ω,

$$E_{P_a}(\phi_{\Delta\otimes\sigma}, s)_w = \sum_{\gamma\in M_a^w(F)\backslash M_a(F)} \int_{[N_a^w]} \int_{N_{a,w}(\mathbb{A})} \lambda_s\phi(w^{-1}\gamma n'n''g)\,\mathrm{d}n'\,\mathrm{d}n''.$$

Then the constant term is expressed as

$$E_{P_a}(\phi_{\Delta\otimes\sigma}, s) = E_{P_a}(\phi_{\Delta\otimes\sigma}, s)_{\mathrm{Id}} + E_{P_a}(\phi_{\Delta\otimes\sigma}, s)_{\omega}. \tag{2-2}$$

We will calculate each of these two terms in the following two subsections.

2B. Id-*term*. Write

$$n(X, Y, Z, W) = \begin{pmatrix} I_a & X & Y & Z & W \\ & I_{a(b-1)} & & & Z' \\ & & I & & Y' \\ & & & I_{a(b-1)} & X' \\ & & & & I_a \end{pmatrix} \in N_a,$$

where X', Y' and Z' are uniquely determined by X, Y and Z. Note that $P_{ab} \cap M_a\backslash M_a \cong P^{n-a}_{a(b-1)}\backslash G_{n-a}$. The Id-term of the constant term is

$$E_{P_a}(\phi_{\Delta\otimes\sigma}, s)_{\mathrm{Id}}(g) = \sum_{\gamma\in P^{n-a}_{a(b-1)}(F)\backslash G_{n-a}(F)} \int_{[N_a]} \lambda_s\phi(\gamma ng)\,\mathrm{d}n, \tag{2-3}$$

where $[N_a] := N_a(F)\backslash N_a(\mathbb{A})$. The integral can be calculated as follows:

$$\begin{aligned}\int_{[N_a]} \lambda_s\phi(\gamma ng)\,\mathrm{d}n &= \int_{[N_a]} \lambda_s\phi(n\gamma g)\,\mathrm{d}n \\ &= \int_{[M_{a\times a(b-1)}]} \int_{[N_{ab}\cap N_a]} \lambda_s\phi(n'n(X)\gamma g)\,\mathrm{d}n'\,\mathrm{d}X \\ &= \int_{[M_{a\times a(b-1)}]} \lambda_s\phi(n(X)\gamma g)\,\mathrm{d}X.\end{aligned}$$

Here $[Z] := Z(F)\backslash Z(\mathbb{A})$ for $Z = N_a$, $M_{a\times a(b-1)}$, and $N_{ab}\cap N_a$, respectively. We denote by $n(X)$ the element $n(X, 0, 0, 0)$ with $X \in M_{a\times a(b-1)}$.

Let us understand the last integral

$$\int_{[M_{a\times a(b-1)}]} \phi(n(X)g)\,\mathrm{d}X. \tag{2-4}$$

Recall that the Levi subgroup M_{ab} is isomorphic to $\mathrm{R}_{F'/F}(\mathrm{GL}_{ab}) \times G_m$. We denote its elements by (x, h) with $x \in \mathrm{R}_{F'/F}(\mathrm{GL}_{ab})$ and $h \in G_m$. We fix $g \in G_n(\mathbb{A})$. Then the function

$$x \mapsto \phi((x, 1)g)$$

is an automorphic function in the space of the residual representation $\mathcal{E}_{(\tau,b)}$ of $\mathrm{GL}_{ab}(\mathbb{A}_{F'})$. Consider the standard maximal parabolic subgroup

$$Q_{a,a(b-1)} = L_{a,a(b-1)} U_{a,a(b-1)}$$

of GL_{ab} associated to the partition $ab = a + a(b-1)$. Then the integral (2-4) is the constant term of $\phi((x, 1)g)$ (as an automorphic form in x) along the maximal parabolic subgroup $Q_{a,a(b-1)}$, which is denoted by $\phi_{Q_{a,a(b-1)}}$.

Let $P_{a,a(b-1)}$ be a standard parabolic subgroup of G_n whose Levi subgroup $M_{a,a(b-1)}$ is isomorphic to

$$\mathrm{R}_{F'/F}\mathrm{GL}_a \times \mathrm{R}_{F'/F}\mathrm{GL}_{a(b-1)} \times G_m$$

and whose unipotent radical is $N := N_{a,a(b-1)}$. We denote by (t, r, h) the element $\mathrm{diag}(t, r, h, \hat{r}^{-1}, \hat{t}^{-1})$ in $M_{a,a(b-1)}(\mathbb{A})$.

Lemma 2.1. *The constant term $\lambda_s \phi_{Q_{a,a(b-1)}}$ belongs to the space*

$$A\big(N_{a,a(b-1)}(\mathbb{A}) M_{a,a(b-1)}(F)\backslash G_n(\mathbb{A})\big)_{\tau|\cdot|_{F'}^{s-(b-1)/2} \otimes \Delta(\tau,b-1)|\cdot|_{F'}^{s+1/2} \otimes \sigma}.$$

Here $|\cdot|_{F'} = |\det|_{\mathbb{A}_{F'}}$; and F' is E if G_n is unitary, and is F otherwise.

Proof. Let $\boldsymbol{K} = \Pi_v K_v$ be the standard choice of maximal compact subgroup of $G_n(\mathbb{A})$ such that the Iwasawa decomposition

$$G_n(\mathbb{A}) = P_{a,a(b-1)}(\mathbb{A})\boldsymbol{K}$$

holds. It suffices to show that for all $k \in \boldsymbol{K}$, the constant term $\lambda_s \phi_{Q_{a,a(b-1)}}((t, r, h)k)$ belongs to the space of automorphic forms

$$A\big(M_{a,a(b-1)}(F)\backslash M_{a,a(b-1)}(\mathbb{A})\big)_{\tau|\cdot|_{F'}^{s-(b-1)/2} \otimes \Delta(\tau,b-1)|\cdot|_{F'}^{s+1/2} \otimes \sigma},$$

where $t \in \mathrm{GL}_a(\mathbb{A}_{F'})$, $r \in \mathrm{GL}_{a(b-1)}(\mathbb{A}_{F'})$ and $h \in G_m(\mathbb{A})$.

By the definition (2-4), we have

$$\phi_{Q_{a,a(b-1)}}((t, r, h)k) = \int_{[M_{a\times a(b-1)}]} \phi(n(X)(t, r, h)k)\,\mathrm{d}X.$$

Since the function $\phi_k(m) := m^{-\rho_{P_{ab}}} \phi(mk)$, for $m \in M_{ab}(\mathbb{A})$, is an automorphic form in $A(M_{ab}(F)\backslash M_{ab}(\mathbb{A}))_{\Delta(\tau,b)\otimes\sigma}$ for all $k \in \boldsymbol{K}$, without loss of generality, we can assume that

$$\phi_k((t, r, h)) = \phi_{k,\Delta(\tau,b)}((t, r)) \otimes \phi_{k,\sigma}(h),$$

where the function $\phi_{k,\Delta(\tau,b)} \in A(\mathrm{GL}_{ab}(F')\backslash \mathrm{GL}_{ab}(\mathbb{A}_{F'}))_{\Delta(\tau,b)}$ and the function $\phi_{k,\sigma} \in A(G_m(F)\backslash G_m(\mathbb{A}))_\sigma$. Therefore, we obtain

$$\phi_{Q_{a,a(b-1)}}((t, r, h)k) = (\phi_{k,\Delta(\tau,b)})_{Q_{a,a(b-1)}}((t, r)) \otimes \phi_{k,\sigma}(h),$$

where $(\phi_{k,\Delta(\tau,b)})_{Q_{a,a(b-1)}}$ is the constant term of $\phi_{k,\Delta(\tau,b)}$ along the parabolic subgroup $Q_{a,a(b-1)}$ of GL_{ab}.

By [Jiang and Liu 2012, Lemma 4.1], the constant term $(\phi_{k,\Delta(\tau,b)})_{Q_{a,a(b-1)}}$ belongs to the space

$$A\big(U_{a,a(b-1)}(\mathbb{A})L_{a,a(b-1)}(F)\backslash \mathrm{GL}_{ab}(\mathbb{A}_{F'})\big)_{|\cdot|_{F'}^{-(b-1)/2}\tau\otimes|\cdot|_{F'}^{1/2}\Delta(\tau,b-1)}.$$

It follows that the function $\lambda_s\phi_{Q_{a,a(b-1)}}(g)$ belongs to the space

$$A\big(N_{a,a(b-1)}(\mathbb{A})M_{a,a(b-1)}(F)\backslash G_n(\mathbb{A})\big)_{\tau|\cdot|_{F'}^{s-(b-1)/2}\otimes\Delta(\tau,b-1)|\cdot|_{F'}^{s+1/2}\otimes\sigma}. \qquad \square$$

According to Lemma 2.1, we restrict the Id-term $E^n_{ab,P^n_a}(\phi_{\Delta\otimes\sigma},s)_{\mathrm{Id}}$ to the subgroup $I_a\times G_{n-a}(\mathbb{A})$ of the Levi subgroup $\mathrm{GL}_a(\mathbb{A}_{F'})\times G_{n-a}(\mathbb{A})$ and obtain

(2-5)
$$\begin{aligned} E^n_{ab,P^n_a}(\phi_{\Delta\otimes\sigma},s)_{\mathrm{Id}}((I_a,h)) &= \sum_{\gamma\in P^{n-a}_{a(b-1)}(F)\backslash G_{n-a}(F)} \lambda_s\phi_{Q^{ab}_{a,a(b-1)}}\big(\mathrm{diag}(I_a,\gamma h,I_a)\big)\\ &= E^{n-a}_{a(b-1)}\big(\lambda_{-1/2}(i^*_{n-a}\phi_Q)_{\Delta(\tau,b-1)\otimes\sigma},s+\tfrac{1}{2}\big)(h),\end{aligned}$$

where $|\cdot|_{F'}:=|\cdot|_{\mathbb{A}_{F'}}$ and the restriction $i^*_{n-a}\phi_Q=i^*_{n-a}\phi_{Q^{ab}_{a,a(b-1)}}$ to $G_{n-a}(\mathbb{A})$ is an automorphic function in the space

$$A\big(N^{n-a}_{a(b-1)}(\mathbb{A})M^{n-a}_{a(b-1)}(F)\backslash G_{n-a}(\mathbb{A})\big)_{\Delta(\tau,b-1)|\cdot|_{F'}^{1/2}\otimes\sigma}.$$

2C. ***ω-term.*** It is easy to see that

$$N^{\omega}_a=\big\{n(0,0,Z,0)\mid Z\in M_{a\times a(b-1)}\big\}.$$

We denote by $\tilde{n}(Z)$ the element $n(0,0,Z,0)$. The coset $M^{\omega}_a(F)\backslash M_a(F)$ is isomorphic to $P_{a(b-1)}(F)\backslash G_{n-a}(F)$. Therefore, we have

$$\begin{aligned} &E_{P_a}(\phi_{\Delta\otimes\sigma},s)_{\omega}(g)\\ &\qquad=\sum_{\gamma\in P^{n-a}_{a(b-1)}(F)\backslash G_{n-a}(F)}\int_{N_{a,\omega}(\mathbb{A})}\int_{[M_{a\times a(b-1)}]}\lambda_s\phi(\omega^{-1}\gamma\tilde{n}(Z)ng)\,\mathrm{d}Z\,\mathrm{d}n\\ &\qquad=\sum_{\gamma\in P^{n-a}_{a(b-1)}(F)\backslash G_{n-a}(F)}\int_{N_{a,\omega}(\mathbb{A})}\int_{[M_{a(b-1)\times a}]}\lambda_s\phi(n(Z)\omega^{-1}n\gamma g)\,\mathrm{d}Z\,\mathrm{d}n,\end{aligned}$$

where $[M_{a(b-1)\times a}]:=M_{a(b-1)\times a}(F)\backslash M_{a(b-1)\times a}(\mathbb{A})$, and $n(Z)$ is the element

$$\begin{pmatrix} I_{a(b-1)} & Z & & & \\ & I_a & & & \\ & & I & & \\ & & & I_a & Z' \\ & & & & I_{a(b-1)}\end{pmatrix} \quad \text{for} \quad Z\in M_{a(b-1)\times a}.$$

We denote the inner integration by

$$\tilde{\phi}(g) := \int_{[M_{a(b-1)\times a}]} \phi(n(Z)g)\,\mathrm{d}Z. \tag{2-6}$$

Let $Q_{a(b-1),a} := L_{a(b-1),a}U_{a(b-1),a}$ be a standard parabolic subgroup of GL_{ab} whose unipotent radical $U_{a(b-1),a}$ embedded into G_n consists of all the elements $n(Z)$. Moreover, the standard parabolic subgroup $P_{a(b-1),a} = M_{a(b-1),a}N_{a(b-1),a}$ of G_n has the property that $M_{a(b-1),a} = L_{a(b-1),a} \times G_m$ and $N_{a(b-1),a} = U_{a(b-1),a}N_{ab}$.

Lemma 2.2. *The function $\lambda_s\tilde{\phi}$ is an automorphic function in the space*

$$A\big(N_{a(b-1),a}(\mathbb{A})M_{a(b-1),a}(F)\backslash G_n(\mathbb{A})\big)_{|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau\otimes\sigma}.$$

Here $|\cdot|_{F'}$ is as defined in Lemma 2.1.

Proof. The proof is similar to the proof of Lemma 2.1. For all $k \in \boldsymbol{K}$, the function $\phi_k(m) := m^{-\rho_{P_{ab}}}\phi(mk)$, for $m \in M_{ab}(\mathbb{A})$, is an automorphic form in the space $A(M_{ab}(F)\backslash M_{ab}(\mathbb{A}))_{\Delta(\tau,b)\otimes\sigma}$. We may assume that

$$\phi_k((t,r,h)) = \phi_{k,\Delta(\tau,b)}((t,r)) \otimes \phi_{k,\sigma}(h),$$

where $t \in \mathrm{GL}_{a(b-1)}(\mathbb{A}_{F'})$, $r \in \mathrm{GL}_a(\mathbb{A}_{F'})$ and $h \in G_m(\mathbb{A})$. Then

$$\tilde{\phi}_k((t,r,h)) = \big[(\phi_{k,\Delta(\tau,b)})_{Q_{a(b-1),a}}((t,r)) \otimes \phi_{k,\sigma}(h)\big].$$

By [Jiang and Liu 2012, Lemma 4.1], the constant term $(\phi_{k,\Delta(\tau,b)})_{Q_{a(b-1),a}}$ is an automorphic function in the space

$$A\big(U_{a(b-1),a}(\mathbb{A})L_{a(b-1),a}(F)\backslash \mathrm{GL}_{ab}(\mathbb{A}_{F'})\big)_{|\cdot|_{F'}^{-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{(b-1)/2}\tau}.$$

This is enough to deduce the lemma. □

Next, following the notation of [Mœglin and Waldspurger 1995, II.1.6], we consider the intertwining operator

$$M(\omega,\cdot) := M\big(\omega, |\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau\otimes\sigma\big), \tag{2-7}$$

which is defined by

$$\big(M(\omega,\cdot)\lambda_s\tilde{\phi}\big)(g) := \int_{N_{a,a(b-1)}\cap\,\omega N_{a(b-1),a}\omega^{-1}(F)\backslash N_{a,a(b-1)}(\mathbb{A})} \lambda_s\tilde{\phi}(\omega^{-1}ng)\,\mathrm{d}n.$$

Now, plug this intertwining operator into the ω-term and obtain

$$E_{P_a}(\phi_{\Delta\otimes\sigma},s)_\omega(g) = \sum_{\gamma\in P^{n-a}_{a(b-1)}(F)\backslash G_{n-a}(F)} \big(M(\omega,\cdot)\lambda_s\tilde{\phi}\big)(\gamma g).$$

By [Mœglin and Waldspurger 1995, Proposition II.1.6], the intertwining operator $M(\omega,\cdot)$ maps

$$A\big(N_{a(b-1),a}(\mathbb{A})M_{a(b-1),a}(F)\backslash G_n(\mathbb{A})\big)_{|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau\otimes\sigma}$$

to

$$A\big(N_{a,a(b-1)}(\mathbb{A})M_{a,a(b-1)}(F)\backslash G_n(\mathbb{A})\big)_{|\cdot|_{F'}^{-(s+(b-1)/2)}\tilde{\tau}^*\otimes|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes\sigma},$$

where $\tau^*=\tau$ if $F'=F$, and $\tau^*=\tau^\iota$ if $F'=E$, with ι being the nontrivial element in the Galois group $\Gamma_{E/F}$. Therefore, the restriction of the ω-term $E^n_{ab,P_a}(\phi_{\Delta\otimes\sigma},s)_\omega$ to the subgroup $I_a\times G_{n-a}(\mathbb{A})$ of the Levi subgroup $\mathrm{GL}_a(\mathbb{A}_{F'})\times G_{n-a}(\mathbb{A})$ is equal to

$$E^{n-a}_{a(b-1)}\big(\lambda_{1/2}(i^*_{n-a}\circ M(\omega,\cdot))\tilde{\phi},s-\tfrac{1}{2}\big)(h). \tag{2-8}$$

Combining the results of Sections 2B and 2C, we achieve an induction formula of the constant term.

Proposition 2.3. *The constant term $E^n_{ab,P_a}(\phi_{\Delta(\tau,b)\otimes\sigma},s)$ restricted to the subgroup $I_a\times G_{n-a}(\mathbb{A})$ of the Levi subgroup $\mathrm{GL}_a(\mathbb{A}_{F'})\times G_{n-a}(\mathbb{A})$ is expressed as the identity*

$$\begin{aligned}E^n_{ab,P_a}(\phi_{\Delta\otimes\sigma},s)((I_a,h)) &= E^{n-a}_{a(b-1)}\big(\lambda_{-1/2}(i^*_{n-a}\phi_{Q^{ab}_{a,a(b-1)}})_{\Delta(\tau,b-1)\otimes\sigma},s+\tfrac{1}{2}\big)(h)\\ &\quad + E^{n-a}_{a(b-1)}\big(\lambda_{1/2}(i^*_{n-a}\circ M(\omega,\cdot))\tilde{\phi},s-\tfrac{1}{2}\big)(h),\end{aligned} \tag{2-9}$$

which holds for all s with $\mathrm{Re}(s)$ large, and then is extended to $s\in\mathbb{C}$ by meromorphic continuation. Here

$$M(\omega,\cdot):=M\big(\omega,|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau\otimes\sigma\big),$$

$|\cdot|_{F'}:=|\cdot|_{\mathbb{A}_{F'}}$, *and $\tilde{\phi}$ is defined in (2-6). Note that F' is E if G_n is a unitary group, and is F otherwise.*

3. A normalized induction formula

In this section, we keep the assumption on G_n as in Section 2 and calculate normalizing factors for the relevant intertwining operators involved in the functional equation of Eisenstein series and in the induction formula (2-9). This leads to an induction formula for normalized Eisenstein series. As we remarked in the introduction of this paper, we have to assume that σ is an irreducible generic cuspidal automorphic representation of $G_m(\mathbb{A})$ if $m>0$.

3A. ***Normalized Eisenstein series ($m > 0$).*** We assume that $m > 0$, and recall the definitions of ρ and ρ^- in (1-2) and (1-3):

$$\rho := \begin{cases} \text{Asai} & \text{if } G_n = \mathrm{U}_{2n}, \\ \text{Asai} \otimes \delta & \text{if } G_n = \mathrm{U}_{2n+1}, \\ \mathrm{Sym}^2 & \text{if } G_n = \mathrm{SO}_{2n+1}, \\ \Lambda^2 & \text{if } G_n = \mathrm{Sp}_{2n} \text{ or } \mathrm{SO}_{2n}, \end{cases}$$

and

$$\rho^- := \begin{cases} \text{Asai} \otimes \delta & \text{if } G_n = \mathrm{U}_{2n}, \\ \text{Asai} & \text{if } G_n = \mathrm{U}_{2n+1}, \\ \Lambda^2 & \text{if } G_n = \mathrm{SO}_{2n+1}, \\ \mathrm{Sym}^2 & \text{if } G_n = \mathrm{Sp}_{2n} \text{ or } \mathrm{SO}_{2n}. \end{cases}$$

It follows [Ginzburg et al. 2011, Remark (3)] that

$$L(s, \tau \times \tau^*) = L(s, \tau, \rho) L(s, \tau, \rho^-),$$

where $\tau^* = \tau$ if $F' = F$ and $\tau^* = \tau^\iota$ if $F' = E$, where the involution ι is the nontrivial element in the Galois group $\Gamma_{E/F}$.

In order to normalize the Eisenstein series, we consider the normalization of the intertwining operator $M(\omega', |\cdot|_{F'}^s \Delta \otimes \sigma)(\phi)$ with

$$\omega' = (-1)^{ab} \begin{pmatrix} & & I_{ab} \\ & I & \\ \pm I_{ab} & & \end{pmatrix}.$$

By the general theory of Eisenstein series and intertwining operators [Langlands 1976; Mœglin and Waldspurger 1995, Chapter VI; Shahidi 2010, Theorem 6.1.7], both $E(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ and $M(\omega', |\cdot|_{F'}^s \Delta \otimes \sigma)$ can be extended to meromorphic functions of $s \in \mathbb{C}$, and the Eisenstein series $E(\phi_{\Delta\otimes\sigma}, s)$ has the functional equation

$$E(\phi_{\Delta(\tau,b)\otimes\sigma}, s) = E\big(M(\omega', |\cdot|_{F'}^s \Delta(\tau, b) \otimes \sigma)(\phi), -s\big). \tag{3-1}$$

If $\mathrm{Re}(s) = 0$, then $E(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic.

For any factorizable function $\phi = \bigotimes_v \phi_v$, we write

$$M\big(\omega', |\cdot|_{F'}^s \Delta \otimes \sigma\big)(\phi) = \prod_v M\big(\omega', |\cdot|_{F'_v}^s \Delta_v \otimes \sigma_v\big)(\phi_v).$$

By [Shahidi 2010, Theorem 6.3.1], for each local place v, define

$$N'\big(\omega', |\cdot|_{F'_v}^s \Delta(\tau_v, b) \otimes \sigma_v\big)(\phi_v) = \frac{1}{r'\big(\omega', |\cdot|_{F'_v}^s \Delta(\tau_v, b) \otimes \sigma_v\big)} M\big(\omega', |\cdot|_{F'_v}^s \Delta(\tau_v, b) \otimes \sigma_v\big)(\phi_v),$$

where the local normalizing factor

$$r'\big(\omega', |\cdot|^s_{F'_v}\Delta(\tau_v, b)\otimes\sigma_v\big) = \frac{L(s, \Delta(\tau_v, b)\times\sigma_v)}{L(s+1, \Delta(\tau_v, b)\times\sigma_v)\varepsilon(s, \Delta(\tau_v, b)\times\sigma_v, \psi_v)} \times \frac{L(2s, \Delta(\tau_v, b), \rho)}{L(2s+1, \Delta(\tau_v, b), \rho)\varepsilon(2s, \Delta(\tau_v, b), \rho, \psi_v)}.$$

Define $r'\big(\omega', |\cdot|^s_{F'}\Delta(\tau, b)\otimes\sigma\big) = \prod_v r'\big(\omega', |\cdot|^s_{F'_v}\Delta(\tau_v, b)\otimes\sigma_v\big)$. The global normalized intertwining operator is

$$N'\big(\omega', |\cdot|^s_{F'}\Delta\otimes\sigma\big) = \prod_v N'\big(\omega', |\cdot|^s_{F'_v}\Delta(\tau_v, b)\otimes\sigma_v\big).$$

For the global (complete) L-functions, we have

$$\begin{aligned} L\big(s, \Delta(\tau, b)\times\sigma\big) &= \prod_{i=1}^{b} L\Big(s+\frac{2i-b-1}{2}, \tau\times\sigma\Big), \\ L\big(s, \Delta(\tau, b), \rho\big) &= \prod_{i=1}^{b} L(s+b-2i+1, \tau, \rho) \\ &\quad\times \prod_{1\le i<j\le b} L\big(s+b-(i+j)+1, \tau\otimes\tau^*\big). \end{aligned} \tag{3-2}$$

Hence the quotient of complete L-functions has the property that

$$\frac{L\big(s, \Delta(\tau, b)\times\sigma\big)L\big(2s, \Delta(\tau, b), \rho\big)}{L\big(s+1, \Delta(\tau, b)\times\sigma\big)L\big(2s+1, \Delta(\tau, b), \rho\big)}$$

is equal to

$$\frac{\prod_{i=1}^{\lceil b/2\rceil} L\big(f_{b,i}(s), \tau, \rho\big)\prod_{i=1}^{\lfloor b/2\rfloor} L\big(f_{b,i}(s)+1, \tau, \rho^-\big)L\Big(s-\frac{b-1}{2}, \tau\times\sigma\Big)}{\prod_{i=1}^{\lceil b/2\rceil} L\big(e_{b,i}(s)+1, \tau, \rho\big)\prod_{i=1}^{\lfloor b/2\rfloor} L\big(e_{b,i}(s), \tau, \rho^-\big)L\Big(s+\frac{b+1}{2}, \tau\times\sigma\Big)},$$

where $e_{b,i}(s) := 2s+b+1-2i$, $f_{b,i}(s) := 2s-b-1+2i$. Define

$$\alpha_b(s) := \prod_{i=1}^{\lceil b/2\rceil} L\big(f_{b,i}(s), \tau, \rho\big)\prod_{i=1}^{\lfloor b/2\rfloor} L\big(f_{b,i}(s)+1, \tau, \rho^-\big)L\Big(s-\frac{b-1}{2}, \tau\times\sigma\Big),$$

$$\beta_b(s) := \prod_{i=1}^{\lceil b/2\rceil} L\big(e_{b,i}(s)+1, \tau, \rho\big)\prod_{i=1}^{\lfloor b/2\rfloor} L\big(e_{b,i}(s), \tau, \rho^-\big)L\Big(s+\frac{b+1}{2}, \tau\times\sigma\Big),$$

and

$$\varepsilon_b(s) := \prod_{i=1}^{\lceil b/2 \rceil} \varepsilon\big(f_{b,i}(s), \tau, \rho\big) \prod_{i=1}^{\lfloor b/2 \rfloor} \varepsilon\big(f_{b,i}(s)+1, \tau, \rho^-\big)\varepsilon\Big(s - \frac{b-1}{2}, \tau\times\sigma\Big).$$

Finally define the global normalizing factor by

$$r\big(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma\big) = r\big(\omega', |\cdot|_{F'}^s \Delta(\tau,b)\otimes\sigma\big) = \frac{\alpha_b(s)}{\beta_b(s)\varepsilon_b(s)} \tag{3-3}$$

and the normalized global intertwining operator by

$$N\big(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma\big) = \frac{\varepsilon_b(s) N'\big(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma\big)}{\varepsilon\big(s, \Delta(\tau,b)\times\sigma\big)\varepsilon\big(2s, \Delta(\tau,b), \rho\big)}. \tag{3-4}$$

Then we have that

$$N\big(\omega', |\cdot|_{F'}^s \Delta(\tau,b)\otimes\sigma\big) = \frac{1}{r\big(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma\big)} M\big(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma\big).$$

Meanwhile, we use $\beta_b(s)$ to normalize the Eisenstein series

$$E_{ab}^{n,*}(\phi_{\Delta(\tau,b)\otimes\sigma}, s) := \beta_b(s) E_{ab}^{n}(\phi_{\Delta(\tau,b)\otimes\sigma}, s). \tag{3-5}$$

By the functional equation (3-1) of $E_{ab}^{n}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$, the normalized Eisenstein series $E_{ab}^{n,*}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ satisfies the functional equation

$$E_{ab}^{n,*}(\phi_{\Delta(\tau,b)\otimes\sigma}, s) = E_{ab}^{n,*}\big(N(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma)(\phi), -s\big). \tag{3-6}$$

In fact,

$$\begin{aligned} E^*(\phi_{\Delta\otimes\sigma}, s) &= \beta_b(s) E\big(M(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma)(\phi), -s\big) \\ &= \beta_b(s)\cdot r\big(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma\big) E\big(N(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma)(\phi), -s\big) \\ &= \frac{\beta_b(s)\cdot r\big(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma\big)}{\beta_b(-s)} E^*\big(N(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma)(\phi), -s\big). \end{aligned}$$

Since $\alpha_b(s) = \varepsilon_b(s)\beta_b(-s)$, we have

$$\beta_b(s)\cdot r\big(\omega', |\cdot|_{F'}^s \Delta\otimes\sigma\big) = \beta_b(-s).$$

From this we deduce the functional equation (3-6).

We remark that when $b = 1$, it is easy to show that for $\mathrm{Re}(s) > 0$, the normalized global intertwining operator

$$N\big(\omega', |\cdot|_{F'}^s \Delta(\tau,b)\otimes\sigma\big)$$

is holomorphic for all choice of data, and nonzero for some choice of data.

In fact, if $b = 1$ and $\tau\otimes\sigma$ is a generic representation, then the normalized local intertwining operator $N(\omega', \cdot)$ is holomorphic and nonzero by Theorem 11.1 of

[Cogdell et al. 2004]. The proof uses the Langlands functorial transfers of σ from G_n to the corresponding general linear groups, the Ramanujan type estimate for cuspidal automorphic forms on general linear groups [Luo et al. 1999], and the structure of generic unitary dual for classical groups over all local fields [Lapid et al. 2004]. Hence the result for $b = 1$ holds for F-quasisplit unitary groups with the same proof [Cogdell et al. 2011].

However, when $b > 1$, we are not able to prove the above properties for the normalized global intertwining operator $N(\omega', \cdot)$, so that we are not able to control the poles at $\mathrm{Re}(s) < 0$ of the normalized Eisenstein series through the functional equation (3-6). We will leave this issue for our future consideration.

3B. ***Normalization of $M(\omega, \cdot)$ with $m > 0$.*** In order to normalize the global intertwining operator

$$M(\omega, \cdot) := M\big(\omega, |\cdot|_{F'}^{s-1/2}\Delta(\tau, b-1) \otimes |\cdot|_{F'}^{s+(b-1)/2}\tau \otimes \sigma\big),$$

as defined in (2-7), we decompose it into a composition of two intertwining operators

$$M\big(\omega, |\cdot|_{F'}^{s-1/2}\Delta(\tau, b-1) \otimes |\cdot|_{F'}^{s+(b-1)/2}\tau \otimes \sigma\big) = M(\omega_1, \cdot) \circ M(\omega_2, \cdot),$$

where

$$\omega_1 = \begin{pmatrix} & I_a & & & \\ I_{a(b-1)} & & & & \\ & & I & & \\ & & & & I_{a(b-1)} \\ & & & I_a & \end{pmatrix}$$

and

$$\omega_2 = (-1)^a \begin{pmatrix} I_{a(b-1)} & & & & \\ & & & I_a & \\ & & I & & \\ & \pm I_a & & & \\ & & & & I_{a(b-1)} \end{pmatrix}.$$

More precisely, $M(\omega_1, \cdot)$ and $M(\omega_2, \cdot)$ are standard intertwining operators of the following types: $M(\omega_2, \cdot)$ maps from the space

$$A\big(N_{a(b-1),a}(\mathbb{A})M_{a(b-1),a}(F)\backslash G_n(\mathbb{A})\big)_{|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau\otimes\sigma}$$

to the space

$$A\big(N_{a(b-1),a}(\mathbb{A})M_{a(b-1),a}(F)\backslash G_n(\mathbb{A})\big)_{|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{-s-(b-1)/2}\tilde{\tau}^*\otimes\sigma},$$

and $M(\omega_1, \cdot)$ maps from the space

$$A\big(N_{a(b-1),a}(\mathbb{A})M_{a(b-1),a}(F)\backslash G_n(\mathbb{A})\big)_{|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{-s-(b-1)/2}\tilde{\tau}^*\otimes\sigma}$$

to the space

$$A\big(N_{a(b-1),a}(\mathbb{A})M_{a(b-1),a}(F)\backslash G_n(\mathbb{A})\big)_{|\cdot|_{F'}^{-s-(b-1)/2}\tilde{\tau}^*\otimes|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes\sigma}.$$

The standard Langlands–Shahidi normalizing factors for $M(\omega_1,\cdot)$ and $M(\omega_2,\cdot)$ are given by $r(\omega_1,\cdot)$ and $r(\omega_2,\cdot)$, where

$$r(\omega_1,\cdot)=\frac{L\Big(2s+\dfrac{b}{2}-1,\,\Delta(\tau,b-1)\times\tau^*\Big)}{L\Big(2s+\dfrac{b}{2},\,\Delta(\tau,b-1)\times\tau^*\Big)\varepsilon\Big(2s+\dfrac{b}{2}-1,\,\Delta(\tau,b-1)\times\tau^*\Big)}$$

and $r(\omega_2,\cdot)$ is

$$\frac{L\Big(s+\dfrac{b-1}{2},\,\tau\times\sigma\Big)L\big(2s+b-1,\,\tau,\,\rho\big)}{L\Big(s+\dfrac{b+1}{2},\,\tau\times\sigma\Big)L\big(2s+b,\,\tau,\,\rho\big)\varepsilon\Big(s+\dfrac{b-1}{2},\,\tau\times\sigma\Big)\varepsilon\big(2s+b-1,\,\tau,\,\rho\big)}.$$

We define

$$\begin{aligned}M(\omega_1,\cdot)&=r(\omega_1,\cdot)N(\omega_1,\cdot),\\ M(\omega_2,\cdot)&=r(\omega_2,\cdot)N(\omega_2,\cdot),\\ r(\omega,\cdot)&=r(\omega_1,\cdot)r(\omega_2,\cdot),\end{aligned}$$

and

$$M(\omega,\cdot)=r(\omega,\cdot)N(\omega,\cdot). \tag{3-7}$$

It follows that

$$N(\omega,\cdot)=N(\omega_1,\cdot)\circ N(\omega_2,\cdot).$$

Proposition 3.1. *Assume that* $b>1$. *For* $\mathrm{Re}(s)>0$, *the normalized global intertwining operator*

$$N(\omega,\cdot)=N\big(\omega,|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau\otimes\sigma\big)$$

is holomorphic for all choices of data, and nonzero for some choice of data. For $\mathrm{Re}(s)=0$, *it is holomorphic.*

Proof. First we show that the normalized intertwining operators $N(\omega_i,\cdot)$ for $i=1,2$ are holomorphic and nonzero at $\mathrm{Re}(s)\geq 0$.

Indeed, by Theorem 11.1 in [Cogdell et al. 2004] for orthogonal and symplectic group cases, $N(\omega_2,\cdot)$ is holomorphic for all choices of data and nonzero for some choice of data, when $\mathrm{Re}(s+(b-1)/2)\geq 0$. For even and odd unitary group cases, the same result follows from Proposition 9.4 in [Kim and Krishnamurthy 2005] and Proposition 5 in [Kim and Krishnamurthy 2004].

For the normalized intertwining operators $N(\omega_1,\cdot)$, it is essentially the intertwining operator for general linear groups, which is considered in [Mœglin and Waldspurger 1989]. We write it as an eulerian product

$$N(\omega_1,\cdot)=\prod_v N_v(\omega_1,\cdot),$$

with $\tau=\bigotimes_v \tau_v$. Since τ_v is unitary and generic, we can assume that

$$\tau_v=|\cdot|^{\nu_1}\mathrm{St}(\tau_1,a_1)\times|\cdot|^{\nu_2}\mathrm{St}(\tau_2,a_2)\times\cdots\times|\cdot|^{\nu_r}\mathrm{St}(\tau_r,a_r),$$

where $-\frac{1}{2}<\nu_i<\frac{1}{2}$ for all i and $\mathrm{St}(\tau_i,a_i)$ are Steinberg representations for some supercuspidal representations τ_i and integers a_i, and nonlinked. Write $e(\tau_v)=2\inf\bigl\{\frac{1}{2}-|\nu_i|,1\le i\le r\bigr\}$ (referring to I.10 in [Mœglin and Waldspurger 1989]). It follows that

$$e(\Delta(\tau_v,b-1))=e(\tau_v).$$

By Proposition I.10 in [Mœglin and Waldspurger 1989], $N_v(\omega_1,\cdot)$ is holomorphic and nonzero when $\mathrm{Re}\bigl(s-\frac{1}{2}-(-s-(b-1)/2)\bigr)>-e(\tau_v)$ at all local places v. In particular, they are holomorphic and nonzero at $\mathrm{Re}(s)\ge 0$, and so is the normalized global intertwining operator $N(\omega_1,\cdot)$. Hence $N(\omega,\cdot)=N(\omega_1,\cdot)\circ N(\omega_2,\cdot)$ is holomorphic for all choices of data when $\mathrm{Re}(s)\ge 0$.

We notice that for $\mathrm{Re}(s)>0$, $N(\omega_1,\cdot)$ as a GL_{ab}-intertwining operator is an isomorphism, and hence $N(\omega,\cdot)=N(\omega_1,\cdot)\circ N(\omega_2,\cdot)$ is nonzero for some choice of data. □

By substituting the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma},s)$ in (3-5) and the normalized intertwining operator $N(\omega,\cdot)$ in (3-7) into the induction formula (2-9) in Proposition 2.3, we obtain

$$\begin{aligned}E^{n,*}_{ab,P_a}&(\phi_{\Delta\otimes\sigma},s)((I_a,h))\\&=\frac{\beta_b(s)}{\beta_{b-1}\bigl(s+\frac{1}{2}\bigr)}E^{n-a,*}_{a(b-1)}\Bigl(\lambda_{-1/2}\bigl(i^*_{n-a}\phi_{Q^{ab}_{a,a(b-1)}}\bigr)_{\Delta(\tau,b-1)\otimes\sigma},s+\tfrac{1}{2}\Bigr)(h)\\&\qquad+\frac{\beta_b(s)\cdot r(\omega,\cdot)}{\beta_{b-1}\bigl(s-\frac{1}{2}\bigr)}E^{n-a,*}_{a(b-1)}\bigl(\lambda_{1/2}(i^*_{n-a}\circ N(\omega,\cdot))\tilde{\phi},s-\tfrac{1}{2}\bigr)(h).\end{aligned}$$

Using a similar calculation as in (3-2), it is easy to verify that

$$\frac{\beta_b(s)}{\beta_{b-1}\bigl(s+\frac{1}{2}\bigr)}=L\bigl(2s+1,\tau,\rho^{(-)^{b+1}}\bigr),$$

$$\frac{\beta_b(s)}{\beta_{b-1}\bigl(s-\frac{1}{2}\bigr)}r(\omega,\cdot)=\frac{L\bigl(2s,\tau,\rho^{(-)^{b+1}}\bigr)}{\varepsilon'_b(s)},$$

where

$$\varepsilon_b'(s) := \varepsilon\Big(s+\frac{b-1}{2},\tau\times\sigma\Big)\varepsilon\big(2s+b-1,\tau,\rho\big)\varepsilon\Big(2s+\frac{b}{2}-1,\Delta(\tau,b-1)\times\tau^*\Big).$$

Therefore, for $b>1$, we obtain the following normalized induction formula.

Proposition 3.2 (induction formula). *Let G_n be the classical groups as defined in Section 2. Assume that $m>0$ and σ is an irreducible generic cuspidal automorphic representation of $G_n(\mathbb{A})$. For $b>1$, the following formula holds*:

$$\begin{aligned}\text{(3-8)}\quad & E^{n,*}_{ab,P_a}(\phi_{\Delta\otimes\sigma},s)((I_a,h))\\ &= L\big(2s+1,\tau,\rho^{(-)^{b+1}}\big)E^{n-a,*}_{a(b-1)}\big(\lambda_{-1/2}(i^*_{n-a}\phi_Q),s+\tfrac{1}{2}\big)(h)\\ &\quad+\frac{L\big(2s,\tau,\rho^{(-)^{b+1}}\big)}{\varepsilon_b'(s)}E^{n-a,*}_{a(b-1)}\big(\lambda_{1/2}(i^*_{n-a}\circ N(\omega,\cdot))\tilde{\phi},s-\tfrac{1}{2}\big)(h),\end{aligned}$$

where $\lambda_{-1/2}(i^*_{n-a}\phi_Q) := \lambda_{-1/2}\big(i^*_{n-a}\phi_{Q^{ab}_{a,a(b-1)}}\big)_{\Delta(\tau,b-1)\otimes\sigma}$.

3C. *Normalization for the case of $m=0$.* In this section, we consider the case of $m=0$. Due to the similarity between the cases of $m=0$ and $m>0$, we will just briefly sketch the result here. We continue to use the notation and references (which will not be mentioned) introduced in previous sections.

Note that when $m=0$, $G_n=\mathrm{SO}_{2n}$ must be F-split. In this case, we divide the G_n into two types: Type (1), $G_n=\mathrm{Sp}_{2n}$ and U_{2n+1}, and Type (2), $G_n=\mathrm{SO}_{2n+1}$, SO_{2n}, and U_{2n}.

In order to normalize the Eisenstein series, we consider the intertwining operator $M(\omega',|\cdot|^s_{F'}\Delta)(\phi)$ with

$$\omega' = (-1)^{ab}\begin{pmatrix} & & I_{ab}\\ & I & \\ \pm I_{ab} & & \end{pmatrix},$$

where the I in the middle is the identity matrix of order at most one, that is, it either is 1 or disappears, depending on the structure of G_n.

For any factorizable function $\phi=\bigotimes_v\phi_v$, we write

$$M(\omega',|\cdot|^s_{F'}\Delta)(\phi)=\prod_v M(\omega',|\cdot|^s_{F'_v}\Delta_v)(\phi_v),$$

and for each local place v, define

$$N'\big(\omega',|\cdot|^s_{F'_v}\Delta(\tau_v,b)\big)(\phi_v)=\frac{1}{r'\big(\omega',|\cdot|^s_{F'_v}\Delta(\tau_v,b)\big)}M\big(\omega',|\cdot|^s_{F'_v}\Delta(\tau_v,b)\big)(\phi_v),$$

where the local normalizing factor $r'\big(\omega',|\cdot|^s_{F'_v}\Delta(\tau_v,b)\big)$ is defined as follows.

When G_n is of Type (1), define

$$r'\big(\omega', |\cdot|^s_{F'_v}\Delta(\tau_v, b)\big) := \frac{L(s, \Delta(\tau_v, b))}{L(s+1, \Delta(\tau_v, b))\varepsilon(s, \Delta(\tau_v, b), \psi_v)} \times \frac{L(2s, \Delta(\tau_v, b), \rho)}{L(2s+1, \Delta(\tau_v, b), \rho)\varepsilon(2s, \Delta(\tau_v, b), \rho, \psi_v)};$$

and when G_n is of Type (2), define

$$r'\big(\omega', |\cdot|^s_{F'_v}\Delta(\tau_v, b)\big) = \frac{L(2s, \Delta(\tau_v, b), \rho)}{L(2s+1, \Delta(\tau_v, b), \rho)\varepsilon(2s, \Delta(\tau_v, b), \rho, \psi_v)}.$$

Then we define $r'\big(\omega', |\cdot|^s_{F'}\Delta(\tau, b)\big) = \prod_v r'\big(\omega', |\cdot|^s_{F'_v}\Delta(\tau_v, b)\big)$. The global normalized intertwining operator is

$$N'(\omega', |\cdot|^s_{F'}\Delta) = \prod_v N'\big(\omega', |\cdot|^s_{F'_v}\Delta(\tau_v, b)\big).$$

We calculate the L-functions as in (3-2) and obtain

$$\frac{L\big(s, \Delta(\tau, b)\big)}{L\big(s+1, \Delta(\tau, b)\big)} = \frac{L\Big(s - \dfrac{b-1}{2}, \tau\Big)}{L\Big(s + \dfrac{b+1}{2}, \tau\Big)}$$

and

$$\frac{L(2s, \Delta(\tau, b), \rho)}{L(2s+1, \Delta(\tau, b), \rho)} = \frac{\displaystyle\prod_{i=1}^{\lceil b/2\rceil} L(f_{b,i}(s), \tau, \rho) \prod_{i=1}^{\lfloor b/2\rfloor} L(f_{b,i}(s)+1, \tau, \rho^-)}{\displaystyle\prod_{i=1}^{\lceil b/2\rceil} L(e_{b,i}(s)+1, \tau, \rho) \prod_{i=1}^{\lfloor b/2\rfloor} L(e_{b,i}(s), \tau, \rho^-)},$$

where $e_{b,i}(s) := 2s+b+1-2i,\ f_{b,i}(s) := 2s-b-1+2i$.

When G_n is of Type (1), define

$$\alpha_b(s) = \prod_{i=1}^{\lceil b/2\rceil} L\big(f_{b,i}(s), \tau, \rho\big) \prod_{i=1}^{\lfloor b/2\rfloor} L\big(f_{b,i}(s)+1, \tau, \rho^-\big) L\Big(s - \frac{b-1}{2}, \tau\Big),$$

$$\beta_b(s) = \prod_{i=1}^{\lceil b/2\rceil} L\big(e_{b,i}(s)+1, \tau, \rho\big) \prod_{i=1}^{\lfloor b/2\rfloor} L\big(e_{b,i}(s), \tau, \rho^-\big) L\Big(s + \frac{b+1}{2}, \tau\Big),$$

$$\varepsilon_b(s) = \prod_{i=1}^{\lceil b/2\rceil} \varepsilon\big(f_{b,i}(s), \tau, \rho\big) \prod_{i=1}^{\lfloor b/2\rfloor} \varepsilon\big(f_{b,i}(s)+1, \tau, \rho^-\big) \varepsilon\Big(s - \frac{b-1}{2}, \tau\Big),$$

and then define

$$N(\omega', |\cdot|^s_{F'}\Delta) = \frac{\varepsilon_b(s) N'(\omega', |\cdot|^s_{F'}\Delta)}{\varepsilon(s, \Delta(\tau, b), \psi)\varepsilon(2s, \Delta(\tau, b), \rho, \psi)}.$$

When G_n is of Type (2), define

$$\alpha_b(s) = \prod_{i=1}^{\lceil b/2\rceil} L(f_{b,i}(s), \tau, \rho) \prod_{i=1}^{\lfloor b/2\rfloor} L(f_{b,i}(s)+1, \tau, \rho^-),$$

$$\beta_b(s) = \prod_{i=1}^{\lceil b/2\rceil} L(e_{b,i}(s)+1, \tau, \rho) \prod_{i=1}^{\lfloor b/2\rfloor} L(e_{b,i}(s), \tau, \rho^-),$$

$$\varepsilon_b(s) = \prod_{i=1}^{\lceil b/2\rceil} \varepsilon(f_{b,i}(s), \tau, \rho) \prod_{i=1}^{\lfloor b/2\rfloor} \varepsilon(f_{b,i}(s)+1, \tau, \rho^-),$$

and then define

$$N(\omega', |\cdot|_{F'}^s\Delta) = \frac{\varepsilon_b(s)N'(\omega', |\cdot|_{F'}^s\Delta)}{\varepsilon(2s, \Delta(\tau, b), \rho, \psi)}.$$

Now we define the normalizing factor by

$$r(\omega', |\cdot|_{F'}^s\Delta) := \frac{\alpha_b(s)}{\beta_b(s)\varepsilon_b(s)}.$$

Then

$$N\big(\omega', |\cdot|_{F'}^s\Delta(\tau, b)\big) = \frac{1}{r\big(\omega', |\cdot|_{F'}^s\Delta(\tau, b)\big)} M\big(\omega', |\cdot|_{F'}^s\Delta(\tau, b)\big).$$

Remark 3.3. The terms $\alpha_b(s)$ and $\beta_b(s)$ correspond to the terms $a_b(s)$ and $b_b(s)$ in [Brenner 2009, Section 4.2]. We correct the definition of $b_b(s)$ in [Brenner 2009] here.

We use $\beta_b(s)$ to normalize the Eisenstein series

$$E_{ab}^{n,*}(\phi_{\Delta(\tau,b)}, s) := \beta_b(s) E_{ab}^n(\phi_{\Delta(\tau,b)}, s). \tag{3-9}$$

Then, similarly to (3-6), we have the functional equation for the normalized Eisenstein series:

$$E_{ab}^{n,*}(\phi_{\Delta(\tau,b)}, s) = E_{ab}^{n,*}\big(N(\omega', |\cdot|_{F'}^s\Delta)(\phi), -s\big). \tag{3-10}$$

Next, we normalize the intertwining operator as defined in (2-7),

$$M(\omega, \cdot) := M\big(\omega, |\cdot|_{F'}^{s-1/2}\Delta(\tau, b-1) \otimes |\cdot|_{F'}^{s+(b-1)/2}\tau\big),$$

by

$$N(\omega, \cdot) := N\big(\omega, |\cdot|_{F'}^{s-1/2}\Delta(\tau, b-1) \otimes |\cdot|_{F'}^{s+(b-1)/2}\tau\big) = \frac{M(\omega, \cdot)}{r(\omega, \cdot)}, \tag{3-11}$$

with $r(\omega,\cdot)$ defined as follows. When G_n is of Type (1), define

$$\begin{aligned} r(\omega,\cdot) &= r\big(\omega, |\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau\big) \\ &= \frac{L\Big(s+\frac{b-1}{2},\tau\Big)L\big(2s+b-1,\tau,\rho\big)L\Big(2s+\frac{b}{2}-1,\Delta(\tau,b-1)\times\tau^*\Big)}{L\Big(s+\frac{b+1}{2},\tau\Big)L\big(2s+b,\tau,\rho\big)L\Big(2s+\frac{b}{2},\Delta(\tau,b-1)\times\tau^*\Big)} \\ &\quad\times\frac{1}{\varepsilon\Big(s+\frac{b-1}{2},\tau\Big)\varepsilon\big(2s+b-1,\tau,\rho\big)\varepsilon\Big(2s+\frac{b}{2}-1,\Delta(\tau,b-1)\times\tau^*\Big)}; \end{aligned}$$

and when G_n is of Type (2), define

$$\begin{aligned} r(\omega,\cdot) &= r\big(\omega, |\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau\big) \\ &= \frac{L\big(2s+b-1,\tau,\rho\big)L\Big(2s+\frac{b}{2}-1,\Delta(\tau,b-1)\times\tau^*\Big)}{L\big(2s+b,\tau,\rho\big)L\Big(2s+\frac{b}{2},\Delta(\tau,b-1)\times\tau^*\Big)} \\ &\quad\times\frac{1}{\varepsilon\big(2s+b-1,\tau,\rho\big)\varepsilon\Big(2s+\frac{b}{2}-1,\Delta(\tau,b-1)\times\tau^*\Big)}, \end{aligned}$$

where $\tau^*=\tau$ if $F'=F$ and $\tau^*=\tau^\iota$ if $F'=E$, with ι being the nontrivial element in the Galois group $\Gamma_{E/F}$.

The following, corresponding to Proposition 3.1, is also true when $m=0$.

Proposition 3.4. *For* $\mathrm{Re}(s)>0$ *and* $b>1$, *the normalized global intertwining operator* $N(\omega,|\cdot|_{F'}^{s-1/2}\Delta(\tau,b-1)\otimes|\cdot|_{F'}^{s+(b-1)/2}\tau)$ *is holomorphic for all choices of data, and nonzero for some choice of data.*

The proof follows from that of Proposition 3.1, and we omit the details here.

By substituting the normalized Eisenstein series $E_{ab}^{n,*}(\phi_{\Delta(\tau,b)},s)$ in (3-9) and the normalized intertwining operator $N(\omega,\cdot)$ in (3-11) into the induction formula (2-9) in Proposition 2.3, we obtain

$$\begin{aligned} &E_{ab,P_a}^{n,*}(\phi_\Delta,s)((I_a,h)) \\ &\quad= \frac{\beta_b(s)}{\beta_{b-1}\big(s+\frac12\big)}E_{a(b-1)}^{n-a,*}\big(\lambda_{-1/2}(i_{n-a}^*\phi_{Q_{a,a(b-1)}^{ab}})_{\Delta(\tau,b-1)},s+\tfrac12\big)(h) \\ &\qquad\qquad+\frac{\beta_b(s)\cdot r(\omega,\cdot)}{\beta_{b-1}\big(s-\frac12\big)}E_{a(b-1)}^{n-a,*}\big(\lambda_{1/2}(i_{n-a}^*\circ N(\omega,\cdot))\tilde{\phi},s-\tfrac12\big)(h). \end{aligned}$$

Using a similar calculation as in (3-2), it is easy to verify that

$$\frac{\beta_b(s)}{\beta_{b-1}\big(s+\frac{1}{2}\big)} = L\big(2s+1, \tau, \rho^{(-)^{b+1}}\big),$$
$$\frac{\beta_b(s)}{\beta_{b-1}\big(s-\frac{1}{2}\big)} r(\omega, \cdot\,) = \frac{L\big(2s, \tau, \rho^{(-)^{b+1}}\big)}{\varepsilon'_b(s)},$$

where $\varepsilon'_b(s)$ is defined as follows. When G_n is of Type (1), define $\varepsilon'_b(s)$ to be the product

$$\varepsilon\big(2s+b-1, \tau, \rho\big)\varepsilon\Big(2s+\frac{b}{2}-1, \Delta(\tau, b-1)\times\tau^*\Big)\varepsilon\Big(s+\frac{b-1}{2}, \tau\Big);$$

and when G_n is of Type (2), define

$$\varepsilon'_b(s) = \varepsilon\big(2s+b-1, \tau, \rho\big)\varepsilon\Big(2s+\frac{b}{2}-1, \Delta(\tau, b-1)\times\tau^*\Big).$$

Therefore, for $b>1$, we obtain the following normalized induction formula, which is similar to Proposition 3.2.

Proposition 3.5. *With notation as defined above, for $b>1$, the following formula holds*:

$$\begin{aligned}(3\text{-}12)\quad & E^{n,*}_{ab,P_a}(\phi_\Delta, s)((I_a, h)) \\ &= L\big(2s+1, \tau, \rho^{(-)^{b+1}}\big)E^{n-a,*}_{a(b-1)}\big(\lambda_{-1/2}(i^*_{n-a}\phi_Q), s+\tfrac{1}{2}\big)(h) \\ &\quad + \frac{L\big(2s, \tau, \rho^{(-)^{b+1}}\big)}{\varepsilon'_b(s)} E^{n-a,*}_{a(b-1)}\big(\lambda_{1/2}(i^*_{n-a}\circ N(\omega, \cdot\,))\tilde{\phi}, s-\tfrac{1}{2}\big)(h),\end{aligned}$$

where $\lambda_{-1/2}(i^*_{n-a}\phi_Q) := \lambda_{-1/2}\big(i^*_{n-a}\phi_{Q^{ab}_{a,a(b-1)}}\big)_{\Delta(\tau, b-1)}$.

4. Proof of Theorem 1.2 ($m>0$)

We are going to prove Theorem 1.2 for the case where $m>0$ using the normalized induction formula given in Proposition 3.2. From now on, we only consider symplectic group and F-quasisplit special orthogonal group cases.

4A. ***Case of $b=1$.*** The case of $b=1$ is the starting step of our proof by induction. Assume that $s\in\mathbb{C}$ with $\mathrm{Re}(s)>0$.

By Equation (3-5), we normalize $E^n_a(\phi_{\tau\otimes\sigma}, s)$ as follows:

$$E^{n,*}_a(\phi_{\tau\otimes\sigma}, s) = L(s+1, \tau\times\sigma)L(2s+1, \tau, \rho)E^n_a(\phi_{\tau\otimes\sigma}, s).$$

By [Mœglin and Waldspurger 1995, Proposition II.1.7], the constant term of the Eisenstein series $E^n_a(\phi_{\tau\otimes\sigma}, s)$ along a standard parabolic subgroup P' is always zero unless $P'=P_a$. In the case of $P'=P_a$, we have

$$E^n_{a,P_a}(\phi_{\tau\otimes\sigma}, s) = \lambda_s\phi_{\tau\otimes\sigma}(g) + M(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)(\lambda_s\phi).$$

By Lemma I.4.10 of [Mœglin and Waldspurger 1995], the Eisenstein series $E^n_a(\phi_{\tau\otimes\sigma}, s)$ has a pole at some point s_0 if and only if the constant term of $E^n_{a,P_a}(\phi_{\tau\otimes\sigma}, s)$ has a pole at s_0, and hence if and only if the term

$$M(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)(\lambda_s\phi)$$

has a pole at s_0, since the first term $\lambda_s\phi_{\tau\otimes\sigma}(g)$ is holomorphic. By our normalization, we have

$$M(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)(\lambda_s\phi) = r(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)N(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)(\phi),$$

and for $\mathrm{Re}(s) > 0$, by [Cogdell et al. 2004, Theorem 11.1], the normalized global intertwining operator $N(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)$ is holomorphic for all choice of data and nonzero for some choice of data. Thus, it reduces to checking the existence of the pole at $s = s_0$ of the global normalizing factor $r(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)$.

Recall from (3-3) that the global normalizing factor $r(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)$ in this case is

$$\frac{L(s, \tau\times\sigma)L(2s, \tau, \rho)}{L(s+1, \tau\times\sigma)L(2s+1, \tau, \rho)\varepsilon(s, \tau\times\sigma)\varepsilon(2s, \tau, \rho)}.$$

Since both $\varepsilon(s, \tau\times\sigma)$ and $\varepsilon(2s, \tau, \rho)$ are holomorphic and nonzero, the poles of the global normalizing factor $r(\omega', |\cdot|^s_{F'}\tau\otimes\sigma)$ at $s = s_0 > 0$ with $\mathrm{Re}(s_0) > 0$ are the same as the poles of the quotient

$$\frac{L(s, \tau\times\sigma)L(2s, \tau, \rho)}{L(s+1, \tau\times\sigma)L(2s+1, \tau, \rho)}$$

at $s = s_0 > 0$ with $\mathrm{Re}(s_0) > 0$.

Since σ is generic and τ is self-dual, by the global Langlands functorial transfer from G_n to a general linear group [Cogdell et al. 2004] and the analytic property of the complete L-functions of the Rankin–Selberg convolution [Cogdell and Piatetski-Shapiro 2004; Mœglin and Waldspurger 1989], we deduce that the complete L-function $L(s, \tau\times\sigma)$ is holomorphic at all $s\in\mathbb{C}$ except for a possible simple pole at $s = 0$ or 1, and is nonzero when $\mathrm{Re}(s) \le 0$ or $\mathrm{Re}(s) \ge 1$. Such a pole occurs if and only if τ occurs as an isobaric summand in the image of σ under the Langlands functorial transfer [Cogdell et al. 2004].

On the other hand, by [Grbac 2011], based on the work of Arthur [2013] on the classification of the discrete spectrum of $G_n(\mathbb{A})$, the complete L-function $L(s, \tau, \rho)$ is holomorphic at all $s\in\mathbb{C}$ except for a possible simple pole at $s = 0$ or 1, and is nonzero when $\mathrm{Re}(s) \le 0$ or $\mathrm{Re}(s) \ge 1$. Such a pole occurs if and only if τ can descend to an irreducible generic cuspidal automorphic representation of a classical group determined by ρ [Ginzburg et al. 2011].

Hence, when $\mathrm{Re}(s) > 0$, the denominator $L(s+1, \tau\times\sigma)L(2s+1, \tau, \rho)$ is holomorphic and nonzero, and the numerator $L(s, \tau\times\sigma)L(2s, \tau, \rho)$ is holomorphic except for a possible simple pole at $s=\frac{1}{2}$ or $s=1$. This proves the theorem for the case of $b=1$. We summarize the above as the following.

Proposition 4.1 (case $b=1$ of Theorem 1.2). *Let G_n be the symplectic group or the F-quasisplit special orthogonal group. Let τ be an irreducible unitary cuspidal automorphic representation of* $\mathrm{GL}_a(\mathbb{A})$ *and let σ be an irreducible generic cuspidal automorphic representation of $G_m(\mathbb{A})$. The normalized Eisenstein series $E_a^{n,*}(\phi_{\tau\otimes\sigma},s)$ is holomorphic at* $\mathrm{Re}(s)\geq 0$, *except at $s=\frac{1}{2}$ and $s=1$, where it has possible simple poles. Moreover*:

(1) *$E_a^{n,*}(\phi_{\tau\otimes\sigma},s)$ has a simple pole at $s=\frac{1}{2}$ if and only if $L(s,\tau,\rho)$ has a pole at $s=1$, and $L(\frac{1}{2},\tau\times\sigma)\neq 0$.*

(2) *$E_a^{n,*}(\phi_{\tau\otimes\sigma},s)$ has a simple pole at $s=1$ if and only if $L(s,\tau\times\sigma)$ has a pole at $s=1$.*

In particular, $E_a^{n,*}(\phi_{\tau\otimes\sigma},s)$ is holomorphic at $\mathrm{Re}(s)>0$ *if τ is not self-dual.*

Proposition 4.1 includes the case of $m=0$, which is proved in [Grbac 2011].

We remark that by the functional equation for the normalized Eisenstein series (3-6), one deduces the analytic properties at $\mathrm{Re}(s)<0$, since when $b=1$, the normalized intertwining operator occurring in the functional equation is holomorphic for $\mathrm{Re}(s)>0$ and is a nonzero operator. At $\mathrm{Re}(s)=0$, it is holomorphic (Corollary 4.3).

4B. ***Case of b > 1.*** This general case of Theorem 1.2 is proved by using the normalized induction formula (Proposition 3.2) and the case of $b=1$ (Proposition 4.1). One technical point is to prove that the normalized Eisenstein series $E_{ab}^{n,*}(\phi_{\tau\otimes\sigma},s)$ is holomorphic at $s=0$ (Corollary 4.3), which is a consequence of the following.

Proposition 4.2. *Let G_n be the symplectic group or the F-quasisplit special orthogonal group. Assume that σ is an irreducible generic* (*or tempered if nongeneric*) *cuspidal automorphic representation of $G_m(\mathbb{A})$. If $\beta_b(s)$ has a pole at $s=0$, then the pole at $s=0$ of $\beta_b(s)$ must be simple and $E_{ab}^{n}(\phi_{\Delta(\tau,b)\otimes\sigma},s)$ must vanish at $s=0$.*

Proof. Note first that by [Arthur 2013], the Langlands–Shahidi normalization works for intertwining operators with tempered induced data at $s=0$. Hence we allow here that σ could be any irreducible tempered cuspidal automorphic representation if it is not generic.

Assume that $\beta_b(s)$, as defined in (1-1) or in Section 3A with more detail, has a pole at $s=0$. It implies that when $b=1$,

$$\beta_1(s)=L(2s+1,\tau,\rho)L(s+1,\tau\times\sigma)$$

has a pole at $s=0$, and when $b>1$, the only factor in $\beta_b(s)$ to have a possible pole at $s=0$ is $L(s,\tau,\rho^{(-)^{b+1}})$.

When $b=1$, $L(2s+1,\tau,\rho)$ and $L(s+1,\tau\times\sigma)$ both have at most a simple pole at $s=0$, but they cannot happen at the same time, since $L(s+1,\tau\times\sigma)$ having a simple pole at $s=0$ implies that $L(s,\tau,\rho^{-})$ has a pole at $s=1$. Therefore, $\beta_1(s)$ has at most a simple pole at $s=0$. When $b>1$, $L(s,\tau,\rho^{(-)^{b+1}})$ has at most a simple pole at $s=0$. So, for $b>1$, $\beta_b(s)$ also has at most a simple pole at $s=0$. Hence, $\beta_b(s)$ has at most a simple pole at $s=0$ for all $b\geq 1$. Now the assumption that $\beta_b(s)$ has a pole at $s=0$ implies that $\mathrm{ord}_{s=0}(\beta_b(s))=1$ for any $b\geq 1$, that is, $\beta_b(s)$ has a simple pole at $s=0$.

By the functional equation (3-1) and the normalized functional equation (3-6), we have

$$E^n_{ab}(\phi_{\Delta\otimes\sigma},s)=r(\omega',|\cdot|^s_F\Delta\otimes\sigma)E^n_{ab}\big(N(\omega',|\cdot|^s_F\Delta\otimes\sigma)(\phi),-s\big),$$

with

$$r(\omega',|\cdot|^s_F\Delta\otimes\sigma)=\frac{\beta_b(-s)}{\beta_b(s)}.$$

By the above discussion on the pole at $s=0$ of $\beta_b(s)$, it is clear that

$$r(\omega',|\cdot|^s_F\Delta\otimes\sigma)|_{s=0}=(-1)^{\mathrm{ord}_{s=0}(\beta_b(s))}=-1,$$

and hence

$$E^n_{ab}(\phi_{\Delta\otimes\sigma},s)|_{s=0}=-E^n_{ab}\big(N(\omega',|\cdot|^s_F\Delta\otimes\sigma)(\phi),-s\big)\big|_{s=0}. \tag{4-1}$$

So it suffices to show that the normalized intertwining operator $N(\omega',|\cdot|^s_F\Delta\otimes\sigma)$ is an identity map at $s=0$. We deduce this fact from the work of Arthur.

Arthur [2013, Corollary 7.3.5] proved that for the tempered or generic representation that has the Arthur parameter such that $\beta_b(s)$ has a simple pole at $s=0$, the normalized intertwining operator at $s=0$ has the identity

$$\lambda(\omega')\iota(\omega')\circ\mathcal{N}(\omega',\Delta\otimes\sigma)=\mathrm{Id},$$

where $\lambda(\omega')$ is the λ-factor (see for example [Keys and Shahidi 1988, Section 2]), $\iota(\omega')$ is a canonical map from $\omega'\Delta(\tau,b)\otimes\sigma$ to $\Delta(\tau,b)\otimes\sigma$ defined by Arthur [2013], and $\mathcal{N}(\omega',\Delta\otimes\sigma)$ is the evaluation at $s=0$ of the normalized intertwining operator from the induced representation $\mathrm{I}(\Delta\otimes\sigma,s)$ to $\mathrm{I}(\omega'\Delta\otimes\sigma,-s)$ (the vector-valued induced representations). The intertwining operator $N(\omega',|\cdot|^s_F\Delta\otimes\sigma)$ considered in this paper is a map from the space of automorphic forms

$$A\big(N_{ab}(\mathbb{A})M_{ab}(F)\backslash G(\mathbb{A})\big)_{|\cdot|^s_F\Delta\otimes\sigma}$$

to the space

$$A\big(N_{ab}(\mathbb{A})M_{ab}(F)\backslash G(\mathbb{A})\big)_{|\cdot|^{-s}_F\omega'\Delta\otimes\sigma}.$$

Note that by the strong multiplicity one theorem for GL_n and the definition of the isotypic component $A\big(N_{ab}(\mathbb{A})M_{ab}(F)\backslash G(\mathbb{A})\big)_{|\cdot|_F^s\Delta\otimes\sigma}$, this subspace depends only on the equivalence class of Δ, but not on its realization Δ in the space of automorphic forms. Therefore, we have the following relation between the two versions of the normalized intertwining operators at $s=0$:

$$N(\omega', \Delta\otimes\sigma) = \iota(\omega')\circ\mathcal{N}(\omega', \Delta\otimes\sigma).$$

Since the global λ-factor is trivial (see [Keys and Shahidi 1988, Section 2]), we have the following identity at $s=0$:

$$E^n_{ab}\big(N(\omega', |\cdot|_F^s\Delta\otimes\sigma)(\phi), 0\big) = E^n_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, 0).$$

By comparing with the identity (4-1), we obtain that $E^n_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ vanishes at $s=0$. This completes the proof. $\square$

Following from the definition of the normalized Eisenstein series, we have:

Corollary 4.3. *Let G_n be the symplectic group or the F-quasisplit special orthogonal group. Assume that σ is an irreducible generic* (*or tempered if nongeneric*) *cuspidal automorphic representation of $G_m(\mathbb{A})$. The normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic at the point $s=0$.*

By using Corollary 4.3 and the normalized induction formula in Proposition 3.2, we are able to prove Theorem 1.2 for the case of $b>1$, that is, to determine the location of possible poles of the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ for $b>1$. To do so, we consider the following four cases:

(1) $L(s,\tau,\rho)$ has a pole at $s=1$, and $L\big(\frac{1}{2},\tau\times\sigma\big)\neq 0$;

(2) $L(s,\tau,\rho)$ has a pole at $s=1$, and $L\big(\frac{1}{2},\tau\times\sigma\big)=0$;

(3) $L(s,\tau,\rho^-)$ has a pole at $s=1$, and $L(s,\tau\times\sigma)$ has a pole at $s=1$;

(4) $L(s,\tau,\rho^-)$ has a pole at $s=1$, and $L(s,\tau\times\sigma)$ is holomorphic at $s=1$.

We define the sets of possible poles according to the four cases:

$$X^+_{b,\tau,\sigma} := \begin{cases} \left\{\hat{0},\dots,\dfrac{b-2}{2},\dfrac{b}{2}\right\} & \text{in Case (1);}\\ \left\{\hat{0},\dots,\dfrac{b-4}{2},\dfrac{b-2}{2}\right\} & \text{in Case (2);}\\ \left\{\hat{0},\dots,\dfrac{b-1}{2},\dfrac{b+1}{2}\right\} & \text{in Case (3);}\\ \left\{\hat{0},\dots,\dfrac{b-3}{2},\dfrac{b-1}{2}\right\} & \text{in Case (4).} \end{cases}$$

When $b=1$, the set $X^+_{1,\tau,\sigma}$ is equal to the set $\left\{\frac{1}{2}\right\}$ in Case (1); is empty in Case (2); is equal to the set $\{1\}$ in Case (3); and is empty in Case (4). Hence the set $X^+_{b,\tau,\sigma}$ is the set of possible poles of the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ for $b = 1$ and $\mathrm{Re}(s) > 0$, by Proposition 4.1.

It is clear that when $b = 2$, the set $X^{+}_{b,\tau,\sigma}$ is also empty in Case (2). Note that we omit 0 in the set $X^{+}_{b,\tau,\sigma}$, since the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic at $s = 0$ (Corollary 4.3).

Here is the case of $b > 1$ and $m > 0$ of Theorem 1.2. The proof of this theorem for $\mathrm{Re}(s) \geq \frac{1}{2}$ is given by an induction argument, while the proof of this theorem for $0 < \mathrm{Re}(s) < \frac{1}{2}$ needs the Arthur classification [2013] of the discrete spectrum, which is stated here and will be proved in Section 6C.

Proposition 4.4 (case $0 < \mathrm{Re}(s) < \frac{1}{2}$ of Theorem 1.2). *Let G_n be the symplectic group or the F-quasisplit special orthogonal group. Assume that the irreducible cuspidal automorphic representation σ of $G_m(\mathbb{A})$ is generic and the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ is self-dual. Then $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic for $0 < \mathrm{Re}(s) < \frac{1}{2}$.*

With Propositions 4.1 and 4.4, Corollary 4.3, and the normalized induction formula (3-8), we are able to prove the following.

Theorem 4.5 (case $b > 1$ and $m > 0$ of Theorem 1.2). *Let G_n be the symplectic group or the F-quasisplit special orthogonal group. Assume that the irreducible cuspidal automorphic representation σ of $G_m(\mathbb{A})$ is generic and the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ is self-dual. Then $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic for $\mathrm{Re}(s) \geq 0$ except at $s = s_0 \in X^{+}_{b,\tau,\sigma}$, where it may have possibly at most simple poles.*

Proof. By Corollary 4.3 and Proposition 4.4, $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ is holomorphic at $0 \leq \mathrm{Re}(s) < \frac{1}{2}$, and hence we assume that $\mathrm{Re}(s) \geq \frac{1}{2}$ in the following discussion.

When $b = 1$, it is Proposition 4.1. We may assume that $b > 1$ and use the normalized induction formula (3-8):

$$\begin{aligned} &E^{n,*}_{ab,P_a}(\phi_{\Delta\otimes\sigma}, s)((I_a, h)) \\ &\qquad = L\big(2s+1, \tau, \rho^{(-)^{b+1}}\big) E^{n-a,*}_{a(b-1)}\big(\lambda_{-1/2}(i^{*}_{n-a}\phi_Q), s+\tfrac{1}{2}\big)(h) \\ &\qquad\qquad + \frac{L\big(2s, \tau, \rho^{(-)^{b+1}}\big)}{\varepsilon'_b(s)} E^{n-a,*}_{a(b-1)}\big(\lambda_{1/2}(i^{*}_{n-a} \circ N(\omega, \cdot))\tilde{\phi}, s-\tfrac{1}{2}\big)(h). \end{aligned}$$

When $\mathrm{Re}(s) \geq \frac{1}{2}$, the term

$$L\big(2s+1, \tau, \rho^{(-)^{b+1}}\big) E^{n-a,*}_{a(b-1)}\big(\lambda_{-1/2}(i^{*}_{n-a}\phi_Q), s+\tfrac{1}{2}\big)(h)$$

is holomorphic except for possible simple poles at $s_0 \in X^{+}_{b-1,\tau,\sigma} + \frac{1}{2}$, by the induction assumption.

The term

$$\frac{L\big(2s,\tau,\rho^{(-)^{b+1}}\big)}{\varepsilon_b'(s)} E^{n-a,*}_{a(b-1)}\big(\lambda_{1/2}(i^*_{n-a}\circ N(\omega,\cdot))\tilde{\phi}, s-\tfrac{1}{2}\big)(h)$$

is holomorphic for $\mathrm{Re}(s)\geq 1$ except for possible simple poles at $X^+_{b-1,\tau,\sigma}-\frac{1}{2}$, by the induction assumption, while at $\frac{1}{2}<\mathrm{Re}(s)<1$, it is holomorphic by Proposition 4.4. At $s=\frac{1}{2}$,

$$E^{n-a,*}_{a(b-1)}\big(\lambda_{1/2}(i^*_{n-a}\circ N(\omega,\cdot))\tilde{\phi}, s-\tfrac{1}{2}\big)(h)$$

is holomorphic by Corollary 4.3, while the L-function $L(2s,\tau,\rho^{(-)^{b+1}})$ may have a simple pole according to the classification of four cases on the parity of b, the type of τ, and the type of G_n in the Introduction.

Hence $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma},s)$ is holomorphic for $\mathrm{Re}(s)\geq\frac{1}{2}$ except for possible simple poles at $s_0\in\big(X^+_{b-1,\tau,\sigma}+\frac{1}{2}\big)\cup\big(X^+_{b-1,\tau,\sigma}-\frac{1}{2}\big)$ with $\mathrm{Re}(s_0)\geq\frac{1}{2}$. It is easy to check that

$$X^+_{b,\tau,\sigma}=\big(X^+_{b-1,\tau,\sigma}+\tfrac{1}{2}\big)\cup\big[\big(X^+_{b-1,\tau,\sigma}-\tfrac{1}{2}\big)\backslash\{0\}\big].$$

The theorem follows. □

This completes the proof of Theorem 1.2 for the case of $m>0$. We conclude this section with the following remarks.

(1) Theorem 1.2 holds for the F-quasisplit unitary groups if Corollary 4.3 is proven for the F-quasisplit unitary groups, which is done since Arthur's work has been extended to the F-quasisplit unitary groups [Mok 2012]. The extension of Arthur's classification of the discrete spectrum for F-quasisplit unitary groups will also imply that the complete Asai (and twisted Asai) L-functions are holomorphic in $0<s<1$ (as in [Grbac 2011] for symplectic or F-split special orthogonal groups), which is one of the key ingredients in the proof of Theorem 1.2 for $b=1$ and $m>0$.

(2) Theorem 1.2 is also expected to hold when σ is nongeneric, but tempered. The technical issue is the normalization of the local intertwining operators at all local places. At p-adic local fields, one can use Mœglin's work [2008; 2010]. Since her work at archimedean local places is not general enough to cover our cases, one needs more work, which will be considered in our future work.

(3) The current version of Theorem 1.2 is sufficient for our applications to the constructions of endoscopy correspondences considered in [Jiang 2011; 2012].

5. Proof of Theorem 1.2 ($m=0$)

In this case ($m=0$), G_n is either a symplectic group or an F-split special orthogonal group. When $b=1$, Theorem 1.2 for $m=0$ is given in [Grbac 2011, Theorem 3.1].

For $b > 1$, the proof of case $m = 0$ requires analogous results to Proposition 4.2 and Corollary 4.3, which are stated below. By the definition of $\beta_b(s)$ in this case as in Section 3C, the same proof works here.

Proposition 5.1. *Let G_n be a symplectic group or an F-split special orthogonal group. Assume that $b > 1$ and $m = 0$. If $\beta_b(s)$ has a pole at $s = 0$, then $E^n_{ab}(\phi_{\Delta(\tau,b)}, s)$ vanishes at $s = 0$. Moreover, the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)}, s)$ is holomorphic at the point $s = 0$.*

To determine the location of possible poles of the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)}, s)$ for $b > 1$, we consider the following four cases:

(1) $L(s, \tau, \rho)$ has a pole at $s = 1$, and $L\left(\frac{1}{2}, \tau\right) \neq 0$ if G_n is of Type (1);

(2) if G_n is of Type (1), then $L(s, \tau, \rho)$ has a pole at $s = 1$ and $L\left(\frac{1}{2}, \tau\right) = 0$;

(3) if G_n is of Type (1), then $L(s, \tau, \rho^-)$ has a pole at $s = 1$ and $L(s, \tau)$ has a pole at $s = 1$ (this case occurs only if $a = 1$ and τ is the trivial character of $\mathrm{GL}_1(\mathbb{A})$);

(4) $L(s, \tau, \rho^-)$ has a pole at $s = 1$, and $L(s, \tau)$ is holomorphic at $s = 1$ if G_n is of Type (1).

Note that in Type (1), $G_n = \mathrm{Sp}_{2n}$, and in Type (2), $G_n = \mathrm{SO}_{2n+1}$ or SO_{2n}. When $a = 1$ and τ is a quadratic character of $\mathrm{GL}_1(\mathbb{A})$, [Kudla and Rallis 1990; 1994] treat the case when $G_n = \mathrm{Sp}_{2n}$ or SO_{2n}.

Similarly, we define the sets of possible poles according to the four cases:

$$X^+_{b,\tau} := \begin{cases} \left\{\hat{0}, \dots, \frac{b-2}{2}, \frac{b}{2}\right\}, & \text{in Case (1);} \\ \left\{\hat{0}, \dots, \frac{b-4}{2}, \frac{b-2}{2}\right\}, & \text{in Case (2);} \\ \left\{\hat{0}, \dots, \frac{b-1}{2}, \frac{b+1}{2}\right\}, & \text{in Case (3);} \\ \left\{\hat{0}, \dots, \frac{b-3}{2}, \frac{b-1}{2}\right\}, & \text{in Case (4).} \end{cases}$$

We also omit 0 because $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)}, s)$ is holomorphic at $s = 0$ (Proposition 5.1).

Now the same inductive argument proves Theorem 1.2 for the case of $m = 0$ and $b > 1$. We omit the details here.

Theorem 5.2 (case $m = 0$ of Theorem 1.2). *Let G_n be a symplectic group or F-quasisplit orthogonal group. Assume that the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ is self-dual. Then the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)}, s)$ is holomorphic for $\mathrm{Re}(s) \geq 0$ except possibly at most simple poles at $s = s_0 \in X^+_{b,\tau}$.*

6. Residual representations and Arthur parameters

In this section, we assume that G_n is either symplectic or orthogonal, since we will use results from [Arthur 2013]. Based on Theorem 1.2, we will check the square-integrability for the residues at $s_0 \in X^+_{b,\tau,\sigma}$ of the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ (including the case of $m = 0$) in Section 6A, and write down the Arthur parameters for those square-integrable residual representations if they are nonzero in Section 6B. Based on Sections 6A and 6B, we prove Proposition 4.4 using the Arthur classification [2013] of discrete spectrum. Finally, we investigate the conditions for the nonvanishing of those residual representations.

6A. ***Square-integrability.*** We recall that $P_{a^b,m} = M_{a^b,m}N_{a^b,m}$ is the standard parabolic subgroup of G_n whose Levi subgroup is isomorphic to $\mathrm{GL}_a^{\times b} \times G_m$. Simply denote by $\Delta_b := \Delta_{M_{a^b,m}}$ the set of restricted simple roots that can be described as follows.

Let $\{e_i \mid 1 \le i \le b\}$ be the natural set of coordinates on $\mathrm{Re}\,\mathfrak{a}^*_{M_{a^b,m}}$. If G_m is not trivial, then

$$\Delta_b = \{e_1 - e_2, e_2 - e_3, \dots, e_{b-1} - e_b, e_b\}.$$

If G_m is trivial, then $\Delta_b = \Delta_0$, where Δ_0 is the set of simple roots of $R(T_0, G_b)$.

Recall the notation in Section I.3 of [Mœglin and Waldspurger 1995]. Let ϕ be an automorphic function and let $\Pi_0(M, \phi)$ be the cuspidal support of ϕ along $P = MN$. The cuspidal exponent $\mathrm{Re}(\pi)$ for π in $\Pi_0(M, \phi)$ is realized as a vector in terms of the basis $\{e_i \mid 1 \le i \le b\}$. Denote the cuspidal exponent of ϕ by

$$e(\phi) = \bigl\{\mathrm{Re}(\pi) \bigm| \text{for all } \pi \in \Pi_0(M, \phi) \text{ and for all } P = MN\bigr\}.$$

Let $e(s_0, b, \tau, \sigma)$ be the set of cuspidal exponents of the residues of the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ at $s = s_0$ belonging to the set $X^+_{b,\tau,\sigma}$.

By the square-integrability criterion [Mœglin and Waldspurger 1995, Lemma I.4.11], the residues of the Eisenstein series are square-integrable if and only if each character of cuspidal support can be written in the form

$$\sum_{\alpha\in\Delta_M} x_\alpha \alpha,$$

with coefficients $x_\alpha \in \mathbb{R}$, $x_\alpha < 0$. Moreover, in our cases the criterion is equivalent to, for all $\sum_{i=1}^{b} c_i e_i$ in $e(s_0, b, \tau, \sigma)$,

$$\sum_{i=1}^{j} c_i < 0 \quad \text{for all} \quad 1 \le j \le b. \tag{6-1}$$

Theorem 6.1 (square-integrability). *Let $s_0 \in \mathbb{C}$ such that* $\mathrm{Re}(s_0)$ *is in* $(0, (b+1)/2]$. *Assume that the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ has a simple pole* *at $s = s_0$. Then the residue of $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ at s_0 is square-integrable except at $s_0 = (b-1)/2$ in Case (3).*

Proof. The theorem is proved by induction on b. The key step in the proof is to determine the cuspidal exponents in $e(s_0, b, \tau, \sigma)$ by applying the induction formula (3-8), Lemma 2.1, and Lemma 2.2.

First, when $b = 1$, by Section 4A, if the Eisenstein series has a pole at $s_0 > 0$, then the cuspidal exponent of the residue of the Eisenstein series is $-s_0$ and the residue is square-integrable. By Proposition 4.1, the Eisenstein series is holomorphic at $\mathrm{Re}(s) \geq 0$ except at $s = \frac{1}{2}$ or $s = 1$. In these cases, the cuspidal exponent of the residues of the Eisenstein series is $-s_0 = -\frac{1}{2}$ or -1. Then $e(s_0, 1, \tau, \sigma)$ satisfies the condition (6-1) and the residues are square-integrable. Hence, the statement is true for $b = 1$.

Next, we assume that the statement holds for $b - 1$ and show that it is also true for b by induction.

By the induction formula (3-8), we have to consider the cuspidal exponents of the two terms

$$L\big(2s+1, \tau, \rho^{(-)^{b+1}}\big)E^{n-a,*}_{a(b-1)}\big(\cdot\,, s+\tfrac{1}{2}\big)$$

and

$$L\big(2s, \tau, \rho^{(-)^{b+1}}\big)E^{n-a,*}_{a(b-1)}\big(\cdot\,, s-\tfrac{1}{2}\big).$$

If $(b-1)/2 < \mathrm{Re}(s) \leq (b+1)/2$, the first term $E^{n-a,*}_{a(b-1)}\big(\cdot\,, s+\frac{1}{2}\big)$ is holomorphic. Since $b \geq 2$ and $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ has a pole at s_0, the second term $E^{n-a,*}_{a(b-1)}\big(\cdot\,, s-\frac{1}{2}\big)$ has a pole at s_0. By Lemma 2.2, the set $e(s_0, b, \tau, \sigma)$ of the cuspidal exponents equals

$$\left\{\left(-s_0 - \frac{b-1}{2}, c_1, \ldots, c_{b-1}\right) \,\middle|\, (c_1, \ldots, c_{b-1}) \in e\big(s_0 - \tfrac{1}{2}, b-1, \tau, \sigma\big)\right\}.$$

By induction, $e\big(s_0 - \frac{1}{2}, b-1, \tau, \sigma\big)$ satisfies the condition (6-1). It follows that $e(s_0, b, \tau, \sigma)$ also satisfies the condition (6-1). Hence the residue of the Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ at s_0 is square-integrable.

Next we consider the points at $0 < \mathrm{Re}(s) \leq (b-1)/2$. By the normalized induction formula (3-8), Lemma 2.1, and Lemma 2.2, the set of cuspidal exponents $e(s_0, b, \tau, \sigma)$ is a subset of the union

$$\begin{aligned} &\left\{\left(s_0 - \frac{b-1}{2}, c_1, \ldots, c_{b-1}\right) \,\middle|\, (c_1, \ldots, c_{b-1}) \in e\big(s_0 + \tfrac{1}{2}, b-1, \tau, \sigma\big)\right\} \\ &\cup \left\{\left(-s_0 - \frac{b-1}{2}, c_1, \ldots, c_{b-1}\right) \,\middle|\, (c_1, \ldots, c_{b-1}) \in e\big(s_0 - \tfrac{1}{2}, b-1, \tau, \sigma\big)\right\}. \end{aligned} \tag{6-2}$$

When $s_0 = \frac{1}{2}$, the set $e\big(s_0 - \frac{1}{2}, b-1, \tau, \sigma\big)$ needs some explanation. If $s_0 = \frac{1}{2}$ and $E^{n-a,*}_{a(b-1)}\big(\phi_{\Delta(\tau,b-1)\otimes\sigma}, s - \frac{1}{2}\big)$ in the second term vanishes at $s = s_0$, then the second term is holomorphic at $s=\frac{1}{2}$. Hence we do not need to consider this set of cuspidal exponents when we only consider the square-integrability for nonzero residues. On the other hand, if $E^{n-a,*}_{a(b-1)}\big(\phi_{\Delta(\tau,b-1)\otimes\sigma}, s-\frac{1}{2}\big)$ in the second term does not vanish at s_0, by Section 1B, then the cuspidal exponent of $E^{n-a,*}_{a(b-1)}\big(\cdot\,, s-\frac{1}{2}\big)$ at s_0 is

$$e\big(s_0-\tfrac{1}{2}, b-1, \tau, \sigma\big)=\Big\{\Big(\frac{2-b}{2}, \frac{4-b}{2}, \dots, \frac{b-2}{2}\Big)\Big\}.$$

When $s_0=(b-1)/2$ in Case (3), the residue of the first term of the normalized induction formula is nonzero due to the nonvanishing of the residue of

$$E^{n-a,*}_{a(b-1)}(\phi_{\Delta(\tau,b-1)\otimes\sigma}, s)$$

at $s=b/2$ in Theorem 6.2. Then $e(s_0, b, \tau, \sigma)$ contains the set

$$\Big\{\Big(s_0-\frac{b-1}{2}, c_1, \dots, c_{b-1}\Big) \,\Big|\, (c_1, \dots, c_{b-1})\in e\big(s_0+\tfrac{1}{2}, b-1, \tau, \sigma\big)\Big\},$$

which does not satisfy the condition (6-1), but satisfies $\sum_{i=1}^{j} c_i\le 0$.

When $s_0=(b-1)/2$, but not in Case (3), then the first term in the induction formula, $E^{n-a,*}_{a(b-1)}\big(\cdot\,, s+\frac{1}{2}\big)$, is holomorphic at $s=(b-1)/2$. Thus, only the second term $E^{n-a,*}_{a(b-1)}\big(\cdot\,, s-\frac{1}{2}\big)$ has a possible pole at $s=(b-1)/2$. Then $e(s_0, b, \tau, \sigma)$ equals

$$\Big\{\Big(-s_0-\frac{b-1}{2}, c_1, \dots, c_{b-1}\Big) \,\Big|\, (c_1, \dots, c_{b-1})\in e\big(s_0-\tfrac{1}{2}, b-1, \tau, \sigma\big)\Big\},$$

whose vectors satisfy the square-integrability criterion.

For $s_0<(b-1)/2$, we have $s_0-\frac{1}{2}<(b-2)/2$ and $s_0+\frac{1}{2}<b/2$. Since $-s_0-(b-1)/2<0$ and $s_0-(b-1)/2<0$, by induction, each vector in the set (6-2) satisfies the square-integrability criterion. This completes the proof. □

6B. ***Arthur parameters.*** From Theorem 6.1, the residual representations of $G_n(\mathbb{A})$ generated by the residues of the (normalized) Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ at $s=s_0$ belonging to the set $X^+_{b,\tau,\sigma}$ belong to the discrete spectrum of the space of automorphic forms on $G_n(\mathbb{A})$, except one case when $s_0=(b-1)/2$ for Case (3). Denote the residual representation by $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0}$.

We will figure out the Arthur parameters for those square-integrable residual representations $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0}$ if they are nonzero. Note that the nonvanishing conditions for those residual representations will be studied in the next subsection. We do this case by case for $s_0\in X^+_{b,\tau,\sigma}$.

We assume σ is an irreducible cuspidal automorphic representation of $G_m(\mathbb{A})$ with tempered global Arthur parameter ψ_σ [2013].

Case (1): In this case, when $m>0$, the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ has the property that $L(s,\tau,\rho)$ has a simple pole at

$s=1$ and $L\big(\frac{1}{2},\tau\times\sigma\big)\neq 0$, where ρ is the symmetric square representation of $\mathrm{GL}_a(\mathbb{C})$ if $G_n=\mathrm{SO}_{2n+1}$, and is the exterior square representation of $\mathrm{GL}_a(\mathbb{C})$ if G_n is Sp_{2n} or SO_{2n}.

We consider the residual representation $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0}$ of $G_n(\mathbb{A})$ at $s_0=(b-2j)/2$ with $j=0,1,\dots,[(b-1)/2]$. According to [Arthur 2013], the global Arthur parameter ψ attached to the residual representation $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0}$ of $G_n(\mathbb{A})$ is

$$\psi=\psi_{\Delta(\tau,b)\otimes\sigma,(b-2j)/2}=\big(\tau,2(b-j)\big)\boxplus(\tau,2j)\boxplus\psi_\sigma, \tag{6-3}$$

with $j=0,1,\dots,[(b-1)/2]$. Note that when G_n is SO_{2n+1}, τ is of orthogonal type; and when G_n is Sp_{2n} or SO_{2n}, τ is of symplectic type. Thus $\psi_{\Delta(\tau,b)\otimes\sigma,(b-2j)/2}$ is a global Arthur parameter for G_n. When $m=0$, we have

$$\psi=\psi_{\Delta(\tau,b)\otimes\sigma,(b-2j)/2}=\begin{cases}(\tau,2(b-j))\boxplus(\tau,2j) & \text{if } G_n\neq\mathrm{Sp}_{2n},\\ (\tau,2(b-j))\boxplus(\tau,2j)\boxplus(1_{\mathrm{GL}_1(\mathbb{A})},1) & \text{if } G_n=\mathrm{Sp}_{2n}.\end{cases}$$

Case (2): This case is the same as Case (1), and the only difference is that $s_0=(b-2j)/2$ with $j=1,2,\dots,[(b-1)/2]$. Hence when $m>0$, the global Arthur parameter ψ attached to the residual representation $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0}$ of $G_n(\mathbb{A})$ is

$$\psi=\psi_{\Delta(\tau,b)\otimes\sigma,(b-2j)/2}=\big(\tau,2(b-j)\big)\boxplus(\tau,2j)\boxplus\psi_\sigma, \tag{6-4}$$

with $j=1,2,\dots,[(b-1)/2]$; and when $m=0$, we have

$$\psi=\psi_{\Delta(\tau,b)\otimes\sigma,(b-2j)/2}=\begin{cases}\big(\tau,2(b-j)\big)\boxplus(\tau,2j) & \text{if } G_n\neq\mathrm{Sp}_{2n},\\ (\tau,2(b-j))\boxplus(\tau,2j)\boxplus(1_{\mathrm{GL}_1(\mathbb{A})},1) & \text{if } G_n=\mathrm{Sp}_{2n}.\end{cases}$$

Case (3): In this case, when $m>0$, the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ has the property that $L(s,\tau,\rho^-)$ has a simple pole at $s=1$ and $L(s,\tau\times\sigma)$ also has a simple pole at $s=1$, where ρ^- is the exterior square representation of $\mathrm{GL}_a(\mathbb{C})$ if $G_n=\mathrm{SO}_{2n+1}$, and is the symmetric square representation of $\mathrm{GL}_a(\mathbb{C})$ if G_n is Sp_{2n} or SO_{2n}. Following [Arthur 2013], the global tempered Arthur parameter for the irreducible cuspidal automorphic representation σ is

$$\psi_\sigma=(\tau,1)\boxplus\psi', \tag{6-5}$$

where ψ' is a global Arthur parameter that is the complement of $(\tau,1)$ in ψ_σ.

We consider the residual representation $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0}$ of $G_n(\mathbb{A})$ at

$$s_0=\frac{b+1-2j}{2}$$

with $j=0,1,\dots,[b/2]$. According to [Arthur 2013], the global Arthur parameter ψ attached to the residual representation $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0}$ of $G_n(\mathbb{A})$ is

$$\psi=\psi_{\Delta(\tau,b)\otimes\sigma,(b+1-2j)/2}=(\tau,2b+1-2j)\boxplus(\tau,2j-1)\boxplus\psi_\sigma, \tag{6-6}$$

with $j = 2, 3, \dots, [b/2]$. If $j = 0$, we have

$$\psi = \psi_{\Delta(\tau,b)\otimes\sigma,(b+1)/2} = (\tau, 2b+1) \boxplus \psi'. \tag{6-7}$$

Note that when $j = 1$, the residual representation $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0=(b-1)/2}$ of $G_n(\mathbb{A})$ is not square-integrable (Theorem 6.1). Note that when G_n is SO_{2n+1}, τ is of symplectic type; and when G_n is Sp_{2n} or SO_{2n}, τ is of orthogonal type. Hence $\psi_{\Delta(\tau,b)\otimes\sigma,(b-2j)/2}$ with $j = 0$ or $j = 2, 3, \dots, [b/2]$ is a global Arthur parameter for G_n.

When $m = 0$, this case only occurs if $a = 1$ and τ is the trivial representation of $\mathrm{GL}_1(\mathbb{A})$. If $j = 2, 3, \dots, [b/2]$, we have

$$\psi = \psi_{\Delta(\tau,b)\otimes\sigma,(b+1-2j)/2} = (\tau, 2b+1-2j) \boxplus (\tau, 2j-1) \boxplus (1_{\mathrm{GL}_1(\mathbb{A})}, 1).$$

If $j = 0$, according to the definition of the four cases, G_n must be Sp_{2n} and $\psi = \psi_{\Delta(\tau,b)\otimes\sigma,(n+1)/2} = (\tau, 2n+1)$.

Case (4): This case is similar to Case (3). The only difference is that

$$s_0 = \frac{b+1-2j}{2}$$

with $j = 1, 2, \dots, [b/2]$. Hence when $m > 0$, the global Arthur parameter ψ attached to the residual representation $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,s_0}$ of $G_n(\mathbb{A})$ is

$$\psi = \psi_{\Delta(\tau,b)\otimes\sigma,(b+1-2j)/2} = (\tau, 2b+1-2j) \boxplus (\tau, 2j-1) \boxplus \psi_\sigma, \tag{6-8}$$

with $j = 1, 2, 3, \dots, [b/2]$, and when $m = 0$, we have

$$\begin{aligned} \psi &= \psi_{\Delta(\tau,b)\otimes\sigma,(b+1-2j)/2} \\ &= \begin{cases} (\tau, 2b+1-2j) \boxplus (\tau, 2j-1) & \text{if } G_n \neq \mathrm{Sp}_{2n}, \\ (\tau, 2b+1-2j) \boxplus (\tau, 2j-1) \boxplus (1_{\mathrm{GL}_1(\mathbb{A})}, 1) & \text{if } G_n = \mathrm{Sp}_{2n}. \end{cases} \end{aligned}$$

Note that the residual representation $\mathscr{E}_{\Delta(\tau,b)\otimes\sigma,(b-1)/2}$ of $G_n(\mathbb{A})$ ($j = 1$) in this case belongs to the discrete spectrum of $G_n(\mathbb{A})$.

6C. ***Proof of Proposition 4.4.*** Proposition 4.4 follows from the discussion on square-integrability in Section 6A and the discussion on the global Arthur parameter in Section 6B. In fact, if there is an s_0 such that $0 < \mathrm{Re}(s_0) < \frac{1}{2}$, such that the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ has a pole at $s = s_0$, then by Theorem 6.1, the residue at $s = s_0$ must be square-integrable, and hence the residual representation contributes to the discrete spectrum. On the other hand, by the Arthur classification [2013] of the discrete spectrum, there is no global Arthur parameter for G_n that parametrizes such a residual representation. Hence the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$ must be holomorphic at $0 < \mathrm{Re}(s) < \frac{1}{2}$. This proves Proposition 4.4.

6D. ***Nonvanishing conditions.*** When $b=1$, the nonvanishing of the residues of the normalized Eisenstein series has been discussed in Proposition 4.1. In the following, we assume that $b>1$.

For $s=s_0\in X^+_{b,\tau,\sigma}$, the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma},s)$ has a pole at $s=s_0$ if one of its constant terms has a pole at $s=s_0$. The normalized induction formula (3-8) says

$$\begin{aligned}E^{n,*}_{ab,P_a}(\phi_{\Delta\otimes\sigma},s)((I_a,h))\\ &= L\big(2s+1,\tau,\rho^{(-)^{b+1}}\big)E^{n-a,*}_{a(b-1)}\big(\lambda_{-1/2}i^*_{n-a}\phi_Q,s+\tfrac{1}{2}\big)(h)\\ &\qquad+\frac{L\big(2s,\tau,\rho^{(-)^{b+1}}\big)}{\varepsilon'_b(s)}E^{n-a,*}_{a(b-1)}\big(\lambda_{1/2}(i^*_{n-a}\circ N(\omega,\cdot))\tilde{\phi},s-\tfrac{1}{2}\big)(h).\end{aligned}$$

Hence s_0 has the property that $s_0\in\big(X^+_{b-1,\tau,\sigma}+\frac{1}{2}\big)\cup\big(X^+_{b-1,\tau,\sigma}-\frac{1}{2}\big)$ and $s_0>0$.

By the discussion of the global Arthur parameters in Section 6B, if both

$$E^{n-a,*}_{a(b-1)}\big(\lambda_{-1/2}i^*_{n-a}\phi_Q,s+\tfrac{1}{2}\big)(h)$$

and

$$E^{n-a,*}_{a(b-1)}\big(\lambda_{1/2}(i^*_{n-a}\circ N(\omega,\cdot))\tilde{\phi},s-\tfrac{1}{2}\big)(h)$$

are nonzero, they cannot be proportional to each other since they have different Langlands parameters. Hence the problem reduces to verifying the nonvanishing of either of the two terms. Note that from the definition, both $\lambda_{-1/2}i^*_{n-a}\phi_Q$ and $\lambda_{1/2}(i^*_{n-a}\circ N(\omega,\cdot))\tilde{\phi}$ give general sections in the corresponding space, respectively. Therefore, the existence of the poles of the normalized Eisenstein series $E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma},s)$ at $s\in X^+_{b,\tau,\sigma}$ follows from the existence of the poles of the normalized Eisenstein series $E^{n,*}_{a(b-1)}(\phi_{\Delta(\tau,b-1)\otimes\sigma},s)$ at $s\in X^+_{b-1,\tau,\sigma}$.

By repeating the argument, this reduces to the case of b being as small as possible. The discussion will be given for each of the four cases.

Case (1): In this case, the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ has the property that $L(s,\tau,\rho)$ has a simple pole at $s=1$ and

$$L\big(\tfrac{1}{2},\tau\times\sigma\big)\neq 0,$$

where ρ is the symmetric square representation of $\mathrm{GL}_a(\mathbb{C})$ if $G_n=\mathrm{SO}_{2n+1}$, and the exterior square representation of $\mathrm{GL}_a(\mathbb{C})$ if G_n is Sp_{2n} or SO_{2n}. In this case, the smallest possible value of b is $b=1$. The existence of the pole at the only value $s=\frac{1}{2}$ is treated in the first case in Proposition 4.1.

Case (2): In this case, the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ has the property that $L(s,\tau,\rho)$ has a simple pole at $s=1$ and

$$L\big(\tfrac{1}{2},\tau\times\sigma\big)=0,$$

where ρ is the symmetric square representation of $\mathrm{GL}_a(\mathbb{C})$ if $G_n = \mathrm{SO}_{2n+1}$, and the exterior square representation of $\mathrm{GL}_a(\mathbb{C})$ if G_n is Sp_{2n} or SO_{2n}. In this case, the smallest possible value of b is $b = 3$, which leads us to consider the existence of the pole at $s = \frac{1}{2}$. The normalized induction formula is

$$\begin{aligned} E^{n,*}_{3a,P_a}(\phi_{\Delta\otimes\sigma}, s)((I_a, h)) &= L(2s+1, \tau, \rho) E^{n-a,*}_{2a}\big(\lambda_{-1/2} i^*_{n-a}\phi_Q, s+\tfrac{1}{2}\big)(h) \\ &\quad + \frac{L(2s, \tau, \rho)}{\varepsilon'_3(s)} E^{n-a,*}_{2a}\big(\lambda_{1/2}(i^*_{n-a} \circ N(\omega, \cdot))\tilde{\phi}, s-\tfrac{1}{2}\big)(h). \end{aligned}$$

It follows from Theorem 1.2 that the first term

$$L(2s+1, \tau, \rho) E^{n-a,*}_{2a}\big(\lambda_{-1/2} i^*_{n-a}\phi_Q, s+\tfrac{1}{2}\big)(h)$$

is holomorphic at $s = \frac{1}{2}$, since $L(2s+1, \tau, \rho)$ is holomorphic at $s = \frac{1}{2}$ and $s + \frac{1}{2} = 1$ does not belong to the empty set $X^+_{\tau,2,\sigma}$ in the case of the normalized Eisenstein series $E^{n-a,*}_{2a}\big(\lambda_{-1/2} i^*_{n-a}\phi_Q, s+\frac{1}{2}\big)(h)$.

The second term

$$\frac{L(2s, \tau, \rho)}{\varepsilon'_3(s)} E^{n-a,*}_{2a}\big(\lambda_{1/2}(i^*_{n-a} \circ N(\omega, \cdot))\tilde{\phi}, s-\tfrac{1}{2}\big)(h)$$

has a simple pole at $s = \frac{1}{2}$ if and only if the normalized Eisenstein series

$$E^{n-a,*}_{2a}(\phi_{\Delta(\tau,2)\otimes\sigma}, s)$$

is not identically zero at $s = 0$.

Case (3): In this case, the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ has the property that $L(s, \tau, \rho^-)$ has a simple pole at $s = 1$ and that $L(s, \tau \times \sigma)$ also has a simple pole at $s = 1$, where ρ^- is the exterior square representation of $\mathrm{GL}_a(\mathbb{C})$ if $G_n = \mathrm{SO}_{2n+1}$, and the symmetric square representation of $\mathrm{GL}_a(\mathbb{C})$ if G_n is Sp_{2n} or SO_{2n}. In this case, the smallest possible value of b is $b = 1$. The existence of the pole at the only value $s = 1$ is treated in the second case in Proposition 4.1.

Case (4): In this case, the irreducible unitary cuspidal automorphic representation τ of $\mathrm{GL}_a(\mathbb{A})$ has the property that $L(s, \tau, \rho^-)$ has a simple pole at $s = 1$ and $L(s, \tau \times \sigma)$ is holomorphic at $s = 1$, where ρ^- is the exterior square representation of $\mathrm{GL}_a(\mathbb{C})$ if $G_n = \mathrm{SO}_{2n+1}$, and the symmetric square representation of $\mathrm{GL}_a(\mathbb{C})$ if G_n is Sp_{2n} or SO_{2n}. In this case, the smallest possible value of b is $b = 2$, which leads us to consider the existence of the pole at $s = \frac{1}{2}$. The normalized induction formula is

$$
\begin{aligned}
&E^{n,*}_{2a,P_a}(\phi_{\Delta\otimes\sigma}, s)((I_a, h)) \\
&\qquad = L(2s+1, \tau, \rho^-) E^{n-a,*}_a\big(\lambda_{-1/2} i^*_{n-a}\phi_Q, s+\tfrac{1}{2}\big)(h) \\
&\qquad\qquad + \frac{L(2s, \tau, \rho^-)}{\varepsilon'_2(s)} E^{n-a,*}_a\big(\lambda_{1/2}(i^*_{n-a}\circ N(\omega, \cdot))\tilde{\phi}, s-\tfrac{1}{2}\big)(h).
\end{aligned}
$$

It follows from Theorem 1.2 that the first term

$$L(2s+1, \tau, \rho^-) E^{n-a,*}_a\big(\lambda_{-1/2} i^*_{n-a}\phi_Q, s+\tfrac{1}{2}\big)$$

is holomorphic at $s = \frac{1}{2}$ with the same argument as in Case (2). The second term

$$\frac{L(2s, \tau, \rho^-)}{\varepsilon'_2(s)} E^{n-a,*}_a\big(\lambda_{1/2}(i^*_{n-a}\circ N(\omega, \cdot))\tilde{\phi}, s-\tfrac{1}{2}\big)$$

has a simple pole at $s = \frac{1}{2}$ if and only if the normalized Eisenstein series

$$E^{n-a,*}_a(\phi_{\tau\otimes\sigma}, s)$$

does not vanish identically at $s = 0$.

The above discussion leads to the following theorem.

Theorem 6.2. *With the notation above, the following hold.*

(1) *Assume that $L\big(\frac{1}{2}, \tau\times\sigma\big) \neq 0$ and $L(s, \tau, \rho)$ has a simple pole at $s = 1$. The normalized Eisenstein series*

$$E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$$

has a simple pole at each $s \in X^+_{b,\tau,\sigma}$, which is defined in Case (1).

(2) *Assume that $L\big(\frac{1}{2}, \tau\times\sigma\big) = 0$ and $L(s, \tau, \rho)$ has a simple pole at $s = 1$. The normalized Eisenstein series*

$$E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$$

has a simple pole at each $s \in X^+_{b,\tau,\sigma}$, which is defined in Case (2), if and only if the normalized Eisenstein series

$$E^{n-a,*}_{2a}(\phi_{\Delta(\tau,2)\otimes\sigma}, s)$$

is not identically zero at $s = 0$.

(3) *Assume that $L(s, \tau, \rho^-)$ and $L(s, \tau\times\sigma)$ have a simple pole at $s = 1$. The normalized Eisenstein series*

$$E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$$

has a simple pole at each $s \in X^+_{b,\tau,\sigma}$, which is defined in Case (3).

(4) *Assume that $L(s, \tau, \rho^-)$ has a simple pole at $s = 1$ and $L(s, \tau \times \sigma)$ is holomorphic at $s = 1$. The normalized Eisenstein series*

$$E^{n,*}_{ab}(\phi_{\Delta(\tau,b)\otimes\sigma}, s)$$

has a simple pole at each $s \in X^+_{b,\tau,\sigma}$, which is defined in Case (4), if and only if the normalized Eisenstein series

$$E^{n-a,*}_{a}(\phi_{\tau\otimes\sigma}, s)$$

does not vanish identically at $s = 0$.

Remark 6.3. The nonvanishing results may also apply to the case of $m = 0$ accordingly, but we omit the discussion here. Finally, it is natural to expect that the results discussed in this section hold for quasisplit unitary groups.

Acknowledgements

We would like to thank the referee for his/her careful reading of the previous version of this paper and very helpful comments and suggestions. We would also like to thank Erez Lapid for helpful conversations and communications on certain technical issues of this paper.

References

[Arthur 2013] J. Arthur, *The endoscopic classification of representations: orthogonal and symplectic groups*, Colloquium Publ., Amer. Math. Soc., Providence, 2013. To appear; preprint available at http://www.claymath.org/cw/arthur/pdf/Book.pdf.

[Brenner 2009] E. Brenner, "Analytic properties of residual Eisenstein series, I", *J. Ramanujan Math. Soc.* **24**:4 (2009), 359–413. MR 2011a:11102 Zbl 1207.11060

[Cogdell and Piatetski-Shapiro 2004] J. W. Cogdell and I. I. Piatetski-Shapiro, "Remarks on Rankin–Selberg convolutions", pp. 255–278 in *Contributions to automorphic forms, geometry, and number theory* (Baltimore, MD, 2002), edited by H. Hida et al., Johns Hopkins University Press, Baltimore, MD, 2004. MR 2005d:11075 Zbl 1080.11038

[Cogdell et al. 2004] J. W. Cogdell, H. H. Kim, I. I. Piatetski-Shapiro, and F. Shahidi, "Functoriality for the classical groups", *Publ. Math. Inst. Hautes Études Sci.* **99** (2004), 163–233. MR 2006a:22010 Zbl 1090.22010

[Cogdell et al. 2011] J. W. Cogdell, I. I. Piatetski-Shapiro, and F. Shahidi, "Functoriality for the quasisplit classical groups", pp. 117–140 in *On certain L-functions* (West Lafayette, IN, 2007), edited by J. Arthur et al., Clay Math. Proc. **13**, Amer. Math. Soc., Providence, RI, 2011. MR 2012f:22036 Zbl 1235.22025

[Ginzburg 2008] D. Ginzburg, "Endoscopic lifting in classical groups and poles of tensor L-functions", *Duke Math. J.* **141**:3 (2008), 447–503. MR 2009c:22020 Zbl 1195.11068

[Ginzburg et al. 2011] D. Ginzburg, S. Rallis, and D. Soudry, *The descent map from automorphic representations of* $\mathrm{GL}(n)$ *to classical groups*, World Scientific, Singapore, 2011. MR 2012g:22020 Zbl 1233.11056

[Ginzburg et al. 2012] D. Ginzburg, D. Jiang, and D. Soudry, "On correspondences between certain automorphic forms on $\mathrm{Sp}_{4n}(\mathbb{A})$ and $\widetilde{\mathrm{Sp}}_{2n}(\mathbb{A})$", *Israel J. Math.* **192**:2 (2012), 951–1008. MR 3009748 Zbl pre06135123

[Grbac 2011] N. Grbac, "On the residual spectrum of split classical groups supported in the Siegel maximal parabolic subgroup", *Monatsh. Math.* **163**:3 (2011), 301–314. MR 2805875 Zbl 05913913

[Jiang 1998] D. Jiang, "The first term identities for Eisenstein series", *J. Number Theory* **70**:1 (1998), 67–98. MR 99h:11052 Zbl 0913.11021

[Jiang 2011] D. Jiang, "Construction of endoscopy transfers for classical groups", in *Automorphic forms: New directions*, Oberwolfach Reports, EMS **8**:1, March 2011. Available at http://www.math.umn.edu/~dhjiang/Papers/oberwolfach11.pdf.

[Jiang 2012] D. Jiang, "Automorphic integral transforms for classical groups I: endoscopy correspondences", preprint, 2012. To appear in *Proceedings of the Conference on Automorphic Forms and Related Geometry: Assessing the legacy of I. I. Piatetski-Shapiro*, edited by J. Cogdell et al. arXiv 1212.6525

[Jiang and Liu 2012] D. Jiang and B. Liu, "On Fourier coefficients of automorphic forms of $\mathrm{GL}(n)$", *Int. Math. Res. Not.* **2012** (2012), article ID rns153. doi: 10.1093/imrn/rns153.

[Keys and Shahidi 1988] C. D. Keys and F. Shahidi, "Artin L-functions and normalization of intertwining operators", *Ann. Sci. École Norm. Sup.* (4) **21**:1 (1988), 67–89. MR 89k:22034 Zbl 0654.10030

[Kim and Krishnamurthy 2004] H. H. Kim and M. Krishnamurthy, "Base change lift for odd unitary groups", pp. 116–125 in *Functional analysis VIII* (Dubrovnik, 2003), edited by D. Bakić et al., Various Publ. Ser. (Aarhus) **47**, Aarhus Univ., Aarhus, 2004. MR 2006a:11154 Zbl 1146.11313

[Kim and Krishnamurthy 2005] H. H. Kim and M. Krishnamurthy, "Stable base change lift from unitary groups to GL_n", *Int. Math. Res. Pap.* **2005**:1 (2005), 1–52. MR 2006d:22028 Zbl 1146.11314

[Kudla and Rallis 1990] S. S. Kudla and S. Rallis, "Poles of Eisenstein series and L-functions", pp. 81–110 in *Festschrift in honor of I. I. Piatetski-Shapiro on the occasion of his sixtieth birthday, II* (Tel Aviv, 1989), edited by S. Gelbart et al., Israel Math. Conf. Proc. **3**, Weizmann, Jerusalem, 1990. MR 94e:11054 Zbl 0712.11029

[Kudla and Rallis 1994] S. S. Kudla and S. Rallis, "A regularized Siegel–Weil formula: the first term identity", *Ann. of Math.* (2) **140**:1 (1994), 1–80. MR 95f:11036 Zbl 0818.11024

[Langlands 1976] R. P. Langlands, *On the functional equations satisfied by Eisenstein series*, Lecture Notes in Mathematics **544**, Springer, Berlin, 1976. MR 58 #28319 Zbl 0332.10018

[Lapid et al. 2004] E. Lapid, G. Muić, and M. Tadić, "On the generic unitary dual of quasisplit classical groups", *Int. Math. Res. Not.* **2004**:26 (2004), 1335–1354. MR 2005b:22021 Zbl 1079.22015

[Luo et al. 1999] W. Luo, Z. Rudnick, and P. Sarnak, "On the generalized Ramanujan conjecture for $\mathrm{GL}(n)$", pp. 301–310 in *Automorphic forms, automorphic representations, and arithmetic* (Fort Worth, TX, 1996), edited by R. S. Doran et al., Proc. Sympos. Pure Math. **66**, Part 2, Amer. Math. Soc., Providence, RI, 1999. MR 2000e:11072 Zbl 0965.11023

[Mœglin 2008] C. Mœglin, "Formes automorphes de carré intégrable non cuspidales", *Manuscripta Math.* **127**:4 (2008), 411–467. MR 2010i:11071 Zbl 05520091

[Mœglin 2010] C. Mœglin, "Holomorphie des opérateurs d'entrelacement normalisés à l'aide des paramètres d'Arthur", *Canad. J. Math.* **62**:6 (2010), 1340–1386. MR 2012c:22024 Zbl 1204.22012

[Mœglin and Waldspurger 1989] C. Mœglin and J.-L. Waldspurger, "Le spectre résiduel de $\mathrm{GL}(n)$", *Ann. Sci. École Norm. Sup.* (4) **22**:4 (1989), 605–674. MR 91b:22028 Zbl 0696.10023

[Mœglin and Waldspurger 1995] C. Mœglin and J.-L. Waldspurger, *Spectral decomposition and Eisenstein series*, Cambridge Tracts in Mathematics **113**, Cambridge University Press, Cambridge, 1995. MR 97d:11083 Zbl 0846.11032

[Mok 2012] C. P. Mok, "Endoscopic classification of representations of quasi-split unitary groups I", preprint, 2012. arXiv 1206.0882

[Shahidi 2010] F. Shahidi, *Eisenstein series and automorphic L-functions*, Amer. Math. Soc. Colloq. Publ. **58**, Amer. Math. Soc., Providence, RI, 2010. MR 2012d:11119 Zbl 1215.11054

Received January 27, 2012. Revised July 20, 2012.

DIHUA JIANG
SCHOOL OF MATHEMATICS
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN 55455
UNITED STATES
dhjiang@math.umn.edu

BAIYING LIU
SCHOOL OF MATHEMATICS
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN 55455
UNITED STATES
liuxx969@umn.edu

LEI ZHANG
DEPARTMENT OF MATHEMATICS
BOSTON COLLEGE
CARNEY HALL, 140 COMMONWEALTH AVE
CHESTNUT HILL, MA 02467
UNITED STATES
lei.zhang.2@bc.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.125

HARMONIC MAPS ON DOMAINS WITH PIECEWISE LIPSCHITZ CONTINUOUS METRICS

HAIGANG LI AND CHANGYOU WANG

We study harmonic maps $(\Omega, g) \to (N, h)$, where $\Omega \subset \mathbb{R}^n$ is a bounded domain divided into two pieces, the Riemannian metric g is Lipschitz in each piece, and (N, h) is a closed Riemannian submanifold of $\mathbb{R}^k$. We prove the partial regularity of stationary harmonic maps, and the global Lipschitz and piecewise $C^{1,\alpha}$-regularity of weakly harmonic maps from (Ω, g) to manifolds (N, h) that support convex distance square functions.

1. Introduction

Throughout this paper we assume that $\Omega = \Omega^+ \cup \Omega^- \cup \Gamma$ is a bounded domain of $\mathbb{R}^n$ decomposed into two subdomains Ω^+ and Ω^- by a $C^{1,1}$-hypersurface Γ, and that g is a piecewise Lipschitz metric on Ω, satisfying $g \in C^{0,1}(\Omega^+) \cap C^{0,1}(\Omega^-)$ and discontinuous at every $x \in \Gamma$. For example, let $\Omega = B_1 \subset \mathbb{R}^n$ be the unit ball, $\Gamma = B_1 \cap \{x = (x', 0) \in \mathbb{R}^n\}$, and

$$\bar{g}(x) = \begin{cases} g_0 & \text{if } x \in B_1^+ = \{x^n > 0\} \cap B_1, \\ k g_0 & \text{if } x \in B_1^- = \{x^n < 0\} \cap B_1, \end{cases}$$

where g_0 is the standard metric on $\mathbb{R}^n$ and k ($\neq 1$) is a positive constant. Let $(N, h) \hookrightarrow \mathbb{R}^k$ be an l-dimensional, smooth compact Riemannian manifold without boundary, isometrically embedded in the Euclidean space $\mathbb{R}^k$.

Motivated by the recent studies on elliptic systems arising from composite materials (see [Li and Nirenberg 2003]) and the periodic homogenization theory in calculus of variations (see [Avellaneda and Lin 1987] and [Lin and Yan 2003]), we are interested in the regularity issue of harmonic maps from (Ω, g) to (N, h).

In order to describe the problem, let's recall some notations. Throughout this paper, we use the Einstein convention for summation. For the metric $g = g_{ij}\, dx^i\, dx^j$, let $(g^{ij}) = (g_{ij})^{-1}$, and $dv_g = \sqrt{g}\, dx (= \sqrt{\det(g_{ij})}\, dx)$ be the volume form of g. For $1 < p < +\infty$, define the Sobolev space

$$W^{1,p}(\Omega, N) = \left\{ u : \Omega \to \mathbb{R}^k \;\middle|\; u(x) \in N \text{ a.e. } x \in \Omega,\ E_p(u, g) = \int_\Omega |\nabla u|_g^p\, dv_g < \infty \right\},$$

MSC2010: 35J50, 58E20.
Keywords: piecewise continuous metric, harmonic map, regularity.

where

$$|\nabla u|_g^2 \equiv g^{ij}\left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_j}\right\rangle$$

is the energy density of u with respect to g, and $\langle \cdot\,, \cdot \rangle$ denotes the inner product in $\mathbb{R}^k$. Denote $W^{1,2}(\Omega, N)$ by $H^1(\Omega, N)$. Now let's recall the definition of stationary harmonic maps.

Definition 1.1. A map $u \in H^1(\Omega, N)$ is called a (weakly) harmonic map if it is a critical point of $E_2(\cdot\,, g)$, i.e., if u satisfies

$$\Delta_g u + A(u)(\nabla u, \nabla u)_g = 0 \quad \text{in } \Omega \tag{1-1}$$

in the sense of distributions. Here

$$\Delta_g = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x_i}\left(\sqrt{g}g^{ij}\frac{\partial}{\partial x_j}\right)$$

is the Laplace–Beltrami operator on (Ω, g), $A(\cdot)(\cdot\,, \cdot)$ is the second fundamental form of $(N, h) \hookrightarrow \mathbb{R}^k$, and $A(u)(\nabla u, \nabla u)_g = g^{ij} A(u)(\partial u/\partial x_i, \partial u/\partial x_j)$.

Definition 1.2. A (weakly) harmonic map $u \in H^1(\Omega, N)$ is called a stationary harmonic map if, in addition, it is a critical point of $E_2(\cdot\,, g)$ with respect to the following domain variations:

$$\frac{d}{dt}\bigg|_{t=0}\int_\Omega \left|\nabla u^t\right|_g^2\, dv_g = 0, \quad \text{with } u^t(x) = u\big(F_t(x)\big), \tag{1-2}$$

where $F(t, x) := F_t(x) \in C^1\big([-\delta, \delta], C^1(\Omega, \Omega)\big)$, for some small $\delta > 0$, is a family of diffeomorphisms that satisfies

$$\begin{cases} F_0(x) = x & \text{for } x \in \Omega, \\ F_t(x) = x & \text{for } (x, t) \in \partial\Omega \times [-\delta, \delta], \\ F_t\big(\overline{\Omega^\pm}\big) \subset \overline{\Omega^\pm} & \text{for } t \in [-\delta, \delta]. \end{cases} \tag{1-3}$$

In particular, $F_t(\Gamma) \subset \Gamma$ for $0 \le t \le \delta$.

It is readily seen that any minimizing harmonic map from (Ω, g) to (N, h) is a stationary harmonic map. Definition 1.2 implies that a stationary harmonic map on (Ω, g) is a stationary harmonic map on $(\Omega^\pm, g)$. Since $g \in C^{0,1}(\Omega^\pm)$, we can see that u satisfies an energy monotonicity inequality on $\Omega^\pm$. We will show in Section 2 that a stationary harmonic map on (Ω, g) also satisfies an energy monotonicity inequality in Γ under the condition (1-4) below.

The first result is concerned with the (partial) Lipschitz and (partial) piecewise $C^{1,\alpha}$-regularity of stationary harmonic maps. In this context, we are able to extend the well-known partial regularity theorem of stationary harmonic maps on domains with smooth metrics, due to Hélein [2002], Evans [1991], and Bethuel [1993]. More precisely:

Theorem 1.1. *Let $u \in H^1(\Omega, N)$ be a stationary harmonic map on (Ω, g). Suppose that g satisfies the following jump condition on Γ for $n \geq 3$: for any $x \in \Gamma$, there exists a positive constant $k(x) \neq 1$ such that*

$$\lim_{\substack{y\in\Omega^+\\ y\to x}} g(y) = k(x) \lim_{\substack{y\in\Omega^-\\ y\to x}} g(y). \tag{1-4}$$

There exists a closed set $\Sigma \subset \Omega$, with $H^{n-2}(\Sigma) = 0$, such that $u \in \mathrm{Lip}_{\mathrm{loc}}(\Omega \setminus \Sigma, N)$, and for some $0 < \alpha < 1$, $u \in C^{1,\alpha}_{\mathrm{loc}}\big((\Omega^+ \cup \Gamma) \setminus \Sigma, N\big) \cap C^{1,\alpha}_{\mathrm{loc}}\big((\Omega^- \cup \Gamma) \setminus \Sigma, N\big)$.

The jump condition is needed for both energy monotonicity inequalities for u and the piecewise $C^{1,\alpha}$-regularity of u.

We point out that in dimension $n = 2$, since the energy monotonicity inequality automatically holds for H^1-maps, Theorem 1.1 holds for any weakly harmonic map from domains of piecewise $C^{0,1}$-metrics, i.e., any weakly harmonic map on domains with piecewise Lipschitz continuous metrics satisfying (1-4) is both Lipschitz continuous and piecewise $C^{1,\alpha}$ for some $0 < \alpha < 1$.

Weakly harmonic maps from domains with smooth metrics into Riemannian manifolds may not enjoy partial regularity properties in dimensions $n \geq 3$; see [Rivière 1995]. Here we consider weakly harmonic maps on domains with piecewise Lipschitz continuous metrics into a Riemannian manifold (N, h), on which $d_N^2(\cdot, p)$ is convex for $p \in N$. Such Riemannian manifolds N include those with nonpositive sectional curvatures and geodesic convex balls in Riemannian manifolds. In particular, we extend the classical regularity theorems on harmonic maps on domains with smooth metrics, due to [Eells and Sampson 1964] and [Hildebrandt et al. 1977].

Theorem 1.2. *Let g satisfy the conditions of Theorem 1.1. Assume that on the universal cover $(\widetilde{N}, \tilde{h})$ of (N, h),[1] the square of distance function $d_{\widetilde{N}}^2(\cdot, p)$ is convex for any $p \in \widetilde{N}$. If $u \in H^1(\Omega, N)$ is a weakly harmonic map, then $u \in \mathrm{Lip}_{\mathrm{loc}}(\Omega, N)$, and for some $0 < \alpha < 1$, $u \in C^{1,\alpha}_{\mathrm{loc}}(\Omega^+ \cup \Gamma, N) \cap C^{1,\alpha}_{\mathrm{loc}}(\Omega^- \cup \Gamma, N)$.*

The idea for the proof of Theorem 1.1 is motivated in [Evans 1991] and [Bethuel 1993]. However, there are several new technical difficulties:

(i) Establishing an almost energy monotonicity inequality for stationary harmonic maps in (Ω, g). This is achieved by observing that an exact monotonicity inequality holds at any $x \in \Gamma$, see Section 2 below.

[1] Here the covering map $\Pi : \widetilde{N} \to N$ is a Riemannian submersion.

(ii) Establishing a Hodge decomposition in $L^p(B, \mathbb{R}^n)$, for any $1 < p < +\infty$, on a ball $B = B_r(0)$, equipped with certain piecewise continuous metrics g. More precisely, we need to show that the solution of

$$\begin{cases} \dfrac{\partial}{\partial x_i}\Big(a_{ij}\dfrac{\partial v}{\partial x_j}\Big) = \operatorname{div} f & \text{in } B, \\ v = 0 & \text{on } \partial B \end{cases}$$

enjoys a $W^{1,p}$-estimate: for any $1 < p < +\infty$,

$$\|\nabla v\|_{L^p(B)} \le C\|f\|_{L^p(B)}$$

provided that $(a_{ij}) \in C(\overline{B^\pm}) \cap C(B^\delta)$ for some $\delta > 0$, is uniformly elliptic, but is discontinuous on $\partial B^+ \setminus B^\delta$, where $B^\delta = \{x \in B : \operatorname{dist}(x, \partial B) \le \delta\}$. This follows from a recent theorem in [Byun and Wang 2010; Dong and Kim 2010]; see also [Dong and Kim 2011a; 2011b] and Section 3 below.

(iii) Employing the moving frame method to establish the decay estimate in suitable Morrey spaces under a smallness condition, analogous to [Ishizuka and Wang 2008]. To obtain Lipschitz and piecewise $C^{1,\alpha}$-regularity, we compare the harmonic map system with an elliptic system with piecewise constant coefficients and perform a hole-filling argument, similar to [Giaquinta and Hildebrandt 1982].

The paper is organized as follows. In Section 2, we derive an almost energy monotonicity inequality. In Section 3, we show the global $W^{1,p}$ $(1 < p < \infty)$ estimate for elliptic systems with certain piecewise continuous coefficients, and a Hodge decomposition theorem. In Section 4, we adapt the moving frame method of [Hélein 2002] and [Bethuel 1993] to establish an ϵ-Hölder continuity. In Section 5, we establish both Lipschitz and piecewise $C^{1,\alpha}$ regularity for Hölder continuous harmonic maps. In Section 6, we consider harmonic maps into manifolds supporting convex distance square functions and prove Theorem 1.2.

2. Energy monotonicity inequality

This section is devoted to the derivation of energy monotonicity inequalities for stationary harmonic maps from (Ω, g) to (N, h).

Theorem 2.1. *Under the same assumptions as in Theorem 1.1, there exist $C > 0$ and $r_0 > 0$, depending only on Ω, Γ, and g, such that if $u \in W^{1,2}(\Omega, N)$ is a stationary harmonic map on (Ω, g), then for any $x_0 \in \Omega$, there holds*

$$s^{2-n}\int_{B_s(x_0)} |\nabla u|_g^2\, dv_g \le e^{Cr} r^{2-n}\int_{B_r(x_0)} |\nabla u|_g^2\, dv_g \tag{2-1}$$

for all $0 < s \le r \le \min\{r_0, \operatorname{dist}(x_0, \partial\Omega)\}$.

Since $g \in C^{0,1}(\Omega^{\pm})$, there are $C > 0$ and $r_0 > 0$ such that (2-1) holds for any $x_0 \in \Omega^{\pm}$ and $0 < s \le r \le \min\{r_0, \operatorname{dist}(x_0, \partial\Omega^{\pm})\}$; see [Hélein 2002]. In particular, (2-1) holds for any $x_0 \in \Omega \setminus \Gamma^{r_0}$ and $0 < s \le r \le \min\{r_0, \operatorname{dist}(x_0, \partial\Omega)\}$, where $\Gamma^{r_0} = \{x \in \Omega : \operatorname{dist}(x, \Gamma) \le r_0\}$ is the r_0-neighborhood of Γ. To show (2-1) for $x_0 \in \Gamma^{r_0}$, it suffices to consider the case $x_0 \in \Gamma$.

It follows from the assumption on Γ and g that there exists $r_0 > 0$ such that for any $x_0 \in \Gamma$ there exists a $C^{1,1}$-diffeomorphism $\Phi_0 : B_1 \to B_{r_1}(x_0)$, where $r_1 = \min\{r_0, \operatorname{dist}(x_0, \partial\Omega)\}$, such that

$$\begin{cases} \Phi_0(B_1^{\pm}) = \Omega^{\pm} \cap B_{r_1}(x_0), \\ \Phi_0(\Gamma_1) = \Gamma \cap B_{r_1}(x_0), \ \text{ where } \Gamma_1 = \{x \in B_1 : x_n = 0\}. \end{cases}$$

Define $\tilde{u}(x) = u(\Phi_0(x))$ and $\tilde{g}(x) = \Phi_0^*(g)(x)$ for $x \in B_1$. Then it is readily seen that $\tilde{g}$ is piecewise $C^{0,1}$, with Γ as its discontinuity set, and satisfies (1-4) on Γ_1. (In fact, since

$$\Phi_0^*(g)_{ij}(x) = g_{kl}(\Phi_0(x)) \frac{\partial \Phi_0^k}{\partial x_i}(x) \frac{\partial \Phi_0^l}{\partial x_j}(x),$$

condition (1-4) implies that

$$\lim_{\substack{y \in \Omega^+ \\ y \to x}} \Phi_0^* g(y) = k(\Phi_0(x)) \lim_{\substack{y \in \Omega^- \\ y \to x}} \Phi_0^* g(y)$$

for any $x \in \Gamma_1$.) It is also easy to see that, if $u : (B_{r_1}(x_0), g) \to (N, h)$ is a stationary harmonic map, so $\tilde{u} : (B_1, \tilde{g}) \to (N, h)$.

Thus we may assume that $\Omega = B_1$, that g is a piecewise $C^{0,1}$-metric which satisfies (1-4) on the set of discontinuity Γ_1, and that $u : (B_1, g) \to (N, h)$ is a stationary harmonic map. It suffices to establish (2-1) in $B_{1/2}$. We first derive a stationarity identity for u.

Proposition 2.2. *Let $u \in W^{1,2}(B_1, N)$ be a stationary harmonic map on (B_1, g). Then*

$$\int_{B_1} \left(2g^{ij} \left\langle \frac{\partial u}{\partial x_k}, \frac{\partial u}{\partial x_j} \right\rangle Y_i^k - |\nabla u|_g^2 \operatorname{div} Y \right) \sqrt{g}\, dx = \int_{B_1} \frac{\partial}{\partial x_k} \big(\sqrt{g} g^{ij} \big) Y^k \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_j} \right\rangle dx \tag{2-2}$$

for all $Y = (Y^1, \dots, Y^{n-1}, Y^n) \in C_0^1(B_1, \mathbb{R}^n)$ satisfying

$$Y^n(x) \begin{cases} \ge 0 & \text{for } x^n > 0, \\ = 0 & \text{for } x^n = 0, \\ \le 0 & \text{for } x^n < 0, \end{cases} \tag{2-3}$$

where $Y_i^k = \partial Y^k / \partial x_i$ and $\operatorname{div} Y = \sum_{i=1}^n \partial Y^i / \partial x_i$.

Proof. Let $Y \in C_0^1(B_1, \mathbb{R}^n)$ satisfy (2-3). Then there exists $\delta > 0$ such that $F_t(x) = x + tY(x),\ t \in [-\delta, \delta]$, is a family of diffeomorphisms from B_1 to B_1 satisfying the condition (1-3). Hence

$$\begin{aligned}0 &= \frac{d}{dt}\Big|_{t=0} \int_{B_1} \left|\nabla u(F_t(x))\right|_g^2 \, dv_g \\ &= \frac{d}{dt}\Big|_{t=0} \left(\int_{B_1^+} \left|\nabla u(F_t(x))\right|_g^2 \, dv_g + \int_{B_1^-} \left|\nabla u(F_t(x))\right|_g^2 \, dv_g \right).\end{aligned}$$

Set $G_t = F_t^{-1}$, for $t \in [-\delta, \delta]$. Direct calculations yield

$$\begin{aligned}&\frac{d}{dt}\Big|_{t=0} \int_{B_1^\pm} \left|\nabla (u(F_t(x))\right|_g^2 \, dv_g \\ &= \frac{d}{dt}\Big|_{t=0} \int_{B_1^\pm} \sqrt{g(x+tY(x))}\, g^{ij}(x+tY(x)) \left\langle \frac{\partial u}{\partial x_k}, \frac{\partial u}{\partial x_l} \right\rangle \\ &\qquad\qquad\qquad \times (x+tY(x))(\delta_{ki} + tY_i^k)(\delta_{lj} + tY_j^l)\, dx \\ &= \int_{B_1^\pm} \sqrt{g(x)}\, g^{ij}(x) \left\langle \frac{\partial u}{\partial x_k}, \frac{\partial u}{\partial x_l} \right\rangle (\delta_{ki} Y_j^l + \delta_{lj} Y_i^k)\, dx \\ &\qquad + \int_{B_1^\pm} \frac{d}{dt}\Big|_{t=0} \left(g^{ij}(G_t(x)) \sqrt{g(G_t(x))}\, JG_t(x) \right) \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_j} \right\rangle dx \\ &= \int_{B_1^\pm} \left(2g^{ij} \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_l} \right\rangle Y_j^l - g^{ij} \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_j} \right\rangle \operatorname{div} Y \right) \sqrt{g}\, dx \\ &\qquad - \int_{B_1^\pm} \frac{\partial}{\partial x_k} \left(\sqrt{g}\, g^{ij} \right) Y^k \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_j} \right\rangle dx,\end{aligned}$$

where we have used the equalities

$$\begin{cases} \dfrac{d}{dt}\Big|_{t=0} JG_t(x) = -\operatorname{div} Y, \\ \dfrac{d}{dt}\Big|_{t=0} G_t(x) = -Y(x), \\ \dfrac{d}{dt}\Big|_{t=0} \left(g^{ij}(G_t(x)) \sqrt{g(G_t(x))} \right) = -\dfrac{\partial}{\partial x_k} \left(\sqrt{g}\, g^{ij} \right) Y^k. \end{cases}$$

This completes the proof. □

Proposition 2.3. *Let $u \in W^{1,2}(B_1, N)$ be a stationary harmonic map on (B_1, g). There exists $C > 0$ such that*:

(i) *For any $x_0 = (x_0', x_0^n) \in B_{1/2} \setminus \Gamma_1$, there exists $0 < R_0 \le \min\left\{\frac{1}{4}, |x_0^n|\right\}$ such that*

$$r^{2-n} \int_{B_r(x_0)} |\nabla u|_g^2 \, dv_g \le e^{CR} R^{2-n} \int_{B_R(x_0)} |\nabla u|_g^2 \, dv_g \quad \textit{if } 0 < r \le R < R_0. \tag{2-4}$$

(ii) *For any* $x_0 \in B_{1/2}$,

$$r^{2-n}\int_{B_r(x_0)} |\nabla u|_g^2\,dv_g \le e^{CR}R^{2-n}\int_{B_R(x_0)} |\nabla u|_g^2\,dv_g \quad \textit{if } 0<r\le R\le \tfrac14. \tag{2-5}$$

Proof. (i) By choosing $Y \in C_c^\infty(B_1^+,\mathbb{R}^n)$ or $Y \in C_c^\infty(B_1^-,\mathbb{R}^n)$, we conclude that u is a stationary harmonic map on (B_1^+, g) and (B_1^-, g). Hence the monotonicity inequality (2-4) holds; see [Hélein 2002].

(ii) Step 1. We first consider the case where $x_0 \in \Gamma_1$. Without loss of generality, we can assume that $x_0 = (0', 0)$. For $\epsilon > 0$ and $0 < r \le \frac12$, let $Y_\epsilon(x) = x\eta_\epsilon(x)$, where $\eta_\epsilon(x) = \eta_\epsilon(|x|) \in C_0^\infty(B_1)$ satisfies

$$0\le \eta_\epsilon\le 1,\quad \eta_\epsilon'\le 0,\quad |\eta_\epsilon'|\le \frac{2}{\epsilon},\quad \eta_\epsilon(s)=\begin{cases}1 & \text{for } 0\le s\le r-\epsilon,\\ 0 & \text{for } s\ge r.\end{cases} \tag{2-6}$$

Then

$$(Y_\epsilon)_i^j = \delta_{ij}\eta_\epsilon(|x|)+\eta_\epsilon'(|x|)\frac{x^i x^j}{|x|}. \tag{2-7}$$

Substituting Y_ϵ into the right side of (2-2), and using

$$\left|\frac{\partial}{\partial x_k}\big(\sqrt{g}g^{ij}\big)\right| \le C \quad \text{for a.e. } x\in B_1\setminus\Gamma_1,$$

we have

$$\begin{aligned}\left|\int_{B_1}\frac{\partial}{\partial x_k}\big(\sqrt{g}g^{ij}\big)Y_\epsilon^k\left\langle\frac{\partial u}{\partial x_i},\frac{\partial u}{\partial x_j}\right\rangle dx\right| &\le Cr\int_{B_r}|\nabla u|^2\,dx\\ &\le Cr\int_{B_r}|\nabla u|_g^2\,dv_g.\end{aligned} \tag{2-8}$$

Substituting (2-7) into the left side of (2-2), we obtain

$$\begin{aligned}\int_{B_1}&\left(2g^{ij}\left\langle\frac{\partial u}{\partial x_j},\frac{\partial u}{\partial x_k}\right\rangle (Y_\epsilon)_i^k-|\nabla u|_g^2\operatorname{div}Y_\epsilon\right)\sqrt{g}\,dx\\ &=(2-n)\int_{B_1}|\nabla u|_g^2\eta_\epsilon(|x|)\sqrt{g}\,dx-\int_{B_1}|\nabla u|_g^2|x|\eta_\epsilon'(|x|)\sqrt{g}\,dx\\ &\qquad+\int_{B_1}2g^{ij}\left\langle\frac{\partial u}{\partial x_i},\frac{\partial u}{\partial x_k}\right\rangle\frac{x^k x^j}{|x|}\eta_\epsilon'(|x|)\sqrt{g}\,dx.\end{aligned} \tag{2-9}$$

Define $\bar g$ by

$$\bar g(x',x^n)=\begin{cases}\lim_{y\to 0,\ y^n\ge 0} g(y) & \text{if } x^n\ge 0,\\ \lim_{y\to 0,\ y^n<0} g(y) & \text{if } x^n<0.\end{cases}$$

Then we have

$$|g(x)-\bar g(x)|\le C|x| \quad \text{for all } x\in B_1. \tag{2-10}$$

Further, by (1-4) we can assume

$$\bar{g}(x) = \begin{cases} g_0 & \text{if } x^n \geq 0, \\ k g_0 & \text{if } x^n < 0 \quad (k \neq 1). \end{cases}$$

Hence we can write

$$\int_{B_1} 2g^{ij} \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_k} \right\rangle \frac{x^k x^j}{|x|} \eta'_\epsilon(|x|) \sqrt{g}\, dx = I_\epsilon + II_\epsilon. \tag{2-11}$$

where

$$I_\epsilon = 2 \int_{B_1} \bar{g}^{ij} \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_k} \right\rangle \frac{x^k x^j}{|x|} \eta'_\epsilon(|x|) \sqrt{g}\, dx,$$
$$II_\epsilon = 2 \int_{B_1} (g^{ij} - \bar{g}^{ij}) \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_k} \right\rangle \frac{x^k x^j}{|x|} \eta'_\epsilon(|x|) \sqrt{g}\, dx.$$

Since

$$\bar{g}^{ij} \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial u}{\partial x_k} \right\rangle \frac{x^k x^j}{|x|} = \begin{cases} |x| |\partial u / \partial r|^2 & \text{if } x^n \geq 0, \\ (1/k)|x| |\partial u / \partial r|^2 & \text{if } x^n < 0 \end{cases}$$

is nonnegative in B_1 and $\eta'_\epsilon(|x|) \leq 0$, we have $I_\epsilon \leq 0$. For II_ϵ, by (2-10) we have

$$|II_\epsilon| \leq C r^2 \int_{B_r} |\nabla u|_g^2 \, |\eta'_\epsilon|(|x|)\, dv_g. \tag{2-12}$$

Putting these estimates first into (2-11) and then into (2-9), and finally combining (2-9) and (2-8) with (2-2), we obtain, after taking ϵ to zero,

$$\begin{aligned}(2-n) \int_{B_r} |\nabla u|_g^2 \, dv_g + r \int_{\partial B_r} |\nabla u|_g^2 \sqrt{g}\, dH^{n-1} \\ \geq -C \left(r \int_{B_r} |\nabla u|_g^2 \, dv_g + r^2 \int_{\partial B_r} |\nabla u|_g^2 \sqrt{g} dH^{n-1} \right).\end{aligned} \tag{2-13}$$

It is not hard to see that (2-13) implies

$$\frac{d}{dr} \left(e^{Cr} r^{2-n} \int_{B_r} |\nabla u|_g^2 \, dv_g \right) \geq 0,$$

so that (2-5) holds when $x_0 \in B_{1/2}$.

Step 2. To show (2-5) in the general case, it suffices to consider $x_0 \in B_{1/2} \setminus \Gamma_1$ such that

$$|B_R(x_0) \cap B_1^+| > 0 \quad \text{and} \quad |B_R(x_0) \cap B_1^-| > 0.$$

For simplicity, assume $x_0 \in B_1^-$. We consider two cases:

Suppose $d(x_0, \Gamma_1) = |x_0^n| \geq \frac{1}{4} R$. Then:

- If $R \geq r \geq \frac{1}{4}R$, it is easy to see that

$$r^{2-n} \int_{B_r(x_0)} |\nabla u|_g^2 \, dv_g \leq 4^{n-2} R^{2-n} \int_{B_R(x_0)} |\nabla u|_g^2 \, dv_g.$$

- If $0 < r < \frac{1}{4}R$ $(\leq d(x_0, \Gamma_1))$, we have $B_{R/4}(x_0) \subset B_1^-$, so (2-4) implies

$$\begin{aligned} r^{2-n} \int_{B_r(x_0)} |\nabla u|_g^2 \, dv_g &\leq e^{CR} \left(\frac{R}{4}\right)^{2-n} \int_{B_{R/4}(x_0)} |\nabla u|_g^2 \, dv_g \\ &\leq e^{CR} R^{2-n} \int_{B_R(x_0)} |\nabla u|_g^2 \, dv_g. \end{aligned}$$

Suppose instead that $d(x_0, \Gamma_1) = |x_0^n| < \frac{1}{4}R$. Then:

- If $R \geq r \geq \frac{1}{4}R$, then

$$r^{2-n} \int_{B_r(x_0)} |\nabla u|_g^2 \, dv_g \leq 4^{n-2} R^{2-n} \int_{B_R(x_0)} |\nabla u|_g^2 \, dv_g.$$

- If $0 < r \leq d(x_0, \Gamma_1) = |x_0^n| < \frac{1}{4}R$, then by setting $\bar{x}_0 = (x_0^1, \ldots, x_0^{n-1}, 0)$ we have

$$B_r(x_0) \subset B_{|x_0^n|}(x_0) \subset B_{2|x_0^n|}(\bar{x}_0) \subset B_{R/2}(\bar{x}_0) \subset B_R(x_0),$$

so that (2-5) yields

$$\begin{aligned} r^{2-n} \int_{B_r(x_0)} |\nabla u|_g^2 \, dv_g &\leq |x_0^n|^{2-n} \int_{B_{|x_0^n|}(x_0)} |\nabla u|_g^2 \, dv_g \\ &\leq 2^{n-2} (2|x_0^n|)^{2-n} \int_{B_{2|x_0^n|}(\bar{x}_0)} |\nabla u|_g^2 \, dv_g \\ &\leq 2^{n-2} e^{CR} \left(\frac{R}{2}\right)^{2-n} \int_{B_{R/2}(\bar{x}_0)} |\nabla u|_g^2 \, dv_g \\ &\leq e^{CR} R^{2-n} \int_{B_R(x_0)} |\nabla u|_g^2 \, dv_g. \end{aligned}$$

- If $d(x_0, \Gamma_1)$ $(= |x_0^n|) \leq r < \frac{1}{4}R$, then we have

$$B_r(x_0) \subset B_{2r}(\bar{x}_0) \subset B_{R/2}(\bar{x}_0) \subset B_R(x_0),$$

so that (2-5) yields

$$r^{2-n}\int_{B_r(x_0)}|\nabla u|_g^2\,dv_g \le 2^{n-2}(2r)^{2-n}\int_{B_{2r}(\bar{x}_0)}|\nabla u|_g^2\,dv_g$$
$$\le 2^{n-2}e^{CR}\left(\frac{R}{2}\right)^{2-n}\int_{B_{R/2}(\bar{x}_0)}|\nabla u|_g^2\,dv_g$$
$$\le e^{CR}R^{2-n}\int_{B_R(x_0)}|\nabla u|_g^2\,dv_g.$$

Therefore (2-5) is proved in all cases. □

3. $W^{1,p}$-estimate for elliptic equations with piecewise continuous coefficients

In this section, we will provide the global $W^{1,p}$-estimate for elliptic equations with piecewise continuous coefficients. The proof is a slight modification of that of [Dong and Kim 2010] (see also [Dong and Kim 2011a; 2011b]) or [Byun and Wang 2010]. As a corollary, we will establish the Hodge decomposition theorem (Theorem 3.2) for piecewise continuous metrics g, a crucial ingredient to prove Theorem 1.1.

For a ball $B = B_r(0) \subset \mathbb{R}^n$, set $B^\epsilon = \{x \in B : \mathrm{dist}(x, \partial B) \le \epsilon\}$ for $\epsilon > 0$. Let $(a_{ij}(x))_{1\le i,j\le n}$ be bounded measurable, uniformly elliptic on B; i.e., there exist $0 < \lambda \le \Lambda < +\infty$ such that

$$\lambda|\xi|^2 \le a_{ij}(x)\xi_i^\alpha\xi_\beta^j \le \Lambda|\xi|^2 \ \text{ a.e. } x \in B \quad \text{for all } \xi \in \mathbb{R}^n. \tag{3-1}$$

Theorem 3.1. *Assume (a_{ij}) satisfies (3-1), and there exists $\epsilon > 0$ such that $(a_{ij}) \in C\left(\overline{B^\pm}\right) \cap C\left(B^\epsilon\right)$ and is discontinuous on $\partial B^+ \setminus B^\epsilon$. For $p \in (1, +\infty)$, let $f \in L^p(B, \mathbb{R}^n)$. Then there exists a unique weak solution $v \in W_0^{1,p}(B, \mathbb{R}^n)$ to*

$$\begin{cases} \displaystyle\sum_{i,j}\frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial v}{\partial x_j}\right) = \sum_i \frac{\partial f_i}{\partial x_i} & \text{in } B,\\ u = 0 & \text{on } \partial B, \end{cases} \tag{3-2}$$

that satisfies

$$\|\nabla v\|_{L^p(B)} \le C\,\|f\|_{L^p(B)} \tag{3-3}$$

for some $C > 0$ depending only on p and (a_{ij}).

Proof. By (3-1), we see that for any $\delta > 0$, there exists $R = R(\delta) > 0$ such that the coefficient function (a_{ij}) satisfies the (δ, R)-vanishing of codimension-one conditions (2.5) and (2.6) of [Byun and Wang 2010, p. 2562]; see also [Dong and Kim 2010; 2011a; 2011b]. In fact, we have

$$\lim_{r\downarrow 0}\ \max_{x_0=(x_0',x_0^n)\in\overline{B}}\left\|a_{ij}(x',x^n) - a_{ij}(x_0',x^n)\right\|_{L^\infty(B_r((x_0',x_0^n)))} = 0.$$

Therefore Theorem 3.1 follows directly from [Byun and Wang 2010, Theorem 2.2, p. 2653]. □

As an immediate consequence of Theorem 3.1, we have the following Hodge decomposition on B equipped with certain piecewise continuous metrics g.

Theorem 3.2. *Let $\bar{g}$ be a piecewise continuous metric on B such that $\bar{g}$ is continuous on $B^{\pm}$ and on B^{δ} for some $\delta > 0$, and is discontinuous on $\partial B^{+} \setminus B^{\delta}$. Then for any $p \in (1, +\infty)$ and $F = (F_1, \ldots, F_n) \in L^p(B, \mathbb{R}^n)$, there exist $G \in W_0^{1,p}(B)$ and $H \in L^p(B, \mathbb{R}^n)$ such that*

$$F = \nabla G + H, \quad \operatorname{div}_{\bar{g}} H \left(:= \frac{1}{\sqrt{\bar{g}}} \frac{\partial}{\partial x_i} (\sqrt{\bar{g}} \bar{g}^{ij} H_j) \right) = 0 \quad \text{in } B. \tag{3-4}$$

Further, there exists $C = C(p, n, \bar{g}) > 0$ such that

$$\|\nabla G\|_{L^p(B)} + \|H\|_{L^p(B)} \le C \, \|F\|_{L^p(B)}. \tag{3-5}$$

Proof. For $1 \le i, j \le n$, set $a_{ij} = \sqrt{\bar{g}} \bar{g}^{ij}$ on B. Then (a_{ij}) satisfies the conditions of Theorem 3.1, so that there exists a unique solution $G \in W_0^{1,p}(B)$ to

$$\begin{cases} \dfrac{\partial}{\partial x_i} \left(\sqrt{\bar{g}} \bar{g}^{ij} \dfrac{\partial G}{\partial x_j} \right) = \dfrac{\partial}{\partial x_i} \left(\sqrt{\bar{g}} \bar{g}^{ij} F_j \right) & \text{in } B, \\ G = 0 & \text{on } \partial B, \end{cases} \tag{3-6}$$

and

$$\|\nabla G\|_{L^p(B)} \le C \left\| \sqrt{\bar{g}} \bar{g}^{ij} F_j \right\|_{L^p(B)} \le C \, \|F\|_{L^p(B)}.$$

Set $H = F - \nabla G$. Then we have

$$\operatorname{div}_{\bar{g}} H = \frac{1}{\sqrt{\bar{g}}} \frac{\partial}{\partial x_i} \left(\sqrt{\bar{g}} \bar{g}^{ij} \left(F_j - \frac{\partial G}{\partial x_j} \right) \right) = 0 \quad \text{on } B,$$

and

$$\|H\|_{L^p(B_{1/2})} \le \|F\|_{L^p(B_{1/2})} + \|\nabla G\|_{L^p(B)} \le C \, \|F\|_{L^p(B)}.$$

This completes the proof. □

4. Hölder continuity

In this section, we will prove that any stationary harmonic map on (B_1, g), with $g \in C^{0,1}(B_1^{\pm} \cup \Gamma_1)$, is Hölder continuous provided that $\int_{B_1} |\nabla u|_g^2 \, dv_g$ is sufficiently small. The idea is based on suitable modifications of the original argument in [Bethuel 1993] (see also [Ishizuka and Wang 2008]), thanks to both the energy monotonicity inequality and the Hodge decomposition theorem established in the previous two sections. More precisely:

Theorem 4.1. *There exist $\epsilon_0 > 0$ and $\alpha_0 \in (0, 1)$, depending only on n, g, such that if the metric $g \in C^{0,1}(B_1^{\pm} \cup \Gamma_1)$ satisfies the condition (1-4) on Γ_1, and $u \in W^{1,2}(B_1, N)$ is a stationary harmonic map satisfying*

$$r_0^{2-n} \int_{B_{r_0}(x_0)} |\nabla u|_g^2 \, dv_g \le \epsilon_0^2 \tag{4-1}$$

for some $x_0 \in B_{1/2}$ and $0 < r_0 \le \frac{1}{4}$, then $u \in C^{\alpha_0}(B_{r_0/2}(x_0), N)$, and

$$[u]_{C^{\alpha_0}(B_{r_0/2}(x_0))} \le C(r_0, \epsilon_0). \tag{4-2}$$

Proof of Theorem 4.1. The proof is based on suitable modifications of [Bethuel 1993; Ishizuka and Wang 2008]. First, observe that if $x_0 = (x_0', x_0^n) \in B^{\pm}$, it follows from the monotonicity inequality (2-5) that we may assume (4-1) holds for some $0 < r_0 < |x_0^n|$. Then the ϵ_0-regularity theorem in [Bethuel 1993] (see [Ishizuka and Wang 2008] for domains with $C^{0,1}$metrics) implies that for some $0 < \alpha_0 < 1$, $u \in C^{\alpha_0}(B_{r_0/2}(x_0))$ and (4-2) holds. Hence it suffices to consider the case $x_0 = (x_0', 0) \in \Gamma_{1/2}$. By translation and scaling, we may assume $x_0 = (0, 0)$ and proceed as follows.

Step 1. As in [Bethuel 1993; Hélein 2002; Ishizuka and Wang 2008], we assume that there exists an orthonormal frame on $u^*TN|_{B_1}$. For $0 < \theta < \frac{1}{2}$, to be determined later, let $\{e_\alpha\}_{\alpha=1}^l \subset W^{1,2}(B_{2\theta}, \mathbb{R}^k)$ be a Coulomb gauge orthonormal frame of $u^*TN|_{B_{2\theta}}$; that is,

$$\begin{aligned} &\operatorname{div}_g(\langle \nabla e_\alpha, e_\beta \rangle) = 0 \quad \text{in } B_{2\theta} \quad (1 \le \alpha, \beta \le l), \\ &\sum_{\alpha=1}^{l} \int_{B_{2\theta}} |\nabla e_\alpha|_g^2 \, dv_g \le C \int_{B_{2\theta}} |\nabla u|_g^2 \, dv_g. \end{aligned} \tag{4-3}$$

For $1 \le \alpha \le l$, consider $\langle \nabla\left((u - u_{2\theta})\eta\right), e_\alpha \rangle$, where $u_{2\theta} = ⨍_{B_{2\theta}} u$ is the average of u on $B_{2\theta}$, and $\eta \in C_0^\infty(B_1)$ satisfies

$$0 \le \eta \le 1; \quad \eta = 1 \text{ in } B_\theta; \quad \eta = 0 \text{ outside } B_{7\theta/4}; \quad |\nabla \eta| \le \frac{2}{\theta}.$$

Define the metric $\tilde{g}$ on $B_{2\theta}$ by

$$\tilde{g}(x) = \eta(x) g(x) + (1 - \eta(x)) g_0(x), \quad x \in B_{2\theta}.$$

Then it is easy to see that

$$\tilde{g} \equiv g \text{ on } B_\theta; \quad \tilde{g} \equiv g_0 \text{ outside } B_{7\theta/4}; \quad \tilde{g} \in C(\overline{B_{2\theta}^{\pm}}) \cap C(B_{2\theta} \setminus B_{7\theta/4}).$$

In particular, $\tilde{g}$ satisfies the conditions of Theorem 3.2. Hence, by Theorem 3.2, for

$1 < p < n/(n-1)$, there exist $\phi_\alpha \in W_0^{1,p}(B_{2\theta})$ and $\psi_\alpha \in L^p(B_{2\theta})$ such that

$$\begin{aligned} &\langle \nabla((u-u_{2\theta})\eta), e_\alpha \rangle = \nabla\phi_\alpha + \psi_\alpha, \quad \operatorname{div}_{\tilde{g}}(\psi_\alpha) = 0 \quad \text{in } B_{2\theta}, \\ &\|\nabla\phi_\alpha\|_{L^p(B_{2\theta})} + \|\psi_\alpha\|_{L^p(B_{2\theta})} \lesssim \|\nabla((u-u_{2\theta})\eta)\|_{L^p(B_{2\theta})} \lesssim \|\nabla u\|_{L^p(B_{2\theta})}. \end{aligned} \tag{4-4}$$

Since u satisfies the harmonic map equation (1-1), we have

$$\operatorname{div}_g(\langle \nabla u, e_\alpha \rangle) = g^{ij}\nabla_i u \langle \nabla_j e_\alpha, e_\beta \rangle\, e_\beta \quad \text{in } B_1. \tag{4-5}$$

Thus we obtain

$$\Delta_g \phi_\alpha = g^{ij}\nabla_i u \langle \nabla_j e_\alpha, e_\beta \rangle\, e_\beta \quad \text{in } B_\theta. \tag{4-6}$$

Decompose $\phi_\alpha = \phi_\alpha^{(1)} + \phi_\alpha^{(2)}$, where $\phi_\alpha^{(1)}$ solves

$$\begin{cases} \Delta_g \phi_\alpha^{(1)} = 0 & \text{in } B_\theta, \\ \phi_\alpha^{(1)} = \phi_\alpha & \text{on } \partial B_\theta, \end{cases} \tag{4-7}$$

and $\phi_\alpha^{(2)}$ solves

$$\begin{cases} \Delta_g \phi_\alpha^{(2)} = g^{ij}\nabla_i u \langle \nabla_j e_\alpha, e_\beta \rangle\, e_\beta & \text{in } B_\theta, \\ \phi_\alpha^{(2)} = 0 & \text{on } \partial B_\theta. \end{cases} \tag{4-8}$$

<u>Step 2</u>: Estimation of $\phi_\alpha^{(1)}$. We will need the Morrey space defined, for arbitrary $E \subset \mathbb{R}^n$, by

$$M^{p,p}(E) := \Big\{ f : E \to \mathbb{R} \,\Big|\, \|f\|^p_{M^{p,p}(E)} := \sup_{B_r(x)\subset\mathbb{R}^n} \Big\{ r^{p-n} \int_{B_r(x)\cap E} |f|^p\, dx \Big\} < +\infty \Big\}.$$

It is well-known (see [Gilbarg and Trudinger 1983]) that $\phi_\alpha^{(1)} \in C^{\alpha_0}(B_\theta)$ for some $\alpha_0 \in (0,\ 1)$, and for any $0 < r \le \theta/2$,

$$\big[\phi_\alpha^{(1)}\big]^p_{C^{\alpha_0}(B_{r/2})} \lesssim \theta^{p-n} \int_{B_\theta} |\nabla\phi_\alpha^{(1)}|^p\, dx \le C\theta^{p-n} \int_{B_{2\theta}} |\nabla u|^p\, dx, \tag{4-9}$$

and

$$(\tau\theta)^{p-n} \int_{B_{\tau\theta}} |\nabla\phi_\alpha^{(1)}|^p \le C\tau^{p\alpha_0} \|\nabla u\|_{M^{p,p}(B_1)} \quad \text{for all } 0 < \tau < 1, \tag{4-10}$$

<u>Step 3</u>: Estimation of $\phi_\alpha^{(2)}$. Denote by $\mathscr{H}^1(\mathbb{R}^n)$ the Hardy space on $\mathbb{R}^n$ and $BMO(E)$ the BMO space on E for any open set $E \subset \mathbb{R}^n$. By (4.13) of [Ishizuka and Wang 2008, p. 435], for $p' = p/(p-1) > n$, there exists $h \in W_0^{1,p'}(B_\theta)$, with $\|\nabla h\|_{L^{p'}(B_\theta)} = 1$, such that

$$\|\nabla\phi_\alpha^{(2)}\|_{L^p(B_\theta)} \le C \int_{B_\theta} \langle \nabla\phi_\alpha^{(2)}, \nabla h \rangle_g\, dv_g.$$

Using (4-8), (4-4), and the duality between $\mathscr{H}^1$ and BMO, we show that

$$
\begin{aligned}
\text{(4-11)} \qquad \|\nabla\phi_\alpha^{(2)}\|_{L^p(B_\theta)} &\le C\int_{B_\theta}\sqrt{g}g^{ij}\big\langle\nabla_i u, \langle\nabla_j e_\alpha, e_\beta\rangle\big\rangle(e_\beta h)\,dx\\
&= -C\int_{B_\theta}\sqrt{g}g^{ij}\langle\nabla_j e_\alpha, e_\beta\rangle\nabla_i(e_\beta h)u\,dx\\
&\le C\big\|\sqrt{g}g^{ij}\langle\nabla_j e_\alpha, e_\beta\rangle\nabla_i(e_\beta h)\big\|_{\mathscr{H}^1(\mathbb{R}^n)}[u]_{BMO(B_\theta)}\\
&\lesssim \|\sqrt{g}g^{ij}\langle\nabla_j e_\alpha, e_\beta\rangle\|_{L^2(B_\theta)}\|\nabla(e_\beta h)\|_{L^2(B_\theta)}[u]_{BMO(B_\theta)}\\
&\lesssim \|\nabla u\|_{L^2(B_{2\theta})}\|\nabla u\|_{M^{p,p}(B_1)}\cdot\theta^{n/p-n/2}.
\end{aligned}
$$

(Here, to go from the third line to the fourth, we used that $h \in W_0^{1,p'}(B_\theta)$ and that $\operatorname{div}_g\langle\nabla e_\alpha, e_\beta\rangle$ vanishes in B_θ, so $\sqrt{g}g^{ij}\langle\nabla_j e_\alpha, e_\beta\rangle\nabla_i(e_\beta h) \in \mathscr{H}^1(\mathbb{R}^n)$ and

$$\big\|\sqrt{g}g^{ij}\langle\nabla_j e_\alpha, e_\beta\rangle\nabla_i(e_\beta h)\big\|_{\mathscr{H}^1(\mathbb{R}^n)} \le C\,\big\|\sqrt{g}g^{ij}\langle\nabla_j e_\alpha, e_\beta\rangle\big\|_{L^2(B_\theta)}\,\|\nabla_i(e_\beta h)\|_{L^2(B_\theta)}.$$

This last factor satisfies

$$\|\nabla(e_\beta h)\|_{L^2(B_\theta)} \le \|\nabla e_\beta\|_{L^2(B_\theta)}\|h\|_{L^\infty(B_\theta)} + \|\nabla h\|_{L^p(B_\theta)}\theta^{n/p-n/2} \le C\theta^{n/p-n/2},$$

since the Sobolev embedding implies (because $p' > n$) that $h \in C^{1-n/p'}(B_\theta)$ and

$$\|h\|_{L^\infty(B_\theta)} \le C\theta^{1-n/p'}.$$

Finally, the estimate $[u]_{BMO(B_\theta)} \le C\|\nabla u\|_{M^{p,p}(B_1)}$ is a consequence of the Poincaré inequality.)

Putting the estimates of $\phi_\alpha^{(1)}$ and $\phi_\alpha^{(2)}$ together, we obtain that, for all $0 < \tau < 1$,

$$\text{(4-12)} \qquad \left((\tau\theta)^{p-n}\int_{B_{\tau\theta}}|\nabla\phi_\alpha|^p dx\right)^{1/p} \le C\big(\tau^{\alpha_0} + \tau^{1-n/p}\epsilon_0\big)\|\nabla u\|_{M^{p,p}(B_1)}.$$

Step 4: Estimation of ψ_α. Since $\operatorname{div}_{\tilde g}(\psi_\alpha) = 0$ on $B_{2\theta}$, we have

$$
\begin{aligned}
\int_{B_{2\theta}}|\psi_\alpha|_{\tilde g}^2\,dv_{\tilde g} &= \int_{B_{2\theta}}\langle\psi_\alpha+\nabla\phi_\alpha, \psi_\alpha\rangle_{\tilde g}\,dv_{\tilde g}\\
&= \int_{B_{2\theta}}\big\langle\langle\nabla((u-u_{2\theta})\eta), e_\alpha\rangle, \psi_\alpha\big\rangle_{\tilde g}\,dv_{\tilde g}\\
&= -\int_{B_{2\theta}}(u-u_{2\theta})\eta\langle\nabla e_\alpha, \psi_\alpha\rangle_{\tilde g}\,dv_{\tilde g}\\
&\lesssim \big\|\sqrt{\tilde g}\,\tilde g^{ij}\nabla_i e_\alpha\psi_\alpha^j\big\|_{\mathscr{H}^1}\,[(u-u_{2\theta})\eta]_{BMO}\\
&\lesssim \|\psi_\alpha\|_{L^2(B_{2\theta})}\|\nabla e_\alpha\|_{L^2(B_{2\theta})}\,[(u-u_{2\theta})\eta]_{BMO}\\
&\lesssim \|\nabla u\|_{L^2(B_{2\theta})}\|\psi_\alpha\|_{L^2(B_{2\theta})}\|\nabla u\|_{M^{p,p}(B_1)},
\end{aligned}
$$

where we have used the inequality

$$[(u-u_{2\theta})\eta]_{BMO} \le C\,[u]_{BMO(B_{2\theta})} \le C\,\|\nabla u\|_{M^{p,p}(B_1)}\,.$$

This, combined with Hölder's inequality, implies

$$\left(\theta^{p-n}\int_{B_\theta}|\psi_\alpha|^p\right)^{1/p} \le\ C\epsilon_0\,\|\nabla u\|_{M^{p,p}(B_1)}\,. \tag{4-13}$$

<u>Step 5</u>: Decay estimation of ∇u. Putting (4-12) and (4-13) together, we have that, for some $0<\alpha_0<1$,

$$\left((\tau\theta)^{p-n}\int_{B_{\tau\theta}}|\nabla u|^p\right)^{1/p} \le C(\epsilon_0+\tau^{\alpha_0}+\tau^{1-n/p}\epsilon_0)\|\nabla u\|_{M^{p,p}(B_1)} \tag{4-14}$$

for any $0<\tau<1$ and $0<\theta<\frac{1}{2}$. Now we claim that for some $\alpha_0\in(0,\ 1)$, we have

$$\|\nabla u\|_{M^{p,p}(B_{\tau/4})} \le C(\epsilon_0+\tau^{\alpha_0}+\tau^{1-n/p}\epsilon_0)\,\|\nabla u\|_{M^{p,n-p}(B_1)} \tag{4-15}$$

for all $0<\tau<1$. To show this, let $B_s(y)\subset B_{\tau/4}$. We divide into three cases:

(a) $y\in B_{\tau/4}\cap B^{\pm}$ and $s<|y^n|$. As remarked at the beginning of the proof, for some $0<\alpha_0<1$ we have

$$\begin{aligned}
\left(s^{p-n}\int_{B_s(y)}|\nabla u|^p\right)^{1/p} &\le C\left(\frac{s}{|y^n|}\right)^{\alpha_0}\left(|y^n|^{p-n}\int_{B_{|y^n|}(y)}|\nabla u|^p\right)^{1/p}\\
&\le C\left(\frac{s}{|y^n|}\right)^{\alpha_0}\left((2|y^n|)^{p-n}\int_{B_{2|y^n|}(y',0)}|\nabla u|^p\right)^{1/p}\\
&\le C\left(\left(\frac{\tau}{2}\right)^{p-n}\int_{B_{\tau/2}(y',0)}|\nabla u|^p\right)^{1/p}\ \text{(since } |y^n|\le\tau/4)\\
&\le C(\epsilon_0+\tau^{\alpha_0}+\tau^{1-n/p}\epsilon_0)\|\nabla u\|_{M^{p,p}(B_1)}\ \text{(by (4-14))}.
\end{aligned}$$

(b) $y\in B_{\tau/4}\cap B^{\pm}$ and $s\ge|y^n|$. Then $B_s(y)\subset B_{|y^n|+s}(y',0)\subset B_{2s}(y',0)$. Hence

$$\begin{aligned}
\left(s^{p-n}\int_{B_s(y)}|\nabla u|^p\right)^{1/p} &\le 2^{n/p-1}\left((2s)^{p-n}\int_{B_{2s}(y',0)}|\nabla u|^p\right)^{1/p}\\
&\le C(\epsilon_0+\tau^{\alpha_0}+\tau^{1-n/p}\epsilon_0)\,\|\nabla u\|_{M^{p,p}(B_1)}\ \ \text{(by (4-14))}.
\end{aligned}$$

(c) $y\in B_{\tau/4}\cap\Gamma_1$, i.e., $y^n=0$. Then it follows directly from (4-14) that

$$\left(s^{p-n}\int_{B_s(y)}|\nabla u|^p\right)^{1/p} \le C(\epsilon_0+\tau^{\alpha_0}+\tau^{1-n/p}\epsilon_0)\,\|\nabla u\|_{M^{p,p}(B_1)}\,.$$

Combining (a), (b) and (c) together and taking the supremum over all $B_s(y)\subset B_{\tau/4}$, we obtain (4-15).

It is clear that by first choosing τ and then ϵ sufficiently small, we can arrange that

$$\|\nabla u\|_{M^{p,p}(B_{\tau/4})} \le \tfrac{1}{2}\, \|\nabla u\|_{M^{p,p}(B_1)}\, .$$

Iterating this inequality finitely many times yields that there exists $\alpha_1 \in (0,\ 1)$ such that for any $x \in B_{1/4}$ and $0 < r \le \frac{1}{2}$, it holds

$$r^{p-n} \int_{B_r(x)} |\nabla u|^p\ dx \le C\, r^{p\alpha_1}\, \|\nabla u\|^p_{M^{p,p}(B_1)}\, .$$

This implies $u \in C^{\alpha_1}(B_{1/2})$ by Morrey's lemma. The proof is now completed. □

5. Lipschitz and piecewise $C^{1,\alpha}$-regularity

In this section, we will first establish Lipschitz and piecewise $C^{1,\alpha}$-regularity for stationary harmonic maps on domains with piecewise $C^{0,\,1}$-metrics, under a smallness condition of energy. Then we will prove Theorem 1.1.

Theorem 5.1. *There exist $\epsilon_0 > 0$ and $\beta_0 \in (0,\ 1)$, depending only on n and g, such that if the metric $g \in C^{0,1}(B_1^{\pm} \cup \Gamma_1)$ satisfies the condition (1-4) on Γ_1, and $u \in W^{1,2}(B_1, N)$ is a stationary harmonic map on (B_1, g) satisfying*

$$r_0^{2-n} \int_{B_{r_0}(x_0)} |\nabla u|_g^2\, dv_g \le \epsilon_0^2 \tag{5-1}$$

for some $x_0 \in B_{1/2}$ and $0 < r_0 \le \frac{1}{4}$, then $u \in C^{1,\beta_0}\big(B_{r_0/2}(x_0) \cap \overline{B^{\pm}}, N\big)$, and $u \in C^{0,1}\big(B_{r_0/2}(x_0), N\big)$.

Proof. The proof is based on the hole filling argument and the freezing coefficient method. It is divided into two steps.

Step 1: $u \in C^{\alpha}(B_{3r_0/4}(x_0), N)$ for any $0 < \alpha < 1$. To see this, first recall Theorem 4.1 implies that there exists $0 < \alpha_0 < \frac{2}{3}$ such that $u \in C^{\alpha_0}(B_{7r_0/8}(x_0))$ and for any $y \in B_{7r_0/8}(x_0)$, it holds

$$s^{2-n} \int_{B_s(y)} |\nabla u|^2\, dx \le C \left(\frac{s}{r}\right)^{2\alpha_0} r^{2-n} \int_{B_r(y)} |\nabla u|^2\, dx, \quad 0 < s \le r < \frac{r_0}{8}, \tag{5-2}$$

and

$$\mathrm{osc}_{B_r(y)} u \le Cr^{\alpha_0}, \quad 0 < r < \frac{r_0}{8}. \tag{5-3}$$

For $y \in B_{7r_0/8}(x_0)$ and $0 < r < r_0/8$, let $v : B_r(y) \to \mathbb{R}^k$ solve

$$\begin{cases} \Delta_g v = 0 & \text{in } B_r(y), \\ v = u & \text{on } \partial B_r(y). \end{cases} \tag{5-4}$$

By the maximum principle and (5-3), we then have

$$\mathrm{osc}_{B_r(y)} v \le \mathrm{osc}_{\partial B_r(y)} u \le C r^{\alpha_0}.$$

Moreover, since $g \in C^{0,1}(B_1^\pm \cup \Gamma_1)$, it follows from [Li and Nirenberg 2003, Theorem 1.1] that $v \in C^{0,1}\big(B_{r/2}(y), \mathbb{R}^k\big)$ and $v \in C^{1,\beta}\big(B_{r/2}(y) \cap \overline{B^\pm}, \mathbb{R}^k\big)$ for any $0 < \beta < 1$.

Multiplying (1-1) and (5-4) by $u - v$, subtracting one result from the other and integrating over $B_r(y)$, we obtain

$$\int_{B_r(y)} |\nabla(u-v)|^2\, dx \lesssim \int_{B_r(y)} |\nabla u|^2 |u-v| \lesssim r^{n-2+3\alpha_0}.$$

This, combined with

$$\int_{B_{r/2}(y)} |\nabla v|^2\, dx \le C \|\nabla v\|^2_{L^\infty(B_{r/2}(y))} r^n,$$

implies

$$\Big(\frac{r}{2}\Big)^{2-n} \int_{B_{r/2}(y)} |\nabla u|^2\, dx \le C\big(\|\nabla v\|^2_{L^\infty(B_{r/2}(y))}\, r^2 + r^{3\alpha_0}\big) \le C r^{3\alpha_0}.$$

This, combined with Morrey's lemma, yields $u \in C^{3\alpha_0/2}(B_{7r_0/8}(x_0))$. Repeating this argument, we can show that $u \in C^\alpha(B_{3r_0/4}(x_0))$ for any $0 < \alpha < 1$, and

$$r^{2-n} \int_{B_r(y)} |\nabla u|^2\, dx \le C r^{2\alpha} \quad \text{for all } y \in B_{3r_0/4}(x_0),\ 0 < r < \frac{r_0}{4}. \tag{5-5}$$

Step 2: There exists $0 < \beta_0 < 1$ such that $u \in C^{1,\beta_0}\big(B_{r_0/2}(x_0) \cap \overline{B^\pm}, N\big)$. There are two cases to consider:

Case I: $x_0 = (x_0', x_0^n) \in B_1^\pm$. We may assume $0 < r_0 < |x_0^n|$, so that $B_{r_0}(x_0) \subset B^\pm$. For $B_r(x) \subset B_{r_0}(x_0)$, let $v : B_r(x) \to \mathbb{R}^k$ solve

$$\begin{cases} \Delta_g v = 0 & \text{in } B_r(x), \\ v = u & \text{on } \partial B_r(x). \end{cases} \tag{5-6}$$

Then by (5-5), for any $\frac{2}{3} < \alpha < 1$,

$$\int_{B_r(x)} |\nabla(u-v)|^2\, dx \le C \int_{B_r(x)} |\nabla u|^2 |u-v|\, dx \le C\, r^{3\alpha+n-2}. \tag{5-7}$$

Also, since $g \in C^{0,1}(B_{r_0}(x_0))$, we have for any $0 < \beta < 1$ that $v \in C^{1,\beta}(B_{r/2}(x))$ and

$$\fint_{B_s(x)} |\nabla v - (\nabla v)_{B_s(x)}|^2\, dx \le C\Big(\frac{s}{r}\Big)^{2\beta} \fint_{B_r(x)} |\nabla u - (\nabla u)_{B_r(x)}|^2\, dx, \tag{5-8}$$

for $< s \le r/2$. (Here $⨍_E f = \frac{1}{|E|}\int_E f\,dx$.) Note that (5-8) also holds trivially for $r/2 \le s \le r$. Combining (5-7) and (5-8) we obtain, for any $0 < \theta < 1$,

$$\begin{aligned}⨍_{B_{\theta r}(x)} \left|\nabla u - (\nabla u)_{B_{\theta r}(x)}\right|^2 dx &\le 2\left(⨍_{B_{\theta r}(x)} |\nabla u - \nabla v|^2\,dx + ⨍_{B_{\theta r}(x)} \left|\nabla v - (\nabla v)_{B_{\theta r}(x)}\right|^2 dx\right)\\ &\le C\left(\theta^{2\beta} ⨍_{B_r(x)} \left|\nabla u - (\nabla u)_{B_r(x)}\right|^2 dx + \theta^{-n} r^{3\alpha-2}\right).\end{aligned}$$

For $(3\alpha-2)/2 < \beta_0 < \beta$, let $0 < \theta_0 < 1$ be such that $C\theta_0^{2\beta} = \theta_0^{2\beta_0}$. Then

$$⨍_{B_{\theta_0 r}(x)} \left|\nabla u - (\nabla u)_{B_{\theta_0 r}(x)}\right|^2 dx \le \theta_0^{2\beta_0} ⨍_{B_r(x)} \left|\nabla u - (\nabla u)_{B_r(x)}\right|^2 dx + C r^{3\alpha-2}. \tag{5-9}$$

Iterating (5-9) m-times, $m \ge 1$, yields

$$\begin{aligned}⨍_{B_{\theta_0^m r}(x)} &\left|\nabla u - (\nabla u)_{B_{\theta_0 r}(x)}\right|^2 dx\\ &\le \left(\theta_0^m\right)^{2\beta_0} ⨍_{B_r(x)} \left|\nabla u - (\nabla u)_{B_r(x)}\right|^2 dx + C(\theta_0^m r)^{3\alpha-2} \sum_{j=1}^m \theta_0^{j(2\beta_0-(3\alpha-2))}\\ &\le (\theta_0^m)^{3\alpha-2}\left(⨍_{B_r(x)} \left|\nabla u - (\nabla u)_{B_r(x)}\right|^2 dx + C r^{3\alpha-2}\right).\end{aligned} \tag{5-10}$$

This clearly implies that $\nabla u \in C^{3\alpha/2-1}(B_{r_0}(x_0))$.

Case II: $x_0 = (x_0', 0) \in \Gamma_1$. For simplicity, we assume $x_0' = 0$. Define $\bar g$ on B_1 by

$$\bar g(x) = \begin{cases} \lim_{t\downarrow 0^+} g(0', t) & \text{if } x \in B_1^+\\ \lim_{t\uparrow 0^-} g(0', t) & \text{if } x \in B_1^-.\end{cases}$$

Then we have

$$|g(x) - \bar g(x)| \le C|x|, \quad x \in B_1. \tag{5-11}$$

Moreover, by suitable dilations and rotations of the coordinate system, (1-4) implies that there exists a positive constant $k \ne 1$ such that

$$\bar g(x) = (1 + (k-1)\chi_{B_1^-}(x))g_0, \quad x \in B_1,$$

where $\chi_{B_1^-}$ is the characteristic function of B_1^-.

For $0 < r < r_0/2$, let $v : B_r(0) \to \mathbb{R}^k$ solve

$$\begin{cases} \Delta_{\bar g} v = 0 & \text{in } B_r(0),\\ v = u & \text{on } \partial B_r(0).\end{cases} \tag{5-12}$$

Then we have

$$\operatorname{osc}_{B_r(0)} v \le \operatorname{osc}_{B_r(0)} u \le C r^{\alpha}, \quad \int_{B_r(0)} |\nabla v|^2\,dx \le C \int_{B_r(0)} |\nabla u|^2 \le C r^{n-2+2\alpha}.$$

Multiplying (1-1) and (5-12) by $u - v$ and integrating over $B_r(0)$, we obtain

$$\begin{aligned}
\int_{B_r(0)} &|\nabla(u-v)|^2\,dx \\
&\le \int_{B_r(0)} g^{ij}(u-v)_i(u-v)_j\sqrt{g}\,dx \\
&\le C\int_{B_r(0)} |\nabla u|^2|u-v|\,dx + \int_{B_r(0)} |\sqrt{g}g^{ij} - \sqrt{\bar g}\bar g^{ij}||v_i||(u-v)_j|\,dx \\
&\le C \operatorname{osc}_{B_r(0)} v \int_{B_r(0)} |\nabla u|^2\,dx + Cr^2\int_{B_r(0)} |\nabla v|^2 + \frac{1}{2}\int_{B_r(0)} |\nabla(u-v)|^2\,dx \\
&\le Cr^{n-2+3\alpha} + Cr^{n+\alpha} + \frac{1}{2}\int_{B_r(0)} |\nabla(u-v)|^2\,dx.
\end{aligned}$$

This implies

$$\int_{B_r(0)} |\nabla(u-v)|^2\,dx \le Cr^{n-2+3\alpha}. \tag{5-13}$$

It is well-known that $v \in C^\infty\big(\overline{B_s^{\pm}(0)}\big)$ for any $0 < s < r$. In fact, (5-12) is equivalent to:

$$\frac{\partial}{\partial x_i}\left((1 + (k^{n/2} - 1)\chi_{B_1^-})\frac{\partial v}{\partial x_i}\right) = 0 \quad \text{in } B_r(0), \tag{5-14}$$

we conclude

(i) $\partial v/\partial x_n$ satisfies the jump property on Γ_1:

$$\lim_{x_n \downarrow 0^+} \frac{\partial v}{\partial x_n}(x', x_n) = k^{n/2} \lim_{x_n \uparrow 0^-} \frac{\partial v}{\partial x_n}(x', x_n) \quad \text{for all } (x', 0) \in \Gamma_1 \cap B_r(0).$$

(ii) $\nabla^\alpha v \in C^0(B_r(0))$ for any multiindex $\alpha = (\alpha_1, \dots, \alpha_{n-1}, 0)$.

(iii) $\nabla v \in L^\infty(B_s(0))$ for any $0 < s < r$, and

$$\|\nabla v\|^2_{L^\infty(B_{r/2}(0))} \le Cr^{2-n}\int_{B_r(0)} |\nabla u|^2. \tag{5-15}$$

For $f : B_r(0) \to \mathbb{R}^k$, set

$$\widetilde{D} f := \Big(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_{n-1}}, (1 + (k^{n/2} - 1)\chi_{B_1^-})\frac{\partial f}{\partial x_n}\Big), \tag{5-16}$$

and denote by $(\widetilde{D}f)_s = ⨍_{B_s(0)} \widetilde{D}f\,dx$ the average of $\widetilde{D}f$ over $B_s(0)$. Then, for any $0 < \beta < 1$,

$$⨍_{B_s(0)} \left|\widetilde{D}v - (\widetilde{D}v)_s\right|^2 dx \le C\left(\frac{s}{r}\right)^{2\beta} ⨍_{B_r(0)} \left|\widetilde{D}u - (\widetilde{D}u)_r\right|^2 dx \quad \text{for all } 0 < s \le r.$$

Combining this with (5-13) yields, for any $0 < \theta < 1$,

$$⨍_{B_{\theta r}(0)} \left|\widetilde{D}u - (\widetilde{D}u)_{\theta r}\right|^2 dx \le C\theta^{2\beta} ⨍_{B_r(0)} \left|\widetilde{D}u - (\widetilde{D}u)_r\right|^2 dx + C\theta^{-n} r^{3\alpha-2}.$$

As in case I, iterations of this inequality yield, for any $0 < s \le r$,

$$⨍_{B_s(0)} \left|\widetilde{D}u - (\widetilde{D}u)_s\right|^2 dx \le C\left(\frac{s}{r}\right)^{3\alpha-2} ⨍_{B_r(0)} \left|\widetilde{D}u - (\widetilde{D}u)_r\right|^2 dx + Cs^{3\alpha-2}.$$

This, combined with case I, implies that for any $B_r(x) \subset B_{r_0}(x_0)$ and $0 < s \le r$,

$$⨍_{B_s(x)} \left|\widetilde{D}u - (\widetilde{D}u)_{x,s}\right|^2 dx \le C\left(\frac{s}{r}\right)^{3\alpha-2} ⨍_{B_r(x)} \left|\widetilde{D}u - (\widetilde{D}u)_{x,r}\right|^2 dx + Cs^{3\alpha-2},$$

where $(\widetilde{D}u)_{x,s}$ denotes the average of $\widetilde{D}u$ over $B_s(x)$. It is readily seen that the preceding inequality yields $u \in C^{1,3\alpha/2-1}(B_{r_0/2}(x_0) \cap \overline{B_1^{\pm}})$ and $u \in C^{0,1}(B_{r_0/2}(x_0))$. This completes the proof. □

Proof of Theorem 1.1 . Define the singular set

$$\Sigma = \left\{ x \in \Omega : \varliminf_{r \to 0} r^{2-n} \int_{B_r(x)} |\nabla u|^2\,dx \ge \epsilon_0^2 \right\}.$$

Then by a covering argument we have $H^{n-2}(\Sigma) = 0$; see [Evans and Gariepy 1992]. For any $x_0 \in \Omega \setminus \Sigma$, there exists $0 < r_0 < \operatorname{dist}(x_0, \partial\Omega)$ such that

$$r_0^{2-n} \int_{B_{r_0}(x)} |\nabla u|^2\,dx \le \epsilon_0^2.$$

Hence by Theorems 2.1, 4.1, and 5.1, we have

$$u \in C^{1,\alpha}(B_{r_0/2}(x_0) \cap \overline{\Omega^{\pm}}, N) \quad \text{and} \quad u \in C^{0,1}(B_{r_0/2}(x_0), N),$$

for some $0 < \alpha < 1$. In particular, we have

$$\lim_{r \downarrow 0} r^{2-n} \int_{B_r(x)} |\nabla u|^2\,dx = 0 \quad \text{for all } x \in B_{r_0/2}(x_0),$$

so that $B_{r_0/2}(x_0) \cap \Sigma = \varnothing$, i.e., Σ is closed. This completes the proof. □

6. Harmonic maps to manifolds supporting convex distance square functions

In this section, we consider weakly harmonic maps u from (Ω, g), with g the piecewise Lipschitz continuous metric as in Theorem 1.1, to (N, h), whose universal cover $(\widetilde{N}, \tilde{h})$ supports a convex distance square function $d^2_{\widetilde{N}}(\cdot\,,p)$ for any $p \in \widetilde{N}$. We will establish both the global Lipschitz continuity and piecewise $C^{1,\alpha}$-regularity for such harmonic maps u. This can be viewed as a generalization of the well-known regularity theorem of Eells and Sampson [1964] and Hildebrand, Kaul and Widman [Hildebrandt et al. 1977].

The crucial step is the following theorem on Hölder continuity.

Theorem 6.1. *Assume that the metric g is bounded measurable on Ω, i.e., there exist two constants $0 < \lambda < \Lambda < +\infty$ such that $\lambda \mathbb{I}_n \le g(x) \le \Lambda \mathbb{I}_n$ for a.e. $x \in \Omega$, and the universal cover $(\widetilde{N}, \tilde{h})$ of (N, h) supports a convex distance square function $d^2_{\widetilde{N}}(\cdot\,,p)$ for any $p \in \widetilde{N}$. If $u \in H^1(\Omega, N)$ is a weakly harmonic map, then there exists $\alpha \in (0,\ 1)$ such that $u \in C^\alpha(\Omega, N)$.*

Proof. Here we sketch a proof that is based on modifications of that in [Lin 1997]. Similar ideas have been used by Evans in his celebrated work [1982] and by Caffarelli [1982] for quasilinear systems under smallness conditions. First, by lifting $u : \Omega \to N$ to a harmonic map $\tilde{u} : \Omega \to \widetilde{N}$, we may assume $(N, h) = (\widetilde{N}, \tilde{h})$ and $d^2_N(\cdot\,,p)$ is convex on N for any $p \in N$.

We first claim that

$$\Delta_g d^2(u, p) \ge 0. \tag{6-1}$$

In fact, by the chain rule of harmonic maps (see [Jost 1991]), we have

$$\Delta_g d^2(u, p) = \nabla_u d^2(u, p)(\Delta_g u) + \nabla^2_u d^2(u, p)(\nabla u, \nabla u)_g.$$

Since $\Delta_g u \perp T_u N$, $\nabla_u d^2(u, p) \in T_u N$, the first term in the right side vanishes. By the convexity of d^2_N, the second term in the right side satisfies

$$\nabla^2_u d^2(u, p)(\nabla u, \nabla u)_g \ge 0.$$

Since $u \in H^1(\Omega, N)$, by suitably choosing $p \in N$ and applying Poincaré inequality and Harnack's inequality, (6-1) implies $u \in L^\infty_{\text{loc}}(\Omega, N)$.

For a set $E \subset N$, let $\operatorname{diam}_N E$ denote the diameter of E with respect to the distance function $d_N(\cdot\,,\cdot)$. For any ball $B_r(x) \subset \Omega$, we want to show that $u \in C^\alpha(B_{r/2}(x))$ for some $0 < \alpha < 1$. To do it, set $C_r := \operatorname{diam}_N u(B_r(x))$. We may assume $C_r > 0$ (otherwise, u is constant on $B_r(x)$ and we are done). Now we want to show that there exists $0 < \delta_0 = \delta_0(N) \le \frac{1}{2}$ such that

$$\operatorname{diam}_N u(B_{\delta_0 r}(x)) \le \tfrac{1}{2} C_r. \tag{6-2}$$

Since $u_r(y) = u(x+ry) : B_1(0) \to N$ is a harmonic map $(B_1(0), g_r)$, with $g_r(y) = g(x+ry)$, we may, for simplicity, assume $x = 0$ and $r = 2$. For any $0 < \epsilon < \frac{1}{2}$, since $u(B_1) \subset N$ is a bounded set, there exists $m = m(\epsilon) \geq 1$ such that $u(B_1)$ is covered by m balls $B^1, \dots, B^m$ of radius ϵC_1.

Claim. *There exists sufficiently small* $\epsilon > 0$ *such that* $u(B_{1/2})$ *can be covered by at most* $(m-1)$ *balls among* $B^1, \dots, B^m$.

To see this, let $x_i \in B_1$ such that $B^i \subset B_{2\epsilon\, C_1}(p_i)$, $p_i = u(x_i)$, for $1 \leq i \leq m$. Let $1 \leq m' \leq m$ be the maximum number of points in $\{p_i\}_{i=1}^m$ such that the distance between any two of them is at least $C_1/32$. Thus the sets $B_{C_1/16}(p_i)$, for $1 \leq i \leq m'$, cover $u(B_1)$. For convenience, set $U_i = u^{-1}\big(B^N(p_i, C_1/16)\big)$, the notation $B^N(x, R)$ referring to the ball in N with center x and radius R. We will show that there exists $i_0 \in \{1, \dots, m'\}$ such that

$$\tfrac{1}{4}C_1^2 \leq \sup_{x \in B_2} d_N^2(u(x), p_{i_0}) \leq C_1^2, \tag{6-3}$$

and

$$H^n(U_{i_0} \cap B_{1/2}) \geq c_0, \tag{6-4}$$

for some universal constant $c_0 > 0$. Indeed, since $B_{1/2} \subset \bigcup_{i=1}^{m'} U_i$, we have

$$\sum_{i=1}^{m'} H^n(U_i \cap B_{1/2}) \geq H^n(B_{1/2}).$$

Hence there exists $i_0 \in \{1, \dots, m'\}$ such that

$$H^n(U_i \cap B_{1/2}) \geq c_0 := \frac{1}{m'} H^n(B_{1/2}).$$

This implies (6-4). Now (6-3) follows from the triangle inequality.

Next we define

$$f(x) := \sup_{z \in B_1} d_N^2(u(z), p_{i_0}) - d_N^2(u(x), p_{i_0}), \quad x \in B_1.$$

It is clear that $f \geq 0$ in B_1, and (6-1) implies $\Delta_g f \leq 0$ in B_1. By Moser's Harnack inequality, we have

$$\begin{aligned}
\inf_{B_{1/2}} f &\geq C \fint_{B_1} f \geq C \int_{B_{1/2}} f \geq C \int_{B_{1/2} \cap U_{i_0}} f \\
&\geq C\big(\sup_{B_1} d_N^2(u, p_{i_0}) - \sup_{B_1 \cap U_{i_0}} d_N^2(u, p_{i_0})\big) H^n\big(B_{1/2} U_{i_0}\big) \\
&\geq C\left(\tfrac{1}{4}C_1^2 - \tfrac{1}{256}C_1^2\right) c_0 =: \theta_0^2 C_1^2
\end{aligned}$$

for some universal constant $\theta_0 > 0$. This implies

$$\sup_{z\in B_1} d_N(u(z), p_{i_0}) - \sup_{z\in B_{1/2}} d_N(u(z), p_{i_0}) \geq \theta_0 C_1 = (1-\theta_0)C_1. \tag{6-5}$$

Now we argue that the claim follows from (6-5). For, otherwise, we would have $u(B_{1/2}) \cap B_{2\epsilon C_1}(p_j) \neq \varnothing$ for all $1 \leq j \leq m$. Let $z_0 \in B_1$ be such that

$$\epsilon C_1 + d_N(u(z_0), p_{i_0}) \geq \sup_{B_1} d_N(u(z), p_{i_0}).$$

Since $u(B_1) \subset \bigcup_{i=}^m B_{2\epsilon C_1}(p_i)$, there exists $p_{i_1} \in \{p_1, \ldots, p_m\}$ such that $u(z_0) \in B_{2\epsilon C_1}(p_{i_1})$. Since $u(B_{1/2}) \cap B_{2\epsilon C_1}(p_{i_1}) \neq \varnothing$, there exists $z_1 \in B_{1/2}$ such that $u(z_1) \in B_{2\epsilon C_1}(p_{i_1})$. Therefore we have $d_N(u(z_1), u(z_0)) \leq 2\epsilon C_1$. Therefore we have

$$\begin{aligned}\sup_{z\in B_1} d_N(u(z), p_{i_0}) - \sup_{z\in B_{1/2}} d_N(u(z), p_{i_0}) &\leq \epsilon C_1 + d_N(u(z_0), p_{i_0}) - d_N(u(z_1), p_{i_0}) \\ &\leq \epsilon C_1 + d_N(u(z_0), u(z_1)) \leq 3\epsilon C_1.\end{aligned}$$

This contradicts (6-5) if $\epsilon > 0$ is chosen to be sufficiently small.

From this claim, we have either

(i) $\operatorname{diam}_N u(B_{1/2}) \leq \frac{1}{2} C_1$ — in which case (6-2) holds with $\delta_0 = \frac{1}{2}$ — or

(ii) $\operatorname{diam}_N u(B_{1/2}) > \frac{1}{2} C_1$.

Then we consider $v(x) = u(x/2) : B_1 \to N$ and conclude:

- v is a harmonic map on $(B_1, g_{1/2})$, with the metric $g_{1/2}(x) = g(x/2)$.
- $\frac{1}{2} C_1 < \operatorname{diam}_N v(B_1) \leq C_1$.
- $v(B_1)$ is covered by at most $m-1$ balls $B_1, \ldots, B^{m-1}$ of radius ϵC_1.

Thus the claim is applicable to v so that $u(B_{1/4}) = v(B_{1/2})$ can be covered by at most $m-2$ balls among $B^1, \ldots, B^{m-1}$.

If $\operatorname{diam}_N v(B_{1/2}) \leq \frac{1}{2} C_1$, we are done. Otherwise, we can repeat the above argument. It is clear that the process can at most be repeated m times, and the process will not be stopped at step $k_0 \leq m$ unless $\operatorname{diam}_N u(B_{2^{-k_0}}) \leq \frac{1}{2} C_1$. Thus (6-2) is proven.

It is readily seen that iteration of (6-2) implies Hölder continuity. □

Proof of Theorem 1.2. First, by Theorem 6.1 and the argument from Section 4, we can show that for some $0 < \alpha < 1$,

$$\int_{B_r(x)} |\nabla u|^2 \, dx \leq C r^{n-2+2\alpha} \quad \text{for all } B_r(x) \subset \Omega.$$

Then we can follow the proof of (5-2) to show that $u \in C^{0,1}(\Omega) \cap C^{1,\alpha}(\Omega^{\pm} \cup \Gamma, N)$. □

Acknowledgements

Part of this work was completed while Li visited the University of Kentucky. He thanks the Department of Mathematics for its hospitality. Li was partially supported by SRFDPHE (20100003120005) and NSFC in China (11071020, 11126038) and Program for Changjiang Scholars and Innovative Research Team in University in China (IRT0908). Wang is partially supported by NSF grant 1000115.

References

[Avellaneda and Lin 1987] M. Avellaneda and F.-H. Lin, "Compactness methods in the theory of homogenization", *Comm. Pure Appl. Math.* **40**:6 (1987), 803–847. MR 88i:35019 Zbl 0632.35018

[Bethuel 1993] F. Bethuel, "On the singular set of stationary harmonic maps", *Manuscripta Math.* **78**:4 (1993), 417–443. MR 94a:58047 Zbl 0792.53039

[Byun and Wang 2010] S.-S. Byun and L. Wang, "Elliptic equations with measurable coefficients in Reifenberg domains", *Adv. Math.* **225**:5 (2010), 2648–2673. MR 2012a:35046 Zbl 1198.35068

[Caffarelli 1982] L. A. Caffarelli, "Regularity theorems for weak solutions of some nonlinear systems", *Comm. Pure Appl. Math.* **35**:6 (1982), 833–838. MR 83m:35052 Zbl 0496.35035

[Dong and Kim 2010] H. Dong and D. Kim, "Elliptic equations in divergence form with partially BMO coefficients", *Arch. Ration. Mech. Anal.* **196**:1 (2010), 25–70. MR 2011b:35086 Zbl 1206.35249

[Dong and Kim 2011a] H. Dong and D. Kim, "Parabolic and elliptic systems in divergence form with variably partially BMO coefficients", *SIAM J. Math. Anal.* **43**:3 (2011), 1075–1098. MR 2012h:35141 Zbl 1230.35045

[Dong and Kim 2011b] H. Dong and D. Kim, "L_p solvability of divergence type parabolic and elliptic systems with partially BMO coefficients", *Calc. Var. Partial Differential Equat.* **40**:3-4 (2011), 357–389. MR 2011m:35393 Zbl 1216.35160

[Eells and Sampson 1964] J. Eells, Jr. and J. H. Sampson, "Harmonic mappings of Riemannian manifolds", *Amer. J. Math.* **86** (1964), 109–160. MR 29 #1603 Zbl 0122.40102

[Evans 1982] L. C. Evans, "Classical solutions of fully nonlinear, convex, second-order elliptic equations", *Comm. Pure Appl. Math.* **35**:3 (1982), 333–363. MR 83g:35038 Zbl 0469.35022

[Evans 1991] L. C. Evans, "Partial regularity for stationary harmonic maps into spheres", *Arch. Rational Mech. Anal.* **116**:2 (1991), 101–113. MR 93m:58026 Zbl 0754.58007

[Evans and Gariepy 1992] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1992. MR 93f:28001 Zbl 0804.28001

[Giaquinta and Hildebrandt 1982] M. Giaquinta and S. Hildebrandt, "A priori estimates for harmonic mappings", *J. Reine Angew. Math.* **336** (1982), 124–164. MR 84b:58035

[Gilbarg and Trudinger 1983] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Second ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **224**, Springer, Berlin, 1983. MR 86c:35035 Zbl 0562.35001

[Hélein 2002] F. Hélein, *Harmonic maps, conservation laws and moving frames*, Second ed., Cambridge Tracts in Mathematics **150**, Cambridge University Press, Cambridge, 2002. Translated from the 1996 French original, With a foreword by James Eells. MR 2003g:58024 Zbl 1010.58010

[Hildebrandt et al. 1977] S. Hildebrandt, H. Kaul, and K.-O. Widman, "An existence theorem for harmonic mappings of Riemannian manifolds", *Acta Math.* **138**:1-2 (1977), 1–16. MR 55 #6478 Zbl 0356.53015

[Ishizuka and Wang 2008] W. Ishizuka and C. Y. Wang, "Harmonic maps from manifolds of L^∞-Riemannian metrics", *Calc. Var. Partial Differential Equat.* **32**:3 (2008), 387–405. MR 2009f:35064 Zbl 1154.58005

[Jost 1991] J. Jost, *Two-dimensional geometric variational problems*, Pure and Applied Mathematics (New York), John Wiley & Sons Ltd., Chichester, 1991. A Wiley-Interscience Publication. MR 92h:58045 Zbl 0729.49001

[Li and Nirenberg 2003] Y. Li and L. Nirenberg, "Estimates for elliptic systems from composite material", *Comm. Pure Appl. Math.* **56**:7 (2003), 892–925. Dedicated to the memory of Jürgen K. Moser. MR 2004k:35097 Zbl 1125.35339

[Lin 1997] F. H. Lin, "Analysis on singular spaces", pp. 114–126 in *Collection of papers on geometry, analysis and mathematical physics*, World Sci. Publ., River Edge, NJ, 1997. MR 99j:53059 Zbl 1028.58019

[Lin and Yan 2003] F. Lin and X. Yan, "A type of homogenization problem", *Discrete Contin. Dyn. Syst.* **9**:1 (2003), 1–30. MR 2003j:35021

[Rivière 1995] T. Rivière, "Everywhere discontinuous harmonic maps into spheres", *Acta Math.* **175**:2 (1995), 197–226. MR 96k:58059 Zbl 0898.58011

Received March 7, 2012. Revised October 19, 2012.

HAIGANG LI
SCHOOL OF MATHEMATICAL SCIENCES AND LABORATORY OF MATHEMATICS AND COMPLEX SYSTEMS, MINISTRY OF EDUCATION
BEIJING NORMAL UNIVERSITY
BEIJING, 100875
CHINA
hgli@bnu.edu.cn

CHANGYOU WANG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF KENTUCKY
LEXINGTON, KY 40506
UNITED STATES
cywang@ms.uky.edu
http://www.ms.uky.edu/~cywang

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.151

q-HYPERGEOMETRIC DOUBLE SUMS AS MOCK THETA FUNCTIONS

JEREMY LOVEJOY AND ROBERT OSBURN

In memory of Basil Gordon

Recently, Bringmann and Kane established two new Bailey pairs and used them to relate certain q-hypergeometric series to real quadratic fields. We show how these pairs give rise to new mock theta functions in the form of q-hypergeometric double sums. We also prove an identity between one of these sums and two classical mock theta functions introduced by Gordon and McIntosh.

1. Introduction

A *Bailey pair* relative to a is a pair of sequences $(\alpha_n, \beta_n)_{n\geq 0}$ satisfying

$$\beta_n = \sum_{k=0}^{n} \frac{\alpha_k}{(q)_{n-k}(aq)_{n+k}}. \tag{1-1}$$

Here we have used the standard q-hypergeometric notation,

$$(a)_n = (a;q)_n = \prod_{k=1}^{n}(1-aq^{k-1}), \tag{1-2}$$

valid for $n \in \mathbb{N} \cup \{\infty\}$. The *Bailey lemma* states that if (α_n, β_n) is a Bailey pair relative to a, then so is (α'_n, β'_n), where

$$\alpha'_n = \frac{(b)_n(c)_n(aq/bc)^n}{(aq/b)_n(aq/c)_n}\alpha_n, \qquad \beta'_n = \sum_{k=0}^{n} \frac{(b)_k(c)_k(aq/bc)_{n-k}(aq/bc)^k}{(aq/b)_n(aq/c)_n(q)_{n-k}}\beta_k. \tag{1-3}$$

Inserting (1-3) into (1-1) with $n \to \infty$ gives

$$\sum_{n\geq 0}(b)_n(c)_n(aq/bc)^n\beta_n = \frac{(aq/b)_\infty(aq/c)_\infty}{(aq)_\infty(aq/bc)_\infty}\sum_{n\geq 0}\frac{(b)_n(c)_n(aq/bc)^n}{(aq/b)_n(aq/c)_n}\alpha_n, \tag{1-4}$$

MSC2010: primary 33D15; secondary 05A30, 11F03, 11F37.
Keywords: q-hypergeometric series, mock theta functions, Appell–Lerch series, Hecke-type double sums.

valid whenever both sides converge. For more on Bailey pairs, including historical perspectives and recent advances, see [Andrews 1986b, Chapter 3; 2001], or [Warnaar 2001].

In a recent study of multiplicative q-series, Bringmann and Kane [2011] established two new and interesting Bailey pairs. They showed that (a_n, b_n) is a Bailey pair relative to 1, where

$$a_{2n} = (1-q^{4n})q^{2n^2-2n}\sum_{j=-n}^{n-1} q^{-2j^2-2j}, \tag{1-5}$$

$$a_{2n+1} = -(1-q^{4n+2})q^{2n^2}\sum_{j=-n}^{n} q^{-2j^2}, \tag{1-6}$$

$$b_n = \frac{(-1)^n(q;q^2)_{n-1}}{(q)_{2n-1}}\chi(n \neq 0), \tag{1-7}$$

and (α_n, β_n) is a Bailey pair relative to q, where

$$\alpha_{2n} = \frac{1}{1-q}\left(q^{2n^2+2n}\sum_{j=-n}^{n-1} q^{-2j^2-2j} + q^{2n^2}\sum_{j=-n}^{n} q^{-2j^2}\right), \tag{1-8}$$

$$\alpha_{2n+1} = -\frac{1}{1-q}\left(q^{2n^2+4n+2}\sum_{j=-n}^{n} q^{-2j^2} + q^{2n^2+2n}\sum_{j=-n-1}^{n} q^{-2j^2-2j}\right), \tag{1-9}$$

$$\beta_n = \frac{(-1)^n(q;q^2)_n}{(q)_{2n+1}}. \tag{1-10}$$

These closely resemble Bailey pairs related to seventh order mock theta functions [Andrews 1986a], but surprisingly no q-series obtained by a direct substitution of either (1-5)–(1-7) or (1-8)–(1-10) in (1-4) is a genuine mock theta function. For example, it turns out that substituting (1-5)–(1-7) in (1-4) with $b, c \to \infty$ yields

$$\frac{-q}{(-q)_\infty}\omega(q)$$

where $\omega(q)$ is one of the third order mock theta functions. The presence of the infinite product means that this is not a mock theta function but a *mixed mock modular form*.

Recall that mock theta functions are q-series which were introduced by Ramanujan in his last letter to G. H. Hardy on January 12, 1920. Until 2002, it was not known how these functions fit into the theory of modular forms. Thanks to work of Zwegers [2002] and Bringmann and Ono [2006; 2010], we now know that each of Ramanujan's examples of mock theta functions is the holomorphic part of a weight $\frac{1}{2}$ harmonic weak Maass form $f(\tau)$ (as usual, $q := e^{2\pi i\tau}$ where $\tau = x+iy \in \mathbb{H}$). Following [Zagier 2009], the holomorphic part of any weight k harmonic weak

Maass form f is called a mock modular form of weight k. If $k = \frac{1}{2}$ and the image of f under the operator $\xi_k := 2iy^k\overline{\partial}/\partial\overline{\tau}$ is a unary theta function, then the holomorphic part of f is called a mock theta function. Specializations of the Appell–Lerch series

$$m(x, q, z) := \frac{1}{j(z, q)} \sum_{r\in\mathbb{Z}} \frac{(-1)^r q^{\binom{r}{2}} z^r}{1 - q^{r-1}xz}$$

are perhaps the most well-known and most important class of mock theta functions [Zagier 2009; Zwegers 2002]. Here $x, z \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$ with neither z nor xz an integral power of q, and

$$j(x, q) := (x)_\infty (q/x)_\infty (q)_\infty.$$

For more on mock theta functions, their remarkable history and modern developments, see [Ono 2009] and [Zagier 2009].

The goal of this paper is to obtain genuine mock theta functions from the Bailey pairs of Bringmann and Kane by first moving a step along the Bailey chain. Applying (1-3) to (a_n, b_n) with $(b, c) \to (-1, \infty)$ and to (α_n, β_n) with $(b, c) \to (-q, \infty)$, we obtain the Bailey pairs recorded in the following two lemmas.

Lemma 1.1. *The pair* (a'_n, b'_n) *is a Bailey pair relative to* 1, *where*

$$a'_{2n} = 2(1 - q^{2n}) q^{4n^2-n} \sum_{j=-n}^{n-1} q^{-2j^2-2j},$$

$$a'_{2n+1} = -2(1 - q^{2n+1}) q^{4n^2+3n+1} \sum_{j=-n}^{n} q^{-2j^2},$$

$$b'_n = \frac{1}{(-q)_n} \sum_{j=1}^{n} \frac{(-1)_j (q; q^2)_{j-1} (-1)^j q^{\binom{j+1}{2}}}{(q)_{n-j} (q)_{2j-1}}.$$

Lemma 1.2. *The pair* (α'_n, β'_n) *is a Bailey pair relative to* q, *where*

$$\alpha'_{2n} = \frac{1}{1-q} \left(q^{4n^2+3n} \sum_{j=-n}^{n-1} q^{-2j^2-2j} + q^{4n^2+n} \sum_{j=-n}^{n} q^{-2j^2} \right),$$

$$\alpha'_{2n+1} = -\frac{1}{1-q} \left(q^{4n^2+7n+3} \sum_{j=-n}^{n} q^{-2j^2} + q^{4n^2+5n+1} \sum_{j=-n-1}^{n} q^{-2j^2-2j} \right),$$

$$\beta'_n = \frac{1}{(-q)_n} \sum_{j=0}^{n} \frac{(-q)_j (q; q^2)_j (-1)^j q^{\binom{j+1}{2}}}{(q)_{n-j} (q)_{2j+1}}.$$

With our main result, we present four mock theta functions arising from the Bailey pairs in Lemmas 1.1 and 1.2. Define

$$\begin{aligned}\theta_{n,p}(x,y,q)&\\ :=\frac{1}{\bar{J}_{0,np(2n+p)}}&\sum_{r^*=0}^{p-1}\sum_{s^*=0}^{p-1} q^{n\binom{r-(n-1)/2}{2}+(n+p)(r-(n-1)/2)(s+(n+1)/2)+n\binom{s+(n+1)/2}{2}}\\ &\times\frac{(-x)^{r-(n-1)/2}(-y)^{s+(n+1)/2}J^3_{p^2(2n+p)}\,j\big(-q^{np(s-r)}x^n/y^n,q^{np^2}\big)}{j\big(q^{p(2n+p)r+p(n+p)/2}(-y)^{n+p}/(-x)^n,q^{p^2(2n+p)}\big)}\\ &\times\frac{j\big(q^{p(2n+p)(r+s)+p(n+p)}x^py^p,q^{p^2(2n+p)}\big)}{j\big(q^{p(2n+p)s+p(n+p)/2}(-x)^{n+p}/(-y)^n,q^{p^2(2n+p)}\big)},\end{aligned}$$

where $r:=r^*+\{(n-1)/2\}$ and $s:=s^*+\{(n-1)/2\}$ with $0\le\{\alpha\}<1$ denoting the fractional part of α. Also, $J_m:=J_{m,3m}$ with $J_{a,m}:=j(q^a,q^m)$, and $\bar{J}_{a,m}:=j(-q^a,q^m)$.

Theorem 1.3. *The following are mock theta functions*:

$$\begin{aligned}\text{(1-11)}\quad \mathcal{W}_1(q)&:=\sum_{n\ge j\ge 1}\frac{(-1)_j(q;q^2)_{j-1}(-1)^jq^{n^2+\binom{j+1}{2}}}{(-q)_n(q)_{n-j}(q)_{2j-1}}\\ &=4m(-q^{17},q^{48},-1)-4q^{-5}m(-q,q^{48},-1)-\frac{2q^2\theta_{3,2}(q^5,q^5,q)}{j(q,q^3)},\end{aligned}$$

$$\begin{aligned}\text{(1-12)}\quad \mathcal{W}_2(q)&:=\sum_{n\ge j\ge 1}\frac{(q;q^2)_n(-1)_j(q;q^2)_{j-1}(-1)^{n+j}q^{\binom{j+1}{2}}}{(-q)_n(q)_{n-j}(q)_{2j-1}}\\ &=4m(-q,q^8,-1)+\frac{2q\theta_{1,2}(-q^2,-q^2,q)}{j(-1,q)},\end{aligned}$$

$$\begin{aligned}\text{(1-13)}\quad \mathcal{W}_3(q)&:=\sum_{n\ge j\ge 1}\frac{(q;q^2)_n(-1)_j(q^2;q^4)_{j-1}(-1)^{n+j}q^{n^2+j^2+j}}{(-q^2;q^2)_n(q^2;q^2)_{n-j}(q^2;q^2)_{2j-1}}\\ &=4m(-q,q^{12},-1)+\frac{2q^3\theta_{1,1}(-q^7,-q^7,q^4)}{j(-q,q^4)},\end{aligned}$$

$$\begin{aligned}\text{(1-14)}\quad \mathcal{W}_4(q)&:=\sum_{n\ge j\ge 0}\frac{(-q)_j(q;q^2)_j(-1)^jq^{n^2+n+\binom{j+1}{2}}}{(-q)_n(q)_{n-j}(q)_{2j+1}}\\ &=-2q^{-4}m(-q^5,q^{48},-1)-2q^{-2}m(-q^{11},q^{48},-1)+\frac{\theta_{3,2}(q^3,q^3,q)}{j(q,q^3)}.\end{aligned}$$

It should be noted that the series defining $\mathcal{W}_2(q)$ does not converge. However, similar to the sixth order mock theta function $\mu(q)$ [Andrews and Hickerson 1991], the sequence of even partial sums and the sequence of odd partial sums both converge. We define $\mathcal{W}_2(q)$ as the average of these two values.

To prove Theorem 1.3 we first use the Bailey machinery to express the $\mathcal{W}_i$ in terms of Hecke-type double sums $f_{a,b,c}(x,y,q)$, where

$$f_{a,b,c}(x,y,q) := \sum_{\mathrm{sg}(r)=\mathrm{sg}(s)} \mathrm{sg}(r)(-1)^{r+s} x^r y^s q^{a\binom{r}{2}+brs+c\binom{s}{2}}. \tag{1-15}$$

Here $x, y \in \mathbb{C}^*$ and $\mathrm{sg}(r) := 1$ for $r \geq 0$ and $\mathrm{sg}(r) := -1$ for $r < 0$. Then we apply recent results of Hickerson and Mortenson [2012] to express the Hecke-type double sums as Appell–Lerch series $m(x,q,z)$ (up to the addition of weakly holomorphic modular forms).

We highlight one connection to classical mock theta functions. Namely, we express the multisum (1-12) in terms of the "eighth order" mock theta functions $S_1(q)$ and $T_1(q)$, defined by (see [Gordon and McIntosh 2000])

$$S_1(q) := \sum_{n\geq 0} \frac{q^{n(n+2)}(-q;q^2)_n}{(-q^2;q^2)_n} \quad \text{and} \quad T_1(q) := \sum_{n\geq 0} \frac{q^{n(n+1)}(-q^2;q^2)_n}{(-q;q^2)_{n+1}}.$$

Corollary 1.4. *We have the identity*

$$\mathcal{W}_2(q) = 2qT_1(q) - qS_1(q).$$

Similar identities involving mock theta functions and multisums were given in [Andrews 2007, Section 13], and more could be deduced from [Bringmann et al. 2010, Theorem 2.4].

The paper proceeds as follows. Some background material on Hecke-type double sums and Appell–Lerch series is collected in Section 2, and Theorem 1.3 and Corollary 1.4 are established in Section 3.

2. Preliminaries

We recall some relevant preliminaries. The most important is a result which allows us to convert from the Hecke-type double sums (1-15) to Appell–Lerch series. Define

$$\begin{aligned} &g_{a,b,c}(x,y,q,z_1,z_0) \\ &:= \sum_{t=0}^{a-1} (-y)^t q^{c\binom{t}{2}} j(q^{bt}x, q^a) m\left(-q^{a\binom{b+1}{2}-c\binom{a+1}{2}-t(b^2-ac)} \frac{(-y)^a}{(-x)^b}, q^{a(b^2-ac)}, z_0\right) \\ &\quad + \sum_{t=0}^{c-1} (-x)^t q^{a\binom{t}{2}} j(q^{bt}y, q^c) m\left(-q^{c\binom{b+1}{2}-a\binom{c+1}{2}-t(b^2-ac)} \frac{(-x)^c}{(-y)^b}, q^{c(b^2-ac)}, z_1\right). \end{aligned} \tag{2-1}$$

Following Hickerson and Mortenson, we use the term "generic" to mean that the parameters do not cause poles in the Appell–Lerch series or in the quotients of theta functions.

Theorem 2.1 [Hickerson and Mortenson 2012, Theorem 0.3]. *Let n and p be positive integers with $(n, p) = 1$. For generic $x, y \in \mathbb{C}^*$,*

$$f_{n,n+p,n}(x, y, q) = g_{n,n+p,n}(x, y, q, -1, -1) + \theta_{n,p}(x, y, q).$$

We shall also require certain facts about $j(x, q)$, $m(x, q, z)$ and $f_{a,b,c}(x, y, q)$. From the definition of $j(x, q)$, we have

$$j(q^n x, q) = (-1)^n q^{-\binom{n}{2}} x^{-n} j(x, q) \tag{2-2}$$

where $n \in \mathbb{Z}$ and

$$j(x, q) = j(q/x, q) = -xj(x^{-1}, q). \tag{2-3}$$

Next, some relevant properties of the sum $m(x, q, z)$ are given in the following (see (2.2b) of Proposition 2.1 and Theorem 2.3 in [Hickerson and Mortenson 2012]).

Proposition 2.2. *For generic $x, z, z_0 \in \mathbb{C}^*$, we have*

$$m(x, q, z) = x^{-1} m(x^{-1}, q, z^{-1}), \tag{2-4}$$

$$m(x, q, z) = m(x, q, z_0) + \frac{z_0 J_1^3 j(z/z_0, q) j(xzz_0, q)}{j(z_0, q) j(z, q) j(xz_0, q) j(xz, q)}. \tag{2-5}$$

Finally, two important transformation properties of $f_{a,b,c}(x, y, q)$ are given in the following (see Propositions 5.1 and 5.2 in [Hickerson and Mortenson 2012]).

Proposition 2.3. *For $x, y \in \mathbb{C}^*$, we have*

$$\begin{aligned} f_{a,b,c}(x, y, q) &= f_{a,b,c}(-x^2 q^a, -y^2 q^c, q^4) - x f_{a,b,c}(-x^2 q^{3a}, -y^2 q^{c+2b}, q^4) \\ &\quad - y f_{a,b,c}(-x^2 q^{a+2b}, -y^2 q^{3c}, q^4) + xyq^b f_{a,b,c}(-x^2 q^{3a+2b}, -y^2 q^{3c+2b}, q^4), \end{aligned} \tag{2-6}$$

$$f_{a,b,c}(x, y, q) = -\frac{q^{a+b+c}}{xy} f_{a,b,c}(q^{2a+b}/x, q^{2c+b}/y, q). \tag{2-7}$$

3. Proof of Theorem 1.3

Proof of Theorem 1.3. Recall that the goal is to express each double sum q-series in terms of Appell–Lerch series. For (1-11), apply Lemma 1.1 and let $b, c \to \infty$ in

(1-4) to obtain

$$\begin{aligned}
\mathcal{W}_1(q) &= \sum_{n\geq 0} q^{n^2} b_n'(q) = \frac{1}{(q)_\infty} \sum_{n\geq 0} q^{n^2} a_n'(q) \\
&= \frac{1}{(q)_\infty}\Big(\sum_{n\geq 0} q^{4n^2} a_{2n}'(q) + \sum_{n\geq 0} q^{4n^2+4n+1} a_{2n+1}'(q)\Big) \\
&= \frac{2}{(q)_\infty}\Big(\sum_{n\geq 0} q^{8n^2-n} \sum_{j=-n}^{n-1} q^{-2j^2-2j} - \sum_{n\geq 0} q^{8n^2+n} \sum_{j=-n}^{n-1} q^{-2j^2-2j} \\
&\qquad - \sum_{n\geq 0} q^{8n^2+7n+2} \sum_{j=-n}^{n} q^{-2j^2} - \sum_{n\geq 0} q^{8n^2+9n+3} \sum_{j=-n}^{n} q^{-2j^2}\Big).
\end{aligned}$$

After replacing n with $-n$ in the second sum and n with $-n-1$ in the fourth sum, we let $n=(r+s+1)/2$, $j=(r-s-1)/2$ in the first two sums and $n=(r+s)/2$, $j=(r-s)/2$ in the latter two sums to find

$$\begin{aligned}
\mathcal{W}_1(q) &= \frac{2q^2}{(q)_\infty}\bigg(\Big(\sum_{\substack{r,s\geq 0\\ r\not\equiv s \ (\mathrm{mod}\ 2)}} - \sum_{\substack{r,s<0\\ r\not\equiv s \ (\mathrm{mod}\ 2)}}\Big) q^{\frac{3}{2}r^2+5rs+\frac{7}{2}r+\frac{3}{2}s^2+\frac{7}{2}s} \\
&\qquad - \Big(\sum_{\substack{r,s\geq 0\\ r\equiv s \ (\mathrm{mod}\ 2)}} - \sum_{\substack{r,s<0\\ r\equiv s \ (\mathrm{mod}\ 2)}}\Big) q^{\frac{3}{2}r^2+5rs+\frac{7}{2}r+\frac{3}{2}s^2+\frac{7}{2}s}\bigg) \\
&= -\frac{2q^2}{(q)_\infty}\bigg(\Big(\sum_{r,s\geq 0} - \sum_{r,s<0}\Big) q^{\frac{3}{2}r^2+5rs+\frac{7}{2}r+\frac{3}{2}s^2+\frac{7}{2}s}\bigg) \\
&= -\frac{2q^2}{(q)_\infty} f_{3,5,3}(q^5,q^5,q).
\end{aligned}$$

By Theorem 2.1, (2-1), (2-2) and (2-3), we have

$$\begin{aligned}
&f_{3,5,3}(q^5,q^5,q) \\
&= -2q^{-2} j(q,q^3) m(-q^{17},q^{48},-1) + 2q^{-7} j(q,q^3) m(-q,q^{48},-1) + \theta_{3,2}(q^5,q^5,q)
\end{aligned}$$

and so

$$\mathcal{W}_1(q) = 4m(-q^{17},q^{48},-1) - 4q^{-5} m(-q,q^{48},-1) - \frac{2q^2\theta_{3,2}(q^5,q^5,q)}{j(q,q^3)}.$$

For (1-12), apply Lemma 1.1 and let $b=-\sqrt{q}$ and $c=\sqrt{q}$ in (1-4) to get

$$\mathcal{W}_2(q) = \sum_{n\geq 0} (-1)^n (q;q^2)_n b_n'(q) = \frac{(q;q^2)_\infty}{2(q^2;q^2)_\infty} \sum_{n\geq 0} (-1)^n a_n'(q)$$

$$= \frac{(q;q^2)_\infty}{2(q^2;q^2)_\infty}\Big(\sum_{n\geq 0} a'_{2n}(q) - \sum_{n\geq 0} a'_{2n+1}(q)\Big)$$

$$= \frac{(q;q^2)_\infty}{(q^2;q^2)_\infty}\Big(\sum_{n\geq 0} q^{4n^2-n}\sum_{j=-n}^{n-1} q^{-2j^2-2j} - \sum_{n\geq 0} q^{4n^2+n}\sum_{j=-n}^{n-1} q^{-2j^2-2j}$$

$$+\sum_{n\geq 0} q^{4n^2+3n+1}\sum_{j=-n}^{n} q^{-2j^2} - \sum_{n\geq 0} q^{4n^2+5n+2}\sum_{j=-n}^{n} q^{-2j^2}\Big).$$

As before, we proceed with

$$\mathcal{W}_2(q) = \frac{(q;q^2)_\infty}{(q^2;q^2)_\infty}\Bigg(\Big(\sum_{\substack{r,s\geq 0\\ r\not\equiv s \pmod 2}} - \sum_{\substack{r,s<0\\ r\not\equiv s \pmod 2}}\Big) q^{\frac{1}{2}r^2+3rs+\frac{3}{2}r+\frac{1}{2}s^2+\frac{3}{2}s+1}$$

$$+\Big(\sum_{\substack{r,s\geq 0\\ r\equiv s \pmod 2}} - \sum_{\substack{r,s<0\\ r\equiv s \pmod 2}}\Big) q^{\frac{1}{2}r^2+3rs+\frac{3}{2}r+\frac{1}{2}s^2+\frac{3}{2}s+1}\Bigg)$$

$$= \frac{(q;q^2)_\infty}{(q^2;q^2)_\infty}\Bigg(\Big(\sum_{r,s\geq 0} - \sum_{r,s<0}\Big) q^{\frac{1}{2}r^2+3rs+\frac{3}{2}r+\frac{1}{2}s^2+\frac{3}{2}s+1}\Bigg)$$

$$= \frac{q(q;q^2)_\infty}{(q^2;q^2)_\infty} f_{1,3,1}(-q^2,-q^2,q).$$

By Theorem 2.1, (2-1) and (2-2), we have

$$f_{1,3,1}(-q^2,-q^2,q) = 2q^{-1} j(-1,q) m(-q,q^8,-1) + \theta_{1,2}(-q^2,-q^2,q)$$

and so

$$\mathcal{W}_2(q) = 4m(-q,q^8,-1) + \frac{2q\theta_{1,2}(-q^2,-q^2,q)}{j(-1,q)}.$$

For (1-13), apply Lemma 1.1 and let $b = q$, $c \to \infty$ and $q \to q^2$ in (1-4) to get

$$\mathcal{W}_3(q) = \sum_{n\geq 0}(-1)^n (q;q^2)_n q^{n^2} b'_n(q^2) = \frac{(q;q^2)_\infty}{(q^2;q^2)_\infty}\sum_{n\geq 0}(-1)^n q^{n^2} a'_n(q^2)$$

$$= \frac{(q;q^2)_\infty}{(q^2;q^2)_\infty}\Big(\sum_{n\geq 0} q^{4n^2} a'_{2n}(q^2) - \sum_{n\geq 0} q^{4n^2+4n+1} a'_{2n+1}(q^2)\Big)$$

$$= \frac{2(q;q^2)_\infty}{(q^2;q^2)_\infty}\Bigg(\sum_{n\ge 0} q^{12n^2-2n}\sum_{j=-n}^{n-1} q^{-4j^2-4j} - \sum_{n\ge 0} q^{12n^2+2n}\sum_{j=-n}^{n-1} q^{-4j^2-4j} + \sum_{n\ge 0} q^{12n^2+10n+3}\sum_{j=-n}^{n} q^{-4j^2} - \sum_{n\ge 0} q^{12n^2+14n+5}\sum_{j=-n}^{n} q^{-4j^2}\Bigg).$$

So,

$$\begin{aligned}\mathcal{W}_3(q) &= \frac{2(q;q^2)_\infty}{(q^2;q^2)_\infty}\Bigg(\Bigg(\sum_{\substack{r,s\ge 0\\ r\not\equiv s \ (\mathrm{mod}\ 2)}} - \sum_{\substack{r,s<0\\ r\not\equiv s \ (\mathrm{mod}\ 2)}}\Bigg) q^{2r^2+8rs+5r+2s^2+5s+3} \\ &\qquad + \Bigg(\sum_{\substack{r,s\ge 0\\ r\equiv s \ (\mathrm{mod}\ 2)}} - \sum_{\substack{r,s<0\\ r\equiv s \ (\mathrm{mod}\ 2)}}\Bigg) q^{2r^2+8rs+5r+2s^2+5s+3}\Bigg) \\ &= \frac{2(q;q^2)_\infty}{(q^2;q^2)_\infty}\Bigg(\Bigg(\sum_{r,s\ge 0} - \sum_{r,s<0}\Bigg) q^{2r^2+8rs+5r+2s^2+5s+3}\Bigg) \\ &= \frac{2q^3(q;q^2)_\infty}{(q^2;q^2)_\infty} f_{1,2,1}(-q^7,-q^7,q^4).\end{aligned}$$

By Theorem 2.1, (2-1), (2-2) and (2-3), we have

$$f_{1,2,1}(-q^7,-q^7,q^4) = 2q^{-3}j(-q,q^4)m(-q,q^{12},-1) + \theta_{1,1}(-q^7,-q^7,q^4)$$

and so

$$\mathcal{W}_3(q) = 4m(-q,q^{12},-1) + \frac{2q^3\theta_{1,1}(-q^7,-q^7,q^4)}{j(-q,q^4)}.$$

Finally, for (1-14), apply Lemma 1.2 and let b, $c \to \infty$ in (1-4) to get

$$\begin{aligned}\mathcal{W}_4(q) &= \sum_{n\ge 0} q^{n^2+n}\beta'_n(q) = \frac{(1-q)}{(q)_\infty}\sum_{n\ge 0} q^{n^2+n}\alpha'_n(q) \\ &= \frac{(1-q)}{(q)_\infty}\Bigg(\sum_{n\ge 0} q^{4n^2+2n}\alpha'_{2n}(q) + \sum_{n\ge 0} q^{4n^2+6n+2}\alpha'_{2n+1}(q)\Bigg) \\ &= \frac{1}{(q)_\infty}\Bigg(\sum_{n\ge 0} q^{8n^2+5n}\sum_{j=-n}^{n-1} q^{-2j^2-2j} + \sum_{n\ge 0} q^{8n^2+3n}\sum_{j=-n}^{n} q^{-2j^2} \\ &\qquad - \sum_{n\ge 0} q^{8n^2+13n+5}\sum_{j=-n}^{n} q^{-2j^2} - \sum_{n\ge 0} q^{8n^2+11n+3}\sum_{j=-n-1}^{n} q^{-2j^2-2j}\Bigg).\end{aligned}$$

After replacing n with $-n-1$ in the third and fourth sums, we let $n = (r+s+1)/2$, $j = (r-s-1)/2$ in the first and fourth sums and $n = (r+s)/2$, $j = (r-s)/2$ in the second and third sums to get

$$\begin{aligned}\mathcal{W}_4(q) &= \frac{1}{(q)_\infty}\Bigg(\Bigg(\sum_{\substack{r,s\geq 0\\ r\not\equiv s \pmod 2}} - \sum_{\substack{r,s<0\\ r\not\equiv s \pmod 2}}\Bigg)q^{\frac{3}{2}r^2+5rs+\frac{13}{2}r+\frac{3}{2}s^2+\frac{13}{2}s+5} \\ &\qquad + \Bigg(\sum_{\substack{r,s\geq 0\\ r\equiv s \pmod 2}} - \sum_{\substack{r,s<0\\ r\equiv s \pmod 2}}\Bigg)q^{\frac{3}{2}r^2+5rs+\frac{3}{2}r+\frac{3}{2}s^2+\frac{3}{2}s}\Bigg) \\ &= \frac{1}{(q)_\infty}\big(2q^{13}f_{3,5,3}(-q^{25},-q^{29},q^4) \\ &\qquad + f_{3,5,3}(-q^9,-q^9,q^4)+q^{11}f_{3,5,3}(-q^{25},-q^{25},q^4)\big) \\ &= \frac{1}{(q)_\infty}f_{3,5,3}(q^3,q^3,q),\end{aligned}$$

where in the last step we have used (2-6) and (2-7). By Theorem 2.1, (2-1), (2-2), (2-3) and (2-4), we have

$$\begin{aligned}&f_{3,5,3}(q^3,q^3,q) \\ &= -2q^{-4}j(q,q^3)m(-q^5,q^{48},-1)-2q^{-2}j(q,q^3)m(-q^{11},q^{48},-1)+\theta_{3,2}(q^3,q^3,q)\end{aligned}$$

and so

$$\mathcal{W}_4(q) = -2q^{-4}m(-q^5,q^{48},-1)-2q^{-2}m(-q^{11},q^{48},-1)+\frac{\theta_{3,2}(q^3,q^3,q)}{j(q,q^3)}. \quad \square$$

Proof of Corollary 1.4. Equations (4.36) and (4.38) of [Hickerson and Mortenson 2012] state that

$$S_1(q) = -2q^{-1}m(-q,q^8,-1)+\frac{\bar{J}_{3,8}J_{2,8}^2}{qJ_{1,8}^2}, \qquad T_1(q) = q^{-1}m(-q,q^8,q^6).$$

By (1-12), (2-3) and (2-5), the claim is equivalent to the identity

$$\frac{\bar{J}_{3,8}J_{2,8}^2}{J_{1,8}^2}+\frac{2q\theta_{1,2}(-q^2,-q^2,q)}{j(-1,q)} = \frac{-2j(q^8,q^{24})^3j(-q^6,q^8)}{j(q^6,q^8)j(-1,q^8)j(-q^7,q^8)}.$$

We have verified this identity using a MAPLE program [Garvan 2010]. $\square$

Acknowledgements

The authors would like to thank Song Heng Chan and Renrong Mao for their helpful comments.

References

[Andrews 1986a] G. E. Andrews, "The fifth and seventh order mock theta functions", *Trans. Amer. Math. Soc.* **293**:1 (1986), 113–134. MR 87f:33011 Zbl 0593.10018

[Andrews 1986b] G. E. Andrews, *q-series: Their development and application in analysis, number theory, combinatorics, physics, and computer algebra*, CBMS Regional Conference Series in Mathematics **66**, American Mathematical Society, Providence, RI, 1986. MR 88b:11063 Zbl 0594.33001

[Andrews 2001] G. E. Andrews, "Bailey's transform, lemma, chains and tree", pp. 1–22 in *Special functions 2000: Current perspective and future directions* (Tempe, AZ), edited by J. Bustoz et al., NATO Sci. Ser. II Math. Phys. Chem. **30**, Kluwer, Dordrecht, 2001. MR 2004i:33027 Zbl 1005.33005

[Andrews 2007] G. E. Andrews, "Partitions, Durfee symbols, and the Atkin–Garvan moments of ranks", *Invent. Math.* **169**:1 (2007), 37–73. MR 2008d:05013 Zbl 1214.11116

[Andrews and Hickerson 1991] G. E. Andrews and D. Hickerson, "Ramanujan's "lost" notebook, VII: The sixth order mock theta functions", *Adv. Math.* **89**:1 (1991), 60–105. MR 92i:11027 Zbl 0739.11042

[Bringmann and Kane 2011] K. Bringmann and B. Kane, "Multiplicative q-hypergeometric series arising from real quadratic fields", *Trans. Amer. Math. Soc.* **363**:4 (2011), 2191–2209. MR 2012f:11208 Zbl 1228.11157

[Bringmann and Ono 2006] K. Bringmann and K. Ono, "The $f(q)$ mock theta function conjecture and partition ranks", *Invent. Math.* **165**:2 (2006), 243–266. MR 2007e:11127 Zbl 1135.11057

[Bringmann and Ono 2010] K. Bringmann and K. Ono, "Dyson's ranks and Maass forms", *Ann. of Math.* (2) **171**:1 (2010), 419–449. MR 2011e:11165 Zbl 05712731

[Bringmann et al. 2010] K. Bringmann, J. Lovejoy, and R. Osburn, "Automorphic properties of generating functions for generalized rank moments and Durfee symbols", *Int. Math. Res. Not.* **2010**:2 (2010), 238–260. MR 2011c:11159 Zbl 1230.05034

[Garvan 2010] F. G. Garvan, "Theta function supplement — preliminary version", online software package, 2010, www.math.ufl.edu/~fgarvan/qmaple/theta-supplement/.

[Gordon and McIntosh 2000] B. Gordon and R. J. McIntosh, "Some eighth order mock theta functions", *J. London Math. Soc.* (2) **62**:2 (2000), 321–335. MR 2001m:33026 Zbl 1031.11007

[Hickerson and Mortenson 2012] D. Hickerson and E. Mortenson, "Hecke-type double sums, Appell–Lerch sums, and mock theta functions, I", preprint, 2012. arXiv 1208.1421

[Ono 2009] K. Ono, "Unearthing the visions of a master: Harmonic Maass forms and number theory", pp. 347–454 in *Current developments in mathematics* (Cambridge, MA, 2008), edited by D. Jerison et al., International Press, Somerville, MA, 2009. MR 2010m:11060 Zbl 1229.11074

[Warnaar 2001] S. O. Warnaar, "50 years of Bailey's lemma", pp. 333–347 in *Algebraic combinatorics and applications* (Gößweinstein, 1999), edited by A. Betten et al., Springer, Berlin, 2001. MR 2002g:33020 Zbl 0972.11003

[Zagier 2009] D. Zagier, "Ramanujan's mock theta functions and their applications (after Zwegers and Ono–Bringmann)", pp. 143–164 Astérisque **326**, 2009. MR 2011h:11049 Zbl 1198.11046

[Zwegers 2002] S. Zwegers, *Mock theta functions*, Ph.D. thesis, Universiteit Utrecht, 2002. Zbl 1194.11058 arXiv 0807.4834

Received April 11, 2012. Revised September 19, 2012.

JEREMY LOVEJOY
CNRS, LIAFA
UNIVERSITÉ DENIS DIDEROT - PARIS 7
CASE 7014
75205 PARIS CEDEX 13
FRANCE
lovejoy@liafa.univ-paris-diderot.fr

ROBERT OSBURN
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY COLLEGE DUBLIN
BELFIELD
DUBLIN 4
IRELAND
robert.osburn@ucd.ie

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.163

MONIC REPRESENTATIONS AND GORENSTEIN-PROJECTIVE MODULES

XIU-HUA LUO AND PU ZHANG

Dedicated to the memory of Hua Feng

Given a finite-dimensional algebra A over a field k, and a finite acyclic quiver Q, let $\Lambda = A \otimes_k kQ$, where kQ is the path algebra of Q over k. Then the category Λ-mod of Λ-modules is equivalent to the category $\mathrm{Rep}(Q, A)$ of representations of Q over A. This yields the notion of monic representations of Q over A. We denote the full subcategory of $\mathrm{Rep}(Q, A)$ consisting of monic representations of Q over A by $\mathrm{Mon}(Q, A)$. It is proved that $\mathrm{Mon}(Q, A)$ has Auslander–Reiten sequences.

The main result of this paper explicitly describes the Gorenstein-projective Λ-modules via the monic representations plus an extra condition. As a corollary, we prove the equivalence of three conditions: A is self-injective; Gorenstein-projective Λ-modules are exactly the monic representations of Q over A; $\mathrm{Mon}(Q, A)$ is a Frobenius category.

1. Introduction

Let A be an Artin algebra, and A-mod the category of finitely generated left A-modules. *A complete A-projective resolution* is an exact sequence of finitely generated projective A-modules

$$P^\bullet = \quad \cdots \to P^{-1} \to P^0 \xrightarrow{d^0} P^1 \to \cdots$$

such that $\mathrm{Hom}_A(P^\bullet, A)$ is also exact. A module $M \in A$-mod is *Gorenstein-projective* if there exists a complete A-projective resolution $P^\bullet$ such that $M \cong \mathrm{Ker}\, d^0$. Let $\mathcal{P}(A)$ be the full subcategory of A-mod of projective modules, and $\mathcal{GP}(A)$ the full subcategory of A-mod of Gorenstein-projective modules. Then

$$\mathcal{P}(A) \subseteq \mathcal{GP}(A) \subseteq {}^\perp A = \{X \in A\text{-mod} \mid \mathrm{Ext}^i_A(X, A) = 0 \text{ for all } i \geq 1\}.$$

The authors were supported by the NSF of China (11271251) and the Doctoral Fund of the Ministry of Education of China (20120073110058). Pu Zhang is the corresponding author.
MSC2010: primary 16G10; secondary 16E65, 16G50, 16G60.
Keywords: representations of a quiver over an algebra, monic representations, Gorenstein-projective modules.

It is clear that $\mathcal{GP}(A) = A$-mod if and only if A is self-injective. If A is of finite global dimension, $\mathcal{GP}(A) = \mathcal{P}(A)$ (but the converse is *not* true); and if A is *a Gorenstein algebra* (that is, inj.dim $_A A < \infty$ and inj.dim $A_A < \infty$), then $\mathcal{GP}(A) = {}^{\perp}A$ (but the converse is *not* true); see, for example, [Enochs and Jenda 2000, Corollary 11.5.3]. This class of modules enjoys more stable properties than the usual projective modules (see [Auslander and Bridger 1969], where it was called a module of G-dimension zero); it becomes a main ingredient in the relative homological algebra [Enochs and Jenda 1995; 2000] and in the representation theory of algebras (see [Auslander and Reiten 1991a; 1991b; Beligiannis 2005; Gao and Zhang 2010; Iyama et al. 2011], for example), and plays a central role in the Tate cohomology of algebras (see [Avramov and Martsinkovsky 2002; Buchweitz 1987], for example). An important feature is that $\mathcal{GP}(A)$ is a Frobenius category with relative projective-injective objects being projective A-modules, and hence the stable category $\underline{\mathcal{GP}(A)}$ of $\mathcal{GP}(A)$ modulo $\mathcal{P}(A)$ is a triangulated category. By [Buchweitz 1987; Happel 1991], the singularity category of a Gorenstein algebra A is triangle equivalent to $\underline{\mathcal{GP}(A)}$. Thus explicitly constructing all the Gorenstein-projective modules is a fundamental problem, and is useful to all of these applications.

On the other hand, the submodule category has been extensively studied by C. M. Ringel and M. Schmidmeier [2006; 2008a; 2008b]; see also [Simson 2007]. By [Kussin et al. 2012] it is also related to the singularity category; see also [Chen 2011]. It turns out that the category of the Gorenstein-projective modules is closely related to the submodule category (see [Li and Zhang 2010; Xiong and Zhang 2012]), or, in general, to the monomorphism category [Zhang 2011]. The present paper explores such a relation in a more general set-up.

Given a finite-dimensional algebra A over a field k, and a finite acyclic quiver Q (here "acyclic" means that Q has no oriented cycles), let

$$\Lambda = A \otimes_k kQ,$$

where kQ is the path algebra of Q over k. We call Λ the path algebra of a finite quiver Q over A. As in the case of $A = k$, Λ-mod is equivalent to the category $\mathrm{Rep}(Q, A)$ of representations of Q over A. This interpretation permits us to introduce the so-called monic representations of Q over A. See Definition 2.2. Let $\mathrm{Mon}(Q, A)$ be the full subcategory of $\mathrm{Rep}(Q, A)$ consisting of monic representations of Q over A. Then $\mathrm{Mon}(Q, A)$ is a resolving, functorially finite subcategory of $\mathrm{Rep}(Q, A)$, and hence has Auslander–Reiten sequences (see Theorem 3.1). The main result of this paper, Theorem 5.1, explicitly describes all the Gorenstein-projective Λ-modules, via the monic representations of Q over A plus an extra condition. We emphasize that here Λ is not necessarily Gorenstein. By our main result, if we know all the Gorenstein-projective A-modules, we know all the Gorenstein-projective Λ-modules, and, in this way, we give an inductive construction of the Gorenstein-projective modules.

The proof of Theorem 5.1 use induction on $|Q_0|$ and a description of the Gorenstein-projective modules over the triangular extension of two algebras via a bimodule which is projective in both sides (Theorem 4.1). As a corollary, we see that A is self-injective if and only if $\mathcal{GP}(\Lambda) = \mathrm{Mon}(Q, A)$, and if and only if $\mathrm{Mon}(Q, A)$ is a Frobenius category (Corollary 6.1). As another corollary, if Q has an arrow, $\mathcal{P}(\Lambda) = \mathrm{Mon}(Q, A)$ if and only if Λ is hereditary (Corollary 6.3).

2. Monic representations of a quiver over an algebra

Throughout this section k is a field, Q a finite quiver, and A a finite-dimensional k-algebra. We consider the path algebra AQ of Q over A, describe its module category, and introduce the concept of monic representations of Q over A. In Subsections 2A–2D, Q is not assumed to be acyclic if not otherwise stated.

2A. Given a finite quiver

$$Q = (Q_0, Q_1, s, e),$$

let $\mathcal{P}$ be the set of paths of Q. We write the conjunction of paths from right to left. If $p = \alpha_l \cdots \alpha_1 \in \mathcal{P}$ with $\alpha_i \in Q_1$, $l \geq 1$, and $e(\alpha_i) = s(\alpha_{i+1})$ for $1 \leq i \leq l-1$, we call l the length of p and denote it by $l(p)$, and define the starting vertex $s(p) = s(\alpha_1)$ and the ending vertex $e(p) = e(\alpha_l)$. We denote a vertex i by e_i, and regard it as a path of length 0, with $s(e_i) = i = e(e_i)$. Let kQ be the path algebra of Q over k. It is well-known that the category kQ-mod of finite-dimensional kQ-modules is equivalent to the category $\mathrm{Rep}(Q, k)$ of finite-dimensional representations of Q over k; see, for example, [Ringel 1984, p. 44].

2B. Let $\Lambda = AQ$ be the free left A-module with basis $\mathcal{P}$. An element of AQ is written as a finite sum $\sum_{p\in\mathcal{P}} a_p p$, where $a_p \in A$ and $a_p = 0$ for all but finitely many p. Then Λ is a k-algebra, with multiplication bilinearly given by

$$(a_p p)(b_q q) = (a_p b_q)(pq),$$

where $a_p b_q$ is the product in A, and pq is the product in kQ. We have isomorphisms $\Lambda \cong A \otimes_k kQ \cong kQ \otimes_k A$ of k-algebras, and we call $\Lambda = AQ$ *the path algebra of Q over A*.

For example, if $Q = \underset{n}{\bullet} \to \cdots \to \underset{1}{\bullet}$, the algebra Λ is given by the upper triangular matrix algebra of A:

$$T_n(A) = \begin{pmatrix} A & A & \cdots & A & A \\ 0 & A & \cdots & A & A \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A & A \\ 0 & 0 & \cdots & 0 & A \end{pmatrix},$$

In general, if Q is acyclic and Q_0 is labeled as $1, \dots, n$ in such a way that $j > i$ whenever there is an arrow $\alpha : j \to i$ in Q_1, then

$$kQ \cong \begin{pmatrix} k & k^{m_{21}} & k^{m_{31}} & \cdots & k^{m_{n1}} \\ 0 & k & k^{m_{32}} & \cdots & k^{m_{n2}} \\ 0 & 0 & k & \cdots & k^{m_{n3}} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & k \end{pmatrix}_{n\times n}, \tag{2-1}$$

where m_{ji} is the number of paths from j to i and $k^{m_{ji}}$ is the direct sum of m_{ji} copies of k, and hence

$$\Lambda \cong \begin{pmatrix} A & A^{m_{21}} & A^{m_{31}} & \cdots & A^{m_{n1}} \\ 0 & A & A^{m_{32}} & \cdots & A^{m_{n2}} \\ 0 & 0 & A & \cdots & A^{m_{n3}} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & A \end{pmatrix}_{n\times n}. \tag{2-2}$$

2C. By definition, *a representation X of Q over A* is a datum

$$X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1),$$

where X_i is an A-module for each $i \in Q_0$ and $X_\alpha : X_{s(\alpha)} \to X_{e(\alpha)}$ is an A-map for each $\alpha \in Q_1$. It is *a finite-dimensional representation* if each X_i is finite-dimensional. We call X_i *the i-th branch* of X. A morphism f from representation X to representation Y is a datum $(f_i, i \in Q_0)$, where $f_i : X_i \to Y_i$ is an A-map for each $i \in Q_0$, such that, for each arrow $\alpha : j \to i$, the diagram

$$\begin{array}{ccc} X_j & \xrightarrow{f_j} & Y_j \\ \downarrow{\scriptstyle X_\alpha} & & \downarrow{\scriptstyle Y_\alpha} \\ X_i & \xrightarrow{f_i} & Y_i \end{array} \tag{2-3}$$

commutes. We call f_i *the i-th branch* of f. If $p = \alpha_l \cdots \alpha_1 \in \mathcal{P}$ with $\alpha_i \in Q_1$, $l \geq 1$, and $e(\alpha_i) = s(\alpha_{i+1})$ for $1 \leq i \leq l-1$, we put X_p to be the A-map $X_{\alpha_l} \cdots X_{\alpha_1}$. Denote by $\mathrm{Rep}(Q, A)$ the category of finite-dimensional representations of Q over A. A morphism $f = (f_i, i \in Q_0)$ in $\mathrm{Rep}(Q, A)$ is a monomorphism (epimorphism, isomorphism) if and only if f_i is injective (surjective, an isomorphism) for each $i \in Q_0$.

Lemma 2.1. *Let Λ be the path algebra of Q over A. Then we have an equivalence Λ-mod $\cong \mathrm{Rep}(Q, A)$ of categories.*

We omit the proof of Lemma 2.1, which is similar to the case of $A = k$; see [Auslander et al. 1995, Theorem 1.5, p. 57; Ringel 1984, p. 44]. Throughout this paper we will identify a Λ-module with a representation of Q over A. Under this identification, a Λ-module X is a representation $(X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$ of Q

over A, where $X_i = (1e_i)X$, 1 is the identity of A, and the A-action on X_i is given by $a(1e_i)x = (1e_i)(ae_i)x$ for all $x \in X$ and $a \in A$; and $X_\alpha : X_{s(\alpha)} \to X_{e(\alpha)}$ is the A-map given by the left action by $1\alpha \in \Lambda$. On the other hand, a representation $(X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$ of Q over A is a Λ-module $X = \bigoplus_{i \in Q_0} X_i$, with the Λ-action on X given by

$$(ap)(x_i) = \begin{cases} 0 & \text{if } s(p) \neq i, \\ ax_i & \text{if } p = e_i, \\ aX_p(x_i) \in X_{e(p)} & \text{if } s(p) = i \text{ and } l(p) \geq 1, \end{cases}$$

for all $a \in A$, $p \in \mathcal{P}$, $x_i \in X_i$. Let $f : X \to Y$ be a morphism in $\mathrm{Rep}(Q, A)$. Then $\mathrm{Ker}\, f$ and $\mathrm{Coker}\, f$ can be explicitly written out. For example, $\mathrm{Coker}\, f = (\mathrm{Coker}\, f_i, \widetilde{Y}_\alpha, i \in Q_0, \alpha \in Q_1)$, where, for each arrow $\alpha : j \to i$,

$$\widetilde{Y}_\alpha : \mathrm{Coker}\, f_j \to \mathrm{Coker}\, f_i$$

is the A-map induced by Y_α; see (2-3). A sequence of morphisms

$$0 \longrightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \longrightarrow 0$$

in $\mathrm{Rep}(Q, A)$ is exact if and only if each

$$0 \longrightarrow X_i \xrightarrow{f_i} Y_i \xrightarrow{g_i} Z_i \longrightarrow 0$$

is exact in A-mod, for $i \in Q_0$.

In the following, if Q_0 is labeled as $1, \ldots, n$, we also write a representation X of Q over A as

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}_{(X_\alpha,\ \alpha \in Q_1)},$$

and a morphism in $\mathrm{Rep}(Q, A)$ as

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}.$$

2D. The following is a central notion of this paper.

Definition 2.2. A representation $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$ of Q over A is *a monic representation*, or a monic Λ-module, if, for each $i \in Q_0$, the A-map

$$(X_\alpha)_{\substack{\alpha \in Q_1 \\ e(\alpha) = i}} : \bigoplus_{\substack{\alpha \in Q_1 \\ e(\alpha) = i}} X_{s(\alpha)} \to X_i$$

is injective, or, equivalently, if the following two conditions are satisfied.

(m1) For each $\alpha \in Q_1$, the map $X_\alpha : X_{s(\alpha)} \to X_{e(\alpha)}$ is injective.

(m2) For each $i \in Q_0$, there holds $\sum_{\substack{\alpha \in Q_1 \\ e(\alpha)=i}} \operatorname{Im} X_\alpha = \bigoplus_{\substack{\alpha \in Q_1 \\ e(\alpha)=i}} \operatorname{Im} X_\alpha$.

Denote by $\operatorname{Mon}(Q, A)$ the full subcategory of $\operatorname{Rep}(Q, A)$ consisting of monic representations of Q over A. We call $\operatorname{Mon}(Q, A)$ *the monomorphism category of A over Q*.

If Q is a quiver in which, for any vertex i, there is at most one arrow ending at i, condition (m2) vanishes. For example, if $Q = \bullet \to \bullet$, then $\operatorname{Mon}(Q, A)$ is called *the submodule category of A* in [Ringel and Schmidmeier 2006; 2008a]. If

$$Q = \underset{n}{\bullet} \to \cdots \to \underset{1}{\bullet},$$

$\operatorname{Mon}(Q, A)$ is called *the filtered chain category of A* in [Arnold 2000; Simson 2007].

2E. Let Q be a finite acyclic quiver, A a finite-dimensional algebra, and $\Lambda = A \otimes_k kQ$. Throughout this paper, we label the vertices of Q as $1, 2, \dots, n$, in such a way that if there is an arrow from j to i, then $j > i$. Denote by $P(i)$ the indecomposable projective kQ-module at $i \in Q_0$. It is clear that $P(i) \in \operatorname{Mon}(Q, k)$; it follows that $M \otimes_k P(i) \in \operatorname{Mon}(Q, A)$ for $M \in A$-mod. Thus we have the functors

$$- \otimes_k P(i) : A\text{-mod} \to \operatorname{Mon}(Q, A), \quad -_i : \operatorname{Rep}(Q, A) \to A\text{-mod}$$

(by taking the i-th branch).

We also need the adjoint pair $(- \otimes_k P(i), -_i)$.

Lemma 2.3. *For each object $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1) \in \Lambda$-mod and each A-module M, we have isomorphisms of abelian groups, which are natural in both positions*

$$\operatorname{Hom}_\Lambda(M \otimes_k P(i), X) \cong \operatorname{Hom}_A(M, X_i) \tag{2-4}$$

for all $i \in Q_0$.

Proof. For $f = (f_j, j \in Q_0) \in \operatorname{Hom}_\Lambda(M \otimes_k P(i), X)$, we have $f_i \in \operatorname{Hom}_A(M, X_i)$. Since $M \otimes_k P(i) = (M \otimes_k e_j kQ e_i, \mathrm{id}_M \otimes \alpha, j \in Q_0, \alpha \in Q_1)$, it follows from the commutative diagram (2-3) that

$$f_j = \begin{cases} 0 & \text{if there are no paths from } i \text{ to } j, \\ m \otimes_k p \mapsto X_p f_i(m) & \text{if there is a path } p \text{ from } i \text{ to } j. \end{cases} \tag{2-5}$$

By (2-5) we see that $f \mapsto f_i$ gives an injective map

$$\operatorname{Hom}_\Lambda(M \otimes_k P(i), X) \to \operatorname{Hom}_A(M, X_i).$$

This map is also surjective, since for a given $f_i \in \operatorname{Hom}_A(M, X_i)$, $f = (f_j, j \in Q_0)$ given by (2-5) is indeed a morphism in $\operatorname{Rep}(Q, A)$ from $M \otimes_k P(i)$ to X. □

Proposition 2.4. (i) *The indecomposable projective Λ-modules have the form $P \otimes_k P(i)$, where P is an indecomposable projective A-module, and $P(i)$ is the indecomposable projective kQ-module at $i \in Q_0$.*

(ii) *The indecomposable projective objects in* $\mathrm{Mon}(Q, A)$ *are exactly the indecomposable projective Λ-modules.*

(iii) *If I is an indecomposable injective A-module and $P(i)$ is the indecomposable projective kQ-module at $i \in Q_0$, $I \otimes_k P(i)$ is an indecomposable injective object in* $\mathrm{Mon}(Q, A)$.

Proof. (i) As a direct summand of the regular Λ-module ${}_\Lambda\Lambda$, we see that $P \otimes_k P(i)$ is a projective Λ-module, and each projective Λ-module has this form. By (2-4) we have

$$\mathrm{End}_\Lambda(P \otimes_k P(i)) \cong \mathrm{Hom}_A(P, (P \otimes_k P(i))_i) = \mathrm{End}_A(P),$$

from which we see that $P \otimes_k P(i)$ is indecomposable.

(ii) Note that $P \otimes_k P(i) \in \mathrm{Mon}(Q, A)$. By (i) we know that it is an indecomposable projective object in $\mathrm{Mon}(Q, A)$. On the other hand, it is clear that $\mathrm{Mon}(Q, A)$ is closed under taking subobjects, as a consequence any indecomposable projective object in $\mathrm{Mon}(Q, A)$ has this form.

(iii) Note that $I \otimes_k P(i)$ is an indecomposable object in $\mathrm{Mon}(Q, A)$. Put $L = D(A_A) \otimes_k kQ$, where $D = \mathrm{Hom}_k(-, k)$. It suffices to prove that L is an injective object in $\mathrm{Mon}(Q, A)$, by induction on $|Q_0|$. We write $L = (L_i, L_\alpha, i \in Q_0, \alpha \in Q_1)$.

Let Q' be the quiver obtained from Q by deleting a sink vertex 1, L' the representation in $\mathrm{Rep}(Q', A)$ obtained from L by deleting the branch L_1. We observe that $L' = D(A_A) \otimes_k kQ'$, and by inductive hypothesis L' is an injective object in $\mathrm{Mon}(Q', A)$.

Let $0 \to X \xrightarrow{f} Y \xrightarrow{g} Z \to 0$ be an exact sequence in $\mathrm{Mon}(Q, A)$, with $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$, and $h : X \to L$ a morphism in $\mathrm{Rep}(Q, A)$. Let X' be the representation in $\mathrm{Rep}(Q', A)$ obtained from X by deleting the branch X_1, and similarly for Y', Z'. Then we have an exact sequence

$$0 \longrightarrow X' \xrightarrow{f'} Y' \xrightarrow{g'} Z' \longrightarrow 0$$

in $\mathrm{Mon}(Q', A)$, where f' is the morphism in $\mathrm{Rep}(Q', A)$ obtained from f by deleting the branch f_1, and similarly for g' and for $h' : X' \to L'$. Since L' is an injective object in $\mathrm{Mon}(Q', A)$, by definition we have a morphism

$$u' = \begin{pmatrix} u_2 \\ \vdots \\ u_n \end{pmatrix} : Y' \to L'$$

in $\mathrm{Rep}(Q', A)$ such that $h' = u' f'$. It suffices to construct an A-map

$$u_1 : Y_1 \to L_1$$

such that $u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} : Y \to L$ is a morphism in $\mathrm{Rep}(Q, A)$, and that $h_1 = u_1 f_1$.

First, we have an A-map $u_1' : Y_1 \to L_1$ such that the diagram

$$\begin{tikzcd} X_1 \arrow[r, hook, "f_1"] \arrow[d, "h_1"'] & Y_1 \arrow[dl, "u_1'"] \\ L_1 & \end{tikzcd} .$$

commutes. Consider the A-map

$$(L_\alpha u_{s(\alpha)} - u_1' Y_\alpha)_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} : \bigoplus_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} Y_{s(\alpha)} \to L_1.$$

Since we have the exact sequence of A-modules

$$0 \longrightarrow \bigoplus_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} X_{s(\alpha)} \xrightarrow{\mathrm{diag}(f_{s(\alpha)})} \bigoplus_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} Y_{s(\alpha)} \xrightarrow{\mathrm{diag}(g_{s(\alpha)})} \bigoplus_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} Z_{s(\alpha)} \longrightarrow 0,$$

and since

$$\begin{aligned} (L_\alpha u_{s(\alpha)} - u_1' Y_\alpha)_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} \circ \mathrm{diag}(f_{s(\alpha)}) &= (L_\alpha u_{s(\alpha)} f_{s(\alpha)} - u_1' Y_\alpha f_{s(\alpha)})_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} \\ &= (L_\alpha u_{s(\alpha)} f_{s(\alpha)} - u_1' f_1 X_\alpha)_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} \\ &= (L_\alpha h_{s(\alpha)} - h_1 X_\alpha)_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} \\ &= 0, \end{aligned}$$

where the second equality follows from the fact that $f : X \to Y$ is a morphism in $\mathrm{Rep}(Q, A)$, it follows that $(L_\alpha u_{s(\alpha)} - u_1' Y_\alpha)_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}}$ factors through $\mathrm{diag}(g_{s(\alpha)})$. That is, there is an A-map

$$v_1 : \bigoplus_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} Z_{s(\alpha)} \to L_1,$$

such that

$$(L_\alpha u_{s(\alpha)} - u_1' Y_\alpha)_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} = v_1 \circ \mathrm{diag}(g_{s(\alpha)}).$$

Since L_1 is an injective A-module and

$$(Z_\alpha)_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} : \bigoplus_{\substack{\alpha\in Q_1 \\ e(\alpha)=1}} Z_{s(\alpha)} \to Z_1$$

is an injective A-map, it follows that there is an A-map $w_1 : Z_1 \to L_1$, such that $v_1 = w_1 \circ (Z_\alpha)_{\substack{\alpha\in Q_1\\ e(\alpha)=1}}$. So we have

$$(L_\alpha u_{s(\alpha)} - u_1' Y_\alpha)_{\substack{\alpha\in Q_1\\ e(\alpha)=1}} = w_1 \circ (Z_\alpha)_{\substack{\alpha\in Q_1\\ e(\alpha)=1}} \circ \operatorname{diag}(g_{s(\alpha)}) = (w_1 g_1 Y_\alpha)_{\substack{\alpha\in Q_1\\ e(\alpha)=1}},$$

where the second equality follows from the fact that $g : Y \to Z$ is a morphism in $\operatorname{Rep}(Q, A)$. This means that for each $\alpha \in Q_1$ with $e(\alpha) = 1$ we have

$$L_\alpha u_{s(\alpha)} - u_1' Y_\alpha = w_1 g_1 Y_\alpha. \tag{2-6}$$

Now put $u_1 = u_1' + w_1 g_1 : Y_1 \to L_1$. Then (2-6) together with the inductive hypothesis implies that

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} : Y \to L$$

is a morphism in $\operatorname{Rep}(Q, A)$. It is clear that

$$u_1 f_1 = (u_1' + w_1 g_1) f_1 = u_1' f_1 = h_1.$$

This completes the proof. □

2F. Recall from [Auslander and Reiten 1991a] that a full subcategory $\mathcal{X}$ of A-mod is *resolving* if $\mathcal{X}$ contains all projective A-modules and $\mathcal{X}$ is closed under extensions, kernels of epimorphisms, and direct summands. It is straightforward to verify that $\operatorname{Mon}(Q, A)$ is closed under extensions, kernels of epimorphisms, and direct summands. By Proposition 2.4 we have the following.

Corollary 2.5. *For a finite acyclic quiver Q and a finite-dimensional algebra A,* $\operatorname{Mon}(Q, A)$ *is a resolving subcategory of* $\operatorname{Rep}(Q, A)$.

2G. There is another similar but different notion. Let $A = kQ/I$ be a finite-dimensional k-algebra, where I is an admissible ideal of kQ. An I-bounded representations of Q over k is a datum $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$, where X_i is a k-space for each $i \in Q_0$, and $X_\alpha : X_{s(\alpha)} \to X_{e(\alpha)}$ is a k-linear map for each $\alpha \in Q_1$, such that $\sum_{p\in\mathcal{P}} c_p X_p = 0$ for each element $\sum_{p\in\mathcal{P}} c_p p \in I$, where $l(p) \geq 2$ and $c_p \in k$. An I-bounded representation $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$ of Q over k is *a monic representation*, if for each $i \in Q_0$ the k-linear map

$$(X_\alpha)_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} : \bigoplus_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} X_{s(\alpha)} \to X_i$$

is injective. Let $\operatorname{Rep}(Q, I, k)$ be the category of finite-dimensional I-bounded representations of Q over k. There is an equivalence of categories between A-mod and $\mathrm{Rep}(Q, I, k)$; see [Auslander et al. 1995, Proposition 1.7, p. 60; Ringel 1984, p. 45]. Let $\mathrm{Mon}(Q, I, k)$ denote the full subcategory of $\mathrm{Rep}(Q, I, k)$ of I-bounded monic representations Q over k. Then $\mathrm{Mon}(Q, 0, k) = \mathrm{Mon}(Q, k)$.

Proposition 2.6. *Let $A = kQ/I$ be a finite-dimensional k-algebra, where I is an admissible ideal of kQ. Then $\mathcal{P}(A) \subseteq \mathrm{Mon}(Q, I, k)$ if and only if A is hereditary.*

Proof. If A is hereditary, $I = 0$. It is clear $\mathcal{P}(kQ) \subseteq \mathrm{Mon}(Q, 0, k)$.

Conversely, if $I \neq 0$, take an element $\sum_{p \in \mathcal{P}} c_p p \in I$ with $l(p) \geq 2$ and $c_p \in k$. We may assume that all the paths p with $c_p \neq 0$ have the same starting vertex j and the same ending vertex i. Consider the projective A-module $P(j) = Ae_j$. As an I-bounded representation of Q over k, we write $P(j)$ as

$$P(j) = (e_t kQe_j,\ f_\alpha, t \in Q_0, \alpha \in Q_1).$$

Let $\alpha_1, \dots, \alpha_m$ be all the arrows of Q ending at i. We claim that

$$(f_{\alpha_v})_{1 \leq v \leq m} : \bigoplus_{1 \leq v \leq m} e_{s(\alpha_v)} kQe_j \to e_i kQe_j$$

is not injective, where f_{α_v} is the k-linear map given by the left multiplication by α_v. Since each path from j to i must go through some α_v, and $\sum_{p \in \mathcal{P}} c_p f_p = 0$, it follows that

$$\sum_{1 \leq v \leq m} \dim_k(e_{s(\alpha_v)} kQe_j) > \dim_k(e_i kQe_j).$$

This justifies the claim, that is, $P(j) \notin \mathrm{Mon}(Q, I, k)$. □

Now, let $\Lambda = A \otimes_k kQ$ be the path algebra of Q over A. Assume that Λ is of the form $\Lambda = kQ'/I'$, where Q' is a finite quiver and I' is an admissible ideal of kQ'. We emphasize that, in general,

$$\mathrm{Mon}(Q, A) \neq \mathrm{Mon}(Q', I', k).$$

In fact, $\mathcal{P}(\Lambda) \subseteq \mathrm{Mon}(Q, A)$ (Proposition 2.4); but generally $\mathcal{P}(\Lambda) \subseteq \mathrm{Mon}(Q', I', k)$ is not true, as Proposition 2.6 shows. This is the reason why we do not use the notation $\mathrm{Mon}(\Lambda)$.

3. Functorial finiteness of $\mathrm{Mon}(Q, A)$ in $\mathrm{Rep}(Q, A)$

The aim of this section is to prove the following.

Theorem 3.1. *Let Q be a finite acyclic quiver, and A a finite-dimensional algebra. Then $\mathrm{Mon}(Q, A)$ is functorially finite in $\mathrm{Rep}(Q, A)$ and $\mathrm{Mon}(Q, A)$ has Auslander–Reiten sequences.*

The idea of the proof given below is essentially due to Ringel and Schmidmeier [2008a] for the case of $Q = \bullet \to \bullet$. The same result for the case of

$$Q = \underset{n}{\bullet} \to \cdots \to \underset{1}{\bullet}$$

has been obtained in [Moore 2010; Zhang 2011].

3A. Let Q be a finite acyclic quiver. Remember we label the vertices of Q as $1, 2, \ldots, n$, such that if there is an arrow from j to i, $j > i$. So vertex 1 is a sink. Denote by $\mathcal{P}(\to i)$ the set of all the paths p with ending vertex $e(p) = i$ and $l(p) \geq 1$.

For $X \in \operatorname{Rep}(Q, A)$ and $i \in Q_0$, put K_i to be the kernel of the A-map

$$(X_\alpha)_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} : \bigoplus_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} X_{s(\alpha)} \to X_i.$$

Fix an injective envelope $\delta_i : K_i \hookrightarrow IK_i$ of K_i. Then there is an A-map

$$(\varphi_\alpha)_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} : \bigoplus_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} X_{s(\alpha)} \to IK_i$$

such that the diagram

$$\begin{tikzcd} K_i \arrow[r, hook] \arrow[d, hook, "\delta_i"'] & \bigoplus\limits_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} X_{s(\alpha)} \arrow[dl, "(\varphi_\alpha)_{\substack{\alpha\in Q_1\\ e(\alpha)=i}}"] \\ IK_i & \end{tikzcd} \tag{3-1}$$

commutes for each $i \in Q_0$. We construct a representation

$$\operatorname{rMon}(X) = (\operatorname{rMon}(X)_i, \operatorname{rMon}(X)_\alpha, i \in Q_0, \alpha \in Q_1) \in \operatorname{Rep}(Q, A)$$

as follows. For each $i \in Q_0$, define

$$\operatorname{rMon}(X)_i = X_i \oplus IK_i \oplus \bigoplus_{p\in\mathcal{P}(\to i)} IK_{s(p)}. \tag{3-2}$$

(Note that if i is a source, by definition $\operatorname{rMon}(X)_i = X_i$, and that if $p_1, \ldots, p_m$ are all the paths in $\mathcal{P}(\to i)$ with the same starting vertex j, the $\underbrace{IK_j \oplus \cdots \oplus IK_j}_{m}$ is a direct summand of $\bigoplus_{p\in\mathcal{P}(\to i)} IK_{s(p)}$.)

For each arrow $\alpha : j \to i$, define

$$\operatorname{rMon}(X)_\alpha : X_j \oplus IK_j \oplus \bigoplus_{p\in\mathcal{P}(\to j)} IK_{s(p)} \to X_i \oplus IK_i \oplus \bigoplus_{q\in\mathcal{P}(\to i)} IK_{s(q)}$$

to be the A-map given by

$$x_j + k_j + \sum_{p\in\mathcal{P}(\to j)} k_{s(p)} \mapsto X_\alpha(x_j) + \varphi_\alpha(x_j) + k_j + \sum_{p\in\mathcal{P}(\to j)} k_{s(\alpha p)}, \tag{3-3}$$

where $x_j \in X_j$, $k_j \in IK_j$, $k_{s(p)} \in IK_{s(p)}$. Note that $s(p) = s(\alpha p)$, and that $k_{s(\alpha p)}$ is just $k_{s(p)}$. Also note that at the right side of (3-3), k_j and $\sum_{p\in\mathcal{P}(\to j)} k_{s(\alpha p)}$ belong to different direct summands of $\bigoplus_{q\in\mathcal{P}(\to i)} IK_{s(q)}$.

Lemma 3.2. *For* $X \in \mathrm{Rep}(Q, A)$, *we have* $\mathrm{rMon}(X) \in \mathrm{Mon}(Q, A)$.

Proof. For each $i \in Q_0$, let $\alpha_1, \dots, \alpha_m$ be all the arrows ending at i. By definition we only need to prove that the A-map

$$(\mathrm{rMon}(X)_{\alpha_1}, \dots, \mathrm{rMon}(X)_{\alpha_m}) : \bigoplus_{1\le j\le m} \mathrm{rMon}(X)_{s(\alpha_j)} \to \mathrm{rMon}(X)_i$$

is injective. This is clear by (3-1)–(3-3). For completeness we include a justification.

Suppose $z_j = x_{s(\alpha_j)} + k_{s(\alpha_j)} + (\sum_{p\in\mathcal{P}(\to s(\alpha_j))} k_{s(p)}) \in \mathrm{rMon}(X)_{s(\alpha_j)}$, $j = 1, \dots, m$, and $\sum_{1\le j\le m} \mathrm{rMon}(X)_{\alpha_j}(z_j) = 0$. Then by (3-3) we have

$$0 = \sum_{1\le j\le m} X_{\alpha_j}(x_{s(\alpha_j)}) + \sum_{1\le j\le m} \varphi_{\alpha_j}(x_{s(\alpha_j)}) + \sum_{1\le j\le m} k_{s(\alpha_j)} + \sum_{1\le j\le m} \sum_{p\in\mathcal{P}(\to s(\alpha_j))} k_{s(\alpha_j p)}$$
$$\in X_i \oplus IK_i \oplus \bigoplus_{q\in\mathcal{P}(\to i)} IK_{s(q)}.$$

Thus

$$\sum_{1\le j\le m} X_{\alpha_j}(x_{s(\alpha_j)}) = 0, \qquad \sum_{1\le j\le m} \varphi_{\alpha_j}(x_{s(\alpha_j)}) = 0,$$

and $k_{s(\alpha_j)} = 0 = k_{s(\alpha_j p)}$ for all $j = 1, \dots, m$ and all $p \in \mathcal{P}(\to s(\alpha_j))$. Note that $\sum_{1\le j\le m} X_{\alpha_j}(x_{s(\alpha_j)}) = 0$ implies

$$\begin{pmatrix} x_{s(\alpha_1)} \\ \vdots \\ x_{s(\alpha_m)} \end{pmatrix} \in K_i.$$

By (3-1) we have

$$\delta_i \begin{pmatrix} x_{s(\alpha_1)} \\ \vdots \\ x_{s(\alpha_m)} \end{pmatrix} = \sum_{1\le j\le m} \varphi_{\alpha_j}(x_{s(\alpha_j)}) = 0.$$

Since δ_i is injective, we have $x_{s(\alpha_j)} = 0$ for $j = 1, \dots, m$. Thus $z_j = 0$ for $j = 1, \dots, m$. This completes the proof. $\square$

3B. Let $\mathcal{X}$ be a full subcategory of A-mod. Recall from [Auslander and Reiten 1991a] that *a right $\mathcal{X}$-approximation of M* is a morphism $f : X \to M$ with $X \in \mathcal{X}$ such that the induced homomorphism $\mathrm{Hom}_A(X', X) \to \mathrm{Hom}_A(X', M)$ is surjective for each $X' \in \mathcal{X}$. If every object M admits a right $\mathcal{X}$-approximation, $\mathcal{X}$ is called *a contravariantly finite subcategory in* A-mod. Dually one has the concept of *a covariantly finite subcategory in* A-mod. If $\mathcal{X}$ is both contravariantly and covariantly finite in A-mod, $\mathcal{X}$ is *a functorially finite subcategory in* A-mod.

Proposition 3.3. *Let Q be a finite acyclic quiver, and A a finite-dimensional algebra. Then* $\mathrm{Mon}(Q, A)$ *is contravariantly finite in* $\mathrm{Rep}(Q, A)$.

More precisely, let $X \in \mathrm{Rep}(Q, A)$, $f = (f_i, i \in Q_0) :$ $\mathrm{rMon}(X) \to X$, where $f_i :$ $\mathrm{rMon}(X)_i \to X_i$ is the canonical projection. Then f is a right $\mathrm{Mon}(Q, A)$-*approximation of X.*

Proof. We use induction to prove that f is a right $\mathrm{Mon}(Q, A)$-approximation of X. The assertion trivially holds if $|Q_0| = 1$. Suppose that the assertion holds for the quivers Q with $|Q_0| = n - 1$. Assume that $|Q_0| = n$ and that

$$g = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} : Y \to X$$

is a morphism in $\mathrm{Rep}(Q, A)$ with $Y \in \mathrm{Mon}(Q, A)$. We need to prove that there is a morphism

$$h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} : Y \to \mathrm{rMon}(X)$$

in $\mathrm{Rep}(Q, A)$ such that $g = fh$.

Let Q' be the quiver obtained from Q by deleting vertex 1, X' the representation in $\mathrm{Rep}(Q', A)$ obtained from X by deleting the branch X_1, and Y' the representation in $\mathrm{Mon}(Q', A)$ obtained from Y by deleting the branch Y_1. Then by definition $\mathrm{rMon}(X')$ is exactly the representation in $\mathrm{Mon}(Q', A)$ obtained from $\mathrm{rMon}(X)$ by deleting the branch $\mathrm{rMon}(X)_1$. Further,

$$\begin{pmatrix} f_2 \\ \vdots \\ f_n \end{pmatrix} : \mathrm{rMon}(X') \to X' \quad \text{and} \quad \begin{pmatrix} g_2 \\ \vdots \\ g_n \end{pmatrix} : Y' \to X'$$

are morphisms in $\mathrm{Rep}(Q', A)$. By the inductive hypothesis there is a morphism

$$\begin{pmatrix} h_2 \\ \vdots \\ h_n \end{pmatrix} : Y' \to \mathrm{rMon}(X')$$

in $\mathrm{Rep}(Q', A)$, such that

$$\begin{pmatrix} g_2 \\ \vdots \\ g_n \end{pmatrix} = \begin{pmatrix} f_2 \\ \vdots \\ f_n \end{pmatrix} \begin{pmatrix} h_2 \\ \vdots \\ h_n \end{pmatrix}.$$

Let $\alpha_1, \ldots, \alpha_m$ be all the arrows ending at 1. Since

$$(Y_{\alpha_1}, \ldots, Y_{\alpha_m}) : \bigoplus_{1 \le j \le m} Y_{s(\alpha_j)} \to Y_1$$

is an injective A-map and $IK_1 \oplus \left(\bigoplus_{p \in \mathcal{P}(\to 1)} IK_{s(p)}\right)$ is an injective A-module, it follows that there is a map

$$\eta : Y_1 \to IK_1 \oplus \bigoplus_{p \in \mathcal{P}(\to 1)} IK_{s(p)}$$

such that the diagram

$$\begin{tikzcd} \bigoplus\limits_{1 \le j \le m} Y_{s(\alpha_j)} \arrow[r, "(Y_{\alpha_1}, \ldots, Y_{\alpha_m})"] \arrow[d, "\tilde{h}"'] & Y_1 \arrow[d, "\eta"] \\ \bigoplus\limits_{1 \le j \le m} \mathrm{rMon}(X)_{s(\alpha_j)} \arrow[r, "(B_1, \ldots, B_m)"] & IK_1 \oplus \bigoplus\limits_{p \in \mathcal{P}(\to 1)} IK_{s(p)} \end{tikzcd}$$

commutes, where $\tilde{h} = \mathrm{diag}(h_{s(\alpha_1)}, \ldots, h_{s(\alpha_m)})$ and, for each $j = 1, \ldots, m$,

$$B_j : \mathrm{rMon}(X)_{s(\alpha_j)} \to IK_1 \oplus \bigoplus_{p \in \mathcal{P}(\to 1)} IK_{s(p)}$$

is the A-map given by

$$x_{s(\alpha_j)} + k_{s(\alpha_j)} + \sum_{p \in \mathcal{P}(\to s(\alpha_j))} k_{s(p)} \mapsto \varphi_{\alpha_j}(x_{s(\alpha_j)}) + k_{s(\alpha_j)} + \sum_{p \in \mathcal{P}(\to s(\alpha_j))} k_{s(\alpha_j p)}$$

for

$$\begin{aligned} x_{s(\alpha_j)} + k_{s(\alpha_j)} + \sum_{p \in \mathcal{P}(\to s(\alpha_j))} k_{s(p)} \in \mathrm{rMon}(X)_{s(\alpha_j)} \\ = X_{s(\alpha_j)} \oplus IK_{s(\alpha_j)} \oplus \bigoplus_{p \in \mathcal{P}(\to s(\alpha_j))} IK_{s(p)}. \end{aligned}$$

For $y \in Y_{s(\alpha_j)}$, suppose

$$h_{s(\alpha_j)}(y) = x_{s(\alpha_j)} + k_{s(\alpha_j)} + \sum_{p \in \mathcal{P}(\to s(\alpha_j))} k_{s(p)} \in \mathrm{rMon}(X)_{s(\alpha_j)}.$$

Then we have

$$\begin{aligned}
\mathrm{rMon}(X)_{\alpha_j} h_{s(\alpha_j)}(y) &= X_{\alpha_j}(x_{s(\alpha_j)}) + \varphi_{\alpha_j}(x_{s(\alpha_j)}) + k_{s(\alpha_j)} + \sum_{p\in\mathcal{P}(\to s(\alpha_j))} k_{s(\alpha_j p)} \\
&= X_{\alpha_j}(x_{s(\alpha_j)}) + B_j h_{s(\alpha_j)}(y) \\
&= X_{\alpha_j}(f_{s(\alpha_j)} h_{s(\alpha_j)}(y)) + B_j h_{s(\alpha_j)}(y) \\
&= X_{\alpha_j} g_{s(\alpha_j)}(y) + B_j h_{s(\alpha_j)}(y) \\
&= g_1 Y_{\alpha_j}(y) + \eta Y_{\alpha_j}(y),
\end{aligned}$$

where the last equality uses the fact that $g: Y \to X$ is a morphism in $\mathrm{Rep}(Q, A)$.

Now we define $h_1 : Y_1 \to \mathrm{rMon}(X)_1$ to be the A-map given by

$$h_1(y) = g_1(y) + \eta(y)$$

for each $y \in Y_1$. From the computation above we have $\mathrm{rMon}(X)_{\alpha_j} h_{s(\alpha_j)} = h_1 Y_{\alpha_j}$ for $j = 1, \ldots, m$. It follows that

$$h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} : Y \to \mathrm{rMon}(X)$$

is a morphism in $\mathrm{Rep}(Q, A)$. Since $f_1 : \mathrm{rMon}(X)_1 \to X_1$ is the canonical projection, we have $f_1\eta = 0$ and $f_1 g_1 = g_1$, and hence $fh = g$. This completes the proof. □

3C. ***Proof of Theorem 3.1.*** By Corollary 2.5 and Proposition 3.3 we know that $\mathrm{Mon}(Q, A)$ is a resolving, contravariantly finite subcategory of $\mathrm{Rep}(Q, A)$, and hence $\mathrm{Mon}(Q, A)$ is functorially finite in $\mathrm{Rep}(Q, A)$; see [Krause and Solberg 2003, Corollary 2.6(i)]. It follows that $\mathrm{Mon}(Q, A)$ has Auslander–Reiten sequences, by [Auslander and Smalø 1981, Theorem 2.4]. □

4. Gorenstein-projective modules over the upper triangular matrix algebras

4A. Let A and B be rings, M an A-B-bimodule, and $\Lambda = \left(\begin{smallmatrix} A & M \\ 0 & B \end{smallmatrix}\right)$ the upper triangular matrix ring, where the addition and multiplication are given by the ones of matrices. We assume that Λ is an Artin algebra [Auslander et al. 1995, p. 72], and consider finitely generated Λ-modules. A Λ-module can be identified with a triple $\left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)_\phi$, or simply $\left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)$ if ϕ is clear, where $X \in A\text{-mod}$, $Y \in B\text{-mod}$, and $\phi : M \otimes_B Y \to X$ is an A-map. A Λ-map $\left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right)_\phi \to \left(\begin{smallmatrix} X' \\ Y' \end{smallmatrix}\right)_{\phi'}$ can be identified with a pair $\left(\begin{smallmatrix} f \\ g \end{smallmatrix}\right)$, where $f \in \mathrm{Hom}_A(X, X')$, $g \in \mathrm{Hom}_B(Y, Y')$ are such that the diagram

$$\begin{tikzcd}
M \otimes_B Y \arrow[r, "\phi"] \arrow[d, "\mathrm{id}\otimes g"'] & X \arrow[d, "f"] \\
M \otimes_B Y' \arrow[r, "\phi'"] & X'
\end{tikzcd}$$

commutes. A sequence of Λ-maps

$$0 \to \binom{X_1}{Y_1}_{\phi_1} \xrightarrow{\binom{f_1}{g_1}} \binom{X_2}{Y_2}_{\phi_2} \xrightarrow{\binom{f_2}{g_2}} \binom{X_3}{Y_3}_{\phi_3} \to 0$$

is exact if and only if

$$0 \longrightarrow X_1 \xrightarrow{f_1} X_2 \xrightarrow{f_2} X_3 \longrightarrow 0$$

is an exact sequence of A-maps, and

$$0 \longrightarrow Y_1 \xrightarrow{g_1} Y_2 \xrightarrow{g_2} Y_3 \longrightarrow 0$$

is an exact sequence of B-maps. The indecomposable projective Λ-modules are exactly

$$\binom{P}{0} \quad \text{and} \quad \binom{M \otimes_B Q}{Q}_{\mathrm{id}},$$

where P runs over indecomposable projective A-modules and Q runs over indecomposable projective B-modules.

Note that an algebra Λ is of the form above if and only if there is an idempotent decomposition $1 = e + f$ such that $f\Lambda e = 0$; and in this case

$$\Lambda = \begin{pmatrix} e\Lambda e & e\Lambda f \\ 0 & f\Lambda f \end{pmatrix}.$$

4B. The following result describes the Gorenstein-projective Λ-modules, if ${}_AM$ and M_B are projective modules.

Theorem 4.1. *Let*

$$\Lambda = \begin{pmatrix} A & M \\ 0 & B \end{pmatrix}$$

be an Artin algebra, M an A-B-bimodule such that ${}_AM$ and M_B are projective modules. Then

$$\binom{X}{Y}_{\phi} \in \mathcal{GP}(\Lambda)$$

if and only if $\phi : M \otimes_B Y \to X$ is injective, $\operatorname{Coker}\phi \in \mathcal{GP}(A)$, *and $Y \in \mathcal{GP}(B)$. In this case, $X \in \mathcal{GP}(A)$ if and only if $M \otimes_B Y \in \mathcal{GP}(A)$.*

Note that here Λ is not assumed to be Gorenstein: this will be important to the main result in the next section. The same result under the assumption that Λ is Gorenstein can be found in [Xiong and Zhang 2012, Corollary 3.3] (however, the proof there cannot be generalized to the non-Gorenstein case). The same corollary implies that, if Λ is Gorenstein in Theorem 4.1, $\binom{X}{Y}_{\phi} \in \mathcal{GP}(\Lambda)$ implies $X \in \mathcal{GP}(A)$.

Proof of Theorem 4.1. The last assertion is easy: it follows from the exact sequence

$$0 \longrightarrow M \otimes_B Y \xrightarrow{\phi} X \longrightarrow \operatorname{Coker} \phi \longrightarrow 0$$

and the fact that $\mathcal{GP}(A)$ is closed under extensions and the kernels of epimorphisms; see, for example, [Holm 2004].

We next prove the "if" part of the first equivalence in the theorem. We assume that $\phi : M \otimes_B Y \to X$ is injective, $\operatorname{Coker} \phi \in \mathcal{GP}(A)$, and $Y \in \mathcal{GP}(B)$. Then we have a complete B-projective resolution

$$Q^\bullet = \quad \cdots \longrightarrow Q^{-1} \longrightarrow Q^0 \xrightarrow{d'^0} Q^1 \longrightarrow \cdots \tag{4-1}$$

with $Y = \operatorname{Ker} d'^0$, and a complete A-projective resolution

$$P^\bullet = \quad \cdots \longrightarrow P^{-1} \longrightarrow P^0 \xrightarrow{d^0} P^1 \longrightarrow \cdots \tag{4-2}$$

with $\operatorname{Coker} \phi = \operatorname{Ker} d^0$. Since M_B is projective, we get the following exact sequences of A-modules:

$$\begin{gathered} 0 \to M \otimes_B Y \to M \otimes_B Q^0 \to M \otimes_B Q^1 \to \cdots, \\ 0 \to \operatorname{Coker} \phi \to P^0 \to P^1 \to \cdots. \end{gathered}$$

Since ${}_A M$ is projective, $M \otimes_B Q^i$ is a projective A-module for each $i \geq 0$. Since $\operatorname{Ext}^1_A(\operatorname{Coker} \phi, M \otimes_B Q^0) = 0$, it follows from the exact sequence

$$0 \to M \otimes_B Y \xrightarrow{\phi} X \to \operatorname{Coker} \phi \to 0$$

that the map $M \otimes_B Y \to M \otimes_B Q^0$ factors through ϕ. So, by a version of the horseshoe lemma, we see that there is an exact sequence of A-modules

$$0 \to X \to P^0 \oplus (M \otimes_B Q^0) \xrightarrow{\partial^0} P^1 \oplus (M \otimes_B Q^1) \to \cdots \tag{4-3}$$

with

$$\partial^i = \begin{pmatrix} d^i & 0 \\ \sigma^i & \mathrm{id} \otimes_B d'^i \end{pmatrix}, \quad \sigma^i : P^i \to M \otimes_B Q^i$$

for all $i \in \mathbb{Z}$, such that the diagram

$$\begin{tikzcd} 0 \arrow[r] & M \otimes_B Y \arrow[r] \arrow[d, "\phi"] & M \otimes_B Q^0 \arrow[r, "\mathrm{id} \otimes_B d'^0"] \arrow[d, "\binom{0}{\mathrm{id}}"'] & M \otimes_B Q^1 \arrow[r] \arrow[d, "\binom{0}{\mathrm{id}}"] & \cdots \\ 0 \arrow[r] & X \arrow[r] & P^0 \oplus (M \otimes_B Q^0) \arrow[r, "\partial^0"] & P^1 \oplus (M \otimes_B Q^1) \arrow[r] & \cdots \end{tikzcd} \tag{4-4}$$

commutes. By the same argument we get the following commutative diagram with exact rows:

$$
\begin{tikzcd}
\cdots \arrow[r] & M\otimes_B Q^{-2} \arrow[r, "\mathrm{id}\otimes_B d'^{-2}"] \arrow[d, "\binom{0}{\mathrm{id}}"] & M\otimes_B Q^{-1} \arrow[r] \arrow[d, "\binom{0}{\mathrm{id}}"'] & M\otimes_B Y \arrow[r] \arrow[d, "\phi"] & 0 \\
\cdots \arrow[r] & P^{-2}\oplus(M\otimes_B Q^{-2}) \arrow[r, "\partial^{-2}"] & P^{-1}\oplus(M\otimes_B Q^{-1}) \arrow[r] & X \arrow[r] & 0.
\end{tikzcd} \tag{4-5}
$$

Putting (4-4) and (4-5) together, we get the exact sequence of projective Λ-modules

$$
L^\bullet = \quad \cdots \longrightarrow \begin{pmatrix} P^{-1}\oplus(M\otimes_B Q^{-1}) \\ Q^{-1}\end{pmatrix} \longrightarrow \begin{pmatrix} P^{0}\oplus(M\otimes_B Q^{0}) \\ Q^{0}\end{pmatrix}_{\binom{0}{\mathrm{id}}} \xrightarrow{\binom{\partial^0}{d'^0}} \begin{pmatrix} P^{1}\oplus(M\otimes_B Q^{1}) \\ Q^{1}\end{pmatrix} \longrightarrow \cdots \tag{4-6}
$$

with $\operatorname{Ker}\binom{\partial^0}{d'^0} = \binom{X}{Y}_\phi$.

For each projective A-module P, $\operatorname{Hom}_\Lambda\big(L^\bullet, \binom{P}{0}\big) \cong \operatorname{Hom}_A(P^\bullet, P)$ is exact, since $P^\bullet$ is a complete projective resolution. For each projective B-module Q, since $Q^\bullet$ is a complete projective resolution, $\operatorname{Hom}_B(Q^\bullet, Q)$ is exact. Since $M\otimes_B Q$ is projective, $\operatorname{Hom}_A(P^\bullet, M\otimes_B Q)$ is exact. Note that

$$\operatorname{Hom}_\Lambda\Big(L^\bullet, \binom{M\otimes_B Q}{Q}\Big) \cong \operatorname{Hom}_A(P^\bullet, M\otimes_B Q)\oplus \operatorname{Hom}_B(Q^\bullet, Q);$$

here the direct sum only means that each term of the complex at the left side is a direct sum of terms of complexes at the right side, that is, it does not mean a direct sum of complexes; in fact, the complex at the right side has differentials

$$\begin{pmatrix} \operatorname{Hom}_A(d^i, M\otimes_B Q) & \operatorname{Hom}_A(\sigma^i, M\otimes_B Q) \\ 0 & \operatorname{Hom}_B(d'^i, Q)\end{pmatrix}.$$

By the canonical exact sequence of complexes

$$0\to \operatorname{Hom}_A(P^\bullet, M\otimes_B Q) \xrightarrow{\binom{\mathrm{id}}{0}} \operatorname{Hom}_\Lambda\Big(L^\bullet, \binom{M\otimes_B Q}{Q}\Big) \xrightarrow{(0\ \mathrm{id})} \operatorname{Hom}_B(Q^\bullet, Q)\to 0,$$

we know that

$$\operatorname{Hom}_\Lambda\Big(L^\bullet, \binom{M\otimes_B Q}{Q}\Big)$$

is also exact. We conclude that $L^\bullet$ is a complete Λ-projective resolution, and hence $\binom{X}{Y}_\phi$ is a Gorenstein-projective Λ-module.

Conversely, assume that $\binom{X}{Y}_\phi \in \mathcal{GP}(\Lambda)$. Then there is a complete Λ-projective resolution (4-6) with

$$\operatorname{Ker}\binom{\partial^0}{d'^0} = \binom{X}{Y}_\phi.$$

Then we get an exact sequence (4-1) of projective B-modules with $\operatorname{Ker} d'^0 = Y$, and the exact sequence

$$V^\bullet = \quad \cdots \to P^{-1} \oplus (M \otimes_B Q^{-1}) \to P^0 \oplus (M \otimes_B Q^0) \xrightarrow{\partial^0} P^1 \oplus (M \otimes_B Q^1) \to \cdots \tag{4-7}$$

of projective A-modules with $\operatorname{Ker} \partial^0 = X$. Since M_B is projective, it follows that $M \otimes_B Q^\bullet$ is exact. Since $\binom{\partial^i}{d'^i}$ is a Λ-map, by (4-6) we know that ∂^i is of the form

$$\partial^i = \begin{pmatrix} d^i & 0 \\ \sigma^i & \mathrm{id} \otimes_B d'^i \end{pmatrix},$$

where $\sigma^i : P^i \to M \otimes_B Q^i$ for all $i \in \mathbb{Z}$, and

$$P^\bullet = \quad \cdots \longrightarrow P^{-1} \longrightarrow P^0 \xrightarrow{d^0} P^1 \longrightarrow \cdots$$

is a complex. By the canonical exact sequence of complexes

$$0 \longrightarrow M \otimes_B Q^\bullet \xrightarrow{\binom{\mathrm{id}}{0}} V^\bullet \xrightarrow{(0\ \mathrm{id})} \operatorname{Hom}_B(Q^\bullet, Q) P^\bullet \longrightarrow 0,$$

we see that $P^\bullet$ is also exact.

From (4-6) we have the following commutative diagram with exact rows and columns:

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & M \otimes_B Y \arrow[r] \arrow[d, "\phi"] & M \otimes_B Q^0 \arrow[r] \arrow[d, "\binom{0}{\mathrm{id}}"] & M \otimes_B Q^1 \arrow[r] \arrow[d, "\binom{0}{\mathrm{id}}"] & \cdots \\
0 \arrow[r] & X \arrow[r] \arrow[d] & P^0 \oplus (M \otimes_B Q^0) \arrow[r] \arrow[d, "(\mathrm{id},0)"] & P^1 \oplus (M \otimes_B Q_1) \arrow[r] \arrow[d, "(\mathrm{id},0)"] & \cdots \\
0 \arrow[r] & \operatorname{Coker} \phi \arrow[r] \arrow[d] & P^0 \arrow[r] \arrow[d, "d^0"] & P^1 \arrow[r] \arrow[d] & \cdots \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

Thus $\phi : M \otimes_B Y \longrightarrow X$ is injective and $\operatorname{Ker} d^0 \cong \operatorname{Coker} \phi$. For each projective A-module P, since

$$\operatorname{Hom}_\Lambda\bigl(L^\bullet, \tbinom{P}{0}\bigr) \cong \operatorname{Hom}_A(P^\bullet, P)$$

and $L^\bullet$ is a complete projective resolution, it follows that $P^\bullet$ is a complete projective resolution, and hence $\operatorname{Coker} \phi$ is a Gorenstein-projective A-module.

For each projective B-module Q, since $P^\bullet$ is a complete projective resolution, it follows that $\mathrm{Hom}_A(P^\bullet, M \otimes_B Q)$ is exact. Since $L^\bullet$ is a complete projective resolution, it follows that

$$\mathrm{Hom}_\Lambda\left(L^\bullet, \begin{pmatrix} M \otimes_B Q \\ Q \end{pmatrix}\right) \cong \mathrm{Hom}_A(P^\bullet, M \otimes_B Q) \oplus \mathrm{Hom}_B(Q^\bullet, Q)$$

is exact (again, the direct sum does not mean a direct sum of complexes). By the same argument we know that $\mathrm{Hom}_B(Q^\bullet, Q)$ is exact. It follows that Y is a Gorenstein-projective B-module. □

5. Main result

5A. The aim of this section is to prove the following characterization of Gorenstein-projective Λ-modules, where Λ is the path algebra of a finite acyclic quiver over a finite-dimensional algebra. We emphasize that here Λ is not assumed to be Gorenstein.

Theorem 5.1. *Let Q be a finite acyclic quiver, and A a finite-dimensional algebra over a field k. Let $\Lambda = A \otimes_k kQ$, and $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$ be a Λ-module. Then $X \in \mathcal{GP}(\Lambda)$ if and only if $X \in \mathrm{Mon}(Q, A)$ and X satisfies this condition*:

(G) *for each $i \in Q_0$, X_i and the quotient* $X_i \Big/ \bigoplus_{\substack{\alpha \in Q_1 \\ e(\alpha)=i}} \mathrm{Im}\, X_\alpha$ *lie in* $\mathcal{GP}(A)$.

Example 5.2. (i) Taking

$$Q = \underset{n}{\bullet} \to \cdots \to \underset{1}{\bullet}$$

in Theorem 5.1, we get that a $T_n(A)$-module $X = (X_i, \phi_i)$ is Gorenstein-projective if and only if each ϕ_i is injective and that each X_i is a Gorenstein-projective A-module and each $\mathrm{Coker}\, \phi_i$ is a Gorenstein-projective A-module. Under the assumption that A is Gorenstein, this result has been obtained in [Zhang 2011, Corollary 4.1]; the case for $n = 2$ was treated in [Li and Zhang 2010, Theorem 1.1(i)]; see also [Iyama et al. 2011, Proposition 3.6(i)].

(ii) Let Λ be the k-algebra given by quiver

$$\begin{tikzcd} \underset{3}{\bullet} \arrow[loop above, "\lambda_3"] \arrow[r, "\beta"] & \underset{1}{\bullet} \arrow[loop above, "\lambda_1"] & \underset{2}{\bullet} \arrow[loop above, "\lambda_2"] \arrow[l, "\alpha"'] \end{tikzcd}$$

with relations λ_1^2, λ_2^2, λ_3^2, $\alpha\lambda_2 - \lambda_1\alpha$, $\beta\lambda_3 - \lambda_1\beta$. Then

$$\Lambda = A \otimes_k kQ = \begin{pmatrix} A & A & A \\ 0 & A & 0 \\ 0 & 0 & A \end{pmatrix},$$

where Q is the quiver

$$\underset{3}{\bullet} \longrightarrow \underset{1}{\bullet} \longleftarrow \underset{2}{\bullet},$$

and $A = k[x]/\langle x^2\rangle$. Let k be the simple A-module, and $\sigma : k \hookrightarrow A$ the inclusion. By Theorem 5.1, the following Λ-modules lie in $\mathrm{GP}(\Lambda)$:

$$\begin{aligned}
&(X_1 = A,\ X_2 = 0,\ X_3 = 0,\ X_\alpha = 0 = X_\beta),\\
&(X_1 = A,\ X_2 = A,\ X_3 = 0,\ X_\alpha = \mathrm{id},\ X_\beta = 0),\\
&(X_1 = A,\ X_2 = 0,\ X_3 = A,\ X_\alpha = 0,\ X_\beta = \mathrm{id}),\\
&(X_1 = k,\ X_2 = 0,\ X_3 = 0,\ X_\alpha = 0 = X_\beta),\\
&(X_1 = k,\ X_2 = k,\ X_3 = 0,\ X_\alpha = \mathrm{id},\ X_\beta = 0),\\
&(X_1 = k,\ X_2 = 0,\ X_3 = k,\ X_\alpha = 0,\ X_\beta = \mathrm{id}),\\
&(X_1 = A,\ X_2 = k,\ X_3 = 0,\ X_\alpha = \sigma,\ X_\beta = 0),\\
&(X_1 = A,\ X_2 = 0,\ X_3 = k,\ X_\alpha = 0,\ X_\beta = \sigma),\\
&\big(X_1 = A \oplus k,\ X_2 = k,\ X_3 = k,\ X_\alpha = \binom{0}{\mathrm{id}},\ X_\beta = \binom{\sigma}{\mathrm{id}}\big).
\end{aligned}$$

In fact this is the complete list of the pairwise nonisomorphic indecomposable Gorenstein-projective Λ-modules. Also by Theorem 5.1,

$$(Y_1 = A, Y_2 = k, Y_3 = k, Y_\alpha = \sigma = Y_\beta) \notin \mathcal{GP}(\Lambda).$$

For a description of all the pairwise nonisomorphic indecomposable Gorenstein-projective Λ-modules see [Ringel and Zhang 2011], where Λ is the path algebra of an arbitrary acyclic quiver over $A = k[x]/\langle x^2\rangle$.

5B. We prove Theorem 5.1 by using Theorem 4.1 and induction on $|Q_0|$.

Remember we label Q_0 as $1, \dots, n$, in such a way that $j > i$ if $\alpha : j \to i$ is in Q_1. Thus n is a source of Q. Denote by Q' the quiver obtained from Q by deleting vertex n, and $\Lambda' = A \otimes_k kQ'$. Let $P(n)$ be the indecomposable projective (left) kQ-module at vertex n. Put $P = A \otimes_k \operatorname{rad} P(n)$. Clearly P is a Λ'-A-bimodule and $\Lambda = \begin{pmatrix} \Lambda' & P \\ 0 & A \end{pmatrix}$; compare (2-2).

Since kQ is hereditary, $\operatorname{rad} P(n)$ is a projective kQ'-module, and hence $P = A \otimes_k \operatorname{rad} P(n)$ is a (left) projective Λ'-module, and a (right) projective A-module (since as a right A-module, P is a direct sum of copies of A_A). So we can apply Theorem 4.1. For this, we write a Λ-module $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$ as $X = \binom{X'}{X_n}_\phi$, where $X' = (X_i, X_\alpha, i \in Q'_0, \alpha \in Q'_1)$ is a Λ'-module, and

$$\phi : P \otimes_A X_n \to X'$$

is a Λ'-map. The explicit expression of ϕ is given in the proof of Lemma 5.4. We keep all these notations of Q', Λ', $P(n)$, P, X' and ϕ throughout this section.

5C. By a direct translation from Theorem 4.1 in this special case, we have:

Lemma 5.3. *Let* $X = \binom{X'}{X_n}_\phi$ *be a* Λ*-module. Then* $X \in \mathcal{GP}(\Lambda)$ *if and only if* X *satisfies the following conditions*:

(i) $X_n \in \mathcal{GP}(A)$.

(ii) $\phi : P \otimes_A X_n \to X'$ *is injective.*

(iii) $\operatorname{Coker} \phi \in \mathcal{GP}(\Lambda')$.

For each $i \in Q_0'$, put $\mathcal{A}(n \to i)$ to be the set of arrows from n to i; and $\mathcal{P}(n \to i)$ the set of paths from n to i. For an integer $m \geq 0$ and a module M, let M^m denote the direct sum of m copies of M.

Lemma 5.4. *Let* $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$ *be a* Λ*-module. If* X_β *is injective for each* $\beta \in Q_1'$, $\phi : P \otimes_A X_n \to X'$ *is injective if and only if* X_α *is injective for all* $\alpha \in Q_1$, *and* $\sum_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p = \bigoplus_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p$ *for all* Q_0'.

Proof. For $i \in Q_0'$, set $m_i = |\mathcal{P}(n \to i)|$. As a kQ'-module, $\operatorname{rad} P(n)$ can be written as

$$\begin{pmatrix} k^{m_1} \\ \vdots \\ k^{m_{n-1}} \end{pmatrix}$$

(see (2-1) and Section 5B), hence we have isomorphisms of Λ'-modules

$$P \otimes_A X_n \cong (\operatorname{rad} P(n) \otimes_k A) \otimes_A X_n \cong \operatorname{rad} P(n) \otimes_k X_n \cong \begin{pmatrix} X_n^{m_1} \\ \vdots \\ X_n^{m_{n-1}} \end{pmatrix}.$$

Let $\mathcal{P}(n \to i) = \{p_1, \ldots, p_{m_i}\}$. Then ϕ is of the form

$$\begin{pmatrix} \phi_1 \\ \vdots \\ \phi_{n-1} \end{pmatrix} : P \otimes_A X_n \cong \begin{pmatrix} X_n^{m_1} \\ \vdots \\ X_n^{m_{n-1}} \end{pmatrix} \to \begin{pmatrix} X_1 \\ \vdots \\ X_{n-1} \end{pmatrix},$$

where $\phi_i = (X_{p_1}, \ldots, X_{p_{m_i}}) : X_n^{m_i} \to X_i$ (for the meaning of X_p see Section 2C). So ϕ is injective if and only if ϕ_i is injective for each $i \in Q_0'$, and if and only if

$$\sum_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p = \bigoplus_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p \quad \text{and} \quad X_p \text{ is injective for all } p \in \mathcal{P}(n \to i).$$

From this and the assumption the assertion follows. □

Lemma 5.5. *Let* $X = \binom{X'}{X_n}_\phi$ *be a monic* Λ*-module.*

(1) *For each* $i \in Q_0'$ *there holds* $\sum_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p = \bigoplus_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p$.

(2) $\phi : P \otimes_A X_n \to X'$ *is injective.*

(3) $\operatorname{Coker}\phi = \big(X_i/\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p,\ \widetilde{X}_\alpha,\ i\in Q'_0,\ \alpha\in Q'_1\big)$, *where, for each* $\alpha: j\to i$ *in* Q'_1,

$$\widetilde{X}_\alpha : X_j\Big/\bigoplus_{q\in\mathcal{P}(n\to j)} \operatorname{Im} X_q \to X_i\Big/\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p$$

is the A*-map induced by* X_α.

Proof. By Lemma 5.4 and its proof, it suffices to prove (1). For each $i\in Q'_0$, set $l_i=0$ if $\mathcal{P}(n\to i)$ is empty, and $l_i=\max\{l(p)\mid p\in\mathcal{P}(n\to i)\}$ otherwise, where $l(p)$ is the length of p. We use induction on l_i. If $l_i=0$, (1) trivially holds. Suppose $l_i\geqslant 1$. Let $\sum_{p\in\mathcal{P}(n\to i)} X_p(x_{n,p})=0$ for $x_{n,p}\in X_n$. Since

$$\sum_{p\in\mathcal{P}(n\to i)-\mathcal{A}(n\to i)} \operatorname{Im} X_p = \sum_{\substack{\alpha\in Q'_1\\ e(\alpha)=i}} X_\alpha\Big(\sum_{q\in\mathcal{P}(n\to s(\alpha))} \operatorname{Im} X_q\Big),$$

we have

$$\begin{aligned}0=\sum_{p\in\mathcal{P}(n\to i)} X_p(x_{n,p}) &= \sum_{\alpha\in\mathcal{A}(n\to i)} X_\alpha(x_{n,\alpha}) + \sum_{p\in\mathcal{P}(n\to i)-\mathcal{A}(n\to i)} X_p(x_{n,p})\\ &= \sum_{\alpha\in\mathcal{A}(n\to i)} X_\alpha(x_{n,\alpha}) + \sum_{\substack{\beta\in Q'_1\\ e(\beta)=i}} X_\beta\Big(\sum_{q\in\mathcal{P}(n\to s(\beta))} X_q(x_{n,\beta q})\Big).\end{aligned}$$

By (m2) in Definition 2.2 we know that $X_\alpha(x_{n,\alpha})=0$ for $\alpha\in\mathcal{A}(n\to i)$, and

$$X_\beta\Big(\sum_{q\in\mathcal{P}(n\to s(\beta))} X_q(x_{n,\beta q})\Big)=0$$

for $\beta\in Q'_1$ with $e(\beta)=i$. So $\sum_{q\in\mathcal{P}(n\to s(\beta))} X_q(x_{n,\beta q})=0$ by condition (m1) in Definition 2.2. Since $l_{s(\beta)}<l_i$ for each $\beta\in Q'_1$ with $e(\beta)=i$, it follows from the inductive hypothesis that $X_q(x_{n,\beta q})=0$ for $\beta\in Q'_1$, $e(\beta)=i$, and $q\in\mathcal{P}(n\to s(\beta))$. This proves (1) and the lemma. □

Lemma 5.6. *Let* $X=\binom{X'}{X_n}_\phi$ *be a monic* Λ*-module. Then* $\operatorname{Coker}\phi$ *is a monic* Λ'*-module.*

Proof. We need to prove that, for each $i\in Q'_0$, the Λ'-map

$$(\widetilde{X}_\alpha)_{\substack{\alpha\in Q'_1\\ e(\alpha)=i}} : \bigoplus_{\substack{\alpha\in Q'_1\\ e(\alpha)=i}} \Big(X_{s(\alpha)}\Big/\bigoplus_{q\in\mathcal{P}(n\to s(\alpha))} \operatorname{Im} X_q\Big) \to X_i\Big/\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p$$

is injective. For this, assume that

$$\sum_{\substack{\alpha\in Q'_1\\ e(\alpha)=i}} \widetilde{X}_\alpha(\overline{x_{s(\alpha),\alpha}})=0,$$

where $\overline{x_{s(\alpha),\alpha}}$ is the image of $x_{s(\alpha),\alpha} \in X_{s(\alpha)}$ in $X_{s(\alpha)}/\bigoplus_{q\in\mathcal{P}(n\to s(\alpha))} \operatorname{Im} X_q$. Then

$$\sum_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha}) \in \bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p.$$

So there are $x_{n,p} \in X_n$ such that

$$\sum_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha}) = \sum_{p\in\mathcal{P}(n\to i)} X_p(x_{n,p}).$$

Thus

$$\begin{aligned} 0 &= \sum_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha}) - \sum_{p\in\mathcal{P}(n\to i)} X_p(x_{n,p}) \\ &= \sum_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha}) - \sum_{\beta\in\mathcal{A}(n\to i)} X_\beta(x_{n,\beta}) - \sum_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} X_\alpha\Big(\sum_{q\in\mathcal{P}(n\to s(\alpha))} X_q(x_{n,\alpha q})\Big) \\ &= \sum_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha} - \sum_{q\in\mathcal{P}(n\to s(\alpha))} X_q(x_{n,\alpha q})) - \sum_{\beta\in\mathcal{A}(n\to i)} X_\beta(x_{n,\beta}). \end{aligned}$$

Using the assumption on X, we get

$$x_{s(\alpha),\alpha} = \sum_{q\in\mathcal{P}(n\to s(\alpha))} X_q(x_{n,\alpha q}),$$

that is, $\overline{x_{s(\alpha),\alpha}} = 0$. □

Lemma 5.7. *Let $X = \binom{X'}{X_n}_\phi$ be a monic Λ-module satisfying (G). Then*

$$\Big(X_i\Big/\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p\Big)\Big/\Big(\bigoplus_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} \operatorname{Im} \widetilde{X}_\alpha\Big)$$

is a Gorenstein-projective A-module for all $i \in Q_0'$.

Proof. Since

$$\bigoplus_{p\in\mathcal{P}(n\to i)-\mathcal{A}(n\to i)} \operatorname{Im} X_p \subseteq \sum_{\substack{\beta\in Q_1 \\ e(\beta)=i}} \operatorname{Im} X_\beta,$$

it follows that

$$\sum_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} \operatorname{Im} \widetilde{X}_\alpha = \Big(\sum_{\substack{\alpha\in Q_1' \\ e(\alpha)=i}} \operatorname{Im} X_\alpha + \bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p\Big)\Big/\Big(\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p\Big) \tag{5-1}$$

$$= \Big(\sum_{\substack{\beta\in Q_1\\ e(\beta)=i}} \operatorname{Im} X_\beta + \bigoplus_{p\in\mathcal{P}(n\to i)-\mathcal{A}(n\to i)} \operatorname{Im} X_p \Big) \Big/ \Big(\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p \Big)$$
$$= \Big(\sum_{\substack{\beta\in Q_1\\ e(\beta)=i}} \operatorname{Im} X_\beta \Big) \Big/ \Big(\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p \Big)$$
$$= \Big(\bigoplus_{\substack{\beta\in Q_1\\ e(\beta)=i}} \operatorname{Im} X_\beta \Big) \Big/ \Big(\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p \Big)$$

(the last equality following by (m2) in Definition 2.2). Hence the desired quotient is $X_i/\bigoplus_{\substack{\beta\in Q_1\\ e(\beta)=i}} \operatorname{Im} X_\beta$, which is Gorenstein-projective by (G). ☐

Lemma 5.8. *Let $X = \binom{X'}{X_n}_\phi$ be a monic Λ-module satisfying (G). Then*

$$X_i \Big/ \bigoplus_{p\in\mathcal{P}(n\to j)} \operatorname{Im} X_p$$

is a Gorenstein-projective A-module for each $i \in Q_0'$.

Proof. We prove the assertion by using induction on l_i, which is defined in the proof of Lemma 5.5. If $i \in Q_0'$ with $l_i = 0$, the assertion follows from (G).

Suppose $l_i \geqslant 1$. Since $\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p \subseteq \bigoplus_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} \operatorname{Im} X_\alpha$, we have the exact sequence

$$0 \longrightarrow \Big(\bigoplus_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} \operatorname{Im} X_\alpha \Big) \Big/ \Big(\bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p \Big)$$
$$\longrightarrow X_i \Big/ \bigoplus_{p\in\mathcal{P}(n\to i)} \operatorname{Im} X_p \longrightarrow X_i \Big/ \bigoplus_{\substack{\alpha\in Q_1\\ e(\alpha)=i}} \operatorname{Im} X_\alpha \longrightarrow 0,$$

and by (G) the last term on the second line is Gorenstein-projective. It suffices to prove that the term on the first line is Gorenstein-projective. By (5-1) this term is $\bigoplus_{\substack{\alpha\in Q_1'\\ e(\alpha)=i}} \operatorname{Im} \widetilde{X}_\alpha$. By Lemma 5.6 each $\widetilde{X}_\alpha$ is injective, and it follows that

$$\operatorname{Im} \widetilde{X}_\alpha \cong X_j \Big/ \bigoplus_{p\in\mathcal{P}(n\to j)} \operatorname{Im} X_p,$$

where $j = s(\alpha)$. Since $l_j < l_i$, the conclusion of the lemma follows from the inductive hypothesis. ☐

Lemma 5.9. *The sufficiency in Theorem 5.1 holds. That is, if*

$$X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$$

is a monic Λ-module satisfying (G), X is Gorenstein-projective.

Proof. Using induction on $n = |Q_0|$, the assertion clearly holds for $n = 1$. Suppose that the assertion holds for $n-1$ with $n \geq 2$. It suffices to prove that X satisfies Lemma 5.3(i)–(iii).

Condition (i) is contained in (G); and condition (ii) follows from Lemma 5.5(2). By Lemma 5.6 $\operatorname{Coker}\phi$ is a monic Λ'-module; and by Lemmas 5.7 and 5.8 we know that $\operatorname{Coker}\phi$ satisfies (G). It follows from the inductive hypothesis that condition (iii) is satisfied. $\square$

Lemma 5.10. *Let $X = (X_i, X_\alpha, i \in Q_0, \alpha \in Q_1)$ be a Λ-module with X_n a Gorenstein-projective A-module. Then $P \otimes_A X_n$ is a Gorenstein-projective Λ'-module, where P is defined in Section 5B.*

Proof. Let $P(n)$ be the indecomposable projective kQ-module at vertex n. Writing $\operatorname{rad} P(n)$ as a representation of Q' over k, we have

$$\operatorname{rad} P(n) = (k^{m_i}, f_\alpha, i \in Q'_0, \alpha \in Q'_1),$$

where $m_i = |\mathcal{P}(n \to i)|$ for each $i \in Q'_0$. By the construction of $P(n)$ we know that $\operatorname{rad} P(n)$ has the following three properties:

(1) Each $f_\alpha : k^{m_{s(\alpha)}} \to k^{m_{e(\alpha)}}$ is injective.

(2) For each $i \in Q'_0$,

$$\sum_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} \operatorname{Im} f_\alpha = \bigoplus_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} \operatorname{Im} f_\alpha.$$

(3) For each $i \in Q'_0$, $k^{m_i}/\big(\bigoplus_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} \operatorname{Im} f_\alpha\big)$ and $k^{|\mathcal{A}(n \to i)|}$ are isomorphic as k-spaces.

It follows that

$$\begin{aligned} P \otimes_A X_n &\cong (\operatorname{rad} P(n) \otimes_k A) \otimes_A X_n \\ &\cong \operatorname{rad} P(n) \otimes_k X_n = (X_n^{m_i}, f_\alpha \otimes_k \operatorname{id}_{X_n}, i \in Q'_0, \alpha \in Q'_1). \end{aligned}$$

By (1), (2), and (3) we clearly see that $P \otimes_A X_n$ is a monic Λ'-module satisfying (G); for example, by (3) we know that

$$X_n^{m_i} \Big/ \bigoplus_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} \operatorname{Im}(f_\alpha \otimes_k \operatorname{id}_{X_n}) \cong X_n^{|\mathcal{A}(n \to i)|}$$

is a Gorenstein-projective A-module. Now the result follows from Lemma 5.9. $\square$

5D. ***Proof of Theorem 5.1.*** By Lemma 5.9 it remains to prove necessity, namely, if X is a Gorenstein-projective Λ-module, X is a monic Λ-module satisfying (G). Using induction on $n = |Q_0|$, the assertion is clear for $n = 1$. Suppose that the assertion holds for $n-1$ with $n \geq 2$. We write X as $\left(\begin{smallmatrix} X' \\ X_n \end{smallmatrix}\right)_\phi$. Then X satisfies conclusions (i)–(iii) of Lemma 5.3.

By (i) and Lemma 5.10 we know that $P \otimes_A X_n$ is a Gorenstein-projective Λ'-module. Then, by (ii) and (iii), we know that $X' \in \mathcal{GP}(\Lambda')$, since $\mathcal{GP}(\Lambda')$ is closed under extensions. By the inductive hypothesis X' is a monic Λ'-module satisfying (G). Hence:

(1) X_β is injective for each $\beta \in Q'_1$.

(2) X_i is Gorenstein-projective for each $i \in Q'_0$.

(3) X_α is injective for each $\alpha \in Q_1$.

(4) $\sum_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p = \bigoplus_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p$ for all $i \in Q'_0$.

We get (3) and (4) from (1), condition (ii), and Lemma 5.4.

Since $\operatorname{Coker} \phi = \left(X_i / \bigoplus_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p,\ \widetilde{X}_\alpha,\ i \in Q'_0,\ \alpha \in Q'_1\right)$ is a Gorenstein-projective Λ'-module, it follows from the inductive hypothesis that the following properties hold:

(5) For each $\alpha \in Q'_1$, $\widetilde{X}_\alpha$ is injective.

(6) $\sum_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} \operatorname{Im} \widetilde{X}_\alpha = \bigoplus_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} \operatorname{Im} \widetilde{X}_\alpha$, for all $i \in Q'_0$.

Claim 1: *X satisfies* (m2) *in Definition 2.2*.

Indeed, suppose

$$\sum_{\substack{\alpha \in Q_1 \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha}) = 0. \tag{5-2}$$

Since

$$\sum_{\substack{\alpha \in Q_1 \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha}) = \sum_{\alpha \in \mathcal{A}(n \to i)} X_\alpha(x_{s(\alpha),\alpha}) + \sum_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha}),$$

it follows that

$$\begin{aligned} \sum_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} \widetilde{X}_\alpha(\overline{x_{s(\alpha),\alpha}}) &= \sum_{\substack{\alpha \in Q'_1 \\ e(\alpha)=i}} X_\alpha(x_{s(\alpha),\alpha}) + \bigoplus_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p \\ &= -\sum_{\alpha \in \mathcal{A}(n \to i)} X_\alpha(x_{s(\alpha),\alpha}) + \bigoplus_{p \in \mathcal{P}(n \to i)} \operatorname{Im} X_p = 0, \end{aligned}$$

where we used (5-2) for the second equality.

Then by (6) we have $\widetilde{X}_\alpha(\overline{x_{s(\alpha),\alpha}})=0$; and by (5) we know $\overline{x_{s(\alpha),\alpha}}=0$ for each $\alpha\in Q_1'$ with $e(\alpha)=i$. This means that there are $x_{n,q}\in X_n$ such that

$$x_{s(\alpha),\alpha}=\sum_{q\in\mathcal{P}(n\to s(\alpha))}X_q(x_{n,q})\ \in \sum_{q\in\mathcal{P}(n\to s(\alpha))}\operatorname{Im}X_q$$

for each $\alpha\in Q_1'$ with $e(\alpha)=i$. By (5-2) we have

$$0\ =\sum_{\alpha\in\mathcal{A}(n\to i)}X_\alpha(x_{n,\alpha})+\sum_{\substack{\alpha\in Q_1'\\ e(\alpha)=i}}X_\alpha\Bigg(\sum_{q\in\mathcal{P}(n\to s(\alpha))}X_q(x_{n,q})\Bigg).$$

By (4) we know that $X_\alpha(x_{n,\alpha})=0$ for all $\alpha\in\mathcal{A}(n\to i)$, and that $X_\alpha X_q(x_{n,q})=0$ for all $\alpha\in Q_1'$ with $e(\alpha)=i$ and $q\in\mathcal{P}(n\to s(\alpha))$. Thus $X_\alpha(x_{s(\alpha),\alpha})=0$, for all $\alpha\in Q_1$ with $e(\alpha)=i$. This proves Claim 1.

Claim 2: $X_i/\bigoplus_{\substack{\beta\in Q_1\\ e(\beta)=i}}\operatorname{Im}X_\beta$ *is a Gorenstein-projective A-module for each $i\in Q_0$.*

Indeed, since $\operatorname{Coker}\phi$ is a Gorenstein-projective Λ'-module, by the inductive hypothesis we know that

$$\Bigg(X_i\Big/\bigoplus_{p\in\mathcal{P}(n\to i)}\operatorname{Im}X_p\Bigg)\Big/\bigoplus_{\substack{\alpha\in Q_1'\\ e(\alpha)=i}}\operatorname{Im}\widetilde{X}_\alpha$$

is a Gorenstein-projective A-module: it is exactly the desired module by (5-1).

Now, (3) and Claim 1 mean that X is a monic Λ-module; and (2), together with conclusion (i) of Lemma 5.3 and Claim 2, means that X satisfies (G). □

6. Corollaries

6A. For the definition of a Frobenius category in the sense of [Quillen 1973], we refer to [Happel 1988, p. 11; Keller 1990, Appendix A]. As a consequence of Theorem 5.1 and Proposition 2.4, we get the following characterization of self-injectivity.

Corollary 6.1. *Let A be a finite-dimensional algebra, and Q a finite acyclic quiver. Then the following are equivalent*:

(i) *A is self-injective.*

(ii) $\mathcal{GP}(A\otimes_k kQ)=\operatorname{Mon}(Q,A)$.

(iii) $\operatorname{Mon}(Q,A)$ *is a Frobenius category.*

Proof. (i) $\Longrightarrow$ (ii): If A is self-injective, every A-module is Gorenstein-projective, and hence (ii) follows from Theorem 5.1. The implication (ii)$\Longrightarrow$ (iii) is well-known.

(iii) $\Longrightarrow$ (i): Take a sink of Q, say vertex 1, and consider $D(A_A)\otimes_k P(1)$. By Proposition 2.4 (iii) it is an injective object in $\mathrm{Mon}(Q, A)$, and hence, by assumption, it is a projective object in $\mathrm{Mon}(Q, A)$. By Proposition 2.4(ii) we know that $D(A_A)$, the first branch of $D(A_A)\otimes_k P(1)$, is a projective A-module, that is, A is self-injective. $\square$

Let $D^b(\Lambda)$ be the bounded derived category of Λ, and $K^b(\mathcal{P}(\Lambda))$ the bounded homotopy category of $\mathcal{P}(\Lambda)$. By definition the singularity category $D^b_{sg}(\Lambda)$ of Λ is the Verdier quotient $D^b(\Lambda)/K^b(\mathcal{P}(\Lambda))$. Buchweitz [1987, Theorem 4.4.1] proved that if Λ is Gorenstein, there is a triangle-equivalence $D^b_{sg}(\Lambda)\cong \underline{\mathcal{GP}(\Lambda)}$, where $\underline{\mathcal{GP}(\Lambda)}$ is the stable category of $\mathcal{GP}(\Lambda)$ modulo $\mathcal{P}(\Lambda)$; see also [Happel 1991, Theorem 4.6]. Note that if A is Gorenstein, $\Lambda = A\otimes_k kQ$ is Gorenstein; see [Auslander and Reiten 1991b, Proposition 2.2]. So we have the following.

Corollary 6.2. *Let A be a finite-dimensional Gorenstein algebra, and Q a finite acyclic quiver. Let $\Lambda = A\otimes_k kQ$. Then there is a triangle-equivalence $D^b_{sg}(\Lambda)\cong \underline{\mathcal{GP}(\Lambda)}$. In particular, if A is self-injective, then there is a triangle-equivalence $D^b_{sg}(\Lambda)\cong \underline{\mathrm{Mon}(Q, A)}$.*

6B. Recall the tensor product $Q\otimes Q'$ of two finite quivers Q and Q' (not necessarily acyclic). By definition $Q\otimes Q'$ is the quiver with

$$(Q\otimes Q')_0 = Q_0\times Q'_0 \quad\text{and}\quad (Q\otimes Q')_1 = (Q_1\times Q'_0)\cup(Q_0\times Q'_1).$$

More explicitly, if $\alpha : i\to j$ is an arrow of Q, then, for each vertex $t'\in Q'_0$, there is an arrow $(\alpha, t') : (i, t')\to(j, t')$ of $Q\otimes Q'$; and if $\beta' : s'\to t'$ is an arrow of Q', then, for each vertex $i\in Q_0$, there is an arrow $(i, \beta') : (i, s')\to(i, t')$ of $Q\otimes Q'$.

Let $A = kQ/I$ and $B = kQ'/I'$ be two finite-dimensional k-algebras, where Q and Q' are finite quivers (not necessarily acyclic), and I, I' are admissible ideals of kQ, kQ', respectively. Then

$$A\otimes_k B\cong k(Q\otimes Q')/I\,\square\, I',$$

where $I\,\square\, I'$ is the ideal of $k(Q\otimes Q')$ generated by $(I\times Q'_0)\cup(Q_0\times I')$ and the elements

$$(\alpha, t')(i, \beta') - (j, \beta')(\alpha, s'),$$

where $\alpha : i\to j$ is an arrow of Q, and $\beta' : s'\to t'$ is an arrow of Q'. See, for example, [Leszczyński 1994]. Note that $I\,\square\, I'$ may not be zero even if $I = 0 = I'$. We have proved this:

Fact. *$A\otimes_k B$ is hereditary (that is, $I\,\square\, I' = 0$) if and only if either $A\cong k^{|Q_0|}$ as algebras and $I' = 0$, or $B\cong k^{|Q'_0|}$ as algebras and $I = 0$.*

6C. One can describe when Λ is hereditary via $\mathrm{Mon}(Q, A)$.

Corollary 6.3. *Let A be a finite-dimensional basic algebra over an algebraically closed field k, Q a finite acyclic quiver with $|Q_1| \neq 0$, and $\Lambda = A \otimes_k kQ$. Then $\mathcal{P}(\Lambda) = \mathrm{Mon}(Q, A)$ if and only if Λ is hereditary.*

Proof. Without loss of generality we may assume that A is connected (an algebra is connected if it cannot be a product of two nonzero algebras).

If $\Lambda = A \otimes_k kQ$ is hereditary, then, by the fact above and the assumption on Q, we have $A = k$, and hence $\mathrm{Mon}(Q, k) = \mathcal{GP}(kQ)$ by Theorem 5.1. It follows that

$$\mathrm{Mon}(Q, A) = \mathcal{GP}(kQ) = \mathcal{P}(kQ) = \mathcal{P}(\Lambda).$$

Conversely, if $A \neq k$, A is not semisimple since A is assumed to be connected and basic and k is assumed to be algebraically closed. It follows that there is a nonprojective A-module M. Take a sink of Q, say vertex 1, and consider Λ-module $X = M \otimes_k P(1)$, where $P(1)$ is the simple projective kQ-module at vertex 1. Then $X \in \mathrm{Mon}(Q, A)$, but $X \notin \mathcal{P}(\Lambda)$. □

References

[Arnold 2000] D. M. Arnold, *Abelian groups and representations of finite partially ordered sets*, CMS Books in Mathematics **2**, Springer, New York, 2000. MR 2001g:16030 Zbl 0959.16011

[Auslander and Bridger 1969] M. Auslander and M. Bridger, *Stable module theory*, Memoirs of the American Mathematical Society **94**, American Mathematical Society, Providence, R.I., 1969. MR 42 #4580 Zbl 0204.36402

[Auslander and Reiten 1991a] M. Auslander and I. Reiten, "Applications of contravariantly finite subcategories", *Adv. Math.* **86**:1 (1991), 111–152. MR 92e:16009 Zbl 0774.16006

[Auslander and Reiten 1991b] M. Auslander and I. Reiten, "Cohen–Macaulay and Gorenstein Artin algebras", pp. 221–245 in *Representation theory of finite groups and finite-dimensional algebras* (Bielefeld, 1991), edited by G. O. Michler and C. M. Ringel, Progr. Math. **95**, Birkhäuser, Basel, 1991. MR 92k:16018 Zbl 0776.16003

[Auslander and Smalø 1981] M. Auslander and S. O. Smalø, "Almost split sequences in subcategories", *J. Algebra* **69**:2 (1981), 426–454. MR 82j:16048a Zbl 0457.16017

[Auslander et al. 1995] M. Auslander, I. Reiten, and S. O., *Representation theory of Artin algebras*, Cambridge Studies in Advanced Mathematics **36**, Cambridge University Press, 1995. MR 96c:16015 Zbl 0834.16001

[Avramov and Martsinkovsky 2002] L. L. Avramov and A. Martsinkovsky, "Absolute, relative, and Tate cohomology of modules of finite Gorenstein dimension", *Proc. London Math. Soc.* (3) **85**:2 (2002), 393–440. MR 2003g:16009 Zbl 1047.16002

[Beligiannis 2005] A. Beligiannis, "Cohen–Macaulay modules, (co)torsion pairs and virtually Gorenstein algebras", *J. Algebra* **288**:1 (2005), 137–211. MR 2006i:16017 Zbl 1119.16007

[Buchweitz 1987] R.-O. Buchweitz, "Maximal Cohen–Macaulay modules and Tate cohomology over Gorenstein rings", Unpub. manu., 1987, Available at http://tinyurl.com/cohen-mac1986-pdf.

[Chen 2011] X.-W. Chen, "The stable monomorphism category of a Frobenius category", *Math. Res. Lett.* **18**:1 (2011), 125–137. MR 2012e:18024 Zbl 06026606

[Enochs and Jenda 1995] E. E. Enochs and O. M. G. Jenda, "Gorenstein injective and projective modules", *Math. Z.* **220**:4 (1995), 611–633. MR 97c:16011 Zbl 0845.16005

[Enochs and Jenda 2000] E. E. Enochs and O. M. G. Jenda, *Relative homological algebra*, de Gruyter Expositions in Mathematics **30**, Walter de Gruyter, Berlin, 2000. MR 2001h:16013 Zbl 0952.13001

[Gao and Zhang 2010] N. Gao and P. Zhang, "Gorenstein derived categories", *J. Algebra* **323**:7 (2010), 2041–2057. MR 2011f:18017 Zbl 1222.18005

[Happel 1988] D. Happel, *Triangulated categories in the representation theory of finite-dimensional algebras*, London Mathematical Society Lecture Note Series **119**, Cambridge University Press, 1988. MR 89e:16035 Zbl 0635.16017

[Happel 1991] D. Happel, "On Gorenstein algebras", pp. 389–404 in *Representation theory of finite groups and finite-dimensional algebras* (Bielefeld, 1991), edited by G. O. Michler and C. M. Ringel, Progr. Math. **95**, Birkhäuser, Basel, 1991. MR 92k:16022 Zbl 0759.16007

[Holm 2004] H. Holm, "Gorenstein homological dimensions", *J. Pure Appl. Algebra* **189**:1-3 (2004), 167–193. MR 2004k:16013 Zbl 1050.16003

[Iyama et al. 2011] O. Iyama, K. Kato, and J.-I. Miyachi, "Recollement of homotopy categories and Cohen–Macaulay modules", *J. K-Theory* **8**:3 (2011), 507–542. MR 2012k:18018 Zbl 1251.18008

[Keller 1990] B. Keller, "Chain complexes and stable categories", *Manuscripta Math.* **67**:4 (1990), 379–417. MR 91h:18006 Zbl 0753.18005

[Krause and Solberg 2003] H. Krause and Ø. Solberg, "Applications of cotorsion pairs", *J. London Math. Soc.* (2) **68**:3 (2003), 631–650. MR 2004k:16028 Zbl 1061.16006

[Kussin et al. 2012] D. Kussin, H. Lenzing, and H. Meltzer, "Nilpotent operators and weighted projective lines", 2012. To appear in *J. Reine Angew. Math.* arXiv 1002.3797

[Leszczyński 1994] Z. Leszczyński, "On the representation type of tensor product algebras", *Fund. Math.* **144**:2 (1994), 143–161. MR 95b:16013 Zbl 0817.16008

[Li and Zhang 2010] Z.-W. Li and P. Zhang, "A construction of Gorenstein-projective modules", *J. Algebra* **323**:6 (2010), 1802–1812. MR 2011d:16012 Zbl 1210.16011

[Moore 2010] A. Moore, "The Auslander and Ringel–Tachikawa theorem for submodule embeddings", *Comm. Algebra* **38**:10 (2010), 3805–3820. MR 2011k:16040 Zbl 1237.16015

[Quillen 1973] D. Quillen, "Higher algebraic K-theory, I", pp. 85–147 in *Algebraic K-theory, 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

[Ringel 1984] C. M. Ringel, *Tame algebras and integral quadratic forms*, Lecture Notes in Mathematics **1099**, Springer, Berlin, 1984. MR 87f:16027 Zbl 0546.16013

[Ringel and Schmidmeier 2006] C. M. Ringel and M. Schmidmeier, "Submodule categories of wild representation type", *J. Pure Appl. Algebra* **205**:2 (2006), 412–422. MR 2006i:16025 Zbl 1147.16019

[Ringel and Schmidmeier 2008a] C. M. Ringel and M. Schmidmeier, "The Auslander–Reiten translation in submodule categories", *Trans. Amer. Math. Soc.* **360**:2 (2008), 691–716. MR 2008j:16058 Zbl 1154.16011

[Ringel and Schmidmeier 2008b] C. M. Ringel and M. Schmidmeier, "Invariant subspaces of nilpotent linear operators, I", *J. Reine Angew. Math.* **614** (2008), 1–52. MR 2009d:16016 Zbl 1145.16005

[Ringel and Zhang 2011] C. M. Ringel and P. Zhang, "Representations of quivers over the algebras of dual numbers", preprint, 2011. arXiv 1112.1924

[Simson 2007] D. Simson, "Representation types of the category of subprojective representations of a finite poset over $K[t]/(t^m)$ and a solution of a Birkhoff type problem", *J. Algebra* **311**:1 (2007), 1–30. MR 2009b:16040 Zbl 1123.16010

[Xiong and Zhang 2012] B.-L. Xiong and P. Zhang, "Gorenstein-projective modules over triangular matrix Artin algebras", *J. Algebra Appl.* **11**:4 (2012), Article ID 1250066, 14. MR 2959415 Zbl pre06078475

[Zhang 2011] P. Zhang, "Monomorphism categories, cotilting theory, and Gorenstein-projective modules", *J. Algebra* **339** (2011), 181–202. MR 2012k:16034 Zbl pre06009303

Received May 10, 2012.

XIU-HUA LUO
DEPARTMENT OF MATHEMATICS
SHANGHAI JIAO TONG UNIVERSITY
DONGCHUAN ROAD 800
SHANGHAI, 200240
CHINA
x.h.luo@sjtu.edu.cn

PU ZHANG
DEPARTMENT OF MATHEMATICS
SHANGHAI JIAO TONG UNIVERSITY
DONGCHUAN ROAD 800
SHANGHAI, 200240
CHINA
pzhang@sjtu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.195

HELICOIDAL FLAT SURFACES IN HYPERBOLIC 3-SPACE

ANTONIO MARTÍNEZ, JOÃO PAULO DOS SANTOS AND KETI TENENBLAT

A flat surface in hyperbolic space $\mathbb{H}^3$ is determined by a harmonic function as well as by its meromorphic data. In this paper, helicoidal flat surfaces in $\mathbb{H}^3$ are considered. A complete classification of the helicoidal flat fronts is given in terms of their hyperbolic Gauss maps as well as by means of linear harmonic functions. A family of examples that provides the classification of the helicoidal flat fronts is included. Moreover, it is shown that a flat surface in $\mathbb{H}^3$ that corresponds to a linear harmonic function is locally congruent to a helicoidal flat front or to a peach front.

1. Introduction

The study of flat surfaces in hyperbolic 3-space has received much attention in the last few years, mainly because Gálvez, Martínez and Milán [Gálvez et al. 2000] have shown that flat surfaces in hyperbolic 3-space admit a Weierstrass representation formula in terms of meromorphic data as in the theory of minimal surfaces in $\mathbb{R}^3$.

It is known that the only complete examples are the horospheres and the hyperbolic cylinders (see [Spivak 1979]). Thus, a study of flat surfaces with singularities became essential for the advancement of the theory. An important contribution was given in [Kokubu et al. 2005; 2004], where an extension of the Weierstrass representation for flat surfaces with admissible singularities was introduced. Such surfaces are called *flat fronts*.

Helicoidal surfaces arise as a generalization of rotational surfaces and conical surfaces. They are invariant by a subgroup of the group of space isometries, called the *helicoidal group*, defined by a translation composed with rotation around an axis in the same direction. Rotational flat surfaces were classified in [Kokubu et al. 2004] in terms of meromorphic data.

The main purpose of this paper is to give a complete classification of the helicoidal flat surfaces in $\mathbb{H}^3$ in terms of meromorphic data as well as by means of linear

Research partially supported by the Ministério de Educação, CAPES/DGU Proc. No. 246/11, Brazil; Ministerio de Ciencia e Innovación MTM2010-1982, Spain; and Junta de Andalucía Grant FQM325.
MSC2010: 53A35, 53C42.
Keywords: helicoidal surfaces, flat surfaces, flat fronts, hyperbolic space, conformal representation.

harmonic functions. We construct a family of examples, which we call classifying examples, that provide the complete classification of the helicoidal flat surfaces. These results extend those obtained previously in [Kokubu et al. 2004] for rotational flat surfaces. Moreover, we characterize the flat fronts in $\mathbb{H}^3$ that correspond to linear harmonic functions.

The paper is organized as follows:

In Section 2, we give a brief description of helicoidal flat surfaces in $\mathbb{H}^3$ and present two particular cases of this class of surfaces, namely, the rotational and the conical flat surfaces.

In Section 3, we recall the well known result that in a neighborhood of a nonumbilic point, any flat surface in $\mathbb{H}^3$ admits a local parametrization that diagonalizes both the first and second fundamental forms determined by a (euclidean) harmonic function. We then present the conformal representation for flat fronts described in [Corro et al. 2010; Kokubu et al. 2005] and use it to characterize when a complex parameter diagonalizes both the first and second fundamental forms in terms of the hyperbolic Gauss maps. Moreover, we relate the harmonic function to the hyperbolic Gauss maps.

In Section 4, we describe a family of flat fronts that we call *classifying examples* whose hyperbolic Gauss maps are determined by a nonzero complex number, and we obtain the corresponding harmonic function.

Finally, in Section 5, we prove that a flat front in $\mathbb{H}^3$ is helicoidal if and only if it is locally congruent to one of the classifying examples. Moreover, we obtain a complete classification of the helicoidal flat fronts in terms of their hyperbolic Gauss maps as well as by means of suitable linear harmonic functions. We conclude by showing that any flat surface in $\mathbb{H}^3$ that corresponds to a linear harmonic function is locally congruent either to a helicoidal flat surface or to a so-called peach front.

2. Helicoidal surfaces

Helicoidal surfaces arise as a natural generalization of rotational surfaces. They are invariant under a one-parameter group of isometries obtained by composing a translation in a given direction with a rotation about an axis in the same direction. We consider the *half-space model* of the hyperbolic 3-space, that is,

$$\mathbb{H}^3 = \{(y_1, y_2, y_3) \in \mathbb{R}^3 : y_3 > 0\}$$

endowed with the metric

$$\langle\,,\rangle = \frac{1}{y_3^2}\left(dy_1^2 + dy_2^2 + dy_3^2\right),$$

with ideal boundary $\mathbb{C}_\infty = \{(y_1, y_2, 0) \mid y_1, y_2 \in \mathbb{R}\} \cup \{\infty\}$.

The helicoidal group relative to the y_3-axis is given as the composition

$$h_t = \begin{pmatrix} e^{\beta t} & 0 & 0 \\ 0 & e^{\beta t} & 0 \\ 0 & 0 & e^{\beta t} \end{pmatrix} \begin{pmatrix} \cos \alpha t & -\sin \alpha t & 0 \\ \sin \alpha t & \cos \alpha t & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

of a rotation around the y_3-axis with *angular pitch* α with a hyperbolic translation of *ratio* β.

Every helicoidal surface can be generated by a suitable curve $\gamma : I \to \mathbb{H}^3$ by taking the composition

$$\psi(t, s) = (h_t \circ \gamma)(s). \tag{2-1}$$

Notice that the curve γ is chosen suitably so that (2-1) is a regular surface.

In order to have the helicoidal surface (2-1) flat, one has to require conditions on the curve γ, as in the following particular cases:

(i) Rotational flat surfaces ($\beta = 0$). We start with a curve parametrized by arc length on the plane $\{y_2 \equiv 0\}$. It follows from the Gauss equation that the remaining coordinates, y_1 and y_3, must satisfy the differential equation

$$\left(\frac{y_1}{y_3}\right)'' \frac{y_3}{y_1} = 0$$

[do Carmo and Dajczer 1983; Spivak 1979], giving us the relation $y_1(s) = (as + b) y_3(s)$.

(ii) Conical flat surfaces ($\alpha = 0$). In this case, we just have a movement of translation. We can start with a curve on the horosphere $\{y_3 = c\}$, where $c > 0$ is a constant. Assuming that $c = 1$, we consider the curve

$$\gamma(s) = \big(r(s) \cos\theta(s), r(s) \sin\theta(s), 1\big),$$

parametrized by the arc length, that is,

$$(r')^2 + (r\theta')^2 = 1.$$

Then the surface is flat if and only if one has the following expression for r:

$$r(s) = \sqrt{(as + b)^2 - 1}.$$

Therefore, we see that helicoidal surfaces arise as a generalization of the rotational and the conical surfaces, which are well known flat examples. We want to describe all the helicoidal flat surfaces in hyperbolic 3-space.

3. Conformal representation

In this section, we characterize the flat surfaces in $\mathbb{H}^3$ by means of their first and second fundamental forms. We start by recalling that on a neighborhood of a nonumbilic point, any flat surface in $\mathbb{H}^3$ admits a local parametrization that diagonalizes both the first and second fundamental forms, which are determined by a (euclidean) harmonic function. We then consider the conformal representation for flat fronts and characterize a complex parameter that diagonalizes both the first and second fundamental forms in terms of the hyperbolic Gauss maps. We also relate the harmonic function to these maps.

It is well known that on a neighborhood of a nonumbilical point, a flat surface in $\mathbb{H}^3$ can be parametrized by lines of curvature, so that the first and second fundamental forms are given by

$$\mathrm{I} = \cosh^2\phi(u,v)(du)^2 + \sinh^2\phi(u,v)(dv)^2, \tag{3-1}$$

$$\mathrm{II} = \sinh\phi(u,v)\cosh\phi(u,v)\big((du)^2 + (dv)^2\big), \tag{3-2}$$

where ϕ is a harmonic function, that is, $\phi_{uu} + \phi_{vv} = 0$ (for details, see [Tenenblat 1998, Theorem 2.4 and Corollary 2.7]). We will show, in Section 5, that a helicoidal flat surface in $\mathbb{H}^3$ is characterized as a flat surface whose first and second fundamental forms are given by (3-1) and (3-2) where ϕ is linear, that is,

$$\phi(u,v) = au + bv + c, \tag{3-3}$$

and a, b and c are real numbers such that $(a,\ b,\ c) \neq (0,\ \pm 1,\ 0)$.

We will use the conformal representation for flat surfaces in $\mathbb{H}^3$ introduced in [Gálvez et al. 2000]. Let Σ be a 2-manifold and $\psi : \Sigma \to \mathbb{H}^3$ a flat immersion. It follows from the Gauss equation that the second fundamental form $d\sigma^2$ is definite, and hence Σ is orientable and inherits a canonical Riemann surface structure such that the second fundamental form $d\sigma^2$ is hermitian. This canonical Riemann surface structure provides a conformal representation for the immersion ψ that allows one to recover any flat surface in $\mathbb{H}^3$ in terms of holomorphic data (see [Gálvez et al. 2000; Kokubu et al. 2004] for details). Throughout this paper, we will regard Σ as a Riemann surface with the conformal structure determined by the second fundamental form $d\sigma^2$.

For any $p \in \Sigma$, there exist $g(p),\ g^*(p) \in \mathbb{C}_\infty$ distinct points in the ideal boundary such that the oriented normal geodesic at $\psi(p)$ is the geodesic in $\mathbb{H}^3$ starting from $g^*(p)$ towards $g(p)$. The maps $g,\ g^* : \Sigma \to \mathbb{C}_\infty$ are called the *hyperbolic Gauss maps*, and it is proved in [Gálvez et al. 2000] that, for flat surfaces, they are holomorphic when we regard $\mathbb{C}_\infty$ as the Riemann sphere.

Kokubu et al. [2004] extended the conformal representation given by Gálvez et al. [2000] for *flat fronts*, that is, flat immersions with some admissible singularities occurring where the first fundamental form degenerates. They showed how to recover flat fronts in terms of the hyperbolic Gauss maps and how these maps are well defined through the singularities. Reformulating the results in Theorem 2.11 and Proposition 2.5 of [Kokubu et al. 2004] to the upper half-space model, we have the following theorem (see [Corro et al. 2010]):

Theorem 1. *Let g and g^* be nonconstant meromorphic functions on a Riemann surface Σ such that $g(p) \neq g^*(p)$ for all $p \in \Sigma$. Assume that*

(1) *all the poles of the* 1*-form* $\dfrac{dg}{g-g^*}$ *are of order* 1, *and*
(2) $\mathrm{Re} \displaystyle\int_\gamma \frac{dg}{g-g^*} = 0$ *for each loop γ on Σ.*

Set

$$\xi := c \exp \int \frac{dg}{g-g^*}, \qquad c \in \mathbb{C} \setminus \{0\}. \tag{3-4}$$

Then the map $\psi = (\psi_1, \psi_2, \psi_3) : \Sigma \to \mathbb{H}^3$ given by

$$\psi_1 + \mathrm{i}\psi_2 = g - \frac{|\xi|^4 (g-g^*)}{|\xi|^4 + |g-g^*|^2}, \qquad \psi_3 = \frac{|\xi|^2 |g-g^*|^2}{|\xi|^4 + |g-g^*|^2} \tag{3-5}$$

is a flat front. Moreover, if we consider the 1*-forms*

$$\omega = -\frac{1}{\xi^2} g_z \, dz, \quad \theta = \frac{\xi^2}{(g-g^*)^2} g_z^* \, dz, \tag{3-6}$$

where z is a complex parameter, then the first and second fundamental forms are represented as

$$\mathrm{I} = (\omega + \bar{\theta})(\bar{\omega} + \theta), \tag{3-7}$$

$$\mathrm{II} = |\theta|^2 - |\omega|^2. \tag{3-8}$$

The next proposition provides a necessary and sufficient condition on the functions g and g^* in order to diagonalize the first and second fundamental forms simultaneously:

Proposition 2. *Let Σ be a flat front in $\mathbb{H}^3$ given as in Theorem 1. A complex parameter for Σ, $\eta = u + \mathrm{i}v$, diagonalizes the first and second fundamental forms simultaneously as in (3-1) and (3-2) if and only if*

$$\frac{g_\eta g_\eta^*}{(g-g^*)^2} = -\frac{1}{4}, \tag{3-9}$$

where $(\,\cdot\,)_\eta$ is the derivative with respect to η. In this case, the harmonic function ϕ

is given by

$$e^{2\phi} = \frac{|g^*_\eta||\xi|^4}{|g_\eta||g-g^*|^2} = \frac{|\xi|^4}{4|g_\eta|^2}. \tag{3-10}$$

Proof. It follows from (3-6) and (3-7) that

$$I = |\omega|^2 + |\theta|^2 + 2\,\mathrm{Re}(\omega\theta), \tag{3-11}$$

where

$$\omega\theta = -\frac{g_\eta g^*_\eta (d\eta)^2}{(g-g^*)^2},$$

since g and g^* are holomorphic functions on the parameter η.

By writing

$$\frac{g_\eta g^*_\eta}{(g-g^*)^2} = A + \mathrm{i}B,$$

we have $\mathrm{Re}(\theta\omega) = -A(du^2 - dv^2) + 2B\,du\,dv$. Then if η diagonalizes the first and second fundamental forms as in (3-1) and (3-2), we must have $B = 0$. Therefore $g_\eta g^*_\eta/(g-g^*)^2$ is real and holomorphic, which implies it must be a constant function.

If we write

$$\frac{g_\eta g^*_\eta}{(g-g^*)^2} = c_g \tag{3-12}$$

and use equations (3-6) and (3-11), we have the first fundamental form as in (3-1) if and only if

$$\frac{|g_\eta|^2}{|\xi|^4} + \frac{|\xi|^4|g^*_\eta|^2}{|g-g^*|^4} - 2c_g = \cosh^2\phi, \qquad \frac{|g_\eta|^2}{|\xi|^4} + \frac{|\xi|^4|g^*_\eta|^2}{|g-g^*|^4} + 2c_g = \sinh^2\phi.$$

Hence, $c_g = -\frac{1}{4}$ and (3-9) is proved.

Now we prove the expression (3-10). With this value for c_g, using equations (3-9) and (3-12), we rewrite the expressions above as

$$\begin{aligned} \frac{|g_\eta||g-g^*|^2}{4|\xi|^4|g^*_\eta|} + \frac{|\xi|^4|g^*_\eta|}{4|g_\eta||g-g^*|^2} + \frac{1}{2} &= \cosh^2\phi, \\ \frac{|g_\eta||g-g^*|^2}{4|\xi|^4|g^*_\eta|} + \frac{|\xi|^4|g^*_\eta|}{4|g_\eta||g-g^*|^2} - \frac{1}{2} &= \sinh^2\phi. \end{aligned} \tag{3-13}$$

Considering $\lambda = \dfrac{|g_\eta||g-g^*|^2}{|\xi|^4|g^*_\eta|}$, we conclude that

$$\left(\lambda + \frac{1}{\lambda}\right) = e^{2\phi} + e^{-2\phi}. \tag{3-14}$$

If we now consider the second fundamental form, we have

$$\begin{aligned} \mathrm{II} &= |\theta|^2 - |\omega|^2 \\ &= \left(\frac{|\xi|^4 |g^*_\eta|}{4|g-g^*|^2|g_\eta|} - \frac{|g-g^*|^2|g_\eta|}{4|\xi|^4|g^*_\eta|} \right)(du^2+dv^2) = \frac{1}{4}\left(\frac{1}{\lambda} - \lambda\right)(du^2+dv^2). \end{aligned}$$

Therefore, it follows from (3-2) that we must have

$$\left(\frac{1}{\lambda} - \lambda\right) = e^{2\phi} - e^{-2\phi}. \tag{3-15}$$

Combining (3-14) and (3-15), we conclude that

$$e^{2\phi} = \frac{1}{\lambda} = \frac{|\xi|^4|g^*_\eta|}{|g_\eta|\,|g-g^*|^2}. \qquad \square$$

Corollary 3. *Let Σ be a flat front in $\mathbb{H}^3$. Two complex parameters for Σ, z and w, diagonalize the first and second fundamental forms if and only if $w = \pm z + c$, where $c \in \mathbb{C}$ is a constant.*

4. Classifying examples

In this section, we present an important class of examples of flat fronts whose hyperbolic Gauss maps are determined by a nonzero complex number. We call them *classifying examples*. We prove that if a flat front Σ corresponds to a harmonic function ϕ, then ϕ is linear if and only if Σ is locally congruent to one of the classifying examples or to the *peach front*, which is a flat front presented in [Kokubu et al. 2005].

Theorem 4. *For each $z_0 \in \mathbb{C}^* = \mathbb{C} \setminus \{0\}$, consider $g : \mathbb{C} \to \mathbb{C}^*$, the holomorphic function given by*

$$g(z) = e^{(\varepsilon \sinh z_0) z} \tag{4-1}$$

and $g^* = e^{2z_0} g$. Then there exists a flat front $\psi_{z_0} : \mathbb{C} \to \mathbb{H}^3$ whose singular set is

$$\mathcal{S} = \{ z \in \mathbb{C} \mid \mathrm{Re}[(\varepsilon \cosh z_0) z] = 0 \}, \qquad \varepsilon^2 = -1.$$

Also, the first and second fundamental forms of the flat immersion $\psi_{z_0} : \mathbb{C} \setminus \mathcal{S} \to \mathbb{H}^3$ can be written as in (3-1) and (3-2), where $\phi_{z_0}(z, \bar{z})$ is either a nonzero constant or

$$\phi_{z_0}(u, v) = -\mathrm{Re}[(\varepsilon \cosh z_0) z] = au + bv.$$

Proof. From the definitions of g and g^* and since $z_0 \neq 0$, it follows immediately that g and g^* are meromorphic nonconstant functions and $g \neq g^*$. Besides, we have

$$\frac{dg}{g - g^*} = -\frac{\varepsilon\, dz}{2e^{z_0}},$$

which implies that the conditions (1) and (2) of Theorem 1 are satisfied. On the other hand,

$$\frac{g_z g_z^*}{(g-g^*)^2} = \frac{e^{2z_0}(g_z)^2}{(1-e^{2z_0})^2 g^2} = \frac{\varepsilon^2 e^{2z_0} \sinh^2 z_0}{(1-e^{2z_0})^2} = -\frac{1}{4}.$$

By Theorem 1 and Proposition 2, there exists a flat front $\psi_{z_0} : \mathbb{C} \to \mathbb{H}^3$, given by $\psi_{z_0} = (\psi_1, \psi_2, \psi_3)$, with ψ_1, ψ_2, ψ_3 as in (3-4) and (3-5). Its first and second fundamental forms are as in (3-1) and (3-2). From the definition of g and g^* and Equations (3-4) and (3-10), it follows that $\phi_{z_0}(z, \bar{z})$ is either a nonzero constant if $\cosh z_0 = 0$, that is, $e^{z_0} = \pm\mathrm{i}$, or

$$\phi(z, \bar{z}) = -\operatorname{Re}(\varepsilon \cosh z_0 z) = au + bv, \qquad a, b \in \mathbb{R}.$$

In this last case, the singular set of ψ_{z_0} is the straight line given by

$$\mathscr{S} = \{z \in \mathbb{C} \mid \operatorname{Re}(\varepsilon \cosh z_0 z) = 0\}. \qquad \square$$

Choosing all the possible values for z_0 in Theorem 4, we obtain a family of examples that will provide the complete classification of the helicoidal flat surfaces. We will visualize the examples in the Poincaré ball model for $\mathbb{H}^3$.

(i) *Rotational flat fronts*. These flat fronts are obtained when $e^{2z_0} \in \mathbb{R}$. The hyperbolic cylinder (Figure 1, left) is obtained when $e^{2z_0} = -1$. When $e^{2z_0} < 0$ with $e^{2z_0} \neq -1$ we have the *hourglass* (Figure 1, center), and for $e^{2z_0} > 0$ we have the *snowman* (Figure 1, right).

(ii) *Conical flat fronts*. This flat front (Figure 2) is obtained when $e^{2z_0} = \pm\mathrm{i}$. In this case the invariance is only by the movement of translation.

(iii) *Properly helicoidal flat fronts*. The cases not mentioned above are invariant by the two movements, the rotational movement and the translation (Figure 3).

The class of examples obtained from Theorem 4 will be called *classifying examples*.

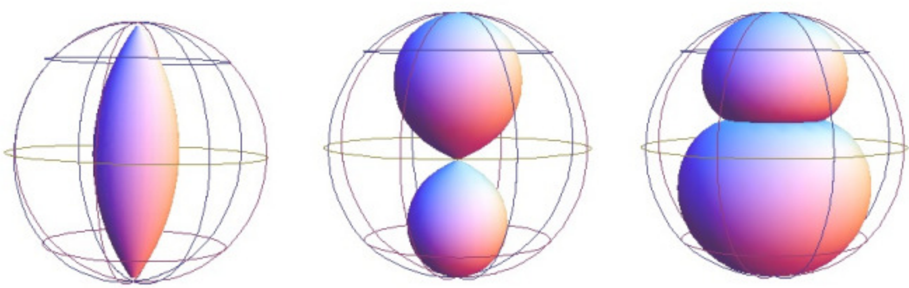


Figure 1. Rotational flat fronts: cylinder (left), hourglass (center), and snowman (right).

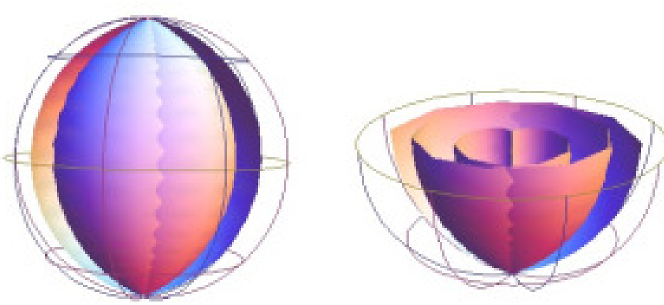


Figure 2. Conical flat fronts.

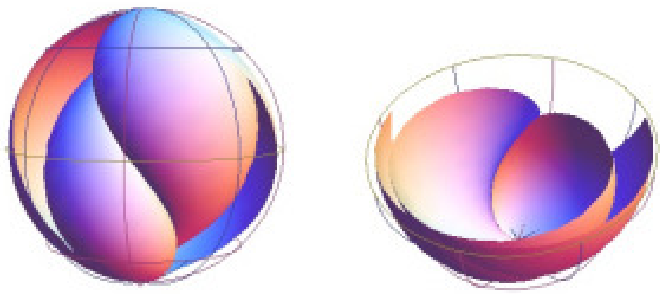


Figure 3. Properly helicoidal flat fronts.

Remark 5. The example given in [Kokubu et al. 2005], called the peach front (Figure 4), is a case where the hyperbolic Gauss maps satisfy $g^* = g - 1$, and it can be parametrized using Theorem 1 as

$$(\psi_1, \psi_2, \psi_3) = \left(\pm\frac{v}{2} - \frac{e^{\pm 2v}}{e^{\pm 2v}+1}, \mp\frac{u}{2}, \frac{e^{\pm v}}{e^{\pm 2v}+1}\right),$$

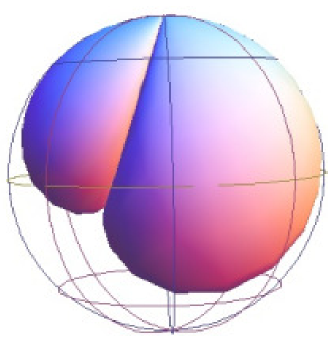


Figure 4. The peach front.

where the first and second fundamental forms are given as (3-1) and (3-2) with $\phi(u, v) = \pm v$. Observe that this value of ϕ can be viewed as

$$\phi(z, \bar{z}) = -\operatorname{Re}(\varepsilon \cosh z_0 z),$$

with $z_0 = 0$.

Theorem 6. *Let $\Sigma \subset \mathbb{H}^3$ be a flat front with a complex parameter $z = u + \mathrm{i}v$ that diagonalizes the first and second fundamental forms simultaneously as in (3-1) and (3-2). Then $\phi(u, v) = au + bv + c$ if and only if the corresponding flat front is locally congruent to one of the classifying examples or to a peach front.*

Proof. Since ϕ and $z = u + \mathrm{i}v$ determine the first and second fundamental forms, it is clear that any flat immersion such that $\phi(z, \bar{z}) = -\operatorname{Re}(\varepsilon \cosh z_0 z)$, for some $z_0 \in \mathbb{C}^*$, must be locally congruent to one of the classifying examples.

On the other hand, as we observed in Remark 5, the peach front has a parametrization that diagonalizes the first and second fundamental forms with $\phi(u, v)$ as above, with $z_0 = 0$. □

5. Characterization

In this section, we prove that a flat front in $\mathbb{H}^3$ is helicoidal if and only if it is locally congruent to one of the classifying examples presented in the previous section. Moreover, we obtain a complete classification of the helicoidal flat fronts in terms of their hyperbolic Gauss maps, as well as by means of linear harmonic functions. As a consequence of Theorem 6, we prove that any flat surface in $\mathbb{H}^3$ that corresponds to a linear harmonic function is locally congruent either to a helicoidal flat surface or to a peach front.

Theorem 7. *A flat front in $\mathbb{H}^3$ is helicoidal if and only if it is locally congruent to one of the classifying examples.*

We split our proof into two lemmas. The first lemma will establish that every classifying example has the geometric property that it is invariant by a helicoidal group of isometries, that is, it is a helicoidal flat front. In the second, we will show that, given any helicoidal flat front, there exists a rigid motion of $\mathbb{H}^3$ such that its hyperbolic Gauss maps satisfy $g^* = e^{2z_0} g$, where z_0 is a nonzero complex number. Once we establish these two lemmas, the proof of Theorem 7 will follow as a consequence of Proposition 2 and Theorem 6.

Lemma 8. *Every classifying example is a helicoidal flat front.*

Proof. The classifying examples were obtained by using the method of producing flat fronts given by Theorem 4. Given such a flat front, its hyperbolic Gauss maps g and $g*$ satisfy $g* = e^{2z_0} g$ and $g = e^{(\varepsilon \sinh z_0) z}$, where z_0 is a nonzero complex number, that is, $1-e^{z_0}\neq 0$. We want to obtain the immersion in $\mathbb{H}^3$ of the flat front, associated to g and $g*$, by using Theorem 1. Since $g* = e^{2z_0}g$, we have

$$g-g^* = (1-e^{2z_0})g. \tag{5-1}$$

Setting $g = Re^{\mathrm{i}v}$, it follows from (3-4) that

$$\xi = c\exp\Big(\frac{\log R+\mathrm{i}v}{1-e^{2z_0}}\Big).$$

From now on, we adopt the notation

$$\frac{1}{1-e^{2z_0}} = x_0+\mathrm{i}y_0.$$

Then we have

$$|\xi|^2 = |c|^2 e^{2(x_0\log R - y_0 v)}. \tag{5-2}$$

We can now obtain the flat front given by (3-5). Using (5-1) and (5-2), we have

$$\psi_1+\mathrm{i}\psi_2 = \left(1-\frac{|c|^4e^{4(x_0\log R-y_0v)}(x_0-\mathrm{i}y_0)}{|c|^4e^{4(x_0\log R-y_0v)}\big(x_0^2+y_0^2\big)+e^{(2\log R)}}\right)Re^{\mathrm{i}v}, \tag{5-3}$$

$$\psi_3 = \frac{|c|^2e^{2(x_0\log R-y_0v+\log R)}}{\big(x_0^2+y_0^2\big)|c|^4e^{4(x_0\log R-y_0v)}+e^{2\log R}}. \tag{5-4}$$

Simplifying (5-4), we have

$$\begin{aligned}\psi_3 &= \frac{|c|^2Re^{(2x_0\log R-2y_0v+\log R)}}{\big(x_0^2+y_0^2\big)|c|^4e^{4(x_0\log R-y_0v)}+e^{2\log R}}\\ &= \frac{|c|^2R}{\big(x_0^2+y_0^2\big)|c|^4e^{(2x_0\log R-2y_0v-\log R)}+e^{(-2x_0\log R+2y_0v+\log R)}}\\ &= \frac{|c|^2R}{\big(x_0^2+y_0^2\big)|c|^4e^{x}+e^{-x}},\end{aligned} \tag{5-5}$$

where $x = (2x_0-1)\log R - 2y_0v$. Using this fact, we rewrite (5-3):

$$\begin{aligned}&\psi_1+\mathrm{i}\psi_2\\ &\quad= \left(1-\frac{|c|^4e^{(4x_0\log R-4y_0v)}(x_0-\mathrm{i}y_0)}{\big(x_0^2+y_0^2\big)|c|^4e^{(4x_0\log R-4y_0v)}+e^{(2\log R)}}\right)Re^{\mathrm{i}v}\\ &\quad= \left(1-\frac{|c|^4e^{(2x_0\log R-2y_0v-\log R)}(x_0-\mathrm{i}y_0)}{\big(x_0^2+y_0^2\big)|c|^4e^{(2x_0\log R-2y_0v-\log R)}+e^{(-2x_0\log R+2y_0v+\log R)}}\right)Re^{\mathrm{i}v}\\ &\quad= \left(1-\frac{|c|^4(x_0-\mathrm{i}y_0)e^{x}}{\big(x_0^2+y_0^2\big)|c|^4e^{x}+e^{-x}}\right)Re^{\mathrm{i}v}.\end{aligned} \tag{5-6}$$

Now we want to prove that the immersed surface is invariant by the helicoidal group of isometries of $\mathbb{H}^3$. First, let us consider the case when $y_0 = 0$. We can see from (5-5) and (5-6) that this case corresponds to the rotational surfaces. On the other hand, when $y_0 \neq 0$, we can write

$$v = f(R) - \frac{x}{2y_0},$$

where $f(R) = \dfrac{2x_0 - 1}{2y_0} \log R$. With this notation, we obtain

$$\begin{aligned}
\psi_1 &= R\big(c_1(x)\cos f(R) - c_2(x)\sin f(R)\big),\\
\psi_2 &= R\big(c_1(x)\sin f(R) + c_2(x)\cos f(R)\big),
\end{aligned}$$

where c_1 and c_2 are real functions given by

$$\begin{aligned}
c_1(x) &= \frac{\big((x_0^2 + y_0^2 - x_0)|c|^4 e^x + e^{-x}\big)\cos\dfrac{x}{2y_0} + y_0|c|^4 e^x \sin\dfrac{x}{2y_0}}{(x_0^2 + y_0^2)|c|^4 e^x + e^{-x}},\\
c_2(x) &= -\frac{\big((x_0^2 + y_0^2 - x_0)|c|^4 e^x + e^{-x}\big)\sin\dfrac{x}{2y_0} + y_0|c|^4 e^x \cos\dfrac{x}{2y_0}}{(x_0^2 + y_0^2)|c|^4 e^x + e^{-x}}.
\end{aligned}$$

Using the notation $\psi_3 = Rc_3(x)$, we then have the following expression for the front:

$$(\psi_1, \psi_2, \psi_3)(R, x) = R\begin{pmatrix} \cos f(R) & -\sin f(R) & 0\\ \sin f(R) & \cos f(R) & 0\\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} c_1(x)\\ c_2(x)\\ c_3(x) \end{pmatrix}.$$

When $x_0 = \frac{1}{2}$, we have $f(R) = 0$, and consequently there is no rotational movement. On the other hand, if $x_0 \neq \frac{1}{2}$, we consider $f(R) = y$ and write

$$(\psi_1, \psi_2, \psi_3)(y, x) = \exp\Big(\frac{2y_0}{2x_0 - 1}y\Big)\begin{pmatrix} \cos y & -\sin y & 0\\ \sin y & \cos y & 0\\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} c_1(x)\\ c_2(x)\\ c_3(x) \end{pmatrix}.$$

This concludes the proof of the lemma. □

The second lemma will show that any helicoidal flat surface in $\mathbb{H}^3$ is congruent to a surface whose hyperbolic Gauss maps satisfy $g^* = cg$, where c is a complex number.

In order to do so, we will consider an approach closer to the one given in [Ripoll 1989]. We consider hyperbolic space $\mathbb{H}^3$ as a submanifold of the Lorentzian 4-space $\mathbb{L}^4$, endowed with coordinates (x_0, x_1, x_2, x_3) and the inner product $\langle\,,\rangle$ given by

$$\langle\,,\rangle = -dx_0^2 + dx_1^2 + dx_2^2 + dx_3^2.$$

The hyperbolic 3-space $\mathbb{H}^3$ will be the Riemannian 3-submanifold with sectional curvature -1 given by the set

$$\mathbb{H}^3 = \{(x_0, x_1, x_2, x_3) \in \mathbb{L}^4 \mid -x_0^2 + x_1^2 + x_2^2 + x_3^2 = -1,\ x_0 > 0\},$$

with the metric induced by $\mathbb{L}^4$.

We can see that the map

$$(x_0, x_1, x_2, x_3) \to \frac{1}{x_0 + x_3}(x_1, x_2, 1) \tag{5-7}$$

is an isometry between this model and the half-space model. Its inverse is given by

$$(y_1, y_2, y_3) \to \frac{1}{2y_3}\left(1 + \sum_{i=1}^{3} y_i^2,\ 2y_1,\ 2y_2,\ 1 - \sum_{i=1}^{3} y_i^2\right). \tag{5-8}$$

With these maps in mind we consider the helicoidal flat surfaces in $\mathbb{H}^3 \subset \mathbb{L}^4$. Let $O_1(4)$ be the orthogonal group in $\mathbb{L}^4$ given by all linear transformations that preserve $\langle\, ,\rangle$. Now consider $m_t \in O_1(4)$ given by the matrix

$$m_t = \begin{pmatrix} \cosh\beta t & 0 & 0 & \sinh\beta t \\ 0 & \cos\alpha t & -\sin\alpha t & 0 \\ 0 & \sin\alpha t & \cos\alpha t & 0 \\ \sinh\beta t & 0 & 0 & \cosh\beta t \end{pmatrix}.$$

Observe that m_t is a one-parameter subgroup of isometries of $\mathbb{H}^3$ given by a translation

$$\begin{pmatrix} \cosh\beta t & 0 & 0 & \sinh\beta t \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\beta t & 0 & 0 & \cosh\beta t \end{pmatrix}$$

along the geodesic $\gamma : -x_0^2 + x_3^2 = -1$, composed with the rotation

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha t & -\sin\alpha t & 0 \\ 0 & \sin\alpha t & \cos\alpha t & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Also, observe that the geodesic γ we are considering is the image of the y_3-axis by the map (5-8). One verifies that any orbit of m_t intersects the totally geodesic submanifold $P^2 = \{x_3 = 0\}$ just once. Thus, up to congruences, any surface invariant under m_t is generated by a curve in P^2.

In order to obtain the hyperbolic Gauss maps, we now consider $\mathbb{H}^3$ contained in the Lorentzian 4-space $\mathbb{L}^4$. We use the theory developed in [Gálvez et al. 2000], where there is a description of these maps using an identification between $\mathbb{L}^4$ and the set of 2×2 hermitian matrices, $\mathbb{H}\mathrm{erm}(2)$. To see this identification, let $\mathbb{N}^3$ be the half part of the light cone such that $x_0 > 0$,

$$\mathbb{N}^3 = \left\{(x_0, x_1, x_2, x_3) \in \mathbb{L}^4 \,\middle|\, -x_0^2 + x_1^2 + x_2^2 + x_3^2 = 0,\, x_0 > 0\right\}.$$

If we associate to each $v \in \mathbb{N}^3$ the half-line $[v]$, we obtain a partition of $\mathbb{N}^3$, and the ideal boundary $\mathbb{S}^2_\infty$ of $\mathbb{H}^3$ can be viewed as the quotient of $\mathbb{N}^3$ under the associated equivalence relation. Thus, the induced metric is well defined up to a scalar multiple, where $\mathbb{S}^2_\infty$ receives a natural conformal structure as the quotient $\mathbb{N}^3/\mathbb{R}^+$. In this approach, as we can see in [Gálvez et al. 2000], the hyperbolic Gauss maps of an immersion $\psi : S \to \mathbb{H}^3$ with unit normal vector field N are given by

$$g = [\psi + N] \quad \text{and} \quad g^* = [\psi - N]. \tag{5-9}$$

We use the identification between $\mathbb{L}^4$ and the set of 2×2 hermitian matrices $\mathbb{H}\mathrm{erm}(2)$, where the point (x_0, x_1, x_2, x_3) is identified with the matrix

$$\begin{pmatrix} x_0 + x_3 & x_1 + \mathrm{i}x_2 \\ x_1 - \mathrm{i}x_2 & x_0 - x_3 \end{pmatrix}. \tag{5-10}$$

Once we have the coordinates of $\psi + N$ and $\psi - N$ in $\mathbb{L}^4$, we find their corresponding matrices and write them as

$$\psi + N = \begin{pmatrix} A\bar{A} & A\bar{B} \\ \bar{A}B & B\bar{B} \end{pmatrix} \quad \text{and} \quad \psi - N = \begin{pmatrix} C\bar{C} & C\bar{D} \\ \bar{C}D & D\bar{D} \end{pmatrix}. \tag{5-11}$$

Therefore, we have the hyperbolic Gauss maps given by

$$g = \frac{A}{B} \quad \text{and} \quad g^* = \frac{C}{D} \tag{5-12}$$

(see [Gálvez et al. 2000] for more details).

With this approach, we are able to establish and prove the second lemma:

Lemma 9. *Let $\psi : \sigma \to \mathbb{H}^3$ be an immersion of a helicoidal flat surface in $\mathbb{H}^3$. Then there is a rigid motion of $\mathbb{H}^3$ such that its hyperbolic Gauss maps g and g^* satisfy $g^* = e^{2z_0} g$, where $z_0 \neq 0$ is a complex number.*

Proof. We start with a helicoidal surface immersed in $\mathbb{H}^3$. Then considering the half-space model for $\mathbb{H}^3$, there exists a rigid motion of $\mathbb{H}^3$ that takes the axis of the helicoidal surface into the y_3-axis. Then by considering the isometry (5-8), between the half-space model and $\mathbb{H}^3 \subset \mathbb{L}^4$, up to a rigid motion of $\mathbb{H}^3$, we may consider the immersion ψ of the helicoidal surface as

$$\psi(t, s) = m_t(\gamma(s)),$$

where γ is a curve in P^2 parametrized by arc length.

In order to describe the hyperbolic Gauss maps g and g^*, we need to obtain the maps $\psi + N$ and $\psi - N$. A normal unit vector field is given by

$$N = \frac{\boxtimes(\psi, \psi_t, \psi_s)}{|\boxtimes(\psi, \psi_t, \psi_s)|},$$

where $\boxtimes(\psi, \psi_t, \psi_s)$ is the Lorentzian vector product between ψ, ψ_t, ψ_s. If we write $\gamma(s) = (x_0(s), x_1(s), x_2(s), 0)$, we have $\psi_t(t, s) = m_t(v(s))$, with the vector v given by $v(s) = (0, -\alpha x_2(s), \alpha x_1(s), \beta x_0(s))$. This fact and the orthogonality of m_t enable us to conclude that

$$N(t, s) = m_t(\eta(s)),$$

where $\eta(s) = \dfrac{\boxtimes(\gamma(s), v(s), \gamma'(s))}{|\boxtimes(\gamma(s), v(s), \gamma'(s))|}$. Therefore

$$\psi + N = m_t(\gamma + \eta), \quad \psi - N = m_t(\gamma - \eta).$$

From (5-11)–(5-12), we have

$$g(s, t) = g_0(s) e^{(\beta + \mathrm{i}\alpha)t} \quad \text{and} \quad g^*(s, t) = g_0^*(s) e^{(\beta + \mathrm{i}\alpha)t}.$$

Now we see that g/g^* is a function only of the variable s, and as was proved in [Gálvez et al. 2000], g and g^* are holomorphic when the surface is flat. Therefore, g/g^* is a holomorphic function that depends only on one variable, which implies

$$g = \omega_0 g^*,$$

where $\omega_0 \in \mathbb{C}$ is a constant. It follows from (5-11)–(5-12) and the fact that ψ and N are orthogonal that $\omega_0 \neq 1$. Therefore, there exists $z_0 \in \mathbb{C}^*$ such that $g = e^{2z_0} g^*$. □

Proof of Theorem 7. One direction of the proof is given by Lemma 8, that is, every classifying example is a helicoidal flat front. Conversely, given any helicoidal flat surface in $\mathbb{H}^3$, it follows from Lemma 9 that it is congruent to a surface whose hyperbolic Gauss maps must satisfy $g = e^{2z_0} g^*$, where $z_0 \neq 0$. Moreover, using Proposition 2, we can choose a complex parameter $\eta = u + \mathrm{i}v$ such that (3-9) holds. Therefore, locally g is given by

$$g = e^{\varepsilon(\sinh z_0)\eta}.$$

Then it follows from (3-10) that ϕ must be linear. Theorem 6 implies that this helicoidal flat front is locally congruent to one of the classifying examples. □

As a consequence of Theorem 7, Proposition 2, Theorem 4 and the definition of the classifying examples, we have a complete classification of the helicoidal flat fronts in terms of their hyperbolic Gauss maps, determined by a nonzero complex number.

Theorem 10. *A flat front in $\mathbb{H}^3$ is helicoidal if and only if up to a rigid motion of $\mathbb{H}^3$, there exists a complex parameter η such that its hyperbolic Gauss maps g and g^* are meromorphic functions given by*

$$g = e^{\varepsilon(\sinh z_0)\eta} \quad \textit{and} \quad g^* = e^{2z_0} g,$$

where z_0 is a nonzero complex number and $\varepsilon^2 = 1$.

As a consequence of Theorems 4, 7 and 10, we get the following two results formulated in term of harmonic functions.

Theorem 11. *A flat front in $\mathbb{H}^3$ is helicoidal if and only if there exists a local parametrization by lines of curvature in a neighborhood of a nonsingular point, such that the first and second fundamental forms are given by (3-1) and (3-2), where*

$$\phi = au + bv + c \quad \textit{and} \quad (a, b, c) \neq (0, \pm 1, 0).$$

Theorem 12. *Let Σ be a flat front in $\mathbb{H}^3$ with a local parametrization in a neighborhood of a nonsingular and nonumbilic point, such that the first and second fundamental forms are diagonal and given by (3-1) and (3-2), where ϕ is a (euclidean) harmonic function. Then ϕ is linear, that is, $\phi = au + bv + c$ if and only if Σ is locally congruent either to a helicoidal flat front or to a peach front.*

References

[do Carmo and Dajczer 1983] M. do Carmo and M. Dajczer, "Rotation hypersurfaces in spaces of constant curvature", *Trans. Amer. Math. Soc.* **277**:2 (1983), 685–709. MR 85b:53055 Zbl 0518.53059

[Corro et al. 2010] A. V. Corro, A. Martínez, and F. Milán, "Complete flat surfaces with two isolated singularities in hyperbolic 3-space", *J. Math. Anal. Appl.* **366**:2 (2010), 582–592. MR 2011b:53028 Zbl 1190.53055

[Gálvez et al. 2000] J. A. Gálvez, A. Martínez, and F. Milán, "Flat surfaces in the hyperbolic 3-space", *Math. Ann.* **316**:3 (2000), 419–435. MR 2002b:53013 Zbl 1003.53047

[Kokubu et al. 2004] M. Kokubu, M. Umehara, and K. Yamada, "Flat fronts in hyperbolic 3-space", *Pacific J. Math.* **216**:1 (2004), 149–175. MR 2005f:53021 Zbl 1078.53009

[Kokubu et al. 2005] M. Kokubu, W. Rossman, K. Saji, M. Umehara, and K. Yamada, "Singularities of flat fronts in hyperbolic space", *Pacific J. Math.* **221**:2 (2005), 303–351. MR 2006k:53102 Zbl 1110.53044

[Ripoll 1989] J. B. Ripoll, "Helicoidal minimal surfaces in hyperbolic space", *Nagoya Math. J.* **114** (1989), 65–75. MR 91a:53015 Zbl 0699.53067

[Spivak 1979] M. Spivak, *A comprehensive introduction to differential geometry*, vol. 2, 2nd ed., Publish or Perish, Berkeley, CA, 1979. MR 82g:53003b Zbl 0439.53002

[Tenenblat 1998] K. Tenenblat, *Transformations of manifolds and applications to differential equations*, Pitman Monographs and Surveys in Pure and Applied Mathematics **93**, Longman, Harlow, 1998. MR 2002b:37111 Zbl 0912.58045

Received February 29, 2012. Revised April 20, 2012.

ANTONIO MARTÍNEZ
DEPARTAMENTO DE GEOMETRÍA Y TOPOLOGÍA
UNIVERSIDAD DE GRANADA
E-18071 GRANADA
SPAIN

amartine@ugr.es

JOÃO PAULO DOS SANTOS
DEPARTAMENTO DE MATEMÁTICA
UNIVERSIDADE DE BRASÍLIA
70910-900 BRASÍLIA
BRAZIL

joaopsantos@unb.br

KETI TENENBLAT
DEPARTAMENTO DE MATEMÁTICA
UNIVERSIDADE DE BRASÍLIA
70910-900 BRASÍLIA
BRAZIL

k.tenenblat@mat.unb.br

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.213

ON A GALOIS CONNECTION BETWEEN THE SUBFIELD LATTICE AND THE MULTIPLICATIVE SUBGROUP LATTICE

JOHN K. MCVEY

Given finite fields $F < E$, we present a collection of subgroups $C \leq E^\times$ and establish, to each C, a Galois connection between the intermediate field lattice $\mathscr{E} = \{L \mid F \leq L \leq E\}$ and C's subgroup lattice. Our main result is that, in all but an extremely limited and completely determined family, the closed subset of $\mathscr{E}$ is $\mathscr{E}$ itself, establishing a natural bijection between $\mathscr{E}$ and the lattice $\{L \cap C \mid L \in \mathscr{E}\}$. As an application, we use this bijection to calculate the set of degrees for the complex-valued irreducible representations of the split extension $C \rtimes \mathrm{Gal}(E/F)$.

1. Introduction

In §3 of [McVey 2004], generalizing results in §5 of [Riedl 1999], we worked towards (among other things) a better understanding of the groups $C \rtimes \mathrm{Gal}(E/F)$ for finite fields $F < E$, where $C < E^\times$ is the subgroup of order $|E^\times : F^\times|$. While working to generalize those results further, we discovered a Galois connection which itself is worthy of further study. This paper's intent is to record the Galois connection as well as the research that motivated its initial study. The primary assertion of the Main Theorem is that, but for a completely determined and rather limited family, the intermediate field lattice $\mathscr{E} = \{L \mid F \leq L \leq E\}$ is itself one of the two closed subsets in the Galois connection, thereby determining a canonical bijection between $\mathscr{E}$ and the other closed set $\{L \cap C \mid L \in \mathscr{E}\}$. As to the motivating research, we use this bijection to calculate the degrees of the irreducible complex representations of the aforementioned split extension $C \rtimes \mathrm{Gal}(E/F)$, showing every integer allowed by Itô's theorem is a degree.

2. Towards the Galois connection

Our focus in this paper is on monotone Galois connections. To avoid confusion between monotone and antitone connections, we define the term and present the basic relevant results. Two monotone nondecreasing functions $f : \mathscr{A} \to \mathscr{B}$ and $g : \mathscr{B} \to \mathscr{A}$ on partially ordered sets $(\mathscr{A}, \leq)$ and $(\mathscr{B}, \leq)$ form a *monotone Galois*

MSC2010: primary 06A15; secondary 20C15.
Keywords: Galois correspondence, lattice, character degree, finite field.

connection if

$$f(a) \le b \iff a \le g(b)$$

over all $a \in \mathscr{A}$ and $b \in \mathscr{B}$. The function f is the *lower adjoint* and g is the *upper adjoint*. The *closed sets* $\mathscr{A}_0$ and $\mathscr{B}_0$ of $\mathscr{A}$ and $\mathscr{B}$ respectively are defined by $\mathscr{A}_0 = g(\mathscr{B})$ and $\mathscr{B}_0 = f(\mathscr{A})$, and satisfy

$$\mathscr{A}_0 = \{a \in \mathscr{A} \mid g \circ f(a) = a\} = g \circ f(\mathscr{A}),$$
$$\mathscr{B}_0 = \{b \in \mathscr{B} \mid f \circ g(b) = b\} = f \circ g(\mathscr{B}).$$

The functions f and g are inverse bijections between the sets $\mathscr{A}_0$ and $\mathscr{B}_0$.

Turning now specifically to our setting of finite fields $F < E$, label by π the set of primes which divide $|F^\times| = |F| - 1$. The collection of groups to which the Galois connection applies consists of all subgroups $C \le E^\times$ for which the index $|E^\times : C|$ is a π-number (thus naturally generalizing results in [Riedl 1999] where hypotheses guaranteed $|E^\times : C| = |F^\times|$). Fixing a group C, the upper adjoint is very easy to describe; it is the function "intersect with C".

As to the lower adjoint, define the *F-closure* $\widehat{X}$ of a subset $X \subseteq E$ to be the smallest subfield of E which contains $X \cup F$. In other words, $\widehat{X}$ is the intersection of all fields L satisfying $X \cup F \subseteq L \le E$. It should be obvious that F-closure actually is a closure operator (i.e.,

$$\widehat{X} \supseteq X \quad \text{and} \quad \widehat{\widehat{X}} = \widehat{X}$$

over all subsets $X \subseteq E$), and that a Galois automorphism $\sigma \in \mathrm{Gal}(E/F)$ centralizes X if and only if it centralizes $\widehat{X}$. The partially ordered sets in our Galois connection are the lattices

$$\mathscr{E} = \{L \mid F \le L \le E \text{ is a field}\} \quad \text{and} \quad \mathscr{C} = \{D \mid D \le C \text{ is a group}\}, \tag{1}$$

ordered by inclusion. The functions $X \mapsto X \cap C$ and $X \mapsto \widehat{X}$ are obviously monotone. Given $D \in \mathscr{C}$ and $L \in \mathscr{E}$, and noting that $L \cap C = L^\times \cap C \in \mathscr{C}$, we have

$$\widehat{D} \subseteq L \iff D \subseteq L \iff D \subseteq L \cap C,$$

showing that $\widehat{\cdot}$ is a lower adjoint while $(\cdot) \cap C$ is an upper adjoint. Therefore, as $\mathscr{A}_0 = g(\mathscr{B})$, the closed subset of $\mathscr{C}$ is $\mathscr{C}_0 = \{L \cap C \mid L \in \mathscr{E}\}$.

We are now ready to state the Main Theorem. All but the last two sentences were proven in the above discussion. Those last two sentences are the true content of the theorem, and their proof is at the end of this section.

Main Theorem. *Let $F < E$ be finite fields and label by π the set of primes dividing $|F| - 1$. Let C be a subgroup of $E^\times$ whose index $|E^\times : C|$ is a π-number. Given the partially ordered sets defined by (1), the functions $\widehat{\cdot} : \mathscr{C} \to \mathscr{E}$ and $(\cdot) \cap C : \mathscr{E} \to \mathscr{C}$*

are respectively the lower and upper adjoints of a monotone Galois connection, and thus provide inverse bijections between the closed subsets $\mathscr{C}_0 \subseteq \mathscr{C}$ *and* $\mathscr{E}_0 \subseteq \mathscr{E}$. *The closed subset* $\mathscr{C}_0$ *of* $\mathscr{C}$ *is the lattice* $\mathscr{C}_0 = \{L \cap C \mid L \in \mathscr{E}\}$. *If* $|F|$ *is a Mersenne prime,* $|E : F|$ *is even, and* 4 *does not divide* $|C|$, *then the closed subset* $\mathscr{E}_0$ *of* $\mathscr{E}$ *is the set* $\mathscr{E}_0 = \mathscr{E} \setminus \{K\}$ *where* $|K : F| = 2$. *Otherwise,* $\mathscr{E} = \mathscr{E}_0$.

Our argument for the as yet unproven portion of the Main Theorem relies fundamentally on number theory. We ask the reader to recall Zsigmondy's prime theorem, as it is the foundation for what follows.

Theorem 1 [Zsigmondy 1892]. *Let* a, b, n *be positive integers and assume* a, b *are coprime and not both* 1. *Then,* $a^n - b^n$ *has a prime divisor which does not divide* $a^k - b^k$ *for integers* $0 < k < n$, *except when either*

$$n = 6 \text{ and } \{a, b\} = \{1, 2\} \quad \text{or} \quad n = 2 \text{ and } a + b \text{ is a } 2\text{-power.}$$

Aside from specifying $\{a, b\}$ as $\{q, 1\}$ with q a prime-power, the main point behind Corollary 2 is that the order of the quantifiers changed (from '$\exists$ prime $\forall\, k$' in Zsigmondy's theorem to '$\forall\, k\ \exists$ prime' in the corollary).

Corollary 2. *Let* $n > 1$ *be an integer and* q *a power of a prime. For each integer* k *with* $0 < k < n$, *there is a prime which divides* $q^n - 1$ *and not* $q^k - 1$, *except when* q *is a Mersenne prime and* $n = 2$. *Conversely, when* q *is a Mersenne prime, every prime dividing* $q^2 - 1$ *divides* $q - 1$.

Proof. As stated previously, Zsigmondy's theorem provides a universal prime (over all k) unless we are in one of the exceptional cases. First, assume $n = 6$ and $q = 2$, in which case $q^n - 1 = 2^6 - 1 = 63 = 3^2 \cdot 7$. It suffices to check that 3 divides none of $1 = 2^1 - 1$, $7 = 2^3 - 1$, and $31 = 2^5 - 1$, and that 7 divides neither $3 = 2^2 - 1$ nor $15 = 2^4 - 1$.

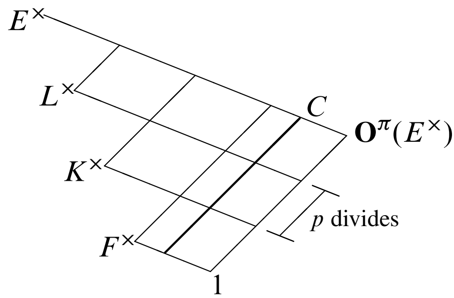
In the other exceptional case, $n = 2$ and $q + 1$ is a 2-power. However, Catalan's conjecture (proven in [Mihăilescu 2004]) says that the integer equation $x^a - y^b = 1$ with $a, b > 1$ only has the solution $3^2 - 2^3$. Because $q + 1$ is a 2-power already, q itself must be prime, hence a Mersenne prime.

As to the converse, when q is a Mersenne prime, the only prime dividing $q + 1$ is 2, which necessarily divides $(q + 1) - 2 = q - 1$. As $q^2 - 1 = (q + 1)(q - 1)$, the result follows. ☐

We now leave number theory and move to algebra proper. Our first algebraic goal is a lemma which shows how the number theory embedded in the previous corollary can be applied to finite fields.

Lemma 3. *Let* $F \le K \le L \le E$ *be finite fields. For the set* π *of prime divisors of* $|F^\times|$, *let* C *be a subgroup of* $E^\times$ *whose index is a* π*-number. If the prime* p *divides* $|L^\times|$ *and not* $|K^\times|$, *then* p *divides* $|L \cap C : K \cap C|$.

Proof. The following picture provides insight into this proof.



That p does not divide $|K^\times|$ implies p does not divide $|F^\times| = q-1$. Since

$$|L^\times : L \cap C| = |L^\times C : C| \text{ divides } |E^\times : C|,$$

which is a π-number and thus coprime to p, necessarily p divides $|L \cap C|$. As p does not divide $|K^\times|$, it also does not divide $|K \cap C|$. □

Theorem 4. *Let q be a prime-power, $e > 1$ an integer, and π the set of primes dividing $q-1$. Label $F = \mathbb{F}_q$ and $E = \mathbb{F}_{q^e}$, and let C be a subgroup of $E^\times$ whose index $|E^\times : C|$ is a π-number. Then, for all fields $F \leq L \leq E$, the equality $L = \widehat{L \cap C}$ holds, except when the following conditions are all satisfied.*

(1) *q is a Mersenne prime.*

(2) *e is even.*

(3) $L = \mathbb{F}_{q^2}$.

(4) 4 *does not divide* $|C|$.

When these simultaneously hold, $L \cap C = F \cap C$, so $\widehat{L \cap C} = F < L$.

Proof. Fix the field L. Obviously, the set $L \cap C$ is a subset both of $\widehat{L \cap C}$ and of C. Therefore,

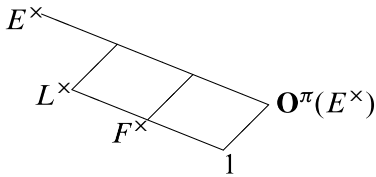
$$L \cap C \subseteq \widehat{L \cap C} \cap C \subseteq \hat{L} \cap C = L \cap C,$$

and we have equality throughout. Applying (the contrapositive of) Lemma 3 to $K = \widehat{L \cap C}$, the equality $K \cap C = L \cap C$ shows that every prime dividing $|L^\times|$ divides $|K^\times|$. Labelling $|L^\times| = q^n - 1$ and $|K^\times| = q^k - 1$, either $k = n$ (and we are done) or we are in the exceptional case of Corollary 2.

Henceforth, assume $n = 2$ and q is a Mersenne prime. As e is a multiple of n, it is even. Write $q = 2^p - 1$. Consequently,

- $|F^\times| = q - 1 = 2(2^{p-1} - 1)$, which has 2-part exactly 2, while
- $|L^\times| = q^2 - 1 = 2^p |F^\times|$, which has 2-part 2^{p+1}.

In particular, the $2'$-part of $|L^\times|$ is exactly the $2'$-part of $|F^\times|$.



We now split the argument as to whether or not 4 divides $|C|$. If 4 divides $|C|$, then because 4 also divides $|L^\times|$, it divides $|L \cap C|$. However, 4 does not divide $|F^\times|$, so $L \cap C \not\leq F^\times$. Accordingly, $F < \widehat{L \cap C} \leq L$, and $F = \mathbb{F}_q$ being a maximal subfield of $L = \mathbb{F}_{q^2}$ shows $L = \widehat{L \cap C}$.

When 4 does not divide $|C|$, the 2-part of $|L \cap C|$ divides $|F^\times|$. Generally, the $2'$-part of $|L \cap C|$ divides $|L^\times|_{2'} = |F^\times|_{2'}$. It follows that $|L \cap C|$ divides $|F^\times|$. Since $E^\times$ is cyclic, this shows $L \cap C \leq F^\times$, so $L \cap C = F \cap C$. □

With the above result in place, we use that $\mathscr{B}_0 = f \circ g(\mathscr{B})$ to conclude

$$\mathscr{E}_0 = \{\widehat{L \cap C} \mid L \in \mathscr{E}\}.$$

Meanwhile, the F-closure $\widehat{L \cap C}$ equals L but for the one exception $\mathbb{F}_{q^2}$ when q is Mersenne, e is even, and 4 fails to divide $|C|$. This finishes the proof of the Main Theorem.

3. Application to degrees

Our concluding section presents the computations for the character degree set of the split extension $C \rtimes \mathrm{Gal}(E/F)$ when $\mathscr{E} = \mathscr{E}_0$. We emphasize once more that this result was the principal impetus for our study of this Galois connection. All standard notations and conventions regarding character theory are taken from [Isaacs 1976]. The following generalizes Theorem 3.2 in [McVey 2004], and the proof here is fundamentally the same as is presented there, the main modification being the use of Theorem 4.

Theorem 5. *Fix a prime-power q and an exponent $1 < e \in \mathbb{Z}$, and label by F the field $\mathbb{F}_q$, by E the field $\mathbb{F}_{q^e}$, and by π the set of primes dividing $q-1$. Let $\Gamma = \mathrm{Gal}(E/F)$, and fix $C \leq E^\times$ under the assumption $|E^\times : C|$ is a π-number. If q is Mersenne and e is even, assume 4 divides $|C|$. Then, Γ normalizes C and*

$$\mathrm{cd}(C\Gamma) = \{n \mid n \text{ divides } e\}.$$

Proof. Because $E^\times$ is cyclic, every subgroup is characteristic. In particular, C is fixed (setwise) under every field automorphism of E, so Γ normalizes C. As C is cyclic, $\mathrm{Irr}(C)$ contains only linear characters and forms a cyclic group under multiplication. Let $\lambda \in \mathrm{Irr}(C)$ be a generator, noting λ is both faithful and a homomorphism. In summary, $\lambda(d_1) = \lambda(d_2)$ implies $d_1 = d_2$, $\lambda^m(d) = \lambda(d^m)$, and $\lambda^\tau(d^\tau) = \lambda(d)$ for all $d, d_1, d_2 \in C$, $m \in \mathbb{Z}$, and $\tau \in \Gamma$. Recalling $C \lhd C\Gamma$ is abelian, Itô's theorem says every degree in $\mathrm{cd}(C\Gamma)$ divides $|C\Gamma : C| = |\Gamma| = e$.

Conversely, fix a divisor n of e, and we will demonstrate an irreducible character of $C\Gamma$ whose degree is n. Let σ be a generator of Γ, and label $\Phi = \langle \sigma^n \rangle$, observing that $|\Gamma : \Phi| = n$. Let L be the fixed field for σ^n in the (usual) Galois correspondence for E over F. Hence, $\Phi = \mathrm{Gal}(E/L)$ and σ^n fixes the subgroup $L \cap C$ of C. For some generator $c \in C$, let $L \cap C = \langle c^m \rangle$.

We claim the stabilizer of λ^m in Γ is Φ. Given the claim, the stabilizer of λ^m in $C\Gamma$ is $C\Phi$, and λ^m extends to a character $\varphi \in \mathrm{Irr}(C\Phi)$ through for example Corollary 11.22 in [Isaacs 1976]. Also, φ induces irreducibly to $C\Gamma$ by Clifford correspondence. Therefore,

$$n = |\Gamma : \Phi| = |C\Gamma : C\Phi| = |C\Gamma : C\Phi|\varphi(1) = \varphi^{C\Gamma}(1) \in \mathrm{cd}(C\Gamma).$$

As n was an arbitrary divisor of e, we will have shown the result.

Given $\tau \in \Gamma$ and recalling λ is faithful, the equalities

$$(\lambda^m)^\tau(d) = \lambda^m(d^{\tau^{-1}}) = \lambda\big((d^{\tau^{-1}})^m\big) = \lambda\big((d^m)^{\tau^{-1}}\big)$$

and

$$\lambda^m(d) = \lambda(d^m)$$

imply that τ centralizes λ^m (the left ends are equal) if and only if τ^{-1} centralizes d^m for every $d \in C$ (the right ends are equal). The latter happens exactly when τ^{-1} centralizes $\langle c^m \rangle = L \cap C$, which occurs if and only if τ^{-1} centralizes $\widehat{L \cap C}$. As $L = \widehat{L \cap C}$ (Theorem 4), this is equivalent to $\tau^{-1} \in \mathrm{Gal}(E/L) = \Phi$. $\square$

In closing, we would be remiss in not mentioning an application of Theorem 5 to a remark made in [Lewis 2001]. For the subsequent, we use the notation of [Lewis 2001]. In the paragraph preceding Lemma 3.4, Dr. Lewis made the comment that "…every divisor of m occurs in $\mathrm{cd}(G/V)$", but that particular conclusion was superfluous to Lemma 3.4, so it went unproven. Reading through the first two and a half paragraphs of that proof, V can be viewed as the additive group of the field $\mathbb{F}_{q^m}$, K/Z acts on $\mathbb{F}_{q^m}$ by multiplication as if it were a subgroup of $\mathbb{F}_{q^m}^\times$, and the quotient H/K behaves as a Galois group. Lastly, the hypotheses to Example 2.4 imply m is coprime to $(q^m - 1)/(q - 1)$. Hence, our result applies to the group H/Z, and Lewis' claim about the degrees is an immediate corollary of Theorem 5 and the relations $G/V \cong H$ and $\mathrm{cd}(H) = \mathrm{cd}(H/Z)$.

References

[Isaacs 1976] I. M. Isaacs, *Character theory of finite groups*, Pure and Applied Mathematics **69**, Academic Press, New York, 1976. MR 57 #417 Zbl 0337.20005

[Lewis 2001] M. L. Lewis, "Solvable groups whose degree graphs have two connected components", *J. Group Theory* **4**:3 (2001), 255–275. MR 2002g:20015 Zbl 0998.20009

[McVey 2004] J. K. McVey, "Prime divisibility among degrees of solvable groups", *Comm. Algebra* **32**:9 (2004), 3391–3402. MR 2005i:20012 Zbl 1077.20009

[Mihăilescu 2004] P. Mihăilescu, "Primary cyclotomic units and a proof of Catalan's conjecture", *J. Reine Angew. Math.* **572** (2004), 167–195. MR 2005f:11051 Zbl 1067.11017

[Riedl 1999] J. M. Riedl, "Character degrees, class sizes, and normal subgroups of a certain class of p-groups", *J. Algebra* **218**:1 (1999), 190–215. MR 2000f:20025 Zbl 0932.20015

[Zsigmondy 1892] K. Zsigmondy, "Zur Theorie der Potenzreste", *Monatsh. Math. Phys.* **3**:1 (1892), 265–284. MR 1546236 JFM 24.0176.02

Received June 2, 2012. Revised December 10, 2012.

JOHN K. MCVEY
DEPARTMENT OF MATHEMATICS
KENT STATE UNIVERSITY
MATHEMATICS AND COMPUTER SCIENCE BUILDING 233
SUMMIT STREET
KENT, OH 44242
UNITED STATES

jmcvey@math.kent.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.221

SOME CHARACTERIZATIONS OF CAMPANATO SPACES VIA COMMUTATORS ON MORREY SPACES

SHAOGUANG SHI AND SHANZHEN LU

We give some creative characterizations of Campanato spaces via the boundedness of commutators associated with the Calderón–Zygmund singular integral operator by some new methods instead of the sharp maximal function theorem.

1. Introduction and main results

Let $-n/p \le \beta < 1$ and $1 \le p < \infty$. A locally integrable function f is said to belong to the Campanato spaces $C^{p,\beta}(\mathbb{R}^n)$ if

$$\|f\|_{C^{p,\beta}(\mathbb{R}^n)} = \sup_Q \|f\|_{C^{p,\beta}(Q)} := \sup_Q \frac{1}{|Q|^{\beta/n}} \left(\frac{1}{|Q|}\int_Q |f - f_Q|^p \, dx\right)^{1/p} < \infty,$$

where $f_Q = \frac{1}{|Q|}\int_Q f(x)\,dx$, Q denotes any cube contained in $\mathbb{R}^n$ and $|Q|$ is the Lebesgue measure of Q.

Campanato spaces are useful tools in the regularity theory of PDEs as a result of their better structures, which allow us to give an integral characterization of the spaces of Hölder continuous functions. This leads to a generalization of the classical Sobolev embedding theorem (see, e.g., [Lemarié-Rieusset 2007; Lu 1995; 1998]). It is also well known that $C^{1,1/p-1}$ is the dual space of Hardy space $H^p(\mathbb{R}^n)$ when $0 < p < 1$ (see [Triebel 1992]). For a recent account of the theory on $C^{p,\beta}(\mathbb{R}^n)$, we refer the reader to [Duong et al. 2007; Lin et al. 2011; Nakai 2006; Yang et al. 2010].

It's obvious that $\beta = 0$ implies $C^{p,0}(\mathbb{R}^n) = BMO(\mathbb{R}^n)$ with the norm

$$\|f\|_{BMO(\mathbb{R}^n)} = \sup_Q \frac{1}{|Q|}\int_Q |b - b_Q|\,dx. \tag{1-1}$$

Corresponding author: Shaoguang Shi. This work was partially supported by NSF of China (Grant Nos. 10931001 and 11271175), NSF of Shandong Province (Grant No. ZR2012AQ026), the Key Laboratory of Mathematics and Complex System (Beijing Normal University), Ministry of Education, China.

MSC2010: primary 42B20; secondary 42B25.

Keywords: Calderón–Zygmund singular integral, commutator, Morrey space, weight.

When $0<\beta<1$ and $1\le p<\infty$, we have $C^{p,\beta}(\mathbb{R}^n)=\mathrm{Lip}_\beta(\mathbb{R}^n)$ (see [DeVore and Sharpley 1984; Janson et al. 1983]) with the equivalent norm

$$\|f\|_{\mathrm{Lip}_\beta(\mathbb{R}^n)} \approx \sup_Q \frac{1}{|Q|^{1+\beta/n}} \int_Q |f-f_Q| \approx \sup_Q \left(\frac{1}{|Q|^{1+q\beta/n}} \int_Q |f-f_Q|^q\right)^{1/q}, \tag{1-2}$$

where $1\le q\le\infty$ and $\mathrm{Lip}_\beta(\mathbb{R}^n)$ is the Lipschitz functional space.

When $-n/p\le\beta<0$, there are several stages in the study of $C^{p,\beta}(\mathbb{R}^n)$. Let Ω be a connected open set of $\mathbb{R}^n$. Denote by $\overline{\Omega}$ the closure of Ω, and by $\operatorname{diam}\Omega$ the diameter of Ω. For any $x_0\in\mathbb{R}^n$ and $l\in(0,\infty)$, set $B(x_0,l)=\{x\in\mathbb{R}^n:|x-x_0|<l\}$ and $\Omega(x_0,l)=B(x_0,l)\cap\Omega$. The following space was first introduced by Morrey [1938] to investigate the local behavior of solutions to the second order elliptic PDE

$$\|f\|_{M^{p,\beta}(\Omega)} = \sup_{\substack{x_0\in\overline{\Omega}\\ l\in(0,\operatorname{diam}\Omega)}} \frac{1}{|\Omega(x_0,l)|^{\beta/n}} \left(\frac{1}{|\Omega(x_0,l)|}\int_{\Omega(x_0,l)} |f|^p\right)^{1/p},$$

where $f\in L^p_{\mathrm{loc}}(\Omega)$, $1\le p<\infty$ and $-n/p\le\beta<0$. When $\Omega=\mathbb{R}^n$, $M^{p,\beta}(\mathbb{R}^n)$ is the classical Morrey space, whose norm is defined by

$$\|f\|_{M^{p,\beta}(\mathbb{R}^n)} = \sup_Q \|f\|_{M^{p,\beta}(Q)} = \sup_Q \frac{1}{|Q|^{\beta/n}} \left(\frac{1}{|Q|}\int_Q |f|^p\right)^{1/p}.$$

$M^{p,\beta}(\mathbb{R}^n)$ is an expansion of $L^p(\mathbb{R}^n)$ in the sense that $M^{p,-n/p}(\mathbb{R}^n)=L^p(\mathbb{R}^n)$. Similarly, for $1\le p<\infty$, $-n/p\le\beta<0$, a function $f\in L^p_{\mathrm{loc}}(\Omega)$ is said to belong to the Campanato space $C^{p,\beta}(\Omega)$ if

$$\|f\|_{C^{p,\beta}(\Omega)} = \sup_{\substack{x_0\in\overline{\Omega}\\ l\in(0,\operatorname{diam}\Omega)}} \frac{1}{|\Omega(x_0,l)|^{\beta/n}} \left(\frac{1}{|\Omega(x_0,l)|}\int_{\Omega(x_0,l)} |f-f_{\Omega(x_0,l)}|^p\right)^{1/p} < \infty.$$

Campanato [1963] proved that, if $\operatorname{diam}\Omega<\infty$ and there exists a positive constant C such that

$$|\Omega(x_0,l)|\ge Cl^n, \tag{1-3}$$

for every $x_0\in\overline{\Omega}$ and $l\in(0,\operatorname{diam}\Omega)$, then

$$M^{p,\beta}(\Omega)=C^{p,\beta}(\Omega). \tag{1-4}$$

(For more accounts about (1-4), see [Rupflin 2008], for example.) Throughout this paper, the letter C stands for a positive constant which may vary from line to line. When $\operatorname{diam}\Omega=\infty$ (i.e., Ω is unbounded, as when $\Omega=\mathbb{R}^n$, for example), Sakamoto and Yabuta [1999] pointed out that when $1\le p<\infty$ and $\beta\in[-n/p,0)$, $C^{p,\beta}(\mathbb{R}^n)$

is equivalent to $M^{p,\beta}(\mathbb{R}^n)$. But Lin [2009] gave a counterexample to verify that when $1 \le p < \infty$ and $\beta \in [-n/p, 0)$, we have

$$M^{p,\beta}(\mathbb{R}^n) \subsetneq C^{p,\beta}(\mathbb{R}^n), \tag{1-5}$$

which implies that the statement in [Sakamoto and Yabuta 1999] may be inaccurate. More precisely, on account of the remark above, we have

$$C^{p,\beta}(\mathbb{R}^n) \begin{cases} = BMO(\mathbb{R}^n) & \text{for } \beta = 0, \\ = \mathrm{Lip}_\beta(\mathbb{R}^n) & \text{for } 0 < \beta < 1, \\ \supset M^{p,\beta}(\mathbb{R}^n) & \text{for } -n/p < \beta < 0. \end{cases}$$

Let T be a linear operator and b a suitable function. For a proper function f, the commutator T_b is defined by

$$T_b(f) := bTf - T(bf).$$

In this paper, we give some characterizations of $C^{p,\beta}(\mathbb{R}^n)$ in terms of the boundedness of T_b, where T is the Calderón–Zygmund singular integral operator

$$Tf(x) = \text{p.v.}\int_{\mathbb{R}^n} K(x-y)f(y)\,dy;$$

here $K \in C^\infty(S^{n-1})$ is a Calderón–Zygmund kernel that satisfies

$$K(x) = K(x/|x|)/|x|^n \quad \text{for } |x| \neq 0 \tag{1-6}$$

and

$$\int_{S^{n-1}} K = 0. \tag{1-7}$$

For more on the theory of the Calderón–Zygmund singular integral operator T, see [Grafakos 2004; Janson 1978; Lu 2011; Lu et al. 2007; Stein 1970], for example.

There are many classical works about the characterizations of Campanato spaces by the boundedness of T_b on Lebesgue spaces. Coifman, Rochberg and Weiss [Coifman et al. 1976] gave a characterization of $BMO(\mathbb{R}^n)$ in terms of the commutator T_b:

$$b \in BMO(\mathbb{R}^n) \iff T_b : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n) \text{ if } 1 < p < \infty.$$

Janson [1978] gave a characterization of $\mathrm{Lip}_\beta(\mathbb{R}^n)$ by the (L^p, L^q)-boundedness of the commutator T_b: If $0 < \beta < 1$, then

$$b \in \mathrm{Lip}_\beta(\mathbb{R}^n) \iff T_b : L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n) \text{ if } 1 < p < q < \infty \text{ and } 1/q = 1/p - \beta/n.$$

Paluszyński [1995] gave a new characterization of $\mathrm{Lip}_\beta(\mathbb{R}^n)$ by the $(L^p, \dot{F}^{\beta}_{p,\infty})$-boundedness of the commutator T_b: If $0 < \beta < 1$, then

$$b \in \mathrm{Lip}_\beta(\mathbb{R}^n) \iff T_b : L^p(\mathbb{R}^n) \to \dot{F}^{\beta}_{p,\infty}(\mathbb{R}^n) \text{ if } 1 < p < \infty,$$

where $\dot{F}^{\beta}_{p,\infty}(\mathbb{R}^n)$ is the homogeneous Triebel–Lizorkin space with the equivalent norm

$$\|f\|_{\dot{F}^{\beta}_{p,\infty}(\mathbb{R}^n)} \approx \left\| \sup_Q \frac{1}{|Q|^{1+\beta/n}} \int_Q |b - b_Q| \right\|_{L^p}.$$

As a natural extension of Lebesgue space, it is interesting to know whether Campanato spaces can be characterized by the boundedness of T_b on Morrey spaces.

Ding [1997] characterized $BMO(\mathbb{R}^n)$ by the $(M^{p,\beta}(\mathbb{R}^n), M^{p,\beta}(\mathbb{R}^n))$-boundedness of T_b:

$$b \in BMO(\mathbb{R}^n) \iff T_b : M^{p,\beta}(\mathbb{R}^n) \to M^{p,\beta}(\mathbb{R}^n) \text{ if } 1 < p < \infty,\ -n/p \le \beta < 0.$$

In the rest of this paper, we shall establish the characterizations of other cases of Campanato spaces — namely, $\mathrm{Lip}_\beta(\mathbb{R}^n)$ for $0 < \beta < 1$ and $M^{p,\beta}(\mathbb{R}^n)$ for $-n/p \le \beta < 0$ — using certain boundedness properties of T_b on Morrey spaces.

Now, we formulate our first result as follows:

Theorem 1.1. *Let* $1 < p < \infty,\ 0 < \alpha < 1,\ -n/p \le \beta < 0,\ 1 + p\beta/n < p/q$, $1/q = 1/p - \alpha/n$ *and* $\tilde{\beta} = (q-p)/p + q\beta/n$. *The following statements are equivalent*:

(1) $b \in \mathrm{Lip}_\alpha(\mathbb{R}^n)$.

(2) T_b *is a bounded operator from* $M^{p,\beta}(\mathbb{R}^n)$ *to* $M^{q,\tilde{\beta}}(\mathbb{R}^n)$.

We say that a nonnegative function f belongs to the reverse Hölder class RH^r if for any $Q \subset \mathbb{R}^n$ and $1 < r < \infty$ we have

$$\left(\frac{1}{|Q|} \int_Q |f|^r \, dx \right)^{1/r} \le \frac{C}{|Q|} \int_Q |f| \, dx.$$

When $r = \infty$, we say that $f \in RH^\infty$ if $f \in L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$ and there exists a constant C such that

$$\|f\|_{L^\infty(Q)} := \sup_{Q \ni x} |f(x)| \le \frac{C}{|Q|} \int_Q |f| \, dx. \tag{1-8}$$

For $1 < r < \infty$, it is easy to see that $RH^\infty = \bigcup_{r>1} RH^r$. Reverse Hölder classes contain many kinds of functions. For example, if $P(x)$ is a polynomial and $\gamma > 0$, then $f(x) = |P(x)|^\gamma \in RH^\infty$(see [Fefferman 1983]). (For more theories about RH^r, see [Cruz-Uribe and Neugebauer 1995; Harboure et al. 1998], for example.)

Theorem 1.2. *Assume* $\max\{1, n/(1-\beta)\} < p < \infty$, $-n/p \le \beta < 0,\ 1 < p_i < \infty$ $(i = 1, 2)$, $p_1 \in \mathbb{N}$ *even*, $-n/p_i \le \beta_i < 0,\ 1/p = 1/p_1 + 1/p_2$ *and* $\beta = \beta_1 + \beta_2$. *If* Ω *satisfies* (1-3) *and* $\mathrm{diam}\, \Omega < \infty$, *the following statements are equivalent*:

(1) $b \in M^{p_1,\beta_1}(\Omega)$.

(2) *If* $b \in RH^{\infty}$, T_b *is a bounded operator from* $M^{p_2,\beta_2}(\Omega)$ *to* $M^{p,\beta}(\Omega)$.

The advantage of using the assumption $\operatorname{diam}\Omega < \infty$ lies in the fact that the equivalent norm of (1-4) is used in the proof of Theorem 1.2. If $\Omega = \mathbb{R}^n$, we can obtain the following characterizations of Campanato spaces:

Theorem 1.3. *Assume* $\max\{1, n/(1-\beta)\} < p < \infty$, $-n/p \le \beta < 0$, $1 < p_i < \infty$ $(i = 1, 2)$, $p_1 \in \mathbb{N}$ *even*, $-n/p_i \le \beta_i < 0$, $1/p = 1/p_1 + 1/p_2$ *and* $\beta = \beta_1 + \beta_2$. *The following statements are equivalent*:

(1) $b \in C^{p_1,\beta_1}(\mathbb{R}^n)$.

(2) T_b *is a bounded operator from* $M^{p_2,\beta_2}(\mathbb{R}^n)$ *to* $C^{p,\beta}(\mathbb{R}^n)$ *if* b *further satisfies that there exists a constant* $C > 0$ *such that for any* $Q \subset \mathbb{R}^n$,

$$\sup_Q |b - b_Q| \le \frac{C}{|Q|}\int_Q |b - b_Q|. \tag{1-9}$$

Remark 1. Inequalities (1-8) and (1-9) can be thought of as a form of mean value equality. Besides polynomial functions, mean value equalities also characterize harmonic functions (see [Gilbarg and Trudinger 1983]).

Remark 2. Solutions to a large class of elliptic second order PDEs satisfy the mean value inequality. Therefore, Theorem 1.2 and Theorem 1.3 can give characterizations of the space of solutions to some second order elliptic PDEs. Take Laplace's equation, for example. If b is a solution to the equation

$$\Delta u = 0, \tag{1-10}$$

where Δ is the Laplace operator and u is a function defined on the bounded domain $\Omega \subset \mathbb{R}^n$, then b satisfies (1-9); see [Gilbarg and Trudinger 1983, Theorem 2.1]. Therefore, if the commutator T_b associated to b is bounded from $M^{p_2,\beta_2}(\mathbb{R}^n)$ to $C^{p,\beta}(\mathbb{R}^n)$, then the space of solutions to (1-10) is the Campanato space $C^{p_1,\beta_1}(\mathbb{R}^n)$.

Remark 3. We emphasize that the methods in dealing with $C^{p,\beta}$ when $\beta < 0$ are quite different from that of $\beta \ge 0$, and there are essential difficulties in establishing the characterizations of Campanato spaces on Morrey spaces when $\beta < 0$. Therefore, we set up Theorem 1.3 under the condition that the symbol of the commutator satisfies the mean value inequality. Condition (1-9) in Theorem 1.3 was intrinsic to the proof of the converse characterizations of $C^{p_1,\beta_1}(\mathbb{R}^n)$. Of course, there are essential differences between the ideas in the proof of Theorem 1.2 and Theorem 1.3 and that of [Janson 1978] and [Paluszyński 1995], where the sharp maximal function theorem were used.

Our theorems provide natural and intrinsic characterizations of Campanato spaces on Morrey spaces. It is also worth pointing out that our paper is the first work on the problem of commutators whose symbol belongs to Morrey spaces. Our viewpoints will shed some new lights on characterizations of Campanato spaces via commutators formed by other operators on Morrey spaces, such as fractional integrals, oscillatory integral operators and Hardy–Littlewood–Paley operators. Besides Euclidean space, characterizations of Campanato spaces on other spaces can similarly be considered, such as on homogeneous groups. Partly inspired by [Janson 1978] and [Paluszyński 1995], we prove Theorems 1.1–1.3 in Section 2.

2. Proof of the main results

For the proofs we need some lemmas about the estimates of operators on Morrey spaces.

Lemma 2.1 [Chiarenza and Frasca 1987]. *Let* $1 < p < n/\alpha$, $0 < \alpha < n$, $1/q = 1/p - \alpha/n$, $0 < 1 + p\beta/n < p/q$, $-n/p \le \beta < 0$ *and* $\tilde{\beta} = (q-p)/p + q\beta/n$. *Then the fractional integral operator*

$$I_\alpha f(x) = \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}}\, dy$$

is bounded from $M^{p,\beta}(\mathbb{R}^n)$ *to* $M^{q,\tilde{\beta}}(\mathbb{R}^n)$.

Lemma 2.2 [Komori and Shirai 2009]. *Let* $1 < p < \infty$ *and* $-n/p \le \beta < 0$. *Then* T *is bounded from* $M^{p,\beta}(\mathbb{R}^n)$ *to* $M^{p,\beta}(\mathbb{R}^n)$.

Proof of Theorem 1.1. (1) $\Rightarrow$ (2). Together, (1-2) and (1-6) imply

$$\begin{aligned} |T_b f(x)| &\le \int_{\mathbb{R}^n} |b(x)-b(y)|\,|K(x-y)|\,|f(y)|\, dy \\ &\le \int_{\mathbb{R}^n} \frac{|f(y)|}{|x-y|^{n-\alpha}}\, dy \le I_\alpha(|f(x)|). \end{aligned}$$

Therefore, T_b is bounded from $M^{p,\beta}(\mathbb{R}^n)$ to $M^{q,\tilde{\beta}}(\mathbb{R}^n)$ by Lemma 2.1.

(2) $\Rightarrow$ (1). The proof consists of the construction of a proper commutator. We follow [Janson 1978] in choosing $z_0 \ne 0$ and $\delta > 0$ such that $1/K(z)$ can be expressed in the neighborhood $|z - z_0| < \sqrt{n}\delta$ as the absolute convergent Fourier series

$$\frac{1}{K(z)} = \sum a_n e^{i v_n \cdot z},$$

where the exact form of the vectors v_n is irrelevant. Set $z_1 = \delta^{-1} z_0$. If $|z - z_1| < \sqrt{n}$, it follows from (1-6) that

$$\frac{1}{K(z)} = \frac{\delta^{-n}}{K(\delta z)} = \delta^{-n} \sum a_n e^{i v_n \cdot \delta z}. \tag{2-1}$$

Choose now any cube $Q = Q(x_0, r)$. Set $y_0 = x_0 - r z_1$ and $Q' = Q(y_0, r)$. Thus, if $x \in Q$ and $y \in Q'$,

$$\left|\frac{x-y}{r}-z_1\right| \le \left|\frac{x-x_0}{r}-\frac{y-y_0}{r}\right| \le \sqrt{n}.$$

Denoting $s(x)=\operatorname{sgn}(b(x)-b_{Q'})$, by (2-1) we have

$$\begin{aligned}
\int_Q |b(x)-b_{Q'}|\,dx &= \int_Q (b(x)-b_{Q'})s(x)\,dx \\
&= \frac{1}{|Q'|}\int_Q\int_{Q'}(b(x)-b(y))s(x)\,dy\,dx \\
&= \frac{1}{r^n}\iint_{\mathbb{R}^n}(b(x)-b(y))s(x)\frac{r^nK(x-y)}{K((x-y)/r)}\chi_Q(x)\chi_{Q'}(y)\,dy\,dx \\
&= C\sum a_n \iint_{\mathbb{R}^n}(b(x)-b(y))K(x-y)e^{i(\delta/r)v_n\cdot x} \\
&\qquad\qquad \times s(x)\chi_Q(x)e^{-i(\delta/r)v_n\cdot y}\chi_{Q'}(y)\,dy\,dx.
\end{aligned}$$

Taking

$$g_n(y)=e^{-i(\delta/r)v_n\cdot y}\chi_{Q'}(y) \quad\text{and}\quad h_n(x)=e^{i(\delta/r)v_n\cdot x}s(x)\chi_Q(x),$$

we obtain

$$\begin{aligned}
\int_Q |b(x)-b_{Q'}|\,dx &= C\sum a_n\iint_{\mathbb{R}^n}(b(x)-b(y))K(x-y)g_n(y)h_n(x)\,dy\,dx \\
&= C\sum a_n\int_{\mathbb{R}^n}T_bg_n(x)h_n(x)\,dx \\
&\le C\sum |a_n|\int_{\mathbb{R}^n}|T_bg_n(x)|\,|h_n(x)|\,dx \\
&\le C\sum |a_n|\int_{\mathbb{R}^n}|T_bg_n(x)|\,dx.
\end{aligned}$$

Applying the Hölder inequality to $\int_{\mathbb{R}^n}|T_bg_n(x)|\,dx$, we have

$$\begin{aligned}
\int_{\mathbb{R}^n}|T_bg_n(x)|\,dx &\le |Q|^{1+(\alpha+\beta)/n}\|T_bg_n\|_{M^{q,\tilde\beta}(\mathbb{R}^n)} \\
&\le C|Q|^{1+(\alpha+\beta)/n}\|g_n\|_{M^{p,\beta}(\mathbb{R}^n)} \\
&\le C|Q|^{1+\alpha/n},
\end{aligned}$$

since $\|g_n\|_{M^{p,\beta}(\mathbb{R}^n)}=|Q|^{-\beta/n}$. Thus we have obtained

$$\frac{1}{|Q|^{1+\alpha/n}}\int_Q|b(x)-b_{Q'}|\,dx \le C,$$

which completes the proof of Theorem 1.1 by (1-2). □

Theorem 1.2 is a restatement of Theorem 1.3 when $C^{p,\beta}$ spaces over domains with finite volume, so we give the proof of Theorem 1.3 first. Again, we begin with some lemmas that are essential to our analysis.

Lemma 2.3. *Let p, p_1, p_2, β, β_1, β_2 and b be the same as in Theorem 1.3. Then*

$$\|(b-b_Q)f\chi_Q\|_{L^p(\mathbb{R}^n)} \le |Q|^{1/p+\beta/n}\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\|f\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}.$$

This follows from Hölder's inequality:

$$\begin{aligned}\|(b-b_Q)f\chi_Q\|_{L^p(\mathbb{R}^n)} &\le \left(\int_Q |b-b_Q|^{p_1}\right)^{1/p_1}\left(\int_Q |f|^{p_2}\right)^{1/p_2}\\ &\le |Q|^{1/p(1+p\beta/n)}\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\|f\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}.\end{aligned}$$

Lemma 2.4. *Suppose that $Q_* \subset Q$ and $b \in C^{p_1,\beta_1}(\mathbb{R}^n)$ with $1 < p_1 < \infty$ and $-n/p_1 \le \beta_1 < 0$. Then the following estimate holds*:

$$|b_{Q_*} - b_Q| \le C\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}|Q_*|^{\beta_1/n}. \tag{2-2}$$

We divide the proof into two cases.

Case 1: Suppose $Q_* \subset Q \subseteq 2Q_*$. Hölder's inequality yields

$$\begin{aligned}|b_{Q_*} - b_Q| &\le \frac{1}{|Q_*|}\int_{Q_*}|b-b_Q| + \frac{1}{|Q|}\int_Q |b-b_Q|\\ &\le C\frac{1}{|Q|}\int_Q |b-b_Q| \le C\left(\int_Q |b-b_Q|^{p_1}\right)^{1/p_1}|Q|^{-1/p_1}\\ &\le C\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}|Q_*|^{\beta_1/n}.\end{aligned}$$

Case 2: Suppose $2Q_* \subset Q$. Choose a sequence of nested cubes

$$Q_* =: Q_1 \subset Q_2 \subset \cdots Q =: Q_{m+1},$$

with $|Q_{i+1}| = 2^n|Q_i|$ for $1 \le i \le m$. By the results of Case 1, we have

$$\begin{aligned}|b_{Q_*} - b_Q| &= |b_{Q_1} - b_{Q_2} + b_{Q_2} - \cdots + b_{Q_m} - b_{Q_{m+1}}|\\ &\le \sum_{i=1}^m |b_{Q_i} - b_{Q_{i+1}}| \le \|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\sum_{i=1}^m 2^{(i+1)\beta_1}|Q_*|^{\beta_1/n}\\ &\le C\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}|Q_*|^{\beta_1/n},\end{aligned}$$

which is (2-2).

The following imbedding theorem for L^p spaces over domains with finite volume is very useful in the analysis of inequality, which you can find in any book about Sobolev spaces (see [Adams and Fournier 2003], for example).

Lemma 2.5. *Suppose that* $|\Omega| = \int_\Omega 1\,dx < \infty$ *and* $1 \le p \le q \le \infty$. *If* $f \in L^q(\Omega)$, *then* $f \in L^p(\Omega)$ *and*

$$\|f\|_{L^p(\Omega)} \le C|\Omega|^{1/p-1/q}\|f\|_{L^q(\Omega)}.$$

Proof of Theorem 1.3. (1) $\Rightarrow$ (2). For a cube $Q = Q(x_Q, r) \subset \mathbb{R}^n$ and $y \in Q$, take $f \in M^{p_2,\beta_2}(\mathbb{R}^n)$ and set $f_1 = f\chi_{2Q}$ and $f_2 = f - f_1$. After noticing that

$$T_b f = T_{(b-b_Q)} f,$$

we have

$$\begin{aligned}
&\left(\frac{1}{|Q|^{1+p\beta/n}}\int_Q |T_b f - (T_b f)_Q|^p\right)^{1/p}\\
&\qquad = \left(\frac{1}{|Q|^{1+p\beta/n}}\int_Q |T_{(b-b_Q)} f - (T_{(b-b_Q)} f)_Q|^p\right)^{1/p}\\
&\qquad \le \left(\frac{1}{|Q|^{1+p\beta/n}}\int_Q |T_{(b-b_Q)} f - (b-b_Q) f_2(x_Q)|^p\right)^{1/p}\\
&\qquad \le I + II + III,
\end{aligned}$$

where

$$\begin{aligned}
I &:= \left(\frac{1}{|Q|^{1+p\beta/n}}\int_Q |(b-b_Q)Tf|^p\right)^{1/p},\\
II &:= \left(\frac{1}{|Q|^{1+p\beta/n}}\int_Q |T(b-b_Q)f_1|^p\right)^{1/p},\\
III &:= \left(\frac{1}{|Q|^{1+p\beta/n}}\int_Q \left|(T(b-b_Q)f_2)(y) - (T(b-b_Q)f_2)(x_Q)\right|^p\right)^{1/p}.
\end{aligned}$$

Hölder's inequality and Lemma 2.2 imply

$$\begin{aligned}
I &= \frac{1}{|Q|^{1/p+\beta/n}}\left(\int_Q |(b-b_Q)Tf|^p\right)^{1/p}\\
&\le \frac{1}{|Q|^{1/p+\beta/n}}\left(\int_Q |b-b_Q|^{p_1}\right)^{1/p_1}\left(\int_Q |Tf|^{p_2}\right)^{1/p_2}\\
&\le C\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\|Tf\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}\\
&\le C\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\|f\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}.
\end{aligned}$$

From Lemma 2.3, it follows that

$$II \le \frac{1}{|Q|^{1/p+\beta/n}}\|(b-b_Q)f_1\|_{L^p} \le C\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\|f\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}.$$

We now turn to the estimate for the term III. From (1-6), it may be concluded that

$$\begin{aligned}
&\left|(T(b-b_Q)f_2)(y)-T((b-b_Q)f_2)(x_Q)\right|\\
&\qquad=\left|\int_{R^n}(K(y-z)-K(x_Q-z))(b(z)-b_Q)f_2(z)\,dz\right|\\
&\qquad\le C\int_{(2Q)^c}\frac{|y-x_Q|}{|z-x_Q|^{n+1}}|b(z)-b_Q||f(z)|\,dz\\
&\qquad\le C\sum_{k=2}^{\infty}\int_{2^kQ\backslash 2^{k-1}Q}\frac{1}{2^k|2^kQ|}(|b(z)-b_{2^kQ}|+|b_Q-b_{2^kQ}|)|f(z)|\,dz\\
&\qquad\le C\sum_{k=2}^{\infty}\frac{1}{2^k|2^kQ|}\int_{2^kQ}(|b(z)-b_{2^kQ}|+|b_Q-b_{2^kQ}|)|f(z)|\,dz,
\end{aligned}$$

which yields

$$\begin{aligned}
III&\le\left(\frac{1}{|Q|^{1+p\beta/n}}\int_Q\left|\sum_{k=2}^{\infty}\frac{1}{2^k|2^kQ|}\int_{2^kQ}|b(z)-b_{2^kQ}||f(z)|\,dz\right|^p\right)^{1/p}\\
&\quad+\left(\frac{1}{|Q|^{1+p\beta/n}}\int_Q\left|\sum_{k=2}^{\infty}\frac{1}{2^k|2^kQ|}\int_{2^kQ}|b_Q-b_{2^kQ}||f(z)|\,dz\right|^p\right)^{1/p}\\
&=:III_1+III_2.
\end{aligned}$$

With repeated application of Lemma 2.3 and the L^p-boundedness of the Hardy–Littlewood maximal operator M, we can obtain

$$\begin{aligned}
III_1&\le\frac{1}{|Q|^{1/p+\beta/n}}\left\|\sum_{k=2}^{\infty}\frac{1}{2^k|2^kQ|}\int_{2^kQ}|b(z)-b_{2^kQ}||f(z)|\,dz\right\|_{L^p}\\
&\le\frac{1}{|Q|^{1/p+\beta/n}}\sum_{k=2}^{\infty}\frac{1}{2^k}\left\|\frac{1}{|2^kQ|}\int_{2^kQ}|b(z)-b_{2^kQ}||f(z)|\,dz\right\|_{L^p}\\
&\le\sum_{k=2}^{\infty}\frac{1}{2^k}\frac{1}{|Q|^{1/p+\beta/n}}\left\|M(|b-b_{2^kQ}||f|)\right\|_{L^p}\\
&\le\sum_{k=2}^{\infty}\frac{1}{2^k}\frac{1}{|Q|^{1/p+\beta/n}}\left\||b-b_{2^kQ}||f|\right\|_{L^p}\\
&\le\sum_{k=2}^{\infty}\frac{1}{2^{k(1-n/p-\beta)}}\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\|f\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}\\
&\le C\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\|f\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}.
\end{aligned}$$

Applying Lemma 2.4 and Lemma 2.5 to III_2, we have

$$
\begin{aligned}
III_2 &\le \sum_{k=2}^{\infty} \frac{1}{2^k} \left(\frac{1}{|Q|^{1+p\beta/n}} \int_Q \left| \frac{1}{|2^k Q|} \int_{2^k Q} |b_Q - b_{2^k Q}| |f(z)|\, dz \right|^p dx \right)^{1/p} \\
&\le \sum_{k=2}^{\infty} \frac{1}{2^k} \|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)} \left(\frac{|Q|^{\beta_1/n}}{|Q|^{1+p\beta/n}} \int_Q \left| \frac{1}{|2^k Q|} \int_{2^k Q} |f(z)|\, dz \right|^p dx \right)^{1/p} \\
&\le \sum_{k=2}^{\infty} \frac{1}{2^k} \|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)} \frac{1}{|Q|^{1/p+\beta_2/n}} \|M(|f|)\|_{L^p} \\
&\le \sum_{k=2}^{\infty} \frac{1}{2^k} \|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)} \frac{1}{|Q|^{1/p+\beta_2/n}} \|f\|_{L^p} \\
&\le \sum_{k=2}^{\infty} \frac{1}{2^{k(1-n/p-\beta_2)}} \|b\|_{C^{p_1,\beta_1}} \|f\|_{M^{p_2,\beta_2}} \\
&\le C \|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)} \|f\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}.
\end{aligned}
$$

Thus, we have obtained $III \le C\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)} \|f\|_{M^{p_2,\beta_2}(\mathbb{R}^n)}$. Our proof ends with the definition of $\|\cdot\|_{C^{p,\beta}(\mathbb{R}^n)}$.

(2) $\Rightarrow$ (1). We first claim that for fixed $Q \subset \mathbb{R}^n$, $b \in C^{p_1,\beta_1}(Q)$ and $f \in M^{p_2,\beta_2}(Q)$ with $\|f\|_{M^{p_2,\beta_2}(Q)} = |Q|^{-\beta_2/n}$, we have

$$
\|T_b^m f\|_{C^{p,\beta}(Q)} \le C|Q|^{\beta_1(m-1)/n} \|f\|_{M^{p_2,\beta_2}(Q)} \le C|Q|^{(\beta_1 m-\beta)/n}, \tag{2-3}
$$

where T_b^m is the m-th ($m \in Z^+$) commutator defined by

$$
T_b^m f(x) = \text{p.v.} \int (b(x)-b(y))^m K(x-y) f(y)\, dy.
$$

We shall prove (2-3) by induction. The case $m = 1$ is trivial. We now assume that for any $b \in C^{p_1,\beta_1}(Q)$, we have

$$
\|T_b^{m-1} f\|_{C^{p,\beta}(Q)} \le C|Q|^{\beta_1(m-2)/n} \|f\|_{M^{p_2,\beta_2}(Q)} \le C|Q|^{(\beta_1(m-1)-\beta)/n}. \tag{2-4}
$$

Next, we show the case m. We now observe that

$$
\begin{aligned}
|T_b^m f(x)| &= \left| \int (b(x)-b(y))^{m-1} K(x-y) f(y)(b(x)-b(y))\, dy \right| \\
&\le \left| \int (b(x)-b(y))^{m-1} K(x-y) f(y)(b(x)-b_Q)\, dy \right| \\
&\quad + \left| \int (b(x)-b(y))^{m-1} K(x-y) f(y)(b(y)-b_Q)\, dy \right| \\
&\le (|b-b_Q| |T_b^{m-1} f|)(x) + |T_b^{m-1}((b-b_Q)f)(x)| =: J_1 + J_2.
\end{aligned}
$$

Equation (2-4) enables us to estimate J_1 as

$$\begin{aligned}\|J_1\|_{C^{p,\beta}(Q)} &\leq \||b-b_Q||T_b^{m-1}f|\|_{C^{p,\beta}(Q)}\\ &\leq \|b-b_Q\|_{L^\infty}\|T_b^{m-1}f\|_{C^{p,\beta}(Q)}\\ &\leq \frac{1}{|Q|}\int_Q |b-b_Q|\,dx\,|Q|^{\beta_1(m-2)/n}\|f\|_{M^{p_2,\beta_2}(Q)}\\ &\leq C|Q|^{\beta_1(m-1)/n}\|f\|_{M^{p_2,\beta_2}(Q)} \leq C|Q|^{(\beta_1 m-\beta)/n}.\end{aligned}$$

With repeated application of (2-4), we have

$$\begin{aligned}\|J_2\|_{C^{p,\beta}(Q)} &= \|T_b^{m-1}((b-b_Q)f)\|_{C^{p,\beta}(Q)}\\ &\leq C|Q|^{\beta_1(m-2)/n}\|(b-b_Q)f\|_{M^{p_2,\beta_2}(Q)}\\ &\leq C|Q|^{\beta_1(m-2)/n}\|b-b_Q\|_{L^\infty}\|f\|_{M^{p_2,\beta_2}(Q)}\\ &\leq C|Q|^{\beta_1(m-1)/n}\|b\|_{C^{p_1,\beta_1}(\mathbb{R}^n)}\|f\|_{M^{p_2,\beta_2}(Q)}.\\ &\leq C|Q|^{(\beta_1 m-\beta)/n}.\end{aligned}$$

We come back to the proof of (2) $\Rightarrow$ (1). The rest of the proof proceeds similarly to that of Theorem 1.1. We apply the same argument again with $s(x)$ replaced by $\operatorname{sgn}(b(x)-b_Q)^{p_1}$ to obtain

$$\int_Q |b-b_Q|^{p_1}\,dx \leq C\sum a_n \int_Q |T_b^{p_1}(g_n)|\,dx. \tag{2-5}$$

Combining (2-3) and observing $g_n \in M^{p_2,\beta_2}(Q)$ with the norm $\|g_n\|_{M^{p_2,\beta_2}(Q)} = |Q|^{-\beta_2/n}$, we estimate (2-5) as

$$\begin{aligned}\sum a_n \int_Q |T_b^{p_1}(g_n)|\,dx &\leq C|Q|^{1+\beta/n}\|T_b^{p_1}(g_n)\|_{C^{p,\beta}(Q)}\\ &\leq C|Q|^{1+p_1\beta_1/n+\beta_2/n}\|g_n\|_{M^{p_2,\beta_2}(Q)}\\ &\leq C|Q|^{1+p_1\beta_1/n}.\end{aligned}$$

and then take the supremum of Q, completing the proof of Theorem 1.3. □

Proof of Theorem 1.2. This proof can be handled in much the same way as that of Theorem 1.3, using (1-5) and replacing $b-b_Q$ by b in the proof of (2-3). □

Acknowledgements

The authors thank the anonymous referees cordially for their valuable suggestions on this paper.

References

[Adams and Fournier 2003] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, 2nd ed., Pure and Applied Mathematics **140**, Elsevier/Academic Press, Amsterdam, 2003. MR 2009e:46025 Zbl 1098.46001

[Campanato 1963] S. Campanato, "Proprietà di hölderianità di alcune classi di funzioni", *Ann. Scuola Norm. Sup. Pisa* (3) **17**:1-2 (1963), 175–188. MR 27 #6119 Zbl 0121.29201

[Chiarenza and Frasca 1987] F. Chiarenza and M. Frasca, "Morrey spaces and Hardy–Littlewood maximal function", *Rend. Mat. Appl.* (7) **7**:3-4 (1987), 273–279. MR 90f:42017 Zbl 0717.42023

[Coifman et al. 1976] R. R. Coifman, R. Rochberg, and G. Weiss, "Factorization theorems for Hardy spaces in several variables", *Ann. of Math.* (2) **103**:3 (1976), 611–635. MR 54 #843 Zbl 0326.32011

[Cruz-Uribe and Neugebauer 1995] D. Cruz-Uribe and C. J. Neugebauer, "The structure of the reverse Hölder classes", *Trans. Amer. Math. Soc.* **347**:8 (1995), 2941–2960. MR 95m:42026 Zbl 0851.42016

[DeVore and Sharpley 1984] R. A. DeVore and R. C. Sharpley, "Maximal functions measuring smoothness", *Mem. Amer. Math. Soc.* **47**:293 (1984), viii+115. MR 85g:46039 Zbl 0529.42005

[Ding 1997] Y. Ding, "A characterization of BMO via commutators for some operators", *Northeast. Math. J.* **13**:4 (1997), 422–432. MR 2000d:42006 Zbl 0903.42005

[Duong et al. 2007] X. T. Duong, J. Xiao, and L. Yan, "Old and new Morrey spaces with heat kernel bounds", *J. Fourier Anal. Appl.* **13**:1 (2007), 87–111. MR 2007k:42027 Zbl 1133.42017

[Fefferman 1983] C. L. Fefferman, "The uncertainty principle", *Bull. Amer. Math. Soc.* (*N.S.*) **9**:2 (1983), 129–206. MR 85f:35001 Zbl 0526.35080

[Gilbarg and Trudinger 1983] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, 2nd ed., Grundlehren der Mathematischen Wissenschaften **224**, Springer, Berlin, 1983. MR 86c:35035 Zbl 0562.35001

[Grafakos 2004] L. Grafakos, *Classical and modern Fourier analysis*, Pearson Education, Upper Saddle River, NJ, 2004. MR 2449250 Zbl 1148.42001

[Harboure et al. 1998] E. Harboure, O. Salinas, and B. Viviani, "Reverse-Hölder classes in the Orlicz spaces setting", *Studia Math.* **130**:3 (1998), 245–261. MR 2000c:42023 Zbl 0905.42012

[Janson 1978] S. Janson, "Mean oscillation and commutators of singular integral operators", *Ark. Mat.* **16**:2 (1978), 263–270. MR 80j:42034 Zbl 0404.42013

[Janson et al. 1983] S. Janson, M. Taibleson, and G. Weiss, "Elementary characterizations of the Morrey–Campanato spaces", pp. 101–114 in *Harmonic analysis* (Cortona, 1982), edited by G. Mauceri et al., Lecture Notes in Math. **992**, Springer, Berlin, 1983. MR 85k:46033 Zbl 0521.46022

[Komori and Shirai 2009] Y. Komori and S. Shirai, "Weighted Morrey spaces and a singular integral operator", *Math. Nachr.* **282**:2 (2009), 219–231. MR 2010h:42019 Zbl 1160.42008

[Lemarié-Rieusset 2007] P. G. Lemarié-Rieusset, "The Navier–Stokes equations in the critical Morrey–Campanato space", *Rev. Mat. Iberoam.* **23**:3 (2007), 897–930. MR 2009g:35221 Zbl 1227.35230

[Lin 2009] H. B. Lin, "Some remarks on Morrey spaces", *Beijing Shifan Daxue Xuebao* **45**:2 (2009), 111–116. MR 2010f:42042 Zbl 1212.42036

[Lin et al. 2011] H. B. Lin, E. Nakai, and D. Yang, "Boundedness of Lusin-area and g_λ^* functions on localized Morrey–Campanato spaces over doubling metric measure spaces", *J. Funct. Spaces Appl.* **9**:3 (2011), 245–282. MR 2012k:42047 Zbl 1231.42017

[Lu 1995] G. Lu, "Embedding theorems on Campanato–Morrey spaces for vector fields and applications", *C. R. Acad. Sci. Paris Sér. I Math.* **320**:4 (1995), 429–434. MR 95k:46051 Zbl 0842.46019

[Lu 1998] G. Lu, "Embedding theorems on Campanato–Morrey spaces for vector fields on Hörmander type", *Approx. Theory Appl.* (*N.S.*) **14**:1 (1998), 69–80. MR 99h:46053 Zbl 0916.46026

[Lu 2011] S. Lu, "Some results and problems on commutators", *Front. Math. China* **6**:5 (2011), 821–833. MR 2012h:42024 Zbl 1227.42016

[Lu et al. 2007] S. Lu, Y. Ding, and D. Yan, *Singular integrals and related topics*, World Scientific, Hackensack, NJ, 2007. MR 2009g:42023 Zbl 1124.42011

[Morrey 1938] C. B. Morrey, Jr., "On the solutions of quasi-linear elliptic partial differential equations", *Trans. Amer. Math. Soc.* **43**:1 (1938), 126–166. MR 1501936 Zbl 0018.40501

[Nakai 2006] E. Nakai, "The Campanato, Morrey and Hölder spaces on spaces of homogeneous type", *Studia Math.* **176**:1 (2006), 1–19. MR 2007j:46054 Zbl 1121.46031

[Paluszyński 1995] M. Paluszyński, "Characterization of the Besov spaces via the commutator operator of Coifman, Rochberg and Weiss", *Indiana Univ. Math. J.* **44**:1 (1995), 1–17. MR 96j:42006 Zbl 0838.42006

[Rupflin 2008] M. Rupflin, "What is a… Morrey Space? What is a… Campanato Space?", preprint, Zürich Graduate School in Mathematics, 2008, http://www.zgsm.ch/fileadmin/math/store/fs08/2902/Morrey2.pdf.

[Sakamoto and Yabuta 1999] M. Sakamoto and K. Yabuta, "Boundedness of Marcinkiewicz functions", *Studia Math.* **135**:2 (1999), 103–142. MR 2000g:42027 Zbl 0930.42009

[Stein 1970] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series **30**, Princeton University Press, Princeton, NJ, 1970. MR 44 #7280 Zbl 0207.13501

[Triebel 1992] H. Triebel, *Theory of function spaces, II*, Monographs in Mathematics **84**, Birkhäuser, Basel, 1992. MR 93f:46029 Zbl 0763.46025

[Yang et al. 2010] D. Yang, D. Yang, and Y. Zhou, "Localized Morrey–Campanato spaces on metric measure spaces and applications to Schrödinger operators", *Nagoya Math. J.* **198** (2010), 77–119. MR 2012a:46060 Zbl 1214.46019

Received May 5, 2012. Revised December 19, 2012.

SHAOGUANG SHI
DEPARTMENT OF MATHEMATICS
LINYI UNIVERSITY
LINYI 276005
CHINA
AND
SCHOOL OF MATHEMATICAL SCIENCES
BEIJING NORMAL UNIVERSITY
BEIJING, 100875

shishaoguang@mail.bnu.edu.cn

SHANZHEN LU
SCHOOL OF MATHEMATICAL SCIENCES
BEIJING NORMAL UNIVERSITY
BEIJING, 100875
CHINA

lusz@bnu.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 264, No. 1, 2013
dx.doi.org/10.2140/pjm.2013.264.235

THE SIEGEL–WEIL FORMULA FOR UNITARY GROUPS

SHUNSUKE YAMANA

We extend the Siegel–Weil formula for unitary groups of hermitian forms over a skew field with involution of the second kind.

Introduction

The Siegel–Weil formula is an identity between an Eisenstein series and an integral of a theta function. After Weil [1965] proved such an identity when both sides of the identity are absolutely convergent, Kudla and Rallis [1988a; 1988b; 1994] extended it for symplectic groups beyond the range of absolute convergence. Their results were extended to almost all classical groups by several authors, of which we mention the following sample: [Tan 1998; Ichino 2004; 2007; Gan and Takeda 2011; Yamana 2011; 2013; Gan 2000]. In this paper we discuss the last case that has to be considered in the theory of classical dual pairs over a number field, namely, unitary groups of hermitian forms over a skew field with involution of the second kind.

Let E/F be a quadratic extension of number fields and D a division algebra with center E, of dimension δ^2 over E and provided with an antiautomorphism ρ of order two under which F is the fixed subfield of E. Let $\mathbb{A}$ and $\mathbb{A}_E$ be the rings of adeles of F and E, respectively. Let $\mathcal{W}$ be a left D-vector space of dimension $2n$ with a nondegenerate skew hermitian form that has a complete polarization, and V a right D-vector space of dimension m with a nondegenerate hermitian form. Let G and H be the unitary groups of $\mathcal{W}$ and V, respectively.

Let α_E denote the standard norm of $\mathbb{A}_E^\times$. A character of $\mathbb{A}_E^\times$ is called principal if it is a complex power of α_E. We denote by P the maximal parabolic subgroup of G that stabilizes a maximal isotropic subspace of $\mathcal{W}$. Note that P has a Levi decomposition $P=MN$ with $M\simeq \mathrm{GL}_n(D)$. For any unitary character χ of $\mathbb{A}_E^\times/E^\times$ and for any $s\in\mathbb{C}$, we consider the representation $I(s,\chi)=\mathrm{Ind}_{P(\mathbb{A})}^{G(\mathbb{A})}\chi\alpha_E^s$ induced from the character $m\mapsto\chi(\nu(m))\alpha_E(\nu(m))^s$, where ν is the reduced norm viewed

The author is supported by JSPS Grant-in-Aid for Research Activity Start-up 24840033. This work is partially supported by the JSPS Institutional Program for Young Researcher Overseas Visits "Promoting international young researchers in mathematics and mathematical sciences led by OCAMI".
MSC2010: primary 11F27; secondary 11F70.
Keywords: Eisenstein series, Siegel–Weil formula, degenerate principal series, doubling method.

as a character of the algebraic group $\mathrm{GL}_n(D)$ and the induction is normalized so that $I(s,\chi)$ is naturally unitarizable when s is pure imaginary. For any holomorphic section $f^{(s)}$ of $I(s,\chi)$, the Eisenstein series

$$E(g; f^{(s)}) = \sum_{\gamma\in P(F)\backslash G(F)} f^{(s)}(\gamma g)$$

is absolutely convergent for $\Re s > \delta n/2$ and has a meromorphic continuation to the whole s-plane. We denote by χ^0 the restriction of χ to $\mathbb{A}^\times$, by $\rho(\chi)$ the character defined by $\rho(\chi)(x) = \chi(x^\rho)$ for $x \in \mathbb{A}_E^\times$, and by $\epsilon_{E/F}$ the quadratic character of $\mathbb{A}^\times/F^\times$ associated to the extension E/F. The following theorem was proven in [Tan 1999] when $\delta = 1$.

Theorem 1. *Let $f^{(s)}$ be a holomorphic section of $I(s,\chi)$.*

(1) *If $\chi\rho(\chi)$ is not principal, then $E(g; f^{(s)})$ is entire.*

(2) *If $\chi = \rho(\chi)^{-1}$, then the poles of $E(g; f^{(s)})$ in $\Re s > -\frac{1}{2}$ are at most simple and can only occur in the set*

$$\left\{ \frac{\delta(n-j)}{2} \;\middle|\; j\in\mathbb{Z},\ 0\le j<n,\ \chi^0 = \epsilon_{E/F}^{\delta j} \right\}.$$

Fix a nontrivial additive character ψ of $\mathbb{A}/F$ and a character χ_V of $\mathbb{A}_E^\times/E^\times$ such that $\chi_V^0 = \epsilon_{E/F}^{\delta m}$. The group $G(\mathbb{A})\times H(\mathbb{A})$ acts on the Schwartz space $\mathscr{S}(V^n(\mathbb{A}))$ of $V^n(\mathbb{A})$ via the Weil representation ω_{ψ,V,χ_V}. Let $S(V^n(\mathbb{A}))$ be the subspace of $\mathscr{S}(V^n(\mathbb{A}))$ consisting of functions that correspond to polynomials in the Fock model at every archimedean place of F.

The theta function associated to $\Phi \in S(V^n(\mathbb{A}))$ is defined by

$$\Theta(g,h;\Phi) = \sum_{x\in V^n(F)} \omega_{\psi,V,\chi_V}(g)\Phi(h^{-1}x)$$

for $g\in G(\mathbb{A})$ and $h \in H(\mathbb{A})$. By Weil's criterion [1965], the integral

$$I(g;\Phi) = \int_{H(F)\backslash H(\mathbb{A})} \Theta(g,h;\Phi)\,dh$$

is absolutely convergent for all Φ either if $r = 0$ or if $m - r > n$, where r is the dimension of a maximal totally isotropic subspace of $V(F)$. When $m \le n$ and $r > 0$, the integral diverges in general, but extends uniquely to a $G(\mathbb{A})$-intertwining, $H(\mathbb{A})$-invariant map on $S(V^n(\mathbb{A}))$ in light of the regularization introduced by Kudla and Rallis [1994].

For $\Phi \in S(V^n(\mathbb{A}))$ we define a section $f_\Phi^{(s)}$ of $I(s,\chi_V)$ by

$$f_\Phi^{(s)}(g) = |a(g)|^{s-s_0}\omega_{\psi,V,\chi_V}(g)\Phi(0),$$

where $g \in G(\mathbb{A})$, $s_0 = \delta(m-n)/2$ and the quantity $|a(g)|$ is defined in the notation section below.

Theorem 2. *If $m \le n$ or if $m - r > n$, then for all $\Phi \in S(V^n(\mathbb{A}))$ the series $E(g; f_\Phi^{(s)})$ is holomorphic at $s = s_0$ and*

$$E(g; f_\Phi^{(s)})|_{s=s_0} = \varkappa I(g; \Phi),$$

where

$$\varkappa = \begin{cases} 2 & \text{if } m \le n, \\ 1 & \text{if } m - r > n. \end{cases}$$

Theorem 2 was proven in [Weil 1965] if $m > 2n$, and in [Tan 1998; Ichino 2004; 2007; Yamana 2011] if $\delta = 1$. The proof requires only slight technical modifications once all of the necessary local facts have been established. The group $G(F_v)$ is isomorphic to the quasisplit unitary group $U(\delta n, \delta n)$ or an inner form of $\mathrm{GL}_{2\delta n}(F_v)$, depending on whether v remains prime or splits in E. The former case has already been discussed in [Kudla and Sweet 1997; Ichino 2007; Lee and Zhu 1998], and the latter case is discussed in Section 1. Coupled with the doubling method, the Siegel–Weil formula relates the theory of theta liftings to the theory of automorphic L-functions. We study the doubling zeta integral for inner forms of general linear groups in the Appendix.

Notation

Let (D, E, F, ρ) be as in the introduction. The restriction of ρ to E, which we denote also by ρ, is the nontrivial automorphism of E over F. For a matrix x with entries in D, let $x^* = {}^t x^\rho$ be the conjugate transpose of x. If x is a square matrix, then $\nu(x)$ and $\tau(x)$ stand for its reduced norm and reduced trace to E.

Fix a natural number n and put $n' = \delta n$. Let $\mathcal{W} = D^{2n}$ be a left D-vector space with the skew hermitian form

$$\langle x, y \rangle = x J y^*, \qquad J = \begin{pmatrix} 0 & \mathbf{1}_n \\ -\mathbf{1}_n & 0 \end{pmatrix}$$

for $x, y \in \mathcal{W}$. Let V be a right D-vector space of dimension m equipped with a nondegenerate hermitian form $(\ ,\ )$. We denote by G (resp. H) the group of all D-linear transformations of $\mathcal{W}$ (resp. V) that leave $\langle\ ,\ \rangle$ (resp. $(\ ,\ )$) invariant. Put $s_0 = \delta(m-n)/2$.

We write P for the stabilizer in G of the maximal isotropic subspace of $\mathcal{W}$ defined by the vanishing of all but the last n coordinates. Let

$$\mathrm{Her}_n = \{x \in \mathrm{M}_n(D) \mid x^* = x\}$$

be the F-subvariety of $n \times n$ hermitian matrices. The group G has a maximal parabolic subgroup $P = MN$ given by

$$M = \left\{ m(a) = \begin{pmatrix} a & 0 \\ 0 & (a^{-1})^* \end{pmatrix} \;\middle|\; a \in \mathrm{GL}_n(D) \right\},$$

$$N = \left\{ n(b) = \begin{pmatrix} \mathbf{1}_n & b \\ 0 & \mathbf{1}_n \end{pmatrix} \;\middle|\; b \in \mathrm{Her}_n \right\}.$$

Let K be the standard maximal compact subgroup of $G(\mathbb{A})$. For any character χ of $\mathbb{A}_E^\times / E^\times$, the representation $I(s, \chi) = I_{n'}(s, \chi)$ is realized on the space of right K-finite functions $f^{(s)} : G(\mathbb{A}) \to \mathbb{C}$ satisfying

$$f^{(s)}\big(m(a)n(b)g\big) = \chi(\nu(a))\alpha_E(\nu(a))^{s+n'/2} f^{(s)}(g)$$

for all $a \in \mathrm{GL}_n(D(\mathbb{A}))$, $b \in \mathrm{Her}_n(\mathbb{A})$ and $g \in G(\mathbb{A})$. We define $|a(g)|$ by writing $g = pk \in G(\mathbb{A})$ with $p = m(a)n(b) \in P(\mathbb{A})$ and $k \in K$, and taking $|a(g)| = \alpha_E(\nu(a))$.

1. Degenerate principal series representations

For each place v of F, let F_v be the v-completion of F and set $E_v = E \otimes_F F_v$ and $D_v = D \otimes_F F_v$. A division algebra D with center E admits an involution of the second kind if and only if D_v is isomorphic to $\mathrm{M}_\delta(E_v)$ whenever v remains prime in E, and D_v is isomorphic to a direct sum of mutually opposite simple algebras whose centers are F_v whenever v splits in E (see [Scharlau 1985, Theorem 10.2.4]).

In the local setting we will depart slightly from our previous notation. Fix a place v of F and suppress it from the notation. Thus E is a quadratic étale algebra over the local field F, D an algebra whose center is E, ρ an involution of D whose restriction to E is the nontrivial automorphism of E over F, V a free right D-module of rank m, and $(\,,\,) : V \times V \to D$ an F-bilinear map satisfying the following conditions:

- for $a, b \in D$ and $x, y \in V$,

$$(x, y)^\rho = (y, x), \qquad (xa, yb) = a^\rho (x, y) b;$$

- $(x, V) = 0$ implies that $x = 0$.

Let H be the unitary group of V. Let $G = \{g \in \mathrm{GL}_{2n}(D) \mid gJg^* = J\}$. For any quasicharacter χ of $E^\times$, let $I(s, \chi)$ be the analogous local induced representation of G. By Morita context, it is enough to consider the case where the triple (D, E, ρ) belongs to the following two types:

- $D = E$ is a quadratic extension of F and ρ generates $\mathrm{Gal}(E/F)$;
- $D = \boldsymbol{D} \oplus \boldsymbol{D}^{\mathrm{op}}$, $E = F \oplus F$ and $(x, y)^\rho = (y, x)$, where $\boldsymbol{D}$ is a division algebra central over F and $\boldsymbol{D}^{\mathrm{op}}$ is its opposite algebra.

The rank of D as a module over E is a square of a natural number that will be denoted by δ. Note that $n' = \delta n$ remains intact after the change in notation.

We fix a nontrivial additive character ψ of F and a character χ_V of $E^\times$ that satisfies $\chi_V^0 = \epsilon_{E/F}^{\delta m}$. Then $G \times H$ acts on the Schwartz space $\mathscr{S}(V^n)$ via the Weil representation ω_{ψ,V,χ_V}. Note that it depends on the data ψ, $(\ ,\ )$ and χ_V (compare [Kudla 1994]). When F is a p-adic field, put $S(V^n) = \mathscr{S}(V^n)$. When $F = \mathbb{R}$ or $\mathbb{C}$, let $\mathfrak{g}$ be the complexified Lie algebra of G and $S(V^n)$ the subspace of $\mathscr{S}(V^n)$ that corresponds to the space of polynomials in the Fock model of ω_{ψ,V,χ_V}. In the archimedean case we only consider admissible representations of the pair $(\mathfrak{g}, K)$, although we will allow ourselves to speak of a representation of the group G. We write $R(V, \chi_V) = R_{n'}(V, \chi_V)$ for the image of the intertwining map

$$S(V^n) \to I(s_0, \chi_V), \quad \Phi \mapsto f_\Phi^{(s_0)}(g) = \omega_{\psi,V,\chi_V}(g)\Phi(0).$$

We extend $f_\Phi^{(s_0)}$ to the standard section $f_\Phi^{(s)}$ of $I(s, \chi_V)$.

We discuss the case $E = F \oplus F$. Put

$$e_1 = (1, 0), \quad e_2 = (0, 1), \quad V_1 = Ve_1, \quad V_2 = Ve_2.$$

We regard V_1 as a right $\boldsymbol{D}$-module and V_2 as both a right $\boldsymbol{D}^{\mathrm{op}}$-module and a left $\boldsymbol{D}$-module. Since $(V_i, V_i) = 0$ for $i = 1, 2$, the spaces V_1 and V_2 are paired nondegenerately against each other by $(\ ,\ )$, and so an antiisomorphism

$$\jmath : \mathrm{End}(V_1, \boldsymbol{D}) \to \mathrm{End}(V_2, \boldsymbol{D}^{\mathrm{op}})$$

is defined by

$$(ax, y) = (x, \jmath(a)y), \quad a \in \mathrm{End}(V_1, \boldsymbol{D}),\ x \in V_1,\ y \in V_2.$$

We obtain

$$H = \bigl\{(a, \jmath(a)^{-1}) \in \mathrm{GL}(V_1, \boldsymbol{D}) \times \mathrm{GL}(V_2, \boldsymbol{D}^{\mathrm{op}}) \bigm| a \in \mathrm{GL}(V_1, \boldsymbol{D})\bigr\}.$$

Thus projection onto the first or second factor induces an isomorphism of H onto $\mathrm{GL}(V_1, \boldsymbol{D})$ or $\mathrm{GL}(V_2, \boldsymbol{D}^{\mathrm{op}})$, respectively. For any nonnegative integer j we write $G'_j = \mathrm{GL}_j(\boldsymbol{D})$. Observe that

$$G = \{(g, J^{-1}\,{}^t g^{-1} J) \mid g \in G'_{2n}\}.$$

Through projection onto the first factor, we identify H with G'_m, G with G'_{2n}, and $P = MN$ with

$$M = \left\{\begin{pmatrix} a & 0 \\ 0 & d\end{pmatrix} \,\middle|\, a, d \in \mathrm{GL}_n(\boldsymbol{D})\right\}, \quad N = \left\{\begin{pmatrix} \mathbf{1}_n & b \\ 0 & \mathbf{1}_n\end{pmatrix} \,\middle|\, b \in \mathrm{M}_n(\boldsymbol{D})\right\}.$$

We write $\nu = \nu_j$ for the reduced norm of $\mathrm{M}_j(\boldsymbol{D})$ and τ for the reduced trace of $\mathrm{M}_j(\boldsymbol{D})$. Let $\alpha_F(x) = |x|_F$ denote the normalized absolute value of $x \in F^\times$. When we write $\chi=(\chi_1,\chi_2)$, the representation $I(s,\chi)$ is translated to

$$I(s,\chi)=\mathrm{Ind}_P^{G'_{2n}}\big((\chi_1\alpha_F^s)\circ\nu_n\boxtimes(\chi_2\alpha_F^s)^{-1}\circ\nu_n\big).$$

If $E=F\oplus F$, then since χ_V is of the form (μ,μ^{-1}), we may assume that $\chi_V=1$ by twisting, and we write $I(s)=I(s,1)$ and $R(V)=R(V,1)$. The Weil representation $\omega_{j,k}$ of the dual pair (G'_j,G'_k) can be taken to be the action on $\mathscr{S}(\mathrm{M}_{k,j}(\boldsymbol{D}))$ given by

$$\omega_{j,k}(a,b)\phi(x)=\alpha_F(\nu_j(a))^{\delta k/2}\alpha_F(\nu_k(b))^{-\delta j/2}\phi(b^{-1}xa)$$

for $a\in G'_j$ and $b\in G'_k$. Note that the integral

$$(\phi,\phi')=\int_{\mathrm{M}_{k,j}(\boldsymbol{D})}\phi(u)\overline{\phi'(u)}\,du,\qquad \phi,\phi'\in\mathscr{S}(\mathrm{M}_{k,j}(\boldsymbol{D}))$$

defines a $G'_j\times G'_k$ invariant positive definite hermitian form on $\omega_{j,k}$. The two models of the Weil representation $\omega_{2n,m}\simeq\omega_{\psi,V,1}$ are related by the partial Fourier transform

$$\mathscr{F}\phi(x,y)=\int_{\mathrm{M}_{m,n}(\boldsymbol{D})}\phi((x,z))\psi(-\tau(z\,{}^t y))\,dz \tag{1-1}$$

for $x\in\mathrm{M}_{m,n}(\boldsymbol{D})$ and $y\in\mathrm{M}_{m,n}(\boldsymbol{D}^{\mathrm{op}})$. In the p-adic case we write $\mathbb{O}$ for the maximal compact subring of $\boldsymbol{D}$ and put $K_n=\mathrm{GL}_n(\mathbb{O})$. In the archimedean case we set

$$K_n=\{g\in G'_n\mid {}^t\bar{g}g=\mathbf{1}_n\},$$

denoting the conjugate transpose of $x\in\mathrm{M}_n(\boldsymbol{D})$ by ${}^t\bar{x}$, where $\bar{\cdot}$ denotes the complex conjugate or the quaternion conjugate. We denote by $f_0^{(s)}$ a unique section of $I(s)$ that is identically 1 on K_{2n}.

Lemma 1.1. *If $E=F\oplus F$, then $R(V)$ contains $f_0^{(s_0)}$.*

Proof. In the p-adic case, we let $\phi_{j,k}$ be the characteristic functions of $\mathrm{M}_{j,k}(\mathbb{O})$. In the archimedean case we let

$$\phi_{j,k}(x)=e^{-\pi\mathrm{Tr}_{F/\mathbb{R}}(\tau({}^t\bar{x}x))},$$

assuming that $\psi(\,\cdot\,)=e^{2\pi\sqrt{-1}\mathrm{Tr}_{F/\mathbb{R}}(\,\cdot\,)}$. Put $\Phi=\mathscr{F}\phi_{2n,m}$. Then $f_\Phi^{(s_0)}$ is nonzero and right invariant under K_{2n}. □

The local intertwining operator is defined analogously by

$$M(s,\chi)f^{(s)}(g)=\int_{\mathrm{Her}_n(F)}f^{(s)}(Jn(b)g)\,db.$$

We define holomorphic sections and standard sections similarly. We write χ^0 for the restriction of χ to $F^\times$. Put

$$a(s,\chi)=a_{n'}(s,\chi)=\prod_{j=1}^{n'} L\big(2s-j+1,\chi^0\cdot\epsilon_{E/F}^{n'+j}\big),$$

$$b(s,\chi)=b_{n'}(s,\chi)=\prod_{j=1}^{n'} L\big(2s+j,\chi^0\cdot\epsilon_{E/F}^{n'+j}\big).$$

A normalized intertwining operator $M^*(s,\chi)$ is defined by setting

$$M^*(s,\chi)=a(s,\chi)^{-1}M(s,\chi).$$

Lemma 1.2. *The operator $M^*(s,\chi)$ is entire.*

Proof. When E/F is a quadratic extension of p-adic fields, Lemma 1.2 is proven in Proposition 3.2 of [Kudla and Sweet 1997]. The proof is completely analogous when $E/F=\mathbb{C}/\mathbb{R}$. Note that Proposition 3A.6 of the same work applies also to this case by a global consideration, namely, by applying (24) of [Lapid and Rallis 2005] with base field $\mathbb{Q}$ and $S=\{\infty\}$.

We suppose that $E=F\oplus F$. For $\phi\in\mathscr{S}(\mathrm{M}_n(\boldsymbol{D}))$ we define a section $f_\phi^{(s)}$ of $I(s,\chi)$ by requiring that $\operatorname{supp}(f_\phi^{(s)})\subset PJN$ and $f_\phi^{(s)}(g)=\phi(b)$ if $g=Jn(b)$ for $b\in\mathrm{Her}_n(F)$. As explained in [Piatetski-Shapiro and Rallis 1987b; Kudla and Sweet 1997], all we have to do is to show that the ratio $a(s,\chi)^{-1}M(s,\chi)f_\phi^{(s)}(J)$ is entire. One can easily observe that

$$M(s,\chi)f_\phi^{(s)}(J)=Z^{GJ}\left(2s-\frac{n'}{2},\phi,\chi^0\circ\nu_n\right),$$

where the right-hand side is the zeta integral studied in [Weil 1974; Godement and Jacquet 1972] (see the Appendix). Our claim follows at once, as the Godement–Jacquet L-factor

$$L^{GJ}\left(2s-\frac{n'-1}{2},\chi^0\circ\nu_n\right)$$

divided by the factor $a(s,\chi)$ is entire. □

For $\beta\in\mathrm{Her}_n(F)$, let ψ_β be the character of N defined by $\psi_\beta(n(b))=\psi(\tau(\beta b))$. Notice that $\tau(\beta b)\in F$. The Fourier transform of a Schwartz function $f\in\mathscr{S}(N)$ is defined by

$$\hat f(\beta)=\int_N f(u)\psi_\beta(u)\,du.$$

For each integer $j\le n'$, we define the subvariety Her_n^j of $\mathrm{Her}_n(F)$ by

$$(E\not\simeq F\oplus F)\qquad \mathrm{Her}_n^j=\big\{\beta\in\mathrm{M}_n(E)\;\big|\;{}^t\beta^\rho=\beta,\ \operatorname{rank}_E\beta\le j\big\},$$

$$(E=F\oplus F)\qquad \mathrm{Her}_n^j=\big\{(\beta,{}^t\beta)\in\mathrm{M}_n(\boldsymbol{D})\oplus\mathrm{M}_n(\boldsymbol{D}^{\mathrm{op}})\;\big|\;\delta(\operatorname{rank}_{\boldsymbol{D}}\beta)\le j\big\}.$$

Definition 1.3. We say that a representation π of G has rank at most j if $f \in \mathscr{S}(N)$ acts by zero on π whenever $\hat{f}$ vanishes on Her_n^j. We say that π is of rank j if in addition j is a multiple of δ and π does not have rank less than j.

For any H-module π, we write π_H for the maximal quotient of π on which H acts trivially. Let $\mathscr{H}_r$ be a split hermitian space of dimension $2r$, that is, $\mathscr{H}_r$ has a D-basis consisting of $2r$ elements e_i, f_i such that

$$(e_i, e_j) = (f_i, f_j) = 0, \quad (e_i, f_j) = \delta_{ij}.$$

Proposition 1.4. *Assume that $m \le n$. Let $U = V \oplus \mathscr{H}_{n-m}$.*

(1) *$R(V, \chi_V)$ is irreducible and unitarizable.*

(2) *$R(V, \chi_V)$ is isomorphic to $S(V^n)_H$.*

(3) *If E/F is a quadratic extension of p-adic fields, then $R(V, \chi_V)$ is of rank m.*

(4) *$R(U, \chi_V)$ has a unique irreducible quotient that is isomorphic to $R(V, \chi_V)$.*

(5) *$M^*(-s_0, \chi_V)$ maps $R(U, \chi_V)$ onto $R(V, \chi_V)$.*

(6) *$b(s, \chi_V)M^*(s, \chi_V) f_\Phi^{(s)}$ is holomorphic at $s = s_0$ for every $\Phi \in S(V^n)$.*

Proof. When $D = E$, these results are known (see [Li 1989; Mœglin et al. 1987; Kudla and Sweet 1997; Lee and Zhu 1998; Yamana 2011]). We may suppose that $E = F \oplus F$ and $\delta > 1$.

For $0 \le i \le k$, let $P_i^k = M_i^k N_i^k$ be the maximal parabolic subgroup of G'_k given by

$$P_i^k = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in G'_k \;\middle|\; a \in G'_{k-i},\ b \in \mathrm{M}_{k-i,i}(\boldsymbol{D}),\ d \in G'_i \right\},$$

$\bar{P}_i^k$ its opposite parabolic subgroup, and r_i the representation of $G'_i \times G'_i$ on $\mathscr{S}(G'_i)$ given by

$$r_i(g_1, g_2)\phi(g) = \phi(g_2^{-1} g g_1), \quad (\phi \in \mathscr{S}(G'_i),\ g, g_1, g_2 \in G'_i).$$

In the archimedean case the representation $I(s)$ is studied extensively in [Lee 2007; Sahi 1995; Zhang 1995]. From their results we know the module structure of $I(s_0)$ and the set of K-types of each of its irreducible constituents, which combined with the technique explained in [Kudla and Rallis 1990a] prove (1), (2). We consider the nonarchimedean case. By Lemma 3.III.2 of [Mœglin et al. 1987], the representation $\omega_{2n,m}$ has a filtration

$$0 \subset S_m \subset \cdots \subset S_1 \subset S_0 = \omega_{2n,m}$$

with successive quotients

$$S_i / S_{i+1} \simeq \mathrm{Ind}_{P_i^{2n} \times \bar{P}_i^m}^{G'_{2n} \times G'_m} \mu_i,$$

where μ_i is the representation of $P_i^{2n} \times \bar{P}_i^m$ on $\mathscr{S}(G'_i)$ given by

$$\mu_i(p,p')\phi = \alpha_F\big(\nu(a)^{m-i}\nu(a')^{i-2n}\nu(d)^{m-i+2n}\nu(d')^{i-m-2n}\big)^{\delta/2} r_i(d,d')\phi,$$

where

$$p = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in P_i^{2n}, \quad p' = \begin{pmatrix} a' & 0 \\ c' & d' \end{pmatrix} \in \bar{P}_i^m, \quad \phi \in \mathscr{S}(G'_i).$$

Let $\mathbb{1}_j$ denote the trivial representation of G'_j. For $0 \le i < m$ and an admissible representation π of G'_{2n}, the Frobenius reciprocity gives

$$\mathrm{Hom}_{G'_{2n}\times G'_m}\big(S_i/S_{i+1}, \pi \otimes \mathbb{1}_m\big) \simeq \mathrm{Hom}_{M_i^{2n}\times M_i^m}\Big((\pi^\vee)_{N_i^{2n}} \otimes \delta_{P_i^m}^{1/2}, \mu_i^\vee\Big),$$

where $\delta_{P_i^m}$ is the modulus function on P_i^m and $(\pi^\vee)_{N_i^{2n}}$ is the normalized Jacquet module of $\pi^\vee$ associated to P_i^{2n}. Since the quasicharacters of G'_{m-i} do not match, the space above is zero. Thus $(S_i/S_{i+1})_{G'_m} = 0$, so that the natural map $(S_m)_{G'_m} \to (\omega_{2n,m})_{G'_m}$ is surjective. If χ is a quasicharacter of G'_m and if a distribution T on $\mathscr{S}(G'_m)$ transforms according to χ under the action of $e \times G'_m$, that is,

$$T(r_m(e,h)f) = \chi(\nu(h))T(f)$$

for all $h \in G'_m$, then there is a constant $c \in \mathbb{C}$ such that

$$T(f) = c\int_{G'_m} f(h)\chi(\nu(h))\,dh, \quad f \in \mathscr{S}(G'_m)$$

(see Lemma 3.II.3 of [Mœglin et al. 1987]). It follows that

$$(S_m)_{G'_m} \simeq \mathrm{Ind}_{P_m^{2n}}^{G'_{2n}}(\mathbb{1}_{2n-m} \otimes \mathbb{1}_m).$$

Since $\mathrm{Ind}_{P_m^{2n}}^{G'_{2n}}(\mathbb{1}_{2n-m} \otimes \mathbb{1}_m)$ is irreducible as a representation of G'_{2n} induced from a unitary representation [Sécherre 2009], we have

$$(\omega_{\psi,V,1})_H \simeq \mathrm{Ind}_{P_m^{2n}}^{G'_{2n}}(\mathbb{1}_{2n-m} \otimes \mathbb{1}_m).$$

Thus the map from $(\omega_{\psi,V,1})_H$ to $R(V)$ is injective. This proves (1), (2).

In the p-adic case, Theorem 5.1 of [Mínguez 2009] tells us that $I(s_0)$ has a unique irreducible subrepresentation, which is $R(V)$, and hence $I(-s_0)$ has a unique irreducible quotient. We refer to [Lee 2007] for the archimedean analogue. From Lemma 1.1 we can infer that $f_0^{(-s_0)}$ generates $I(-s_0)$. It follows that $I(-s_0) = R(U)$. The proof of (4) is complete.

To prove (5), (6), it suffices to check that $b(s)M^*(s)f_0^{(s)}$ (resp. $M^*(s)f_0^{(s)}$) are holomorphic and nonzero at $s = s_0$ (resp. $s = -s_0$) in light of [Kudla and Rallis 1988a, Proposition 4.9]. Let $\phi_0 = \phi_{n,n} \in S(\mathrm{M}_n(\boldsymbol{D}))$ be as in the proof of Lemma 1.1. Define $\phi_1 \in S(\mathrm{M}_{n,2n}(\boldsymbol{D}))$ by $\phi_1(x,y) = \phi_0(x)\phi_0(y)$. The sections $\mathfrak{F}_{\phi_1}^{(s)}$ and $\mathfrak{F}_{\hat{\phi}_1}^{(s)}$ are defined in the Appendix. Since $\mathfrak{F}^{(s)}_{\phi_1}$ is right K-invariant, so is $\mathfrak{F}^{(s)}_{\hat{\phi}_1}$ by Lemma A.1. From Propositions 10.7 and 10.8 of [Weil 1974], we know

$$\mathfrak{F}^{(s)}_{\phi_1} = \mathfrak{F}^{(s)}_{\phi_1}(e) \cdot f^{(s)}_0 = Z^{GJ}\Big(2s + \frac{n'}{2}, \phi_0, 1\Big) \cdot f^{(s)}_0 = f^{(s)}_0 \prod_{j=1}^{n} \xi(2s + \delta j)$$

up to multiplication by exponential factors, where $\xi(s) = \zeta(s)$ in the p-adic case, and $\xi(s) = \Gamma(s)$ in the archimedean case. Observe that

$$\begin{aligned}\mathfrak{F}^{(-s)}_{\hat{\phi}_1} &= Z^{GJ}\Big(-2s + \frac{n'}{2}, \hat{\phi}_0, 1\Big) \cdot f^{(s)}_0 \\ &= (-1)^{n(\delta-1)} \gamma^{GJ}\Big(2s - \frac{n'-1}{2}, \mathbb{1}_n, \psi\Big) Z^{GJ}\Big(2s - \frac{n'}{2}, \phi_0, 1\Big) \cdot f^{(s)}_0.\end{aligned}$$

Substituting these into the equality in Lemma A.1, we get

$$M(s) f^{(s)}_0 = f^{(-s)}_0 \prod_{j=1}^{n} \frac{\xi(2s - \delta j + \delta)}{\xi(2s + \delta j)}. \tag{1-2}$$

Now we can easily conclude our proof. □

2. Proof of Theorem 1

Back to the global setup, we write $\mathscr{A}$ for the space of automorphic forms on $G(\mathbb{A})$. For $\beta \in \mathrm{Her}_n(F)$ and $A \in \mathscr{A}$, let

$$A_\beta(g) = \int_{\mathrm{Her}_n(F)\backslash \mathrm{Her}_n(\mathbb{A})} A(n(b)g)\psi(-\tau(\beta b))\, db, \qquad g \in G(\mathbb{A})$$

denote the β-th Fourier coefficient of A. The following lemma can be proven in exactly the same way as in [Kudla and Rallis 1990b; Tan 1999].

Lemma 2.1. *Let $f^{(s)}$ be a holomorphic section of $I(s, \chi)$ and $\beta \in \mathrm{Her}_n(F)$ with $\nu(\beta) \neq 0$.*

1. *$b(s, \chi) E_\beta(g; f^{(s)})$ is holomorphic in $\Re s > -\frac{1}{2}$.*
2. *If $m \geq n$ and β is represented by $V(F)$, then $E_\beta(g; f^{(s)}_\Phi)$ can be made nonzero at $s = s_0$ for a suitable choice of $\Phi \in S(V^n(\mathbb{A}))$.*
3. *If $\chi\rho(\chi)$ is not principal, then $E(g; f^{(s)})$ is entire.*
4. *If $\chi = \rho(\chi)^{-1}$, then the poles of $E(g; f^{(s)})$ in $\Re s > -\frac{1}{2}$ are at most simple and can only occur in the set*
$$\Big\{ \frac{n'-j}{2} \;\Big|\; j \in \mathbb{Z},\ 0 \leq j < n',\ \chi^0 = \epsilon^j_{E/F} \Big\}.$$
5. *If $\chi^0 = \epsilon^{n'+1}_{E/F}$, then $E(g; f^{(s)})|_{s=0}$ is identically zero.*

Definition 2.2. For each integer $l \le n$, we say that $A \in \mathscr{A}$ has rank δl if $A_\beta = 0$ when $\mathrm{rank}_D \beta > l$, but $A_\beta \neq 0$ for some β of rank l. When π is a representation of $G(\mathbb{A})$ realized on a subspace of $\mathscr{A}$, we say that π has rank at most δl if all functions in π have rank at most δl.

We call A singular if it has rank less than δn. The following lemma can be proven in the same way as in the proof of [Howe 1981, Lemma 2.4].

Lemma 2.3. *Let π be a subrepresentation of $\mathscr{A}$. For every integer $l \le n$ the following conditions are equivalent*:

- *π has rank at most δl*;
- *for every place v, $G(F_v)$ acts on π by a representation of rank at most δl*;
- *for at least one place v, $G(F_v)$ acts on π by a representation of rank at most δl.*

In particular, *if $G(F_v)$ acts on π by a representation of rank at most j, then $G(F_v)$ acts on π by a representation of rank at most $\delta\ell$, where $\ell = [j/\delta]$.*

For $s' \in \mathbb{C}$ with $\Re s' > -\frac{1}{2}$, the residue $\mathrm{Res}_{s=s'} E(g; f^{(s)})$ depends only on $f^{(s')}$, and $f^{(s')} \mapsto \mathrm{Res}_{s=s'} E(g; f^{(s)})$ gives a $G(\mathbb{A})$ intertwining map

$$A_{-1}(s') : I(s', \chi) \to \mathscr{A}.$$

Assume that $\chi = \rho(\chi)^{-1}$, assume that j is an integer between 0 and n', assume that $\chi^0 = \epsilon_{E/F}^j$, and assume that j is not divisible by δ. Let $s' = (n' - j)/2$. To complete the proof of Theorem 1, it remains to prove that $A_{-1}(s')$ is zero. Fix a finite inert place v of F. By Theorem 1.2 of [Kudla and Sweet 1997], $I_v(s', \chi_v)$ has a unique irreducible submodule R and

$$I_v(s', \chi_v)/R \simeq \bigoplus_{V_0} R(V_0, \chi_v),$$

where V_0 runs over all equivalence classes of hermitian spaces over E_v of dimension j. Since the image of $A_{-1}(s')$ lies in the space of singular automorphic forms in view of Lemma 2.1(1) and since R is nonsingular, the map $A_{-1}(s')$ factors through the quotient $\bigoplus_{V_0} R(V_0, \chi_v)$ at v. Proposition 1.4(3) shows that $G(F_v)$ acts on the image of $A_{-1}(s')$ by a representation of rank at most j. Put $\ell = [j/\delta]$. Lemma 2.3 shows that $G(F_v)$ acts on the image of $A_{-1}(s')$ by a representation of rank at most $\delta\ell$. Since $\delta\ell < j$, Proposition 1.4(3) forces $A_{-1}(s')$ to be zero.

3. Proof of Theorem 2

Lemma 3.1. *If $m = n$ or if $m - r > n$, then for all $\Phi \in S(V^n(\mathbb{A}))$ and $\beta \in \mathrm{Her}_n(F)$ with $\nu(\beta) \neq 0$,*

$$E_\beta(g; f_\Phi^{(s)})|_{s=s_0} = \varkappa I_\beta(g; \Phi).$$

Proof. The proof can be carried out by the same technique as in that of [Ichino 2004, Proposition 6.2]. We omit the details. □

First we prove Theorem 2 in the case $m-r>n$. Ichino [2007] proved the special case of this result for $\delta=1$ (compare [Kudla and Rallis 1988b; Yamana 2013]). Many of the results there apply word for word in our general case.

If $m>2n$, then $E(g;f_\Phi^{(s_0)})$ converges absolutely and the stated identity was proven by Weil [1965]. We may suppose that $m\le 2n$. Fix $\Phi^0=\bigotimes_v\Phi_v^0\in S(V^n(\mathbb{A}))$. By Theorem 10.6.2 of [Scharlau 1985], there is an inert place w of F such that the Witt index r_w of V_w satisfies $r_w<\delta(r+1)$, where V_w stands for the hermitian space over E_w corresponding to $V(F_w)$. Note that

$$\delta m-r_w>\delta n.$$

We consider the $G(F_w)$-intertwining map

$$A_{-1,w}:S(V_w^{n'})\to\mathscr{A},\quad \Phi_w\mapsto A_{-1}(s_0)(f_\Phi^{(s_0)}),$$

where $\Phi=\Phi_w\otimes\bigl(\bigotimes_{v\neq w}\Phi_v^0\bigr)$. The invariant distribution theorem [Mœglin et al. 1987; Lee and Zhu 1998] asserts that $A_{-1,w}$ factors through the quotient $R(V_w,\chi_{V_w})$. Lemma 2.1(1) shows that $A_{-1,w}(\Phi_w)$ is singular for every $\Phi_w\in S(V_w^{n'})$. If w is finite, then $\delta m=2r_w+2$ and $\delta n=r_w+1$, and hence $R(V_w,\chi_{V_w})$ is irreducible and nonsingular by [Kudla and Sweet 1997, Theorem 1.2], so that $A_{-1,w}$ must be zero. If w is real and ∇ is the element of the universal enveloping algebra of the complexified Lie algebra of $G(F_w)$ defined by (2.1) of [Ichino 2007], then $\nabla A_{-1,w}(\Phi_w)=0$. Since Proposition 2.2 of [Ichino 2007] asserts that $\nabla f_{\Phi_w}^{(s_0)}$ generates the submodule $R(V_w,\chi_{V_w})$ for a suitable choice of Φ_w, the map $A_{-1,w}$ must be zero. Consequently, $E(g;f_\Phi^{(s)})$ is holomorphic at $s=s_0$ for every $\Phi\in S(V^n(\mathbb{A}))$.

Next we consider the K_w-intertwining map

$$A_w:S(V_w^{n'})\to\mathscr{A},\quad \Phi_w\mapsto E(g;f_\Phi^{(s)})|_{s=s_0}-I(g;\Phi),$$

where $\Phi=\Phi_w\otimes\bigl(\bigotimes_{v\neq w}\Phi_v^0\bigr)$. The image of A_w lies in the space of singular automorphic forms by Lemma 3.1. We write $\mathscr{R}_w$ for the subspace of $\mathscr{A}$ spanned by residues $\operatorname{Res}_{s=s_0}E(g;f^{(s)})$, where $f^{(s)}$ is a holomorphic section of $I(s,\chi_V)$ of the form

$$f^{(s)}=f_w^{(s)}\otimes\Bigl(\bigotimes_{v\neq w}f_{\Phi_v^0}^{(s)}\Bigr),\qquad f_w^{(s)}\in I_w(s,\chi_{V_w}).$$

Then A_w induces a $G(F_w)$-intertwining map $R(V_w,\chi_{V_w})\to\mathscr{A}/\mathscr{R}_w$. The remaining part of the proof continues as in Section 3 of [Ichino 2007]. □

Theorem 2 is demonstrated in [Yamana 2011], provided that $\delta=1$ and $m\le n$. Since the proof in our general case can be done by the same technique, we shall omit most of the details. We define the functions $a(s,\chi)$ and $b(s,\chi)$ by taking the complete Hecke L-functions in place of the local abelian L-factors in the definition of $a_v(s,\chi_v)$ and $b_v(s,\chi_v)$. We define a normalized global intertwining operator by

$$M^\circ(s,\chi)=\frac{b(s,\chi)}{a(s,\chi)}M(s,\chi),$$

which is holomorphic in $\Re s>-\frac{1}{2}$ by Lemma 1.2 and (1-2).

Let $\mathscr{C}=\{W_v\}$ be a collection of local hermitian spaces of dimension m over D_v such that W_v is isometric to $V(F_v)$ for almost all v. We form a restricted tensor product $\Pi(\mathscr{C},\chi_V)=\bigotimes_v' R_{n'}(W_v,\chi_{V_v})$, which we can regard as a subrepresentation of $I(s_0,\chi_V)$. The proof of the following result is completely analogous to that of [Kudla and Rallis 1994, Theorem 3.1].

Proposition 3.2. *Assume that $m\le n$. Then*

$$\dim \operatorname{Hom}_{G(\mathbb{A})}(\Pi(\mathscr{C},\chi_V),\mathscr{A})\le 1.$$

If there is no global hermitian space with W_v as its completions, then

$$\dim \operatorname{Hom}_{G(\mathbb{A})}(\Pi(\mathscr{C},\chi_V),\mathscr{A})=0.$$

Next we are going to prove the special case of Theorem 2 in which $m=n$. Let $\mathscr{C}=\{V(F_v)\}$. Since Proposition 1.4(2) shows that the two intertwining maps $\Phi\mapsto E(g;f_\Phi^{(s)})|_{s=0}$ and $\Phi\mapsto I(g;\Phi)$ define elements of the space

$$\operatorname{Hom}_{G(\mathbb{A})}(\Pi(\mathscr{C},\chi_V),\mathscr{A}),$$

they must be proportional by Proposition 3.2. From Lemmas 2.1(2) and 3.1, they are nonvanishing, and the constant of proportionality is determined to be 2. □

We now suppose that $m<n$. Let $\mathscr{C}'$ be a collection of local hermitian spaces of dimension $2n-m$ obtained by adding a split space of suitable dimension to $\mathscr{C}$. By Proposition 1.4(4) and (5), $\Pi(\mathscr{C}',\chi_V)$ has a unique irreducible quotient $\Pi(\mathscr{C},\chi_V)$, and $M^\circ(-s_0,\chi_V)$ induces a nonzero intertwining map $\Pi(\mathscr{C}',\chi_V)\to\Pi(\mathscr{C},\chi_V)$. The same reasoning as in Section 4 of [Yamana 2011] implies the following result:

Proposition 3.3. *Suppose that $m<n$. Let $f^{(s)}$ be a standard section of $I(s,\chi_V)$ such that $f^{(s_0)}\in\Pi(\mathscr{C},\chi_V)$. Put $h^{(-s)}=M^\circ(s,\chi_V)f^{(s)}$.*

(1) *$E(g;f^{(s)})$ is holomorphic at $s=s_0$.*

(2) *$h^{(s)}$ is holomorphic at $s=-s_0$, $h^{(-s_0)}\in\Pi(\mathscr{C}',\chi_V)$, and*

$$\operatorname{Res}_{s=-s_0}E(g;h^{(s)})=-\operatorname{Res}_{s=s_0}\left[\frac{b(s,\chi_V)}{a(s,\chi_V)}\right]E(g;f^{(s)})\big|_{s=s_0}.$$

Lemma 3.4. *If $m<n$, then the image of the map $A_{-1}(-s_0)$ lies in the space of square integrable automorphic forms on $G(\mathbb{A})$.*

Proof. We use [Kudla and Sweet 1997, Proposition 6.2] and follow closely the guideline of the proof of [Kudla and Rallis 1994, Proposition 4.6]. □

Proposition 3.5. *If $m < n$, then the restriction of $A_{-1}(-s_0)$ to $\Pi(\mathscr{C}', \chi_V)$ is zero unless $\mathscr{C}$ is the set of localizations of a global space, in which case it defines a nonzero intertwining map $\Pi(\mathscr{C}, \chi_V) \to \mathscr{A}$.*

Proof. The image of $A_{-1}(-s_0)$ is completely reducible in view of Lemma 3.4. Thus the restriction of $A_{-1}(-s_0)$ to $\Pi(\mathscr{C}', \chi_V)$ must factor through the unique irreducible quotient $\Pi(\mathscr{C}, \chi_V)$. Proposition 3.2 shows that $\Pi(\mathscr{C}, \chi_V)$ makes no contribution unless $\mathscr{C}$ comes from a global space. It remains to check that $A_{-1}(-s_0)$ is nonzero on $\Pi(V, \chi_V)$. From Proposition 3.3(2) this amounts to proving that the holomorphic value $E(g; f_\Phi^{(s)})|_{s=s_0}$ is nonzero for a good choice of $\Phi \in S(V^n(\mathbb{A}))$.

Let $\beta_0 \in \mathrm{Her}_m(F)$ with $\nu(\beta_0) \neq 0$. Put

$$\beta = \begin{pmatrix} \mathbf{0} & 0 \\ 0 & \beta_0 \end{pmatrix} \in \mathrm{Her}_n(F), \qquad G_0 = \left\{ \left(\begin{array}{cc|cc} \mathbf{1}_{n-m} & & & \\ & a & & b \\ \hline & & \mathbf{1}_{n-m} & \\ & c & & d \end{array} \right) \in G \right\}.$$

Define $\Phi_0 \in S(V^m(\mathbb{A}))$ by $\Phi_0(y) = \Phi((0, y))$ for $y \in V^m(\mathbb{A})$. The nonvanishing can be proven by considering the β-th Fourier coefficient of $E(g; f_\Phi^{(s)})$ as in Section 6 of [Yamana 2011] (compare Theorem 4.9 of [Kudla and Rallis 1994]). The exponents of the $n - m + 1$ terms in this Fourier coefficient are distinct at $s = s_0$, so that there can be no cancellations among them. The first term is just the β_0-th Fourier coefficient of the central value of the Eisenstein series on $G_0(\mathbb{A})$ attached to the standard section $f_{\Phi_0}^{(s)}$. Lemma 2.1(2) now completes our proof. □

Corollary 3.6. *Suppose that $m \le n$. Let $f^{(s)}$ be a standard section of $I(s, \chi_V)$ such that $f^{(s_0)} \in \Pi(\mathscr{C}, \chi_V)$. If $\mathscr{C}$ cannot be the set of localizations of any global space, then $E(g; f^{(s)})|_{s=s_0}$ is identically zero.*

Proof. Propositions 3.2, 3.3(2) and 3.5 prove this corollary. □

The regularized Siegel–Weil formula can be deduced from Propositions 3.2 and 3.5.

Theorem 3.7. *Assume that $m < n$. Then there is a nonzero constant c_0 such that if holomorphic sections $f^{(s)}$ of $I(s, \chi_V)$ and $\Phi \in S(V^n(\mathbb{A}))$ satisfy the relation*

$$M^\circ(-s_0, \chi_V) f^{(-s_0)} = f_\Phi^{(s_0)},$$

then we have

$$\mathrm{Res}_{s=-s_0} E(g; f^{(s)}) = c_0 I(g; \Phi).$$

Finally, we prove Theorem 2 when $m < n$. Applying Proposition 3.3(2) and Theorem 3.7 to $h^{(-s)} = M^\circ(s, \chi_V) f_\Phi^{(s)}$, we see that

$$E(g; f_\Phi^{(s)})|_{s=s_0} = cI(g; \Phi),$$

where c is independent of Φ. One can prove that $c = 2$ in exactly the same manner as in Section 6 of [Yamana 2011]. $\square$

Appendix. Zeta integrals for $\mathrm{GL}_n(\boldsymbol{D})$

Let F be a local field of characteristic zero and $\boldsymbol{D}$ a division algebra central and of dimension δ^2 over F. We begin by reviewing the Godement–Jacquet construction of the local factors of representations of $G_n' = \mathrm{GL}_n(\boldsymbol{D})$. The Fourier transform $\hat{\phi} \in \mathscr{S}(\mathrm{M}_{ba}(\boldsymbol{D}))$ of $\phi \in \mathscr{S}(\mathrm{M}_{ab}(\boldsymbol{D}))$ is defined by

$$\hat{\phi}(x) = \int_{\mathrm{M}_{ab}(\boldsymbol{D})} \phi(y)\psi(\tau(xy))\,dy, \qquad x \in \mathrm{M}_{ba}(\boldsymbol{D}),$$

where the Haar measure dy is so chosen that

$$\int_{\mathrm{M}_{ab}(\boldsymbol{D})} \hat{\phi}({}^t y)\,dy = \phi(0).$$

In the archimedean case $S(\mathrm{M}_{ab}(\boldsymbol{D}))$ is the subspace of $\mathscr{S}(\mathrm{M}_{ab}(\boldsymbol{D}))$ as defined on p. 115 of [Godement and Jacquet 1972], and in the p-adic case $S(\mathrm{M}_{ab}(\boldsymbol{D})) = \mathscr{S}(\mathrm{M}_{ab}(\boldsymbol{D}))$.

Let π be an irreducible admissible representation of G_n'. We write $\pi^\vee$ for its admissible dual and denote the standard pairing on $\pi^\vee \boxtimes \pi$ by $\langle\,,\rangle$. For $s \in \mathbb{C}$, $\phi \in \mathscr{S}(\mathrm{M}_n(\boldsymbol{D}))$, $\xi \in \pi$ and $\xi^\vee \in \pi^\vee$ we set

$$Z^{GJ}(s, \phi, \xi \boxtimes \xi^\vee) = \int_{G_n'} \langle \pi(g)\xi, \xi^\vee \rangle \phi(g) |\nu(g)|_F^{s+n'/2}\,dg.$$

This integral converges in some half-plane and extends to a meromorphic function on the whole s-plane satisfying

$$Z^{GJ}(-s, \hat{\phi}, \xi^\vee \boxtimes \xi) = (-1)^{n(\delta-1)} \gamma^{GJ}\big(s + \tfrac{1}{2}, \pi, \psi\big) Z^{GJ}(s, \phi, \xi \boxtimes \xi^\vee).$$

Fix a pair $\chi = (\chi_1, \chi_2)$ of quasicharacters of $F^\times$. Recall $\chi^0 = \chi_1\chi_2$. We attach a section $s \mapsto \mathfrak{F}_\phi^{(s,\chi)}$ to each $\phi \in \mathscr{S}(\mathrm{M}_{n,2n}(\boldsymbol{D}))$ by setting

$$\mathfrak{F}_\phi^{(s,\chi)}(g) = \chi_1(\nu(g)) |\nu(g)|_F^{s+n'/2} \int_{G_n'} \phi((0, t)g) \chi^0(\nu(t)) |\nu(t)|_F^{2s+n'}\,dt.$$

This integral converges absolutely for sufficiently large $\Re s$. Observe that if ϕ belongs to $S(\mathrm{M}_{n,2n}(\boldsymbol{D}))$, then $\mathfrak{F}_\phi^{(s,\chi)} \in I(s, \chi)$ (compare (1-1)). For $\varphi \in \mathscr{S}(\mathrm{M}_{2n,n}(\boldsymbol{D}))$ we define a section $\mathfrak{F}_\varphi^{(s,\chi)}$ of $I(s,\chi)$ to be

$$\chi_2(\nu(g))^{-1}|\nu(g)|_F^{-s-n'/2}\int_{G'_n}\varphi\Bigl(g^{-1}\begin{pmatrix}t\\0\end{pmatrix}\Bigr)\chi^0(\nu(t))|\nu(t)|_F^{2s+n'}\,dt.$$

Lemma A.1. *For each $\phi\in S(\mathrm{M}_{n,2n}(\boldsymbol{D}))$,*

$$M(s,\chi)\mathfrak{F}_\phi^{(s,\chi)}=\frac{(-1)^{n(\delta-1)}\chi_1(-1)^{n'}}{\gamma^{GJ}\Bigl(2s-\dfrac{n'-1}{2},\chi^0\circ\nu_n,\psi\Bigr)}\mathfrak{F}_{\hat\phi}^{(-s,\rho(\chi)^{-1})}.$$

Proof. The case $n=\delta=1$ is discussed in Lemma 14.7.1 of [Jacquet 1972]. The proof is substantially the same. For $g\in G'_{2n}$ we put

$$\Psi_g(t)=\int_{\mathrm{M}_n(\boldsymbol{D})}\phi((t,x)g)\,dx$$

for $t\in\mathrm{M}_n(\boldsymbol{D})$. Then

$$\begin{aligned}
&M(s,\chi)\mathfrak{F}_\phi^{(s,\chi)}(g)\\
&\quad=\int_{\mathrm{M}_n(\boldsymbol{D})}\mathfrak{F}_\phi^{(s,\chi)}\Bigl(\begin{pmatrix}0&\mathbf{1}_n\\\mathbf{1}_n&0\end{pmatrix}\begin{pmatrix}\mathbf{1}_n&x\\0&\mathbf{1}_n\end{pmatrix}g\Bigr)\,dx\\
&\quad=\chi_1\bigl((-1)^{n'}\nu(g)\bigr)|\nu(g)|_F^{s+n'/2}\\
&\qquad\qquad\times\int_{\mathrm{M}_n(\boldsymbol{D})}\int_{G'_n}\phi\Bigl((0,t)\begin{pmatrix}0&\mathbf{1}_n\\\mathbf{1}_n&x\end{pmatrix}g\Bigr)\chi^0(\nu(t))|\nu(t)|_F^{2s+n'}\,dt\,dx\\
&\quad=\chi_1\bigl((-1)^{n'}\nu(g)\bigr)|\nu(g)|_F^{s+n'/2}\int_{\mathrm{M}_n(\boldsymbol{D})}\int_{G'_n}\phi\bigl((t,x)g\bigr)\chi^0\bigl(\nu(t)\bigr)|\nu(t)|_F^{2s}\,dt\,dx\\
&\quad=\chi_1(-1)^{n'}\chi_1\bigl(\nu(g)\bigr)|\nu(g)|_F^{s+n'/2}Z^{GJ}\Bigl(2s-\frac{n'}{2},\Psi_g,\chi^0\circ\nu_n\Bigr).
\end{aligned}$$

Since $\widehat{\Psi_g}(t)=|\nu(g)|_F^{-n'}\hat\phi\Bigl(g^{-1}\begin{pmatrix}t\\0\end{pmatrix}\Bigr)$,

$$\chi_1\bigl(\nu(g)\bigr)|\nu(g)|_F^{s+n'/2}Z^{GJ}\Bigl(\frac{n'}{2}-2s,\widehat{\Psi_g},(\chi^0\circ\nu_n)^{-1}\Bigr)=\mathfrak{F}_{\hat\phi}^{(-s,\rho(\chi)^{-1})}.$$

Lemma A.1 follows from the functional equation of $Z^{GJ}(s,\phi,\chi^0\circ\nu_n)$. □

Fix $A\in\mathrm{GL}_n(\boldsymbol{D})$. For a section $f^{(s)}$ of $I(s,\chi)$, the integral

$$l_A(f^{(s)})=\int_{\mathrm{M}_n(\boldsymbol{D})}f^{(s)}\Bigl(\begin{pmatrix}\mathbf{1}_n&0\\x&\mathbf{1}_n\end{pmatrix}\Bigr)\psi(\tau(Ax))\,dx$$

converges absolutely for $\Re s\gg 0$. In the p-adic case, Karel [1979] has proven that $l_A(f^{(s)})$ admits an entire analytic continuation to the whole s-plane and satisfies a functional equation

$$l_A \circ M(s,\chi) = \chi^0(\nu(A))^{-1}|\nu(A)|_F^{-2s}c(s,\chi,\psi)l_A$$

for some meromorphic function $c(s,\chi,\psi)$. The factor $c(s,\chi,\psi)$ is independent of the choice of A. Analogous results are proven in the archimedean case in [Wallach 1988]. The normalization $M^\dagger(s,\chi)$ of $M(s,\chi)$ is defined so that

$$l_A \circ M^\dagger(s,\chi) = \chi_2(-1)^{n'}\chi^0(\nu(A))^{-1}|\nu(A)|_F^{-2s}l_A.$$

Lemma A.2. *For each* $\Phi \in S(\mathrm{M}_{n,2n}(\boldsymbol{D}))$,

$$M^\dagger(s,\chi)\mathfrak{F}_\Phi^{(s,\chi)} = \chi_2(-1)^{n'}\mathfrak{F}_{\hat{\Phi}}^{(-s,\rho(\chi)^{-1})}.$$

Proof. It is enough to show that

$$l_A\big(\mathfrak{F}_{\hat{\Phi}}^{(-s,\rho(\chi)^{-1})}\big) = \chi^0(\nu(A))^{-1}|\nu(A)|_F^{-2s}l_A\big(\mathfrak{F}_\Phi^{(s,\chi)}\big).$$

Take $\phi_1,\phi_2 \in S(\mathrm{GL}_n(\boldsymbol{D}))$ and define $\Phi \in S(\mathrm{M}_{n,2n}(\boldsymbol{D}))$ by $\Phi(x,y) = \hat{\phi}_1(x)\phi_2(y)$. Then

$$\begin{aligned}
l_A\big(\mathfrak{F}_\Phi^{(s,\chi)}\big) &= \int_{\mathrm{M}_n(\boldsymbol{D})} \mathfrak{F}_\Phi^{(s,\chi)}\left(\begin{pmatrix} \mathbf{1}_n & 0 \\ x & \mathbf{1}_n \end{pmatrix}\right)\psi(\tau(Ax))\,dx \\
&= \int_{\mathrm{M}_n(\boldsymbol{D})}\int_{\mathrm{GL}_n(\boldsymbol{D})} \Phi\left((0,t)\begin{pmatrix} \mathbf{1}_n & 0 \\ x & \mathbf{1}_n \end{pmatrix}\right)\chi^0\big(\nu(t)\big)|\nu(t)|_F^{2s+n'}\,dt\,\psi\big(\tau(Ax)\big)\,dx \\
&= \int_{\mathrm{GL}_n(\boldsymbol{D})} \phi_1(-At^{-1})\phi_2(t)\chi^0\big(\nu(t)\big)|\nu(t)|_F^{2s}\,dt.
\end{aligned}$$

Similarly, $l_A\big(\mathfrak{F}_{\hat{\Phi}}^{(-s,\rho(\chi)^{-1})}\big)$ is equal to

$$\begin{aligned}
&\int_{\mathrm{M}_n(\boldsymbol{D})}\int_{\mathrm{GL}_n(\boldsymbol{D})} \phi_1(-t)\hat{\phi}_2(-xt)\chi^0(\nu(t))^{-1}|\nu(t)|_F^{-2s+n'}\psi(\tau(Ax))\,dt\,dx \\
&\qquad = \int_{\mathrm{GL}_n(\boldsymbol{D})} \phi_1(-t)\phi_2(t^{-1}A)\chi^0(\nu(t))^{-1}|\nu(t)|_F^{-2s}dt \\
&\qquad = \chi^0(\nu(A))^{-1}|\nu(A)|_F^{-2s}l_A(\mathfrak{F}_\Phi^{(s,\chi)}).
\end{aligned}$$

Since both $l_A\big(\mathfrak{F}_\Phi^{(s,\chi)}\big)$ and $l_A\big(\mathfrak{F}_{\hat{\Phi}}^{(-s,\rho(\chi)^{-1})}\big)$ are not identically zero for a suitable choice of ϕ_1 and ϕ_2, the proof is complete. □

The embedding i of $G'_n \times G'_n$ into G'_{2n} is given by

$$(g_1,g_2) \mapsto w_1\begin{pmatrix} g_1 & 0 \\ 0 & g_2 \end{pmatrix}w_1^{-1}, \quad w_1 = \begin{pmatrix} 2^{-1}\cdot\mathbf{1}_n & -2^{-1}\cdot\mathbf{1}_n \\ \mathbf{1}_n & \mathbf{1}_n \end{pmatrix}.$$

Let π be an irreducible admissible representation of G'_n. For $\xi \in \pi$, $\xi^\vee \in \pi^\vee$ and a section $f^{(s)}$ of $I(s,\chi)$, we define the zeta integral by

$$Z(\xi \boxtimes \xi^\vee, f^{(s)}) = \int_{G'_n} \langle \pi(g)\xi, \xi^\vee \rangle f^{(s)}(i(g,e))\, dg,$$

following [Piatetski-Shapiro and Rallis 1987a; Lapid and Rallis 2005]. This integral converges absolutely for $\Re s \gg 0$ and extends to a meromorphic function in s that satisfies the functional equation

$$Z\big(\xi \boxtimes \xi^\vee, M^\dagger(s,\chi) f^{(s)}\big) = \pi(-1)\gamma\big(s+\tfrac{1}{2}, \pi \times \chi, \psi\big) Z(\xi \boxtimes \xi^\vee, f^{(s)}).$$

Lapid and Rallis [2005] demonstrated the special case of the following result for $\delta = 1$ in a different manner. It was pointed out by Wee Teck Gan [2012] that there is a typo in [Lapid and Rallis 2005, (25)].

Proposition A.3. *For any irreducible admissible representation π of G'_n and any pair $\chi = (\chi_1, \chi_2)$ of quasicharacters of $F^\times$,*

$$\gamma(s, \pi \times \chi, \psi) = \gamma^{GJ}(s, \pi \otimes \chi_1, \psi)\gamma^{GJ}(s, \pi^\vee \otimes \chi_2, \psi).$$

Proof. Let $\mathscr{F}^{(s,\chi)}_\Phi$ be the translate of $\mathfrak{F}^{(s,\chi)}_\Phi$ by the element $w_1 \in G'_{2n}$. Then

$$\begin{aligned} Z\big(\xi \boxtimes \xi^\vee, \mathscr{F}^{(s,\chi)}_\Phi\big) &= \int_{G'_n} \langle \pi(g)\xi, \xi^\vee \rangle \chi_1(\nu(g)) |\nu(g)|_F^{s+n'/2} \\ &\quad \times \int_{G'_n} \Phi\left((0,t) w_1 \begin{pmatrix} g & 0 \\ 0 & \mathbf{1}_n \end{pmatrix}\right) \chi^0(\nu(t)) |\nu(t)|_F^{2s+n'}\, dt\, dg \\ &= \int_{G'_n \times G'_n} \big\langle (\pi \otimes \chi_1)(g)\xi, (\pi^\vee \otimes \chi_2)(t)\xi^\vee \big\rangle |\nu(gt)|_F^{s+n'/2} \Phi(g,t)\, dg\, dt. \end{aligned}$$

If $\Phi(x,y)$ is of the form $\phi_1(x)\phi_2(y)$, then the last integral is equal to

$$\big\langle Z^{GJ}(s, \pi \otimes \chi_1, \phi_1)\xi, Z^{GJ}(s, \pi^\vee \otimes \chi_2, \phi_2)\xi^\vee \big\rangle.$$

Piatetski-Shapiro and Rallis [1987a] employ this relation to calculate the unramified local zeta integrals.

We can see by Lemma A.2 that

$$\begin{aligned} Z\big(\xi \boxtimes \xi^\vee, M^\dagger(s,\chi)\mathscr{F}^{(s,\chi)}_\Phi\big) = \chi_2(-1)^{n'} &\int_{G'_n \times G'_n} \hat{\phi}_1(g)\hat{\phi}_2(t) \\ &\times |\nu(gt)|_F^{-s+n'/2} \big\langle (\pi \otimes \chi_1)(g^{-1})\xi, (\pi^\vee \otimes \chi_2)(-t^{-1})\xi^\vee \big\rangle\, dg\, dt. \end{aligned}$$

The stated relation follows upon combining these with the definitions of the gamma factors. $\square$

Let $\chi = 1$. Put $\Delta_s(g) = f_0^{(s-n'/2)}\left(w_1\begin{pmatrix} g & \\ & \mathbf{1}_n\end{pmatrix}\right)$ for $g \in G'_n$. Note that

$$\begin{aligned}\Delta_s(k_1 g k_2) &= f_0^{(s-n'/2)}\left(w_1\begin{pmatrix} k_1 g k_2 & \\ & \mathbf{1}_n\end{pmatrix}\right)\\ &= f_0^{(s-n'/2)}\left(i(k_1,k_1)w_1\begin{pmatrix} g & \\ & \mathbf{1}_n\end{pmatrix}\begin{pmatrix} k_2 & \\ & k_1^{-1}\end{pmatrix}\right) = \Delta_s(g)\end{aligned}$$

for $k_1, k_2 \in K_n$ and $g \in G'_n$. An explicit formula for this function is obtained in [Piatetski-Shapiro and Rallis 1987a, Proposition 6.4] in the case of symplectic or split even orthogonal groups. One can deduce from their argument a formula of the same type for the unit groups of simple algebras.

Lemma A.4. (1) *If F is a p-adic field and $g = k_1 d k_2$ with elements $k_1, k_2 \in K_n$ and $d = \mathrm{diag}[\varpi^{a_1}, \dots, \varpi^{a_n}]$, where ϖ is a generator of the maximal ideal of $\mathbb{O}$, and we put $q = |\nu(\varpi)|_F^{-1}$, then*

$$\Delta_s(g) = q^{-s\sum_{i=1}^n |a_i|}.$$

(2) *Assume that $F = \mathbb{R}$ or $\mathbb{C}$. Put $t = [F:\mathbb{R}]$. If $g = k_1 d k_2$ with $k_1, k_2 \in K_n$ and $d = \mathrm{diag}[d_1, \dots, d_n]$ with positive real numbers d_i, then*

$$\Delta_s(g) = 2^{n\delta t s}\prod_{i=1}^n (d_i^{-1} + d_i)^{-\delta t s}.$$

Lemma A.5. *If $\Re s > \delta(n-1)$, then Δ_s belongs to $L^1(G'_n)$.*

Proof. Put $\sigma = \Re s$. We consider the p-adic case. Proposition 1.5.2 of [Casselman 1995] gives a positive constant c such that

$$\begin{aligned}\int_{G'_n} |\Delta_s(g)|\,dg &\le c \sum_{a_1 \ge a_2 \ge \cdots \ge a_n} q^{-\sigma\sum_{i=1}^n |a_i|}\prod_{j=1}^n q^{\delta(n+1-2j)a_j}\\ &\le c\prod_{j=1}^n \sum_{a_j \in \mathbb{Z}} q^{-\sigma|a_j| + \delta(n+1-2j)a_j}\\ &= c\prod_{j=1}^n\left(\frac{1}{1-q^{\delta(n+1-2j)-\sigma}} + \frac{q^{\delta(2j-n-1)-\sigma}}{1-q^{\delta(2j-n-1)-\sigma}}\right).\end{aligned}$$

The archimedean case can be proven in the same way. $\square$

Lemma A.6. *If $\sigma > 0$, then the function $z \mapsto \Delta_\sigma(zg)$ is integrable over the center Z of G'_n for any $g \in G'_n$. Moreover, there exists a positive constant A_σ depending only on σ such that, for every $g \in G'_n$,*

$$\int_Z \Delta_\sigma(zg)\,dz \le A_\sigma.$$

Proof. In the p-adic case,

$$\int_Z \Delta_\sigma(zg)\,dz = \sum_{j\in\mathbb{Z}} q^{-\sigma\sum_{i=1}^n |a_i+\delta j|} \le \sum_{j\in\mathbb{Z}} q^{-\sigma|j|} = \frac{1+q^{-\sigma}}{1-q^{-\sigma}}.$$

The proof for the archimedean case is completely analogous. □

Recall that π is called square integrable if it admits a unitary central character and its matrix coefficients are square integrable modulo the center. For $(s_1, s_2) \in \mathbb{C}$, we write $I(s_1, s_2) = I(0, (\alpha_F^{s_1}, \alpha_F^{s_2}))$.

Proposition A.7. *If π is square integrable, $\Re s_1, \Re s_2 > -\delta/2$ and $f \in I(s_1, s_2)$, then the integral defining $Z(\xi \boxtimes \xi^\vee, f)$ is absolutely convergent.*

Proof. Put $\sigma = \min\{\Re s_1, \Re s_2\}$. Note that $(\alpha_F \circ \nu_{2n})^{s'} \cdot f_0^{(s)} \in I(s+s', s-s')$. By Lemma A.4, we can majorize $|f((g, e))|$ by $cf_0^{(\sigma)}((g, e))$ for some positive constant c. Our task is to check that for any $\sigma > -\delta/2$,

$$\int_{G'_n} \left|\langle \pi(g)\xi, \xi^\vee\rangle\right| \Delta_{\sigma+n'/2}(g)\,dg$$

is finite. Take a constant σ' so that $0 < \sigma' < \sigma + \delta/2$. The square of this integral is less than or equal to the product of the integrals

$$\int_{G'_n} \Delta_{2\sigma+n'-2\sigma'}(zg)\,dg$$

and

$$\begin{aligned}\int_{G'_n} \left|\langle \pi(g)\xi, \xi^\vee\rangle\right|^2 \Delta_{2\sigma'}(g)\,dg &= \int_{Z\backslash G'_n} \left|\langle \pi(\dot g)\xi, \xi^\vee\rangle\right|^2 \int_Z \Delta_{2\sigma'}(z\dot g)\,dz\,d\dot g\\ &= A_{2\sigma'} \int_{Z\backslash G'_n} \left|\langle \pi(\dot g)\xi, \xi^\vee\rangle\right|^2 d\dot g,\end{aligned}$$

both of which are finite, the first by Lemma A.5 and the second by Lemma A.6. □

Acknowledgements

We would like to thank the National University of Singapore, where this paper was partly written.

References

[Casselman 1995] W. Casselman, "Introduction to the theory of admissible representations of p-adic reductive groups", preprint, 1995.

[Gan 2000] W. T. Gan, "A Siegel–Weil formula for exceptional groups", *J. Reine Angew. Math.* **528** (2000), 149–181. MR 2001k:11089 Zbl 1005.11022

[Gan 2012] W. T. Gan, "Doubling zeta integrals and local factors for metaplectic groups", *Nagoya Math. J.* **208** (2012), 67–95. MR 3006697 Zbl 06127268

[Gan and Takeda 2011] W. T. Gan and S. Takeda, "On the regularized Siegel–Weil formula (the second term identity) and non-vanishing of theta lifts from orthogonal groups", *J. Reine Angew. Math.* **659** (2011), 175–244. MR 2012i:22029 Zbl 05971444

[Godement and Jacquet 1972] R. Godement and H. Jacquet, *Zeta functions of simple algebras*, Lecture Notes in Mathematics **260**, Springer, Berlin, 1972. MR 49 #7241 Zbl 0244.12011

[Howe 1981] R. Howe, "Automorphic forms of low rank", pp. 211–248 in *Noncommutative harmonic analysis and Lie groups* (Marseille, 1980), edited by J. Carmona and M. Vergne, Lecture Notes in Math. **880**, Springer, Berlin, 1981. MR 83j:10033 Zbl 0463.10015

[Ichino 2004] A. Ichino, "A regularized Siegel–Weil formula for unitary groups", *Math. Z.* **247**:2 (2004), 241–277. MR 2005g:11082 Zbl 1071.11022

[Ichino 2007] A. Ichino, "On the Siegel–Weil formula for unitary groups", *Math. Z.* **255**:4 (2007), 721–729. MR 2007m:11063 Zbl 1223.11056

[Jacquet 1972] H. Jacquet, *Automorphic forms on* GL(2)*, II*, Lecture Notes in Mathematics **278**, Springer, Berlin, 1972. MR 58 #27778 Zbl 0243.12005

[Karel 1979] M. L. Karel, "Functional equations of Whittaker functions on p-adic groups", *Amer. J. Math* **101**:6 (1979), 1303–1325. MR 81b:10021 Zbl 0427.22011

[Kudla 1994] S. S. Kudla, "Splitting metaplectic covers of dual reductive pairs", *Israel J. Math.* **87**:1-3 (1994), 361–401. MR 95h:22019 Zbl 0840.22029

[Kudla and Rallis 1988a] S. S. Kudla and S. Rallis, "On the Weil–Siegel formula", *J. Reine Angew. Math.* **387** (1988), 1–68. MR 90e:11059 Zbl 0644.10021

[Kudla and Rallis 1988b] S. S. Kudla and S. Rallis, "On the Weil–Siegel formula, II: The isotropic convergent case", *J. Reine Angew. Math.* **391** (1988), 65–84. MR 90e:11060 Zbl 0644.10022

[Kudla and Rallis 1990a] S. S. Kudla and S. Rallis, "Degenerate principal series and invariant distributions", *Israel J. Math.* **69**:1 (1990), 25–45. MR 91e:22016 Zbl 0708.22005

[Kudla and Rallis 1990b] S. S. Kudla and S. Rallis, "Poles of Eisenstein series and L-functions", pp. 81–110 in *Festschrift in honor of I. I. Piatetski-Shapiro on the occasion of his sixtieth birthday, 2* (Tel Aviv, 1989), edited by S. Gelbart et al., Israel Math. Conf. Proc. **3**, Weizmann, Jerusalem, 1990. MR 94e:11054 Zbl 0712.11029

[Kudla and Rallis 1994] S. S. Kudla and S. Rallis, "A regularized Siegel–Weil formula: the first term identity", *Ann. of Math.* (2) **140**:1 (1994), 1–80. MR 95f:11036 Zbl 0818.11024

[Kudla and Sweet 1997] S. S. Kudla and W. J. Sweet, Jr., "Degenerate principal series representations for $\mathrm{U}(n, n)$", *Israel J. Math.* **98** (1997), 253–306. MR 98h:22021 Zbl 0896.22007

[Lapid and Rallis 2005] E. M. Lapid and S. Rallis, "On the local factors of representations of classical groups", pp. 309–359 in *Automorphic representations, L-functions and applications: progress and prospects*, edited by J. W. Cogdell et al., Ohio State Univ. Math. Res. Inst. Publ. **11**, de Gruyter, Berlin, 2005. MR 2006j:11071 Zbl 1188.11023

[Lee 2007] S. T. Lee, "Covariants of $\mathrm{Sp}_n(\mathbb{C})$ and degenerate principal series of $\mathrm{GL}_n(\mathbb{H})$", *J. Funct. Anal.* **253**:1 (2007), 18–42. MR 2008i:22010 Zbl 1139.22011

[Lee and Zhu 1998] S. T. Lee and C.-B. Zhu, "Degenerate principal series and local theta correspondence", *Trans. Amer. Math. Soc.* **350**:12 (1998), 5017–5046. MR 99c:22021 Zbl 0901.22008

[Li 1989] J.-S. Li, "Singular unitary representations of classical groups", *Invent. Math.* **97**:2 (1989), 237–255. MR 90h:22021 Zbl 0694.22011

[Mínguez 2009] A. Mínguez, "Sur l'irréductibilité d'une induite parabolique", *J. Reine Angew. Math.* **629** (2009), 107–131. MR 2010h:22025 Zbl 1172.22008

[Mœglin et al. 1987] C. Mœglin, M.-F. Vignéras, and J.-L. Waldspurger, *Correspondances de Howe sur un corps p-adique*, Lecture Notes in Mathematics **1291**, Springer, Berlin, 1987. MR 91f:11040 Zbl 0642.22002

[Piatetski-Shapiro and Rallis 1987a] I. Piatetski-Shapiro and S. Rallis, "L-functions for the classical groups", pp. 1–52 in *Explicit constructions of automorphic L-functions*, edited by S. Gelbart et al., Lecture Notes in Math. **1254**, Springer, Berlin, 1987.

[Piatetski-Shapiro and Rallis 1987b] I. Piatetski-Shapiro and S. Rallis, "Rankin triple L functions", *Compos. Math.* **64**:1 (1987), 31–115. MR 89k:11037 Zbl 0637.10023

[Sahi 1995] S. Sahi, "Jordan algebras and degenerate principal series", *J. Reine Angew. Math.* **462** (1995), 1–18. MR 96d:22022 Zbl 0822.22006

[Scharlau 1985] W. Scharlau, *Quadratic and Hermitian forms*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences] **270**, Springer, Berlin, 1985. MR 86k:11022 Zbl 0584.10010

[Sécherre 2009] V. Sécherre, "Proof of the Tadić conjecture (U0) on the unitary dual of $\mathrm{GL}_m(D)$", *J. Reine Angew. Math.* **626** (2009), 187–203. MR 2009k:22027 Zbl 1170.22009

[Tan 1998] V. Tan, "A regularized Siegel–Weil formula on $\mathrm{U}(2, 2)$ and $\mathrm{U}(3)$", *Duke Math. J.* **94**:2 (1998), 341–378. MR 99f:11065 Zbl 0989.11025

[Tan 1999] V. Tan, "Poles of Siegel Eisenstein series on $\mathrm{U}(n, n)$", *Canad. J. Math.* **51**:1 (1999), 164–175. MR 2000e:11073 Zbl 0963.11028

[Wallach 1988] N. R. Wallach, "Lie algebra cohomology and holomorphic continuation of generalized Jacquet integrals", pp. 123–151 in *Representations of Lie groups* (Kyoto, Hiroshima, 1986), edited by K. Okamoto and T. Oshima, Adv. Stud. Pure Math. **14**, Academic Press, Boston, MA, 1988. MR 91d:22014 Zbl 0714.17016

[Weil 1965] A. Weil, "Sur la formule de Siegel dans la théorie des groupes classiques", *Acta Math.* **113** (1965), 1–87. MR 36 #6421 Zbl 0161.02304

[Weil 1974] A. Weil, *Basic number theory*, 3rd ed., Die Grundlehren der Mathematischen Wissenschaften **144**, Springer, New York, 1974. MR 55 #302 Zbl 0326.12001

[Yamana 2011] S. Yamana, "On the Siegel–Weil formula: the case of singular forms", *Compos. Math.* **147**:4 (2011), 1003–1021. MR 2822859 Zbl 05938388

[Yamana 2013] S. Yamana, "On the Siegel–Weil formula for quaternionic unitary groups", *Amer. J. Math* (2013). To appear.

[Zhang 1995] G. K. Zhang, "Jordan algebras and generalized principal series representations", *Math. Ann.* **302**:4 (1995), 773–786. MR 97h:22012 Zbl 0829.22023

Received March 12, 2012.

SHUNSUKE YAMANA
FACULTY OF MATHEMATICS
KYUSHU UNIVERSITY
744, NISHI-KU
MOTOOKA 819-0395
JAPAN

yamana@math.kyushu-u.ac.jp

Guidelines for Authors

Authors may submit manuscripts at msp.berkeley.edu/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 264 No. 1 July 2013
