Pacific Journal of Mathematics

Volume 250 No. 2 April 2011

PACIFIC JOURNAL OF MATHEMATICS

http://www.pjmath.org

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

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

The subscription price for 2011 is US $420/year for the electronic version, and $485/year for print and electronic.
Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and the Science Citation Index.

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840, is published monthly except July and August. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFlow™ from Mathematical Sciences Publishers.

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

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

REALIZING PROFINITE REDUCED SPECIAL GROUPS

VINCENT ASTIER AND HUGO MARIANO

Special groups are an axiomatization of the algebraic theory of quadratic forms over fields. It is known that any finite reduced special group is the special group of some field. We show that any special group that is the projective limit of a projective system of finite reduced special groups is also the special group of some field.

1. Introduction

The theory of special groups is an axiomatization of the algebraic theory of quadratic forms, introduced in [Dickmann and Miraglia 2000]. The class of special groups, together with its morphisms, forms a category. As for other such axiomatizations, the main examples of special groups are provided by fields, in this case by applying the special group functor, which associates to each field F a special group $G(F)$ describing the theory of quadratic forms over F.

The category of special groups is equivalent to that of abstract Witt rings via covariant functors, while the category of reduced special groups is equivalent, via the restriction of the same covariant functors, to the category of reduced abstract Witt rings (see [Dickmann and Miraglia 2000, 1.25 and 1.26]; recall that the special group of a field F is reduced if and only if F is formally real and Pythagorean). The category of reduced special groups is also equivalent, via contravariant functors, to the category of abstract spaces of orderings; see Chapter 3 of the same reference.

The question whether it is possible to realize every (reduced) special group as the special group of some (formally real, Pythagorean) field is still open, but the case of finite reduced special groups (actually of reduced special groups of finite chain length) has been positively answered by the combination of two results: Kula [1979], building on techniques introduced in [Bröcker 1977] for the field case, showed that the product of two finite special groups of (formally real, Pythagorean) fields is still the special group of some (formally real, Pythagorean) field; then Marshall [1980] showed that every finite reduced special group can be constructed from the special group of any real closed field by applying a finite number of times the operations of product and extension. (Marshall's result is actually stated and

MSC2000: 03C65, 11E81.
Keywords: quadratic forms, special groups, projective limit, profinite, representation theorems.

proved for abstract spaces of orderings.) Since the extension of the special group of a (formally real, Pythagorean) field is still the special group of a (formally real, Pythagorean) field, it shows that every finite reduced special group (or reduced special group of finite chain length) is realized as the special group of a field.

After finite reduced special groups, the simplest objects to consider are probably projective limits of finite reduced special groups, that is, profinite reduced special groups. They have already been studied, for example, in [Astier and Tressl 2005; Lira de Lima 1997; Mariano 2003], and notably in [Kula et al. 1984], where the question of the realization of these special groups by fields is considered and where it is shown (as Corollary 4.7) that every profinite reduced special group is *isomorphic to a quotient* of the reduced special group of some field.

In this paper, we improve on this result by showing that every profinite reduced special group is *isomorphic* to the special group of some (necessarily formally real and Pythagorean) field.

2. Preliminaries

Definition 2.1. Let A, B, A', B' be objects in a category $\mathscr{C}$, and let $\lambda : A \to B$, $\lambda' : A' \to B'$ be $\mathscr{C}$-morphisms. Then λ, λ' are said to be *naturally identified* (in symbols, $\lambda \cong \lambda'$) if and only if there are $\mathscr{C}$-isomorphisms $i_A : A \to A'$, $i_B : B \to B'$ such that the following diagram

$$\begin{tikzcd} A \arrow[r, "i_A"] \arrow[d, "\lambda"'] & A' \arrow[d, "\lambda'"] \\ B \arrow[r, "i_B"] & B' \end{tikzcd}$$

commutes. In this case, we also say that λ and λ' are *naturally identified via* i_A, i_B.

On special groups. We assume some familiarity with the theory of special groups, as presented in [Dickmann and Miraglia 2000], and only introduce the following notation:

If G is a special group, $\mathrm{Ssat}(G)$ denotes the poset of saturated subgroups of G, ordered by inclusion. We recall that if $\Delta \in \mathrm{Ssat}(G)$, then G/Δ is a reduced special group if and only if $\Delta \subsetneq G$, if and only if $-1 \notin \Delta$.

Definition 2.2. A profinite reduced special group is the projective limit of a projective system of finite reduced special groups.

If $(G'_i, f'_{ij})_{i \le j \in I}$ is a projective system of finite reduced special groups, where $(I, \le)$ is a downward directed poset, and if G is the projective limit of this system, the fact that G is indeed a special group (with the structure induced by its inclusion in the product $\prod_{i \in I} G'_i$) follows immediately from [Dickmann and Miraglia 2003, Theorem 3.24]. Moreover, as proved in [Lira de Lima 1997, Proposition 1.9.11], it is always possible to describe G as the projective limit of a projective system $(G_i, f_{ij})_{i\le j\in I}$ having the following properties:

(1) For every $i \in I$, G_i is G/Δ_i with Δ_i saturated subgroup of G of finite index;

(2) For every $i \le j \in I$, $\Delta_i \subseteq \Delta_j$ and f_{ij} is the canonical projection of special groups induced by this inclusion.

We briefly sketch the argument: Let $\iota : G \hookrightarrow \prod_{j\in I} G'_j$ be the canonical embedding given by the definition of projective limit, and let $\pi_i : \prod_{j\in I} G'_j \to G'_i$ be the canonical projection. We define $\Delta_i := \ker(\pi_i \circ \iota)$, $G_i := G/\Delta_i$ and, for $i \le j \in I$, f_{ij} to be the canonical projection induced by $\Delta_i \subseteq \Delta_j$. The system $(G_i, f_{ij})_{i\le j\in I}$ is a projective system, whose projective limit is isomorphic to G, via the map $g \in G \mapsto (g.\Delta_i)_{i\in I} \in \varprojlim(G/\Delta_i, f_{ij})_{i\le j\in I}$.

Remark 2.3. If $\mathcal{M} = (M_i, f_{ij})_{i\le j\in I}$ is any projective system, and if $i' \in I$, by restricting this system to the set $I' := \{i \in I \mid i \le i'\}$ we obtain a new system

$$\mathcal{M}' := (M_i, f_{ij})_{i\le j\in I'}.$$

Since I' is coinitial in I, $\mathcal{M}$ and $\mathcal{M}'$ have isomorphic projective limits, and $\mathcal{M}'$ possesses the following extra property:

(3) The index set of the projective system has a maximum element.

Definition 2.4. We call *adequate* a projective system of special groups that satisfies conditions (1), (2) and (3) above.

We will adhere to the following convention throughout this paper: Let $(I, \le)$ be a downward directed poset. If $(I, \le)$ has a maximum element, we will denote it by $\top$, and if $(I, \le)$ has a minimum element (which happens for instance if I is finite), we will denote it by $\bot$.

Let G_0, G_1 be abstract groups and denote by $\pi_0 : G_0 \times G_1 \twoheadrightarrow G_0 : (g_0, g_1) \mapsto g_0$, $\pi_1 : G_0 \times G_1 \twoheadrightarrow G_1 : (g_0, g_1) \mapsto g_1$ the canonical projections and by $\iota_0 : G_0 \rightarrowtail G_0 \times G_1 : g_0 \mapsto (g_0, 1)$, $\iota_1 : G_1 \rightarrowtail G_0 \times G_1 : g_1 \mapsto (1, g_1)$ the canonical injections.

The statements in the next paragraph are straightforward.

Fact 2.5. *Let G_0, G_1 be special groups. Then the canonical map*

$$\begin{aligned} \psi : \mathrm{Ssat}(G_0 \times G_1) &\to \mathrm{Ssat}(G_0) \times \mathrm{Ssat}(G_1) \\ \Delta &\mapsto (\iota_0^{-1}[\Delta], \iota_1^{-1}[\Delta]) = (\pi_0[\Delta], \pi_1[\Delta]) \end{aligned}$$

is an order-preserving bijection, whose inverse is $(\Delta_0, \Delta_1) \overset{\psi^{-1}}{\mapsto} \Delta_0 \times \Delta_1$. *In particular, if* $\Delta \in \mathrm{Ssat}(G_0 \times G_1)$ *and* $(\Delta_0, \Delta_1) := (\iota_0^{-1}[\Delta], \iota_1^{-1}[\Delta])$, *then* $\Delta = \Delta_0 \times \Delta_1$ *and* Δ *is proper if and only if* Δ_0 *or* Δ_1 *is proper. Moreover*:

- *The canonical surjective morphism of special groups*

$$G_0 \times G_1 \twoheadrightarrow G_0/\Delta_0 \times G_1/\Delta_1$$

induces a natural isomorphism of special groups[1]

$$\bar{q}_\Delta : (G_0 \times G_1)/\Delta \xrightarrow{\cong} G_0/\Delta_0 \times G_1/\Delta_1.$$

- *If* $\Delta \subseteq \Delta' \in \mathrm{Ssat}(G_0 \times G_1)$, *then the projection* $(G_0 \times G_1)/\Delta \twoheadrightarrow (G_0 \times G_1)/\Delta'$ *is naturally identified*, *via the isomorphisms* $\bar{q}_\Delta, \bar{q}_{\Delta'}$, *with the* (*product*) *projection* $G_0/\Delta_0 \times G_1/\Delta_1 \twoheadrightarrow G_0/\Delta'_0 \times G_1/\Delta'_1$.

On projective systems of (valued) fields. Let $(I, \leq)$ be a poset. For each $i, j \in I$ such that $i \leq j$ we define $d(i, j) := \max\{\text{length of a chain from } i \text{ to } j\} \in \mathbb{N} \cup \{\infty\}$. If $i \nleq j$ then we set $d(i, j) := -\infty$. Of course, if i and j are comparable, we have $d(i, j) = d(j, i)$ if and only if $i = j$, if and only if $d(i, j) = 0$.

We will often consider $(I, \leq)$ as a directed graph whose vertices are the elements of I, and where there is an edge from i to j if and only if $i \leq j$ and $d(i, j) = 1$.

We first remark that it is possible to describe some projective systems of fields as projective systems whose morphisms are all inclusions.

Remark 2.6. Let $\mathcal{F} := (F_i, f_{ij})_{i \leq j \in I}$ be a projective system of fields over a downward directed poset $(I, \leq)$ with maximum element $\top \in I$. Then there is an isomorphic projective system of fields $\mathcal{F}' = (F'_i, \iota_{ij})_{i \leq j \in I}$ such that, if $i \leq j \in I$, then $F'_i \subseteq F'_j$ and the morphism of fields $\iota_{ij} : F'_i \to F'_j$ is the inclusion. The projective limit of the system $\mathcal{F}$ is thus isomorphic to the intersection of the fields F'_i, $i \in I$.

We briefly sketch the argument. For each $i \in I$, we define $F'_i := f_{i\top}[F_i] \subseteq F_\top$. Since for $i \leq j \in I$, $f_{i\top} = f_{j\top} \circ f_{ij}$, we obtain $F'_i \subseteq F'_j$, so we can define ι_{ij} to be this inclusion. It follows that $\mathcal{F}$ and $\mathcal{F}'$ are isomorphic via the morphisms $(f_{i\top})_{i \in I}$. Therefore: $\varprojlim(F_j, f_{ij})_{i \leq j \in I} \cong \varprojlim(F'_j, \iota_{ij})_{i \leq j \in I} \cong \bigcap_{i \in I} F'_i \subseteq F'_\top$.

The next results lead to Corollary 2.10, which shows that any finite projective system of fields of characteristic zero, whose index set has a maximum element, is isomorphic to the projective system given by the residues of a finite projective system of valued fields. We first fix some notation:

If (K, v) is a valued field, we denote by Kv or by $\overline{K}$ (if there is no risk of confusion about which valuation we consider) the residue field of v, by vK its value group, by O_K the valuation ring associated to v and by M_K its maximal ideal (if there is no ambiguity about the valuation v under consideration). If $a \in O_K$, we denote by av or $\bar{a}$ (once again if there is no risk of confusion) the class of a in the residue field $\overline{K}$. Finally, if v has rank one, K^v denotes a completion of K with respect to v.

[1] That are reduced if Δ is proper or, otherwise, the trivial special group $\{1\}$.

If $\mathcal{F} = (F_i, \xi_{ij})_{i\leq j\in I}$ is a projective system of fields, we denote by $G(\mathcal{F})$ the system $(G(F_i), G(\xi_{ij}))_{i\leq j\in I}$ obtained from $\mathcal{F}$ by applying the special group functor G.

If $\mathcal{F} = ((F_i, v_i), \xi_{ij})_{i\leq j\in I}$ is a projective system of valued fields, we denote by $\operatorname{res}\mathcal{F}$ or $\mathcal{F}v$ the induced residue projective system $(F_i v_i, (\xi v)_{ij})_{i\leq j\in I}$, where the $(\xi v)_{ij}$ are the induced morphisms of fields.

If a projective system of fields or of valued fields is denoted by $(F_i)_{i\in I}$ or $(F_i, v_i)_{i\in I}$, without mention of the morphisms, it means that the morphisms are all inclusions (from a field within all fields with larger index).

Lemma 2.7. *Let (K, v) be a henselian valued field of residue characteristic zero and let L be a subfield of K. Let N be a subfield of $\overline{K}$ such that $\overline{L} \subseteq N \subseteq \overline{K}$. Then there is a field M such that $L \subseteq M \subseteq K$ and $\overline{M} = N$. Moreover, if $[N : \overline{L}]$ is algebraic, respectively finite, then M can be chosen such that $[M : L]$ is algebraic, respectively finite.*

Proof. Write $N = \bar{L}(X)(\alpha_i, i \in \beta)$, where X is a transcendence basis of N over $\bar{L}$ and $(\alpha_i, i \in \beta)$ is a (possibly infinite) tuple of elements that are algebraic over $\bar{L}(X)$, indexed by an ordinal β. Let Y be a set of transcendental elements over L such that $\overline{Y} = X$. By [Engler and Prestel 2005, Corollary 2.2.2], the restriction of v to $L(Y)$ is the Gauss extension of v from L to $L(Y)$. In particular, $\overline{L(Y)} = \bar{L}(X)$.

We now proceed by induction on $k \in \beta$ to find elements $a_i \in K$, $i < k$, such that $\overline{L(Y)(a_i, i < k)} = \bar{L}(X)(\alpha_i, i < k)$.

If $k = 0$ there is nothing to prove since $\overline{L(Y)} = \bar{L}(X)$.

Assume we have found all a_i for $i < k$. Let $N_k = L(Y)(a_i, i < k)$ and $M_k = \bar{L}(X)(\alpha_i, i < k)$. By hypothesis we have $\overline{N}_k = M_k$. Let $P \in O_K[T]$ be a unitary polynomial such that $\overline{P}$ is the minimal polynomial of α_k over $\bar{L}(X)$. Let a_k be a root of P in K such that $\bar{a}_k = \alpha_k$ (it exists since (K, v) is henselian of residue characteristic zero). We have $\overline{N_k(a_k)} \supseteq M_k(\alpha_k)$ and the fundamental inequality [Engler and Prestel 2005, Theorem 3.3.4] tells us that

$$[\overline{N_k(a_k)} : \overline{N_k}] \leq [N_k(a_k) : N_k] (\leq \deg P).$$

Since $[M_k(\alpha_k) : M_k] = \deg P$, it follows that $\overline{N_k(a_k)} = M_k(\alpha_k)$, which is the desired result. □

Definition 2.8. Let (K, v) be a valued field and let $(E_i)_{i<n}$ and $(F_i)_{i<n}$ be two sequences of fields of the same length n. We say that $(F_i)_{i<n}$ is a good residue of $(E_i)_{i<n}$ in (K, v) if

(1) $E_i \subseteq K$ and $F_i \subseteq \overline{K}$ for $i < n$;

(2) For every $A \subseteq \{0, \dots, n-1\}$, $\overline{\langle E_i, i \in A\rangle} = \langle F_i, i \in A\rangle$ (where $\langle L_i, i \in A\rangle$ denotes the compositum of the fields L_i).

Lemma 2.9. *Let (K, v) be a henselian valued field of residue characteristic zero, and let $(E_i)_{i<n}$ and $(F_i)_{i<n}$ be two sequences of fields of length n such that $(F_i)_{i<n}$ is a good residue of $(E_i)_{i<n}$ in (K, v). Let $(F'_i)_{i<m}$ be a sequence of subfields of $\overline{K}$, and let, for $i \in \{0, \dots m-1\}$*

$$A_i = \{j \in \{0, \dots, n-1\} \mid F_j \subseteq F'_i\}.$$

Then there is a sequence $(E'_i)_{i<m}$ of subfields of K such that

(1) *for every $i \in \{0, \dots, m-1\}$ and every $j \in A_i$, $E_j \subseteq E'_i$ and* $\operatorname{trdeg} E'_i \mid E_j = \operatorname{trdeg} F'_i \mid F_j$;

(2) *$(F'_i)_{i<m}$ is a good residue of $(E'_i)_{i<m}$.*

Proof. We will use the following reformulation of Remark 4.1.2(3) in [Engler and Prestel 2005]:

Fact 1. Let (N, w) be a valued field and let $P, Q \in O_N[T]$ and $R \in N[T]$ be such that $P = QR$. Assume that Q is primitive (that is, $w'(Q) = 0$, where w' is the Gauss extension of w to $N[T]$, i.e., $\min_{i \le k} w(a_i) = 0$ if one writes $Q = a_0 + \cdots + a_k T^k$). Then $R \in O_N[T]$.

Proof. Write $P = aP_1$ and $R = cR_1$ with $a, c \in N$ and $P_1, R_1 \in N[T]$ such that $w'(P_1) = w'(R_1) = 0$ (so $P_1, R_1 \in O_N[T]$). Then $w(c) = w'(Q) + w'(cR_1) = w'(QcR_1) = w'(QR) = w'(P) \ge 0$ since $P \in O_N[T]$. This yields $R = cR_1$ with $w'(R) = w(c) + w'(R_1) = w(c) \ge 0$, i.e., $R \in O_N[T]$. □

We next fix some notation. For $A \subseteq \{0, \dots, n-1\}$ we denote by F_A the field $\langle F_i, i \in A\rangle$ and similarly by E_A the field $\langle E_i, i \in A\rangle$.

For $i < m$ let $X_i = \{x_{i1}, \dots, x_{ik_i}\}$ be a transcendence basis of F'_i over $F_{A_i} = \bar{E}_{A_i}$, and let $Y_i = \{y_{i1}, \dots, y_{ik_i}\} \subseteq K$ be a set of transcendental elements over E_{A_i} such that $\overline{Y}_i = X_i$. Note that by [Engler and Prestel 2005, corollary 2.2.2], it implies that the restriction of v to $E_{A_i}(Y_i)$ is the Gauss extension of v from E_{A_i} to $E_{A_i}(Y_i)$. In particular we have $\overline{E_{A_i}(Y_i)} = \overline{E_{A_i}}(X_i) = F_{A_i}(X_i)$ (the last equality holds because $(F_i)_{i<n}$ is a good residue of $(E_i)_{i<n}$).

Write $F'_i = F_{A_i}(X_i)(\alpha_i)$, where $\alpha_i = (\alpha_{ij})_{j \in \beta_i}$ is a (possibly infinite) tuple of elements algebraic over $F_{A_i}(X_i)$. For $i < m$ and $j \in \beta_i$ let $P_{ij} \in O_{E_{A_i}(Y_i)}[T]$ be a unitary polynomial such that $\overline{P}_{ij}$ is the minimal polynomial of α_{ij} over $\overline{E_{A_i}(Y_i)} = F_{A_i}(X_i)$, and let $a_{ij} \in O_K$ be a root of P_{ij} with $\bar{a}_{ij} = \alpha_{ij}$ (a_{ij} exists since (K, v) is henselian of residue characteristic zero). We take for E'_i the field $E_{A_i}(Y_i)(a_i)$, where $a_i = (a_{ij})_{j \in \beta_i}$. The first conclusion of the lemma is obviously satisfied. Let $A \subseteq \{0, \dots, m-1\}$.

Claim. Let L be a subfield of K such that (L, v) is henselian, $\langle F'_i, i \in A\rangle \subseteq \bar{L}$, and $\langle E_j, j \in A_i, i \in A\rangle(Y_i, i \in A) \subseteq L$. Then $a_i \in L$ for every $i \in A$, i.e., $L \supseteq \langle E'_i, i \in A\rangle$.

Proof. Let $i \in A$ and $j \in \beta_i$. Since $\alpha_{ij} \in \bar{L}$ and (L, v) is henselian (of residue characteristic zero), there is $b_{ij} \in O_L$ such that $\bar{b}_{ij} = \alpha_{ij}$ and b_{ij} is a root of P_{ij}. Assume $b_{ij} \neq a_{ij}$. Then we can write $P_{ij}(T) = (T - a_{ij})(T - b_{ij})R(T)$ in $E_{A_i}(Y_i, a_{ij}, b_{ij})$. But P_{ij}, $(T - a_{ij})$, $(T - b_{ij})$ each lie in $O_{E_{A_i}(Y_i, a_{ij}, b_{ij})}[T]$ and $(T - a_{ij})(T - b_{ij})$ is primitive, so by Fact 1 we have $R(T) \in O_{E_{A_i}(Y_i, a_{ij}, b_{ij})}[T]$. Going to the residue field $\overline{K}$ we get $\overline{P}_{ij}(T) = (T - \alpha_{ij})^2 \overline{R}(T)$, so α_{ij} is root of order at least 2 of $\overline{P}_{ij}$, which is impossible since $\overline{P}_{ij}$ is the minimal polynomial of α_{ij} and char $\overline{K} = 0$. So $a_{ij} = b_{ij} \in L$. End of proof of the claim. □

We have $E'_A = \langle E'_i, i \in A\rangle = \langle \langle E_j, j \in A_i\rangle(Y_i)(a_i), i \in A\rangle = \langle E_j, j \in A_i, i \in A\rangle(Y_i, i \in A)(a_i, i \in A)$, and

$$\begin{aligned}\overline{\langle E_j, j \in A_i, i \in A\rangle(Y_i, i \in A)} &= \overline{\langle E_j, j \in A_i, i \in A\rangle}(X_i, i \in A)\\ &= \langle F_j, j \in A_i, i \in A\rangle(X_i, i \in A).\end{aligned}$$

Moreover,

$$\langle F'_i, i \in A\rangle = \langle \langle F_j, j \in A_i\rangle(X_i)(\alpha_i), i \in A\rangle = \langle F_j, j \in A_i, i \in A\rangle(X_i, i \in A)(\alpha_i, i \in A).$$

So $\langle F'_i, i \in A\rangle$ is an algebraic extension of $\overline{\langle E_j, j \in A_i, i \in A\rangle(Y_i, i \in A)}$. In particular (see Lemma 2.7) there is an algebraic extension E'' of

$$\langle E_j, j \in A_i, i \in A\rangle(Y_i, i \in A)$$

(inside K) such that $\overline{E}'' = \langle F'_i, i \in A\rangle$. Let $\tilde{E}$ be the henselian closure of E'' in (K, v). We have $\overline{\tilde{E}} = \langle F'_i, i \in A\rangle$, $E'' \subseteq \tilde{E}$. By the claim, since $\tilde{E}$ is henselian and $\overline{\tilde{E}} \supseteq \langle F'_i, i \in A\rangle$, we have $a_i \in \tilde{E}$ for every $i \in A$. It implies $E'_A \subseteq \tilde{E}$, which gives, taking residues $\overline{E}'_A \subseteq \overline{\tilde{E}} = \langle F'_i, i \in A\rangle$. But by construction of the E'_i we obviously have $\overline{E}'_A \supseteq \langle F'_i, i \in A\rangle$. It follows that $\overline{E}'_A = \langle F'_i, i \in A\rangle$. □

Corollary 2.10. *Let $\mathcal{F} = (F_i)_{i \in I}$ be a finite projective system of fields of characteristic zero and let $\perp$ be the minimum of I. Assume that $(I, \leq)$ has a maximum $\top$ and let $(E_\perp, v_\perp)$ be a valued field such that $E_\perp v_\perp \cong F_\perp$. Then there is a projective system of valued fields $(E_i, v_i)_{i \in I}$ such that $(F_i)_{i \in I} \cong \operatorname{res}(E_i, v_i)_{i \in I}$ and, for every $i \in I$, $\operatorname{trdeg} E_i \mid E_\perp = \operatorname{trdeg} F_i \mid F_\perp$. Moreover:*

- *We can assume that all (E_i, v_i), $i \in I$, are henselian.*
- *If $v_\perp$ has rank one, then we can choose the valuations v_i, $i \in I$, such that they all have rank one.*

Proof. We first show that there is a projective system of fields $\mathcal{F}' = (F'_i)_{i \in I}$ with $\mathcal{F}' \cong \mathcal{F}$ and there is an extension (K, v) of $(E_\perp, v_\perp)$ such that $\overline{K} = F'_\top$, and such that v has rank one if $v_\perp$ has rank one. In particular $\overline{K} \supseteq F'_i$ for every $i \in I$.

Indeed, write $F_\top = F_\perp(X)(\bar{a})$, where X is a set of elements transcendental over $F_\perp$ and $\bar{a}$ is a sequence of elements algebraic over $F_\perp(X)$. Take Y a set of indeterminates with the same cardinality as X and consider the Gauss extension w of $v_\perp$ to $E_\perp(Y)$. Then $\overline{E_\perp(Y)} \cong F_\perp(X)$. Note that w has rank one if $v_\perp$ has rank one. Using now for instance [Endler 1963, Satz 1], we find an (algebraic) extension (K, v) of $(E_\perp(Y), w)$ such that $\overline{K}$ and $F_\top$ are isomorphic via a map which we denote by $h : \overline{K} \to F_\top$ (and with v of rank one if w has rank one). Define $F'_i := h^{-1}[F_i]$. This justifies the claim in the first paragraph of the proof.

To keep notation simple, we assume $\mathcal{F} = \mathcal{F}'$ as above. We construct the valued fields (E_i, v_i) (for $i \neq \perp$) as subfields of K endowed with the restriction of the valuation v. Since the valuation will always be v, we only look for the subfields E_i. Let $\perp$ be the minimum of I. We find the fields E_i by induction on $d(\perp, i)$ (note that $d(\perp, \top) = \max_{j\in I} d(\perp, j)$).

For $l \in \{0, \dots, d(\perp, \top)\}$, let $D_l = \{i \in I \mid d(\perp, i) = l\}$.

If $d(\perp, i) = 0$, then, by hypothesis and by the claim above, we already have the subvalued field $(E_\perp, v_\perp) \subseteq (K, v)$. Note that since $D_0 = \{\perp\}$ the sequence of fields $(F_i)_{i\in D_0}$ is a good residue of $(E_i)_{i\in D_0}$ in (K, v).

Assume we have found a system of fields $(E_i)_{i\in I, d(\perp,i)\leq l}$ such that $\operatorname{res}(E_i, v \restriction E_i) = F_i$ for $i \in I$ so that $d(\perp, i) \leq l$ and $(F_i)_{i\in D_l}$ is a good residue of $(E_i)_{i\in D_l}$ in (K, v). We write $D_{l+1} = \{i_k \mid k < m\}$, then we apply Lemma 2.9 with $(F'_k)_{k<m} = (F_{i_k})_{k<m}$, and obtain in this way a sequence $(E'_k)_{k<m}$. We define the fields E_i for $i \in D_{l+1}$ by $(E_{i_k})_{k<m} = (E'_k)_{k<m}$.

Finally, we can replace $(E_\top, v_\top)$ by one of it henselian closures, and each (E_i, v_i) by its henselian closure inside $(E_\top, v_\top)$. The new residue system is isomorphic to the previously defined residue system, which shows that we can assume that all (E_i, v_i) are henselian. □

3. Main results

Our main result, Corollary 3.3, is a direct consequence of the next two theorems, whose proofs are given in Sections 4 and 5 respectively.

Theorem 3.1. *Let $\mathcal{K} := (K_i, f_{ij})_{i\leq j\in I}$ be a projective system of fields (respectively formally real Pythagorean fields) such that $G(K_i)$ is finite for every $i \in I$. Let $(G_i, \lambda_{ij})_{i\leq j\in I} = G(\mathcal{K})$ and let G be the projective limit of this projective system of finite special groups. Then G is isomorphic to the special group of some field (respectively formally real Pythagorean field).*

Theorem 3.2. *Let $\mathcal{G} := (G_i, \lambda_{ij})_{i\leq j\in I}$ be an adequate projective system of finite reduced special groups* (see Definition 2.4). *Then there is a projective system $\mathcal{K}$ of formally real Pythagorean fields whose morphisms are inclusions, such that $\mathcal{G} \cong G(\mathcal{K})$.*

Now consider a profinite reduced special group G. Say it is the projective limit of the system $\mathcal{G} = (G_i, f_{ij})_{i\leq j\in I}$ of finite reduced special groups. Let i' be any element in I and consider the system $\mathcal{G}'$ equal to $\mathcal{G}$ restricted to indices in $I' := \{i \in I \mid i \leq i'\}$. The special group G is the projective limit of the system $\mathcal{G}'$, whose index set I' has a maximum element $\top = i'$. We can now use the strategy outlined after Definition 2.2 to express G as an adequate projective system whose index set is I'. Applying Theorem 3.2 then Theorem 3.1 now yields:

Corollary 3.3. *Every profinite reduced special group is isomorphic to the special group of some formally real Pythagorean field.*

4. Proof of Theorem 3.1

If $(I, \leq)$ is a downward directed poset and $i \in I$, then $i^{\leftarrow}$ denotes $\{j \in I \mid j \leq i\}$ and $i^{\rightarrow}$ denotes $\{j \in I \mid j \geq i\}$.

We first assume the following reductions:

(1) I has a maximum $\top$ ($I = \top^{\leftarrow}$).

(2) All the K_i, $i \in I$, are subfields of the field $M := K_\top$, and the morphisms $f_{ij} : K_i \to K_j$ are inclusions. In particular, the projective limit of the system $\mathcal{K}$ is isomorphic to the intersection of the fields K_i, $i \in I$.

These assumptions can safely be made because for the original projective system of fields $\mathcal{K} := (K_j, f_{jk})_{j \leq k \in I}$ and for each $i' \in I$ fixed,

(i) the set $i'^{\leftarrow}$ is a coinitial subset of I, and

(ii) if $j \leq i' \in I$, we can identify K_j with the subfield $K'_j := f_{ji'}[K_j]$ of $K_{i'}$, and the morphisms $f_{jk} : K_j \to K_k$ are naturally identified with inclusions $\iota_{jk} : K'_j \hookrightarrow K'_k$.

The reductions above give us

$$\varprojlim (K_j, f_{jk})_{j \leq k \in I} \cong \varprojlim (K_j, f_{jk})_{j \leq k \in i'^{\leftarrow}} \cong \varprojlim (K'_j, \iota_{jk})_{j \leq k \in i'^{\leftarrow}} \\ \cong \bigcap_{j \in i'^{\leftarrow}} f_{ji'}[K_j] \subseteq K_{i'}$$

and

$$G := \varprojlim (G(K_j), G(f_{jk}))_{j \leq k \in I} \cong \varprojlim (G(K_j), G(f_{jk}))_{j \leq k \in i'^{\leftarrow}} \\ \cong \varprojlim (G(K'_j), G(\iota_{jk}))_{j \leq k \in i'^{\leftarrow}}.$$

Now consider the language $L = L_R \cup \{R_i\}_{i \in I}$, where L_R is the language of rings and the R_i are unary relation symbols. We turn M into an L-structure by interpreting each R_i in M by the subfield K_i.

Let N be an $|I|^+$-saturated elementary extension of M in the language L. (See [Chang and Keisler 1990, Chapter 5 and Lemma 5.1.2] or [Hodges 1993, p. 480 and Corollary 10.2.2] for the definition of saturated models and the existence result we just used. Note that this notion of saturation is not linked to the existing one for subgroups of special groups.) Each $F_i := R_i^N$ is a subfield of N, and the fields F_i form a projective system of fields $\mathcal{F}$ (since for $k \leq i, j \in I$ the sentence "$R_k \subseteq R_i \cap R_j$" is in the theory of M). Moreover, for every $i \in I$, $K_i \hookrightarrow F_i$ is an L_R-elementary embedding and therefore induces an isomorphism of special groups

$$G(K_i) \xrightarrow{\cong} G(F_i)$$

(since the special groups $G(K_i)$, being finite, are described in the theory of M). More generally $G(\mathcal{K}) \cong G(\mathcal{F})$, so $G \cong \varprojlim G(\mathcal{K}) \cong \varprojlim G(\mathcal{F})$.

Let $F := \bigcap_{i\in I} F_i$ and define

$$\begin{aligned} \xi :\ G(F) &\to \varprojlim G(\mathcal{F}) \\ a \cdot \dot{F}^2 &\mapsto (a \cdot \dot{F_i}^2)_{i\in I}. \end{aligned}$$

We show that ξ is an isomorphism of special groups, which yields $G \cong G(F)$ as needed (in particular, if the fields K_i, $i \in I$, are formally real Pythagorean, then $F = \bigcap_{i\in I} F_i$ is formally real Pythagorean, since $G(F)$ is a reduced special group).

Step 1. It is clear that ξ is well-defined and is a morphism of groups.

Step 2. ξ is a morphism of special groups. Indeed, it is clear that ξ sends -1 to -1. Let $a \cdot \dot{F}^2, b \cdot \dot{F}^2 \in G(F)$ be such that $a \cdot \dot{F}^2 \in D_{G(F)}\langle 1, b \cdot \dot{F}^2\rangle$. There are then $c, d \in F$ such that, for all $i \in I$, $a = c^2 + bd^2$ in F_i. Then $a \cdot \dot{F_i}^2 \in D_{G(F_i)}\langle 1, b \cdot \dot{F_i}^2\rangle$ for every $i \in I$, and therefore $\xi(a \cdot \dot{F}^2) \in D_{G'}\langle 1, \xi(b \cdot \dot{F}^2)\rangle$.

Step 3. ξ is surjective: Let $a = (a_i \cdot \dot{F}_i^2)_{i\in I} \in \varprojlim G(F_i)$. So for all $i \leq j \in I$, $a_i \cdot \dot{F}_j^2 = a_j \cdot \dot{F}_j^2$. We want $x \in N$ satisfying the set of formulas

$$\Delta := \{x \in F_i\}_{i\in I} \cup \{x = a_i \bmod \dot{F}_i^2\}_{i\in I}.$$

Every finite part of Δ is satisfied in N since $a = (a_i \cdot \dot{F}_i^2)_{i\in I} \in \varprojlim G(F_i)$ (it suffices to take $x = a_k$, where k is less than every one of the indices $i \in I$ occurring in this finite part). By $|I|^+$-saturation, Δ has a solution x in N. Then $\xi(x) = (a_i \cdot \dot{F}_i^2)$.

The rest of the proof relies on the following lemma.

Lemma 4.1. *Let $n \in \mathbb{N}$ and let $P(X_1, \ldots, X_n) \in F[X_1, \ldots, X_n]$. Assume the equation $P(X_1, \ldots, X_n) = 0$ has a solution in every F_i, $i \in I$. Then the same equation has a solution in F.*

Proof. We are looking for $\bar{x} \in N$ such that the set of formulas

$$\Sigma := \{P(\bar{x}) = 0\} \cup \{\bar{x} \in F_i\}_{i\in I}$$

is satisfied in N. Since the F_i, together with the inclusions between them, form a projective system, every finite part of Σ has a solution, and by the $|I|^+$-saturation of N, Σ has a solution in N. $\square$

We go back to proving that ξ is an isomorphism:

Step 4. ξ is injective: Let $a = a \cdot \dot{F}^2 \in G(F)$ be such that $\xi(a) = 1$, i.e., $a \in \dot{F_i}^2$ for every $i \in I$, i.e., the polynomial $X^2 - a$ has a root in each F_i, $i \in I$. By Lemma 4.1, $X^2 - a$ has a root in F, hence $a \in \dot{F}^2$.

Step 5. ξ is a monomorphism of special groups: Let $a, b \in F$ be such that, for every $i \in I$, $a \cdot \dot{F}_i^2 \in D\langle 1, b \cdot \dot{F}_i^2 \rangle$. Let $P(X, Y) = a - (X^2 + bY^2) \in F[X, Y]$. By hypothesis, $P(X, Y) = 0$ has a solution in each F_i, hence a solution in F by Lemma 4.1, which means $a \in D_{G(F)}\langle 1, b\rangle$.

5. Proof of Theorem 3.2

Reducing to a finite projective system. Since $\mathcal{G}$ is adequate, the set $(I, \leq)$ has a maximum element $\top$.

In this subsection we show that it is enough to prove Theorem 3.2 when $\mathcal{G} := (G_i, f_{ij})_{i \leq j \in I}$ is a finite projective system of special groups such that I has a maximum (which we will also denote by $\top$).

Let L be the language $\{0, 1, -, +, \cdot\} \cup \{F_i \mid i \in I\} \cup \{Q_i^g \mid i \in I, g \in G_i\}$, where $0, 1$ are constant symbols, $-$ is a unary function symbol, $+, \cdot$ are binary function symbols and F_i, Q_i^g are unary predicate symbols, for each $i \in I$ and $g \in G_i$. Denote by λ_i the *inverse* of the bijection $g \in G_i \mapsto Q_i^g$, $i \in I$. The projective system of fields we are looking for is a model of the theory Ω consisting of (first-order) L-sentences that are informally described in the four items below:

(1) the interpretation of the unary predicate $F_\top$ is the universe of the L-structure (i.e., $\forall x(F_\top(x))$) and "$(F_\top, 0, 1, +, \cdot)$ is a field";

(2) for every $i \leq j \in I$:
"$F_i \subseteq F_j$" and "$(F_i, 0, 1, +, \cdot)$ is a subfield of the field $(F_\top, 0, 1, +, \cdot)$" (technically speaking, $+$ and $\cdot$ are functional symbols globally defined whose restrictions to F_i give internal operations on F_i);

(3) for every $i \in I$:
"λ_i is an isomorphism of special groups $G(F_i) \to G_i$";

(4) for every $i \leq j \in I$:
"the morphism of special groups induced by the inclusion $F_i \subseteq F_j$ is naturally identified with f_{ij}, via the isomorphisms λ_i, λ_j".

It is clear how to describe the expressions in items (1) and (2) by first-order L-sentences. For the reader's convenience, we add a more explicit description of the L-sentences involved in the two remaining items: the hypothesis that the special groups G_i are all finite ensures that the prescription in item (3) can be encoded by a set of first-order L-sentences.

Item (3): for each $i \in I$:

- for each $g \in G_i$, "$Q_i^g \subseteq \dot{F}_i$" and "$Q_i^g = a\dot{F}_i^2$, for some $a \in \dot{F}_i$";
- for each $g, g' \in G_i$ such that $g \neq g'$, "$Q_i^g \cap Q_i^{g'} = \varnothing$";
- "$\dot{F}_i = \bigcup\{Q_i^g \mid g \in G_i\}$" (as G_i is a finite special group, this can be described by a first-order L-sentence);
- "$1 \in Q_i^1$ and $-1 \in Q_i^{-1}$";
- for each $g, g' \in G_i$, "for each a, a', if $a \in Q_i^g$ and $a' \in Q_i^{g'}$ then $a \cdot a' \in Q_i^{gg'}$";
- for each $g, g' \in G_i$ such that $g' \in D_{G_i}(1, g)$, "for each a, a', if $a \in Q_i^g$ and $a' \in Q_i^{g'}$ then there are $x, y \in F_i$ such that $a' = x^2 + ay^2$";
- for each $g, g' \in G_i$ such that $g' \notin D_{G_i}(1, g)$, "for each a, a', if $a \in Q_i^g$ and $a' \in Q_i^{g'}$ then for all $x, y \in F_i$, $a' \neq x^2 + ay^2$".

Item (4): for each $i \leq j \in I$:

By the axioms above: since $F_i \subseteq F_j$ and we have the partitions

$$\dot{F}_i/\dot{F}_i^2 = \{Q_i^g \mid g \in G_i\} \quad \text{and} \quad \dot{F}_j/\dot{F}_j^2 = \{Q_j^{g'} \mid g' \in G_j\},$$

then for each $g \in G_i$ there is a unique $g' \in G_j$ such that $Q_i^g \subseteq Q_j^{g'}$. In this way we obtain a function $q_{ij} : \dot{F}_i/\dot{F}_i^2 \to \dot{F}_j/\dot{F}_j^2$. Clearly $q_{ij}(a.\dot{F}_i^2) = a.\dot{F}_j^2$, for every $a \in \dot{F}_i$, i.e., q_{ij} is the special group morphism induced by the inclusion $F_i \subseteq F_j$. We add a new list of axioms expressing that $\lambda_j \circ q_{ij} = f_{ij} \circ \lambda_i$. A direct examination of the equivalent condition $q_{ij} = \lambda_j^{-1} \circ f_{ij} \circ \lambda_i$ shows that these axioms must be

$$\text{for each } g \in G_i, \text{ "}Q_i^g \subseteq Q_j^{f_{ij}(g)}\text{".}$$

Using now the compactness theorem (see [Chang and Keisler 1990, Theorem 1.3.22] or [Hodges 1993, Theorem 6.1.1]), to find a model of this theory we only need to find a model of every finite part $\Omega_0 \subseteq \Omega$. Let J be the set of elements of I occurring in this finite part Ω_0, together with $\top$. Since I is downward directed, we can assume that J is also downward directed (taking a larger set J if necessary), that is J has a first element $\perp$. In particular J determines a finite projective system of special groups whose index set that has a maximum and a minimum.

Description of the proof by induction. We therefore assume from now on that the index poset $(I, \leq)$ is finite and that it has a minimum $\perp$ and a maximum $\top$. We find a finite projective system $\mathcal{H}$ of Pythagorean fields of characteristic 0 such that $\mathcal{G} \cong G(\mathcal{H})$ by induction on the construction of $G_\perp$ by products and extensions. For the purpose of the proof, we allow the (nonreduced) special group $\{1\}$ to appear in $\mathcal{G}$.

Recall that since $\mathcal{G}$ is an adequate projective system, the morphisms f_{ij}, $i \leq j \in I$, are quotients by saturated subgroups (see the paragraph after Definition 2.2).

If $G_\perp \cong \{1\}$, then all special groups in the system are trivial and all morphisms are isomorphisms. We can obviously realize such a system by taking $F_i = A$, $i \in I$, where A is any fixed algebraically closed field of characteristic 0.

If $G_\perp \cong \mathbb{Z}_2$, then all special groups in the system are isomorphic to $\mathbb{Z}_2$ or to $\{1\}$ and all morphisms are isomorphisms or naturally identified with $\mathbb{Z}_2 \twoheadrightarrow \{1\}$. We can obviously realize such a system by simply selecting a real closed field R and an algebraically closed field A such that $R \subseteq A$.

If $G_\perp \cong G'_\perp \times G''_\perp$, since all morphisms and special groups in the systems are quotients of $G_\perp$ by (larger and larger) saturated subgroups, and using Fact 2.5, the whole projective system $(G_i, f_{ij})_{i \leq j \in I}$ splits according to the product $G_\perp \cong G'_\perp \times G''_\perp$ into two adequate projective systems of finite special groups:

$$(G'_i, f'_{ij})_{i \leq j \in I} \quad \text{and} \quad (G''_i, f''_{ij})_{i \leq j \in I}. \tag{5-1}$$

(Note that, for each $i \in I$, if G_i is reduced, then either both G'_i and G''_i are reduced or one of them is the trivial special group $\{1\}$ and the other is reduced.) By induction the systems in (5-1) are realized by two projective systems of Pythagorean fields of characteristic 0: $\mathcal{F}' = (F'_i)_{i \in I}$ and $\mathcal{F}'' = (F''_i)_{i \in I}$ (where the morphisms are inclusions), so we just need to "glue" them together. For this we use results from [Kula 1979], which describe how to realize a finite product of finite special groups when each one is already realized. This is achieved in the next subsection.

If $G_\perp \cong G'[H]$, as above, the morphisms of special groups in the projective system are quotients of $G_\perp$ by (larger and larger) saturated subgroups Δ_i. This case is dealt with starting on page 279, using results from [Becher 2002].

Gluing, the product case. The next several pages are taken by the proof of the following result.

Theorem 5.1. *Let $(I, \leq)$ be a finite downward directed index set with first element $\perp$ and last element $\top$. Let $\mathcal{F}' = (F'_i)_{i \in I}$, $\mathcal{F}'' = (F''_i)_{i \in I}$ be finite projective systems of fields of characteristic* 0, *where the morphisms are inclusions and such that for every $i \in I$ $G(F'_i)$ and $G(F''_i)$ are finite special groups. Then there is a finite projective system $\mathcal{F} = (F_i)_{i \in I}$ of fields of characteristic* 0 (*where the morphisms are inclusions*) *such that*

$$G(\mathcal{F}) \cong G(\mathcal{F}') \times G(\mathcal{F}'').$$

Remark 5.2. In this theorem, for each $i \in I$ we have:

(a) F_i is Pythagorean if and only if F'_i *and* F''_i are Pythagorean.

(b) If F_i is Pythagorean, then F_i is formally real if and only if F'_i *or* F''_i is formally real.

We begin with a reformulation of some results from [Kula 1979].

Definition 5.3. Let F be a field equipped with n mutually independent valuations of rank one $v_1, \dots, v_n$, and let f_i be an embedding of F into F^{v_i}, a completion of F with respect to v_i. We say that $(F, f_1, \dots, f_n)$ fulfills the global squares property if, for every $a \in F$,

$$a \in \dot{F}^2 \iff \forall i \in \{1, \dots, n\}\ f_i(a) \in (\dot{F^{v_i}})^2.$$

(Note that the left to right implication always holds.)

Theorem 5.4 [Kula 1979, Corollary 2.5]. *With notation as in Definition 5.3, assume that $(F, f_1, \dots, f_n)$ fulfills the global squares property. Then the map*

$$\begin{aligned} \xi_F : G(F) &\to \textstyle\prod_{i=1}^n G(F^{v_i}) \\ a\dot{F}^2 &\mapsto (f_i(a) \cdot (\dot{F^{v_i}})^2)_{i=1,\dots,n} \end{aligned}$$

is an isomorphism of special groups.

Theorem 5.5 [Kula 1979, Theorem 2.6]. *Let $(L_i, v_{i1}, \dots, v_{in})_{i \in I}$ be a finite projective system of fields equipped with n mutually independent valuations of rank one, and such that I has a maximum element $\top$. Then for every $i \in I$ there is an algebraic extension $\mathscr{E}(L_i)$ of L_i and a morphism of special groups*

$$\eta_i : G(\mathscr{E}(L_i)) \to \prod_{k=1}^n G((L_i)^{v_{ik}})$$

such that

1. *$\mathscr{E}(L_i) \subseteq \mathscr{E}(L_j)$ for every $j \in I$, $j \geq i$;*
2. *$G(\mathscr{E}(L_i)) \underset{\eta_i}{\cong} \prod_{k=1}^n G((L_i)^{v_{ik}})$; and*
3. *the morphism of special groups $\prod_{k=1}^n G((L_i)^{v_{ik}}) \to \prod_{k=1}^n G((L_j)^{v_{jk}})$, given by the product of the morphisms of special groups induced by $(L_i)^{v_{ik}} \subseteq (L_j)^{v_{jk}}$ is naturally identified, via the isomorphisms η_i and η_j, with the morphism of special groups $G(\mathscr{E}(L_i)) \to G(\mathscr{E}(L_j))$ induced by $\mathscr{E}(L_i) \subseteq \mathscr{E}(L_j)$.*

Proof. Since a valuation v_{ik} is the restriction on L_i of the valuation $v_{\top k}$, we drop the first index and simply denote it by v_k. For $k \in \{1, \dots, n\}$ we fix a completion $L_\top^k$ of $L_\top$ with respect to v_k and define, for $i \in I$, L_i^k to be the completion of L_i in $L_\top^k$ with respect to v_k. The systems $(L_i^k)_{i \in I}$, for $k \in \{1, \dots, n\}$, are all projective systems of fields, where the morphisms are the inclusions (since L_i^k is simply the set of limits in $L_\top^k$ of v_k-Cauchy sequences of elements of L_i).

Let K_+ be an algebraic closure of $L_\top$. We define the set

$$\begin{aligned} \mathscr{L} := \bigl\{ &\text{projective systems of fields } (E_i, \iota_{i1}, \dots, \iota_{in})_{i \in I} \\ &\text{such that } L_i \subseteq E_i \subseteq K_+ \text{ with } E_i \mid L_i \text{ algebraic,} \\ &\text{equipped with the } L_i\text{-embeddings of fields } \iota_{ik} : E_i \to L_i^k \text{ for } k = 1, \dots, n \bigr\}. \end{aligned}$$

(Note that the condition that $(E_i, \iota_{i1}, \dots, \iota_{in})_{i\in I}$ is a projective system implies $\iota_{ik} \subseteq \iota_{jk}$ for $i \le j \in I$ and $k \in \{1, \dots, n\}$, which is possible since $L_i^k \subseteq L_j^k$.) We equip $\mathscr{L}$ with the partial ordering

$$(E_i, \iota_{i1}, \dots, \iota_{in})_{i\in I} \le (F_i, \kappa_{i1}, \dots, \kappa_{in})_{i\in I}$$
$$\text{if and only if}$$
$$\text{for every } i \in I \text{ and } k \in \{1, \dots, n\},\ E_i \subseteq F_i \text{ and } \iota_{ik} \subseteq \kappa_{ik}.$$

By Zorn's lemma, $\mathscr{L}$ has a maximal element $(M_i, f_{i1}, \dots, f_{in})_{i\in I}$. We show that, for $j \in I$, $(M_j, f_{j1}, \dots, f_{jn})$ fulfills the global squares property. Let $j \in I$ and let $a \in M_j \setminus \{0\}$ be such that $f_{jk}(a) \in (L_j^k)^{\times 2}$, for $k = 1, \dots, n$. Assume $\sqrt{a} \notin M_j$. Fix a square root $\sqrt{a}$ of a and $\alpha_k \in L_j^k$ such that $\alpha_k^2 = f_{jk}(a)$. Then each morphism f_{jk} can be (properly) extended to $M_j' := M_j(\sqrt{a})$ by sending $\sqrt{a}$ to α_k. Moreover, with $A_j := \{r \in I \mid r \ge j\}$, and since for $r \in A_j$ we have $L_j^k \subseteq L_r^k$, the same reasoning tells us that, for each $r \in A_j$ and $k \in \{1, \dots, n\}$, each morphism f_{rk} can be extended to $M_r' := M_r(\sqrt{a})$ by sending $\sqrt{a}$ to α_k (since $\alpha_k \in L_r^k$). If $r \in I \setminus A_j$, we take $M_r' := M_r$. We obtain in this way $(M_i', f_i', \dots, f_n')_{i\in I}$, a projective system of fields equipped with n morphisms of fields that is (strictly) larger than $(M_i, f_i, \dots, f_n)_{i\in I}$, a contradiction. It follows that $\sqrt{a} \in M_j$ and thus that $(M_j, f_{j1}, \dots, f_{jn})$, for $j \in I$, fulfills the global squares property. If we take $\mathscr{E}(L_i) = M_i$ for $i \in I$, the first conclusion of the theorem then holds, and the second follows by Theorem 5.4, with

$$\begin{aligned} \eta_i : G(M_i) &\overset{\cong}{\to} \prod_{k=1}^{n} G((L_i)^k) \\ a \cdot \dot{M}_i^2 &\mapsto (f_i(a) \cdot (L_i^k)^{\times 2})_{i=1,\dots,n}, \end{aligned}$$

for $i \in I$. The third conclusion is proved in the next lemma. □

Lemma 5.6. *Let the notation be as in Theorem 5.5 and its proof.*

Let $(L, v_1, \dots, v_n) \supseteq (K, v_1 \restriction K, \dots, v_n \restriction K)$ be two fields equipped with n mutually independent valuations of rank one. For $m = 1, \dots, n$ let

- *L^m be a completion of L with respect to v_m and K^m be a completion of K with respect to $v_m \restriction K$ such that $K^m \subseteq L^m$,*
- *f_m be an embedding of K into K^m and g_m be an embedding of L into L^m extending f_m.*

Assume $(K, f_1, \dots, f_n)$ and $(L, g_1, \dots, g_n)$ satisfy the global squares property.

Let $\lambda : \prod_{m=1}^n G(K^m) \to \prod_{m=1}^n G(L^m)$ be the product of the morphisms of special groups induced by the inclusions $K^m \subseteq L^m$ for $m = 1, \dots, n$, and let $\mu : G(K) \to G(L)$ be the morphism of special groups induced by $K \subseteq L$.

Then λ and μ are naturally identified via the isomorphisms ξ_K and ξ_L given by Theorem 5.4.

Proof. By Theorem 5.4 the isomorphism $G(K) \cong G(K^1) \times \cdots \times G(K^n)$ is

$$\begin{aligned} \xi_K : G(K) &\to G(K^1) \times \cdots \times G(K^n) \\ x \cdot \dot{K}^2 &\mapsto (f_1(x) \cdot (\dot{K^1})^2, \ldots, f_n(x) \cdot (\dot{K^n})^2). \end{aligned}$$

Similarly, the isomorphism between $G(L)$ and $G(L^1) \times \cdots \times G(L^n)$ is

$$\begin{aligned} \xi_L : G(L) &\to G(L^1) \times \cdots \times G(L^n) \\ x \cdot \dot{L}^2 &\mapsto (g_1(x) \cdot (\dot{L^1})^2, \ldots, g_n(x) \cdot (\dot{L^n})^2). \end{aligned}$$

Thus $\lambda = \xi_L \circ \mu \circ \xi_K^{-1}$ since $g_m \restriction K = f_m$ for $m = 1, \ldots, n$. $\square$

We now turn our attention to the two finite projective systems of fields $\mathcal{F}' = (F_i')_{i \in I}$ and $\mathcal{F}'' = (F_i'')_{i \in I}$ of characteristic zero introduced in the statement of Theorem 5.1. We first show that we can assume that the fields in $\mathcal{F}'$ and $\mathcal{F}''$ are at most countable and of finite transcendence degree over $\mathbb{Q}$. This is achieved by the following proposition.

Proposition 5.7 [Kula 1979, Proposition 3.1]. *Let $\mathcal{L} := (L_i)_{i \in I}$ be a finite projective system of fields of characteristic* 0 *such that $G(L_i)$ is a finite special group for all $i \in I$. There is a map $\mathcal{F}$, defined on $\{L_i\}_{i \in I}$, satisfying the following properties whenever $i \leq j \in I$:*

(1) *$\mathcal{F}(L_i)$ is a countable subfield of L_i with finite transcendence degree over $\mathbb{Q}$.*
(2) *If $\varphi_i : \mathcal{F}(L_i) \hookrightarrow L_i$ is the inclusion map, then $G(\varphi_i) : G(\mathcal{F}(L_i)) \to G(L_i)$ is an isomorphism of special groups.*
(3) *$\mathcal{F}(L_i) \subseteq \mathcal{F}(L_j)$.*
(4) *If $\lambda_{ij} : G(L_i) \to G(L_j)$ is the morphism of special groups induced by $L_i \subseteq L_j$, then the morphism of special groups $G(\mathcal{F}(L_i)) \to G(\mathcal{F}(L_j))$ induced by $\mathcal{F}(L_i) \subseteq \mathcal{F}(L_j)$ is naturally identified with λ_{ij}, via the isomorphisms $G(\varphi_i)$ and $G(\varphi_j)$.*

Proof. The proof is a trivial extension of Kula's. If L is a field with a finite number of square classes, a representative system of $G(L)$ is a finite subset $R(L) = A \cup B$ of L such that

- $A \subseteq \dot{L}$ and $\dot{L}/\dot{L}^2 = A/\dot{L}^2$;
- For every $a_1, a_2 \in A$ with $a_1 \in D_L\langle 1, a_2\rangle$, there are $b_1, b_2 \in B$ such that $a_1 = b_1^2 + a_2 b_2^2$.

Claim: For every $i \in I$ there is a representative system $R(L_i)$ of L_i such that $R(L_i) \subseteq R(L_j)$ whenever $i \leq j$.

Proof of the claim: Direct by induction on $d(\perp, i)$ (just take a system of representatives of L_i and add to it all the $R(L_j)$ for $\perp \leq j < i$).

Then, just as in [Kula 1979], take for $\mathcal{F}(L_i)$ the algebraic closure of $\mathbb{Q}(R(L_i))$ in L_i. $\square$

The following two propositions show that we can assume that $\operatorname{atd}(F_i') = \operatorname{atd}(F_i'')$ for every $i \in I$, where atd denotes the absolute transcendence degree, i.e., the transcendence degree over $\mathbb{Q}$.

Proposition 5.8 ([Kula 1979], Lemma 3.2). *Let $(L_i)_{i\in I}$ be a finite projective system of countable fields of finite absolute transcendence degree. There is a map $\mathcal{T}$ defined on $\{L_i\}_{i\in I}$ satisfying the following properties whenever $i \le j \in I$:*

(1) *$\mathcal{T}(L_i)$ is a countable field extension of L_i.*

(2) $\operatorname{atd}(\mathcal{T}(L_i)) = \operatorname{atd}(L_i) + 1$.

(3) *If $\tau_i : L_i \hookrightarrow \mathcal{T}(L_i)$ is the inclusion map, then $G(\tau_i) : G(L_i) \to G(\mathcal{T}(L_i))$ is an isomorphism of special groups.*

(4) $\mathcal{T}(L_i) \subseteq \mathcal{T}(L_j)$.

(5) *If $\lambda_{ij} : G(L_i) \to G(L_j)$ is the morphism of special groups induced by $L_i \subseteq L_j$, then the morphism of special groups $G(\mathcal{T}(L_i)) \to G(\mathcal{T}(L_j))$ induced by $\mathcal{T}(L_i) \subseteq \mathcal{T}(L_j)$ is naturally identified with λ_{ij} (via $G(\tau_i)$ and $G(\tau_j)$).*

Proof. For $i \in I$ let $K_i := L_i(x)(\sqrt[2^n]{x})_{n\in\mathbb{N}}$ (x is an indeterminate), and consider on K_i the unique extension v_i of the valuation on $L_i(x)$ determined by the irreducible polynomial x. The K_i, together with their inclusions, form a projective system, and the sets $\Phi_i := \{v_i\}$ satisfy the hypothesis of Theorem 5.5. We now apply the map $\mathcal{E}$ defined in Theorem 5.5 to the projective system of the K_i and get the projective system of the $\mathcal{T}(L_i)$. Since L_i is countable, K_i and $\mathcal{T}(L_i) = \mathcal{E}(K_i)$ are countable. Kula's proof of [Kula 1979, lemma 3.2] shows that the second and third claims of the proposition hold, and the last two hold by Theorem 5.5. $\square$

Proposition 5.9. *There exist finite projective systems $\mathcal{K}' = (K_i')_{i\in I}$ and $\mathcal{K}'' = (K_i'')_{i\in I}$ of fields of characteristic 0 such that*

(1) *$G(\mathcal{K}') \cong G(\mathcal{F}')$ and $G(\mathcal{K}'') \cong G(\mathcal{F}'')$, and*

(2) *for every $i \in I$, $\operatorname{atd}(K_i') = \operatorname{atd}(K_i'') < \infty$.*

Proof. We assume there is some $i \in I$ such that $\operatorname{atd}(F_i') \ne \operatorname{atd}(F_i'')$ and we proceed by induction on $d(\perp, i)$, the maximal length of a chain from $\perp$ to i.

- $d(\perp, i) = 0$, i.e., $i = \perp$. Let $t := \max\{\operatorname{atd}(F_\perp'), \operatorname{atd}(F_\perp'')\}$. We then apply Proposition 5.8 as many times as necessary to the system $\mathcal{F}'$ or $\mathcal{F}''$ (the one that does no realize the maximum), and we obtain two new systems $\mathcal{F}_{(0)}'$ and $\mathcal{F}_{(0)}''$ indexed by I, whose fields of index $\perp$ have same (finite) absolute transcendence degree t.

- $d(\perp, i) = n > 0$. We now proceed by induction on the number of i's with $d(\perp, i) = n$ and $\operatorname{atd}(F'_i) \neq \operatorname{atd}(F''_i)$. We fix one of them: i_1. By induction we can assume that the projective systems $\mathcal{F}'$ and $\mathcal{F}''$ satisfy $\operatorname{atd}(F'_j) = \operatorname{atd}(F''_j)$ for every $j \in I$, $d(\perp, j) < n$. We consider the systems $\mathcal{F}' \restriction i_1^{\to}$ and $\mathcal{F}'' \restriction i_1^{\to}$. By applying Proposition 5.8, we get two new systems $\mathcal{T}'$ and $\mathcal{T}''$, indexed by $i_1^{\to}$ whose fields indexed by i_1 have same absolute transcendence degree. We replace, in $\mathcal{F}'$, respectively $\mathcal{F}''$, the subsystem $\mathcal{F}' \restriction i_1^{\to}$ by $\mathcal{T}'$, respectively $\mathcal{F}'' \restriction i_1^{\to}$ by $\mathcal{T}''$ and we write $\mathcal{F}'_{(1)}$, $\mathcal{F}''_{(1)}$ for the new sets of fields. Since every field has been replaced by a field extension, we still get projective systems of fields and, moreover, $G(\mathcal{F}'_{(1)}) \cong G(\mathcal{F}')$ and $G(\mathcal{F}''_{(1)}) \cong G(\mathcal{F}'')$. Now $\operatorname{atd}(F'_{(1)i_1}) = \operatorname{atd}(F''_{(1)i_1}) < \infty$, and we proceed by induction. □

So, from now on, we assume that our two finite projective systems of fields $\mathcal{F}'$ and $\mathcal{F}''$ consist of countable fields having the same finite transcendence degree over $\mathbb{Q}$ at each index.

Remark 5.10. Let K be a field equipped with two independent valuations v_1 and v_2 and let (L, w_1, w_2) be an extension of (K, v_1, v_2). Then w_1 and w_2 are independent. Indeed, if it were not the case, then w_1 and w_2 would define the same topology on L (see [Engler and Prestel 2005, Theorem 2.3.4]), and therefore the same induced topologies on K, which coincide with the topologies defined by v_1 and v_2. It shows that v_1 and v_2 define the same topology on K, a contradiction since they are independent (again by the theorem just cited).

Lemma 5.11. *There are two henselian valued fields* $(E'_\perp, v')$ *and* $(E''_\perp, v'')$ *both containing* $\mathbb{Q}(X)$, *such that*

(1) v' *and* v'' *are of rank one*,

(2) $E'_\perp v' \cong F'_\perp$ *and* $E''_\perp v'' \cong F''_\perp$,

(3) $\operatorname{atd} E'_\perp = \operatorname{atd} F'_\perp + 1 = \operatorname{atd} F''_\perp + 1 = \operatorname{atd} E''_\perp$,

(4) $v' E'_\perp$ *and* $v'' E''_\perp$ *are divisible*, *and*

(5) *the restrictions of* v' *and* v'' *to* $\mathbb{Q}(X)$ *are independent.*

(In (4), two-divisible is actually enough for our purposes.)

Proof. Let $\{y_1, \dots, y_k\}$ be a finite transcendence basis of $F'_\perp$ over $\mathbb{Q}$, and let E be $\mathbb{Q}(y_1, \dots, y_k)(X)$, equipped with the valuation v determined by the irreducible polynomial $X \in \mathbb{Q}(y_1, \dots, y_k)[X]$. Then $\bar{E} \cong \mathbb{Q}(y_1, \dots, y_k)$, $vE = \mathbb{Z}$ and $F'_\perp$ is isomorphic to an algebraic extension of $\bar{E}$. By [Endler 1963, Satz 1], there is an algebraic extension $E'_\perp$ of E and an extension v' of v to $E'_\perp$ such that $E'_\perp = F'_\perp$ and $v' E'_\perp$ is divisible of rank one.

To construct $(E''_\perp, v'')$, we proceed as above but start with the valuation on $\mathbb{Q}(y_1, \dots, y_k)(X)$ associated to the irreducible polynomial $X - 1$. Obviously, $v'' \restriction \mathbb{Q}(X)$ and $v' \restriction \mathbb{Q}(X)$ are independent over $\mathbb{Q}(X)$. $\square$

We now apply Corollary 2.10 twice (with the valued fields $(E'_\perp, v')$ and $(E''_\perp, v'')$ given by Lemma 5.11), and get two projective systems of henselian valued fields $\mathscr{E}' = (E'_i, v'_i)_{i\in I}$ and $\mathscr{E}'' = (E''_i, v''_i)_{i\in I}$ equipped with valuations of rank one, such that $\operatorname{res}(\mathscr{E}') \cong \mathscr{F}'$ and $\operatorname{res}(\mathscr{E}'') \cong \mathscr{F}''$. Up to renaming the transcendental elements, we can assume that for every $i \in I$ there is a finite set X_i of transcendental elements over $\mathbb{Q}$ and an algebraic closure Q_i of $\mathbb{Q}(X_i)$ such that $E'_i, E''_i \subseteq Q_i$, and such that, for every $i \leq j \in I$ $X_i \subseteq X_j$ and $Q_i \subseteq Q_j$.

Since, for $i \in I$, E'_i and E''_i are both subfields of Q_i, we can consider the projective system of valued fields $(E'_i \cap E''_i, v'_i, v''_i)_{i\in I}$. Note that v'_i and v''_i are independent by Remark 5.10 and Lemma 5.11(5). We recall now the following special case of a result from [Heinemann 1985]:

Theorem 5.12. *Let K be a field equipped with two independent valuations v_1 and v_2. Fix an algebraic closure $\tilde{K}$ of K. Let (H_i, v_i), for $i = 1, 2$, be henselian extensions of (K, v_i) such that $H_1, H_2 \subseteq \tilde{K}$ and $K = H_1 \cap H_2$.*

Then (H_i, v_i) is a henselization of (K, v_i), for $i = 1, 2$.

Applying this result, we obtain that, for every $i \in I$, (E'_i, v'_i) is a henselization of $(E'_i \cap E''_i, v'_i)$ and (E''_i, v''_i) is a henselization of $(E'_i \cap E''_i, v''_i)$. In particular:

(1) $v'(E'_i \cap E''_i)$ and $v'(E'_i \cap E''_i)$ are two-divisible;

(2) $\operatorname{res}(E'_i \cap E''_i, v'_i)_{i\in I} \cong \mathscr{F}'$ and $\operatorname{res}(E'_i \cap E''_i, v''_i)_{i\in I} \cong \mathscr{F}''$;

(3) v' and v'' are independent on $E'_i \cap E''_i$ (by Lemma 5.11.(5) and Remark 5.10).

We now apply Theorem 5.5 to the system $(E'_i \cap E''_i, v'_i, v''_i)_{i\in I}$ and get the system $(\mathscr{E}(E'_i \cap E''_i))_{i\in I}$, which satisfies

$$G((\mathscr{E}(E'_i \cap E''_i))_{i\in I}) \cong (G((E'_i \cap E''_i)^{v'_i}) \times G((E'_i \cap E''_i)^{v''_i}), g'_{ij} \times g''_{ij})_{i\leq j\in I},$$

where g'_{ij}, respectively g''_{ij}, is the map induced by $(E'_i \cap E''_i)^{v'_i} \subseteq (E'_j \cap E''_j)^{v'_j}$, respectively by $(E'_i \cap E''_i)^{v''_i} \subseteq (E'_j \cap E''_j)^{v''_j}$. We claim that this last projective system of (Pythagorean) fields is isomorphic to $(G(F'_i) \times G(F''_i), f'_{ij} \times f''_{ij})_{i\leq j\in I}$. It suffices to check that, for instance, the projective system $(G((E'_i \cap E''_i)^{v'_i}), g'_{ij})_{i\leq j\in I}$ is isomorphic to $(G(F'_i), f'_{ij})_{i\leq j\in I}$. This is the content of the remainder of this section.

Since $((E'_i \cap E''_i)^{v'_i}, v'_i)$ is an immediate extension of $(E'_i \cap E''_i, v'_i)$, we have $\operatorname{res}((E'_i \cap E''_i)^{v'_i}, v'_i)_{i\in I} \cong \mathscr{F}'$, so

$$G(\operatorname{res}((E'_i \cap E''_i)^{v'_i}, v'_i)_{i\in I}) \cong G(\mathscr{F}') = (G(F'_i), f'_{ij})_{i\leq j\in I}$$

and it suffices to show that $(G((E_i' \cap E_i'')^{v_i'}, g_{ij}')_{i \leq j \in I}$ and $G(\mathrm{res}((E_i' \cap E_i'')^{v_i'}, v_i')_{i \in I}))$ are isomorphic.

We are now in position to conclude by using the following adaptation of the Baer-Krull theorem [Dickmann and Miraglia 2000, Theorem 1.33]. Recall that the functor G is well defined, in general, from the category of *unitary commutative rings* into the category of L_{SG}*-structures*.

Lemma 5.13. *Let (K, v) be a valued field, $i : O_K \to K$ be the inclusion and $q : O_K \to \overline{K}$ be the projection on the quotient ($\overline{K} = O_K/M_K$). Suppose that $v(2) = 0$.*

(1) *The L_{SG}-structures $G(K)$ and $G(\overline{K})$ are special groups. The induced L_{SG}-morphism $G(i) : G(O_K) \to G(K)$ is injective and the induced L_{SG}-morphism $G(q) : G(O_K) \to G(\overline{K})$ is surjective.*

(2) *If (K, v) is 2-henselian and $vK = 2vK$, then $G(i) : G(O_K) \to G(K)$ and $G(q) : G(O_K) \to G(\overline{K})$ are L_{SG}-isomorphisms. In particular, the L_{SG}-structure $G(O_K)$ is a special group.*

(3) *If $(K', v') \supseteq (K, v)$ is a valued field extension, then $O_K \subseteq O_{K'}$, $M_K \subseteq M_{K'}$ and the diagram of special groups below is commutative (where the vertical arrows are induced by the field extension).*

$$\begin{tikzcd}
G(\overline{K}) \arrow[d] & G(O_K) \arrow[l, "G(q)"'] \arrow[r, "G(i)"] \arrow[d] & G(K) \arrow[d] \\
G(\overline{K}') & G(O_{K'}) \arrow[l, "G(q')"'] \arrow[r, "G(i')"] & G(K')
\end{tikzcd}$$

Proof. (1) Since $v(2) = 0$, 2 is invertible in the rings K, O_K and $\overline{K}$, and therefore, as K and $\overline{K}$ are fields, the L_{SG}-structures $G(K)$ and $G(\overline{K})$ are special groups ([Dickmann and Miraglia 2000, Theorem 1.32 p.23]). As $q : O_K \to \overline{K}$ is a surjective ring homomorphism, it induces a surjective group homomorphism $\dot{O}_K/\dot{O}_K^2 \to \dot{\overline{K}}/\dot{\overline{K}}^2$ and therefore $G(q) : G(O_K) \to G(\overline{K})$ is a surjective L_{SG}-morphism. Now let $a \in \dot{O}_K$ such that $a.\dot{K}^2 = 1.\dot{K}^2$; i.e.,there is $b \in \dot{K}$ such that $a = b^2$, then $2v(b) = v(a) = 0$ and $b \in \dot{O}_K$; therefore $\ker(G(i)) = \{1.\dot{O}_K^2\}$ and $G(i) : G(O_K) \to G(K)$ is an injective L_{SG}-morphism.

(2) We first prove that $G(q)$ is an L_{SG}-isomorphism.

Let $a \in \dot{O}_K$ such that $q(a).\dot{\overline{K}}^2 = 1.\dot{\overline{K}}^2$ then, as $q : O_K \to \overline{K}$ is a surjective ring homomorphism, there is $b \in \dot{O}_K$ such that $q(a) = q(b^2)$. Consider now the polynomial $P(t) = t^2 - a$ in $O_K[t]$: it is a quadratic monic polynomial such that $q(b) \in \overline{K}$ is a root of $P^q(t) = t^2 - q(a)$ in $\overline{K}$ and this root is simple (since $q(a) \neq 0$ and $\mathrm{char}(\overline{K}) \neq 2$). The hypothesis (K, v) 2-henselian then entails that there is $b' \in O_K$ such that $q(b') = q(b)$ and $P(b') = 0$; i.e., $a = b'^2$, for some $b' \in \dot{O}_K$

(because $a \in \dot{O}_K$). Therefore $\ker(G(q)) = \{1.\dot{O}_K^2\}$ and $G(q)$ is an injective L_{SG}-morphism.

We show that whenever $a \in \dot{O}_K$, we have

$$D_{G(\overline{K})}\langle 1, G(q)(a\dot{O}_K^2)\rangle \subseteq G(q)[D_{G(O_K)}\langle 1, a\dot{O}_K^2\rangle].$$

Since q and $G(q)$ are surjective, it is enough to prove that for any $z \in \dot{O}_K$ such that there are $x, y \in O_K$ with $q(z) = 1.q(x)^2 + q(a).q(y)^2$, there are $z' \in \dot{O}_K$ and $x', y' \in O_K$ with $z' = 1.x'^2 + a.y'^2$ and

$$q(z').\dot{\overline{K}}^2 = q(z).\dot{\overline{K}}^2.$$

We recall that $\dot{O}_K = O_K \setminus M_K$ and we split the proof into four cases:

- $x \in M_K, y \in M_K$: it is not possible because $q(z) \neq 0$.
- $x \in \dot{O}_K, y \in M_K$: then $q(z) = q(x^2)$ and the quadratic monic polynomial $P(t) = t^2 - z$ over O_K has a root in $\overline{K}$, and this root is simple (because $q(z) \neq 0$ and $\mathrm{char}(\overline{K}) \neq 2$). By the hypothesis that (K, v) is 2-henselian, P has then a root x' in O_K. Taking this x' as well as $z' := z$ and $y' := 0$ proves the result.
- $x \in M_K, y \in \dot{O}_K$: then $q(z) = q(ay^2)$ and the polynomial $P(t) = t^2 - a^{-1}z$ has a root in $\overline{K}$ and this root is again simple (because $q(z), q(a) \neq 0$ and $\mathrm{char}(\overline{K}) \neq 2$). Therefore P has a root $y' \in O_K$. Taking this y' together with $z' := z$ and $x' = 0$ proves the result.
- $x \in \dot{O}_K, y \in \dot{O}_K$: then $q((x/y)^2 + a - zy^{-2}) = 0$ and the polynomial $P(t) = t^2 + (a - z')$, with $z' := zy^{-2}$, is a quadratic monic polynomial in $O_K[t]$ such that $q(x/y) \in \overline{K}$ is a root of $P^q(t) = t^2 + q(a - z')$ in $\overline{K}$ and we may suppose this root is simple (because, if not, as $\mathrm{char}(\overline{K}) \neq 2$, then $q(a - zy^{-2}) = 0$ and we can proceed as in the case just above). Then the hypothesis (K, v) 2-henselian entails that there is $x' \in O_K$ such that $0 = P(x') = x'^2 + a - z'$ and $q(x') = q(x/y)$, i.e., such that $z' = 1.x'^2 + a.y'^2$, with $y' := 1$. Therefore $G(q)(z.\dot{O}_K^2) = G(q)(z'.\dot{O}_K^2) \in G(q)[D_{G(O_K)}\langle 1, a\dot{O}_K^2\rangle]$.

We now prove that $G(i)$ is an L_{SG}-isomorphism.

As $vK = 2vK$, for any $a \in \dot{K}$ there is $c \in \dot{K}$ such that $v(ac^2) = 0$, i.e., $ac^2 \in \dot{O}_K$. Therefore $G(i)(ac^2.\dot{O}_K^2) = a.\dot{K}^2$ and $G(i)$ is surjective.

To finish the proof, we must check that for each $a \in \dot{O}_K$, we have

$$D_{G(K)}\langle 1, G(i)(a\dot{O}_K^2)\rangle \subseteq G(i)[D_{G(O_K)}\langle 1, a\dot{O}_K^2\rangle].$$

Note that if $a = -b^2$ for some $b \in \dot{O}_K$ then, as $2 \in \dot{O}_K$, we have $D_{G(O_K)}\langle 1, a\dot{O}_K^2\rangle = \dot{O}_K/\dot{O}_K^2$. Since $G(i)$ is a surjective group homomorphism, we have

$$G(i)[D_{G(O_K)}\langle 1, a\dot{O}_K^2\rangle] = \dot{K}/\dot{K}^2$$

and therefore $D_{G(K)}\langle 1, G(i)(a\dot{O}_K^2)\rangle \subseteq G(i)[D_{G(O_K)}\langle 1, a\dot{O}_K^2\rangle]$. Thus we only have to deal with the case $a \notin -\dot{O}_K^2$ and, again as $G(i)$ is a surjective group homomorphism, it is enough to prove that for any $z \in \dot{O}_K$ such that there are $x, y \in K$ with $z = 1.x^2 + a.y^2$, there are $z' \in \dot{O}_K$ and $x', y' \in O_K$ such that $z' = 1.x'^2 + a.y'^2$ and $z'.\dot{K}^2 = z.\dot{K}^2$.

We split the proof into four cases:

- $x, y \in O_K$. Then we simply take $z' := z$, $x' := x$ and $y' := y$.
- $x \in O_K$ and $y \notin O_K$. Then $y^{-1} \in M_K$ and $x/y \in M_K$. Thus $(x/y)^2 \in M_K$ and $1(x/y)^2 + a = zy^{-2} \in M_K$. This implies $a \in M_K$, a contradiction because $a \in \dot{O}_K = O_K \setminus M_K$.
- $x \notin O_K$ and $y = 0$. Then $z = x^2 \notin O_K$, a contradiction.
- $x \notin O_K$ and $y \neq 0$. Then $z = x^2(1+a(y/x)^2) \in \dot{O}_K$ and $x^{-1} \in M_K$. As $z \in \dot{O}_K$, this implies $(1 + a(y/x)^2) = zx^{-2} \in M_K$, and thus $-a(y/x)^2 \in 1 + M_K$. As (K, v) is 2-henselian and $\operatorname{char}(\overline{K}) \neq 2$, $1 + M_K \subseteq O_K^2$ and as $y \neq 0$, then $-a \in K^2$. But $-a \in \dot{O}_K$, so $-a \in \dot{O}_K \cap K^2 = \dot{O}_K^2$, contradicting the hypothesis $a \in \dot{O}_K \setminus -\dot{O}_K^2$.

(3) It follows directly from the definition of extension of valued fields that the following diagram of (local) rings and (local) homomorphisms is commutative:

$$\begin{tikzcd}
\overline{K} \arrow[d] & O_K \arrow[l, "i"'] \arrow[r, "q"] \arrow[d] & K \arrow[d] \\
\overline{K}' & O_{K'} \arrow[l, "i'"'] \arrow[r, "q'"] & K'
\end{tikzcd}$$

The result follows by applying the functor G to it. $\square$

Under the hypotheses of Lemma 5.13, the last item gives us in particular the commutative diagram

$$\begin{tikzcd}
G(\overline{K}) \arrow[r, "\tau_K"] \arrow[d] & G(K) \arrow[d] \\
G(\overline{K}') \arrow[r, "\tau_{K'}"] & G(K')
\end{tikzcd} \tag{5-2}$$

where the maps $\tau_K := G(q) \circ G(i)^{-1}$ and $\tau_{K'} := G(q') \circ G(i')^{-1}$ are isomorphisms of special groups whenever (K, v) and (K', v') are 2-henselian with 2-divisible value groups, and the vertical maps are induced by the field inclusions.

Since, for $i \leq j \in I$, we have an extension of valued fields $((E'_i \cap E''_i)^{v'_i}, v'_i) \subseteq ((E'_j \cap E''_j)^{v'_j}, v'_j)$ and these two fields are 2-henselian with divisible value groups, we conclude that the diagram

$$\begin{tikzcd} G(\overline{(E'_i \cap E''_i)^{v'_i}}) \arrow[r, "\tau_i"] \arrow[d] & G((E'_i \cap E''_i)^{v'_i}) \arrow[d] \\ G(\overline{(E'_j \cap E''_j)^{v'_j}}) \arrow[r, "\tau_j"] & G((E'_j \cap E''_j)^{v'_j}) \end{tikzcd}$$

is commutative, where the maps τ_i and τ_j are the isomorphisms corresponding to τ_K and $\tau_{K'}$ in (5-2). This concludes the proof of Theorem 5.1.

Gluing, the extension case. Now assume that $G_\perp \cong G'[H]$, the last case discussed on page 269. Here we have $G_i \cong G'[H]/\Delta_i$ for every $i \in I$, where H is a fixed finite group of exponent 2, Δ_i is a saturated subgroup of $G'[H]$ and $\Delta_\perp = \{1\}$. Furthermore, if $i \leq j \in I$ we have $\Delta_i \subseteq \Delta_j$ and f_{ij} is naturally identified with the canonical projection from $G'[H]/\Delta_i$ onto $G'[H]/\Delta_j$.

In view of this, the following theorem is a reformulation of the last case in the induction step (page 269), and this section is devoted to its proof.

Theorem 5.14. *Let $(I, \leq)$ be a finite downward directed index set with first element $\perp$ and last element $\top$. Let G' be a reduced special group and assume that whenever $\mathcal{G} = (G_i, \eta_{ij})_{i \leq j \in I}$ is a projective system of reduced special groups with $G_\perp = G'$, then $\mathcal{G}$ is realized by a projective system of Pythagorean fields of characteristic zero (where the morphisms are inclusions).*

Let H be a finite group of exponent 2 and let $(\Delta_i)_{i \in I}$ be a projective system of saturated subgroups of $G'[H]$, where the morphisms are inclusions. Let $\mathcal{G}'$ be the projective system indexed by I of the special groups $G'[H]/\Delta_i$, where the morphisms are the canonical projections.

Then $\mathcal{G}'$ is realized by a projective system of Pythagorean fields of characteristic zero (where the morphisms are inclusions).

Notation: If G is a special group and H is a group of exponent 2, we will identify G (respectively H) with the subgroup $G \times \{1\}$ (respectively $\{1\} \times H$) in $G[H] = \{(g, h) \mid g \in G, h \in H\}$ and write $g \cdot h$ for the pair (g, h).

As $H \cong H_1 \times H_2$ entails $G'[H] \cong (G'[H_1])[H_2]$, we may assume $\dim_{\mathbb{F}_2} H = 1$, i.e., $H = \{1, h\}$ with $h^2 = 1$ and $h \neq 1$.

We define, for $i \in I$ and $i \leq j \in I$:

$$\Omega_i := \Delta_i \cap G', \quad G''_i := G'/\Omega_i \ \text{ (note that } \Omega_i \subseteq \Omega_j)$$
$$q_{ij} : G''_i \to G''_j \text{ the canonical projection,}$$
$$\Theta_i := \{(g \cdot \Omega_i).w \in G''_i[H] \mid g.w \in \Delta_i\}.$$

The following fact is then easily checked:

Fact 5.15. (1) *Ω_i is a saturated subgroup of G'.*

(2) *Θ_i is a saturated subgroup of $G''_i[H]$ with $G''_i \cap \Theta_i = \{1\}$.*

(3) *The morphism of special groups $q_{ij} \times \mathrm{Id} : G''_i[H] \to G''_j[H]$ is such that $(q_{ij} \times \mathrm{Id})(\Theta_i) \subseteq \Theta_j$ and $(q_{ij} \times \mathrm{Id}) \restriction \Theta_i : \Theta_i \to \Theta_j$ is injective.*

(4) *The map*

$$\begin{array}{rrcl} \omega_i : & G'[H]/\Delta_i & \to & (G'/\Omega_i)[H]/\Theta_i \\ & (g \cdot h)/\Delta_i & \mapsto & ((g/\Omega_i) \cdot h)/\Theta_i \end{array}$$

is an isomorphism of special groups.

(5) *The diagram*

$$\begin{tikzcd} G'[H]/\Delta_i \arrow[r, "f_{ij}"] \arrow[d, "\omega_i"'] & G'[H]/\Delta_j \arrow[d, "\omega_j"] \\ G''_i[H]/\Theta_i \arrow[r, "\widetilde{q_{ij} \times \mathrm{Id}}"] & G''_j[H]/\Theta_j \end{tikzcd}$$

commutes, where $\widetilde{q_{ij} \times \mathrm{Id}}$ is the canonical map induced on the quotients.

Note that by hypothesis, since $G''_\perp = G'$, the projective system $(G''_i, q_{ij})_{i \le j \in I}$ is realized by a system of Pythagorean fields $(K_i)_{i \in I}$ of characteristic zero.

To complete the proof, it is then enough to represent the projective system of special groups $(G''_i[H]/\Theta_i, \widetilde{q_{ij} \times \mathrm{Id}})_{i \le j \in I}$ by some projective system of Pythagorean fields of characteristic zero; this is the content of the following proposition.

Proposition 5.16. *There is a projective system of Pythagorean fields of characteristic zero $(L_i)_{i \in I}$, where the morphisms are inclusions, such that*

$$(G''_i[H]/\Theta_i, \widetilde{q_{ij} \times \mathrm{Id}})_{i \le j \in I} \cong G((L_i)_{i \in I}).$$

The rest of this section now consists in the proof of Proposition 5.16.

Let us denote by γ_{ij} the morphism of special groups induced by $K_i \subseteq K_j$:

$$(G(K_i), \gamma_{ij})_{i \le j \in I} \cong (G''_i, q_{ij})_{i \le j \in I}.$$

We define $M_i = K_i((t))$ for every $i \in I$ and record a well known result:

Lemma 5.17. $\dot{M}_i/\dot{M}_i^2 = \{at^k \cdot \dot{M}_i^2 \mid a \in \dot{K}_i, k \in \{0, 1\}\}$, *and the isomorphism of special groups from $G(M_i)$ to $G(K_i)[H]$ is*

$$\begin{array}{rrcl} \lambda_i : & G(M_i) & \to & G(K_i)[H] \\ & at^k \cdot \dot{M}_i^2 & \mapsto & (a \cdot \dot{K}_i^2)h^k. \end{array}$$

Proof. This is exactly [Dickmann and Miraglia 2000, Theorem 1.33], where the explicit definition of the isomorphism is given at the beginning of the proof on page 28. $\square$

It immediately follows that

$$(G(K_i((t))), \gamma_{ij} \times \mathrm{Id})_{i \le j \in I} \cong (G_i''[H], q_{ij} \times \mathrm{Id})_{i \le j \in I}. \tag{5-3}$$

For each $i \in I$, let Γ_i be the saturated subgroup of $G(K_i((t)))$ that corresponds, by the isomorphisms above, to the saturated subgroup Θ_i of $G_i''[H]$. This then yields

$$(G(K_i((t)))/\Gamma_i, \widetilde{\gamma_{ij} \times \mathrm{Id}})_{i \le j \in I} \cong (G_i''[H]/\Theta_i, \widetilde{q_{ij} \times \mathrm{Id}})_{i \le j \in I}$$

(where $\widetilde{\gamma_{ij} \times \mathrm{Id}}$ denotes the induced map on the quotients), which in turn shows that we only have to find a projective system of fields realizing the system

$$(G(K_i((t)))/\Gamma_i, \widetilde{\gamma_{ij} \times \mathrm{Id}})_{i \le j \in I}.$$

Therefore, to keep notation simple, we may assume that $G(K_i)[H] = G_i''[H]$, $\Gamma_i = \Theta_i$, $\gamma_{ij} \times \mathrm{Id} = q_{ij} \times \mathrm{Id}$, and that $\widetilde{q_{ij} \times \mathrm{Id}}$ is the map from $G(K_i)[H]/\Theta_i$ to $G(K_j)[H]/\Theta_j$ induced by $\gamma_{ij} \times \mathrm{Id} = q_{ij} \times \mathrm{Id}$.

In this vein, for every $i \le j \in I$, we will write $G(M_i) = G_i''[H]$, $q_{ij} \times \mathrm{Id}$ will stand for the morphism of special groups induced by the inclusion $M_i \subseteq M_j$, and the diagram

$$\begin{tikzcd} \dot{M}_i \arrow[r] \arrow[d, "p_i"'] & \dot{M}_j \arrow[d, "p_j"'] \\ G(M_i) \arrow[r, "q_{ij}\times\mathrm{Id}"] & G(M_j) \end{tikzcd} \tag{5-4}$$

is commutative, where p_i and p_j denote the canonical maps.

Define $n_i := \dim_{\mathbb{F}_2} \Theta_i$ for $i \in I$. Note that $\dim_{\mathbb{F}_2} \Theta_i \le \dim_{\mathbb{F}_2} H = 1$, so $n_i \in \{0, 1\}$. Since $(q_{ij} \times \mathrm{Id}) \restriction \Theta_i : \Theta_i \to \Theta_j$ is injective, we have $n_i \le n_j$ whenever $i \le j \in I$. If $n_i = 1$, write $\Theta_i = \{1, a_i h\}$, with $a_i \in G_i''$. In this case, and if $i \le j \in I$, we have $(q_{ij} \times \mathrm{Id})(\Theta_i) = \Theta_j$, so $q_{ij}(a_i) = a_j$.

Lemma 5.18. *There is $b \in \dot{M}_\perp$ such that, for every $i \in I$, $\Theta_i \subseteq \{1, p_i(b)\}$.*

Proof. For every $i \le j \in I$, the map q_{ij} is surjective. In particular the map $q_{\perp\top}$ is surjective and, by diagram (5-4) above, $p_\top(\dot{M}_\perp) = \mathrm{Im}(q_{\perp\top} \times \mathrm{Id}) = G(M_\top)$. Let $b \in \dot{M}_\perp$ be such that $\{1, p_\top(b)\} = \Theta_\top$. Let now $i \in I$ and let $x \in \Theta_i$. Then

$$(q_{i\top} \times \mathrm{Id})(x) \in \Theta_\top = \{1, p_\top(b)\}.$$

If $(q_{i\top} \times \mathrm{Id})(x) = 1$, we get $x = 1 \in \{1, p_i(b)\}$, because $(q_{ij} \times \mathrm{Id})\restriction : \Theta_i \to \Theta_j$ is an injective group homomorphism. If $(q_{i\top} \times \mathrm{Id})(x) = p_\top(b)$, since diagram (5-4) is commutative, we get $p_\top(b) = (q_{i\top} \times \mathrm{Id})(p_i(b))$, so $(q_{i\top} \times \mathrm{Id})(x) = (q_{i\top} \times \mathrm{Id})(p_i(b))$ and we conclude that $x = p_i(b)$. □

We assume from now on that there is $i \in I$ such that $n_i = 1$ (equivalently, $n_\top = 1$). Otherwise $\Theta_j = \{1\}$ for every $j \in I$, and the projective system of fields $(M_i)_{i \in I}$ realizes the projective system of special groups $(G_i''[H]/\Theta_i, \widetilde{q_{ij} \times \mathrm{Id}})_{i \leq j \in I}$.

Recall that an element a of a special group T is called rigid when $a \neq 1$ and $D_T\langle 1, a\rangle = \{1, a\}$ and an element b of T is birigid when b and $-b$ are rigid. If $T = G[H]$, then every element in $G[H] \setminus G$ is birigid (this is essentially the only way to obtain birigid elements in a special group; see [Dickmann and Miraglia 2000, p. 12, Berman's Theorem]).

Since we assume that $n_i = 1$ for some $i \in I$ (in other words $n_\top = 1$), it follows that the element b produced in Lemma 5.18 is birigid in M_i for every $i \in I$.

The next proposition uses the following notation: If K is a field and $a \in K$ then $K(\sqrt[\infty]{a})$ stands for $K(\sqrt[2^n]{a}, n \in \mathbb{N})$.

Proposition 5.19 [Becher 2002, Proposition 8.2]. *Let F be a field, let a be a birigid element in F (i.e., $a \in \dot{F}$ and $a.\dot{F}^2$ is birigid in $G(F)$) and let φ be a quadratic form over F. Let $L := F(\sqrt[\infty]{a})$. Then*

(1) *$\dot{L} = \dot{F}\dot{L}^2$ and $\dot{F} \cap \dot{L}^2 = \dot{F}^2 \cup a\dot{F}^2$;*

(2) *φ is isotropic over L if and only if $\varphi \oplus a\varphi$ is isotropic over F.*

We define, for $i \in I$,

$$L_i = \begin{cases} M_i & \text{if } n_i = 0, \\ M_i(\sqrt[\infty]{b}) & \text{if } n_i = 1. \end{cases}$$

Since $n_i = 1$ implies $n_j = 1$ whenever $i \leq j \in I$, the system $(L_i)_{i \in I}$ is a projective system of fields. Note that the following diagram of fields is obviously commutative (with the natural inclusions as morphisms):

$$\begin{array}{ccc} M_i & \longrightarrow & M_j \\ \downarrow & & \downarrow \\ L_i & \longrightarrow & L_j \end{array}$$

which implies that the induced diagram of special groups is also commutative:

$$\begin{array}{ccc} G(M_i) & \xrightarrow{q_{ij} \times \mathrm{Id}} & G(M_j) \\ \mu_i \downarrow & & \downarrow \mu_j \\ G(L_i) & \xrightarrow{\tau_{ij}} & G(L_j) \end{array} \tag{5-5}$$

where $\mu_i : G(M_i) \to G(L_i)$ is the map induced by $M_i \subseteq L_i$ and τ_{ij} is the map induced by $L_i \subseteq L_j$.

Lemma 5.20. *For $i \in I$, let $\pi_i : G(K_i)[H] \to G(K_i)[H]/\Theta_i$ be the canonical projection. Then μ_i is surjective and there is a unique isomorphism of special groups $\xi_i : G(L_i) \to G(K_i)[H]/\Theta_i$ such that the diagram*

$$\begin{tikzcd} G(M_i) \arrow[r, "\mu_i"] \arrow[d, "\lambda_i"'] & G(L_i) \arrow[d, "\xi_i"] \\ G(K_i)[H] \arrow[r, "\pi_i"] & G(K_i)[H]/\Theta_i \end{tikzcd}$$

is commutative. In particular, L_i is Pythagorean.

Proof. The case $n_i = 0$ is trivial, so we assume $n_i = 1$. To avoid unnecessary notational complications, if K is a field and $x \in \dot{K}$, we simply write $\bar{x}$ for the class of x in $\dot{K}/\dot{K}^2$. By Proposition 5.19(1) we know that μ_i is surjective and that $\ker(\mu_i \circ \lambda_i^{-1}) = \{1, \lambda_i(\bar{b})\} = D_{G(K_i)[H]}\langle 1, \lambda_i(\bar{b})\rangle$. In particular, there is a unique isomorphism of groups $\xi_i : G(L_i) \to G(K_i)[H]/\{1, \lambda_i(\bar{b})\}$ such that the following diagram commutes:

$$\begin{tikzcd} G(K_i)[H] \arrow[r, "\lambda_i^{-1}"] \arrow[d, "\pi_i"'] & G(M_i) \arrow[r, "\mu_i"] & G(L_i) \arrow[lld, "\xi_i"] \\ G(K_i)[H]/\{1, \lambda_i(\bar{b})\} & & \end{tikzcd}$$

We show that ξ_i is an isomorphism of special groups. The image of -1 is clearly -1. Take $\mu_i(\bar{c}), \mu_i(\bar{d}) \in G(L_i)$, where $c, d \in \dot{M}_i$. We have

$$\begin{aligned} \mu_i(\bar{c}) \in D_{G(L_i)}\langle 1, \mu_i(\bar{d})\rangle &\Leftrightarrow c \in D_{L_i}\langle 1, d\rangle \\ &\Leftrightarrow \langle\langle -c, d\rangle\rangle \text{ isotropic over } L_i \\ &\Leftrightarrow \langle\langle -c, d\rangle\rangle \oplus b\langle\langle -c, d\rangle\rangle \text{ isotropic over } M_i, \end{aligned}$$

the last equivalence following from Proposition 5.19(2). Recalling that Pfister forms are isotropic if and only if they are hyperbolic, we continue the chain of equivalences with

$$\begin{aligned} &\Leftrightarrow \langle 1, b\rangle \otimes \langle\langle -c, d\rangle\rangle \text{ isotropic over } M_i \\ &\Leftrightarrow \langle 1, b\rangle \otimes \langle\langle -c, d\rangle\rangle \text{ hyperbolic over } M_i \\ &\Leftrightarrow \langle 1, \lambda_i(\bar{b})\rangle \otimes \langle\langle -\lambda_i(\bar{c}), \lambda_i(\bar{d})\rangle\rangle \text{ hyperbolic in } G(K_i)[H] \\ &\Leftrightarrow \langle 1, \lambda_i(\bar{b})\rangle \otimes \langle\langle -\lambda_i(\bar{c}), \lambda_i(\bar{d})\rangle\rangle \equiv \langle 1, \lambda_i(\bar{b})\rangle \otimes \langle -1, 1, -1, 1\rangle \text{ in } G(K_i)[H] \\ &\Leftrightarrow \langle\langle -\pi_i \circ \lambda_i(\bar{c}), \pi_i \circ \lambda_i(\bar{d})\rangle\rangle \equiv \langle -1, 1, -1, 1\rangle \\ &\qquad\qquad \text{in } G(K_i)[H]/D\langle 1, \lambda_i(\bar{b})\rangle = G(K_i)[H]/\{1, \lambda_i(\bar{b})\}, \end{aligned}$$

the last step following from [Dickmann and Miraglia 2000, Proposition 2.21]. But this last condition is equivalent to $\langle\langle -\pi_i \circ \lambda_i(\bar{c}), \pi_i \circ \lambda_i(\bar{d})\rangle\rangle$ being hyperbolic in

$G(K_i)[H]/\{1, \lambda_i(\bar{b})\}$, and so to $\pi_i \circ \lambda_i(\bar{c})$ lying in $D_{G(K_i)[H]/\{1,\lambda_i(\bar{b})\}}\langle 1, \pi_i \circ \lambda_i(\bar{d})\rangle$. This shows that ξ_i is an isomorphism of special groups. Since $\lambda_i(\bar{b})$ is birigid and $|H| = 2$, we obtain that $G(L_i)$, being isomorphic to $G(K_i)[H]/\{1, \lambda_i(\bar{b})\} \cong G(K_i)$, is a reduced special group or $\{1\}$, which entails that L_i is a Pythagorean field. $\square$

Recall that, using the identifications made after (5-3), we have

$$(\widetilde{q_{ij} \times \mathrm{Id}})((a\dot{K}_i^{\,2})h^k \cdot \Theta_i) = (a\dot{K}_j^{\,2})h^k \cdot \Theta_j \quad \text{for } a \in \dot{K}_i \text{ and } k \in \{0, 1\}.$$

Proposition 5.21. *The diagram*

$$\begin{tikzcd} G(L_i) \arrow[r, "\tau_{ij}"] \arrow[d, "\xi_i"'] & G(L_j) \arrow[d, "\xi_j"] \\ G(K_i)[H]/\Theta_i \arrow[r, "\widetilde{q_{ij} \times \mathrm{Id}}"] & G(K_j)[H]/\Theta_j \end{tikzcd}$$

commutes. In particular, $G((L_i)_{i\in I}) \cong (G_i''[H]/\Theta_i, \widetilde{q_{ij} \times \mathrm{Id}})_{i\le j\in I}$ *and*

$$\varprojlim G((L_i)_{i\in I}) \cong \varprojlim (G_i, f_{ij})_{i\le j\in I}.$$

Proof. Since μ_i and μ_j are surjective by Lemma 5.20, the commutative diagram in that same lemma completely determines ξ_i and ξ_j. Let $z = \mu_i(at^k\dot{M}_i^{\,2}) \in G(L_i)$ (with $a \in \dot{K}_i$ and $k \in \{0, 1\}$). Then

$$\xi_i(z) = \xi_i \circ \mu_i(at^k\dot{M}_i^{\,2}) = \pi_i \circ \lambda_i(at^k\dot{M}_i^{\,2}) = \pi_i((a\dot{K}_i^{\,2})h^k) = ((a\dot{K}_i^{\,2})h^k)\cdot\Theta_i. \tag{5-6}$$

where the second equality comes from Lemma 5.20. Applying this, we obtain $(\widetilde{q_{ij} \times \mathrm{Id}}) \circ \xi_i(z) = (\widetilde{q_{ij} \times \mathrm{Id}})(((a\dot{K}_i^{\,2})h^k) \cdot \Theta_i) = (a\dot{K}_j^{\,2})h^k \cdot \Theta_j$ and

$$\begin{aligned} \xi_j \circ \tau_{ij}(z) &= \xi_j \circ \tau_{ij} \circ \mu_i(at^k\dot{M}_i^{\,2}) \\ &= \xi_j \circ \mu_j \circ (q_{ij} \times \mathrm{Id})(at^k\dot{M}_i^{\,2}) && \text{by diagram (5-5)} \\ &= \xi_j \circ \mu_j(at^k\dot{M}_j^{\,2}) && \text{since } q_{ij} \times \mathrm{Id} \text{ is induced by } M_i \subseteq M_j \\ &= (a\dot{K}_j^{\,2})h^k \cdot \Theta_j && \text{by (5-6),} \end{aligned}$$

which finishes the proof. $\square$

Acknowledgement

Both authors wish to thank the referee for many detailed and helpful comments.

References

[Astier and Tressl 2005] V. Astier and M. Tressl, "Axiomatization of local-global principles for pp-formulas in spaces of orderings", *Arch. Math. Logic* **44**:1 (2005), 77–95. MR 2005k:03088 Zbl 1099.03027

[Becher 2002] K. J. Becher, "Le radical de Kaplansky", Publ. Mat. de Besançon, Univ. Franche-Comté, 2002, available at http://www.math.uni-konstanz.de/∼becher/art/kap.pdf. MR 2004g:11030 Zbl 1161.11328

[Bröcker 1977] L. Bröcker, "Über die Anzahl der Anordnungen eines kommutativen Körpers", *Arch. Math.* (*Basel*) **29**:5 (1977), 458–464. MR 58 #5613

[Chang and Keisler 1990] C. C. Chang and H. J. Keisler, *Model theory*, 3rd ed., Studies in Logic and the Foundations of Math. **73**, North-Holland, Amsterdam, 1990. MR 91c:03026 Zbl 0697.03022

[Dickmann and Miraglia 2000] M. A. Dickmann and F. Miraglia, *Special groups: Boolean-theoretic methods in the theory of quadratic forms*, Mem. Amer. Math. Soc. **689**, Amer. Math. Soc., Providence, RI, 2000. MR 2000j:11052 Zbl 1052.11027

[Dickmann and Miraglia 2003] M. A. Dickmann and F. Miraglia, "Bounds for the representation of quadratic forms", *J. Algebra* **268**:1 (2003), 209–251. MR 2004k:11052 Zbl 1053.11035

[Endler 1963] O. Endler, "Über einen Existenzsatz der Bewertungstheorie", *Math. Ann.* **150** (1963), 54–65. MR 26 #3732 Zbl 0142.28302

[Engler and Prestel 2005] A. J. Engler and A. Prestel, *Valued fields*, Springer, Berlin, 2005. MR 2007a:12005 Zbl 1128.12009

[Heinemann 1985] B. Heinemann, "On finite intersections of "Henselian valued" fields", *Manuscripta Math.* **52**:1-3 (1985), 37–61. MR 87a:12013

[Hodges 1993] W. Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications **42**, Cambridge University Press, Cambridge, 1993. MR 94e:03002 Zbl 0789.03031

[Kula 1979] M. Kula, "Fields with prescribed quadratic form schemes", *Math. Z.* **167**:3 (1979), 201–212. MR 80f:10024 Zbl 0388.10017

[Kula et al. 1984] M. Kula, M. A. Marshall, and A. Sładek, "Direct limits of finite spaces of orderings", *Pacific J. Math.* **112**:2 (1984), 391–406. MR 85f:11023 Zbl 0535.10020

[Lira de Lima 1997] A. Lira de Lima, *Les groupes spéciaux: aspects algébriques et combinatoires de la théorie des espaces d'ordres abstraits*, Ph.D. thesis, Université de Paris 7, Paris, 1997.

[Mariano 2003] H. L. Mariano, *Contribuições à teoria dos grupos especiais*, Ph.D. thesis, Universidade de São Paulo, São Paulo, 2003.

[Marshall 1980] M. Marshall, "Spaces of orderings, IV", *Canad. J. Math.* **32**:3 (1980), 603–627. MR 81m:10035 Zbl 0433.10009

Received February 1, 2010. Revised September 7, 2010.

VINCENT ASTIER
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY COLLEGE DUBLIN
BELFIELD, DUBLIN 4
IRELAND
vincent.astier@ucd.ie

HUGO MARIANO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA DA UNIVERSIDADE DE SÃO PAULO
RUA DO MATÃO, 1010 (CIDADE UNIVERSITÁRIA)
05508-090 SÃO PAULO
BRAZIL
hugomar@ime.usp.br

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

ON FIBERED COMMENSURABILITY

DANNY CALEGARI, HONGBIN SUN AND SHICHENG WANG

This paper initiates a systematic study of the relation of commensurability of surface automorphisms, or equivalently, fibered commensurability of 3-manifolds fibering over S^1. We show that every hyperbolic fibered commensurability class contains a unique minimal element. The situation for toroidal manifolds is more complicated, and we illustrate a range of phenomena that can occur in this context.

1. Introduction

The main purpose of this paper is to study the equivalence relation of *commensurability* of surface automorphisms. Informally, two surface automorphisms are commensurable if they lift to automorphisms of a finite covering surface that have nontrivial common powers. Equivalently, a surface automorphism determines a foliation of a 3-manifold by closed surfaces, and two automorphisms are commensurable if their corresponding 3-manifolds admit common finite covers for which the pulled-back foliations are isotopic. Thus commensurability of surface automorphisms is a special case of the study of commensurability of 3-manifolds equipped with a certain kind of geometric structure; again informally, we call this commensurability relation *fibered commensurability*.

The relation of commensurability of 3-manifolds is well-studied; see, for example, [Thurston 1979, Chapter 6; Borel 1981; Macbeath 1983; Neumann 1997; Behrstock and Neumann 2010]. When studying commensurability in a given context, the most important distinction to make is between those commensurability classes that admit *finitely many* minimal elements, and those that admit *infinitely* many. For example, amongst hyperbolic 3-manifolds, this is precisely the distinction between nonarithmetic and arithmetic commensurability classes; see [Margulis 1991; Borel 1981], for instance. This distinction has a cleaner statement if one is prepared to work in the category of orbifolds: each commensurability class of nonarithmetic hyperbolic 3-manifolds contains a unique minimal element.

Fibered commensurability is more rigid than ordinary commensurability. However, a given 3-manifold can fiber in infinitely many different ways. For Seifert

MSC2000: 57M50.
Keywords: commensurability, fibration, 3-manifold, mapping class group.

manifolds, there is exactly one fibered commensurability class of surface bundles of all closed (resp. with torus boundary) Seifert fibered manifolds whose fiber has negative Euler characteristic, and this class contains infinitely many minimal elements. On the other hand, in the hyperbolic world we obtain:

Theorem 3.1 (Hyperbolic Theorem). *Every commensurability class of hyperbolic fibered pairs contains a unique* (*orbifold*) *minimal element.*

An immediate corollary is that for a fibered hyperbolic 3-manifold M, each fibered commensurability class contains at most finitely many fibrations of M; hence M has either one fibered commensurability class, or infinitely many fibered commensurability classes.

The reducible case is more complicated:

Examples 5.3 and 5.5 (Toroidal cases). *There are examples of graph manifolds with infinitely many fibered commensurability classes*, *and a single graph manifold can fiber in infinitely many ways in a single commensurability class.*

As these results suggest, obstructions to commensurability of surface automorphisms arise from their behavior on pseudo-Anosov orbits, and near their reducing systems. We describe such obstructions in detail.

In Section 2, we give basic definitions and illustrate their meaning, in the special case of commensurability of spherical and toral automorphisms. We recall the Nielsen–Thurston classification of surface automorphisms, and discuss a "normal form" for automorphisms. This standard material may be skipped by the expert.

In Section 3, we study fibered commensurability of hyperbolic manifolds, and prove Theorem 3.1. We also list some commensurability invariants of pseudo-Anosov automorphisms (Lemma 3.10 and Proposition 3.15), and describe examples that illustrate their use.

Finally, Section 4 and Section 5 are devoted to the case of reducible automorphisms, especially of graph manifolds. In Section 4 we define certain numerical commensurability invariants for reducible maps (Theorem 4.3, as well as Proposition 4.11), and give many examples. In Section 5 we give examples of graph manifolds with infinitely many incommensurable fibrations, including one with boundary (Example 5.3) that also admits infinitely many *commensurable* (but nonisomorphic) fibrations, and a closed one (Example 5.5) that admits incommensurable fibrations of the same genus.

2. Fibered commensurability

Basic definitions. Let F be a compact surface. An *automorphism* ϕ of F is an isotopy class of self-homeomorphisms of F. We use the notation (F, ϕ) where ϕ is an automorphism of F.

Remark 2.1. When F has boundary, it is more usual to study isotopy classes of self-homeomorphisms fixed pointwise on the boundary. However, since we are interested in automorphisms which might permute boundary components, we adhere to this nonstandard convention.

One surface automorphism can "cover" another in two distinct ways: either topologically (in the sense that one surface covers the other) or dynamically (in the sense that one automorphism is a power of another). We consider covering in both senses in the sequel. More formally, we make the following definition.

Definition 2.2. A pair $(\tilde{F}, \tilde{\phi})$ *covers* (F, ϕ) if there is a finite cover $\pi : \tilde{F} \to F$ and representative homeomorphisms $\tilde{f}$ and f of $\tilde{\phi}$ and ϕ respectively so that $\pi \circ \tilde{f} = f \circ \pi$ as maps $\tilde{F} \to F$.

Remark 2.3. The relation of covering is transitive: if (F_1, ϕ_1) covers (F_2, ϕ_2), and (F_2, ϕ_2) covers (F_3, ϕ_3), then (F_1, ϕ_1) covers (F_3, ϕ_3). This follows by appealing to a "normal form" for representative homeomorphisms which is compatible with finite covers. This normal form is well-known, and summarized in Theorem 2.14 and Proposition 2.15 below.

An automorphism ϕ of F determines an outer automorphism ϕ_* of $\pi_1(F)$ preserving peripheral subgroups, and by the well-known theorem of Dehn and Nielsen [Nielsen 1927], this correspondence is a bijection. A cover $\tilde{F}$ determines a conjugacy class of subgroups G of $\pi_1(F)$, and an automorphism ϕ of F lifts to an automorphism $\tilde{\phi}$ of $\tilde{F}$ if and only if G and $\phi_*(G)$ are conjugate in $\pi_1(F)$. However, a particular lift $\tilde{\phi}$ depends on a choice of conjugating element. Thus a finite cover of surfaces $\tilde{F} \to F$ might determine zero, one, or many covers of automorphisms $(\tilde{F}, \tilde{\phi}) \to (F, \phi)$ (even if $\tilde{\phi}$ is primitive).

Example 2.4. If $\tilde{F} \to F$ is any finite cover, then (F, id) is covered by $(\tilde{F}, \psi)$ where ψ is any element of the deck group of the cover.

Definition 2.5. Two automorphisms (F_1, ϕ_1) and (F_2, ϕ_2) are *commensurable* if there is a surface $\tilde{F}$, automorphisms $\tilde{\phi}_1$ and $\tilde{\phi}_2$ of $\tilde{F}$, and nonzero integers k_1 and k_2, so that $(\tilde{F}, \tilde{\phi}_i)$ covers (F_i, ϕ_i) for $i = 1, 2$, and if $\tilde{\phi}_1^{k_1} = \tilde{\phi}_2^{k_2}$ as automorphisms of $\tilde{F}$. Moreover say (F_1, ϕ_1) and (F_2, ϕ_2) are *topologically commensurable* if $|k_1| = |k_2| = 1$, and *dynamically commensurable* if $\tilde{F} = F_1 = F_2$.

Commensurability of automorphisms is readily seen to be an equivalence relation, and is the main object of study in this paper.

Statements about surfaces and automorphisms can usefully be translated into statements about 3-manifolds with certain types of foliations. These objects — "fibered pairs", to be defined below — admit natural generalizations to objects called *orbifold* fibered pairs, that are awkward to discuss in the language of surfaces and automorphisms. Certain theorems in this paper are more elegantly stated and proved in this category. A basic reference for the theory of orbifolds is [Thurston 1979, Chapter 13].

Definition 2.6. A *fibered pair* is a pair $(M, \mathscr{F})$ where M is a compact 3-manifold with boundary a union of tori and Klein bottles, and $\mathscr{F}$ is a foliation by compact surfaces. More generally, an *orbifold fibered pair* is a pair $(O, \mathscr{G})$ where O is a compact 3-orbifold, and $\mathscr{G}$ is a foliation of O by compact 2-orbifolds.

At interior points (resp. boundary points) an orbifold fibered pair $(O, \mathscr{G})$ looks locally like the quotient of an open ball in $\mathbb{R}^3$ (resp. a relatively open ball in a vertical half-space) foliated by horizontal planes by a finite group of smooth foliation-preserving homeomorphisms.

A surface automorphism (F, ϕ) determines a fibered pair whose underlying manifold is an F bundle over S^1 with monodromy ϕ, and whose foliation is the foliation by surface fibers (which are all homeomorphic to F). If we want to emphasize its dynamical origin, we use the notation $[F, \phi]$ in the sequel to denote the fibered pair associated to the automorphism (F, ϕ).

If the underlying orbifold O is *good* (i.e., it admits a finite manifold cover) then $(O, \mathscr{G})$ is finitely covered by a pair $(M, \mathscr{F})$ where M is a manifold, and every leaf of $\mathscr{F}$ is a compact surface. After passing to a further 2-fold cover if necessary, we can assume $\mathscr{F}$ is co-orientable, in which case M fibers over S^1 in such a way that the leaves of $\mathscr{F}$ are the fibers.

Definition 2.7. A fibered pair $(\tilde{M}, \tilde{\mathscr{F}})$ *covers* $(M, \mathscr{F})$ if there is a finite covering of manifolds $\pi : \tilde{M} \to M$ such that $\pi^{-1}(\mathscr{F})$ is isotopic to $\tilde{\mathscr{F}}$. Two fibered pairs $(M_1, \mathscr{F}_1)$ and $(M_2, \mathscr{F}_2)$ are *commensurable* if there is a third fibered pair $(\tilde{M}, \tilde{\mathscr{F}})$ that covers both.

If $(M_i, \mathscr{F}_i)$ for $i = 1, 2$ are fibered pairs with co-orientable foliations, then they are commensurable in the sense of Definition 2.7 if and only if the associated surface automorphisms are commensurable. Thus, the category of fibered pairs enlarges the category of surface automorphisms in such a way that the definition of commensurability of a surface automorphism is the same, whichever category we use.

To stress that the definition of commensurability of fibered pairs depends on both the underlying 3-manifold and the foliation, we call this equivalence relation *fibered commensurability*.

The relation of covering is transitive, but it is not yet a partial order because of the existence of automorphisms of finite order. We must take such examples into account in order to define minimal elements with respect to commensurability.

Definition 2.8. We say that two fibered pairs $(M, \mathscr{F})$ and $(N, \mathscr{G})$ are *covering equivalent* if each covers the other. Call a covering equivalence class *minimal* if no representative covers any element of another covering equivalence class.

The relation of covering descends to a transitive relation on covering equivalence classes, and defines a partial order on such classes. Minimal classes are minimal with respect to this partial order.

Remark 2.9. Each covering equivalence class of fibered pairs $[F, \phi]$ contains exactly one fibered pair unless ϕ is periodic. In the periodic case, (F, ϕ) and (G, ψ) are in the same covering equivalent class if and only if $F = G$ and both ϕ and ψ generate the same finite cyclic group. With this understood, in the sequel we are relaxed in our terminology, and use the word "minimal element" when we really mean "minimal class".

Simple cases. For simplicity, we usually restrict attention to the case that F (and therefore M) is closed. However, because of the nature of the theory of surface automorphisms, to really understand this case we are forced to consider surfaces (and 3-manifolds) with boundary, associated to the restrictions of automorphisms to invariant subsurfaces.

Evidently, the sign of $\chi(F)$ is a commensurability invariant of (F, ϕ). In the case of fibered pairs (of good orbifolds), all leaves have the same sign, so we can speak unambiguously about fibered pairs with spherical, Euclidean, or hyperbolic leaves. We first discuss the situation when $\chi(F) \geq 0$.

Example 2.10 (spherical automorphisms). There is only one commensurability class, consisting of the bundles $S^2 \times S^1$ and $S^2 \tilde{\times} S^1$, each foliated by spheres, and $\mathbb{RP}^3 \# \mathbb{RP}^3$, which can be thought of as an S^2 bundle over a mirror orbifold. The elements $S^2 \tilde{\times} S^1$ and $\mathbb{RP}^3 \# \mathbb{RP}^3$ are minimal.

Example 2.11 (toral automorphisms). The mapping class group of a torus is isomorphic to $\mathrm{GL}(2, \mathbb{Z})$, and every automorphism has a linear representative. An automorphism can be periodic, reducible, or Anosov. From elementary linear algebra, automorphisms in different classes are not commensurable. We discuss each case in turn.

(1) Periodic case: there is only one commensurability class; moreover there are exactly two minimal elements, corresponding to the periodic automorphisms of order 4 and 6 on a square and hexagonal torus respectively.

(2) Reducible case: as automorphisms, each map (T, ϕ) is represented by a matrix which can be conjugated into the form

$$\phi \sim \pm \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$$

where $n \neq 0$. So there is only one commensurability class and two minimal elements, corresponding to the conjugacy classes of matrices

$$\phi \sim \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

(3) Anosov case: the resulting Sol manifolds are commensurable if and only if they are fibered commensurable, which occurs if and only if the logarithms of the dilatations of the automorphisms are commensurable as real numbers. Hence there are infinitely many fibered commensurability classes.

Standard form for surface automorphisms. In the remainder of the paper therefore we concentrate on the case of surfaces F with $\chi(F) < 0$. Furthermore, unless we explicitly say to the contrary, all surfaces F are assumed to be compact and connected.

A commensurability between automorphisms restricts to a commensurability between the underlying surfaces. A complete set of commensurability invariants of compact surfaces are the sign of Euler characteristic, and the property of possessing (or not possessing) a nonempty boundary.

Lemma 2.12. *Let F_1 and F_2 be compact surfaces with $\chi < 0$. If both or neither have nonempty boundary, they are commensurable. Otherwise they are incommensurable.*

The proof is elementary; see [Massey 1974], for example. Since every compact surface orbifold with $\chi < 0$ is good, the lemma extends to orbifolds.

Notation 2.13. Suppose Γ (resp. F') is a union of circles (resp. a compact subsurface) in F. Let $F \setminus \Gamma$ (resp. $F \setminus F'$) denote the compact surface obtained by splitting F along Γ (resp. removing $\mathrm{int} F'$, the interior of F').

Recall the Nielsen–Thurston classification of surface automorphisms.

Theorem 2.14 [Thurston 1988; Fathi et al. 1979]. *Let ϕ be an automorphism of a compact surface F. Then the isotopy class of ϕ has a representative (which by abuse of notation we continue to denote by ϕ) so that either*

(1) *ϕ has finite order, and $[F, \phi]$ is a Seifert manifold with $\mathbb{H}^2 \times \mathbb{R}$ geometry; or*

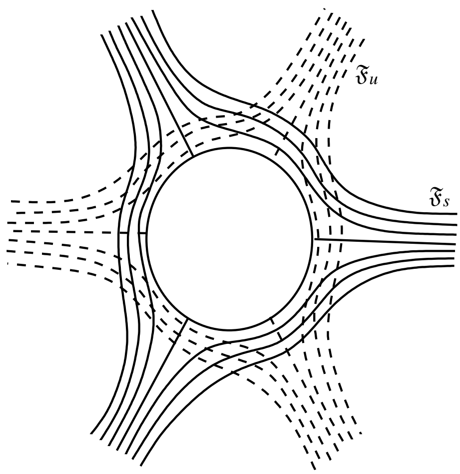
(2) *ϕ is pseudo-Anosov — i.e., F admits a pair of transversely measured singular foliations $\mathfrak{F}_s$ and $\mathfrak{F}_u$ with measures μ_s, μ_u, and there is a real number $\lambda > 1$ called the* dilatation *so that ϕ takes each foliation to itself, stretching μ_u by λ and compressing μ_s by $1/\lambda$ — and the interior of $[F, \phi]$ admits a complete hyperbolic structure of finite volume; or*

(3) *ϕ is reducible — i.e., there is a minimal nonempty embedded 1-manifold Γ in F with a ϕ-invariant tubular neighborhood $N(\Gamma)$ such that on each ϕ-orbit*

of $F \setminus N(\Gamma)$ the restriction of ϕ is either finite order or pseudo-Anosov, and $[F, \phi]$ is a 3-manifold with a JSJ decomposition (whose tori correspond to the ϕ orbits of Γ) into Seifert fibered and hyperbolic pieces.

In the sequel, we will need more precise control over the normal form of ϕ near the boundary of a subsurface on which ϕ is pseudo-Anosov. We say a representative pseudo-Anosov map ϕ on F with boundary is in *standard form* if it satisfies the following two conditions:

(1) Near each boundary circle, two p-pronged measured transverse foliations $(\mathfrak{F}^s, \mu^s)$ and $(\mathfrak{F}^u, \mu^u)$ have the form illustrated here (for the case p=3).



(2) On each ϕ-orbit on ∂F, the restriction of ϕ is periodic.

Proposition 2.15 [Jiang and Guo 1993]. *Each reducible map ϕ as in case* (3) *of Theorem 2.14 can be isotoped into a* standard form; *i.e.*:

(1) *The restriction of ϕ to each pseudo-Anosov orbit of $F \setminus N(\Gamma)$ is in standard form as above.*

(2) *The restriction of ϕ to each periodic orbit of $F \setminus N(\Gamma)$ is periodic.*

This completely fixes the behavior of ϕ on the complement of the regions $N(\Gamma)$. In the sequel we assume that each reducible map ϕ has been isotoped to its standard form in Proposition 2.15. Then for any such ϕ, there is some positive integer l so that ϕ^l is the identity on $\partial(F \setminus N(\Gamma(\phi)))$ and ϕ on $N(\Gamma)$ are Dehn twists along each $\gamma \in \Gamma(\phi)$ relative to $\partial(F \setminus N(\Gamma(\phi)))$.

Definition 2.16. Let ϕ be a reducible map. Say ϕ is *D-type* if it is generated by Dehn twists along components of $\Gamma(\phi)$; say ϕ is *D-type along* $\Gamma(\phi)$ if ϕ restricts to the identity along $\partial N(\Gamma(\phi))$.

Remark 2.17. Note that every ϕ has a power ϕ^l which is D-type along $\Gamma(\phi)$. Moreover, ϕ is a root of D-type, i.e., some power ϕ^l is D-type, if and only if ϕ is periodic on each ϕ-orbit of $F \setminus N(\Gamma)$. Alternatively, every ϕ is either a root of D-type or has pseudo-Anosov ϕ-orbits.

Finally we make the following notational convention. We denote surfaces in general by F, F_i, G and so on, and use $\Sigma_{g,n}$ to denote the surface of genus g with n boundary components. We sometimes abbreviate $\Sigma_{g,0}$ to Σ_g.

Seifert fibered case. Finite order automorphisms are very easy to understand. Suppose (F_1, ϕ_1) and (F_2, ϕ_2) have finite order, so that the manifolds $[F_1, \phi_1]$ and $[F_2, \phi_2]$ are Seifert manifolds with a product geometry. Each $[F_i, \phi_i]$ is finitely covered by a product $F_i \times S^1$. From Lemma 2.12 we can deduce:

Proposition 2.18. *There is exactly one fibered commensurability class of surface bundles of all closed* (*resp. with torus boundary*) *Seifert fibered manifolds whose fiber has negative Euler characteristic. This class contains infinitely many minimal elements.*

Proof. All that needs to be proved is that the class contains infinitely many minimal elements. A key observation is that if $\tilde{\phi}$ is primitive in $\mathrm{MCG}(\tilde{F})$ and has a fixed point near which it acts as a rotation through order p, the same is true of any $\phi \in \mathrm{MCG}(F)$ that it covers. This observation lets us construct infinitely many minimal elements, as follows.

For each genus $g > 1$, let ϕ_g be a maximum-order orientation-preserving periodic map on Σ_g. Then (see [Steiger 1935]) ϕ_g has order $4g+2$ (indeed there is a unique $\mathbb{Z}/(4g+2)\mathbb{Z}$ subgroup of $\mathrm{MCG}(\Sigma_g)$ up to conjugacy) and has exactly one fixed point, one periodic orbit of length 2 and one periodic orbit of length $2g+1$. Clearly (Σ_g, ϕ_g) is primitive, and (Σ_g, ϕ_g) and (Σ_g, ψ) cover each other if and only if $\psi = \phi_g^q$ for q coprime with $4g+2$. Now suppose (Σ_g, ϕ_g) covers (Σ_l, ψ) with $l \neq g$. Of course, we must have $l < g$. On the other hand by the observation above, ψ must have a fixed point near which it acts as a rotation through order $4g+2$, which implies that ψ is a periodic map on Σ_l of order at least $4g+2$, which is impossible. This completes the proof. □

3. Pseudo-Anosov automorphisms

Minimal elements. The most important fact we prove about commensurability of pseudo-Anosov automorphisms — equivalently, of fibered commensurability of hyperbolic fibered pairs — is the existence of finitely many minimal elements in each commensurability class. In fact, working in the orbifold category, the statement is as clean as it could be:

Theorem 3.1. *Every commensurability class of hyperbolic fibered pairs contains a unique* (*orbifold*) *minimal element.*

Remark 3.2. If M is not arithmetic, then the commensurability class of M (in the usual sense) contains a unique minimal element which is some orbifold O. However, if M is arithmetic, no such unique minimal element exists, and the commensurator of $\pi_1(M)$ is dense in PSL(2, $\mathbb{C}$); see [Borel 1981; Margulis 1991].

Remark 3.3. Compare with Proposition 2.18 to see that the hypothesis of "hyperbolic" is essential here (in fact, the hyperbolic world is essentially the only context in which there are unique minimal elements in a commensurability class).

Proof of Theorem 3.1. Let $(M, \mathscr{F})$ be a fibered pair, and after passing to a 2-fold cover if necessary, assume that M fibers over S^1 with fibers the leaves of $\mathscr{F}$. Thus M has the structure of an F-bundle over S^1 with monodromy ϕ, for some compact surface F, and some pseudo-Anosov homeomorphism $\phi : F \to F$. The suspension of the product structure gives a pseudo-Anosov flow X transverse to $\mathscr{F}$, with finitely many closed singular orbits corresponding to the singular points of ϕ. The interior of the manifold M admits a unique complete singular Sol metric for which the leaves of $\mathscr{F}$ are Euclidean surfaces with cone singularities on the singular orbits of X; see [Thurston 1997] or [Fathi et al. 1979] for details.

Pulling back the singular Sol metric on M gives the interior of the universal cover $(\tilde{M}, \tilde{\mathscr{F}})$ the structure of a complete simply connected singular Sol manifold, for which the leaves of $\tilde{\mathscr{F}}$ are singular Euclidean planes, and on which $\pi_1(M)$ acts as a discrete finite covolume group of isometries. Let Λ denote the full group of isometries of $\tilde{M}$ with its singular Sol metric.

Claim: Λ *is itself a lattice, and it preserves the foliation* $\tilde{\mathscr{F}}$.

We show how the theorem follows from this Claim. Since $\pi_1(M) \subset \Lambda$ we have the foliation-preserving covering $p : (M, \mathscr{F}) = (\tilde{M}, \tilde{\mathscr{F}})/\pi_1(M) \to (\tilde{M}, \tilde{\mathscr{F}})/\Lambda$. Since $(M, \mathscr{F})$ is a hyperbolic surface bundle of finite volume, we conclude that $(\tilde{M}, \tilde{\mathscr{F}})/\Lambda$ is an orbifold fiber pair $(O, \mathscr{G})$. Notice that any covering map of fibered pairs $(\tilde{M}, \tilde{\mathscr{F}}) \to (M, \mathscr{F})$ is isotopic to an isometric covering of the interiors in the singular Sol metrics. Then it is easy to see that for any pair $(M', \mathscr{F}')$ commensurable with $(M, \mathscr{F})$ the group $\pi_1(M')$ embeds into Λ in such a way that $(M', \mathscr{F}')$ covers $(O, \mathscr{G})$.

Now we prove the Claim. First, it is evident that Λ preserves the stratification of $\tilde{M}$ into "ordinary" points (those with a neighborhood isometric to an open set in Sol) and singular points (those on the lifts of the singular flowlines of X). Moreover, any isometry between open subsets of Sol must preserve the foliation by Euclidean planes, as can be seen by appealing to the well-known structure of the point stabilizers in Isom(Sol); see [Thurston 1997, Chapter 3], for instance.

Since Λ is equal to the group of isometries of the nonsingular part of $\tilde{M}$, it follows that Λ is a Lie group, by the well-known theorem of Myers and Steenrod [1939]. Hence if Λ is not discrete, it must contain a continuous family of nontrivial isometries. Such isometries can only act on the singular flowlines as translations. Let $\ell(t)$ and $\ell'(t)$ be two such flowlines, parametrized by length in such a way that $\ell(t)$ and $\ell'(t)$ are contained in the same singular Euclidean leaf of $\tilde{M}$, for each t. Assume furthermore that for $|t|$ sufficiently small, the points $\ell(t)$ and $\ell'(t)$ can be joined by a unique (nonsingular) Euclidean geodesic in the singular Euclidean leaf containing them. Then for small t, the length of this Euclidean geodesic as a function of t has the form $\sqrt{e^{2t}x^2+e^{-2t}y^2}$ for fixed x and y; in particular, the length of this Euclidean geodesic is not locally constant, and therefore (since elements of Λ preserve the foliation by singular Euclidean planes) a continuous family of isometries must fix ℓ and ℓ' pointwise. But this implies that $\tilde{M}$ admits no continuous family of nontrivial isometries, and Λ is discrete. Since it contains $\pi_1(M)$, it is therefore a lattice, as claimed. □

Remark 3.4. If F is closed, $\tilde{M}$ with its singular Sol metric and with its hyperbolic metric are quasi-isometric. Consequently if ℓ, ℓ' are two flowlines, the distance function $d(\cdot\,,\cdot)$ is proper on $\ell \times \ell'$ and therefore one obtains another proof that Λ contains no nontrivial continuous family.

Remark 3.5. A fibration of M over a circle is uniquely determined by an element of $H^1(M;\mathbb{Z})$, which is represented by a unique harmonic 1-form α in the hyperbolic metric on M. A cover $(\tilde{M}, \tilde{\mathcal{F}}) \to (M, \mathcal{F})$ pulls back the harmonic 1-form on M to the corresponding harmonic 1-form on $\tilde{M}$ (up to scale), so one can give a slightly different proof of Theorem 3.1 by using the pullback of this 1-form to $\mathbb{H}^3$ and arguing that its set of (projective) symmetries is discrete. Compare with the proof of Theorem 0.1 in [Agol 2006].

The following two corollaries are immediate:

Corollary 3.6. *For any positive constant C, the set of hyperbolic fibered pairs in a commensurability class whose underlying 3-manifold has volume bounded above by C contains only finitely many elements.*

Proof. Such a pair corresponds to a finite index subgroup of the orbifold fundamental group of $(O, \mathcal{G})$ (with notation as in Theorem 3.1) where the index is bounded by $C/\mathrm{vol}(O)$. Since $\pi_1(O)$ is finitely generated, the number of such subgroups is bounded. □

Corollary 3.7. *Suppose M is hyperbolic and fibers over S^1, and rank$(H_1(M)) > 1$. Then M fibers over S^1 in ways representing infinitely many fibered commensurability classes.*

Example 3.8. Suppose (F, ϕ) is pseudo-Anosov. Let c be an essential simple closed curve on F, and let τ_c be a Dehn twist along c. Then the automorphisms $(F, \tau_c^l \circ \phi)$ are hyperbolic for all large l, while the volumes of $[F, \tau_c^l \circ \phi]$ are all bounded by the volume of the cusped manifold $[F, \phi] \setminus (c \times \{0\})$. By Corollary 3.6, there are infinitely many commensurability classes among the $(F, \tau_c^l \circ \phi)$ for large l. Of course, it is easy to see directly in this case that the underlying manifolds fall into infinitely many commensurability classes (in the usual sense); see [Anderson 2002], for instance. We give more substantial examples of incommensurable pseudo-Anosov automorphisms in the next subsection and after.

Remark 3.9. One trivial way to produce a hyperbolic 3-manifold M with many nonisotopic but commensurable fibrations is just to choose a 3-manifold with a large isometry group. We do not know explicit examples of two commensurable fibrations of a single hyperbolic 3-manifold with different genus.

Commensurability invariants. The following is an incomplete list of elementary commensurability invariants for pseudo-Anosov automorphisms:

(1) whether the underlying surface is closed or bounded;

(2) the commensurability class of the underlying 3-manifold of $[F, \phi]$.

(3) the commensurability class of $\log(K)$ where K is the dilatation;

(4) the set of orders of the singular points of the invariant foliations;

For later use we say a few words about (3) and (4). First we make some definitions. For a pseudo-Anosov automorphism (F, ϕ) with a pair of transversely measured singular foliations $\mathfrak{F}_{s,u}$, we use $\lambda(\phi) > 1$ to denote the dilatation of ϕ, and $\delta_n(\phi)$ to denote the number of singularities of degree n, then define $\Delta(\phi)$ to be the (infinite) vector whose coordinates are the $\delta_n(\phi)$.

The first observation to make is that for pseudo-Anosov automorphisms, $\lambda(*)$ is only affected by dynamical coverings, and $\Delta(*)$ is only affected by topological coverings.

Lemma 3.10. *Suppose (F_1, ϕ_1), (F_2, ϕ_2) are two commensurable pseudo-Anosov maps. Then for some $s, s' \in \mathbb{Q}_+$,*

(1) $\log \lambda(\phi_1) = s \log \lambda(\phi_2)$, *and moreover* $\log \lambda(\phi_1) = \log \lambda(\phi_2)$ *if they are topologically commensurable*; *and*

(2) $\Delta(\phi_1) = s' \Delta(\phi_2)$, *and moreover* $\Delta(\phi_1) = \Delta(\phi_2)$ *if they are dynamically commensurable.*

Proof. These facts follow immediately from the definitions (recall Definition 2.5; also, (1) follows from the proof of Proposition 4.11). □

Example 3.11 (bounded–unbounded). Remark 4.3 of [Hironaka 2009] gives an example of a pair of automorphisms $\phi_{(1,3)}$ defined on a genus 2 surface with four boundary components, and $\phi_{(3,4)}$ defined on a closed genus 3 surface with the same dilatation. The commensurability classes of these examples are also distinguished by the orders of the singular points.

Example 3.12. Explicit examples of incommensurable fibrations of the same hyperbolic 3-manifold are straightforward to construct and distinguish by means of Lemma 3.10. For example, in page 4 of [Hironaka 2009], fibrations of the complement of the link 6_2^2 in Rolfsen's tables [1976] are listed, and their singularity sets do not satisfy the commensurability condition in bullet (2) of Lemma 3.10.

Example 3.13. Incommensurable examples may be obtained by branched covers. Start with an Anosov automorphism ϕ of a torus T with dilatation K, and let P be a finite subset of T permuted by ϕ. Let F be obtained as a branched cover of T, branched over P. Then some power of ϕ lifts to an automorphism of F with dilatation a power of K. Different choices of branch orders give rise to incommensurable automorphisms of closed surfaces with the same dilatations, but usually incommensurable singular sets.

One may define a more subtle invariant of commensurability as follows. Let ϕ be a pseudo-Anosov automorphism of F, with measured foliations $\mathfrak{F}_{s,u}$ and projectively invariant transverse measures $\mu_{s,u}$, and singular set S (note that S is finite). For any pair of points p and q (possibly $p = q$) in the singular set, and any homotopy class of paths γ from p to q in the complement $F \setminus S$ we define a number $\ell(\gamma)$ to be the infimum, over all paths γ' from p to q which are homotopic to γ in $F \setminus S$ rel. endpoints, of the product

$$\ell(\gamma) = \inf_{\gamma'} \mu_s(\gamma')\mu_u(\gamma')$$

This number depends on the choice of measures μ_s, μ_u in their projective class, but is well-defined if we normalize the product of measures so that $\int_F d\mu_s d\mu_u = -\chi(F)$.

Definition 3.14. Define the *spectrum* of (F, ϕ) to be the set of numbers $\ell(\gamma)$ as γ varies over nontrivial homotopy classes of paths in $F \setminus S$ as above.

Proposition 3.15. *With the normalization of the product of measures as above, the spectrum is a commensurability invariant. Furthermore, it is strictly positive, and discrete as a subset of* $\mathbb{R}$ (*and is therefore bounded away from zero*).

Proof. By multiplicativity of Euler characteristic, the normalization of the product of measures is compatible under finite covers. Each homotopy class of arcs joining singular points on F lifts to an arc joining singular points in any cover $\tilde{F}$, so the spectrum as defined is a commensurability invariant.

It remains to show that the spectrum is discrete. By the properties of a pseudo-Anosov, we have $\ell(\gamma) = \ell(\phi^i(\gamma))$ for any homotopy class γ and any integer i. To show that the spectrum is discrete, it suffices to show that there are only finitely many ϕ-orbits of homotopy classes γ with $\ell(\gamma) \le C$.

Suppose $K > 1$ is the dilatation of ϕ, and γ' is any path between singular points on F. By the definition of $\mathfrak{F}_{s,u}$, we have $\mu_s(\phi(\gamma')) = K\mu_s(\gamma')$ and $\mu_u(\phi(\gamma')) = K^{-1}\mu_u(\gamma')$. So under the automorphism ϕ, the difference of their logs changes by $2\log K$. It follows that whatever the difference of logs is initially, after a suitable power of ϕ the absolute value of the difference can be taken to be at most $\log(K)$. In other words, there is some integer i so that

$$|\log(\mu_s(\phi^i(\gamma'))) - \log(\mu_u(\phi^i(\gamma')))| \le \log(K).$$

If A and B are positive numbers, a bound on AB and one on $|\log(A) - \log(B)|$ let us bound both A and B. It follows that if $\ell(\gamma) \le C$ then for some i, the homotopy class $\phi^i(\gamma)$ is represented by an arc $\beta = \phi^i(\gamma')$ for which both $\mu_s(\beta)$ and $\mu_u(\beta)$ are bounded, by a constant depending only on C and K. By the discreteness of S, there are only finitely many such relative homotopy classes $\phi^i(\gamma)$, and each of them has a positive ℓ length. So $\ell(\gamma)$ takes only finitely many values in $[0, C]$ (all of them positive). $\square$

Remark 3.16. If Σ is a Riemann surface, any quadratic holomorphic differential α on Σ defines a pair of singular measured foliations, and we can define a spectrum as above for a pair (Σ, α). Multiplying α by a constant also multiplies the spectrum by a constant, so we can normalize to quadratic differentials with $\int_\Sigma |\alpha| = 1$. The set of such pairs (Σ, α) can be identified with the unit cotangent bundle in moduli space. The spectrum (defined as above) is constant on orbits of the Teichmüller flow (see, e.g., [Masur and Tabachnikov 2002] for a definition), and is discrete (by Proposition 3.15) for points on closed orbits of the flow. For general quadratic differentials the spectrum can have accumulation points, or its closure can contain a perfect set, or it can even be dense.

This invariant gives rise to a new way to distinguish commensurability classes of automorphisms.

Example 3.17 (different spectrum). As above, let ϕ be an Anosov automorphism of a torus T (with a flat metric on the torus of total area 1). The set of periodic points is dense, so we can choose two periodic points O, P. The stable and unstable foliations of ϕ give coordinates on T, at least in a neighborhood of O, so that $O = (0, 0)$ and $P = (x, y)$.

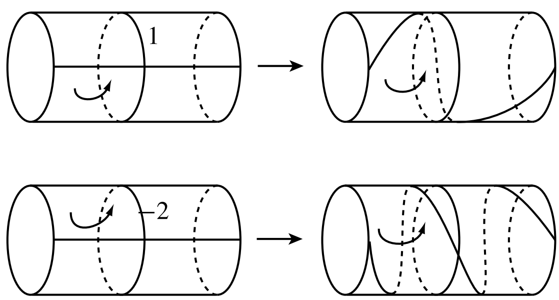
In a suitable cover of T branched over O and P we obtain an automorphism with dilatation a power of K for which the smallest term in the spectrum is at most $|xy|$ times a constant depending only on the combinatorics of the cover. By choosing the periodic point P so that $|xy|$ is sufficiently small, we can ensure that the first term in the spectrum is as close to 0 as we desire, while at the same time fixing the orders of the singular points. By Proposition 3.15, this construction gives rise to infinitely many commensurability classes with commensurable log dilatation and the same combinatorial invariants.

Remark 3.18. Example 3.13 also produces examples of infinitely many (incommensurable) pseudo-Anosov maps with different singular orders but the same spectrum. It is not clear if there exists a pair of pseudo-Anosov maps with incommensurable log dilatations but the same spectrum.

4. Reducible automorphisms

Commensurability invariants of reducible automorphisms. We have assumed that each reducible map is in its standard form as described in Proposition 2.15. We also use the notation from that proposition without comment.

Let A be an oriented annulus A. The mapping class group of A rel. boundary is isomorphic to $\mathbb{Z}$, generated by a positive Dehn twist τ along the core circle. We denote the n-th power of such a Dehn twist by τ_n. Here is an illustration of the cases $n = 1$ and $n = -2$.



Remark 4.1. In this and later figures, the orientation of the surface is indicated by a "cup" shaped arrow, and the numbered circles on the surface indicate the power of a positive Dehn twist (with respect to the given orientation).

For a reducible map ϕ, choose l so that ϕ^l is the identity on $\partial(F \setminus N(\Gamma(\phi)))$. For each component $N(\gamma)$ of $N(\Gamma(\phi))$, where $\gamma \in \Gamma(\phi)$, $N(\gamma)$ has the induced orientation and $\phi^l|\partial N(\gamma)$ is the identity. Then the restriction of ϕ^l to $N(\gamma)$ is the n-th power of a Dehn twist for some integer n. Now define

$$I(\phi^l, \gamma); \quad I(\phi, \gamma) = I(\phi^l, \gamma)/l; \quad a_k(\phi) = \#\{\gamma \in \Gamma(\phi) |\ I(\phi, \gamma) = k\},\ k \in \mathbb{Q}$$

Further, define

$$S(\phi) = \{S \mid S \text{ a component of } F \setminus N(\Gamma(\phi))\}$$

and

$$\Omega(S) = \{\gamma \mid \gamma \text{ a component of } \partial S \setminus \partial F\}.$$

For every $S \in S(\phi)$, define

$$a_{S,k}(\phi) = \#\{\gamma \in \Omega(S) \mid I(\phi,\gamma) = k\}; \quad A(\phi, S) = \left(\sum_{k \in \mathbb{Q}_+} \frac{a_{S,k}(\phi)}{k}, \sum_{k \in \mathbb{Q}_-} \frac{a_{S,k}(\phi)}{-k} \right).$$

The following two numerical invariants are easy to compute:

$$A(\phi) = \frac{1}{2} \sum_{S \in S(\phi)} A(\phi, S) = \left(\sum_{k \in \mathbb{Q}_+} \frac{a_k(\phi)}{k}, \sum_{k \in \mathbb{Q}_-} \frac{a_k(\phi)}{-k} \right),$$

$$\Pi(\phi) = \left\{ \frac{1}{-\chi(S)} A(\phi, S) \;\middle|\; S \in S(\phi) \right\}.$$

We say that two sets of ordered pairs of rational numbers $\{(p_i, q_i)\}$ and $\{(p'_j, q'_j)\}$ are equal *up to a flip*, denoted $\{(p_i, q_i)\} \sim \{(p'_j, q'_j)\}$, if either they are equal, or $\{(p_i, q_i)\} = \{(q'_j, p'_j)\}$. Immediately we have:

Lemma 4.2. *Reversing the orientation of F preserves $A(\phi, S)$, and therefore also $A(\phi)$ and $\Pi(\phi)$, up to a flip.*

We can derive commensurability invariants from $A(\cdot)$ and $\Pi(\cdot)$ as follows:

Theorem 4.3. *Suppose (F_1, ϕ_1), (F_2, ϕ_2) are two reducible maps. If they are commensurable, then for some $s \in \mathbb{Q}_+$,*

$$A(\phi_1) \sim sA(\phi_2) \quad \textit{and} \quad \Pi(\phi_1) \sim s\Pi(\phi_2).$$

We postpone the proof of Theorem 4.3 until page 302.

Remark 4.4. The invariant $\Pi(\cdot)$ is typically better than $A(\cdot)$ at distinguishing commensurability classes (though not always; see Example 4.13). We say that a D-type map is *definite* if it is a product of Dehn twists in the components of $\Gamma(\phi)$ of the same sign. Note that the property of having a power which is definite (along $\Gamma(\phi)$) is a commensurability invariant. The invariant $A(\cdot)$ can distinguish between definite and indefinite maps, but can never distinguish different commensurability classes of definite maps, whereas $\Pi(\cdot)$ can.

Remark 4.5. Both $A(\phi)$ and $\Pi(\phi)$ can be encoded as a polynomial (with fractional exponents), as follows. For any pair of nonnegative rational numbers (p, q), define

$$S(\phi)(p,q) = \left\{ S \in S(\phi) \;\middle|\; \frac{A(\phi, S)}{-\chi(S)} = (p,q) \right\}, \quad \lambda(\phi)_{(p,q)} = \frac{\sum_{S \in S(\phi)(p,q)} \chi(S)}{\chi(F)}.$$

Now define a polynomial pair

$$P(\phi)(x,y) = (P_1(\phi)(x,y), P_2(\phi)(x,y)) = \sum_{(p,q) \in \mathbb{Q}^2} (p,q)\lambda(\phi)_{(p,q)} x^p y^q.$$

One can recover $A(\cdot)$ and $\Pi(\cdot)$ from this polynomial by the formulae

$$\frac{2}{-\chi(F)}A(\phi) = \sum_{(p,q)\in\mathbb{Q}^2}(p,q)\lambda(\phi)_{(p,q)} = P(\phi)(1,1),$$
$$\Pi(\phi) = \{(p,q) \mid \lambda(\phi)_{(p,q)} \neq 0\}.$$

One can show along lines similar to the proof of Theorem 4.3 (in the next subsection) that if two reducible maps (F_1,ϕ_1), (F_2,ϕ_2) are commensurable, then for some $s \in \mathbb{Q}_+$, we have

$$P(\phi_1)(x,y) \sim s\,P(\phi_2)(x^s,y^s).$$

Proof of Theorem 4.3. We need some lemmas, which can be verified immediately from the definitions.

Lemma 4.6. *If ϕ is a reducible map, we have, then for any positive integer k,*

$$I(\phi^k,\gamma) = kI(\phi,\gamma), \quad a_{S,n}(\phi^k) = a_{S,n/k}(\phi), \quad A(\phi^k,S) = \frac{1}{k}A(\phi,S). \tag{4-1}$$

Lemma 4.7. *Suppose two automorphisms ϕ_1 and ϕ_2 on F are isotopic, and two circles γ_1 and γ_2 on F are isotopic. If ϕ_i is D-type along γ_i, $i=1,2$, then $I(\phi_1,\gamma_1) = I(\phi_2,\gamma_2)$.*

Lemma 4.8. $\Pi(\phi)$ *and* $A(\phi)$ *are isotopy invariants.*

Proof of Lemma 4.8. This follows from the definitions, from Lemma 4.7 and from the fact that the reducible system Γ is unique up to isotopy; see Theorem 1 in [Wu 1987], for example. □

Now turning to the proof of Theorem 4.3 proper, suppose (F_1,ϕ_1) and (F_2,ϕ_2) are commensurable. Then there is a surface $\tilde{F}$, automorphisms $\tilde{\phi}_1$ and $\tilde{\phi}_2$ of $\tilde{F}$, and nonzero integers k_1 and k_2, so that $(\tilde{F},\tilde{\phi}_i)$ covers (F_i,ϕ_i) for $i=1,2$, and $\tilde{\phi}_1^{k_1} = \tilde{\phi}_2^{k_2}$ as automorphisms of $\tilde{F}$. Denote the covering $\tilde{F}\to F_i$ by p_i, $i=1,2$. By Lemma 4.2, we may assume that the orientations of $\tilde{F}$, F_1 and F_2 have been chosen so that both p_1, p_2 are orientation-preserving.

Assume that $k_1=k_2=1$ for the moment. By Lemma 4.8, we may assume that $\tilde{\phi}_1 = \tilde{\phi}_2$ as maps in usual sense (rather than in their isotopy class).

Consider the commutative diagram

$$\begin{array}{ccc}
\partial p_1^{-1}(N(\Gamma(\phi_1))) & \xrightarrow{\tilde{\phi}_1^{\,k}|} & \partial p_1^{-1}(N(\Gamma(\phi_1))) \\
{\scriptstyle p_1|}\big\downarrow & & \big\downarrow{\scriptstyle p_1|} \\
\partial N(\Gamma(\phi_1)) & \xrightarrow{\phi_1^k|} & \partial N(\Gamma(\phi_1))
\end{array}$$

where k is chosen so that $\phi_1^k|_{\partial N(\Gamma(\phi_1))} = \mathrm{id}|_{\partial N(\Gamma(\phi_1))}$. It follows that the restriction of $\tilde{\phi_1}^k$ to $\partial p_1^{-1}(N(\Gamma(\phi_1)))$ is a deck transformation of the covering $p_1|$. Since $p_1|$ is a finite covering, by replacing k by a power if necessary, we can assume that $\tilde{\phi_1}^k$ agrees with id on $\partial p_1^{-1}(N(\Gamma(\phi_1)))$ and consequently maps every component of $p_1^{-1}(N(\Gamma(\phi_1)))$ to itself. For such a k, each ϕ_i^k, $\tilde{\phi_i}^k$, $i = 1, 2$ are D-type along their respective reducible systems, where $\Gamma(\tilde{\phi_i}^k) = p_i^{-1}(\Gamma(\phi_i))$.

For each $S_1 \in S(\phi_1)$ and each component $\tilde{S}$ of $p_1^{-1}(S_1)$, there exists a component $S_2 \in S(\phi_2)$, such that $\tilde{S}$ is a component of $p_2^{-1}(S_2)$. Assume $p_i| : \tilde{S} \to S_i$ are l_i-sheeted coverings, for $i = 1, 2$.

Pick a component $\gamma \in \Omega(S_1)$. Suppose that $\{\delta_1, \dots, \delta_t\} = (p_1|\tilde{S})^{-1}(\gamma)$ and that $p_1 : \delta_i \to \gamma$ is a d_i-sheeted covering. Then $\sum_{i=1}^{l} d_i = l_1$.

Under an m-fold covering of annuli, a Dehn twist on the covering annulus projects to the m-th power of a Dehn twist on the image annulus. Consequently $d_i I(\tilde{\phi_1}^k, \delta_i) = I(\phi_1^k, \gamma)$, and by (4-1) we have

$$I(\tilde{\phi_1}^k, \delta_i) = \frac{kI(\phi_1, \gamma)}{d_i} \tag{4-2}$$

and moreover the $I(\tilde{\phi_1}^k, \delta_i)$ all have the same sign as the $I(\phi_1, \gamma)$, $i = 1, \dots, t$ (because p_1 preserves orientation and $k > 0$). Suppose $I(\phi_1, \gamma) \neq 0$. Then by (4-2),

$$\sum_{i=1}^{t} \frac{1}{I(\tilde{\phi_1}^k, \delta_i)} = \sum_{i=1}^{t} \frac{d_i}{kI(\phi_1, \gamma)} = \frac{l_1}{kI(\phi_1, \gamma)} \tag{4-3}$$

Now we sum over circles $\delta \in \Omega(\tilde{S})$ with positive $I(\tilde{\phi}_1, \delta)$:

$$\begin{aligned}
\sum_{l>0} \frac{a_{\tilde{S},l}(\tilde{\phi_1}^k)}{l} &= \sum_{l>0} \frac{\#\{\delta \in \Omega(S_1') | I(\tilde{\phi}^k, \delta) = l\}}{l} \\
&= \sum_{\substack{\delta \in \Omega(\tilde{S}) \\ I(\tilde{\phi}_1, \delta) > 0}} \frac{1}{I(\tilde{\phi_1}^k, \delta)} = \sum_{\substack{\gamma_i \in \Omega(S_1) \\ I(\phi_1, \gamma_i) > 0}} \sum_{\delta \in (p_1|\tilde{S})^{-1}(\gamma_i)} \frac{1}{I(\tilde{\phi_1}^k, \delta)} \\
&= \frac{l_1}{k} \sum_{\substack{\gamma_i \in \Omega(S_1) \\ I(\phi_1, \gamma_i) > 0}} \frac{1}{I(\phi_1, \gamma_i)} = \frac{l_1}{k} \sum_{l>0} \frac{a_{S_1,l}(\phi_1)}{l},
\end{aligned}$$

where the penultimate equality follows from (4-3).

By a similar computation, we have

$$\sum_{l<0} \frac{a_{\tilde{S},l}(\tilde{\phi_1}^k)}{l} = \frac{l_1}{k} \sum_{l<0} \frac{a_{S_1,l}(\phi_1)}{l},$$

and therefore

$$A(\tilde{\phi_i}^k, \tilde{S}) = \frac{l_i}{k} A(\phi_i, S_i), \quad i = 1, 2. \tag{4-4}$$

By (4-1) we have

$$A(\tilde{\phi_1}, \tilde{S}) = kA(\tilde{\phi_1}^k, \tilde{S}) = A(\tilde{\phi_2}, S_2'). \tag{4-5}$$

Since $l_i = \chi(\tilde{S})/\chi(S_i)$, by (4-4) and (4-5), we get

$$\frac{A(\phi_1, S_1)}{-\chi(S_1)} = \frac{A(\tilde{\phi_1}, \tilde{S})}{-\chi(\tilde{S})} = \frac{A(\phi_2, S_2)}{-\chi(S_2)}. \tag{4-6}$$

From the definition of $\Pi(\,\cdot\,)$ we have $\Pi(\phi_2) \subset \Pi(\phi_1)$. By symmetry we have $\Pi(\phi_2) = \Pi(\phi_1)$. Summing over all Γ in the argument above in place of $\Omega(S_1)$, we get similarly

$$\frac{A(\phi_1)}{\chi(F_1)} = \frac{A(\phi_2)}{\chi(F_2)}.$$

From (4-1) we have $\Pi(\phi^k) = \Pi(\phi)/k$ and $A(\phi^k) = A(\phi)/k$ and the proof is complete. □

From the proof above immediately we have:

Corollary 4.9. *If (F_1, ϕ_1) and (F_2, ϕ_2) are topologically commensurable, then*

$$\frac{A(\phi_1)}{\chi(F_1)} \sim \frac{A(\phi_2)}{\chi(F_2)} \quad \textit{and} \quad \Pi(\phi_1) \sim \Pi(\phi_2).$$

Remark 4.10. We remind the reader that our invariants are defined for all reducible maps (and not just D-type examples and their roots). When reducible maps are not the roots of the D-type maps, then they have pseudo-Anosov orbits, and we can combine the invariants defined in Sections 3 and 4. For example, see the proposition below and Example 4.18.

Proposition 4.11. *Suppose (F_1, ϕ_1), (F_2, ϕ_2) are two commensurable reducible maps. Then for some $s \in \mathbb{Q}_+$,*

$$\log\lambda(\phi_1) = s \log\lambda(\phi_2) \quad \textit{and} \quad \Pi(\phi_1) \sim s^{-1}\Pi(\phi_2).$$

Here we think of $\lambda(\phi)$ for a reducible map ϕ as a (possibly empty) *set* of dilatations of the set of restrictions of ϕ to its pseudo-Anosov orbits.

Proof. From the definition of commensurability, there are positive integers k_1 and k_2 such that $(F_1, \phi_1^{k_1})$ and $(F_2, \phi_2^{k_2})$ are topologically commensurable, both covered by $(\tilde{F}, \tilde{\phi})$. Evidently we have $\lambda(\phi_1^{k_1}) = \lambda(\tilde{\phi}) = \lambda(\phi_2^{k_2})$, and therefore $k_1 \log\lambda(\phi_1) = \log\lambda(\phi_1^{k_1}) = \log\lambda(\phi_2^{k_2}) = k_2 \log\lambda(\phi_2)$ and then

$$\log\lambda(\phi_1) = \frac{k_2}{k_1}\log\lambda(\phi_2).$$

On the other hand, by Corollary 4.9 and (4-1), we have

$$\frac{\Pi(\phi_1)}{k_1} = \Pi(\phi_1^{k_1}) \sim \Pi(\phi_2^{k_2}) = \frac{\Pi(\phi_1)}{k_2}$$

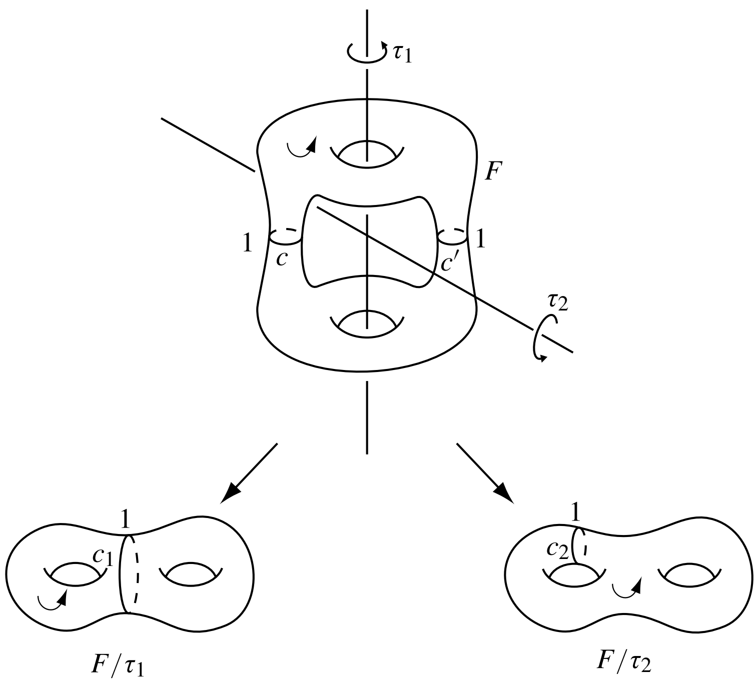
and therefore

$$\Pi(\phi_1) \sim \frac{k_1}{k_2}\Pi(\phi_1).$$

The proposition is proved by setting $s = k_2/k_1$. □

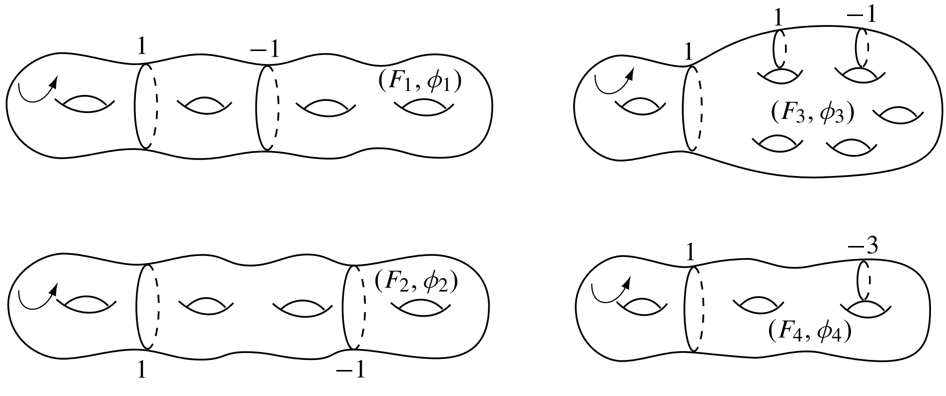
Examples of reducible automorphisms. In this section we give several examples, which illuminate the meaning of the invariants defined above. A D-type map on an oriented F can be indicated pictorially by assigning integers to disjoint essential simple closed curves on a surface; we use this convention in what follows.

Example 4.12. Dehn twists in separating and nonseparating curves (on the same surface) are commensurable. In the figure below, let $\tilde{\phi}$ be a D-type automorphism on a surface F of genus 3 generated by full Dehn twists on circles c and c' as indicated in the figure.



Then $\tilde{\phi}$ is invariant under both π-rotations along τ_1 and τ_2. Hence $\tilde{\phi}$ induces ϕ_i on F/τ_i, where ϕ_i is the Dehn twist along the circle c_i. Since c_1 is separating while c_2 not, ϕ_1 and ϕ_2 are not conjugate. But from the construction they are commensurable.

Example 4.13. This example show that $\Pi(\phi)$ is not always finer than $A(\phi)$. Four automorphisms are depicted here:



By computing $A(\phi)$ and $\Pi(\phi)$, it can be seen that no pair of them are commensurable. Notice that on one hand $A(\phi_1)=A(\phi_2)=(1,1)$ and $\{(1,0),(\frac{1}{2},\frac{1}{2}),(0,\frac{1}{3})\}=\Pi(\phi_1)\neq\Pi(\phi_2)=\{(1,0),(\frac{1}{4},\frac{1}{4}),(0,1)\}$, and on the other hand $(2,1)=A(\phi_3)\neq A(\phi_4)=(1,\frac{1}{3})$ and $\Pi(\phi_3)=\Pi(\phi_4)=\{(1,0),(\frac{1}{3},\frac{2}{9})\}$.

Example 4.14 (minimal elements). Let ϕ_g be a orientation-preserving periodic map on Σ_g of order $4g+2$ which rotates by $\pi/(2g+1)$ around its unique fixed point x_g (see the proof of Proposition 2.18). Remove a ϕ_g-invariant disc at x_g from Σ_g to get $\Sigma_{g,1}$. Connect $\Sigma_{2,1}$ and $\Sigma_{3,1}$ along their boundaries via an annulus A to form a closed surface Σ_5 and define ϕ on F_5 by $\phi|\Sigma_{2,1}=\phi_2|\Sigma_{2,1}$ and $\phi|\Sigma_{3,1}=\phi_3^{-1}|\Sigma_{3,1}$, and then extend to A by a continuous family of rotations through angles from $\frac{\pi}{5}$ to $\frac{\pi}{7}$. The difference in speeds on the boundary components is $\frac{2\pi}{35}$, and it follows that ϕ^{35} is a Dehn twist D_c. By the uniqueness of the reducible system and the argument similar in the proof of Proposition 2.18, one can verify (Σ_5,ϕ) is a minimal element. One can construct infinitely many minimal elements in such a way.

Remark 4.15. One can verify that 35 is the largest order of a root of a Dehn twists on Σ_5. It is amazing that the maximal order of roots of Dehn twist along *nonseparating* curves, which is 11 on Σ_5 (and in general is $2g+1$ in Σ_g), was determined only very recently by several papers; see [Margalit and Schleimer 2009; McCullough and Rajeevsarathy 2009; Monden 2009].

Example 4.16. This example will be used in Section 5. Σ_{kn+1} can be presented as the union of $\Sigma_{1,n}$ and n copies of $\Sigma_{k,1}$ in a in symmetric way so that there is an action $\tau_{n,k}$ of order n which acts freely on the triple $\left(\Sigma_{kn+1},\Sigma_{1,n},\bigcup_1^n\Sigma_{k,1}\right)$.

Let D_c be the positive Dehn twist along one component c of $\partial\Sigma_{1,n}$ and let $\phi_{n,k}$ be the composition of $D_c\circ\tau_{n,k}$. Then one can verify that $D_{n,k}=\phi_{n,k}^n$ is D-type, and is given by the product of a positive Dehn twist along each component of $\partial\Sigma_{1,n}$. For fixed k, the automorphisms $(\Sigma_{kn+1,0}, D_{n,k})$ and $(\Sigma_{km+1}, D_{m,k})$ have a common cover $(\Sigma_{kmn+1}, D_{mn,k})$. Therefore for fixed k, $(\Sigma_{kn+1}, \phi_{n,k})$ are in the same commensurability class for all n.

On the other hand one can verify by inspection that $\Pi(D_{n,k}) = \{(1, 0), (1/(2k-1), 0)\}$. So $(\Sigma_{kn+1}, D_{n,k})$ and $(\Sigma_{k'm+1}, D_{m,k'})$ are not commensurable for $k \neq k'$ by Theorem 4.3.

Example 4.17. Each D-type map (F, ϕ) is commensurable with a D-type map (F', ψ) so that the Dehn twist on each $\gamma \in \Gamma(\psi)$ is a single positive or negative Dehn twist. We can argue as below:

For simplicity, assume F is closed, $S(\phi) = \{S_i, i = 1, \dots, k\}$, and set

$$d_\gamma = |I(\phi, \gamma)|.$$

By replacing ϕ by a power if necessary, we may assume that d_γ is an integer > 1 for each $\gamma \in \Gamma(\phi)$. Then for each i there is a covering $q_i : \tilde{S}_i \to S_i$ such that $q_i| : \tilde{\gamma} \to \gamma$ is of degree d_γ for each component $\gamma \in \partial S$ and each component $\tilde{\gamma}$ in $q_i^{-1}(\gamma)$. One quick way to see this is to attach an orbifold disk D_γ of index d_γ to each $\gamma \in \partial S_i$. The result is 2-dimensional orbifold which is good, since $\chi(S_i) < 0$ and each $d_\gamma > 1$. This orbifold has a manifold cover (see [Thurston 1979, Chapter 13]), and the restriction to S_i gives the required covering $q_i : \tilde{S}_i \to S_i$.

If P is a planar surface of negative Euler characteristic, then for every $n \geq 2$ coprime with the number of components of ∂P, there is a cover $\hat{P} \to P$ of degree n, which restricts to a cover of degree n on each boundary component of $\hat{P}$, and such that $\hat{P}$ is nonplanar. Moreover, every nonplanar surface with negative Euler characteristic has a covering of any given degree which is a covering of degree 1 on each boundary component. So after replacing ϕ by ϕ^n, we can find covers

$$\hat{q}_i : \hat{S}_i \to \tilde{S}_i$$

and a covering of degree $n \prod_{k \neq i} \deg(q_k)$ so that the restriction on each component of $\partial\hat{S}_i$ is a covering of degree exactly n. The coverings $p_i = q_i \circ \hat{q}_i : \hat{S}_i \to S_i$ match compatibly to produce a covering $p : \tilde{F} = \cup\hat{S}_i \to F$ such that $p| : \tilde{\gamma} \to \gamma$ is of degree nd_γ for each $\gamma \in \Gamma(\phi)$ and each component $\tilde{\gamma}$ in $p^{-1}(\gamma)$. Define a D-type map $\tilde{\phi}$ on $\tilde{F}$ with $I(\tilde{\phi}, \tilde{\gamma}) = 1$ if $I(\phi, \gamma) > 0$, and $I(\tilde{\phi}, \tilde{\gamma}) = -1$ otherwise, then $\tilde{\phi}$ covers ϕ (see the paragraph before (4-2) in the proof of Theorem 4.3.)

Now we give an application of Proposition 4.11 to reducible maps which are not roots of D-type maps.

Example 4.18. Let F be a closed oriented surface of genus 2, and c a nonseparating circle in F. Let ϕ be any pseudo Anosov map on $F \setminus c$ with dilatation

$\lambda(\phi) = K$ and twist angle $2\pi r$ near c, $r \in \mathbb{Q}$, and let τ_c be a positive Dehn twist along c. Then:

(1) $\tau^{k_1} \circ \phi$ and $\tau^{k_2} \circ \phi$ are commensurable if and only if $k_1 = k_2$.

(2) $\tau \circ \phi^{k_1}$ and $\tau \circ \phi^{k_2}$ are commensurable if and only if $k_1 = k_2$.

The proofs of (1) and (2) are similar; we only give a proof of (1). Note that

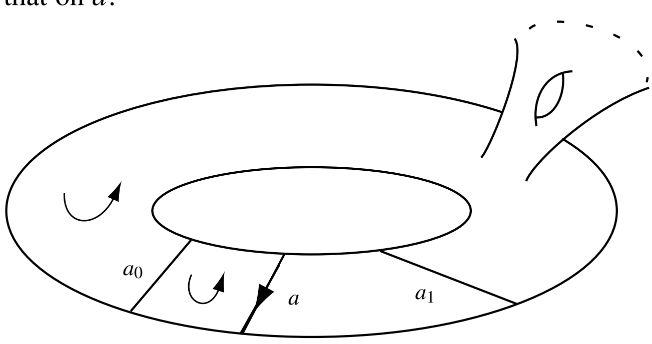
$$\Pi(\tau^k \circ \phi) = \Big(\frac{1}{(k-r)}, 0\Big) \quad \text{and} \quad \lambda(\tau^k \circ \phi) = \lambda(\phi) = K > 1,$$

where r and K depend only on ϕ. If $\tau^{k_1} \circ \phi$ and $\tau^{k_2} \circ \phi$ are commensurable, by Proposition 4.11 and the fact we are considering the automorphism in the same oriented surface F, we should have $\log K = s \log K$ and $1/(k_1 - r) = s^{-1}/(k_2 - r)$ for some $s \in \mathbb{Q}_+$. The first equality implies that $s = 1$, and the second implies $k_1 = k_2$.

5. Commensurable and incommensurable bundles in graph manifolds

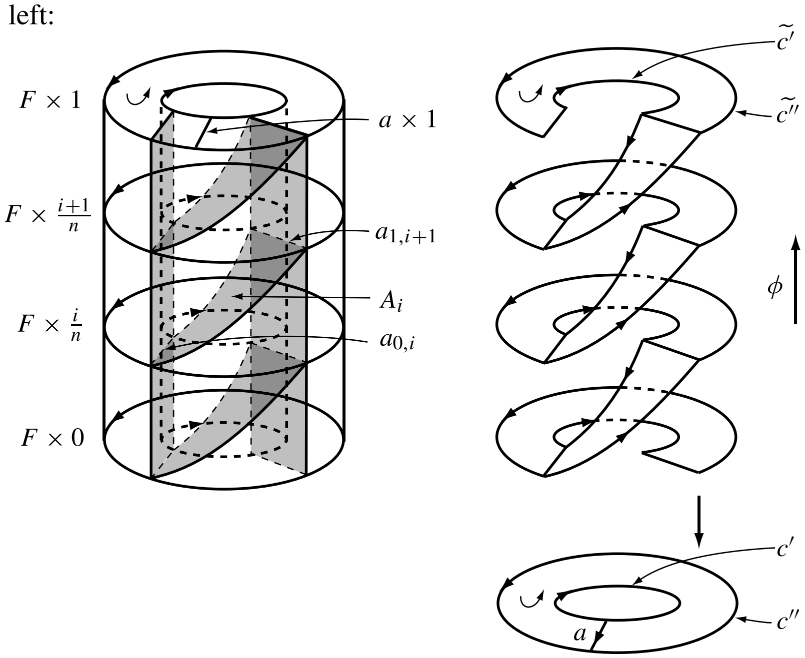
In this section we give two more complicated examples. The first (Example 5.3) is an example of a graph manifold that is the total space of infinitely many incommensurable fibrations, and at the same time fibers in infinitely many ways in the *same* commensurability class. The second (Example 5.5) is an example of a graph manifold that is the total space of infinitely many incommensurable fibrations, including two incommensurable fibrations with the same genus. Both examples depend on a construction that we turn to now.

Primary construction. Let F be a compact oriented surface with the induced orientation on ∂F. Let a be an essential oriented arc on F connecting two different components of ∂F. Let a_0 and a_1 be the two components of the quadrilateral $\partial N(a) \setminus \partial F$ such that the direction on a_0 induced from the orientation on $\partial N(a)$ is parallel to that on a:



Then in $F \times [0, 1]$, the surface $F \times \{\frac{i}{n}\}$ intersects the quadrilateral $a_j \times [0, 1]$ in the arc $a_{j,i} = a_j \times \{\frac{i}{n}\}$ for each integer $n \geq 2$, where $j = 0, 1$ and $i = 0, 1, \dots, n$.

Let $A_1, \ldots, A_n$ be n pairwise disjoint quadrilaterals that are properly embedded in $N(a) \times [0, 1]$ so that A_i is a stair connecting $a_{0,i}$ and $a_{1,i+1}$, as shows in the figure on the left:



Let $F_i = (F \times \frac{i}{n}) \setminus (N(a) \times [0, 1])$ and build a surface $R(a, n) = \bigcup_{i=0}^{n} F_i \cup \bigcup_{l=1}^{n} A_l$ in $F \times [0, 1]$; see right diagram above. A similar surface $R(\alpha, n)$ in $F \times [0, 1]$ can be constructed if we replace a by a disjoint union of essential arcs α on F.

We call the quotient of $R(\alpha, n)$ in $F \times S^1 = [F, \mathrm{id}]$ the *n-floor staircase* along α in $F \times S^1$, or just *n-floor along* α for short, and denote it as $F(\alpha, n)$. Note that the surface $F(\alpha, n)$ is transverse to the S^1 fibers. If α is empty, then $F(\varnothing, n)$ is just n disjoint copies of F in $F \times S^1$.

Let S^1 have the orientation induced from $[0, 1]$. Then both $F \times S^1$ and $\partial F \times S^1$ are oriented. For each component $c \in \partial F$, the torus $c \times S^1$ has product coordinates (c, t). The proof of the following lemma is a routine verification:

Lemma 5.1. *Let $p : F \times S^1 \to F$ be the projection. Suppose that $\alpha \cap c \leq 1$ for each component $c \in \partial F$. Then*:

(1) *$p : F(\alpha, n) \to F$ is a cyclic covering of degree n. Moreover $F(\alpha, n)$ is a surface of genus $1 - k + n(k - 1 + g)$ with $n(\#\partial F - 2k) + 2k$ boundary components*, *where $k = \#\alpha$.*

(2) *$p^{-1}(c)$ is either connected or has n components for each component c of ∂F, and $p^{-1}(c)$ is connected if and only if $\alpha \cap c \neq \varnothing$. Moreover suppose a is an arc in α with tail in c' and head in c'', then $\tilde{c}' = p^{-1}(c')$ has slope $(n, -1)$ and $\tilde{c}'' = p^{-1}(c'')$ has slope $(n, 1)$.*

(3) *Let $\tilde{\tau}$ be the $2\pi/n$-rotation of $F \times S^1$ along the oriented S^1 factor*, *and let $\tilde{c}'$ and $\tilde{c}''$ be as in* (2). *Then τ, the restriction $\tilde{\tau}$ on $F(\alpha, n)$ is a generator of the deck group of the covering in* (1), *which rotates $\tilde{c}'$ and $\tilde{c}''$ through $2\pi/n$ in negative and positive directions respectively*; *see right diagram on the previous page.*

(4) $F \times S^1 = [F, id] = [F(\alpha, n), \tau]$, *and*

$$p_{\alpha,n} : F(\alpha, n) \times S^1 = [F(\alpha, n), \tau^n] \to F \times S^1 = [F, id]$$

is a cyclic covering of degree n.

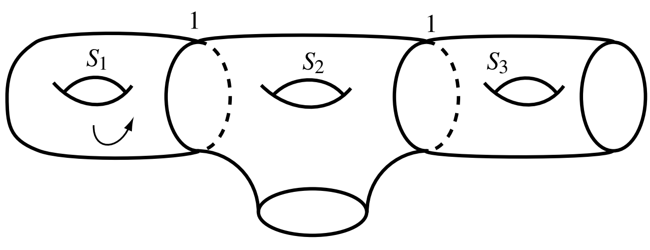
Remark 5.2. We can perform a similar construction for a nonseparating circle γ in F, in which case the description of the boundary is much simpler: each component of ∂F gives rise to precisely n copies of $\partial F(\gamma, n)$.

Example 5.3. We describe a graph manifold that

(1) admits fibrations representing infinitely many fibered commensurability classes, and

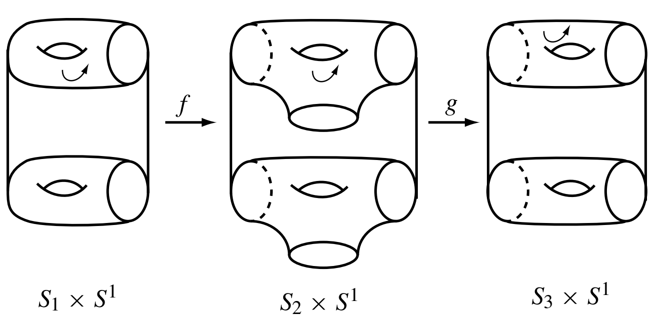
(2) admits infinitely many fibrations representing the *same* fibered commensurability class.

First take $M = [F_1, \phi_1]$, where the oriented surface F_1 and the monodromy ϕ_1 are as shown here:



Note that M has two boundary components and ϕ is D-type and definite.

Another view of M is this:

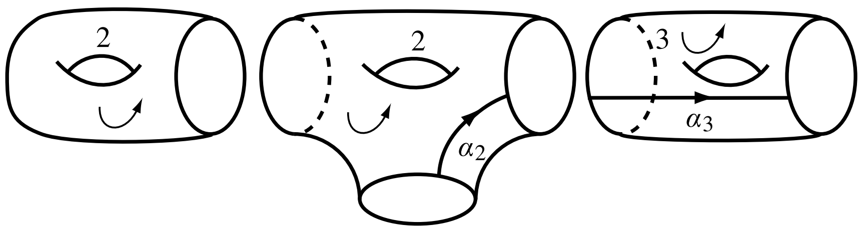


Here every component is of the form $S_i \times S^1$ (depicted in the figure as an $S_i \times I$) for $i = 1, 2, 3$, and two pairs of boundary tori are identified by maps f and g expressed in terms of coordinates by the maps

$$f(1,0) = (-1,0), \quad f(0,1) = (-1,1);$$
$$g(1,0) = (-1,0), \quad g(0,1) = (-1,1).$$

Recall that this notation means that each $(1, 0)$ denotes the homotopy class of some component of some ∂S_i, and each $(0, 1)$ denotes an $S^1 \times *$.

Now we construct another surface fibration of the same underlying manifold $M = [F_2, \phi_2]$ as follows. Pick oriented arcs $\alpha_i \in S_i$, $i = 2, 3$ as follows:



$S_1(\phi, 2)$ $S_2(\alpha_2, 2)$ $S_3(\alpha_3, 3)$

Then construct

$$S_1' = S_1(\varnothing, 2), \quad S_2' = S_2(\alpha_2, 2), \quad S_3' = S_3(\alpha_3, 3)$$

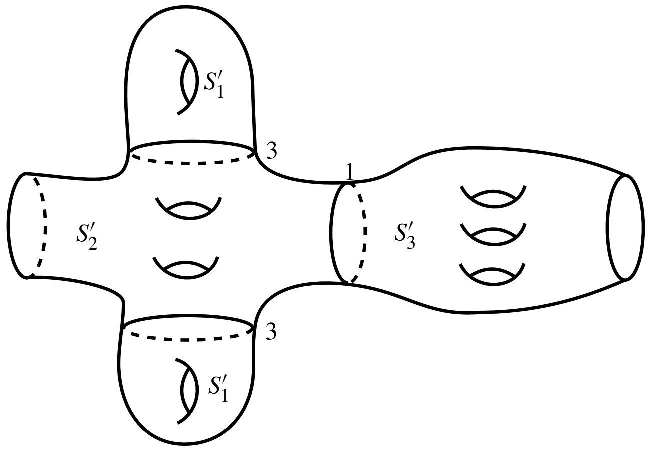
in $S_i \times S^1$, $i = 1, 2, 3$.

By Lemma 5.1(1), it is easy to see that S_1' is two copies of S_1, that S_2' is a surface of genus 2 with 4 boundary components, and that S_3' is a surface of genus 3 with 2 boundary components. By Lemma 5.1(2), we see that $\tilde{c}_2'$ is of slope $(2, 1)$ in $c_2' \times S^1$, and $\tilde{c}_3''$ is of slope $(-3, 1)$ in $c_3'' \times S^1$,

Since g sends $(2, 1)$ to $(-3, 1)$, the maps f and g match S_1', S_2' and S_3' together to produce a new surface F_2 in M. Let τ_i be the generator of the (cyclic) deck group for the covering $p_i : S_i' \to S_i$ given by Lemma 5.1(3). Then τ_1, τ_2, τ_3 have periods 2, 2, 3 respectively. Now the new surface bundle structures $[S_i, \tau_i]$ in $S_i \times S^1$ given by Lemma 5.1(4), $i = 1, 2, 3$, match to produce a new surface bundle structure of M, which we denote by $[F_2, \phi_2]$.

The monodromy map ϕ_2 is a virtual D-type automorphism whose restriction on each S_i' is τ_i. Hence ϕ_2 permutes the two copies of S_1 in F_2. Moreover under this permutation, each copy also undergoes a half-twist relative to S_2'. By Lemma 5.1(3), τ_2 rotates $\tilde{c}_2''$ by π and τ_3 rotate $\tilde{c}_3'$ by $-\frac{2}{3}\pi$ respective along the directions shown in the first figure on page 310. So the relative twist at $S_2' \cap S_3'$ is

$\pi - \frac{2\pi}{3} = \frac{1}{3}\pi$. Now ϕ_2^6 is a D-type automorphism, as shown here:



A direct computation gives

$$\Pi(\phi_1) = \{(1,0), (\tfrac{2}{3}, 0), (\tfrac{1}{2}, 0)\} \quad \text{and} \quad \Pi(\phi_2) = \{(2,0), (\tfrac{5}{3}, 0), (1,0)\}$$

Consequently there is no $s \in \mathbb{Q}$ so that $\Pi(\phi_1) \sim s\Pi(\phi_2)$. By Theorem 4.3, (F_1, ϕ_1) and (F_2, ϕ_2) are not commensurable.

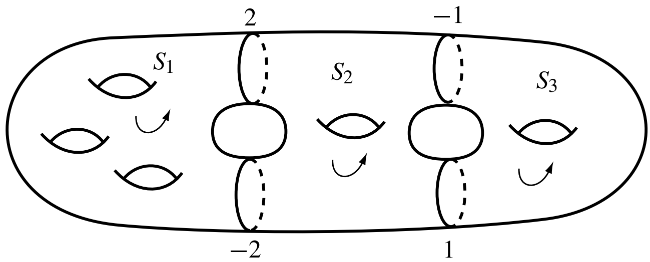
If we perform a similar construction starting from $S_1(\varnothing, n)$, $S_2(\alpha_2, n)$, and $S_3(\alpha_3, n+1)$ in $S_i \times S^1$, $i = 1, 2, 3$, we will get a surface bundle structure $[F_n, \phi_n]$ on M, where ϕ_n is a virtual D-type automorphism and $\phi_n^{n(n+1)}$ is a D-type automorphism, and $\Pi(\phi_n) = \{(n, 0), (\frac{2n+1}{3}, 0), (\frac{n}{2}, 0)\}$. So for any positive integers $i \neq j$, the automorphisms (F_i, ϕ_i), (F_j, ϕ_j) are not commensurable. We have verified that M fibers in infinitely many incommensurable ways.

On the other hand if we start from $S_1(\gamma, n)$, $S_2(\varnothing, n)$ and $S_3(\varnothing, n)$, where γ is a nonseparating circle in S_1, then by Remark 5.2 and the argument above, we can produce a fibration of M with monodromy $(\Sigma_{2n+1,2n}, \phi_{2,n})$, where we adapt the notations in Example 4.16, and use $\Sigma_{2,3} = S_2 \cup S_3$ in place of $\Sigma_{2,1}$. As observed in Example 4.16, the automorphisms $(\Sigma_{2n+1,2n}, \phi_{2,n})$ are commensurable for all n. So M admits infinitely many distinct but commensurable fibrations, as claimed.

Remark 5.4. One can modify the construction in Example 5.3 to a more general setting where the arc connecting two boundary components of F passes through the cores of more than one Dehn twist. For simplicity, consider a D-type map which is either a single positive or negative Dehn twist on each $\gamma \in \Gamma(\phi)$ (compare with Example 4.17). Then one always gets infinitely many fibered commensurability classes unless the $\chi(S_i)$ satisfy a certain linear equation so that the invariants in Section 4 fail to distinguish them, where S_i's are pieces of $F \setminus \Gamma(\phi)$ meeting the arc.

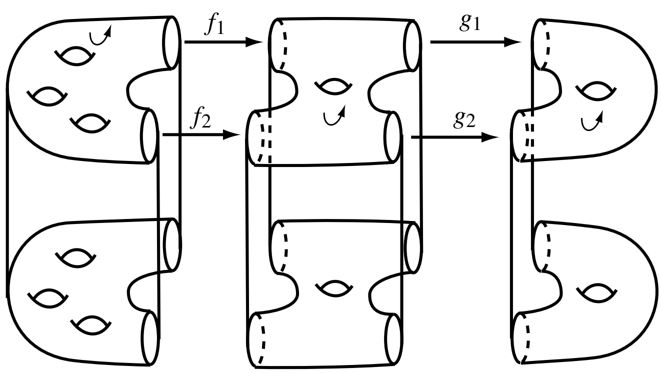
Example 5.5. We now give an example of a *closed* graph manifold which fibers in infinitely many incommensurable ways, including two incommensurable fibrations with fibers of the same genus.

Let $M = [F, \phi]$ be the graph manifold with ϕ as indicated here:



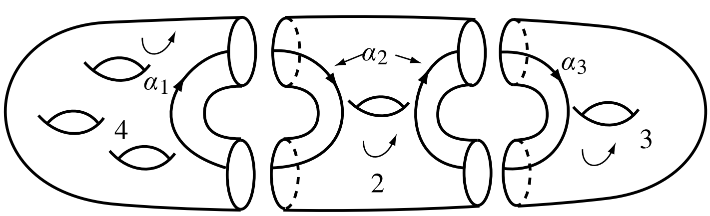
Our discussion of the bottom figure on page 310 applies mutatis mutandis in this case, leading to the following diagram, with gluings given by

$$f_1(1,0)=(-1,0),\quad f_1(0,1)=(2,1);\qquad f_2(1,0)=(-1,0),\quad f_2(0,1)=(-2,1);$$
$$g_1(1,0)=(-1,0),\quad g_1(0,1)=(-1,1);\qquad g_2(1,0)=(-1,0),\quad g_2(0,1)=(1,1).$$



$S_1 \times S^1$ $\quad$ $S_2 \times S^1$ $\quad$ $S_3 \times S^1$

First we construct infinitely many commensurability classes of fibrations of M. Pick oriented arcs $\alpha_i \in S_i$, $i = 1, 2, 3$ as follows:

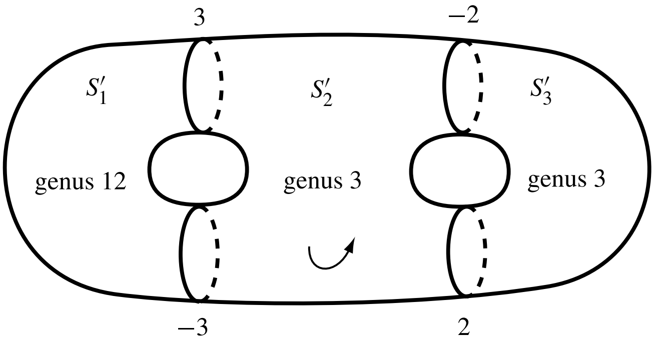


$S_1(\alpha_1, 4)$ $\quad$ $S_2(\alpha_2, 2)$ $\quad$ $S_3(\alpha_3, 3)$

Construct $S_1' = S_1(\alpha_1, 4)$, $S_2' = S_2(\alpha_2, 2)$, $S_3' = S_3(\alpha_3, 3)$ in $S_i \times S^1$, $i = 1, 2, 3$, respectively. Then f_i and g_i, $i = 1, 2$ paste the boundary of S_i' together to produce another bundle structure on M; i.e., we have $M = [\Sigma_{20}, \phi_2]$, where ϕ_2^{12} is a D-type automorphism on the surface of genus 20. We can check that $(\Sigma_{20}, \phi_2^{12})$ has invariant

$$\Pi(\phi_2) = \{(\tfrac{1}{6}, \tfrac{1}{6}), (\tfrac{5}{4}, \tfrac{5}{4}), (1, 1)\}$$

and is as follows:



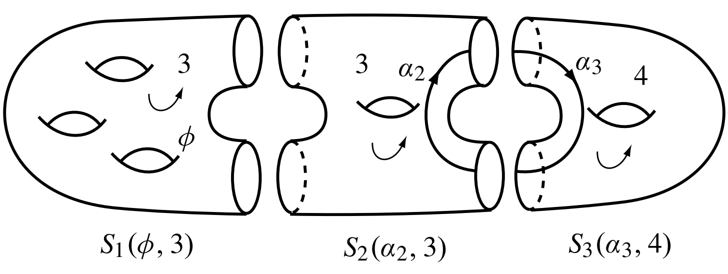
We can perform a similar construction starting from $S_1(\alpha_1, n+2)$, $S_2(\alpha_2, n)$ and $S_3(\alpha_3, n+1)$ in $S_i \times S^1$, $i = 1, 2, 3$, and obtain a surface bundle structure $[\Sigma_{6n+8}, \phi_n]$ on M, where $\phi_n^{n(n+1)(n+2)}$ is a D-type automorphism of a surface of genus $6n+8$ and

$$\Pi(\phi_n) = \{(\tfrac{n}{12}, \tfrac{n}{12}), (\tfrac{3n+4}{8}, \tfrac{3n+4}{8}), (\tfrac{n}{2}, \tfrac{n}{2})\}.$$

So for any positive integers $i \neq j$, (Σ_{6i+8}, ϕ_i), (Σ_{6j+8}, ϕ_j) are incommensurable.

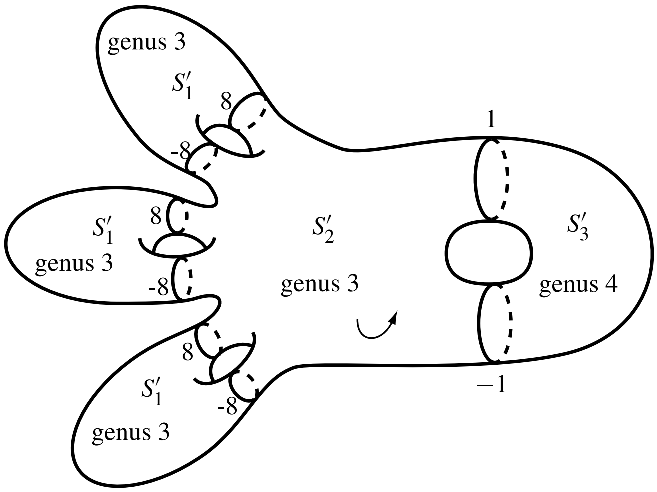
Now we construct another surface bundle structure $[\Sigma_{20}, \psi]$ on M, which is not commensurable with (Σ_{20}, ϕ_2), where ϕ_2 is the automorphism above.

Pick oriented arcs $\alpha_i \in S_i$, $i = 1, 2, 3$ as follows:



and construct $S_1' = S_1(\varnothing, 3)$, $S_2' = S_2(\alpha_2, 3)$, $S_3' = S_3(\alpha_3, 4)$ in $S_i \times S^1$, $i = 1, 2, 3$, respectively. Then f_i and g_i, $i = 1, 2$ glue the boundary of S_i' together to provide M another structure of surface bundle: $M = [\Sigma_{20}, \psi]$, where ψ^{12} is a D-type automorphism on Σ_{20} of genus 20. We can check that (Σ_{20}, ψ^{12}) has invariants

$\Pi(\psi) = \{(\frac{1}{4}, \frac{1}{4}), (\frac{11}{8}, \frac{11}{8}), (\frac{3}{2}, \frac{3}{2})\}$ and is as follows:



By Theorem 4.3 we deduce that (F_2, ψ) and (F_2, ϕ_2) are not commensurable, as claimed.

Acknowledgments

Calegari was partially supported by NSF grant DMS 0707130. Wang was supported by grant No. 10631060 of NSF of China and by Caltech Mathematics Department as a short term scholar. The content of this paper benefited from conversations with Juan Souto. We would like to thank the referee for some helpful suggestions.

References

[Agol 2006] I. Agol, "Virtual betti numbers of symmetric spaces", 2006. arXiv math/0611828

[Anderson 2002] J. W. Anderson, "Incommensurability criteria for Kleinian groups", *Proc. Amer. Math. Soc.* **130**:1 (2002), 253–258. MR 2002f:57031 Zbl 0983.57014

[Behrstock and Neumann 2010] J. Behrstock and W. Neumann, "Quasi-isometric classification of non-geometric 3-manifold groups", 2010. To appear in *J. Reine Angew. Math.* arXiv 1001.0212

[Borel 1981] A. Borel, "Commensurability classes and volumes of hyperbolic 3-manifolds", *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **8**:1 (1981), 1–33. MR 82j:22008 Zbl 0473.57003

[Fathi et al. 1979] A. Fathi, F. Laudenbach, and V. Poenaru, *Travaux de Thurston sur les surfaces*, Astérisque **66–67**, Société Mathématique de France, Paris, 1979. MR 82m:57003 Zbl 1024.01017

[Hironaka 2009] E. Hironaka, "Small dilatation pseudo-Anosov mapping classes coming from the simplest hyperbolic braid", preprint, 2009. arXiv 0909.4517

[Jiang and Guo 1993] B. J. Jiang and J. H. Guo, "Fixed points of surface diffeomorphisms", *Pacific J. Math.* **160**:1 (1993), 67–89. MR 94d:57037 Zbl 0829.55001

[Macbeath 1983] A. M. Macbeath, “Commensurability of co-compact three-dimensional hyperbolic groups”, *Duke Math. J.* **50**:4 (1983), 1245–1253. MR 85f:22013 Zbl 0588.22009

[Margalit and Schleimer 2009] D. Margalit and S. Schleimer, “Dehn twists have roots”, *Geom. Topol.* **13**:3 (2009), 1495–1497. MR 2010d:57019 Zbl 1203.57007

[Margulis 1991] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Math. **17**, Springer, Berlin, 1991. MR 92h:22021 Zbl 0732.22008

[Massey 1974] W. S. Massey, “Finite covering spaces of 2-manifolds with boundary”, *Duke Math. J.* **41** (1974), 875–887. MR 50 #8493 Zbl 0291.57001

[Masur and Tabachnikov 2002] H. Masur and S. Tabachnikov, “Rational billiards and flat structures”, pp. 1015–1089 in *Handbook of dynamical systems*, vol. 1A, North-Holland, Amsterdam, 2002. MR 2003j:37002 Zbl 1057.37034

[McCullough and Rajeevsarathy 2009] D. McCullough and K. Rajeevsarathy, “Roots of Dehn twists”, preprint, 2009. arXiv 0906.1601

[Monden 2009] N. Monden, “On roots of Dehn twists”, preprint, 2009. arXiv 0911.5079

[Myers and Steenrod 1939] S. B. Myers and N. E. Steenrod, “The group of isometries of a Riemannian manifold”, *Ann. of Math.* (2) **40**:2 (1939), 400–416. MR 1503467 Zbl 0021.06303

[Neumann 1997] W. D. Neumann, “Commensurability and virtual fibration for graph manifolds”, *Topology* **36**:2 (1997), 355–378. MR 98d:57007 Zbl 0872.57021

[Nielsen 1927] J. Nielsen, “Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen”, *Acta Math.* **50** (1927), 189–358. Translated as pp. 223–341 in his *Collected mathematical papers*, vol. 1, Birkhäuser, Boston, 1986. Zbl 53.0545.12

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

[Steiger 1935] F. Steiger, “Die maximalen Ordnungen periodischer topologischer Abbildungen geschlossener Flächen in sich”, *Comment. Math. Helv.* **8**:1 (1935), 48–69. MR 1509518

[Thurston 1979] W. P. Thurston, “The geometry and topology of three-manifolds”, lecture notes, Princeton University, 1979, Available at http://msri.org/publications/books/gt3m.

[Thurston 1988] W. P. Thurston, “On the geometry and dynamics of diffeomorphisms of surfaces”, *Bull. Amer. Math. Soc.* (*N.S.*) **19**:2 (1988), 417–431. MR 89k:57023 Zbl 0674.57008

[Thurston 1997] W. P. Thurston, *Three-dimensional geometry and topology*, vol. 1, Princeton Math. Series **35**, Princeton University Press, Princeton, NJ, 1997. MR 97m:57016 Zbl 0873.57001

[Wu 1987] Y. Q. Wu, “Canonical reducing curves of surface homeomorphism”, *Acta Math. Sinica* (*N.S.*) **3**:4 (1987), 305–313. MR 89e:57007 Zbl 0651.57008

Received April 7, 2010.

DANNY CALEGARI
DEPARTMENT OF MATHEMATICS
CALIFORNIA INSITUTE OF TECHNOLOGY
PASADENA CA 91125
UNITED STATES

dannyc@its.caltech.edu
http://www.its.caltech.edu/∼dannyc

HONGBIN SUN
DEPARTMENT OF MATHEMATICS
PRINCETON UNIVERSITY
FINE HALL, ROOM 304
WASHINGTON ROAD
PRINCETON NJ 08544
UNITED STATES
hongbins@math.princeton.edu

SHICHENG WANG
SCHOOL OF MATHEMATICAL SCIENCES
PEKING UNIVERSITY
BEIJING, 100871
CHINA
wangsc@math.pku.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

ON AN OVERDETERMINED ELLIPTIC PROBLEM

LAURENT HAUSWIRTH, FRÉDÉRIC HÉLEIN AND FRANK PACARD

A smooth flat Riemannian manifold is called an *exceptional domain* if it admits positive harmonic functions having vanishing Dirichlet boundary data and constant (nonzero) Neumann boundary data. In analogy with minimal surfaces, a representation formula is derived and applied to the classification of exceptional domains. Some interesting open problems are proposed along the way.

1. Introduction

Given an m-dimensional Riemannian manifold (M, g) and a smooth bounded domain Ω in M, we denote by $\lambda_1(\Omega)$ the first eigenvalue of the Laplace–Beltrami operator under an identically zero Dirichlet boundary condition. The critical points of the functional

$$\Omega \mapsto \lambda_1(\Omega)$$

under the volume constraint $\operatorname{Vol}\Omega = \alpha$, where $\alpha \in (0, \operatorname{Vol} M)$ is fixed, are called *extremal domains*. Smooth extremal domains are characterized by the property that the eigenfunctions associated with the first eigenvalue of the Laplace–Beltrami operator have constant Neumann boundary data [Soufi 2007]. In other words, a smooth domain is extremal if and only if there exists a positive function u_1 and a constant λ_1 such that

$$\Delta_g u_1 + \lambda_1 u_1 = 0$$

in Ω with $u_1 = 0$ and $\nabla_n u_1$ constant on $\partial\Omega$, where n denotes the inward unit normal vector to $\partial\Omega$.

The theory of extremal domains is very reminiscent of the theory of constant mean curvature surfaces or hypersurfaces. To give some credit to this assertion, we recall that J. Serrin [1971] proved that the only compact, smooth, extremal domains in Euclidean space are round balls, paralleling the well-known result of Alexandrov asserting that round spheres are the only (embedded) compact constant mean curvature hypersurfaces in Euclidean space. More recently, F. Pacard and

We thank Romain Dujardin and Charles Favre for useful discussions. F. Pacard is partially supported by the ANR-08-BLANC-0335-01 grant.
MSC2000: primary 35N25, 35J25, 35R35; secondary 53A10.
Keywords: harmonic function, overdetermined elliptic problem, extremal domain, minimal surface.

P. Sicbaldi [2009] proved the existence of extremal domains close to small geodesic balls centered at critical points of the scalar curvature function, paralleling an earlier result of R. Ye [1991], which provides constant mean curvature topological spheres (with high mean curvature) close to small geodesic spheres centered at nondegenerate critical points of the scalar curvature function.

We propose the following:

Definition 1.1. A smooth domain $\Omega \subset \mathbb{R}^m$ is said to be an *exceptional domain* if it supports positive harmonic functions having identically zero Dirichlet boundary data and constant (nonzero) Neumann boundary data. Any such harmonic function is called a *roof function*.

Exceptional domains arise as limits under scaling of sequences of extremal domains, just like minimal surfaces arise as limits under scaling of sequences of constant mean curvature surfaces. As explained above, there is a formal correspondence between extremal domains and constant mean curvature surfaces. In this note, we try to explain that there is also a strong analogy between exceptional domains and minimal surfaces. More generally, we propose:

Definition 1.2. An m-dimensional flat Riemannian manifold M is said to be *exceptional* if it supports positive harmonic functions having identically zero Dirichlet boundary data and constant (nonzero) Neumann boundary data. Any such harmonic function is called a *roof function*.

Our results raise the problem of the classification of (unbounded) smooth m-dimensional exceptional manifolds. In trying to address this classification problem, we provide a Weierstrass-type representation formula for exceptional flat surfaces. When the dimension $m = 2$, we give nontrivial examples of exceptional domains that are embedded in $\mathbb{R}^2$, and we prove a half-space result for exceptional domains that are conformal to a half-plane.

2. A nontrivial example of an exceptional domain in $\mathbb{R}^2$

The property of being an exceptional domain is preserved under the action of the group of similarities of $\mathbb{R}^m$ (generated by isometries and dilations). We first give trivial examples of exceptional domains in $\mathbb{R}^m$:

(i) The half-space $\{x = (x_1, \dots, x_m) \in \mathbb{R}^m : x_1 > 0\}$ is an exceptional domain in $\mathbb{R}^m$, since the function $u(x) = x_1$ is a positive harmonic function with identically zero Dirichlet boundary data and constant Neumann boundary data.

(ii) The complement of a ball of radius 1 in $\mathbb{R}^m$ is an exceptional domain since the function u defined by $u(x) := \log|x|$ when $m = 2$ and by $u(x) := 1 - |x|^{2-m}$ when $m \geq 3$ is positive, harmonic, and has 0 Dirichlet and constant Neumann data on the unit sphere.

(iii) The product $\Omega \times \mathbb{R}^k$ is an exceptional domain in $\mathbb{R}^m$ provided $\Omega \subset \mathbb{R}^{m-k}$ is an exceptional domain in $\mathbb{R}^{m-k}$.

In dimension $m = 2$, there exists (up to a similarity) at least one other exceptional domain. To describe this domain, we make use of the invariance of the Laplace operator under conformal transformations. The idea is that there exists a (somehow natural) unbounded, positive harmonic function U with identically zero Dirichlet boundary condition on an infinite strip in $\mathbb{R}^2$. This function does not have constant Neumann data, but we can then look for a conformal transformation h which has the property that the pullback of the harmonic function U by h has constant Neumann boundary data on the boundary of the image of the strip by h.

To proceed, it is convenient to identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}$.

Proposition 2.1. *The domain* $\Omega := \{w \in \mathbb{C} : |\operatorname{Im} w| < \pi/2 + \cosh(\operatorname{Re} w)\}$ *is an* exceptional domain.

To prove this result, we define the infinite strip

$$S := \{z \in \mathbb{C} : \operatorname{Im} z \in (-\pi/2, \pi/2)\}$$

and the mapping

$$F(z) := z + \sinh z.$$

Observe that $\Omega = F(S)$. The proof of Proposition 2.1 follows from the next two lemmas.

Lemma 2.2. *The mapping F is a conformal diffeomorphism from S into Ω.*

Proof. We can write

$$F(z) - F(z') = (z - z') \int_0^1 \big(1 + \cosh(tz + (1-t)z')\big)\, dt.$$

In particular

$$\langle z - z', F(z) - F(z')\rangle = |z - z'|^2 \left(1 + \int_0^1 \operatorname{Re}\cosh\,(tz + (1-t)z')\, dt\right), \tag{2-1}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{C}$. Now, for all $x + i\,y \in S$, we have

$$\operatorname{Re}\cosh\,(x + i\,y) = \cosh x\ \cos y \geq 0.$$

This, together with (2-1), implies immediately that F restricted to S is injective. Also,

$$|\partial_z \Lambda(z)|^2 = |1 + \cosh z|^2 = (\cosh x + \cos y)^2.$$

Therefore $\partial_z F$ does not vanish in S. Thus F is a local diffeomorphism, and because the mapping F is holomorphic, it is conformal. □

We define the real-valued function u on Ω by the identity

$$u(F(z)) = \operatorname{Re}\cosh z \qquad \text{for all } z \in S.$$

Lemma 2.3. *The function u is harmonic and positive in Ω, vanishes and has constant Neumann boundary data on $\partial\Omega$.*

Proof. The function W defined in $\mathbb{C}$ by $W(z) := \operatorname{Re}\cosh z$ is harmonic. Indeed, as mentioned in the proof of the previous lemma, $W(x+iy) = \cosh x \, \cos y$. Hence W is both harmonic and positive in S, and vanishes on ∂S. The mapping F being a conformal diffeomorphism from S to Ω, we conclude the function u is both harmonic and positive in Ω, and vanishes on $\partial\Omega$. We claim that u has constant Neumann data on $\partial\Omega$. Indeed, by definition,

$$u(F(z)) = \tfrac{1}{2}(\cosh z + \cosh \bar{z}).$$

Since F is holomorphic, differentiation with respect to z yields

$$2\,\partial_z u(F(z)) = \frac{\sinh z}{1+\cosh z}.$$

Therefore

$$|\nabla u|^2(F(z)) = \frac{\cosh x - \cos y}{\cosh x + \cos y},$$

where $z = x+iy$. On $\partial\Omega$, we have $y = \pm\pi/2$ and hence $|\nabla u| \equiv 1$. Since we already know that $u = 0$ on $\partial\Omega$, we conclude that u has constant Neumann boundary data. $\square$

Lemmas 2.2 and 2.3 complete the proof that $\Omega = F(S)$ is an exceptional domain in $\mathbb{R}^2$ with roof function u.

Remark 2.4. We suspect that this example generalizes to any dimension $m \geq 3$: specifically, there should exist a rotationally symmetric exceptional domain in $\mathbb{R}^m$ for all $m \geq 3$.

3. Toward a global representation formula

Let M be an *exceptional flat surface* (an exceptional domain of dimension 2) with smooth boundary ∂M. Let $\tilde{M}$ be its universal cover and let $\partial\tilde{M}$ be the preimage of ∂M by the covering map $\tilde{M} \to M$. In the following, we exclude the uninteresting case where $\partial M = \varnothing$.

By assumption, M is a flat surface. Hence $\tilde{M}$ is naturally endowed with a flat Riemannian metric g and hence with an induced complex structure, which is conformal to the standard one. Also, there exists an orientation-preserving isometric immersion $F : (\tilde{M}, g) \to (\mathbb{C}, g_{\mathbb{C}})$, where $g_{\mathbb{C}}$ is the canonical Euclidean metric on

$\mathbb{C}$; this induces a smooth immersion of $\partial\tilde{M}$, Observe that F is holomorphic and that $\|dF\|_g = 1$ in $\tilde{M} \cup \partial\tilde{M}$. We define the holomorphic $(1, 0)$-form

$$\Phi := dF = \partial_z F\, dz.$$

Observe that Φ does not vanish and admits a smooth extension to $\tilde{M} \cup \partial\tilde{M}$.

We let $u : M \to \mathbb{R}^+$ be a roof function on M and, with a slight abuse of notation, we denote its lift also by $u : \tilde{M} \to \mathbb{R}^+$. The roof function u can be normalized so that

$$\|\nabla u\|_g = 1 \tag{3-1}$$

on ∂M. We consider the harmonic conjugate function $v : \tilde{M} \to \mathbb{R}$ (uniquely defined up to some additive constant) that is the solution of

$$\partial_z(u - i\,v) = 0 \qquad \text{(and hence } \partial_{\bar{z}}(u + i\,v) = 0\,). \tag{3-2}$$

We set

$$U := u + i\,v.$$

Recall that U is a holomorphic function from $\tilde{M}$ into $\mathbb{C}$. The property that u takes positive values in M and vanishes on ∂M can be translated into the fact that U maps $\tilde{M}$ to $\mathbb{C}^+ := \{w \in \mathbb{C} : \operatorname{Re} w > 0\}$ and $\partial\tilde{M}$ to $i\,\mathbb{R}$. Since $\Phi \neq 0$ on $\tilde{M}$, there exists a unique holomorphic function h on $\tilde{M}$ such that $dU = \partial_z U\, dz = h\,\Phi$. We deduce from the fact that u vanishes on $\partial\tilde{M}$ and from (3-1) that $\nabla_n U = 1$, where n denotes the inward unit normal vector to $\partial\tilde{M}$. Hence

$$\|\partial_z U\|_g = 1 \qquad \text{on } \partial\tilde{M}. \tag{3-3}$$

Now, condition (3-1) translates into the fact that $\|\Phi\|_g = \|dF\|_g = 1 = \|dU\|_g$ on $\partial\tilde{M}$. Clearly, this is equivalent to the fact that $|h| = 1$ on $\partial\tilde{M}$. Therefore, we end up with the following data:

(i) An oriented, simply connected complex surface $\tilde{M}$ with smooth boundary $\partial\tilde{M}$.

(ii) A holomorphic function U, defined on $\tilde{M}$, which takes values in $\mathbb{C}^+$ and maps $\partial\tilde{M}$ into $i\,\mathbb{R}$.

(iii) A holomorphic function h, defined on $\tilde{M}$, such that $|h| = 1$ on $\partial\tilde{M}$, and for which the 1-form Φ defined by $\Phi := (1/h)dU$ does not vanish on $\tilde{M}$.

By analogy with the theory of minimal surfaces, we call these data the Weierstrass-type representation formula for exceptional flat surfaces.

Conversely, given a set of such data, we can define the map $F : \tilde{M} \to \mathbb{C}$ by integrating $dF = \Phi$. Thanks to (iii), this map is an immersion and its image is an immersed exceptional flat surface with roof function given by $u = \operatorname{Re} U$. In

Section 4, we will give some explicit examples of such constructions when $\partial\tilde{M}$ is equal to $\partial D \setminus \{\alpha_1, \dots, \alpha_n\}$, where $\alpha_1, \dots, \alpha_n$ is a finite collection of points on $\partial D = S^1$.

Example 3.1. Here is a (rather pathological) illustration of this Weierstrass-type formula. Consider $M = \mathbb{C}^+$, the function $U(z) = z$ and

$$F(z) = \int_0^z e^{-\sinh\zeta}\, d\zeta.$$

Note that $\partial_z F$ is $2i\pi$-periodic, and this implies that $F(z+2i\pi) = F(z)+C$, where the constant C is given explicitly by

$$C := i\int_0^{2\pi} e^{-i\sin s}\, ds.$$

Moreover, for $x > 0$,

$$F(x+iy) = F(iy) + \int_0^x e^{-\sinh(s+iy)}\, ds$$

converges to $+\infty$ as $x \to +\infty$ if $y = 0$, but this quantity is bounded if $|y-\pi| < \pi/2$, and even admits a finite limit as $x \to +\infty$.

Hence, in addition to the regular boundary $F(i\mathbb{R})$, which is a smooth periodic curve, the image of F has a singular boundary: the set of limits of $F(x+iy)$ as u tends to $+\infty$, for the values of y for which this limit exists. The roof function tends to infinity along this singular boundary.

4. Examples of exceptional flat surfaces

Thanks to the Weierstrass-type representation in the previous section, we can give many nontrivial examples of exceptional flat surfaces. We keep the notation from that section.

The construction makes use of an integer $n \in \mathbb{N} \setminus \{0\}$ and the Riemann surface $D = \{z \in \mathbb{C} : |z| < 1\}$. On D, we define the holomorphic functions

$$h(z) = z^{n-1} \quad \text{and} \quad U(z) := \frac{1+z^n}{1-z^n}.$$

The 1-form Φ is given by

$$\Phi(z) := \frac{2n}{(1-z^n)^2}\, dz.$$

Both U and Φ have singularities at the n-th roots of unity. The function F is then obtained by integrating Φ, and the roof function u is defined by $u = \operatorname{Re} U$.

(i) When $n = 1$, we can take

$$F(z) = \frac{1+z}{1-z}.$$

In this case, we simply have $F(D) = \mathbb{C}^+$, and we recover the fact that the half-plane is an exceptional domain. This is the counterpart of the plane in the framework of minimal surfaces.

(ii) When $n = 2$, we can take

$$F(z) = \frac{2z}{1-z^2} + \log\frac{z+1}{z-1}.$$

In this case, the exceptional flat surface found can be isometrically embedded in $\mathbb{C}$, and hence $F(D)$ is an exceptional domain. In fact, $F(D)$ corresponds (up to some similarity) to the domain Ω, which was defined in Proposition 2.1. This exceptional domain is the counterpart of the catenoid.

(iii) Finally when $n \geq 3$, the exceptional flat surfaces we find cannot be isometrically embedded in $\mathbb{C}$ anymore. They are counterparts of the minimal n-noids described in [Jorge and Meeks 1983].

Let us analyze this example further. The function U can be written as

$$U(z) = -\frac{1}{n}\sum_{k=1}^{n}\frac{z+\alpha^k}{z-\alpha^k},$$

where $\alpha := e^{i2\pi/n}$. In particular, $\operatorname{Re} U$ is nothing but a multiple of the sum of the Poisson kernel on the unit disc with poles at $1, \alpha, \dots, \alpha^{n-1}$. Next,

$$dU = z^{n-1}\,\frac{2n}{(1-z^n)^2}\,dz,$$

so the function h is cooked up to counterbalance the zero of dU and ensure that Φ does not vanish in the unit disk, while keeping the condition $|dU|^2 = |\Phi|^2$ on ∂D.

To generalize the example, consider n distinct points $\alpha_1, \dots, \alpha_n \in S^1 \subset \mathbb{C}$ and $a_1, \dots, a_n > 0$. We define

$$U(z) := -\sum_{k=1}^{n} a_k\,\frac{z+\alpha_k}{z-\alpha_k}. \tag{4-1}$$

It is easy to check that $\operatorname{Re} U$ is positive (since each function $z \mapsto -\frac{z+\alpha_k}{z-\alpha_k}$ maps D to $\mathbb{C}^+$) and vanishes on $\partial D \setminus \{\alpha_1, \dots, \alpha_n\}$. We have

$$\prod_{k=1}^{n}(z-\alpha_k)^2\,dU = P(z)\,dz,$$

where P is a polynomial that depends on the choice of points $\alpha_1, \dots, \alpha_n$ and weights $a_1, \dots, a_n$. Assume that P does not vanish on ∂D and denote by $z_1, \dots, z_l$ the roots of P in the unit disc, counted with multiplicity. We simply define

$$h(z) := \prod_{j=1}^{l} \frac{z - z_j}{z\,\bar{z}_j - 1}$$

and the 1-form Φ by $\Phi := (1/h)\, dU$. Integration of Φ yields a $2n$-dimensional family of exceptional flat surfaces immersed in $\mathbb{C}$.

5. A global Weierstrass-type representation

In this section, we show that exceptional flat surfaces whose immersion in $\mathbb{C}$ have finitely many *regular ends* and are locally finite coverings of $\mathbb{C}$ are precisely the examples in the previous section. We use the notations introduced in Section 3, and we set

$$\widehat{M} := M \cup \partial M.$$

We further assume that M is *simply connected* and that $\partial M \neq \varnothing$. In particular, M has the conformal type of the unit disk D, and without loss of generality, we can assume that M is indeed equal to D and consider $\bar{D}$ as a natural compactification of M. We denote by F an orientation preserving, holomorphic, isometric immersion $F : (\widehat{M}, g) \to (\mathbb{C}, g_{\mathbb{C}})$. Recall that $\|dF\|_g = 1$ on ∂M. Some natural hypotheses are needed:

(H-1) *M has finitely many ends.* This means that

$$\partial M = \partial D \setminus \bigcup_{j=1}^{n} E_j = \bigcup_{j=1}^{n} I_j,$$

where each $E_j \subset S^1$ is a closed arc and $I_j \subset S^1$ is an open arc.

(H-2) *F is proper.* This means that $F(w)$ tends to infinity as w tends to $\bigcup_{j=1}^{n} E_j$.

(H-3) *Each end of M is regular.* This means the image of $I_j := (\theta_j^-, \theta_j^+)$ by F is a curve Γ_j asymptotically parallel to fixed directions at infinity. In other words, there exist two unit vectors τ_j^- and $\tau_j^+ \in S^1 \subset \mathbb{C}$ such that

$$\lim_{\theta \in I_j,\, \theta \to \theta_j^\pm} \frac{F(e^{i\theta})}{|F(e^{i\theta})|} = \tau_j^\pm.$$

This is the case, for example, if we assume each Γ_j has finite total curvature.

(H-4) *The mapping F is a locally finite covering.* This means there exists $d \in \mathbb{N}^*$ such that, for any $z \in \mathbb{C}$, the cardinal of $\{\zeta \in M : F(\zeta) = z\}$ is at most d.

We now state the main result of this section.

Theorem 5.1. *Assume that M is a simply connected exceptional flat surface and let $F : M \to \mathbb{C}$ be an isometric immersion. Further assume that* (H-1)–(H-4) *holds and identify M with D. Then there exist $\mu \in \mathbb{R}$, n distinct points $\alpha_1, \dots, \alpha_n \in S^1$ and n constants $a_1, \dots, a_n > 0$ such that*

$$dF = e^{i\mu} \prod_{k=1}^{m} \frac{\bar{z}_k z - 1}{z - z_k} \, dU,$$

where $z_1, \dots, z_m \in \bar{D}$ denote the zeros of dU, counted with multiplicity, and where

$$U(z) := -\sum_{j=1}^{n} a_j \frac{z + \alpha_j}{z - \alpha_j} \quad \text{in } \bar{D}.$$

The proof is divided into a few lemmas and propositions. We start by analyzing the ends E_j and show that they reduce to isolated points $\alpha_1, \dots, \alpha_n$. Next we analyze the behavior of F near the points α_j and show that F does not have any essential singularity there. Then we proceed with the analysis of the function U and show that it has the expected form. The proof is completed with the study of the function h.

As promised, we first analyze the sets E_j:

Lemma 5.2. *Under the assumptions of Theorem 5.1, there exists a finite number of points $\alpha_1, \dots, \alpha_n \in \partial D = S^1$ such that $\widehat{M} = \bar{D} \setminus \{\alpha_1, \dots, \alpha_n\}$.*

Proof. We need to show that each interval E_j is reduced to a point. This essentially follows from the fact that the capacity of E_j vanishes.

Suppose, for a contradiction, that E_j is an arc of positive arc length for some j, and take some $l \in (0, \pi/2)$ and an arc $E \subset E_j$ of length l. Our problem being invariant under the action of fractional linear transformations of the unit disk, we can assume without loss of generality that E is the image of $[-l/2, l/2]$ under $s \mapsto e^{is}$. Reducing l if necessary, we can also assume that the opposite arc $-E$, the image of $[-l/2, l/2]$ under $s \mapsto -e^{is}$, is contained in $S^1 \setminus \bigcup_{j=1}^{n} E_j$.

Recall that for any smooth function defined on (a, b) which satisfies $f(b) = 1$ and $f(a) = 0$, we have

$$1 = f(b) - f(0) = \int_a^b f'(s)\, ds \le \left(\int_a^b (f')^2(s)\, ds \right)^{1/2} \sqrt{b - a}.$$

If in addition, $b - a \le 2$, we conclude that

$$\int_a^b (f')^2(s)\, ds \ge \tfrac{1}{2}.$$

Now assume that we are given a smooth function $f : \bar{D} \to \mathbb{R}$ such that $f = 1$ on E and $f = 0$ on $-E$. Using the previous inequality, we can write

$$\int_D \|\nabla f\|^2_{g_{\mathbb{C}}}\, dx\, dy \geq \int_{D\cap\{|x|<\sin(l/2)\}} |\partial_y f|^2\, dx\, dy \geq \int_{|x|\leq \sin(l/2)} \tfrac{1}{2}\, dx = \sin \frac{l}{2}. \tag{5-1}$$

Given $R > r > 0$, define $\chi : \mathbb{C} \to \mathbb{R}$ by

$$\chi(z) = \begin{cases} 0 & \text{if } |z| \leq r, \\ \dfrac{\log(|z|/r)}{\log(|z|/R)} & \text{if } r \leq |z| \leq R, \\ 1 & \text{if } R \leq |z|, \end{cases}$$

and we define $f : D \to \mathbb{R}$ by $f := \chi \circ F$. Since F is conformal, we can write

$$\int_D \|\nabla f\|^2_{g_{\mathbb{C}}}\, dx\, dy = \int_D \|\nabla f\|^2_g\, d\,\mathrm{vol}_g\, .$$

Now, using (H-4), we conclude that

$$\int_D \|\nabla f\|^2_g\, d\,\mathrm{vol}_g \leq d \int_{\mathbb{C}} \|\nabla \chi\|^2_{g_{\mathbb{C}}}\, dx\, dy = d\, \frac{2\pi}{\log(R/r)}. \tag{5-2}$$

Fixing $r > 0$ large enough, we can ensure that f is identically equal to 0 on $-E$. Using (H-2), we see that f is identically equal to 1 on each E_j, and in particular on E. Therefore f can be used in (5-1), which together with (5-2) yields

$$2\,\pi\, d \geq \sin\frac{l}{2}\, \log\frac{R}{r}$$

independently of $R > r$. Letting R tend to infinity, we get a contradiction, and the proof is complete. □

Therefore, we now know that $E_j := \{\alpha_j\}$. Without loss of generality, we can assume that $\alpha_1, \dots, \alpha_n$ are arranged counterclockwise along S^1. We agree that $\alpha_0 := \alpha_n$ and $\alpha_{n+1} := \alpha_1$, and that for each $j = 1, \dots, n$, the arc I_j is positively oriented and joins α_j to α_{j+1}. We now analyze the singularities of F close to α_j.

Given $j = 1, \dots, n$, we denote by $S(\alpha_j, r)$ the circle of radius $r > 0$ centered at α_j. We define

$$\gamma_j := \bar{D} \cap S(\alpha_j, r),$$

which we assume to be oriented clockwise. The *angle* $\theta_j \in \mathbb{R}$ at α_j is defined by

$$\theta_j := -\lim_{r\to 0} \int_{\gamma_k} F^* d\theta,$$

where $d\theta := \operatorname{Im} dz/z$. Thanks to (H-3), θ_j is well defined, and we have

$$\tau_j^- = e^{i\theta_j}\, \tau_{j-1}^+.$$

Lemma 5.3. *Under the assumption of Theorem 5.1, the function*

$$H_j(z) := (z-\alpha_j)^{\theta_j/\pi}\, F(z)$$

is holomorphic in a neighborhood of α_j in $\bar{D} \setminus \{\alpha_j\}$, and $H_j(\alpha_j) \neq 0$.

Proof. Without loss of generality, we can assume that $\alpha_j = 1$. By right composing F with the conformal transformation $z \mapsto (1-z)/(1+z)$, we can replace D by $\mathbb{C}^+$. Now we define

$$G(z) := F(z)^{-\pi/\theta_j}.$$

Observe that $G(0)=0$ by (H-2). Moreover, (H-3) and the definition of θ_j imply that the image by G of a neighborhood of 0 in $i\,\mathbb{R}$ is a $\mathscr{C}^1$-curve, and hence analytic. In particular, there exists some conformal transformation T such that, for some $r > 0$, the image by $T \circ G$ of $i\,(-r, r)$ is a straight line segment in $i\,\mathbb{R}$. Then it is possible to extend $T \circ G$ to a function $\tilde{G}$ defined on a neighborhood of 0 in $\mathbb{C}$ by setting

$$\tilde{G}(z) = \begin{cases} T(G(z)) & \text{if } \operatorname{Im} z \geq 0, \\ -\overline{T(G(-\bar{z}))} & \text{if } \operatorname{Im} z \leq 0. \end{cases}$$

The resulting $\tilde{G}$ is bounded in a neighborhood of 0 in $\mathbb{C}$ and holomorphic away from 0. It is well known that the singularity is then removable and hence it is holomorphic. Therefore $\tilde{G}$ is actually holomorphic in a neighborhood of 0. In particular, we can write

$$G(z) = z^k\, H(z)$$

near 0, where H is a holomorphic function that does not vanish at 0. Going back to the definition of G, this implies that

$$F(z) = (z-\alpha_j)^{-k\,\theta_j/\pi}\, H_j(z),$$

where H_j is holomorphic in a neighborhood of α_j and does not vanish at α_j. But the definition of θ_j readily implies that $k = 1$. This completes the proof. □

As a corollary, we conclude that

$$H(z) := F(z) \prod_{j=1}^{n} (z-\alpha_j)^{\theta_j/\pi} \tag{5-3}$$

is a bounded holomorphic function in $\overline{D}$. Moreover, since F tends to infinity as z approaches α_j, this implies that $\theta_j > 0$.

We now make use of the fact that M is an exceptional domain, and hence there is a roof function $u : \widehat{M} \to [0, +\infty)$. We can define the holomorphic function $U := u + i\,v$, where $v : \widehat{M} \to \mathbb{R}$ is the (real-valued) harmonic conjugate of u. The purpose of the next result is to show that U is precisely given by (4-1).

Lemma 5.4. *Under the assumptions of Theorem 5.1, there exist n constants a_1, $\dots, a_n > 0$ such that*

$$U(z) = -\sum_{j=1}^{n} a_j \, \frac{z+\alpha_j}{z-\alpha_j}.$$

Proof. First, it is possible to extend the function U to all $\mathbb{C} \setminus \{\alpha_1, \dots, \alpha_n\}$ by defining V to be equal to U in $\bar{D} \setminus \{\alpha_1, \dots, \alpha_n\}$, and

$$V(z) := -\overline{U(1/\bar{z})}$$

when $z \in \mathbb{C} \setminus \bar{D}$. The key observation is that, since $\operatorname{Re} U = 0$ on $\partial D \setminus \{\alpha_1, \dots, \alpha_n\}$, the function V is continuous and in fact holomorphic on $\mathbb{C} \setminus \{\alpha_1, \dots, \alpha_n\}$. Moreover, V converges to $V(\infty) := -\overline{U(0)}$ at infinity.

We proceed with the proof that the function V has no essential singularity at any α_j; it will follow from Picard's theorem. By definition, $\operatorname{Re} V$ vanishes on I_j and is positive in D. Therefore the outward normal derivative of $\operatorname{Re} V$ on I_j is negative. This implies that the tangential derivative of $\operatorname{Im} V$ on I_j does not vanish and hence that $\operatorname{Im} V$ is strictly monotone on each I_j. This shows that there exists some neighborhood $\mathcal{V}$ of α_j in $\mathbb{C}$ such that any element of $i\,\mathbb{R}$ is achieved by V at most twice in $\mathcal{V}$ (that is, at most once on I_j and at most once on I_{j-1} and certainly not in $\mathcal{V} \setminus \partial D$, since V takes values in $\mathbb{C} \setminus i\,\mathbb{R}$ away from ∂D). Picard's big theorem [Conway 1978] then implies that α_j is not an essential singularity of V. Hence α_j is either a removable singularity of V or a pole.

Since $\|\nabla u\|_g \equiv 1$ on ∂M, this forces $|\partial_z U| = |\partial_z F|$ on ∂M, and since $|\partial_z F|$ tends to $+\infty$ at α_j, so does $|\partial_z U|$. Hence all α_j are poles of V.

We are now interested in the zeros of V. Since $\operatorname{Re} V$ takes positive values in D and negative values in $(\mathbb{C} \cup \{\infty\}) \setminus \bar{D}$, we already know that the only possible zeros of V are on ∂D. We have already seen that, along I_j, the function V equals $i\,v$, where v is strictly monotone. Further, since α_{j-1} and α_j are poles of this function, $|V|$ must converge to $+\infty$ as we approach either α_{j-1} or α_j. Because of the continuity of v along each I_j it follows that v vanishes exactly at one point β_j on each I_j. Moreover, this zero is simple: if it had order $k > 1$, the zero set of $\operatorname{Re} V$ near β_j would contain k curves intersecting at β_j, and this would force $\operatorname{Re} V = \operatorname{Re} U$ to vanish in D, in contradiction with our hypothesis.

Finally, we prove that V has only simple poles. We know that V extends meromorphically to a map on $\mathbb{C}P^1 = \mathbb{C} \cup \{\infty\}$ with neither a pole nor a zero at infinity. Furthermore, V has exactly n simple zeros and n poles; hence these poles must be simple. To summarize, V can be written as a linear combination of the constant function and functions of the form $z \mapsto 1/(z-\alpha_j)$. Without loss of generality, this amounts to saying that V can be written as

$$V(z) = a - \sum_{j=1}^{n} a_j \, \frac{z+\alpha_j}{z-\alpha_j},$$

where a and the a_j are complex numbers. Using the fact that, by construction, $V(1/\bar{z}) = -\overline{V(z)}$, we conclude that $a \in i\,\mathbb{R}$ and also that $a_j \in \mathbb{R}$. Moreover, since $\operatorname{Re} U$ is positive, this implies that the a_j are positive real numbers. This completes the proof, since U is defined up to the addition of some element of $i\,\mathbb{R}$. □

We are now in a position to complete our analysis of the function F. Since F is an immersion, $dF \neq 0$ on $\widehat{M}$. Hence there exists a unique holomorphic function h on $\widehat{M}$ such that

$$\partial_z U = h\,\partial_z F \tag{5-4}$$

on $\widehat{M}$. Moreover, since $\|\nabla u\|_g \equiv 1$ on ∂M, this implies that $|h| \equiv 1$ on ∂M. We now analyze the function h, which will complete the proof of Theorem 5.1.

Lemma 5.5. *Under the assumptions of Theorem 5.1, there exists a constant $e^{i\mu} \in \mathbb{R}$ such that the function h defined by* (5-4) *has the form*

$$h(z) = e^{-i\mu} \prod_{k=1}^{m} \frac{z - z_k}{\bar{z}_k z - 1},$$

where $z_1, \dots z_m$ are the zeros of $\partial_z U$ in D counted with multiplicity.

Proof. The function h is holomorphic in D and satisfies $|h|=1$ on $\partial D \setminus \{\alpha_1, \dots, \alpha_n\}$. We can extend h to a holomorphic function H, defined on $(\mathbb{C} \cup \{\infty\}) \setminus \{\alpha_1, \dots, \alpha_n\}$ by setting $H(z) := h(z)$ for all $z \in \overline{D} \setminus \{\alpha_1, \dots, \alpha_n\}$ and

$$\overline{H(z)} := \frac{1}{h(1/\bar{z})} \tag{5-5}$$

for all $z \in \mathbb{C} \setminus \overline{D}$. Clearly H is locally bounded in $\overline{D} \setminus \{\alpha_1, \dots, \alpha_n\}$, and its only singularities in $(\mathbb{C} \cup \{\infty\}) \setminus \overline{D}$ are poles that are the images by $z \mapsto 1/\bar{z}$ of the zeros of h; hence H is meromorphic outside $\{\alpha_1, \dots, \alpha_n\}$. But Lemma 5.3 and (5-3) imply that, near α_j, $|H|$ is bounded by a constant times $|z - \alpha_j|^{-k_j}$ for some $k_j > 0$. Therefore α_j is not an essential singularity of H, and hence H is meromorphic in $\mathbb{C} \cup \{\infty\}$.

Observe that $|H(z)| = 1$ on $\partial D \setminus \{\alpha_1, \dots, \alpha_n\}$, and this implies that the points α_j are not poles of H. Therefore, the singularities α_j of H are removable. Also, we have

$$\Delta\,|H|^2 = 4\partial_z \partial_{\bar{z}}\,|H|^2 = 4\,|\partial_z H|^2 \geq 0,$$

and since $|H| = 1$ on ∂D, the maximum principle implies that $|H| \leq 1$ in D.

Since H is bounded in $\bar{D}$, it does not have poles in this set. This also implies that H has no zeroes in $(\mathbb{C}\cup\{\infty\})\setminus D$, because otherwise H would have poles in $\bar{D}$ by (5-5). Therefore, if $z_1, \ldots, z_m \in D$ denote the zeros of H (counted with multiplicity), the poles of H are given by $1/\bar{z}_1, \ldots, 1/\bar{z}_m$ (also counted with multiplicity). It is then a simple exercise to check that H is of the form

$$H(z) = C \prod_{k=1}^{m} \frac{z - z_k}{\bar{z}_k\, z - 1}$$

for some constant $C \in \mathbb{C}$. Finally, the condition that $|H(z)| = 1$ on ∂D forces $|C| = 1$. This completes the proof. □

6. A Bernstein type result for two-dimensional exceptional domains

We prove the following Bernstein type result for two-dimensional exceptional domains.

Proposition 6.1. *Assume that Ω is a two-dimensional exceptional domain conformal to $\mathbb{C}^+$, and let u be a roof function on Ω. We further assume that $\partial_x u > 0$ in Ω. Then Ω is a half-plane.*

Proof. Since we have assumed that Ω is conformal to $\mathbb{C}^+$, there exists a holomorphic map $\Psi : \mathbb{C}^+ \mapsto \Omega$. We then define

$$H := (\partial_z u) \circ \Psi.$$

The function H is holomorphic in $\mathbb{C}^+$ and does not vanish, since we have assumed that $\partial_x u \neq 0$. Moreover, $|H| \equiv 1$ on $\partial\mathbb{C}^+$. We can write $H = e^{i\Theta}$, where Θ is a holomorphic function defined in $\mathbb{C}^+$ that is real valued on the imaginary axis. This means that $\operatorname{Im}\Theta = 0$ when $\operatorname{Re} z = 0$. Since we have assumed that $\partial_x u > 0$, we also conclude that $\operatorname{Re}\Theta \in (-\pi/2, \pi/2)$.

We can extend Θ as a holomorphic function $\tilde{\Theta}$ in $\mathbb{C}$ as follows:

$$\tilde{\Theta}(z) := \begin{cases} \Theta(z) & \text{if } \operatorname{Re} z \geq 0, \\ \overline{\Theta(-\bar{z})} & \text{if } \operatorname{Re} z < 0. \end{cases}$$

It is easy to check that $\tilde{\Theta}$ is a holomorphic function: in fact, the real part of Θ is extended as an even function of $\operatorname{Re} z$, while the imaginary part of Θ is extended as an odd function of $\operatorname{Re} z$. That $\tilde{\Theta}$ is $\mathscr{C}^1$ is then a consequence of the fact that $\operatorname{Im}\Theta = 0$ on the imaginary axis, while the holomorphicity of Θ follows from the fact that $\partial_x \operatorname{Re}\Theta = 0$ on the imaginary axis of $\mathbb{C}$.

The real part of $\tilde{\Theta}$, being a bounded harmonic function, must be constant. Then $\tilde{\Theta}$, being holomorphic, must itself be constant. But this implies that the gradient of u is constant, and hence the level sets of u are straight lines. This implies that

u only depends on one variable, and hence it is an affine function. This completes the proof. □

Corollary 6.2. *There is no exceptional domain contained in a wedge*

$$\Omega \subset \{z \in \mathbb{C} : \operatorname{Re} z \geq \kappa\, |\operatorname{Im} z|\}$$

for any $\kappa > 0$.

Proof. The proof is by contradiction. If Ω were such an exceptional domain, there would exist on Ω a roof function u. One can apply the moving plane method [Serrin 1971; Gidas et al. 1979] to prove that $\partial_x u > 0$ and hence that $\partial\Omega$ is a graph over the y-axis. Since Ω is contained in a half-plane, there is no bounded, positive, harmonic function on Ω having 0 boundary data on $\partial\Omega$; otherwise one could use an affine function as a barrier to obtain a contradiction. Certainly, $\Omega \cup \partial\Omega$ is conformal to $\bar{D} \setminus E$, where D is the unit disc and E is a closed arc included in S^1. Necessarily, E is reduced to a point, since otherwise we can construct bounded, positive, harmonic functions on E that have 0 boundary data on $S^1 \setminus E$, and these would lift to bounded, positive, harmonic function on Ω, with 0 boundary data, a contradiction. Therefore, we conclude that Ω is conformal to $\mathbb{C}^+$. The assumptions of Lemma 5.5 are fulfilled, and hence we conclude that Ω is a half-plane, which is a contradiction. □

7. Open problems

We have no nontrivial example of an exceptional domain in higher dimensions $\mathbb{R}^m$ for $m \geq 3$, besides the ones described in Section 2. In dimension $m = 2$, it is tempting to conjecture that (up to similarity) the only exceptional domains that can be embedded in $\mathbb{R}^2$ are half-spaces, the complement of a ball and the example discussed in Section 2.

References

[Conway 1978] J. B. Conway, *Functions of one complex variable, I*, Graduate Texts in Math. **11**, Springer, New York, 1978. MR 80c:30003 Zbl 0277.30001

[Gidas et al. 1979] B. Gidas, W.-M. Ni, and L. Nirenberg, "Symmetry and related properties via the maximum principle", *Commun. Math. Phys.* **68**:3 (1979), 209–243. MR 80h:35043 Zbl 0425.35020

[Jorge and Meeks 1983] L. P. Jorge and W. H. Meeks, III, "The topology of complete minimal surfaces of finite total Gaussian curvature", *Topology* **22**:2 (1983), 203–221. MR 84d:53006 Zbl 0517.53008

[Pacard and Sicbaldi 2009] F. Pacard and P. Sicbaldi, "Extremal domains for the first eigenvalue of the Laplace–Beltrami operator", *Ann. Inst. Fourier* (*Grenoble*) **59**:2 (2009), 515–542. MR 2521426 Zbl 1166.53029

[Serrin 1971] J. Serrin, "A symmetry problem in potential theory", *Arch. Ration. Mech. Anal.* **43**:4 (1971), 304–318. MR 48 #11545 Zbl 0222.31007

[Soufi 2007] A. El Soufi and S. Ilias, "Domain deformations and eigenvalues of the Dirichlet Laplacian in a Riemannian manifold", *Illinois J. Math.* **51**:2 (2007), 645–666. MR 2342681 Zbl 1124.49035

[Ye 1991] R. Ye, "Foliation by constant mean curvature spheres", *Pacific J. Math.* **147**:2 (1991), 381–396. MR 92f:53030 Zbl 0722.53022

Received January 4, 2010.

LAURENT HAUSWIRTH
laurent.hauswirth@univ-mlv.fr
UNIVERSITÉ PARIS-EST MARNE-LA-VALLÉE
LABORATOIRE D'ANALYSE ET MATHÉMATIQUES APPLIQUÉES
CITÉ DESCARTES
5 BOULEVARD DESCARTES
CHAMPS-SUR-MARNE
77454 MARNE-LA-VALLÉE CEDEX 2
FRANCE
http://perso-math.univ-mlv.fr/users/hauswirth.laurent

FRÉDÉRIC HÉLEIN
helein@math.jussieu.fr
UNIVERSITÉ PARIS DIDEROT–PARIS 7
INSTITUT DE MATHÉMATIQUES DE JUSSIEU, UMR CNRS 7586
UFR DE MATHÉMATIQUES, CASE 7012
BÂTIMENT CHEVALERET
75205 PARIS CEDEX 13
FRANCE
http://www.math.jussieu.fr/~helein/

FRANK PACARD
INSTITUT UNIVERSITAIRE DE FRANCE ET CENTRE DE MATHÉMATIQUES LAURENT SCHWARTZ
UMR CNRS 7640
ÉCOLE POLYTECHNIQUE
91128 PALAISEAU
FRANCE
frank.pacard@math.polytechnique.fr
http://www.math.polytechnique.fr/~pacard

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

MINIMAL SETS OF A RECURRENT DISCRETE FLOW

HATTAB HAWETE

S. G. Dani, giving a counterexample to a result in a paper of Knight, showed that recurrent transitive flows can admit multiple minimal sets. Here we show that such a phenomenon occurs on a wider scale.

Let (X, T) be a discrete flow, where X is a compact metric space and T is a self-homeomorphism of X. For $x \in X$, the set $\{T^n(x) : n \in \mathbb{Z}\}$ is called the *orbit* of x and is denoted by $O(x, T)$. A set W is a *minimal set* of (X, T) if for all $x \in W$ we have $\overline{O(x, T)} = W$. The study of minimal sets of such a system is a central question in topological dynamics. Zorn's lemma ensures the existence of at least a minimal set of (X, T). If X is a minimal set, (X, T) is called a *minimal flow*.

A point x of X is *recurrent* if $T^{n_k}(x) \to x$ for some sequence $n_k \to +\infty$. When each point of X is recurrent we say that (X, T) is a *recurrent flow*. All periodic points are recurrent. The standard example of a nonperiodic recurrent point is any point in the irrational flow on the circle $\mathbb{S}^1$. Every point in a minimal set is recurrent, so the existence of minimal sets implies the existence of recurrent points.

Knight [1987] purported to prove that, if X is a compact recurrent orbit closure in (X, T), then any pair of orbit closures intersect and, in particular, X contains a unique compact minimal set. Dani [1991] pointed out with a counterexample that this statement is false.

In Theorem 0.1 below we enlarge the class of known counterexamples. More specifically, for any weakly mixing, minimal, uniformly rigid system (X, T) the system $(X \times X, T \times T)$, defined by $(T \times T)(x_1, x_2) = (T(x_1), T(x_2))$ for $(x_1, x_2) \in X \times X$, is a recurrent and transitive system with multiple minimal sets.

(Recall that the a discrete flow (X, T) is called

- *transitive* if there exists $x_0 \in X$ with a dense orbit;
- *ergodic* if for all two open subsets U and V there exits n such that $T^n U \cap V$ is nonempty;
- *weakly mixing* if the discrete flow $(X \times X, T \times T)$ is ergodic;

MSC2010: 54H20.
Keywords: minimal set, discrete flow, uniformly rigid, weakly mixing.

- *uniformly rigid* if there exists a sequence $n_k \to +\infty$ such that

$$\lim_{n_k \to +\infty} \sup_{x \in X} d(T^{n_k} x; x) = 0,$$

where d is the metric on X.)

Minimal uniformly rigid weakly mixing systems exist; see [Glasner and Maon 1989, Proposition 6.5].

Theorem 0.1. *Let (X, T) be a minimal uniformly rigid weakly mixing system. Then $(X \times X, T \times T)$ is transitive and recurrent, and admits infinitely many minimal sets.*

Proof. Let (X, T) be a minimal uniformly rigid weakly mixing system.

<u>Step 1:</u> $(X \times X, T \times T)$ *is transitive.* Since (X, T) is weakly mixing, $(X \times X, T \times T)$ is ergodic. But for discrete flows on compact spaces, ergodicity is equivalent to transitiveness; see [de Vries 1993], for example. Because $X \times X$ is compact, this means that $(X \times X, T \times T)$ is transitive.

<u>Step 2:</u> $(X \times X, T \times T)$ *is recurrent.* Since (X, T) is a uniformly rigid flow, there is a sequence $n_k \to +\infty$ such that

$$\lim_{n_k \to +\infty} \sup_{x \in X} d(T^{n_k} x, x) = 0.$$

For each point (x, y) point of $X \times X$ we have

$$\lim_{n_k \to +\infty} (T \times T)^{n_k}(x, y) = \lim_{n_k \to +\infty} (T^{n_k} x, T^{n_k} y) = (x, y).$$

Thus (x, y) is a recurrent point and so $(X \times X, T \times T)$ is a recurrent discrete flow.

<u>Step 3:</u> *There are infinitely many minimal sets of* $(X \times X, T \times T)$. Define $D_n = \{(x, T^n(x)) : x \in X\}$. Then D_n is an invariant closed set of $(X \times X, T \times T)$. If F is a nonempty closed $(T \times T)$-invariant subset of D_n, then so is its projection, say $p_1(F)$, on the first factor. By the minimality of T we get $p_1(F) = X$, and hence $F = D_n$. Thus D_n is minimal for every n. Since (X, T) is minimal it follows that the D_n are pairwise distinct. □

Remark 0.2. The discrete flow $(X \times X, T \times T)$ does not have fixed points because we chose (X, T) as a minimal discrete flow.

Acknowledgement

The author gratefully acknowledges helpful corrections, comments and suggestions from the referee.

References

[Dani 1991] S. G. Dani, "A remark on recurrent dynamical systems", *J. Indian Math. Soc. (N.S.)* **56**:1-4 (1991), 1–6. MR 93g:58113 Zbl 0864.54033

[Glasner and Maon 1989] S. Glasner and D. Maon, "Rigidity in topological dynamics", *Ergodic Theory Dynam. Systems* **9**:2 (1989), 309–320. MR 90h:54050 Zbl 0661.58027

[Knight 1987] R. A. Knight, "Minimal sets in recurrent discrete flows", *Proc. Amer. Math. Soc.* **100**:1 (1987), 195–198. MR 88f:54076 Zbl 0612.54051

[de Vries 1993] J. de Vries, *Elements of topological dynamics*, Math. and its Appl. **257**, Kluwer, Dordrecht, 1993. MR 94m:54098 Zbl 0783.54035

Received February 19, 2010. Revised June 14, 2010.

HATTAB HAWETE
INSTITUT SUPÉRIEUR D'INFORMATIQUE ET DU MULTIMEDIA
ROUTE DE TUNIS KM 10
B.P. 242
SFAX 3021
TUNISIA
hattab.hawete@yahoo.fr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

TRACE-POSITIVE POLYNOMIALS

IGOR KLEP

In this paper positivity of polynomials in free noncommuting variables in a dimension-dependent setting is considered. That is, the images of a polynomial under finite-dimensional representations of a fixed dimension are investigated. It is shown that unlike in the dimension-free case, every trace-positive polynomial is (after multiplication with a suitable denominator — a Hermitian square of a central polynomial) a sum of a positive semidefinite polynomial and commutators. Together with our previous results this yields the following Positivstellensatz: every trace-positive polynomial is modulo sums of commutators and polynomial identities a sum of Hermitian squares with weights and denominators. Understanding trace-positive polynomials is one of the approaches to Connes' embedding conjecture.

1. Introduction

Interest in positivity questions involving noncommutative polynomials has been recently revived by Helton's seminal paper [2002], in which he proved that a polynomial is a sum of squares if and only if its values in matrices of any size are positive semidefinite. Considering polynomials with *positive trace*, Klep and Schweighofer [2008, Theorem 1.6] observed that Connes' embedding conjecture [1976, Section V, pp. 105–107] on type II_1 von Neumann algebras is equivalent to a problem of describing polynomials whose values at tuples of self-adjoint $d \times d$ matrices (of norm at most 1) have nonnegative trace for every $d \geq 1$. This result is the motivation for the present work. Here we investigate polynomials whose values at tuples of $d \times d$ matrices have nonnegative trace for a *fixed* $d \geq 1$. We show that such a polynomial is (after multiplication with a Hermitian square of a suitable central polynomial) a sum of commutators and of a polynomial whose values at tuples of $d \times d$ matrices are positive semidefinite. The latter were characterized in [Klep and Unger 2010], leading us to the following Positivstellensatz: *every polynomial with nonnegative trace on $d \times d$ matrices is modulo sums of commutators and*

Supported by the Slovenian Research Agency (project no. J1-3608 and program no. P1-0222).
MSC2000: primary 16W10, 13J30; secondary 11E25, 16R50.
Keywords: free algebra, noncommutative polynomial, central simple algebra, (reduced) trace, polynomial identity, involution, central polynomial, quadratic form, free positivity.

polynomial identities for $d \times d$ matrices a sum of Hermitian squares with weights and denominators. See Section 4 for a precise formulation.

The organization of this paper is as follows: Section 2 introduces the main notions and interprets them in full matrix algebras, Section 3 considers these notions for free algebras, while Section 4 presents our main results.

2. Basic notions and a motivating example

Let R be an associative ring with 1 and *involution* $a \mapsto a^*$ (that is, $(a+b)^* = a^*+b^*$, $(ab)^* = b^*a^*$ and $a^{**} = a$ for all $a, b \in R$). We denote by $\operatorname{Sym} R := \{a \in R \mid a = a^*\}$ its set of *symmetric* elements. Elements of the form a^*a and $ab - ba$ ($a, b \in A$) are called *Hermitian squares* and *commutators*, respectively. We introduce an equivalence relation (*cyclic equivalence*) on R by declaring $a \overset{\text{cyc}}{\sim} b$ if and only if $a - b$ is a sum of commutators in R. For notational convenience we write

$$\Sigma^2 R := \Big\{\sum a_i^* a_i \mid a_i \in R\Big\} \subseteq \operatorname{Sym} R, \qquad \Theta^2 R := \Big\{a \in R \mid \exists\, b \in \Sigma^2 R : a \overset{\text{cyc}}{\sim} b\Big\}$$

for the sets of (finite) sums of Hermitian squares, and sums of Hermitian squares and commutators in R, respectively.

Throughout this paper k will denote $\mathbb{R}$ or $\mathbb{C}$.

Matrices. For a concrete example of these notions, consider the ring $R = \mathrm{M}_d(k)$ of real or complex square matrices of a fixed size $d \geq 1$ endowed with the usual (complex conjugate) transposition of matrices, denoted here by $*$. Using $\succeq$ to denote the Löwner partial order (that is, $A \succeq B$ if and only if $A - B$ is positive semidefinite), it is easy to see that for $A \in \mathrm{M}_d(k)$, we have

(A) $A \succeq 0$ if and only if $A \in \Sigma^2 \mathrm{M}_d(k)$;

(B) $\operatorname{tr}(A) = 0$ if and only if $A \overset{\text{cyc}}{\sim} 0$ in $\mathrm{M}_d(k)$;

(C) $\operatorname{tr}(A) \geq 0$ if and only if $A \in \Theta^2 \mathrm{M}_d(k)$.

Let us determine multiplication by which matrices respect these properties.

Lemma 2.1. *Suppose $A \in \mathrm{M}_d(k)$ is such that for all $B \in \mathrm{M}_d(k)$,*

$$B \succeq 0 \quad \Rightarrow \quad AB \succeq 0. \tag{1}$$

Then $A = \lambda$ for some $\lambda \in \mathbb{R}_{\geq 0}$.

Proof. Using (1) with $B = 1$, we obtain $A \succeq 0$. In particular, $A = A^*$. Again by (1), A commutes with all positive semidefinite matrices, hence with all symmetric matrices, which are differences of two positive semidefinite matrices by

$$B = \tfrac{1}{4}(B+1)^2 - \tfrac{1}{4}(B-1)^2.$$

So A is scalar and the desired conclusion follows. $\square$

Lemma 2.2. *Suppose $A \in \mathrm{M}_d(k)$ is such that for all $B \in \mathrm{M}_d(k)$,*

$$\operatorname{tr}(B) = 0 \quad \Rightarrow \quad \operatorname{tr}(AB) = 0. \tag{2}$$

Then $A = \lambda$ for some $\lambda \in k$.

Proof. Write $A = [a_{ij}]_{i,j=1}^{d}$. Let $i \neq j$. Then $B = \lambda E_{ij}$ has zero trace for every $\lambda \in k$. (Here E_{ij} denotes the $d \times d$ *matrix unit* with a one in position (i, j) and zeros elsewhere.) By (2), this implies that $\lambda a_{ij} = \operatorname{tr}(AB) = 0$. Since $\lambda \in k$ was arbitrary, $a_{ij} = 0$.

Now let $B = \lambda(E_{ii} - E_{jj})$. Clearly, $\operatorname{tr}(B) = 0$ and hence

$$\lambda(a_{ii} - a_{jj}) = \operatorname{tr}(AB) = 0.$$

As before, this gives $a_{ii} = a_{jj}$. □

Lemma 2.3. *Suppose $A \in \mathrm{M}_d(k)$ is such that for all $B \in \mathrm{M}_d(k)$,*

$$\operatorname{tr}(B) \geq 0 \quad \Rightarrow \quad \operatorname{tr}(AB) \geq 0. \tag{3}$$

Then $A = \lambda$ for some $\lambda \in \mathbb{R}_{\geq 0}$.

Proof. By Lemma 2.2, A is scalar. In addition to that, $a_{ii} = \operatorname{tr}(AE_{ii}) \geq 0$ by (3), showing that A must be a nonnegative multiple of the identity. □

Likewise we can characterize matrices that map positive semidefinite matrices into matrices with nonnegative trace:

Lemma 2.4. *Suppose $A \in \mathrm{M}_d(k)$ is such that for all $B \in \mathrm{M}_d(k)$,*

$$B \succeq 0 \quad \Rightarrow \quad \operatorname{tr}(AB) \geq 0. \tag{4}$$

In the case $k = \mathbb{R}$, assume moreover that $A = A^*$. Then $A \succeq 0$.

Proof. This is just a restatement of the well-known self-duality of the cone of all positive semidefinite matrices. For $v \in k^d$, let $B = vv^* \succeq 0$. Then

$$0 \leq \operatorname{tr}(AB) = \operatorname{tr}(Avv^*) = \operatorname{tr}(v^*Av) = \langle Av, v\rangle,$$

showing A is positive semidefinite. □

Converses of Lemmas 2.1–2.4 hold as well.

3. Positivity in free algebras

Words and polynomials. Fix $n \in \mathbb{N}$. Let

$$\underline{X} := (X_1, \dots, X_n) \quad \text{and} \quad \underline{X}^* := (X_1^*, \dots, X_n^*)$$

denote tuples of n distinct variables (or letters). By $\langle \underline{X}, \underline{X}^*\rangle$ we denote the free monoid on $\{\underline{X}, \underline{X}^*\}$ (consisting of *words* in $\underline{X}, \underline{X}^*$) and let $k\langle \underline{X}, \underline{X}^*\rangle$ be the semigroup algebra of $\langle \underline{X}, \underline{X}^*\rangle$ over k (consisting of *polynomials* in noncommuting variables $\underline{X}$ and $\underline{X}^*$ with coefficients in k). We endow $k\langle \underline{X}, \underline{X}^*\rangle$ with the involution $p \mapsto p^*$ mapping $X_j \mapsto X_j^*$ and extending complex conjugation on k. Thus $k\langle \underline{X}, \underline{X}^*\rangle$ is the free $*$-algebra on $\underline{X}$ over k.

Cyclic equivalence. It is well known and easy to see that trace-zero matrices are sums of commutators, that is, cyclically equivalent to 0. Cyclic equivalence can also be easily tested in $k\langle \underline{X}, \underline{X}^*\rangle$:

(a) For $v, w \in \langle \underline{X}, \underline{X}^*\rangle$, we have $v \overset{\text{cyc}}{\sim} w$ if and only if there are $v_1, v_2 \in \langle \underline{X}, \underline{X}^*\rangle$ such that $v = v_1 v_2$ and $w = v_2 v_1$. That is, $v \overset{\text{cyc}}{\sim} w$ if and only if w is a cyclic permutation of v.

(b) Polynomials

$$f = \sum_{w \in \langle \underline{X}, \underline{X}^*\rangle} a_w w \quad \text{and} \quad g = \sum_{w \in \langle \underline{X}, \underline{X}^*\rangle} b_w w \quad \text{for } a_w, b_w \in k$$

are cyclically equivalent if and only if for each $v \in \langle \underline{X}, \underline{X}^*\rangle$,

$$\sum_{\substack{w \in \langle \underline{X}, \underline{X}^*\rangle \\ w \overset{\text{cyc}}{\sim} v}} a_w = \sum_{\substack{w \in \langle \underline{X}, \underline{X}^*\rangle \\ w \overset{\text{cyc}}{\sim} v}} b_w. \tag{5}$$

Evaluations and representations. Let $d \in \mathbb{N}$. An n-tuple of matrices $\underline{A} \in (\mathrm{M}_d(k))^n$ gives rise to a *$*$-representation*

$$\mathrm{ev}_{\underline{A}} : k\langle \underline{X}, \underline{X}^*\rangle \to \mathrm{M}_d(k), \quad p \mapsto p(\underline{A}, \underline{A}^*). \tag{6}$$

We are interested in the values of a *fixed* element $f \in k\langle \underline{X}, \underline{X}^*\rangle$ under all these $*$-representations. If the size d of the matrices A_i is free, we talk about *dimension-free* properties; otherwise we call them *dimension-dependent*. We are mostly interested in the latter, but briefly review the former for the sake of completeness.

Dimension-freeness. Free analogs of properties (A) and (B) have been established, while a free version of (C) is closely related to an important open problem on operator algebras due to Connes; see below for further details.

Let $f \in \operatorname{Sym} k\langle \underline{X}, \underline{X}^*\rangle$.

$(\mathrm{A})^{\mathrm{fr}}$ $f(\underline{A}, \underline{A}^*) \succeq 0$ for all $d \in \mathbb{N}$ and all $\underline{A} \in \mathrm{M}_d(k)^n$ if and only if $f \in \Sigma^2 k\langle \underline{X}, \underline{X}^*\rangle$;

$(\mathrm{B})^{\mathrm{fr}}$ $\operatorname{tr}\big(f(\underline{A}, \underline{A}^*)\big) = 0$ for all $d \in \mathbb{N}$ and all $\underline{A} \in \mathrm{M}_d(k)^n$ if and only if $f \overset{\text{cyc}}{\sim} 0$ in $k\langle \underline{X}, \underline{X}^*\rangle$.

Part $(\mathrm{A})^{\mathrm{fr}}$ is due to Helton [2002] (see also [McCullough 2001; McCullough and Putinar 2005]), and $(\mathrm{B})^{\mathrm{fr}}$ is Theorem 2.1 of [Klep and Schweighofer 2008]. (This reference will henceforth be abbreviated as [KS 2008].) See also [Collins and Dykema 2008, Lemma 2.9] for a proof inspired by free probability. For a recent study of trace-positive polynomials in a dimension-free setting see also [Netzer and Thom 2010].

The obvious extension of (C) fails: there are $f \in \operatorname{Sym} k\langle \underline{X}, \underline{X}^*\rangle$ with positive trace everywhere, but still not cyclically equivalent to a sum of Hermitian squares. The following is a variant of the noncommutative Motzkin polynomial from Example 4.4 of [KS 2008] given in free (nonsymmetric) variables.

Example 3.1. Let X denote a single free variable and set

$$M_0 := \\ 3X^4 - 3(XX^*)^2 - 4X^5X^* - 2X^3X^{*3} + 2X^2X^*XX^{*2} + 2X^2X^{*2}XX^* + 2(XX^*)^3.$$

Then the noncommutative Motzkin polynomial is

$$M := 1 + M_0 + M_0^* \in \operatorname{Sym} k\langle X, X^*\rangle.$$

It is trace-nonnegative everywhere since

$$M' := YZ^4Y + ZY^4Z - 3YZ^2Y + 1 \overset{\text{cyc}}{\sim} M\left(\frac{Y + \mathrm{i}Z}{2}, \frac{Y - \mathrm{i}Z}{2}\right) \in k\langle Y, Z\rangle$$

is trace-nonnegative on symmetric matrices; see Example 4.4 of [KS 2008]. Alternatively, $M(X^3, (X^*)^3) \in \Theta^2 k\langle X, X^*\rangle$. On the other hand, $M \notin \Theta^2 k\langle X, X^*\rangle$. (Some of these computations were done with the aid of the computer algebra systems NCSOStools [Cafuta et al. 2010] and NCAlgebra [Helton et al. 2010].)

Connes' embedding conjecture [1976, Section V, pp. 105–107] states that every separable II_1-factor is embeddable in an ultrapower of the hyperfinite II_1-factor. Understanding trace-positive polynomials in the dimension-free setting is the key to this problem, because it is equivalent, by Theorem 1.6 of [KS 2008], to Conjecture 1.5 of the same reference, which we repeat here for convenience:

Conjecture 3.2 (algebraic version of Connes' conjecture). For $f \in \operatorname{Sym} k\langle \underline{X}, \underline{X}^*\rangle$ the following are equivalent:

(i) $\operatorname{tr}\big(f(\underline{A}, \underline{A}^*)\big) \geq 0$ for all $d \in \mathbb{N}$ and all tuples of contractions $\underline{A} \in \mathrm{M}_d(k)^n$;

(ii) for every $\varepsilon \in \mathbb{R}_{>0}$, $f + \varepsilon$ is cyclically equivalent to an element of the form

$$\sum_j s_j^* s_j + \sum_{i,j} p_{ij}^*(1 - X_i^* X_i) p_{ij},$$

where $s_j, p_{ij} \in k\langle \underline{X}, \underline{X}^*\rangle$.

In the sequel we indicate an approach to this problem "from below". That is, we abandon the dimension-free setting and solve a Hilbert 17-type problem characterizing polynomials with nonnegative trace in a dimension-dependent setting. It is our belief that this might constitute an important step towards (a positive or negative resolution of) Connes' embedding conjecture.

4. Dimension-dependent positivity

The properties (A) and (B) for free algebras in a dimension-dependent setting are well understood due to our previous work. Roughly speaking, a trace-zero polynomial is cyclically equivalent to a polynomial identity [Brešar and Klep 2009, Section 4], and a positive semidefinite polynomial is a sum of Hermitian squares with denominators and weights [Klep and Unger 2010, Section 5]. In this section property (C) is explored and we present our main result, a Positivstellensatz characterizing polynomials with nonnegative trace on all tuples of $d \times d$ matrices for *fixed* d. This is done in Section 4C. Before that we recall generic matrices and universal division algebras with involution in Section 4A and take a look at polynomial preservers of the various notions of positivity in Section 4B.

4A. ***Generic matrices and universal division algebras.*** We assume the reader is familiar with the theory of polynomial identities as presented, e.g., in [Procesi 1973; Rowen 1980]. We review the notion of generic matrices and universal division algebras with involution and refer the reader to [Procesi 1976; Procesi and Schacher 1976] for details.

Let $\zeta := (\zeta_{ij}^{(\ell)} \mid 1 \le i, j \le d,\ 1 \le \ell \le n)$ and $\bar{\zeta} := (\bar{\zeta}_{ij}^{(\ell)} \mid 1 \le i, j \le d,\ 1 \le \ell \le n)$ denote commuting variables. To keep the notation uniform, let

$$\underline{\zeta} := \begin{cases} \zeta & \text{if } k = \mathbb{R}, \\ (\zeta, \bar{\zeta}) & \text{if } k = \mathbb{C}. \end{cases}$$

Form the polynomial $*$-algebra $k[\underline{\zeta}]$ that endowed with the involution that extends complex conjugation on k and fixes $\zeta_{ij}^{(\ell)}$ pointwise (if $k = \mathbb{R}$) or sends $\zeta_{ij}^{(\ell)}$ to $\bar{\zeta}_{ij}^{(\ell)}$ (if $k = \mathbb{C}$).

Consider the $d \times d$ matrices

$$Y_\ell := \big[\zeta_{ij}^{(\ell)}\big]_{1 \le i, j \le d} \in \mathrm{M}_d(k[\underline{\zeta}]) \quad \text{for } \ell \in \mathbb{N}.$$

Each Y_ℓ is called a *generic matrix*. The (unital) k-subalgebra of $\mathrm{M}_d(k[\underline{\zeta}])$ generated by the Y_ℓ and their (complex conjugate) transposes is the *ring of generic matrices with involution* $\mathrm{GM}_d(k)$. Equivalently,

$$\mathrm{GM}_d(k) \cong k\langle \underline{X}, \underline{X}^* \rangle / \mathfrak{t}_d,$$

where $\mathfrak{t}_d \subseteq k\langle \underline{X}, \underline{X}^* \rangle$ is the T-ideal of polynomial identities for $d \times d$ matrices.

For $d \ge 2$, the ring $\mathrm{GM}_d(k)$ is a prime PI algebra (see [Procesi and Schacher 1976, Section II]). Hence its central localization is a central simple algebra $\mathrm{UD}_d(k)$ with involution, which we call (by an abuse of notation) the *universal division algebra*. Relating these notions to $*$-representations of the free $*$-algebra is the following commutative diagram: for $d \in \mathbb{N}$ and $\underline{A} \in \mathrm{M}_d(k)^n$, let $R_{\underline{A}}$ denote all the elements of $\mathrm{UD}_d(k)$ that are regular at $\underline{A}$. Then:

$$\begin{tikzcd}
k\langle \underline{X}, \underline{X}^*\rangle \arrow[rr, "\mathrm{ev}_{\underline{A}}"] \arrow[dd, "\pi"'] & & \mathrm{M}_d(k) \\
& & R_{\underline{A}} \arrow[u] \arrow[d, hook] \\
\mathrm{GM}_d(k) \arrow[rr, hook, "\iota"] \arrow[urr, hook] & & \mathrm{UD}_d(k)
\end{tikzcd}$$

For a more geometric viewpoint of the ring of generic matrices and the universal division algebra we refer the reader to [Procesi 1976; Saltman 1999]. The standard textbook on central simple algebras with involution is [Knus et al. 1998].

4B. ***Polynomial preservers.*** In this subsection we present versions of Lemmas 2.1–2.4 in the context of free $*$-algebras. To avoid trivialities, we assume throughout that $d \geq 2$.

Lemma 4.1. *Suppose $f \in k\langle \underline{X}, \underline{X}^*\rangle$ is such that for all $g \in k\langle \underline{X}, \underline{X}^*\rangle$,*

$$g \succeq 0 \text{ on } d \times d \text{ matrices} \quad \Rightarrow \quad fg \succeq 0 \text{ on } d \times d \text{ matrices.} \tag{7}$$

Then f is a central polynomial positive semidefinite on $d \times d$ matrices.

Proof. Using (7) with $g = 1$, we see f is positive semidefinite on $d \times d$ matrices. Thus there is no harm in assuming $f = f^*$.

Again by (7), $fg - gf$ vanishes on all $d \times d$ matrices for all polynomials g of the form $g = h^*h$. That is, $[f, g]$ is a polynomial identity of $d \times d$ matrices. Now the same holds true for all symmetric g, since

$$2[f, g] + [f, g^2] = [f, (1+g)^2]$$

is then a polynomial identity. Hence f commutes (modulo the T-ideal of identities) with all symmetric polynomials.

Every element of $\mathrm{UD}_d(k)$ can be represented as rs^{-1} for some $r, s \in \mathrm{GM}_d(k)$ with $s = s^* \in Z(\mathrm{GM}_d(k))$. Such an element is symmetric if and only if $r = r^*$. So $\pi(f)$ commutes with all symmetric elements of $\mathrm{UD}_d(k)$. By Dieudonné's theorem [1952, Lemma 1], the latter generate $\mathrm{UD}_d(k)$. Hence $\pi(f) \in Z(\mathrm{UD}_d(k))$ and f is indeed a central polynomial.

(Note: once we have established that f commutes with all symmetric polynomials, an easier argument is available if $k = \mathbb{C}$. In this case one immediately obtains that f also commutes with all skew symmetric polynomials as these are all of the form $\mathbb{i}g$ for symmetric g.) $\square$

Lemma 4.2. *Suppose $f \in k\langle \underline{X}, \underline{X}^*\rangle$ is such that for all $g \in k\langle \underline{X}, \underline{X}^*\rangle$,*

$$\mathrm{tr}(g) = 0 \text{ on } d \times d \text{ matrices} \quad \Rightarrow \quad \mathrm{tr}(fg) = 0 \text{ on } d \times d \text{ matrices.} \tag{8}$$

Then f is a central polynomial.

Proof. Let $g = [h_1, h_2]$ for some $h_i \in k\langle \underline{X}, \underline{X}^*\rangle$. Then

$$fg = f[h_1, h_2] = [f, h_1h_2] + [h_1, fh_2] + h_1[h_2, f]. \tag{9}$$

Since $\mathrm{tr}(g) = 0$ on all $d \times d$ matrices, this implies $\mathrm{tr}(h_1[h_2, f]) = 0$ on $d \times d$ matrices. Fix h_2 and denote $r := [h_2, f]$. Then r satisfies

$$\mathrm{tr}(pr) = 0 \text{ on } d \times d \text{ matrices}$$

for all $p \in k\langle \underline{X}, \underline{X}^*\rangle$. Taking $p = -r^*$ leads to $-\mathrm{tr}(r^*r) = 0$, and hence $r = 0$ on all $d \times d$ matrices. That is, r is an identity of $d \times d$ matrices. As $r = [h_2, f]$ and h_2 was arbitrary, this implies f is a central polynomial. □

Lemma 4.3. *Suppose $f \in k\langle \underline{X}, \underline{X}^*\rangle$ is such that for all $g \in k\langle \underline{X}, \underline{X}^*\rangle$,*

$$\mathrm{tr}(g) \geq 0 \text{ on } d \times d \text{ matrices} \quad \Rightarrow \quad \mathrm{tr}(fg) \geq 0 \text{ on } d \times d \text{ matrices.} \tag{10}$$

Then f is a central polynomial positive semidefinite on $d \times d$ matrices.

Proof. If $\mathrm{tr}(g) = 0$, then by (10), $\mathrm{tr}(fg) \geq 0$ and $\mathrm{tr}(-fg) \geq 0$ on $d \times d$ matrices. That is, $\mathrm{tr}(fg) = 0$. Now by Lemma 4.2, f is a central polynomial.

Applying (10) with $g=1$ yields $f(\underline{A}, \underline{A}^*) = \mathrm{tr}\big(f(\underline{A}, \underline{A}^*)\big) \geq 0$ for all $\underline{A} \in \mathrm{M}_d(k)^n$, showing f is positive semidefinite on $d \times d$ matrices. □

Likewise we can characterize polynomials that map positive semidefinite polynomials into trace-nonnegative ones. At the same time this indicates how to build examples of trace-nonnegative polynomials. As we shall see in the next subsection, the procedure is essentially exhaustive.

Lemma 4.4. *Suppose $f \in \mathrm{Sym}\, k\langle \underline{X}, \underline{X}^*\rangle$ is such that for all $g \in k\langle \underline{X}, \underline{X}^*\rangle$,*

$$g \succeq 0 \text{ on } d \times d \text{ matrices} \quad \Rightarrow \quad \mathrm{tr}(fg) \geq 0 \text{ on } d \times d \text{ matrices.} \tag{11}$$

Then f is positive semidefinite on $d \times d$ matrices.

Proof. Assume f is not positive semidefinite on $d \times d$ matrices. Then there exists an n-tuple $\underline{A} = (A_1, \dots, A_n) \in \mathrm{M}_d(k)^n$ with

$$f(\underline{A}, \underline{A}^*) \not\succeq 0. \tag{12}$$

Let $\mathscr{A} \subseteq \mathrm{M}_d(k)$ denote the $*$-subalgebra generated by the $A_1, \dots, A_n$. Since the Hermitian square of a nonzero matrix is not nilpotent, $\mathscr{A}$ is semisimple. By the Artin–Wedderburn theorem, $\mathscr{A}$ is $*$-isomorphic to a direct sum of full matrix algebras. We distinguish two cases.

CASE 1: If $k = \mathbb{C}$, there is a $*$-isomorphism

$$\mathscr{A} \cong \bigoplus_{j=1}^{s} \mathrm{M}_{d_j}(\mathbb{C}) \tag{13}$$

for some $d_j \in \mathbb{C}$, and $\sum_j d_j \le d$. This induces a block diagonalization

$$A_j = \begin{bmatrix} A_{j,1} & & \\ & \ddots & \\ & & A_{j,s} \end{bmatrix}, \quad \text{with } A_{j,k} \in \mathrm{M}_{d_k}(\mathbb{C}).$$

By (12), there is a j such that $\underline{A}_{(j)} = (A_{1,j}, \dots, A_{n,j}) \in \mathrm{M}_{d_j}(\mathbb{C})^n$ satisfies

$$f(\underline{A}_{(j)}, \underline{A}^*_{(j)}) \not\succeq 0.$$

Choose $u \in \mathbb{C}^{d_j}$ with

$$\big\langle f(\underline{A}_{(j)}, \underline{A}^*_{(j)})u, u\big\rangle < 0. \tag{14}$$

There is a $B \in \mathrm{M}_{d_j}(\mathbb{C})$ with $Be_{i,d_j} = u$ for all $i = 1, \dots, d_j$. (Here e_{i,d_j} are the standard basis vectors for $\mathbb{C}^{d_j}$.) By the construction of $\mathscr{A}$ and (13), there is an $h \in \mathbb{C}\langle \underline{X}, \underline{X}^*\rangle$ with $h(\underline{A}_{(j)}, \underline{A}^*_{(j)}) = B$. Let $g = hh^*$. Then

$$\begin{aligned} \operatorname{tr}((fg)(\underline{A}_{(j)}, \underline{A}^*_{(j)})) &= \operatorname{tr}((h^* f h)(\underline{A}_{(j)}, \underline{A}^*_{(j)})) \\ &= \sum_{i=1}^{d_j} \big\langle h^*(\underline{A}_{(j)}, \underline{A}^*_{(j)}) f(\underline{A}_{(j)}, \underline{A}^*_{(j)}) h(\underline{A}_{(j)}, \underline{A}^*_{(j)}) e_{i,d_j}, e_{i,d_j}\big\rangle \\ &= \sum_{i=1}^{d_j} \big\langle f(\underline{A}_{(j)}, \underline{A}^*_{(j)}) Be_{i,d_j}, Be_{i,d_j}\big\rangle \\ &= \sum_{i=1}^{d_j} \big\langle f(\underline{A}_{(j)}, \underline{A}^*_{(j)})u, u\big\rangle < 0. \end{aligned} \tag{15}$$

As this contradicts our assumption (11), we conclude that $f \succeq 0$ on $d \times d$ matrices.

CASE 2: If $k = \mathbb{R}$, the reasoning is the same with a minor technical modification. Let

$$\mathscr{A} \cong \bigoplus_{j=1}^{s} \mathrm{M}_{d_j}(\mathbb{R}) \oplus \bigoplus_{k=1}^{r} \mathrm{M}_{e_k}(\mathbb{C}) \oplus \bigoplus_{\ell=1}^{p} \mathrm{M}_{f_\ell}(\mathbb{H}) \tag{16}$$

for some $d_j, e_k, f_\ell \in \mathbb{N}$.

If there is a tuple $\underline{A} \in \mathrm{M}_{d_j}(\mathbb{R})^n$ with $f(\underline{A}, \underline{A}^*) \not\succeq 0$, we proceed as in Case 1. If there is an $\underline{A} \in \mathrm{M}_{e_k}(\mathbb{C})^n$ with $0 \not\preceq f(\underline{A}, \underline{A}^*) \in \mathrm{M}_{e_k}(\mathbb{C})$, we proceed as follows. Let V be the invariant subspace of $\mathbb{R}^d$ corresponding to the action of $\mathrm{M}_{e_k}(\mathbb{C})$. There is a $u \in V$ with $\langle f(\underline{A}, \underline{A}^*)u, u\rangle < 0$. Pick a basis $\{v_1, \dots, v_{e_k}\}$ of V over $\mathbb{C}$, and let

$B \in M_{e_k}(\mathbb{C})$ satisfy $Bv_j = u$ for all j. Choose $h \in \mathbb{R}\langle \underline{X}, \underline{X}^* \rangle$ with $h(\underline{A}, \underline{A}^*) = B$ and $g = hh^*$. Then the *complex* trace z of $(fg)(\underline{A}, \underline{A}^*)$ is negative by the same computation as in (15). Hence the *real* trace satisfies

$$\operatorname{tr}\big((fg)(\underline{A}, \underline{A}^*)\big) = \frac{z+\bar{z}}{2} < 0.$$

The remaining case of quaternion matrices is dealt with similarly. We leave this as an exercise for the reader. □

It is clear that converses of Lemmas 4.1–4.4 hold true. Also, with the exception of (11), which is satisfied when f is a sum of Hermitian squares, there are no nonconstant dimension-free polynomial preservers.

4C. ***The dimension-dependent tracial Positivstellensatz.*** Our main tool for describing trace-nonnegative polynomials is the following proposition deduced from the properties of the reduced trace [Knus et al. 1998, Section 1] on $\mathrm{UD}_d(k)$.

Proposition 4.5. *For every $f \in k\langle \underline{X}, \underline{X}^* \rangle$ and $d \in \mathbb{N}$ there exists a nonvanishing central polynomial for $d \times d$ matrices, denoted by $c \in k\langle \underline{X}, \underline{X}^* \rangle$, such that cf is cyclically equivalent to a central polynomial. That is,*

$$cf \overset{\text{cyc}}{\sim} c' \tag{17}$$

for some central polynomial c'.

Proof. Consider $F := \iota(\pi(f)) \in \mathrm{UD}_d(k)$. So $\mathrm{Trd}(F) \in Z(\mathrm{UD}_d(k))$, and there is a nonvanishing central polynomial $c_0 \in k\langle \underline{X}, \underline{X}^* \rangle$ and a central polynomial c_0' with

$$\mathrm{Trd}(F) = \pi(c_0')\pi(c_0)^{-1}. \tag{18}$$

Since Trd is $Z(\mathrm{UD}_d(k))$-linear, this yields $\mathrm{Trd}(\pi(c_0 f - c_0')) = 0$. By [Amitsur and Rowen 1994, Theorem 2.4], $\pi(c_0 f - c_0') \overset{\text{cyc}}{\sim} 0$ in $\mathrm{UD}_d(k)$. Clearing denominators shows

$$\pi(cf - c'') \overset{\text{cyc}}{\sim} 0 \tag{19}$$

in $\mathrm{GM}_d(k)$ for a nonvanishing central polynomial c and a central polynomial c''. Lifting (19) to $k\langle \underline{X}, \underline{X}^* \rangle$ gives the desired conclusion: $cf \overset{\text{cyc}}{\sim} c'$. □

Remark 4.6. Instead of the Amitsur–Rowen result used in this proof, we can apply the tracial Nullstellensatz [Brešar and Klep 2009, Theorem 5.2]: once we have established that $\mathrm{Trd}(\pi(c_o f - c_0')) = 0$, by clearing denominators we obtain $\operatorname{tr}(\pi(c_0 c'' f - c_0' c'')) = 0$ for some nonvanishing central polynomial c''. Hence $\pi(c_0 c'' f - c' c'') \overset{\text{cyc}}{\sim} 0$ in $\mathrm{GM}_d(k)$ by [Brešar and Klep 2009, Theorem 5.2]. As before, lifting this relation to $k\langle \underline{X}, \underline{X}^* \rangle$ yields the desired conclusion.

We are now ready to give our main results characterizing trace-nonnegative polynomials.

Theorem 4.7. *Let* $k \in \{\mathbb{R}, \mathbb{C}\}$ *and suppose* $f \in \operatorname{Sym} k\langle \underline{X}, \underline{X}^*\rangle$ *satisfies*

$$\operatorname{tr}\big(f(\underline{A}, \underline{A}^*)\big) \geq 0 \tag{20}$$

for all $\underline{A} \in \mathrm{M}_d(k)^n$. *Then there is a nonvanishing central polynomial for* $d \times d$ *matrices, denoted by* $c \in k\langle \underline{X}, \underline{X}^*\rangle$, *such that* cfc^* *is cyclically equivalent to a polynomial* $g \in k\langle \underline{X}, \underline{X}^*\rangle$ *that is positive semidefinite on* $d \times d$ *matrices*:

$$cfc^* \overset{\text{cyc}}{\sim} g \quad \textit{and} \quad g \succeq 0 \textit{ on } d \times d \textit{ matrices}. \tag{21}$$

Proof. This is a consequence of Proposition 4.5. Indeed, there is a nonvanishing central polynomial c with

$$cf \overset{\text{cyc}}{\sim} c' \tag{22}$$

for a central polynomial c'. Multiplying (22) with c^* (from the right) shows

$$cfc^* \overset{\text{cyc}}{\sim} c'c^*. \tag{23}$$

For any $\underline{A} \in \mathrm{M}_d(k)^n$,

$$\begin{aligned} 0 &\leq \operatorname{tr}\big(c(\underline{A}, \underline{A}^*)f(\underline{A}, \underline{A}^*)c(\underline{A}, \underline{A}^*)^*\big) = \operatorname{tr}\big(c'(\underline{A}, \underline{A}^*)c(\underline{A}, \underline{A}^*)^*\big) \\ &= \operatorname{tr}\big((c'c^*)(\underline{A}, \underline{A}^*)\big) = (c'c^*)(\underline{A}, \underline{A}^*). \end{aligned} \tag{24}$$

So $g := c'c^*$ is a (central) polynomial positive semidefinite on $d \times d$ matrices satisfying

$$cfc^* \overset{\text{cyc}}{\sim} g.$$ □

Remark 4.8. The proof shows that g in Theorem 4.7 can actually be taken to be a central polynomial.

Combining Theorem 4.7 with the dimension-dependent Positivstellensatz for positive semidefinite polynomials ([Procesi and Schacher 1976, Theorem 5.4] or [Klep and Unger 2010, Theorem 5.4]) yields:

Corollary 4.9. *Choose* $\alpha_1, \ldots, \alpha_m \in k\langle \underline{X}, \underline{X}^*\rangle$ *whose images in* $\mathrm{GM}_d(k)$ *form a diagonalization of the quadratic form* $\operatorname{Trd}(x^*x)$ *on* $\mathrm{UD}_d(k)$. *For* $f \in \operatorname{Sym} k\langle \underline{X}, \underline{X}^*\rangle$, *the following are equivalent*:

(i) $\operatorname{tr}\big(f(\underline{A}, \underline{A}^*)\big) \geq 0$ *for every* $\underline{A} \in \mathrm{M}_d(k)^n$.

(ii) *There exists a nonvanishing central polynomial* $c \in k\langle \underline{X}, \underline{X}^*\rangle$, *a polynomial identity* $h \in k\langle \underline{X}, \underline{X}^*\rangle$ *for* $d \times d$ *matrices, and* $p_{i,\varepsilon} \in k\langle \underline{X}, \underline{X}^*\rangle$ *with*

$$cfc^* \overset{\text{cyc}}{\sim} h + \sum_{\varepsilon \in \{0,1\}^m} \underline{\alpha}^{\varepsilon} \sum_i p_{i,\varepsilon}^* p_{i,\varepsilon}. \tag{25}$$

Remark 4.10. For experts we mention that, by applying the reduced trace, we can reformulate (25) as

$$cfc^* \overset{\text{cyc}}{\sim} h + t, \tag{26}$$

where c and h are as above, and t belongs to the preordering in $Z(\mathrm{UD}_d(k))$ generated by the α_j.

If $d = 2$, the weights α_j are superfluous since the reduced trace of a Hermitian square is a sum of Hermitian squares in this case (see [Procesi and Schacher 1976, p. 405] or [Klep and Unger 2010, Section 4]), and Corollary 4.9 simplifies as follows:

Corollary 4.11. *For $f \in \operatorname{Sym} k\langle \underline{X}, \underline{X}^*\rangle$, the following are equivalent*:

(i) $\operatorname{tr}\big(f(\underline{A}, \underline{A}^*)\big) \geq 0$ *for every* $\underline{A} \in \mathrm{M}_2(k)^n$.

(ii) *There exists a nonvanishing central polynomial $c \in k\langle \underline{X}, \underline{X}^*\rangle$, and a polynomial identity $h \in k\langle \underline{X}, \underline{X}^*\rangle$ for 2×2 matrices*, *such that*

$$cfc^* \in h + \Theta^2\, k\langle \underline{X}, \underline{X}^*\rangle. \tag{27}$$

Example 4.12. We finish this presentation with an example showing denominators are necessary for these results to hold. First, the Motzkin polynomial M from Example 3.1 is not cyclically equivalent to a sum of Hermitian squares modulo a T-ideal of identities. Indeed, suppose that

$$M \overset{\text{cyc}}{\sim} h + \sum g_j^* g_j \tag{28}$$

for some $g_j \in k\langle \underline{X}, \underline{X}^*\rangle$ and a polynomial identity $h \in k\langle \underline{X}, \underline{X}^*\rangle$ for $d \times d$ matrices ($d \geq 2$). Then

$$M_{\mathrm{cc}} = \operatorname{tr}\left(M\left(\begin{bmatrix} Y/2 & Z/2 \\ -Z/2 & Y/2 \end{bmatrix}\right)\right) = \sum \operatorname{tr}\left((g_j^* g_j)\left(\begin{bmatrix} Y/2 & Z/2 \\ -Z/2 & Y/2 \end{bmatrix}\right)\right),$$

where $M_{\mathrm{cc}} \in \mathbb{R}[Y, Z]$ denotes the commutative collapse $Y^4Z^2 + Y^2Z^4 - 3Y^2Z^2 + 1$ of the noncommutative variant M' of the Motzkin polynomial (in symmetric variables). Since M_{cc} is not a sum of squares in $\mathbb{R}[Y, Z]$, and the trace of a Hermitian square is a sum of squares, M does not satisfy a relation of the form (28). Hence a denominator is needed in Corollaries 4.9 and 4.11.

A little more work is required to show the necessity of the denominator in Theorem 4.7. Let $d \in \mathbb{N}$ be sufficiently large (at least 127, the dimension of the vector space of all polynomials in X, X^* of degree at most 6). Suppose M is cyclically equivalent to a polynomial g that is positive semidefinite on $d \times d$ matrices. Without loss of generality, $g \in \operatorname{Sym} k\langle \underline{X}, \underline{X}^*\rangle$. Choose g of the smallest possible degree. If this degree is greater than 6, then the highest homogeneous component $g^{(\infty)}$ of g is positive semidefinite on $d \times d$ matrices and at the same time $g^{(\infty)} \overset{\text{cyc}}{\sim} 0$. Hence $\operatorname{tr}(g^{(\infty)}) = 0$ on $d \times d$ matrices, implying that $g^{(\infty)}$ is a polynomial identity. Then

$$M \overset{\text{cyc}}{\sim} (g - g^{(\infty)}),$$

with $g - g^{(\infty)}$ positive semidefinite and of degree smaller than g. This contradicts the minimality of g, so $\deg(g) \le 6$.

Now g is positive semidefinite on $d \times d$ matrices for some $d \ge 127$ and is thus a sum of Hermitian squares by Helton's sum of squares theorem [2002]. But M is not cyclically equivalent to a sum of Hermitian squares by the first part of this example.

Acknowledgments

The author thanks Matej Brešar for enlightening conversations and insightful comments.

References

[Amitsur and Rowen 1994] S. A. Amitsur and L. H. Rowen, "Elements of reduced trace 0", *Israel J. Math.* **87**:1-3 (1994), 161–179. MR 95h:16019 Zbl 0852.16012

[Brešar and Klep 2009] M. Brešar and I. Klep, "Values of noncommutative polynomials, Lie skew-ideals and tracial Nullstellensätze", *Math. Res. Lett.* **16**:4 (2009), 605–626. MR 2010m:16061 Zbl 1189.16021

[Cafuta et al. 2010] K. Cafuta, I. Klep, and J. Povh, "NCSOStools: a computer algebra system for symbolic and numerical computation with noncommutative polynomials", 2010, available at http://www.optimization-online.org/DB_HTML/2010/05/2635.html. To appear in *Optim. Methods Softw.* Software available at http://ncsostools.fis.unm.si/.

[Collins and Dykema 2008] B. Collins and K. Dykema, "A linearization of Connes' embedding problem", *New York J. Math.* **14** (2008), 617–641. MR 2010a:46141 Zbl 1162.46032

[Connes 1976] A. Connes, "Classification of injective factors: Cases II_1, II_∞, III_λ, $\lambda \neq 1$", *Ann. Math.* (2) **104**:1 (1976), 73–115. MR 56 #12908 Zbl 0343.46042

[Dieudonné 1952] J. Dieudonné, "On the structure of unitary groups", *Trans. Amer. Math. Soc.* **72**:3 (1952), 367–385. MR 14,134c Zbl 0046.25301

[Helton 2002] J. W. Helton, "'Positive' noncommutative polynomials are sums of squares", *Ann. Math.* (2) **156**:2 (2002), 675–694. MR 2003k:12002 Zbl 1033.12001

[Helton et al. 2010] J. W. H. Helton, M. C. de Oliveira, R. L. Miller, and M. Stankus, "NCAlgebra: a Mathematica package for doing non-commuting algebra", January 2010, available at http://www.math.ucsd.edu/~ncalg.

[Klep and Schweighofer 2008] I. Klep and M. Schweighofer, "Connes' embedding conjecture and sums of Hermitian squares", *Adv. Math.* **217**:4 (2008), 1816–1837. MR 2009g:46109 Zbl 1184.46055

[Klep and Unger 2010] I. Klep and T. Unger, "The Procesi–Schacher conjecture and Hilbert's 17th problem for algebras with involution", *J. Algebra* **324**:2 (2010), 256–268. MR 2651356 Zbl 05768597

[Knus et al. 1998] M.-A. Knus, A. S. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, Amer. Math. Soc. Colloquium Publ. **44**, AMS, Providence, RI, 1998. MR 2000a:16031 Zbl 0955.16001

[McCullough 2001] S. McCullough, "Factorization of operator-valued polynomials in several non-commuting variables", *Linear Algebra Appl.* **326** (2001), 193–203. MR 1815959 Zbl 0980.47024

[McCullough and Putinar 2005] S. McCullough and M. Putinar, "Noncommutative sums of squares", *Pacific J. Math.* **218**:1 (2005), 167–171. MR 2007j:47026 Zbl 1177.47020

[Netzer and Thom 2010] T. Netzer and A. Thom, "Tracial algebras and an embedding theorem", *J. Funct. Anal.* **259**:11 (2010), 2939–2960. MR 2719281 Zbl 05809715

[Procesi 1973] C. Procesi, *Rings with polynomial identities*, Pure and Applied Mathematics **17**, Marcel Dekker, New York, 1973. MR 51 #3214 Zbl 0262.16018

[Procesi 1976] C. Procesi, "The invariant theory of $n \times n$ matrices", *Adv. Math.* **19**:3 (1976), 306–381. MR 54 #7512 Zbl 0331.15021

[Procesi and Schacher 1976] C. Procesi and M. Schacher, "A non-commutative real Nullstellensatz and Hilbert's 17th problem", *Ann. Math.* (2) **104**:3 (1976), 395–406. MR 55 #5599 Zbl 0347.16010

[Rowen 1980] L. H. Rowen, *Polynomial identities in ring theory*, Pure and Applied Mathematics **84**, Academic Press, New York, 1980. MR 82a:16021 Zbl 0461.16001

[Saltman 1999] D. J. Saltman, *Lectures on division algebras*, CBMS Regional Conference Series in Mathematics **94**, AMS, Providence, RI, 1999. MR 2000f:16023 Zbl 0934.16013

Received December 31, 2009.

IGOR KLEP
UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
JADRANSKA 21
SI-1111 LJUBLJANA
SLOVENIA

and

UNIVERZA V MARIBORU
FAKULTETA ZA NARAVOSLOVJE IN MATEMATIKO
KOROŠKA 160
SI-2000 MARIBOR
SLOVENIA

igor.klep@fmf.uni-lj.si

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

REMARKS ON THE PRODUCT OF HARMONIC FORMS

LIVIU ORNEA AND MIHAELA PILCA

A metric is formal if all products of harmonic forms are again harmonic. The existence of a formal metric implies Sullivan formality of the manifold, and hence formal metrics can exist only in the presence of a very restricted topology. We show that a warped product metric is formal if and only if the warping function is constant and derive further topological obstructions to the existence of formal metrics. In particular, we determine the necessary and sufficient conditions for a Vaisman metric to be formal.

1. Introduction

A fundamental problem in algebraic topology is the reading of the homotopy type of a space in terms of cohomological data. A precise definition of this property was given by Sullivan [1977] and called *formality*. As concerns manifolds, it is known, for example, that all compact Riemannian symmetric spaces and all compact Kähler manifolds are formal. For a recent survey of topological formality, see [Papadima and Suciu 2009].

Sullivan also observed that if a compact manifold admits a metric such that the wedge product of any two harmonic forms is again harmonic, then, by Hodge theory, the manifold is formal. This motivated the following definition:

Definition 1.1 [Kotschick 2001]. A closed manifold is called *geometrically formal* if it admits a formal Riemannian metric.

In particular, the length of any harmonic form with respect to a formal metric is (pointwise) constant. This larger class of metrics having all harmonic (one-)forms of constant length naturally appears in other geometric contexts, for instance in the study of certain systolic inequalities, and has been investigated in [Nagy 2006; Nagy and Vernicos 2004].

Classical examples of geometrically formal manifolds are compact symmetric spaces. In [Kotschick and Terzić 2003; 2011] more general examples are provided,

Both authors are partially supported by CNCSIS grant PNII IDEI contract 529/2009. M. Pilca also acknowledges partial support from SFB/TR 12.
MSC2000: primary 53C25; secondary 53C55, 58A14.
Keywords: formality, harmonic form, warped product, Vaisman manifold, Betti numbers.

both of geometrically formal and of formal but nongeometrically formal homogeneous manifolds.

Geometric formality imposes strong restrictions on the (real) cohomology of the manifold. For example, it is proven in [Kotschick 2001] that a manifold admits a nonformal metric if and only if it is not a rational homology sphere.

In this note, we shall obtain further obstructions to formality. We shall see (Section 2) that if a compact manifold with $b_1 = p \geq 1$ admits a formal metric, and if there exist two vanishing Betti numbers such that the distance between them is not larger than $p+2$, then all the intermediary Betti numbers must be zero too. Also, a conformal class of metrics on an even-dimensional compact manifold with nonzero middle Betti number can contain no more than one formal metric.

Our main concern will be the formality of warped products (Section 2). We will show that a warped product metric on a compact manifold is formal if and only if the warping function is constant. On the way, we shall also provide a proof for the fact (stated in [Kotschick 2001], for instance) that a product of formal metrics is formal.

Unlike Kähler manifolds, which are known to be formal, for the time being, nothing is known about the Sullivan formality of locally conformally Kähler (in particular Vaisman) manifolds. In Section 3 of this note, we shall discuss compact Vaisman manifolds, whose universal cover is a special type of warped product, a Riemannian cone to be precise, and we shall find obstructions to the metric formality of a Vaisman metric. Several computational facts and their proofs are gathered in the Appendix.

2. Geometric formality of warped product metrics

For completeness, and as a first step in the study of geometrically formal warped products, we provide a proof for the formality of Riemannian product formal metrics.

Proposition 2.1. *If (M_1, g_1) and (M_2, g_2) are two compact Riemannian manifolds with formal metrics, then the metric $g = g_1 + g_2$ on the product manifold $M = M_1 \times M_2$ is also formal.*

Proof. Let $\gamma \in \Omega^p M$ and $\gamma' \in \Omega^q M$ be two harmonic forms on M. By Lemma A.2, γ and γ' are given by linear combinations with real coefficients of the basis elements in (A-3). Thus, it is enough to check that the exterior product of any two such basis elements is a harmonic form on M. But

$$\big(\pi_1^*(\alpha) \wedge \pi_2^*(\beta)\big) \wedge \big(\pi_1^*(\alpha') \wedge \pi_2^*(\beta')\big) = (-1)^{|\alpha'||\beta|} \pi_1^*(\alpha \wedge \alpha') \wedge \pi_2^*(\beta \wedge \beta'),$$

which is g-harmonic on M by Lemma A.2 and by the formality of g_1 and g_2 (as $\alpha \wedge \alpha'$ is again a g_1-harmonic form and $\beta \wedge \beta'$ a g_2-harmonic form). $\square$

We now pass to the setting we are mainly interested in, warped products.

Theorem 2.2. *Let (B^n, g_B) and (F^m, g_F) be two compact Riemannian manifolds with formal metrics. Then the warped product metric $g = \pi^*(g_B) + (\varphi \circ \pi)^2 \sigma^*(g_F)$ on $B \times_\varphi F$ is formal if and only if the warping function φ is constant.*

Proof. Let $\beta \in \Omega^p(F)$ be a g_F-harmonic form on F (as $b_m(F) = 1$, there exists at least a harmonic m-form on F). From the equalities (A-4) in the Appendix, it follows that $\sigma^*\beta$ is a g-harmonic form on the warped product $B \times_\varphi F$. If we assume the warped metric g to be formal, it follows in particular that the length of $\sigma^*\beta$ is constant. As g_F is also assumed to be formal, the length of β is constant as well. On the other hand,

$$g(\sigma^*\beta, \sigma^*\beta) = (\varphi \circ \pi)^{2p} g_F(\beta, \beta) \circ \sigma, \tag{2-1}$$

showing that the function φ must be constant.

Conversely, if φ is constant, then the warped product reduces to the Riemannian product between the Riemannian manifolds (B, g_B) and $(F, \varphi^2 g_F)$, which is geometrically formal by Proposition 2.1. □

Remark 2.3. From the above proof we see that Theorem 2.2 holds more generally for metrics having all harmonic forms of constant length.

An interesting question regarding the formal metrics is their existence in a given conformal class. Under a weak topological assumption, we prove that there may exist at most one such formal metric. More precisely, we have

Proposition 2.4. *Let M^{2n} be an even-dimensional compact manifold whose middle Betti number $b_n(M)$ is nonzero. Then, in any conformal class of metrics there is at most one formal metric* (*up to homothety*).

Proof. Let $[g]$ be a class of conformal metrics on M and suppose there are two formal metrics g_1 and $g_2 = e^{2f} g_1$ in $[g]$. The main observation is that in the middle dimension the kernel of the codifferential is invariant at conformal changes of the metric, so that there are the same harmonic forms for all metrics in a conformal class: $\mathcal{H}^n(M, g_1) = \mathcal{H}^n(M, g_2)$. As $b_n(M) \geq 1$ there exists a nontrivial g_1-harmonic (and thus also g_2-harmonic) n-form α on M. The length of α must then be constant with respect to both metrics, which are assumed to be formal and thus we get

$$g_2(\alpha, \alpha) = e^{2nf} g_1(\alpha, \alpha),$$

which shows that f must be constant. □

Using the product construction to ensure that the middle Betti number is nonzero, one can build such examples of formal metrics which are unique in their conformal class.

Other examples are provided by manifolds with "big" first Betti number, as follows from the following property of "propagation" of Betti numbers on geometrically formal manifolds proven in [Kotschick 2001, Theorem 7]: *if* $b_1(M) = p \geq 1$, *then* $b_q(M) \geq \binom{p}{q}$, *for all* $1 \leq q \leq p$. In particular, if $b_1(M^{2n}) \geq n$, then $b_n(M^{2n}) \geq 1$.

Another property of the Betti numbers of geometrically formal manifolds is this:

Proposition 2.5. *Let* M^n *be a compact geometrically formal manifold such that* $b_1(M) = p \geq 1$. *If there exist two vanishing Betti numbers* $b_k(M) = b_{k+l}(M) = 0$, *for some* k *and* l *with* $0 < k + l < n$ *and* $0 < l \leq p + 1$, *then all intermediary Betti numbers must vanish*: $b_i(M) = 0$, *for* $k \leq i \leq k + l$. *In particular*, *if there exists* $k \geq (n - p - 1)/2$ *such that* $b_k(M) = 0$, *then* $b_i(M) = 0$ *for all* $k \leq i \leq n - k$.

Proof. Let $\{\theta_1, \ldots, \theta_p\}$ be an orthogonal basis of g-harmonic 1-forms, where g is a formal metric on M. We first notice that here is no ambiguity in considering the orthogonality with respect to the global scalar product or to the pointwise inner product, because, when restricting ourselves to the space of harmonic forms of a formal metric, these notions coincide. This is mainly due to [Kotschick 2001, Lemma 4], which states that the inner product of any two harmonic forms is a constant function. Thus, if two harmonic forms α and β are orthogonal with respect to the global product, we get

$$0 = (\alpha, \beta) = \int_M \langle \alpha, \beta \rangle \, d\mathrm{vol}_g = \langle \alpha, \beta \rangle \, \mathrm{vol}(M),$$

showing that their pointwise inner product is the zero-function.

It is enough to show that $b_{k+1}(M) = 0$ and then use induction on i. Let α be a harmonic $(k + 1)$-form. By formality, $\theta_1 \wedge \theta_2 \wedge \cdots \wedge \theta_{l-1} \wedge \alpha$ is a harmonic $(k + l)$-form and thus must vanish, since $b_{k+l}(M) = 0$. On the other hand,

$$\theta_j^\sharp \lrcorner \alpha = (-1)^{k(n-k-1)} * (\theta_j \wedge *\alpha)$$

is a harmonic k-form, again by formality. Since $b_k(M) = 0$, it follows that $\theta_j^\sharp \lrcorner \alpha$ vanishes for $1 \leq j \leq p$. Then, since $\{\theta_1, \ldots, \theta_p\}$ are also orthogonal, we obtain

$$0 = \theta_1^\sharp \lrcorner \cdots \lrcorner \theta_{l-1}^\sharp \lrcorner (\theta_1 \wedge \cdots \wedge \theta_{l-1} \wedge \alpha) = \pm |\theta_1|^2 \cdots |\theta_{l-1}|^2 \alpha,$$

which implies that $\alpha = 0$, because each θ_j has nonzero constant length. This shows that $b_{k+1}(M) = 0$. □

3. Geometric formality of Vaisman metrics

A Vaisman manifold is a particular type of locally conformal Kähler (LCK) manifold. It is defined as a Hermitian manifold (M, J, g), of real dimension $n = 2m \geq 4$, whose fundamental 2-form ω satisfies the conditions

$$d\omega = \theta \wedge \omega, \quad \nabla\theta = 0.$$

Here θ is a (closed) 1-form, called the Lee form, and ∇ is the Levi-Civita connection of the LCK metric g (we always consider $\theta \neq 0$, to not include the Kähler manifolds among the Vaisman ones).

Locally, $\theta = df$ and the *local* metric $e^{-f} g$ is Kähler, hence the name LCK. When lifted to the universal cover, these local metrics glue to a global one, which is Kähler and acted on by homotheties by the deck group of the covering.

In the Vaisman case, the universal cover is a Riemannian cone. In fact, compact Vaisman manifolds are closely related to Sasakian ones, as the following structure theorem shows:

Theorem 3.1 [Ornea and Verbitsky 2003]. *Compact Vaisman manifolds are mapping tori over S^1. More precisely, the universal cover $\tilde{M}$ is a metric cone $N \times \mathbb{R}^{>0}$, with N compact Sasakian manifold and the deck group is isomorphic with $\mathbb{Z}$, generated by*

$$(x, t) \mapsto (\lambda(x), t + q)$$

for some $\lambda \in \mathrm{Aut}(N)$, $q \in \mathbb{R}^{>0}$.

This puts compact Vaisman manifolds into the framework of warped products and motivates their consideration here.

Vaisman manifolds are abundant. Any Hopf manifold (quotient of $\mathbb{C}^{\mathbb{N}} \setminus \{0\}$ by the cyclic group generated by a semisimple operator with subunitary eigenvalues) is such, as are its compact complex submanifolds [Verbitsky 2004, Proposition 6.5]. A complete list of compact Vaisman surfaces is given in [Belgun 2000].

On the other hand, examples of LCK manifolds (satisfying only the condition $d\omega = \theta \wedge \omega$ for a closed θ) which cannot admit any Vaisman metric are also known: for example, one type of Inoue surface and the nondiagonal Hopf surface; see [Belgun 2000]. The nondiagonal Hopf surface is particularly relevant for our discussion because it is topologically formal, as are all manifolds having the same cohomology ring as a product of odd spheres.

Being parallel and Killing [Dragomir and Ornea 1998], the Lee field $\theta^\sharp$ is real holomorphic and, together with $J\theta^\sharp$, generates a complex one-dimensional totally geodesic Riemannian foliation $\mathcal{F}$. Note that $\mathcal{F}$ is transversally Kähler, meaning that the transversal part of the Kähler form is closed (for a proof of this result, see [Vaisman 1982, Theorem 3.1]).

In the sequel, the terms *basic* (*foliate*) and *horizontal* refer to $\mathcal{F}$. We recall that a form is called *horizontal* with respect to a foliation $\mathcal{F}$ if its interior product with any vector field tangent to the foliation vanishes and is called *basic* if in addition its Lie derivative along a vector field tangent to the foliation also vanishes. Moreover, we shall use the basic versions of the standard operators acting on $\Omega^*_B(M)$, the space of basic forms: Δ_B is the basic Laplace operator, L_B is the exterior multiplication with the transversal Kähler form and Λ_B its adjoint with respect to the transversal metric. For details on these operators and their properties we refer the reader to [Tondeur 1988, Chapter 12].

Here is the main result of this section. It puts severe restrictions on formal Vaisman metrics.

Theorem 3.2. *Let (M^{2m}, g, J) be a compact Vaisman manifold. The metric g is geometrically formal if and only if $b_p(M) = 0$ for*

$$2 \le p \le 2m-2, \quad b_1(M) = b_{2m-1}(M) = 1,$$

that is, M is a cohomological Hopf manifold.

Proof. Let $\gamma \in \Omega^p(M)$ be a harmonic form on M for some p, $1 \le p \le m-1$. By [Vaisman 1982, Theorem 4.1], γ has the form

$$\gamma = \alpha + \theta \wedge \beta, \tag{3-1}$$

with α and β basic, transversally harmonic and transversally primitive.

Since α is basic, $J\alpha$ is also a basic p-form that is transversally harmonic and transversally primitive:

$$\Delta_B(J\alpha) = 0, \quad \Lambda_B(J\alpha) = 0,$$

because Δ_B and Λ_B both commute with the transversal complex structure J (as the foliation is transversally Kähler). Again from the theorem just cited, by taking $\beta = 0$, it follows that $J\alpha$ is a harmonic form on M: $\Delta(J\alpha) = 0$.

The assumption that g is geometrically formal implies that $\alpha \wedge J\alpha$ is harmonic on M, so that in particular it is coclosed: $\delta(\alpha \wedge J\alpha) = 0$. By [Vaisman 1982] (where the term *transversally effective* is used instead of transversally primitive), this implies that $\alpha \wedge J\alpha$ is transversally primitive: $\Lambda_B(\alpha \wedge J\alpha) = 0$.

Otherwise, by [Grosjean and Nagy 2009, Proposition 2.2], for primitive forms $\eta, \mu \in \Lambda^p V$, where (V, g, J) is any Hermitian vector space, the algebraic relation

$$(\Lambda)^p(\eta \wedge \mu) = (-1)^{(p(p-1))/2} p\langle \eta, J\mu \rangle, \tag{3-2}$$

holds, where J is the extension of the complex structure to $\Lambda^* V$ defined by

$$(J\eta)(v_1, \dots, v_p) := \eta(Jv_1, \dots, Jv_p), \quad \text{for all } \eta \in \Lambda^p V, v_1, \dots, v_p \in V.$$

We apply the formula above to the transversal Kähler geometry and conclude that α vanishes everywhere:

$$0 = (\Lambda_B)^p(\alpha \wedge J\alpha) = (-1)^{(p(p+1))/2} p\langle \alpha, \alpha \rangle.$$

The same argument as above applied to $\beta \in \Omega_B^{p-1}(M)$ shows that β is identically zero if $p \ge 2$. Thus, $\gamma = 0$ for $2 \le p \le m-1$, which proves that

$$b_2(M) = \cdots = b_{m-1}(M) = 0.$$

If $p = 1$, then β is a basic function, which is transversally harmonic, so that β is a constant. Thus γ is a multiple of θ, showing that the space of harmonic 1-forms on M is 1-dimensional: $b_1(M) = 1$.

It remains to show that the Betti number in the middle dimension, $b_m(M)$, also vanishes. This follows from Proposition 2.5 applied to $p = 1$, $k = m - 1$ and $l = 2$.

The converse is clear, since the space of harmonic forms with respect to the Vaisman metric g is spanned by $\{1, \theta, *\theta, d\mathrm{vol}_g\}$ and thus the only product of harmonic forms which is not trivial is $\theta \wedge *\theta = g(\theta, \theta)\, d\mathrm{vol}_g$, which is harmonic because θ has constant length, being a parallel 1-form. □

Remark 3.3. (i) There exist Vaisman manifolds that do not admit any formal Vaisman metric. Indeed, let $f : N \hookrightarrow \mathbb{C}P^n$ be an embedded curve of genus $g > 1$ and let M be the total space of the induced Hopf bundle $f^*(S^1 \times S^{2n+1})$. Then M is Vaisman and $b_1(M) > 1$ [Vaisman 1982], hence, according to 3.2, it does not admit any formal Vaisman metric. Other examples can be found in [Belgun 2000].

(ii) On the other hand, we do not have an example of a topologically formal complex compact manifold, which admits Vaisman metrics, but does not admit geometrically formal Vaisman metrics. This seems to be a difficult open problem.

(iii) In complex dimension 2 the Vaisman condition in Theorem 3.2 is not necessary. Due to the results of Kotschick [2001], the existence of any geometrically formal metric on a non-Kähler surface implies that $b_1 = 1$ and $b_2 = 0$.

(iv) Theorem 3.2 may be considered as an analogue of the following result on the geometric formality of Sasakian manifolds.

Theorem 3.4 [Grosjean and Nagy 2009, Theorem 2.1]. *Let (M^{2n+1}, g) be a compact Sasakian manifold. If the metric g is geometrically formal, then $b_p(M) = 0$ for $1 \le p \le 2n$, that is, M is a real cohomology sphere.*

Appendix: Auxiliary results

Lemma A.1 (characterization of geometric formality). *Let α and β be two harmonic forms on a compact Riemannian manifold (M^n, g). Then $\alpha \wedge \beta$ is harmonic if and only if*

$$\text{(A-1)} \qquad \sum_{i=1}^{n} (e_i \lrcorner \alpha) \wedge \nabla_{e_i}\beta = -(-1)^{|\alpha||\beta|} \sum_{i=1}^{n} (e_i \lrcorner \beta) \wedge \nabla_{e_i}\alpha,$$

where $\{e_i\}_{i=\overline{1,n}}$ is a local orthonormal basis of vector fields. Thus, the metric g is formal if and only if (A-1) *holds for any two g-harmonic forms.*

Proof. Since M is compact, $\alpha \wedge \beta$ is harmonic if and only if it is closed and coclosed. As $\alpha \wedge \beta$ is closed, we have to show that (A-1) is equivalent to $\alpha \wedge \beta$ being coclosed. This is implied by the following:

$$
\begin{aligned}
&\delta(\alpha\wedge\beta)\\
&\quad=-\sum_{i=1}^{n} e_i\lrcorner\nabla_{e_i}(\alpha\wedge\beta)=-\sum_{i=1}^{n} e_i\lrcorner(\nabla_{e_i}\alpha\wedge\beta+\alpha\wedge\nabla_{e_i}\beta)\\
&\quad=\delta\alpha\wedge\beta-(-1)^{|\alpha|}\sum_{i=1}^{n}\nabla_{e_i}\alpha\wedge(e_i\lrcorner\beta)-\sum_{i=1}^{n}(e_i\lrcorner\alpha)\wedge\nabla_{e_i}\beta+(-1)^{|\alpha|}\alpha\wedge\delta\beta\\
&\quad=-(-1)^{|\alpha||\beta|}\sum_{i=1}^{n}(e_i\lrcorner\beta)\wedge\nabla_{e_i}\alpha-\sum_{i=1}^{n}(e_i\lrcorner\alpha)\wedge\nabla_{e_i}\beta. \qquad\square
\end{aligned}
$$

Riemannian products. Let $(M^{n+m}, g) = (M_1^n, g_1) \times (M_2^m, g_2)$. We denote by $\pi_i : M \to M_i$ the natural projections, which are totally geodesic Riemannian submersions.

One may describe the bundle of p-forms on M as follows:

$$\Lambda^p M = \bigoplus_{k=0}^{p} \pi_1^*(\Lambda^k M_1) \otimes \pi_2^*(\Lambda^{p-k} M_2). \tag{A-2}$$

This identification also works for the space of harmonic forms, namely the harmonic forms on (M, g) can be described in terms of the harmonic forms on the factors (M_1, g_1) and (M_2, g_2). To this end let $\mathcal{H}^k(M_i, g_i)$ be the space of harmonic k-forms on M_i and let $b_k(M_i)$ be the Betti numbers of M_i, $i = 1, 2$.

Lemma A.2. *Let $\{\alpha_1^k, \dots, \alpha_{b_k(M_1)}^k\}$ be a basis of $\mathcal{H}^k(M_1, g_1)$ and $\{\beta_1^k, \dots, \beta_{b_k(M_2)}^k\}$ a basis of $\mathcal{H}^k(M_2, g_2)$). Then the forms*

$$\bigl\{\pi_1^*(\alpha_s^k) \wedge \pi_2^*(\beta_l^{p-k}) \mid 1 \le s \le b_k(M_1),\ 1 \le l \le b_{p-k}(M_2),\ 0 \le k \le p\bigr\} \tag{A-3}$$

form a basis of the space of $\mathcal{H}^p(M, g)$, for each $0 \le p \le m+n$.

For a proof, see [Griffiths and Harris 1978, page 105].

Warped products. Let (B^n, g_B) and (F^m, g_F) be two Riemannian manifolds and $\varphi > 0$ be a smooth function on B. Then $M = B \times_\varphi F$ denotes the warped product with the metric $g = \pi^*(g_B) + (\varphi\circ\pi)^2\sigma^*(g_F)$, where $\pi : M \to B$ and $\sigma : M \to F$ are the natural projections.

Let $\{e_i\}_{i=\overline{1,n}}$ be a local orthonormal basis on B and let $\{f_j\}_{j=\overline{1,m}}$ be a local orthonormal basis on F, which we lift to M and thus obtain a local orthonormal basis of M:

$$\Bigl\{\tilde{e}_i, \frac{1}{\varphi\circ\pi}\tilde{f}_j\Bigr\}_{i=\overline{1,n};\, j=\overline{1,m}}.$$

Consider the decomposition $\delta = \delta_1 + \delta_2$ of the codifferential on M, where

$$\delta_1 := -\sum_{i=1}^{n} \tilde{e}_i \lrcorner \nabla_{\tilde{e}_i}, \quad \delta_2 := -\frac{1}{(\varphi\circ\pi)^2}\sum_{j=1}^{m} \tilde{f}_j \lrcorner \nabla_{\tilde{f}_j}.$$

We first determine the commutation relations between the pullback of forms on B and F with δ_1 and δ_2.

Lemma A.3. *For $\alpha \in \Omega^*(B)$ and $\beta \in \Omega^*(F)$, we have*

$$\delta_1(\sigma^*(\beta)) = 0, \qquad \delta_2(\sigma^*(\beta)) = \frac{1}{(\varphi\circ\pi)^2}\sigma^*(\delta^{g_F}(\beta)), \tag{A-4}$$

$$\delta_1(\pi^*(\alpha)) = \pi^*(\delta^{g_B}(\alpha)), \qquad \delta_2(\pi^*(\alpha)) = -\frac{m}{\varphi\circ\pi}\operatorname{grad}(\varphi\circ\pi)\lrcorner\pi^*(\alpha). \tag{A-5}$$

Proof. Let $\beta \in \Omega^{p+1}(F)$. For any tangent vector fields $X_1, \dots, X_p$ to M we obtain

$$\begin{aligned}
&\delta_1(\sigma^*(\beta))(X_1,\dots,X_p)\\
&\quad = -\sum_{i=1}^n (\tilde{e}_i \lrcorner \nabla_{\tilde{e}_i}(\sigma^*\beta))(X_1,\dots,X_p)\\
&\quad = -\sum_{i=1}^n \tilde{e}_i\big(\beta(\sigma_*\tilde{e}_i, \sigma_*X_1,\dots,\sigma_*X_p)\circ\sigma\big) + \sum_{i=1}^n \beta(\sigma_*(\nabla_{\tilde{e}_i}\tilde{e}_i), \sigma_*X_1,\dots,\sigma_*X_p)\\
&\qquad + \sum_{i=1}^n \Big(\beta\big(\sigma_*\tilde{e}_i, \sigma_*(\nabla_{\tilde{e}_i}X_1),\dots,\sigma_*X_p\big) + \dots + \beta\big(\sigma_*\tilde{e}_i, \sigma_*X_1,\dots,\sigma_*(\nabla_{\tilde{e}_i}X_p)\big)\Big)\\
&\quad = 0,
\end{aligned}$$

since $\sigma_*\tilde{e}_i = 0$, because $\tilde{e}_i$ is the lift of a vector field on B and also

$$\sigma_*(\nabla_{\tilde{e}_i}\tilde{e}_i) = \sigma_*(\widetilde{\nabla^{g_B}_{e_i} e_i}) = 0.$$

This proves that $\delta_1(\sigma^*(\beta)) = 0$.

The commutation rule in (A-4) is shown as follows:

$$\begin{aligned}
&(\varphi\circ\pi)^2\delta_2(\sigma^*(\beta))(X_1,\dots,X_p)\\
&\quad = -\sum_{j=1}^m (\tilde{f}_j \lrcorner \nabla_{\tilde{f}_j}(\sigma^*\beta))(X_1,\dots,X_p)\\
&\quad = -\sum_{j=1}^m \tilde{f}_j(\beta(\sigma_*\tilde{f}_j,\sigma_*X_1,\dots,\sigma_*X_p)\circ\sigma) + \sum_{j=1}^m \beta(\sigma_*(\nabla_{\tilde{f}_j}\tilde{f}_j),\sigma_*X_1,\dots,\sigma_*X_p)\circ\sigma\\
&\qquad + \sum_{j=1}^m \big(\beta(\sigma_*\tilde{f}_j, \sigma_*(\nabla_{\tilde{f}_j}X_1),\dots,\sigma_*X_p)\\
&\qquad\qquad + \dots + \beta(\sigma_*\tilde{f}_j, \sigma_*X_1,\dots,\sigma_*(\nabla_{\tilde{f}_j}X_p))\big)\circ\sigma\\
&\quad = -\sum_{j=1}^m f_j(\beta(f_j,\sigma_*X_1,\dots,\sigma_*X_p))\circ\sigma\\
&\qquad + \sum_{j=1}^m \beta(\sigma_*(\widetilde{\nabla^{g_F}_{f_j} f_j} - \frac{g(\tilde{f}_j,\tilde{f}_j)}{\varphi\circ\pi}\operatorname{grad}(\varphi\circ\pi)), \sigma_*X_1,\dots,\sigma_*X_p)\circ\sigma\\
&\qquad + \sum_{j=1}^m \big(\beta(f_j,\sigma_*(\nabla_{\tilde{f}_j}X_1),\dots,\sigma_*X_p) + \dots + \beta(f_j,\sigma_*X_1,\dots,\sigma_*(\nabla_{\tilde{f}_j}X_p))\big)\circ\sigma,
\end{aligned}$$

where we may again assume, without loss of generality, that X_i are lifts of vector fields Z_i on F: $X_i = \tilde{Z}_i$ for $i = 1, \dots, p$. For a tangent vector field Y to B, each of the above terms vanishes, since $\sigma_*(Y) = 0$. We then get

$$\begin{aligned}
&(\varphi \circ \pi)^2 \delta_2(\sigma^*(\beta))(X_1, \dots, X_p) \\
&\quad = -\sum_{j=1}^{m} f_j(\beta(f_j, Z_1, \dots, Z_p)) \circ \sigma + \sum_{j=1}^{m} \beta(\nabla^{g_F}_{f_j} f_j, Z_1, \dots, Z_p) \circ \sigma \\
&\qquad + \sum_{j=1}^{m} [\beta(f_j, \nabla^{g_F}_{f_j} Z_1, \dots, \sigma_* X_p) + \dots + \beta(f_j, Z_1, \dots, \nabla^{g_F}_{f_j} Z_p)] \circ \sigma \\
&\quad = \sigma^*(\delta^{g_F}(\beta))(X_1, \dots, X_p).
\end{aligned}$$

The relations (A-5) can be obtained by similar computations, which we omit here. □

Acknowledgement

We thank D. Kotschick and P.-A. Nagy for very useful comments on the first draft of this note and S. Papadima for a discussion about topological formality. We are also grateful to the referee for his suggestions which improved the structure of the paper and its presentation.

References

[Belgun 2000] F. A. Belgun, "On the metric structure of non-Kähler complex surfaces", *Math. Ann.* **317**:1 (2000), 1–40. MR 2002c:32027 Zbl 0988.32017

[Dragomir and Ornea 1998] S. Dragomir and L. Ornea, *Locally conformal Kähler geometry*, Progress in Mathematics **155**, Birkhäuser, Boston, 1998. MR 99a:53081 Zbl 0887.53001

[Griffiths and Harris 1978] P. Griffiths and J. Harris, *Principles of algebraic geometry*, Pure and Applied Mathematics, Wiley-Interscience, New York, 1978. MR 80b:14001 Zbl 0408.14001

[Grosjean and Nagy 2009] J.-F. Grosjean and P.-A. Nagy, "On the cohomology algebra of some classes of geometrically formal manifolds", *Proc. Lond. Math. Soc.* **98**:3 (2009), 607–630. MR 2010j:53417 Zbl 1166.53043

[Kotschick 2001] D. Kotschick, "On products of harmonic forms", *Duke Math. J.* **107**:3 (2001), 521–531. MR 2002c:53076 Zbl 1036.53030

[Kotschick and Terzić 2003] D. Kotschick and S. Terzić, "On formality of generalized symmetric spaces", *Math. Proc. Cambridge Philos. Soc.* **134**:3 (2003), 491–505. MR 2004c:53071 Zbl 1042.53035

[Kotschick and Terzić 2011] D. Kotschick and S. Terzić, "Geometric formality of homogeneous spaces and of biquotients", *Pacific J. Math.* **249**:1 (2011), 157–176. Zbl 05837651

[Nagy 2006] P.-A. Nagy, "On length and product of harmonic forms in Kähler geometry", *Math. Z.* **254**:1 (2006), 199–218. MR 2007g:53083

[Nagy and Vernicos 2004] P.-A. Nagy and C. Vernicos, "The length of harmonic forms on a compact Riemannian manifold", *Trans. Amer. Math. Soc.* **356**:6 (2004), 2501–2513. MR 2005f:58055 Zbl 1048.53023

[Ornea and Verbitsky 2003] L. Ornea and M. Verbitsky, "Structure theorem for compact Vaisman manifolds", *Math. Res. Lett.* **10**:5-6 (2003), 799–805. MR 2004j:53093 Zbl 1052.53051

[Papadima and Suciu 2009] S. Papadima and A. Suciu, "Geometric and algebraic aspects of 1-formality", *Bull. Math. Soc. Sci. Math. Roumanie* (*N.S.*) **52**:3 (2009), 355–375. MR 2010k:55018 Zbl 1199.55010

[Sullivan 1977] D. Sullivan, "Infinitesimal computations in topology", *Inst. Hautes Études Sci. Publ. Math.* 47 (1977), 269–331. MR 58 #31119 Zbl 0374.57002

[Tondeur 1988] P. Tondeur, *Foliations on Riemannian manifolds*, Springer, New York, 1988. MR 89e:53052 Zbl 0643.53024

[Vaisman 1982] I. Vaisman, "Generalized Hopf manifolds", *Geom. Dedicata* **13**:3 (1982), 231–255. MR 84g:53096 Zbl 0506.53032

[Verbitsky 2004] M. S. Verbitsky, "Theorems on the vanishing of cohomology for locally conformally hyper-Kähler manifolds", pp. 64–91 in Алгебраическая геометрия: методы, связи и приложения, Tr. Mat. Inst. Steklova **246**, 2004. In Russian; translated as pp. 54–78 in *Algebraic geometry: methods, relations, and applications*, Proc. Steklov Inst. Math. **246** (2004). MR 2005h:53071

Received January 19, 2010. Revised October 11, 2010.

LIVIU ORNEA
FACULTY OF MATHEMATICS
UNIVERSITY OF BUCHAREST
STR ACADEMIEI NR 14
RO-010014 BUCHAREST
ROMANIA

and

INSTITUTE OF MATHEMATICS "SIMION STOILOW" OF THE ROMANIAN ACADEMY
CALEA GRIVITEI NR 21
RO-010702 BUCHAREST
ROMANIA

liviu.ornea@imar.ro

MIHAELA PILCA
FAKULTÄT FÜR MATHEMATIK
UNIVERSITY OF REGENSBURG
UNIVERSITÄTSSTR. 31
D-93053 REGENSBURG
GERMANY

and

INSTITUTE OF MATHEMATICS "SIMION STOILOW" OF THE ROMANIAN ACADEMY
CALEA GRIVITEI NR 21
RO-010702 BUCHAREST
ROMANIA

Mihaela.Pilca@mathematik.uni-regensburg.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

STEINBERG REPRESENTATION OF GSp(4): BESSEL MODELS AND INTEGRAL REPRESENTATION OF L-FUNCTIONS

AMEYA PITALE

We obtain explicit formulas for the test vector in the Bessel model, and derive the criteria for existence and uniqueness of Bessel models for the unramified quadratic twists of the Steinberg representation π of $\mathrm{GSp}_4(F)$, where F is a nonarchimedean local field of characteristic zero. We also give precise criteria for the Iwahori spherical vector in π to be a test vector. We apply the formulas for the test vector to obtain an integral representation of the local L-function of π, twisted by any irreducible admissible representation of $\mathrm{GL}_2(F)$. Using results of Furusawa and of Pitale and Schmidt, we derive from this an integral representation for the global L-function of the irreducible cuspidal automorphic representation of $\mathrm{GSp}_4(\mathbb{A})$ obtained from a Siegel cuspidal Hecke newform, with respect to a Borel congruence subgroup of square-free level, twisted by any irreducible cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A})$. A special-value result for this L-function, in the spirit of Deligne's conjecture, is obtained.

1. Introduction

It is known that the representation of the symplectic group obtained from a Siegel modular form is nongeneric, which means that it does not have a Whittaker model. Consequently, one cannot use in this case the techniques or results for generic representations. In such a situation, one introduces the notion of a generalized Whittaker model, now called a Bessel model. These Bessel models have been used to obtain integral representations of L-functions. It is known that, if $\mathbb{A}$ is the ring of adeles of a number field, an automorphic representation of $\mathrm{GSp}_4(\mathbb{A})$ obtained from a Siegel modular form always has some global Bessel model. For the purposes of local calculations, it is often very important to know the precise criteria for the existence of local Bessel models and have explicit formulas. In this paper, we wish to investigate Bessel models for unramified quadratic twists of

MSC2000: primary 11F46; secondary 11F66, 11F67, 11F70.
Keywords: Steinberg representation, Siegel modular forms, L-functions, special values of L-functions.

the Steinberg representation π of $\mathrm{GSp}_4(F)$, where F is any nonarchimedean local field of characteristic zero.

We first briefly explain what a Bessel model is (detailed definitions will be given in Section 3). Let F be a nonarchimedean local field of characteristic zero. Let $U(F)$ be the unipotent radical of the Siegel parabolic subgroup of $\mathrm{GSp}_4(F)$, and θ be any nondegenerate character of $U(F)$. The group $\mathrm{GL}_2(F)$, embedded in the Levi subgroup of the Siegel parabolic subgroup, acts on $U(F)$ by conjugation and, hence, on characters of $U(F)$. Let $T(F) = \mathrm{Stab}_{\mathrm{GL}_2(F)}(\theta)$; then, $T(F)$ is isomorphic to the units of a quadratic algebra L over F. The group $R(F) = T(F)U(F)$ is called the Bessel subgroup of $\mathrm{GSp}_4(F)$ (depending on θ). Let Λ be any character of $T(F)$, and denote by $\Lambda \otimes \theta$ the character obtained on $R(F)$. Let (π, V) be any irreducible admissible representation of $\mathrm{GSp}_4(F)$. A linear functional $\beta : V \to \mathbb{C}$, satisfying $\beta(\pi(r)v) = (\Lambda \otimes \theta)(r)\beta(v)$ for any $r \in R(F)$ and $v \in V$, is called a (Λ, θ)-Bessel functional for π. We say that π has a (Λ, θ)-Bessel model if π is isomorphic to a subspace of smooth functions $B : \mathrm{GSp}_4(F) \to \mathbb{C}$ such that $B(rh) = (\Lambda \otimes \theta)(r)B(h)$ for all $r \in R(F)$ and $h \in \mathrm{GSp}_4(F)$. The existence of a nontrivial Bessel functional is equivalent to the existence of a Bessel model for a representation. If π has a nontrivial (Λ, θ)-Bessel functional β, then a vector $v \in V$ such that $\beta(v) \neq 0$ is called a *test vector* for β.

Prasad and Takloo-Bighash [2007] have obtained, for any irreducible admissible representation π of $\mathrm{GSp}_4(F)$, the criteria to be satisfied by Λ for the existence of a (Λ, θ)-Bessel functional for π. Their method involves the use of theta lifts and distributions. The uniqueness of Bessel functionals has been obtained in [Novo–dvorsky and Piatetski-Shapiro 1973] for many cases; in particular, for any π with a trivial central character. In [Sugano 1985], a test vector is obtained when both the representation π and the character Λ are unramified. In [Saha 2009], a test vector is obtained when $F = \mathbb{Q}_p$, where p is odd and inert in the quadratic field extension L corresponding to $T(\mathbb{Q}_p)$, the representation π is an unramified quadratic twist of the Steinberg representation, and Λ has conductor $1 + p\mathfrak{o}_L$. The explicit formulas of the test vector in the above two cases have been used in [Furusawa 1993; Saha 2009] to obtain an integral representation of the $\mathrm{GSp}_4 \times \mathrm{GL}_2$ L-function, where the GL_2 representation is either unramified or Steinberg.

The main goal of this paper is to obtain explicit formulas for a test vector, whenever a Bessel model for the unramified quadratic twist of the Steinberg representation of $\mathrm{GSp}_4(F)$ exists. In addition to obtaining these formulas, we in fact obtain an independent proof of the criteria for the existence *and* uniqueness of the Bessel models. We also give precise conditions on the character Λ, so that the Iwahori spherical vector in π is a test vector. This is achieved in:

Theorem 3.18. *Let $\pi = \Omega\,\mathrm{St}_{\mathrm{GSp}_4}$ be the Steinberg representation of* $\mathrm{GSp}(F)$, *twisted by an unramified quadratic character Ω. Let Λ be a character of $L^\times$*

such that $\Lambda|_{F^\times} \equiv 1$. *If* L *is a field*, *then* π *has a* (Λ, θ)*-Bessel model if and only if* $\Lambda \neq \Omega \circ N_{L/F}$. *If* L *is not a field*, *then* π *always has a* (Λ, θ)*-Bessel model. In case* π *has a* (Λ, θ)*-Bessel model*, *it is unique. In addition*, *if* π *has a* (Λ, θ)*-Bessel model*, *then the Iwahori spherical vector of* π *is a test vector for the Bessel functional if and only if*

i) Λ *is trivial on* $1 + \mathfrak{P}$ (see (2-1) for the definition of $\mathfrak{P}$), *and*

ii) *in case* $L = F \oplus F$ *and* Λ *is unramified*, *we have* $\Lambda((1, \varpi)) \neq \Omega(\varpi)$, *where* ϖ *is the uniformizer in the ring of integers of* F.

The criterion for the existence of the Bessel model obtained in this theorem is the same as in [Prasad and Takloo-Bighash 2007]. However, the methods used to prove it are very different from those in that paper and in [Novodvorsky and Piatetski-Shapiro 1973].

When the Iwahori spherical vector is a test vector, we use the explicit formula for the test vector to obtain in Theorem 4.3 an integral representation of the local L-function $L(s, \pi \times \tau)$ of the Steinberg representation π of $\mathrm{GSp}_4(F)$, twisted by *any* irreducible admissible representation τ of $\mathrm{GL}_2(F)$. This integral involves a function B in the Bessel model of π, and a Whittaker function $W^\#$ in a certain induced representation of GU(2, 2) related to τ. We wish to remark that, in this paper as well as in other works [Furusawa 1993; Pitale and Schmidt 2009b; 2009c; Saha 2009], the Bessel function B is always chosen to be a “distinguished” vector (spherical if π is unramified, and Iwahori spherical if π is Steinberg) that has the additional property of being a test vector. With this choice of B, we have a systematic way of choosing $W^\#$ (see [Pitale and Schmidt 2009c]) so that the integral is nonzero and gives an integral representation of the L-function. The work so far suggests that, to obtain an integral representation for the L-function with a general irreducible admissible representation π of $\mathrm{GSp}_4(F)$, we will have to choose B to be both a “distinguished” vector in the Bessel model of π *and* a test vector for the Bessel functional. This further highlights the importance of obtaining more information and explicit formulas for test vectors for Bessel models of $\mathrm{GSp}_4(F)$. This is a topic of ongoing work.

Using the local computation mentioned above, together with the archimedean and p-adic calculations from [Furusawa 1993; Pitale and Schmidt 2009c], we obtain in Theorem 5.2 an integral representation of the global L-function $L(s, \pi \times \tau)$ of an irreducible cuspidal automorphic representation π of $\mathrm{GSp}_4(\mathbb{A})$, obtained from a Siegel cuspidal newform with respect to the Borel congruence subgroup of square-free level, twisted by *any* irreducible cuspidal automorphic representation τ of $\mathrm{GL}_2(\mathbb{A})$. When τ corresponds to an elliptic cusp form in $S_l(N, \chi)$, we obtain in Theorem 5.3 algebraicity results for special values of the twisted L-function, in the spirit of Deligne’s conjecture [1979].

2. Steinberg representation of GSp_4

Nonarchimedean setup. Let F be a nonarchimedean local field of characteristic zero. Let $\mathfrak{o}$, $\mathfrak{p}$, ϖ, q be the ring of integers, prime ideal, uniformizer and cardinality of the residue class field $\mathfrak{o}/\mathfrak{p}$, respectively. We fix three elements $a, b, c \in F$ such that $d := b^2 - 4ac \neq 0$. Let

$$L = \begin{cases} F(\sqrt{d}) & \text{if } d \notin F^{\times 2}, \\ F \oplus F & \text{if } d \in F^{\times 2}. \end{cases}$$

In the case when $L = F \oplus F$, we consider F diagonally embedded. If L is a field, we denote by $\bar{x}$ the Galois conjugate of $x \in L$ over F. If $L = F \oplus F$, let $\overline{(x, y)} = (y, x)$. In every case, we let $N(x) = x\bar{x}$ and $\mathrm{tr}(x) = x + \bar{x}$. We shall assume that $a, b \in \mathfrak{o}$ and $c \in \mathfrak{o}^\times$. In addition, we assume that d is the generator of the discriminant of L/F if $d \notin F^{\times 2}$ and $d \in \mathfrak{o}^\times$ if $d \in F^{\times 2}$.

The Legendre symbol $\left(\frac{L}{\mathfrak{p}}\right)$ is set to

$$\left(\frac{L}{\mathfrak{p}}\right) = \begin{cases} -1 & \text{if } d \notin F^{\times 2} \text{ and } d \notin \mathfrak{p} \text{ (the inert case)}, \\ 0 & \text{if } d \notin F^{\times 2} \text{ and } d \in \mathfrak{p} \text{ (the ramified case)}, \\ 1 & \text{if } d \in F^{\times 2} \text{ (the split case)}. \end{cases}$$

If L is a field, then let $\mathfrak{o}_L$ be its ring of integers. If $L = F \oplus F$, then let $\mathfrak{o}_L = \mathfrak{o} \oplus \mathfrak{o}$. Let ϖ_L be the uniformizer of $\mathfrak{o}_L$ if L is a field, and set $\varpi_L = (\varpi, 1)$ if L is not a field. Note that, if $\left(\frac{L}{\mathfrak{p}}\right) \neq -1$, then $N(\varpi_L) \in \varpi\mathfrak{o}^\times$. Let $\alpha \in \mathfrak{o}_L$ be defined by

$$\alpha := \begin{cases} \dfrac{b + \sqrt{d}}{2c} & \text{if } L \text{ is a field}, \\ \left(\dfrac{b + \sqrt{d}}{2c}, \dfrac{b - \sqrt{d}}{2c}\right) & \text{if } L = F \oplus F. \end{cases}$$

We fix in $\mathfrak{o}_L$ the ideal

$$\mathfrak{P} := \mathfrak{p}\mathfrak{o}_L = \begin{cases} \mathfrak{p}_L & \text{if } \left(\frac{L}{\mathfrak{p}}\right) = -1, \\ \mathfrak{p}_L^2 & \text{if } \left(\frac{L}{\mathfrak{p}}\right) = 0, \\ \mathfrak{p} \oplus \mathfrak{p} & \text{if } \left(\frac{L}{\mathfrak{p}}\right) = 1. \end{cases} \tag{2-1}$$

Here, when L is a field extension, $\mathfrak{p}_L$ is the maximal ideal of $\mathfrak{o}_L$. Note that $\mathfrak{P}$ is prime only if $\left(\frac{L}{\mathfrak{p}}\right) = -1$. We have

$$\mathfrak{P}^n \cap \mathfrak{o} = \mathfrak{p}^n \quad \text{for all } n \geq 0.$$

Lemma 2.1 [Pitale and Schmidt 2009b, Lemma 3.1.1]. *With the notation above, the elements* 1 *and* α *constitute an integral basis of* L/F. *There does not exists any* $x \in \mathfrak{o}$ *such that* $\alpha + x \in \mathfrak{P}$.

Steinberg representation. We define the symplectic group $H = \mathrm{GSp}_4$ by

$$H(F) := \{g \in \mathrm{GL}_4(F) : {}^t g J g = \mu_2(g) J,\ \mu_2(g) \in F^\times\}, \quad \text{where } J = \begin{bmatrix} & 1_2 \\ -1_2 & \end{bmatrix}.$$

The maximal compact subgroup is denoted by

$$K^H := \mathrm{GSp}_4(\mathfrak{o}).$$

We define the Iwahori subgroup by

$$\mathrm{I} := \left\{ g \in K^H : g \equiv \begin{bmatrix} * & 0 & * & * \\ * & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix} \pmod{\mathfrak{p}} \right\}.$$

Let Ω be an unramified quadratic character of $F^\times$. Let π be the Steinberg representation of $H(F)$, twisted by the character Ω. This representation is denoted by $\Omega\,\mathrm{St}_{\mathrm{GSp}_4}$. Since we have assumed that Ω is quadratic, we see that π has trivial central character. The Steinberg representation has the property that it is the only representation of $H(F)$ which has a unique (up to a constant) Iwahori fixed vector. The Iwahori Hecke algebra acts on the space of I-invariant vectors. We will next describe the Iwahori Hecke algebra.

Iwahori Hecke algebra. The Iwahori Hecke algebra $\mathscr{H}_{\mathrm{I}}$ of $H(F)$ is the convolution algebra of left and right I-invariant functions on $H(F)$. We refer the reader to [Schmidt 2005, §2.1] for details on the Iwahori Hecke algebra. Here, we recall the two projection operators (projecting onto the Siegel and Klingen parabolic subgroups) and the Atkin–Lehner involution. The unique (up to a constant) Iwahori fixed vector v_0 in π is annihilated by the projection operators and is an eigenvector of the Atkin–Lehner involution.

$$\sum_{w \in \mathfrak{o}/\mathfrak{p}} \pi\left(\begin{bmatrix} 1 & w & & \\ & 1 & & \\ & & 1 & \\ & & -w & 1 \end{bmatrix}\right) v_0 + \pi(s_1) v_0 = 0, \qquad \pi(\eta_0) v_0 = \omega v_0,$$
$$\sum_{y \in \mathfrak{o}/\mathfrak{p}} \pi\left(\begin{bmatrix} 1 & & & \\ & 1 & & \\ y & & 1 & \\ & & & 1 \end{bmatrix}\right) v_0 + \pi(s_2) v_0 = 0. \tag{2-2}$$

Here,

$$s_1 = \begin{bmatrix} & 1 & & \\ 1 & & & \\ & & & 1 \\ & & 1 & \end{bmatrix}, \quad s_2 = \begin{bmatrix} & & 1 & \\ & 1 & & \\ -1 & & & \\ & & & 1 \end{bmatrix}, \quad \eta_0 = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \varpi & & \\ \varpi & & & \end{bmatrix}, \quad \omega = -\Omega(\varpi).$$

3. Existence and uniqueness of Bessel models

for the Steinberg representation

We fix an additive character ψ of F, with conductor $\mathfrak{o}$. Let $a, b \in \mathfrak{o}$ and $c \in \mathfrak{o}^\times$ be as in Section 2, and set

$$S = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}.$$

Then, ψ defines a character θ on

$$U(F) = \left\{ \begin{bmatrix} 1_2 & X \\ & 1_2 \end{bmatrix} : {}^tX = X \right\} \qquad \text{by} \qquad \theta\left(\begin{bmatrix} 1 & X \\ & 1 \end{bmatrix} \right) = \psi(\operatorname{tr}(SX)).$$

Let

$$T(F) := \left\{ g \in \mathrm{GL}_2(F) : {}^tgSg = \det(g)S \right\}. \tag{3-1}$$

Set

$$\xi = \begin{bmatrix} \frac{b}{2} & c \\ -a & \frac{b}{2} \end{bmatrix} \qquad \text{and} \qquad F(\xi) = \{x + y\xi : x, y \in F\}.$$

It can be checked that $T(F)$ equals $F(\xi)^\times$ and is isomorphic to $L^\times$, with the isomorphism given by

$$\begin{bmatrix} x + \frac{b}{2}y & cy \\ -ay & x - \frac{b}{2}y \end{bmatrix} \mapsto \begin{cases} x + y\frac{\sqrt{d}}{2} & \text{if } L \text{ is a field;} \\ \left(x + y\frac{\sqrt{d}}{2},\ x - y\frac{\sqrt{d}}{2}\right) & \text{if } L = F \oplus F. \end{cases} \tag{3-2}$$

We consider $T(F)$ as a subgroup of $H(F)$ via

$$T(F) \ni g \longmapsto \begin{bmatrix} g & \\ & \det(g)\,{}^tg^{-1} \end{bmatrix} \in H(F).$$

Let $R(F) = T(F)U(F)$. We call $R(F)$ the *Bessel subgroup* of $H(F)$ (with respect to the given data a, b, c). Let Λ be any character on $L^\times$ that is trivial on $F^\times$. We will consider Λ as a character on $T(F)$. We have $\theta(t^{-1}ut) = \theta(u)$ for all $u \in U(F)$ and $t \in T(F)$. Hence, the map $tu \mapsto \Lambda(t)\theta(u)$ defines a character of $R(F)$. We denote this character by $\Lambda \otimes \theta$.

As mentioned in the introduction, a linear functional $\beta : V \to \mathbb{C}$, satisfying $\beta(\pi(r)v) = (\Lambda \otimes \theta)(r)\beta(v)$ for any $r \in R(F)$ and $v \in V$, is called a (Λ, θ)-Bessel functional for π. We say that π has a (Λ, θ)-Bessel model if π is isomorphic to a subspace of smooth functions $B : H(F) \to \mathbb{C}$ satisfying

$$B(tuh) = \Lambda(t)\theta(u)\,B(h) \qquad \text{for all } t \in T(F), u \in U(F), h \in H(F). \tag{3-3}$$

The existence of a nonzero (Λ, θ)-Bessel functional for π is equivalent to the existence of a nontrivial (Λ, θ)-Bessel model for π. If π has a nonzero (Λ, θ)-Bessel functional β, then the space $\{B_v : v \in \pi, B_v(h) := \beta(\pi(h)v)\}$ gives a nontrivial (Λ, θ)-Bessel model for π. Conversely, if π has a nontrivial (Λ, θ)-Bessel model

$\{B_v : v \in \pi\}$ then the linear functional $\beta(v) := B_v(1)$ is a nonzero (Λ, θ)-Bessel functional for π. We say that $v \in \pi$ is a test vector for a Bessel functional β if $\beta(v) \neq 0$. Note that a vector $v \in \pi$ is a test vector for β if and only if the corresponding function B_v in the Bessel model satisfies $B_v(1) \neq 0$.

Define the space $B(\Lambda, \theta)^{\mathrm{I}}$ of smooth functions B on $H(F)$ which are right I-invariant, satisfy (3-3) and the following conditions, for any $h \in H(F)$, obtained from (2-2),

$$\sum_{w\in\mathfrak{o}/\mathfrak{p}} B\left(h\begin{bmatrix}1 & w & & \\ & 1 & & \\ & & 1 & \\ & & -w & 1\end{bmatrix}\right) + B(hs_1) = 0, \tag{3-4}$$

$$B(h\eta_0) = \omega B(h), \tag{3-5}$$

$$\sum_{y\in\mathfrak{o}/\mathfrak{p}} B\left(h\begin{bmatrix}1 & & & \\ & 1 & & \\ y & & 1 & \\ & & & 1\end{bmatrix}\right) + B(hs_2) = 0. \tag{3-6}$$

Our aim is to obtain the criteria for existence and uniqueness for (Λ, θ)-Bessel models for π. We state the steps we take to obtain this.

i) Since a function B in $B(\Lambda, \theta)^{\mathrm{I}}$ is right I-invariant and satisfies (3-3) we see that the values of B are completely determined by its values on double coset representatives $R(F)\backslash H(F)/\mathrm{I}$. We obtain these representatives in Proposition 3.3.

ii) In Proposition 3.8, we use the I-invariance of B and (3-3)–(3-6) to obtain necessary conditions to be satisfied by the values of functions in $B(\Lambda, \theta)^{\mathrm{I}}$ on double coset representatives for $R(F)\backslash H(F)/\mathrm{I}$. This gives us $\dim(B(\Lambda, \theta)^{\mathrm{I}}) \leq 1$ in Corollary 3.9.

iii) In Proposition 3.10, we show that the function B with the given values at double coset representatives for $R(F)\backslash H(F)/\mathrm{I}$ (obtained in Proposition 3.8) is well-defined. We show that B satisfies (3-4), (3-5) and (3-6) for all values of $h \in H(F)$ and obtain the criteria for $\dim(B(\Lambda, \theta)^{\mathrm{I}}) = 1$ in Theorem 3.11.

iv) Suppose Λ is such that $\dim(B(\Lambda, \theta)^{\mathrm{I}}) = 1$. If Λ is unitary then we use $0 \neq B \in B(\Lambda, \theta)^{\mathrm{I}}$ to generate a Hecke module V_B. We define an inner product on V_B and show in Proposition 3.15 that V_B is irreducible and provides a (Λ, θ)-Bessel model for π. If Λ is not unitary (this can happen only if L is a split extension of F), then we show that any irreducible, generic, admissible representation of $H(F)$ has a split (Λ, θ)-Bessel model. Since π is generic in the split case, we obtain in Theorem 3.18 the precise criteria for existence and uniqueness of a (Λ, θ)-Bessel model for π.

3.1. ***Double coset decomposition.*** From [Furusawa 1993, (3.4.2)], we have the disjoint double coset decomposition

$$H(F)=\bigsqcup_{l\in\mathbb{Z}}\bigsqcup_{m\geq 0}R(F)h(l,m)K^H,\qquad h(l,m)=\begin{bmatrix}\varpi^{2m+l}&&&\\&\varpi^{m+l}&&\\&&1&\\&&&\varpi^m\end{bmatrix}.$$

It follows from the Bruhat decomposition for $\mathrm{Sp}(4,\mathfrak{o}/\mathfrak{p})$ that

$$\begin{aligned}K^H=\mathrm{I}\sqcup\bigsqcup_{x\in\mathfrak{o}/\mathfrak{p}}\begin{bmatrix}1&&&\\x&1&&\\&&1&-x\\&&&1\end{bmatrix}s_1\mathrm{I}\sqcup\bigsqcup_{x\in\mathfrak{o}/\mathfrak{p}}\begin{bmatrix}1&&x&\\&1&&\\&&1&\\&&&1\end{bmatrix}s_2\mathrm{I}\\
\sqcup\bigsqcup_{x,y\in\mathfrak{o}/\mathfrak{p}}\begin{bmatrix}1&&&\\x&1&&y\\&&1&-x\\&&&1\end{bmatrix}s_1s_2\mathrm{I}\sqcup\bigsqcup_{x,y\in\mathfrak{o}/\mathfrak{p}}\begin{bmatrix}1&&x&y\\&1&y&\\&&1&\\&&&1\end{bmatrix}s_2s_1\mathrm{I}\\
\sqcup\bigsqcup_{x,y,z\in\mathfrak{o}/\mathfrak{p}}\begin{bmatrix}1&&&y\\x&1&y&xy+z\\&&1&-x\\&&&1\end{bmatrix}s_1s_2s_1\mathrm{I}\sqcup\bigsqcup_{x,y,z\in\mathfrak{o}/\mathfrak{p}}\begin{bmatrix}1&&x&y\\&1&y&z\\&&1&\\&&&1\end{bmatrix}s_2s_1s_2\mathrm{I}\\
\sqcup\bigsqcup_{w,x,y,z\in\mathfrak{o}/\mathfrak{p}}\begin{bmatrix}1&&x&y\\w&1&wx+y&wy+z\\&&1&-w\\&&&1\end{bmatrix}s_1s_2s_1s_2\mathrm{I}.\end{aligned}$$

Let $W=\{1,s_1,s_2,s_1s_2,s_2s_1,s_1s_2s_1,s_2s_1s_2,s_1s_2s_1s_2\}$ be the Weyl group of $\mathrm{Sp}_4(F)$ and let the representatives for $\{1,s_1\}\setminus W$ be given by $W^{(1)}=\{1,s_2,s_2s_1,s_2s_1s_2\}$. Observing that

$$h(l,m)\begin{bmatrix}1&&\mathfrak{o}&\mathfrak{o}\\&1&\mathfrak{o}&\mathfrak{o}\\&&1&\\&&&1\end{bmatrix}h(l,m)^{-1}$$

is contained in $R(F)$, we get a preliminary (nondisjoint) decomposition

$$R(F)h(l,m)K^H=\bigcup_{\substack{s\in W^{(1)}\\ w\in\mathfrak{o}/\mathfrak{p}}}\big(R(F)h(l,m)s\mathrm{I}\cup R(F)h(l,m)W_ws_1s\mathrm{I}\big),\tag{3-7}$$

$$\text{with}\quad W_w:=\begin{bmatrix}1&&&\\w&1&&\\&&1&-w\\&&&1\end{bmatrix}.$$

The next lemma gives the condition under which the two double cosets of the form $R(F)h(l,m)s\mathrm{I}$ and $R(F)h(l,m)W_ws_1s\mathrm{I}$ are the same.

Lemma 3.1. *For $w\in\mathfrak{o}/\mathfrak{p}$ and $m\geq 0$, set $\beta_w^m:=a\varpi^{2m}+b\varpi^mw+cw^2$. Let $s\in W^{(1)}$. Then $R(F)h(l,m)s\mathrm{I}=R(F)h(l,m)W_ws_1s\mathrm{I}$ if and only if $\beta_w^m\in\mathfrak{o}^\times$.*

Proof. Suppose $\beta_w^m \in \mathfrak{o}^\times$. Take $y = \varpi^m$, $x = \varpi^m b/2 + cw$ and set

$$g = \begin{bmatrix} x + \frac{b}{2}y & cy \\ -ay & x - \frac{b}{2}y \end{bmatrix}.$$

Then

$$\begin{bmatrix} g & \\ & \det(g)\,{}^t g^{-1} \end{bmatrix} h(l,m) = h(l,m)\, W_w s_1 k,$$

where

$$k = \begin{bmatrix} -\beta_w^m & & & \\ b\varpi^m + cw & c & & \\ & & -c & b\varpi^m + cw \\ & & & \beta_w^m \end{bmatrix} \in \mathrm{I}.$$

Note that for any $s \in W^{(1)}$, we have $s^{-1}ks \in \mathrm{I}$. Using

$$rh(l,m)s = h(l,m)\, W_w s_1 s(s^{-1}ks),$$

we obtain $R(F)h(l,m)s\mathrm{I} = R(F)h(l,m)\, W_w s_1 s\mathrm{I}$, as required. The computation of the converse is straightforward. □

The next lemma describes for which $w \in \mathfrak{o}/\mathfrak{p}$ we have $\beta_w^m \in \mathfrak{o}^\times$.

Lemma 3.2. *For $w \in \mathfrak{o}/\mathfrak{p}$ and $m \geq 0$, set $\beta_w^m := a\varpi^{2m} + b\varpi^m w + cw^2$ as above.*

i) *If $m > 0$, then $\beta_w^m \in \mathfrak{o}^\times$ if and only if $w \in (\mathfrak{o}/\mathfrak{p})^\times$.*

ii) *Let $m = 0$.*

a) *If $\left(\frac{L}{\mathfrak{p}}\right) = -1$, then $\beta_w^0 \in \mathfrak{o}^\times$ for every $w \in \mathfrak{o}/\mathfrak{p}$.*

b) *Let $\left(\frac{L}{\mathfrak{p}}\right) = 0$. Let w_0 be the unique element of $\mathfrak{o}/\mathfrak{p}$ such that $\alpha + w_0 \in \mathfrak{p}_L$, the prime ideal of $\mathfrak{o}_L$. Then $\beta_w^0 \in \mathfrak{o}^\times$ if and only if $w \neq w_0$. In case $\#(\mathfrak{o}/\mathfrak{p})$ is odd, one can take $w_0 = -b/(2c)$.*

c) *Let $\left(\frac{L}{\mathfrak{p}}\right) = 1$. Then $\beta_w^0 \in \mathfrak{o}^\times$ if and only if $w \neq \dfrac{-b+\sqrt{d}}{2c}, \dfrac{-b-\sqrt{d}}{2c}$.*

Proof. Part (i) is clear. For the rest of the lemma, we need the equivalence

$$\beta_w^0 \in \mathfrak{o}^\times \iff \alpha + w \in \mathfrak{o}_L^\times. \tag{3-8}$$

This follows from the identity

$$a + bw + cw^2 = -c(\alpha + w)(\bar{\alpha} + w) = -cN(\alpha + w). \tag{3-9}$$

If $\left(\frac{L}{\mathfrak{p}}\right) = -1$, then $\mathfrak{p}_L = \mathfrak{P}$ and Lemma 2.1 implies that $\alpha + w \in \mathfrak{o}_L^\times$ for all $w \in \mathfrak{o}/\mathfrak{p}$. The equivalence (3-8) gives (ii-a) of the lemma. Let us now assume that $\left(\frac{L}{\mathfrak{p}}\right) = 0$. In this case, the injective map $\iota : \mathfrak{o} \hookrightarrow \mathfrak{o}_L$ gives an isomorphism between the fields $\mathfrak{o}/\mathfrak{p} \simeq \mathfrak{o}_L/\mathfrak{p}_L$. Let $w_0 = -\iota^{-1}(\alpha)$ be the unique element in $\mathfrak{o}/\mathfrak{p}$ such that $\alpha + w_0 \in \mathfrak{p}_L$. In case $\#(\mathfrak{o}/\mathfrak{p})$ is odd, then one can take $w_0 = -b/(2c) \in \mathfrak{o}$ since $\sqrt{d} \in \mathfrak{p}_L$. Then for any $w \in \mathfrak{o}/\mathfrak{p}$, $w \neq w_0$, we have $\alpha + w \in \mathfrak{o}_L^\times$. Now (3-8) gives (ii-b) of the lemma. Next assume that $\left(\frac{L}{\mathfrak{p}}\right) = 1$. Since $\sqrt{d} \in \mathfrak{o}^\times$ by assumption, we have $\alpha \notin \mathfrak{P}$. If $\alpha + w \notin \mathfrak{o}_L^\times$ for some $w \in \mathfrak{o}$, then we have one of $(b \pm \sqrt{d})/(2c) + w$ lies in $\mathfrak{p}$. Hence, we see that the only choices of $w = (w, w)$ such that $\alpha + w \notin \mathfrak{o}_L^\times$ are $w = (-b \pm \sqrt{d})/(2c)$. Note that $\sqrt{d} \in \mathfrak{o}^\times$ implies that $(-b \pm \sqrt{d})/(2c)$ are not equal modulo $\mathfrak{p}$. This completes the proof of the lemma. $\square$

In the case $\left(\frac{L}{\mathfrak{p}}\right) = 0$, (3-9) implies that $\beta^0_{w_0} \in \mathfrak{p}$ but $\beta^0_{w_0} \notin \mathfrak{p}^2$ by Lemma 2.1. The disjointness of all the relevant double cosets can be checked easily. We summarize in the following proposition.

Proposition 3.3. *Let W be the Weyl group of $\mathrm{Sp}_4(F)$ and set*

$$W^{(1)} = \{1, s_2, s_2s_1, s_2s_1s_2\}.$$

If $\left(\frac{L}{\mathfrak{p}}\right) = 0$, let w_0 be the unique element of $\mathfrak{o}/\mathfrak{p}$ such that $\alpha + w_0 \in \mathfrak{p}_L$. If $\#(\mathfrak{o}/\mathfrak{p})$ is odd, then take $w_0 = -b/(2c)$. We have the disjoint double coset decomposition

$$R(F)h(l,m)K^H =$$

$$\begin{cases} \displaystyle\bigsqcup_{s \in W} R(F)h(l,m)s\mathrm{I} & \text{if } m > 0; \\ \displaystyle\bigsqcup_{s \in W^{(1)}} R(F)h(l,0)s\mathrm{I} & \text{if } m = 0,\ \left(\frac{L}{\mathfrak{p}}\right) = -1; \\ \displaystyle\bigsqcup_{s \in W^{(1)}} \big(R(F)h(l,0)s\mathrm{I} \sqcup R(F)h(l,0)\,W_{w_0}s_1s\mathrm{I}\big) & \text{if } m = 0,\ \left(\frac{L}{\mathfrak{p}}\right) = 0; \\ \displaystyle\bigsqcup_{s \in W^{(1)}} \big(R(F)h(l,0)s\mathrm{I} \sqcup R(F)h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s \sqcup R(F)h(l,0)\,W_{\frac{-b-\sqrt{d}}{2c}}s_1s\mathrm{I}\big) & \\ & \text{if } m = 0,\ \left(\frac{L}{\mathfrak{p}}\right) = 1. \end{cases}$$

3.2. ***Necessary conditions for values of $B \in B(\Lambda, \theta)^{\mathrm{I}}$.*** We will now obtain the necessary conditions on the values of $B \in B(\Lambda, \theta)^{\mathrm{I}}$ on the double coset representatives from Proposition 3.3 using the I-invariance of B and (3-3)–(3-6).

Conductor of Λ: We define

$$c(\Lambda) = \min\{m \geq 0 : \Lambda|_{(1+\mathfrak{P}^m)\cap\mathfrak{o}_L^\times} \equiv 1\}. \tag{3-10}$$

Note that $(1 + \mathfrak{P}^m) \cap \mathfrak{o}_L^\times = 1 + \mathfrak{P}^m$ if $m \geq 1$ and $(1 + \mathfrak{P}^m) \cap \mathfrak{o}_L^\times = \mathfrak{o}_L^\times$ if $m = 0$. Also, $c(\Lambda)$ is the conductor of Λ only if $\left(\frac{L}{\mathfrak{p}}\right) = -1$. We set $c(\Lambda) = m_0$. Since Λ is trivial on $F^\times$, we see that $\Lambda|_{(\mathfrak{o}^\times+\mathfrak{P}^{m_0})\cap\mathfrak{o}_L^\times} \equiv 1$. Observe that if L is a field, then we have $L^\times = \langle\varpi_L\rangle.\mathfrak{o}_L^\times$. If $\left(\frac{L}{\mathfrak{p}}\right) = -1$ and $m_0 = 0$, then we have that $\Lambda(\varpi_L) = 1$, since $\varpi_L \in \varpi\,\mathfrak{o}_L^\times$. In case $\left(\frac{L}{\mathfrak{p}}\right) = 0$ and $m_0 = 0$, we see that $\Lambda(\varpi_L) = \pm 1$. In general, if L is a field, we see that Λ is a unitary character since m_0 is finite. On the other hand, if L is not a field, then $L^\times = F^\times \oplus F^\times$ and $\Lambda((x, y)) = \Lambda_1(x)\Lambda_2(y)$, where

Λ_1, Λ_2 are two characters of $F^\times$ satisfying $\Lambda_1.\Lambda_2 \equiv 1$. In this case, m_0 is the conductor of both Λ_1, Λ_2 and the character Λ need not be unitary.

In the next lemma, we will describe some coset representatives, which will be used in the evaluation of certain sums involving the character Λ.

Lemma 3.4. *Let $m \geq 1$. A set of coset representatives for*

$$((\mathfrak{o}^\times + \mathfrak{P}^{m-1}) \cap \mathfrak{o}_L^\times)/(\mathfrak{o}^\times + \mathfrak{P}^m)$$

is given by $\{w + \alpha\varpi^{m-1} : w \in (\mathfrak{o}/\mathfrak{p})^\times\} \cup \{1\}$ if $m \geq 2$ and $\{w + \alpha : w \in \mathfrak{o}/\mathfrak{p},\ w + \alpha \in \mathfrak{o}_L^\times\} \cup \{1\}$ if $m = 1$.

Proof. Let $x + \alpha\varpi^{m-1}y \in (\mathfrak{o}^\times + \mathfrak{P}^{m-1}) \cap \mathfrak{o}_L^\times$, with $x, y \in \mathfrak{o}$. If $m \geq 2$, then $x \in \mathfrak{o}^\times$. If $y \in \mathfrak{p}$, then $x + \alpha\varpi^{m-1}y \in (\mathfrak{o}^\times + \mathfrak{P}^m)$, and hence corresponds to the coset representative 1. Now, we assume that $y \in \mathfrak{o}^\times$. Then, using $y \in \mathfrak{o}^\times + \mathfrak{P}^m$, we see that $x + \alpha\varpi^{m-1}y$ is equivalent to $x/y + \alpha\varpi^{m-1}$ modulo $(\mathfrak{o}^\times + \mathfrak{P}^m)$. Note that $x/y + \alpha\varpi^{m-1} \in \mathfrak{o}_L^\times$ implies that, modulo $\mathfrak{p}$, the element x/y lies in

$$\begin{cases} (\mathfrak{o}/\mathfrak{p})^\times & \text{if } m \geq 2, \\ \mathfrak{o}/\mathfrak{p} & \text{if } m = 1,\ \left(\frac{L}{\mathfrak{p}}\right) = -1 \\ \mathfrak{o}/\mathfrak{p} - \{w_0\} & \text{if } m = 1,\ \left(\frac{L}{\mathfrak{p}}\right) = 0, \\ \mathfrak{o}/\mathfrak{p} - \{(-b \pm \sqrt{d})/(2c)\} & \text{if } m = 1,\ \left(\frac{L}{\mathfrak{p}}\right) = 1. \end{cases} \tag{3-11}$$

This follows from the proof of Lemma 3.2. A calculation shows that if w, w' are equivalent, modulo $\mathfrak{p}$, to (not necessarily the same) elements in the sets defined in (3-11), then

$$w \equiv w' \pmod{\mathfrak{p}} \quad \text{if and only if} \quad (w + \alpha\varpi^{m-1})/(w' + \alpha\varpi^{m-1}) \in \mathfrak{o}^\times + \mathfrak{P}^m.$$

This completes the proof of the lemma. □

Depending on the $c(\Lambda) = m_0$, certain values of B have to be zero. This is obtained in the next lemma.

Lemma 3.5. *For any $l \in \mathbb{Z}$, we have $B(h(l, m)s) = 0$, if any of the following conditions are satisfied.*

i) $m \leq m_0 - 2,\ m_0 \geq 2,\ s = 1;$

ii) $m = 0,\ \left(\frac{L}{\mathfrak{p}}\right) = 1,\ m_0 \geq 1,\ s \in \{W_w s_1 : w = (-b \pm \sqrt{d})/(2c)\};$

iii) $m = 0,\ \left(\frac{L}{\mathfrak{p}}\right) = 0,\ \Lambda = \Omega \circ N_{L/F},\ m_0 = 0,\ s = W_{w_0} s_1 s_2;$

iv) $m = 0,\ \left(\frac{L}{\mathfrak{p}}\right) = -1,\ m_0 = 0,\ s = 1.$

Proof. We illustrate the proof of (i) here. Let $m \leq m_0 - 2$. Let

$$1 + x + \alpha y \in 1 + \mathfrak{P}^{m+1}, \quad \text{with } x, y \in \mathfrak{p}^{m+1},$$

be such that $\Lambda(1+x+\alpha y) \neq 1$. Let

$$k = \begin{bmatrix} c(1+x)+by & cy\varpi^{-m} & & \\ -ay\varpi^{m} & c(1+x) & & \\ & & c(1+x) & ay\varpi^{m} \\ & & -cy\varpi^{-m} & c(1+x)+by \end{bmatrix} \in \mathrm{I}.$$

Then

$$\begin{aligned} B(h(l,m)) &= B(h(l,m)k) \\ &= B\left(\begin{bmatrix} c(1+x)+by & cy & & \\ -ay & c(1+x) & & \\ & & c(1+x) & ay \\ & & -cy & c(1+x)+by \end{bmatrix} h(l,m)\right) \\ &= \Lambda(1+x+\alpha y)\, B(h(l,m)), \end{aligned}$$

which implies that $B(h(l,m)) = 0$, as required. The other cases are computed in a similar manner. $\square$

From Lemmas 3.4 and 3.5(i), we obtain information on certain character sums involving Λ:

Lemma 3.6. *For any* l, *we have*

$$\sum_{w\in(\mathfrak{o}/\mathfrak{p})^\times} \Lambda(w+\alpha\varpi^m)\, B(h(l,m)) + B(h(l,m)) = \begin{cases} 0 & \text{if } 0 < m < m_0, \\ qB(h(l,m)) & \text{if } m \geq m_0,\ m > 0; \end{cases}$$

$$\sum_{\substack{w\in\mathfrak{o}/\mathfrak{p} \\ w+\alpha\in\mathfrak{o}_L^\times}} \Lambda(w+\alpha)\, B(h(l,0)) + B(h(l,0)) = \begin{cases} 0 & \text{if } m_0 \geq 1, \\ \left(q - \left(\frac{L}{\mathfrak{p}}\right)\right) B(h(l,0)) & \text{if } m_0 = 0. \end{cases}$$

Conductor of ψ. Since the conductor of ψ is $\mathfrak{o}$, we obtain the following further vanishing conditions on the values of B.

Lemma 3.7. *For* $m \geq 0$, *we have* $B(h(l,m)s) = 0$ *if one of the following conditions are satisfied*:

i) $l < 0,\ s \in \{1, s_1, s_2, s_2s_1\}$;

ii) $l < -1,\ s \in \{s_1s_2, s_1s_2s_1, s_2s_1s_2, s_1s_2s_1s_2\}$.

For $w \in \mathfrak{o}$, *we have* $B(h(l,0)W_ws) = 0$ *if one of the following conditions are satisfied*:

i) $l < 0,\ s = s_1$;

ii) $l < -1,\ s \in \{s_1s_2, s_1s_2s_1, s_1s_2s_1s_2\}$.

If $\left(\frac{L}{\mathfrak{p}}\right) = 1$ *and* $w = \dfrac{-b\pm\sqrt{d}}{2c}$, *then* $B(h(-1,0)W_ws_1s_2) = 0$.

Proof. We illustrate the proof for the case $m \geq 0$, $l < 0$, $s \in \{1, s_1, s_2, s_2 s_1\}$. For any $\epsilon \in \mathfrak{o}^\times$, set

$$k_s^\epsilon = \begin{bmatrix} 1 &&& \\ & 1 & \epsilon & \\ && 1 & \\ &&& 1 \end{bmatrix} \text{ if } s = 1, s_2 \quad \text{and} \quad k_s^\epsilon = \begin{bmatrix} 1 && \epsilon & \\ & 1 && \\ && 1 & \\ &&& 1 \end{bmatrix} \text{ if } s = s_1, s_2 s_1.$$

Then, for $s \in \{1, s_1, s_2, s_2 s_1\}$ and $\epsilon \in \mathfrak{o}^\times$, we obtain

$$\begin{aligned} B(h(l,m)s) &= B(h(l,m)sk_s^\epsilon) \\ &= B\left(\begin{bmatrix} 1 &&& \\ & 1 & \epsilon\varpi^l & \\ && 1 & \\ &&& 1 \end{bmatrix} h(l,m)s\right) = \psi(c\epsilon\varpi^l)\, B(h(l,m)s). \end{aligned}$$

Since the conductor of ψ is $\mathfrak{o}$, we conclude that $B(h(l,m)s) = 0$ if $l < 0$. The other cases are computed in a similar manner. $\square$

Values of B using (3-4). Substituting $h = h(l,m)s_1$ in (3-4) and using Lemmas 3.1, 3.2 and 3.6, we get, for any l,

$$B(h(l,m)s_1) = \begin{cases} 0 & \text{if } m < m_0 \text{ and } m > 0, \\ -qB(h(l,m)) & \text{if } m \geq m_0 \text{ and } m > 0; \end{cases} \tag{3-12}$$

$$B(h(l,0)W_{w_0}s_1) = \begin{cases} 0 & \text{if } m_0 \geq 1, \\ -qB(h(l,0)) & \text{if } m_0 = 0; \end{cases} \tag{3-13}$$

$$B\big(h(l,0)W_{\frac{-b+\sqrt{d}}{2c}} s_1\big) + B\big(h(l,0)W_{\frac{-b-\sqrt{d}}{2c}} s_1\big) = -(q-1)\, B(h(l,0)) \quad \text{if } m_0 = 0. \tag{3-14}$$

Substituting $h = h(l,m)s_2 s_1$ in (3-4) and using that the conductor of ψ is $\mathfrak{o}$, we get for any l, m

$$B(h(l,m)s_2 s_1) = -\frac{1}{q} B(h(l,m)s_2). \tag{3-15}$$

Substituting $h = h(l,m)s_1 s_2 s_1$ in (3-4) and using that the conductor of ψ is $\mathfrak{o}$, we get for any $m > 0$ and l

$$B(h(l,m)s_1 s_2 s_1) = -\frac{1}{q} B(h(l,m)s_1 s_2). \tag{3-16}$$

Let $\left(\frac{L}{\mathfrak{p}}\right) = 0$. Substituting $h = h(-1,0)W_{w_0}s_1 s_2 s_1$ in (3-4) and using that the conductor of ψ is $\mathfrak{o}$ and $b + 2cw_0 \in \mathfrak{p}$, we get

$$B(h(-1,0)\, W_{w_0}s_1 s_2 s_1) = -\frac{1}{q} B(h(-1,0)\, W_{w_0}s_1 s_2).$$

Let $\left(\frac{L}{\mathfrak{p}}\right)=1$ and $w=(-b\pm\sqrt{d})/(2c)$. Substituting $h=h(l,0)W_w s_1 s_2 s_1$ in (3-4) and using that the conductor of ψ is $\mathfrak{o}$ and $\sqrt{d}\in\mathfrak{o}^\times$, we get for $l\neq -1$

$$B(h(l,m)\,W_w s_1 s_2 s_1)=-\frac{1}{q}B(h(l,m)\,W_w s_1 s_2). \tag{3-17}$$

Values of B using (3-6). Substituting $h=h(l,m)s_2$ in (3-6) and using that the conductor of ψ is $\mathfrak{o}$, we get for any l,m

$$B(h(l,m)\,s_2)=-\frac{1}{q}B(h(l,m)). \tag{3-18}$$

Substituting $h=h(l,m)s_2 s_1 s_2$ in (3-6) and using that the conductor of ψ is $\mathfrak{o}$, we get for $l\neq -1$

$$B(h(l,m)\,s_2 s_1 s_2)=-\frac{1}{q}B(h(l,m)\,s_2 s_1). \tag{3-19}$$

Set

$$w=\begin{cases}0 & \text{if } m>0,\\ w_0 & \text{if } m=0,\ \left(\frac{L}{\mathfrak{p}}\right)=0,\\ \dfrac{-b\pm\sqrt{d}}{2c} & \text{if } m=0,\ \left(\frac{L}{\mathfrak{p}}\right)=1.\end{cases}$$

Substituting $h=h(l,m)W_w s_1 s_2$ in (3-6) and using that the conductor of ψ is $\mathfrak{o}$, we get for $l\neq -1$

$$B(h(l,m)\,W_w s_1 s_2)=-\frac{1}{q}B(h(l,m)\,W_w s_1). \tag{3-20}$$

Substituting $h=h(l,m)\,W_w s_1 s_2 s_1 s_2$ in (3-6) and using that the conductor of ψ is $\mathfrak{o}$, we get for all l,m

$$B(h(l,m)\,W_w s_1 s_2 s_1 s_2)=-\frac{1}{q}B(h(l,m)\,W_w s_1 s_2 s_1). \tag{3-21}$$

Values of B using (3-5). For any l,m,w we have the matrix identities

$$h(l,m)\,s_2 s_1\eta_0=h(l-1,m+1)s_1 s_2 s_1\begin{bmatrix}1&&&\\&-1&&\\&&-1&\\&&&1\end{bmatrix}, \tag{3-22}$$

$$h(l,m)\,W_w s_1 s_2 s_1 s_2\eta_0=h(l+1,m)\,W_w s_1\begin{bmatrix}1&&&\\&1&&\\&&-1&\\&&&-1\end{bmatrix}, \tag{3-23}$$

$$h(l,m)\,s_2 s_1 s_2\eta_0=h(l+1,m)\begin{bmatrix}1&&&\\&1&&\\&&-1&\\&&&-1\end{bmatrix}. \tag{3-24}$$

Hence, by (3-5), we have

$$B(h(l,m)\,s_2s_1) = \omega B(h(l-1,m+1)\,s_1s_2s_1), \tag{3-25}$$

$$B(h(l,m)\,W_w s_1s_2s_1s_2) = \omega B(h(l+1,m)\,W_w s_1), \tag{3-26}$$

$$B(h(l,m)\,s_2s_1s_2) = \omega B(h(l+1,m)). \tag{3-27}$$

Using (3-24) we see that

$$\begin{aligned} B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2\big) &= \omega B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2\eta_0\big) \\ &= \omega B\left(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}\begin{bmatrix}1&&&\\&\varpi&&\\&&\varpi&\\&&&1\end{bmatrix}s_2\right). \end{aligned}$$

Let $x=\sqrt{d}/2+\varpi$, $y=1$, $g=\begin{bmatrix}x+by/2 & cy\\ -ay & x-by/2\end{bmatrix}$, and $r=\begin{bmatrix}g&\\&\det(g)\,{}^tg^{-1}\end{bmatrix}$. We have the matrix identity

$$rh(l,0)\,W_{\frac{-b-\sqrt{d}}{2c}}s_1s_2 = h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}\begin{bmatrix}1&&&\\&\varpi&&\\&&\varpi&\\&&&1\end{bmatrix}s_2k,$$

$$\text{with}\quad k=\begin{bmatrix}\sqrt{d}/c&&&-1\\&-\sqrt{d}/c&1&\\&\varpi&c&\\-\varpi&&&-c\end{bmatrix}\in \mathrm{I}.$$

This gives us

$$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2\big) = \omega\Lambda((\sqrt{d}+\varpi,\varpi))\,B\big(h(l,0)\,W_{\frac{-b-\sqrt{d}}{2c}}s_1s_2\big). \tag{3-28}$$

Summary. Using (3-15), (3-18), (3-19) and (3-27) we get for $l,m\geq 0$

$$B(h(l+1,m)) = -\frac{\omega}{q^3}B(h(l,m)). \tag{3-29}$$

Using (3-12), (3-15), (3-16), (3-18), (3-20), (3-25) and (3-29), we get for $l\geq 0$ and $m\geq m_0-1$

$$B(h(l,m+1)) = \frac{1}{q^4}B(h(l,m)). \tag{3-30}$$

Hence, we conclude that

$$B(h(l,m)) = \begin{cases} 0 & \text{if } l\leq -1 \text{ or } 0\leq m\leq m_0-2,\\ q^{-4(m-m_0+1)}(-\omega q^{-3})^l B(h(0,m_0-1)) & \text{if } l\geq 0 \text{ and } m\geq m_0-1>0,\\ q^{-4m}(-\omega q^{-3})^l B(1) & \text{if } l\geq 0 \text{ and } m\geq m_0=0,1. \end{cases} \tag{3-31}$$

Let $\left(\frac{L}{\mathfrak{p}}\right) = 1$ and $w = (-b \pm \sqrt{d})/(2c)$. Using (3-17), (3-20), (3-21) and (3-26), we get for $l \geq 0$, $B(h(l+1,0)\, W_w s_1) = (-\omega q^{-3}) B(h(l,0)\, W_w s_1)$, which gives us

$$B(h(l,0)\, W_w s_1) = (-\omega q^{-3})^l B(W_w s_1).$$

In addition, if $m_0 = 0$ and $\omega\Lambda((1,\varpi)) = -1$, using (3-14), (3-20) and (3-28), we get for all $l \geq 0$

$$B(h(l,0)) = 0.$$

Summarizing the calculations of the values of B, we obtain

Proposition 3.8. *Let $c(\Lambda) = m_0$. For $l, m \in \mathbb{Z}$, $m \geq 0$, we set*

$$A_{l,m} := \begin{cases} q^{-4(m-m_0+1)}(-\omega q^{-3})^l & \text{if } m_0 \geq 1, \\ q^{-4m}(-\omega q^{-3})^l & \text{if } m_0 = 0, \end{cases}$$

$$C_{m_0} := \begin{cases} B(h(0, m_0-1)) & \text{if } m_0 \geq 1, \\ B(1) & \text{if } m_0 = 0. \end{cases}$$

We have the following necessary conditions on the values of $B \in B(\Lambda,\theta)^{\mathrm{I}}$.

i) *For $m \geq 0$ and any m_0,*

a) $$B(h(l,m)) = \begin{cases} 0 & \text{if } l \leq -1 \text{ or } m \leq m_0-2, \\ A_{l,m}C_{m_0} & \text{if } l \geq 0 \text{ and } m \geq m_0-1. \end{cases}$$

b) $$B(h(l,m)s_2) = \begin{cases} 0 & \text{if } l \leq -1 \text{ or } m \leq m_0-2, \\ -q^{-1}A_{l,m}C_{m_0} & \text{if } l \geq 0 \text{ and } m \geq m_0-1. \end{cases}$$

c) $$B(h(l,m)s_2s_1) = \begin{cases} 0, & \text{if } l \leq -1 \text{ or } m \leq m_0-2, \\ q^{-2}A_{l,m}C_{m_0}, & \text{if } l \geq 0 \text{ and } m \geq m_0-1. \end{cases}$$

d) $$B(h(l,m)s_2s_1s_2) = \begin{cases} 0, & \text{if } l \leq -2 \text{ or } m \leq m_0-2, \\ \omega A_{0,m}C_{m_0}, & \text{if } l = -1 \text{ and } m \geq m_0-1, \\ -q^{-3}A_{l,m}C_{m_0}, & \text{if } l \geq 0 \text{ and } m \geq m_0-1. \end{cases}$$

ii) *For $m > 0$ and any m_0,*

a) $$B(h(l,m)s_1) = \begin{cases} 0 & \text{if } l \leq -1 \text{ or } m \leq m_0-1, \\ -qA_{l,m}C_{m_0} & \text{if } l \geq 0 \text{ and } m \geq m_0. \end{cases}$$

b) $$B(h(l,m)s_1s_2) = \begin{cases} 0 & \text{if } l \leq -2 \text{ or } m \leq m_0-1, \\ -\omega q^3 A_{0,m}C_{m_0}, & \text{if } l = -1 \text{ and } m \geq m_0, \\ A_{l,m}C_{m_0}, & \text{if } l \geq 0 \text{ and } m \geq m_0. \end{cases}$$

c) $$B(h(l,m)s_1s_2s_1) = \begin{cases} 0 & \text{if } l \leq -2 \text{ or } m \leq m_0-1, \\ \omega q^2 A_{0,m}C_{m_0} & \text{if } l = -1 \text{ and } m \geq m_0, \\ -q^{-1}A_{l,m}C_{m_0}, & \text{if } l \geq 0 \text{ and } m \geq m_0. \end{cases}$$

d) $B(h(l,m)s_1s_2s_1s_2) = \begin{cases} 0 & \text{if } l \leq -2 \text{ or } m \leq m_0 - 1, \\ -\omega q A_{0,m} C_{m_0} & \text{if } l = -1 \text{ and } m \geq m_0, \\ q^{-2} A_{l,m} C_{m_0}, & \text{if } l \geq 0 \text{ and } m \geq m_0. \end{cases}$

iii) *Let* $m_0 \geq 1$.

a) *If* $\left(\frac{L}{\mathfrak{p}}\right) = 0$ *and* $s \in \{1, s_2, s_2s_1, s_2s_1s_2\}$*, then, for all* l*,*

$$B(h(l,0)\,W_{w_0}s_1s) = 0.$$

b) *If* $\left(\frac{L}{\mathfrak{p}}\right) = 1$, $s \in \{1, s_2, s_2s_1, s_2s_1s_2\}$ *and* $w = \dfrac{-b \pm \sqrt{d}}{2c}$*, then, for all* l*,*

$$B(h(l,0)\,W_w s_1 s) = 0.$$

iv) *Let* $m_0 = 0$.

a) *If* $\left(\frac{L}{\mathfrak{p}}\right) = -1$ *then* $C_0 = 0$.

b) *Suppose* $\left(\frac{L}{\mathfrak{p}}\right) = 0$*. Then*

1) $$B(h(l,0)\,W_{w_0}s_1) = \begin{cases} 0 & \text{if } l \leq -1, \\ -q A_{l,0} C_0 & \text{if } l \geq 0. \end{cases}$$

2) $$B(h(l,0)\,W_{w_0}s_1s_2) = \begin{cases} 0 & \text{if } l \leq -2, \\ -\omega q^3 C_0 & \text{if } l = -1, \\ A_{l,0} C_0, & \text{if } l \geq 0. \end{cases}$$

3) $$B(h(l,0)\,W_{w_0}s_1s_2s_1) = \begin{cases} 0 & \text{if } l \leq -2, \\ \omega q^2 A_{l+1,0} C_0, & \text{if } l \geq -1. \end{cases}$$

4) $$B(h(l,0)\,W_{w_0}s_1s_2s_1s_2) = \begin{cases} 0 & \text{if } l \leq -2, \\ -\omega q A_{l+1,0} C_0, & \text{if } l \geq -1. \end{cases}$$

c) *Suppose* $\left(\frac{L}{\mathfrak{p}}\right) = 0$ *and* $\Lambda = \Omega \circ N_{L/F}$*. Then* $C_0 = 0$.

d) *Suppose* $\left(\frac{L}{\mathfrak{p}}\right) = 1$*. Then for* $s \in \{1, s_2, s_2s_1, s_2s_1s_2\}$

$$B\big(h(l,0)\,W_{\frac{-b-\sqrt{d}}{2c}}s_1s\big) = \frac{1}{\omega\Lambda((1,\varpi))} B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s\big).$$

e) *Suppose* $\left(\frac{L}{\mathfrak{p}}\right) = 1$ *and* $\omega\Lambda((1,\varpi)) = -1$.

1) $$C_0 = 0.$$

2) $$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1\big) = \begin{cases} 0 & \text{if } l \leq -1, \\ A_{l,0} B(W_{\frac{-b+\sqrt{d}}{2c}}s_1) & \text{if } l \geq 0. \end{cases}$$

3) $$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2\big) = \begin{cases} 0 & \text{if } l \leq -1, \\ -\frac{1}{q} A_{l,0} B\big(W_{\frac{-b+\sqrt{d}}{2c}}s_1\big) & \text{if } l \geq 0. \end{cases}$$

4) $$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2s_1\big)=\begin{cases}0 & \textit{if } l\le -2,\\ -\omega q A_{l+1,0}B\big(W_{\frac{-b+\sqrt{d}}{2c}}s_1\big) & \textit{if } l\ge -1.\end{cases}$$

5) $$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2s_1s_2\big)=\begin{cases}0 & \textit{if } l\le -2,\\ \omega A_{l+1,0}B\big(W_{\frac{-b+\sqrt{d}}{2c}}s_1\big) & \textit{if } l\ge -1.\end{cases}$$

f) *Suppose* $\left(\frac{L}{\mathfrak{p}}\right)=1$ *and* $\omega\Lambda((1,\varpi))\neq -1$. *Set* $\nu=\dfrac{q-1}{1+\omega\Lambda((1,\varpi))}$.

1) $$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1\big)=\begin{cases}0 & \textit{if } l\le -1,\\ -\nu A_{l,0}C_0, & \textit{if } l\ge 0.\end{cases}$$

2) $$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2\big)=\begin{cases}0 & \textit{if } l\le -1,\\ q^{-1}\nu A_{l,0}C_0, & \textit{if } l\ge 0.\end{cases}$$

3) $$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2s_1\big)=\begin{cases}0 & \textit{if } l\le -2,\\ \omega q\nu A_{l+1,0}C_0, & \textit{if } l\ge -1.\end{cases}$$

4) $$B\big(h(l,0)\,W_{\frac{-b+\sqrt{d}}{2c}}s_1s_2s_1s_2\big)=\begin{cases}0 & \textit{if } l\le -2,\\ -\omega\nu A_{l+1,0}C_0 & \textit{if } l\ge -1.\end{cases}$$

Corollary 3.9. *For any character* Λ, *we have*

$$\dim\big(B(\Lambda,\theta)^{\mathrm{I}}\big)\le 1.$$

3.3. ***Well-definedness of B.*** In this section, we will show that a function B on $H(F)$, which is right I-invariant, satisfies (3-3) and with values on the double coset representatives of $R(F)\backslash H(F)/\mathrm{I}$ given by Proposition 3.8, is well defined. Hence, we have to show that

$$r_1sk_1=r_2sk_2 \;\Rightarrow\; B(r_1sk_1)=B(r_2sk_2)$$

for $r_1,r_2\in R(F)$, $k_1,k_2\in \mathrm{I}$ and any double coset representative s. This is obtained in the following proposition.

Proposition 3.10. *Let s be any double coset representative from Proposition 3.3 and the values* $B(s)$ *be as in Proposition 3.8. Let* $t\in T(F)$, $u\in U(F)$ *such that* $s^{-1}tus\in \mathrm{I}$. *Then*

$$\Lambda(t)\theta(u)=1 \quad \textit{or} \quad B(s)=0.$$

Proof. Let

$$t=\begin{bmatrix} g & \\ & \det(g)\,{}^tg^{-1}\end{bmatrix} \quad\text{and}\quad u=\begin{bmatrix}1 & X\\ & 1\end{bmatrix},$$

with

$$g=\begin{bmatrix} x+by/2 & cy\\ -ay & x-by/2\end{bmatrix} \quad\text{and}\quad X={}^tX.$$

First let $s = h(l, m)$. Observe that

$$x + y\frac{\sqrt{d}}{2} = x - \frac{by}{2} + cy\alpha.$$

(In the split case, we consider the same identity with $(x + y\sqrt{d}/2,\ x - y\sqrt{d}/2)$). We assume $s^{-1}tus \in \mathrm{I}$. We see that $x \pm by/2 \in \mathfrak{o}^\times$, $y \in \mathfrak{p}^{m+1}$ and $x + \sqrt{d}y/2 \in \mathfrak{o}^\times + \mathfrak{P}^{m+1}$. Hence, we conclude that $g \in \mathrm{GL}_2(\mathfrak{o})$. This gives us

$$X \in \begin{bmatrix} \mathfrak{p}^{l+2m} & \mathfrak{p}^{l+m} \\ \mathfrak{p}^{l+m} & \mathfrak{p}^{l} \end{bmatrix}.$$

Now looking at the values of $B(h(l, m))$ from Proposition 3.8, we get that either $B(s) = 0$ or $\Lambda(t) = \theta(u) = 1$.

We will illustrate one other case, $s = h(l, 0)W_{w_0}s_1s_2$, since it is the most complicated. Here, w_0 is the unique element of $\mathfrak{o}/\mathfrak{p}$ such that $w_0 + \alpha \notin \mathfrak{o}_L^\times$. If $m_0 \geq 1$ or $l \leq -2$, then we have $B(s) = 0$. Hence, assume that $m_0 = 0$ and $l \geq -1$. Note that $x + y\sqrt{d}/2 = x - by/2 - cw_0y + c(w_0 + \alpha)y$ and $a + bw_0 + cw_0^2 \in \mathfrak{p}$. We see that $s^{-1}tus \in \mathrm{I}$ implies that

$$y \in \mathfrak{o} \qquad \text{and} \qquad x \pm \Big(\frac{b}{2} + cw_0\Big)y \in \mathfrak{o}^\times.$$

Hence, we see that $x + y\frac{\sqrt{d}}{2} \in \mathfrak{o}_L^\times$. This implies that $g \in \mathrm{GL}_2(\mathfrak{o})$ and $\Lambda(t) = 1$. We have

$$\begin{bmatrix} 1 & \\ -w_0 & 1 \end{bmatrix} gX \begin{bmatrix} 1 & -w_0 \\ & 1 \end{bmatrix} \in \begin{bmatrix} \mathfrak{p}^{l} & \mathfrak{p}^{l} \\ \mathfrak{p}^{l} & \mathfrak{p}^{l+1} \end{bmatrix}.$$

If $l \geq 0$, then we get $\theta(u) = 1$, as required. If $l = -1$, then let

$$\begin{bmatrix} 1 & \\ -w_0 & 1 \end{bmatrix} gX \begin{bmatrix} 1 & -w_0 \\ & 1 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \qquad \text{with } x_1, x_2, x_3 \in \varpi^{-1}\mathfrak{o},\ x_4 \in \mathfrak{o}.$$

Set $\epsilon_1 = x + (b/2 + cw_0)y, \epsilon_2 = x - (b/2 + cw_0)y$. Using the fact that X is symmetric and $\beta_{w_0}^0 \in \mathfrak{p}$, we conclude that $x_3\epsilon_1 - x_2\epsilon_2 \in \mathfrak{o}$. Now $\theta(u) = \psi(\mathrm{tr}(SX))$ is equal to

$$\begin{aligned}
&\psi\bigg(\frac{1}{\det(g)}\Big(a\Big(\Big(x - \frac{by}{2}\Big)x_1 - yc(x_3 + w_0x_1)\Big) + b\Big(yax_1 + \Big(x + \frac{by}{2}\Big)(x_3 + w_0x_1)\Big) \\
&\qquad\qquad + c\Big(ya(x_2 + w_0x_1) + \Big(x + \frac{by}{2}\Big)(w_0^2x_1 + w_0(x_2 + x_3) + x_4)\Big)\Big)\bigg) \\
&= \psi\bigg(\frac{1}{\det(g)}\Big(\Big(x + \frac{by}{2}\Big)(x_1\beta_{w_0}^0 + cx_4) + x_2\beta_{w_0}^0yc - x_3\beta_{w_0}^0yc \\
&\qquad\qquad + (x_2\epsilon_2 - x_3\epsilon_1)cw_0 + x_3\epsilon_1(b + 2cw_0)\Big)\bigg) \\
&= 1.
\end{aligned}$$

Here, we have used that $x_3\epsilon_1 - x_2\epsilon_2 \in \mathfrak{o}$, $b+2cw_0 \in \mathfrak{p}$, and ψ is trivial on $\mathfrak{o}$. The other cases are computed in a similar manner. $\square$

3.4. ***Criterion for*** $\mathbf{dim}(\boldsymbol{B}(\boldsymbol{\Lambda},\boldsymbol{\theta})^{\mathbf{I}})=\mathbf{1}$**.** In the previous sections, we have explicitly obtained a well-defined function B, which is right I-invariant and satisfies (3-3). The values of B on the double coset representatives of $R(F)\backslash H(F)/\mathrm{I}$ were obtained, in Proposition 3.8, using one or more of the conditions (3-4)–(3-6). To show that the function B is actually an element of $B(\Lambda,\theta)^{\mathrm{I}}$, we have to show that the conditions (3-4)–(3-6) are satisfied by B for every $h \in H(F)$. In fact, it is sufficient to show that B satisfies these conditions when h is any double coset representative of $R(F)\backslash H(F)/\mathrm{I}$. The computations for checking this are long but not complicated. We will describe the calculation for $h = h(l,m)$ below.

$$B(h(l,m)\eta_0) = B\left(h(l,m)\begin{bmatrix}\varpi&&&\\&\varpi&&\\&&\varpi&\\&&&\varpi\end{bmatrix}h(-1,0)s_2s_1s_2\right)$$
$$= B(h(l-1,m)s_2s_1s_2) = \omega B(h(l,m)).$$

Here, we have used Proposition 3.8 and the identities $A_{l-1,m} = (-\omega q^3)A_{l,m}$. Using the matrix identity

$$\begin{bmatrix}1&&&\\w&1&&\\&&1&-w\\&&&1\end{bmatrix}=\begin{bmatrix}1&w^{-1}&&\\&1&&\\&&1&\\&&-w^{-1}&1\end{bmatrix}s_1\begin{bmatrix}-w&&&\\&-w^{-1}&&\\&&-w^{-1}&\\&&&-w\end{bmatrix}\begin{bmatrix}1&w^{-1}&&\\&1&&\\&&1&\\&&-w^{-1}&1\end{bmatrix}$$

for $w \in \mathfrak{o}$, $w \neq 0$, Lemmas 3.1, 3.2, 3.6 and Proposition 3.8, we get

$$\sum_{w\in\mathfrak{o}/\mathfrak{p}} B(h(l,m)s_1W_ws_1) + B(h(l,m)s_1) = 0.$$

Using the matrix identity

$$\begin{bmatrix}1&&&\\&1&&\\y&&1&\\&&&1\end{bmatrix}=\begin{bmatrix}1&&y^{-1}&\\&1&&\\&&1&\\&&&1\end{bmatrix}s_2\begin{bmatrix}-y&&&\\&1&&\\&&-y^{-1}&\\&&&1\end{bmatrix}\begin{bmatrix}1&&y^{-1}&\\&1&&\\&&1&\\&&&1\end{bmatrix}$$

for $y \in \mathfrak{o}$, $y \neq 0$ and Proposition 3.8, we obtain

$$\sum_{y\in\mathfrak{o}/\mathfrak{p}} B\left(h(l,m)\begin{bmatrix}1&&&\\&1&&\\y&&1&\\&&&1\end{bmatrix}\right) + B(h(l,m)s_2) = 0$$

This shows that, for $h = h(l,m)$, the function B satisfies (3-4)–(3-6), as required. The calculation for other values of h follows in a similar manner. Hence, we get the following theorem.

Theorem 3.11. *Let Λ be a character of $L^\times$. Let $B(\Lambda,\theta)^{\mathrm{I}}$ be the space of smooth functions on $H(F)$, which are right* I*-invariant, satisfy* (3-3) *and the Hecke conditions* (3-4) - (3-6). *Then*

$$\dim\bigl(B(\Lambda,\theta)^{\mathrm{I}}\bigr)=\begin{cases}0 & \textit{if } \Lambda=\Omega\circ N_{L/F} \textit{ and } \left(\frac{L}{\mathfrak{p}}\right)\in\{-1,0\},\\ 1 & \textit{otherwise.}\end{cases}$$

The condition on Λ, in the case $\left(\frac{L}{\mathfrak{p}}\right)\in\{-1,0\}$, follows from cases (iv-a) and (iv-c) of Proposition 3.8.

3.5. ***Existence of a Bessel model.*** We now obtain the existence of a (Λ,θ)-Bessel model for π. When Λ is a unitary character, we act with the Hecke algebra of $H(F)$ on a nonzero function in $B(\Lambda,\theta)^{\mathrm{I}}$. We define an inner product on this Hecke module and also show that the Hecke module has a unique, up to a constant, function which is right I-invariant (the same function that we started with). This leads to the proof that the Hecke module is irreducible and is isomorphic to π, thus giving a (Λ,θ)-Bessel model for π.

When Λ is not unitary (this can happen only if $L=F\oplus F$), we obtain a Bessel model for π using the Whittaker model.

The Hecke module. The Hecke algebra $\mathcal{H}$ of $H(F)$ is the space of all complex-valued functions on $H(F)$ that are locally constant and compactly supported, with the convolution product defined by

$$(f_1*f_2)(g):=\int_{H(F)} f_1(h)\,f_2(h^{-1}g)\,dh \qquad \text{for } f_1,\,f_2\in\mathcal{H},\,g\in H(F).$$

We refer the reader to [Cartier 1979] for details on Hecke algebras of p-adic groups and Hecke modules. Let Λ be a character of $L^\times$ such that $B(\Lambda,\theta)^{\mathrm{I}}\neq 0$. Let $B\in B(\Lambda,\theta)^{\mathrm{I}}$ be the unique, up to a constant, function whose values are described in Proposition 3.8. Define the action of $f\in\mathcal{H}$ on B by

$$(R(f)B)(g):=\int_{H(F)} f(h)\,B(gh)\,dh.$$

This is a finite sum and hence converges for all f. Let

$$V_B:=\{R(f)B: f\in\mathcal{H}\}. \tag{3-32}$$

Since $R(f_1)R(f_2)B=R(f_1*f_2)B$, we see that V_B is a Hecke module. Note that every function in V_B transforms on the left according to $\Lambda\otimes\theta$.

Inner product on Hecke module. We now assume that Λ is a unitary character. Note that, by the comments in the beginning of Section 3.2, if L is a field, then Λ is always unitary. In this case, we will define an inner product on the space V_B.

Lemma 3.12. *The norm $\langle B,B\rangle:=\int_{R(F)\backslash H(F)}|B(h)|^2dh$ is finite.*

Proof. We have

$$\begin{aligned}\langle B, B\rangle &= \sum_{s\in R(F)\backslash H(F)/\mathrm{I}} \int_{R(F)\backslash R(F)s\mathrm{I}} |B(h)|^2 dh \\ &= \sum_{s\in R(F)\backslash H(F)/\mathrm{I}} |B(s)|^2 \int_{\mathrm{I}_s\backslash \mathrm{I}} dh = \sum_{s\in R(F)\backslash H(F)/\mathrm{I}} |B(s)|^2 \frac{\mathrm{vol}(\mathrm{I})}{\mathrm{vol}(\mathrm{I}_s)}.\end{aligned}$$

Here $\mathrm{I}_s := s^{-1}R(F)s\cap \mathrm{I}$. To get the last equality, we argue as in [Pitale and Schmidt 2009b, Lemma 3.7.1]. The volume of I_s can be computed by similar methods to Sections 3.7.1 and 3.7.2 of the same reference. Now, using the values of $B(s)$ from Proposition 3.8 and geometric series, we get the result. □

Let

$$L^2\big(R(F)\backslash H(F), \Lambda\otimes\theta\big) := \left\{\begin{array}{r}\varphi : H(F) \to \mathbb{C} \text{ such that } \varphi \text{ is smooth,} \\ \varphi(rh) = (\Lambda\otimes\theta)(r)\varphi(h) \text{ for } r\in R(F),\ h\in H(F), \\ \text{and } \int_{R(F)\backslash H(F)} |\varphi(h)|^2\, dh < \infty.\end{array}\right\}$$

The previous lemma tells us that $B \in L^2\big(R(F)\backslash H(F), \Lambda\otimes\theta\big)$. It is an easy exercise to see that, in fact, for any $f \in \mathcal{H}$, we have

$$R(f)\, B \in L^2\big(R(F)\backslash H(F), \Lambda\otimes\theta\big).$$

Now, we see that V_B inherits the inner product from $L^2\big(R(F)\backslash H(F), \Lambda\otimes\theta\big)$. For $f_1, f_2 \in \mathcal{H}$, we obtain

$$\langle R(f_1)\, B, R(f_2)\, B\rangle = \int_{R(F)\backslash H(F)} (R(f_1)\, B)(g)\, \overline{(R(f_2)\, B)(g)}\, dg. \tag{3-33}$$

Lemma 3.13. *For $f \in \mathcal{H}$, define $f^* \in \mathcal{H}$ by $f^*(g) = \overline{f(g^{-1})}$. Then, for any $B_1, B_2 \in V_B$,*

$$\langle B_1, R(f)\, B_2\rangle = \langle R(f^*)\, B_1, B_2\rangle.$$

Proof. The lemma follows by a formal calculation. □

Irreducibility of V_B.

Lemma 3.14. *Let V_B^{I} be the subspace of functions in V_B that are right* I*-invariant. Then*

$$\dim(V_B^{\mathrm{I}}) = 1.$$

Proof. We know that V_B^{I} is not trivial since $B \in V_B^{\mathrm{I}}$. Let $\chi_{\mathrm{I}} \in \mathcal{H}$ be the characteristic function of I and set $f_{\mathrm{I}} := \mathrm{vol}(\mathrm{I})^{-1}\chi_{\mathrm{I}}$. Then, by definition, any $B' \in V_B^I$, satisfies $R(f_{\mathrm{I}})\, B' = B'$. Let $f \in \mathcal{H}$ be such that $B' = R(f)\, B = R(f * f_{\mathrm{I}})\, B$. Here, we have used that $B \in V_B^{\mathrm{I}}$. Then

$$B' = R(f_{\mathrm{I}})\, B' = R(f_{\mathrm{I}})\big(R(f * f_{\mathrm{I}})\, B\big) = R(f_{\mathrm{I}} * f * f_{\mathrm{I}})\, B.$$

But $f_{\mathrm{I}} * f * f_{\mathrm{I}} \in \mathscr{H}_{\mathrm{I}}$, the Iwahori Hecke algebra. Since B is an eigenfunction of $\mathscr{H}_{\mathrm{I}}$, we see that $B' \in \mathbb{C}B$. Hence, $\dim(V_B^{\mathrm{I}}) = 1$, as required. □

Proposition 3.15. *Let $\pi = \Omega\,\mathrm{St}_{\mathrm{GSp}_4}$ be the Steinberg representation of $H(F)$, twisted by an unramified quadratic character Ω. Let Λ be a character of $L^\times$ such that* $\dim(B(\Lambda, \theta)^{\mathrm{I}}) = 1$. *Let V_B be as in* (3-32). *If Λ is unitary, then V_B is irreducible and isomorphic to π.*

Proof. We assume, to the contrary, that V_B is reducible. Let W be an $\mathscr{H}$-invariant subspace. Let $W^\perp$ be the complement of W in V_B with respect to the inner product $\langle\ ,\ \rangle$ given in (3-33). Using Lemma 3.13, we see that $W^\perp$ is also $\mathscr{H}$-invariant. Write $B = B_1 + B_2$, with $B_1 \in W$, $B_2 \in W^\perp$. Let f_{I} be as defined in the proof of Lemma 3.14. Since $W, W^\perp$ are $\mathscr{H}$-invariant, we see that $R(f_{\mathrm{I}})B_1 \in W$ and $R(f_{\mathrm{I}})B_2 \in W^\perp$. Since B is right I-invariant, we see that $B_1 = R(f_{\mathrm{I}})B_1$ and $B_2 = R(f_{\mathrm{I}})B_2$. By Lemma 3.14, we obtain, either $B = B_1$ or $B = B_2$. Since V_B is generated by B, we have either $W = V_B$ or $W = 0$. Hence, we see that V_B is an irreducible Hecke module, which contains a unique, up to a constant, vector which is right I-invariant. This uniquely characterizes the Steinberg representation of $H(F)$, and hence, V_B is isomorphic to π. □

Generic representations have split Bessel models. We now assume that Λ is not a unitary character. This can happen only if $L = F \oplus F$. In this case, we will use the fact that $\Omega\mathrm{St}_{\mathrm{GSp}_4}$ is a generic representation. We will now show that any irreducible admissible generic representation of $H(F)$ has a split Bessel model. We believe that this result is known to the experts (for example, see the proof of [Takloo-Bighash 2000, Theorem 2.1]) but we present the details of the proof here.

Let
$$S = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$$
be such that $b^2 - 4ac$ is a square in $F^\times$. One can find a matrix $A \in \mathrm{GL}_2(\mathfrak{o})$ such that
$$S' := {}^t A\, S\, A = \begin{bmatrix} & \frac{1}{2} \\ \frac{1}{2} & \end{bmatrix}.$$
In this case, $T_{S'}(F) := \{g \in \mathrm{GL}_2(F) : {}^t g S' g = \det(g)\, S'\} = A^{-1} T(F)\, A$. The group $T_{S'}(F)$ embedded in $H(F)$ is given by
$$\left\{ \begin{bmatrix} x & & & \\ & y & & \\ & & y & \\ & & & x \end{bmatrix} : x, y \in F^\times \right\}.$$
Let θ' be the character of $U(F)$ obtained from S' and Λ' be the character of $T_{S'}(F)$ obtained from Λ. Then it is easy to see that π has a (Λ, θ)-Bessel model if and

only if it has a (Λ', θ')-Bessel model. So, we will assume that

$$S = \begin{bmatrix} & \frac{1}{2} \\ \frac{1}{2} & \end{bmatrix}.$$

Let (π, V) be an irreducible admissible representation of $H(F)$. For $c_1, c_2 \in F^\times$, consider the character ψ_{c_1,c_2} of the unipotent radical $N_1(F)$ of the Borel subgroup given by

$$\psi_{c_1,c_2}\left(\begin{bmatrix} 1 & x & * & * \\ & 1 & * & y \\ & & 1 & \\ & & -x & 1 \end{bmatrix}\right) = \psi(c_1 x + c_2 y).$$

The representation π of $H(F)$ is called *generic* if $\mathrm{Hom}_{N_1(F)}(\pi, \psi_{c_1,c_2}) \neq 0$. In this case there is an associated Whittaker model $\mathcal{W}(\pi, \psi_{c_1,c_2})$ consisting of functions $H(F) \to \mathbb{C}$ that transform on the left according to ψ_{c_1,c_2}. For $W \in \mathcal{W}(\pi, \psi_{c_1,c_2})$, there is an associated zeta integral

$$Z(s, W) = \int_{F^\times} \int_F W\left(\begin{bmatrix} y & & & \\ & y & & \\ & & 1 & \\ & x & & 1 \end{bmatrix}\right) |y|^{s-3/2} \, dx \, d^\times y.$$

This integral is convergent for $\mathrm{Re}(s) > s_0$, where s_0 is independent of W [Roberts and Schmidt 2007, Proposition 2.6.3]. More precisely, the integral converges to an element of $\mathbb{C}(q^{-s})$, and therefore has meromorphic continuation to all of $\mathbb{C}$. Moreover, there exists an L-factor of the form

$$L(s, \pi) = \frac{1}{Q(q^{-s})}, \qquad Q(X) \in \mathbb{C}[X], \ Q(0) = 1,$$

such that

$$\frac{Z(s, W)}{L(s, \pi)} \in \mathbb{C}[q^{-s}, q^s] \qquad \text{for all } W \in \mathcal{W}(\pi, \psi_{c_1,c_2}). \tag{3-34}$$

(This is proved in [Roberts and Schmidt 2007, Proposition 2.6.4] for π with trivial central character. Also see [Takloo-Bighash 2000, §3.1])

Lemma 3.16. *Let (π, V) be an irreducible admissible generic representation of $H(F)$ with trivial central character. Let σ be a unitary character of $F^\times$, and let $s \in \mathbb{C}$ be arbitrary. Then there exists a nonzero functional $f_{s,\sigma} : V \to \mathbb{C}$ with the following properties.*

i) *For all $x, y, z \in F$ and $v \in V$,*

$$f_{s,\sigma}\left(\pi\left(\begin{bmatrix} 1 & & x & y \\ & 1 & y & z \\ & & 1 & \\ & & & 1 \end{bmatrix}\right) v\right) = \psi(c_1 y) f_{s,\sigma}(v). \tag{3-35}$$

ii) *For all* $x \in F^\times$ *and* $v \in V$,

$$f_{s,\sigma}\left(\pi\left(\begin{bmatrix} x &&& \\ & 1 && \\ && 1 & \\ &&& x \end{bmatrix}\right)v\right) = \sigma(x)^{-1}|x|^{-s+1/2} f_{s,\sigma}(v). \tag{3-36}$$

Proof. We may assume that $V = \mathcal{W}(\pi, \psi_{c_1,c_2})$. Let $s_0 \in \mathbb{R}$ be such that $Z(s, W)$ is absolutely convergent for $\mathrm{Re}(s) > s_0$. Then the integral

$$Z_\sigma(s, W) = \int_{F^\times}\int_F W\left(\begin{bmatrix} y &&& \\ & y && \\ && 1 & \\ & x && 1 \end{bmatrix}\right)|y|^{s-3/2}\sigma(y)\,dx\,d^\times y$$

is also absolutely convergent for $\mathrm{Re}(s) > s_0$, since σ is unitary. Note that these are the zeta integrals for the twisted representation $\sigma\pi$. Therefore, by (3-34), the quotient $Z_\sigma(s, W)/L(s, \sigma\pi)$ is in $\mathbb{C}[q^{-s}, q^s]$ for all $W \in \mathcal{W}(\pi, \psi_{c_1,c_2})$. Now, for $\mathrm{Re}(s) > s_0$, we define

$$f_{s,\sigma}(W) = \frac{Z_\sigma(s, \pi(w)W)}{L(s, \sigma\pi)}, \qquad \text{where } w = \begin{bmatrix} 1 &&& \\ &&& 1 \\ && 1 & \\ & -1 && \end{bmatrix}. \tag{3-37}$$

Straightforward calculations show that (3-35) and (3-36) are satisfied. For general s, since the quotient (3-37) is entire, we can define $f_{s,\sigma}$ by analytic continuation. □

Proposition 3.17. *Let* (π, V) *be an irreducible admissible generic representation of* $H(F)$ *with trivial central character. Then* π *admits a split Bessel functional with respect to any character* Λ *of* $T(F)$ *that satisfies* $\Lambda\big|_{F^\times} \equiv 1$.

Proof. As mentioned earlier, we can take

$$S = \begin{bmatrix} & \frac{1}{2} \\ \frac{1}{2} & \end{bmatrix}.$$

Let $s \in \mathbb{C}$ and σ be a unitary character of $F^\times$ such that

$$\Lambda\left(\begin{bmatrix} x &&& \\ & 1 && \\ && 1 & \\ &&& x \end{bmatrix}\right) = \sigma(x)^{-1}|x|^{-s+1/2} \qquad \text{for all } x \in F^\times.$$

Let $f_{s,\sigma}$ be as in Lemma 3.16. We may assume that $c_1 = 1$, so that $f_{s,\sigma}(\pi(u)v) = \theta(u)v$ for all $u \in U(F)$ by (3-35). We have

$$f_{s,\sigma}\left(\pi\left(\begin{bmatrix} x &&& \\ & 1 && \\ && 1 & \\ &&& x \end{bmatrix}\right)v\right) = \Lambda(x)\, f_{s,\sigma}(v) \qquad \text{for all } x \in F^\times,$$

by (3-36). Since $\Lambda\big|_{F^\times} \equiv 1$ we in fact obtain $f_{s,\sigma}(\pi(t)v) = \Lambda(t) f_{s,\sigma}(v)$ for all $t \in T(F)$. Hence $f_{s,\sigma}$ is a Bessel functional as desired. □

We remark here that, in the split case, for values of $s \in \mathbb{C}$ outside the range of convergence of the zeta integral, we do not have an explicit formula for the Bessel functional. This, in turn, is also reflected in the fact that, when Λ is not unitary, it is not very easy to define an inner product on the space V_B defined in (3-32), although it is known that the Steinberg representation is square-integrable.

Main result on existence and uniqueness of Bessel models.

Theorem 3.18. *Let $\pi = \Omega\,\mathrm{St}_{\mathrm{GSp}_4}$ be the Steinberg representation of $H(F)$, twisted by an unramified quadratic character Ω. Let Λ be a character of $L^\times$ such that $\Lambda|_{F^\times} \equiv 1$. If L is a field, then π has a (Λ, θ)-Bessel model if and only if $\Lambda \neq \Omega \circ N_{L/F}$. If L is not a field, then π always has a (Λ, θ)-Bessel model. In case π has a (Λ, θ)-Bessel model, it is unique.*

In addition, if π has a (Λ, θ)-Bessel model, then the Iwahori spherical vector of π is a test vector for the Bessel functional if and only if Λ satisfies the following conditions.

i) $\Lambda|_{1+\mathfrak{P}} \equiv 1$, *i.e.*, $c(\Lambda) \leq 1$ (*see* (3-10) *for definition of* $c(\Lambda)$).

ii) *If* $\left(\frac{L}{\mathfrak{p}}\right) = 1$ *and* Λ *is unramified*, *then* $\Lambda((1, \varpi)) \neq \Omega(\varpi)$.

Proof. If π has a (Λ, θ)-Bessel model, then it contains a unique vector in $B(\Lambda, \theta)^{\mathrm{I}}$. By Theorem 3.11, the dimension of $B(\Lambda, \theta)^{\mathrm{I}}$ is one, which gives us the uniqueness of Bessel models.

Now we will show the existence of the Bessel model. Let Λ be a character of $L^\times$, with $\Lambda|_{F^\times} \equiv 1$, such that, if L is a field, $\Lambda \neq \Omega \circ N_{L/F}$. We know, by Theorem 3.11, that $\dim(B(\Lambda, \theta)^{\mathrm{I}}) = 1$. If Λ is unitary, Proposition 3.15 tells us that V_B is a (Λ, θ)-Bessel model for π. If Λ is not unitary, we use the fact that π is a generic representation in the split case. Then Proposition 3.17 gives us the result.

The statement regarding the test vector can be deduced from Proposition 3.8 and the fact that a Bessel function B corresponds to a test vector if and only if $B(1) \neq 0$. □

4. Integral representation of the nonarchimedean local L-function

Using the explicit values of the Bessel function obtained in Proposition 3.8, we will now obtain an integral representation of the L-function for the Steinberg representation π of $H(F)$ twisted by any irreducible admissible representation τ of $\mathrm{GL}_2(F)$. For this, we will use the integral obtained in [Furusawa 1993]. We briefly describe the setup.

4.1. ***The unitary group, parabolic induction and the local integral.*** Let $G = \mathrm{GU}(2, 2; L)$ be the unitary similitude group, whose F-points are given by

$$G(F) := \{g \in \mathrm{GL}_4(L) : {}^t\bar{g}Jg = \mu_2(g)J, \mu_2(g) \in F^\times\} \quad \text{where } J = \begin{bmatrix} & 1_2 \\ -1_2 & \end{bmatrix}.$$

Note that $H(F) = G(F) \cap \mathrm{GL}_4(F)$. As a minimal parabolic subgroup we choose the subgroup of all matrices that become upper triangular after switching the last two rows and last two columns. Let P be the standard maximal parabolic subgroup of $G(F)$ with a nonabelian unipotent radical. Let $P = MN$ be the Levi decomposition of P. We have $M = M^{(1)}M^{(2)}$, where

$$M^{(1)}(F) = \left\{ \begin{bmatrix} \zeta &&& \\ & 1 && \\ && \bar{\zeta}^{-1} & \\ &&& 1 \end{bmatrix} : \zeta \in L^\times \right\}, \tag{4-1}$$

$$M^{(2)}(F) = \left\{ \begin{bmatrix} 1 &&& \\ & \alpha && \beta \\ && \mu & \\ & \gamma && \delta \end{bmatrix} \in G(F) \right\}, \tag{4-2}$$

$$N(F) = \left\{ \begin{bmatrix} 1 & z && \\ & 1 && \\ && 1 & \\ && -\bar{z} & 1 \end{bmatrix} \begin{bmatrix} 1 && w & y \\ & 1 & \bar{y} & \\ && 1 & \\ &&& 1 \end{bmatrix} : w \in F,\ y, z \in L \right\}. \tag{4-3}$$

The modular factor of the parabolic P is given by

$$\delta_P\left(\begin{bmatrix} \zeta &&& \\ & 1 && \\ && \bar{\zeta}^{-1} & \\ &&& 1 \end{bmatrix} \begin{bmatrix} 1 &&& \\ & \alpha && \beta \\ && \mu & \\ & \gamma && \delta \end{bmatrix} \right) = \left|N(\zeta)\mu^{-1}\right|^3 \qquad (\mu = \bar{\alpha}\delta - \beta\bar{\gamma}),$$

where $|\cdot|$ is the normalized absolute value on F. Let (τ, V_τ) be an irreducible admissible representation of $\mathrm{GL}_2(F)$, and let χ_0 be a character of $L^\times$ such that $\chi_0\big|_{F^\times}$ coincides with ω_τ, the central character of τ. We assume that V_τ is the Whittaker model of τ with respect to the character ψ^{-c} (we assume that $c \neq 0$). Then the representation $(\lambda, g) \mapsto \chi_0(\lambda)\tau(g)$ of $L^\times \times \mathrm{GL}_2(F)$ factors through $\{(\lambda, \lambda^{-1}) : \lambda \in F^\times\}$, and consequently defines a representation of $M^{(2)}(F)$ on the same space V_τ. Let χ be a character of $L^\times$, considered as a character of $M^{(1)}(F)$. Extend the representation $\chi \times \chi_0 \times \tau$ of $M(F)$ to a representation of $P(F)$ by setting it to be trivial on $N(F)$. If s is a complex parameter, set $I(s, \chi, \chi_0, \tau) = \mathrm{Ind}_{P(F)}^{G(F)}(\delta_P^{s+1/2} \times \chi \times \chi_0 \times \tau)$.

Let (π, V_π) be the twisted Steinberg representation of $H(F)$. We assume that V_π is a Bessel model for π with respect to a character $\Lambda \otimes \theta$ of $R(F)$. Let the characters χ, χ_0 and Λ be related by $\chi(\zeta) = \Lambda(\bar{\zeta})^{-1}\chi_0(\bar{\zeta})^{-1}$. Let $W^\#(\cdot, s)$ be an element of $I(s, \chi, \chi_0, \tau)$ for which the restriction of $W^\#(\cdot, s)$ to the standard maximal compact subgroup of $G(F)$ is independent of s, i.e., $W^\#(\cdot, s)$ is a "flat section" of the family of induced representations $I(s, \chi, \chi_0, \tau)$. By [Pitale and Schmidt 2009b, Lemma 2.3.1], it is meaningful to consider the integral

$$Z(s) = \int_{R(F)\backslash H(F)} W^{\#}(\eta h, s)\, B(h)\, dh, \qquad \eta = \begin{bmatrix} 1 &&& \\ \alpha & 1 && \\ && 1 & -\bar{\alpha} \\ &&& 1 \end{bmatrix}. \tag{4-4}$$

This is the local component of the global integral considered in Section 5.2 below.

4.2. *The* GL_2 *newform.* We define $K^{(0)}(\mathfrak{p}^0) = \mathrm{GL}_2(\mathfrak{o})$ and, for $n > 0$,

$$K^{(0)}(\mathfrak{p}^n) = \mathrm{GL}_2(\mathfrak{o}) \cap \begin{bmatrix} 1+\mathfrak{p}^n & \mathfrak{o} \\ \mathfrak{p}^n & \mathfrak{o}^\times \end{bmatrix}. \tag{4-5}$$

As above, let (τ, V_τ) be a generic, irreducible admissible representation of $\mathrm{GL}_2(F)$ such that V_τ is the ψ^{-c}-Whittaker model of τ. It is well known that V_τ has a unique (up to a constant) vector $W^{(1)}$, called the newform, that is right-invariant under $K^{(0)}(\mathfrak{p}^n)$ for some $n \geq 0$. We then say that τ has conductor $\mathfrak{p}^n$. We normalize $W^{(1)}$ so that $W^{(1)}(1) = 1$. We will need the values of $W^{(1)}$ evaluated at

$$\begin{bmatrix} \varpi^l & \\ & 1 \end{bmatrix},$$

for $l \geq 0$. The following table gives these values (refer to [Schmidt 2002, §2.4]).

τ	$W^{(1)}\left(\begin{bmatrix} \varpi^l & \\ & 1 \end{bmatrix}\right)$
$\alpha \times \beta$ with α and β unramified, $\alpha\beta^{-1} \neq \lvert\cdot\rvert^{\pm 1}$	$q^{-l/2}\dfrac{\alpha(\varpi^{l+1}) - \beta(\varpi^{l+1})}{\alpha(\varpi) - \beta(\varpi)}$
$\alpha \times \beta$ with α unramified, β ramified, $\alpha\beta^{-1} \neq \lvert\cdot\rvert^{\pm 1}$	$\omega_\tau(\varpi^l)\, \alpha(\varpi^{-l})\, q^{-l/2}$
supercuspidal OR ramified twist of Steinberg OR $\alpha \times \beta$ with α, β ramified, $\alpha\beta^{-1} \neq \lvert\cdot\rvert^{\pm 1}$	$\begin{cases} 1 & \text{if } l = 0 \\ 0 & \text{if } l > 0 \end{cases}$
$\Omega' \mathrm{St}_{\mathrm{GL}_2}$, with Ω' unramified	$\Omega'(\varpi^l)\, q^{-l}$

We extend $W^{(1)}$ to a function on $M^{(2)}(F)$ via $W^{(1)}(ag) = \chi_0(a) W^{(1)}(g)$ for $a \in L^\times$, $g \in \mathrm{GL}_2(F)$.

4.3. *Choice of* Λ *and* $W^{\#}$. We will choose a character Λ of $L^\times$ such that π has a (Λ, θ)-Bessel model and the Iwahori spherical vector is a test vector for the Bessel functional. Noting that $\Lambda|_{F^\times}$ is the central character of π and using Theorem 3.18, we impose the following conditions on Λ:

i) $\Lambda|_{F^\times} \equiv 1$.

ii) $c(\Lambda) \leq 1$.

iii) $\Lambda \neq \Omega \circ N_{L/F}$ in case L is a field.

iv) $\omega\Lambda((1, \varpi)) \neq -1$ in case L is not a field and $c(\Lambda) = 0$.

Note that this implies that $\Lambda|_{\mathfrak{o}^\times + \mathfrak{P}} \equiv 1$. For $n \geq 1$, let $\Gamma(\mathfrak{P}^n)$ be the principal congruence subgroup of the maximal compact subgroup $K^G := G(\mathfrak{o})$ of $G(F)$, defined by

$$\Gamma(\mathfrak{P}^n) := \{g \in K^G : g \equiv 1 \pmod{\mathfrak{P}^n}\}.$$

The next lemma will be crucial for the well-definedness of $W^\#$ below.

Lemma 4.1. *Let (τ, V_τ) be a generic, irreducible admissible representation of $\mathrm{GL}_2(F)$ with conductor $\mathfrak{p}^n$, $n \geq 0$. Set $n_0 = \max\{1, n\}$ and let*

$$\hat{m} = \begin{bmatrix} \zeta & & & \\ & a' & & b' \\ & & \mu\bar{\zeta}^{-1} & \\ & c' & & d' \end{bmatrix} \in M(F) \quad \textit{and} \quad \hat{n} = \begin{bmatrix} 1 & z & & \\ & 1 & & \\ & & 1 & \\ & & -\bar{z} & 1 \end{bmatrix} \begin{bmatrix} 1 & & w & y \\ & 1 & \bar{y} & \\ & & 1 & \\ & & & 1 \end{bmatrix} \in N(F).$$

Suppose that $A := \eta^{-1}\hat{m}\hat{n}\eta$ lies in $I\Gamma(\mathfrak{P}^{n_0})$. Then

i) *$c' \in \mathfrak{P}^{n_0}$ and $a'\bar{\zeta}^{-1} \in 1 + \mathfrak{P}^{n_0}$, and*

ii) *for any $\begin{bmatrix} a_1' & b_1' \\ c_1' & d_1' \end{bmatrix} \in \mathrm{GU}(1, 1; L)(F)$, we have*

$$\chi(\zeta) W^{(1)}\left(\begin{bmatrix} a_1' & b_1' \\ c_1' & d_1' \end{bmatrix}\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right) = W^{(1)}\left(\begin{bmatrix} a_1' & b_1' \\ c_1' & d_1' \end{bmatrix}\right).$$

Proof. Using Lemma 2.1, it is easy to show that for $n \geq 0$

$$x \in \mathfrak{o} + \mathfrak{P}^n \text{ and } \alpha x \in \mathfrak{o} + \mathfrak{P}^n \;\Rightarrow\; x \in \mathfrak{P}^n. \tag{4-6}$$

First note that $I\Gamma(\mathfrak{P}^{n_0}) \subset M_4(\mathfrak{o} + \mathfrak{P}^{n_0})$. Looking at the (4, 1), (4, 2) coefficient of A, we see that $c', \alpha c' \in \mathfrak{o} + \mathfrak{P}^{n_0}$. By (4-6), we obtain $c' \in \mathfrak{P}^{n_0}$, as required.

Observe that $\hat{m}\hat{n} \in K^G$ and $c' \in \mathfrak{P}^{n_0} \subset \mathfrak{P}$ implies that $\zeta, a', d' \in \mathfrak{o}_L^\times$. The upper left 2×2 block of A is given by

$$\begin{bmatrix} \zeta + \alpha z\zeta & z\zeta \\ \alpha a' - \alpha(\zeta + \alpha z\zeta)) & a' - \alpha z\zeta \end{bmatrix}.$$

We will repeatedly use the following fact:

$$\text{If } x \in \mathfrak{o} + \mathfrak{P}^{n_0}, \text{ then } x \equiv \bar{x} \pmod{(\alpha - \bar{\alpha})\mathfrak{P}^{n_0}}.$$

Applying this to the matrix entries of A, we get $z\zeta \equiv \bar{z}\bar{\zeta} \pmod{(\alpha - \bar{\alpha})\mathfrak{P}^{n_0}}$, and then

$$\begin{aligned} a' - \bar{a}' &\equiv (\alpha - \bar{\alpha})z\zeta \pmod{(\alpha - \bar{\alpha})\mathfrak{P}^{n_0}}, \\ \zeta - \bar{\zeta} &\equiv (\bar{\alpha} - \alpha)z\zeta \pmod{(\alpha - \bar{\alpha})\mathfrak{P}^{n_0}}. \end{aligned} \tag{4-7}$$

Using $\zeta + \alpha z \zeta \equiv \bar{\zeta} + \bar{\alpha}\bar{z}\bar{\zeta} \pmod{(\alpha - \bar{\alpha})\mathfrak{P}^{n_0}}$ and (4-7), we get from the $(2, 1)$ coefficient of A that

$$(a' - \bar{\zeta})(\alpha - \bar{\alpha}) \equiv 0 \pmod{(\alpha - \bar{\alpha})\mathfrak{P}^{n_0}}.$$

Hence $a' - \bar{\zeta} \equiv 0 \pmod{\mathfrak{P}^{n_0}}$, so that $a'\bar{\zeta}^{-1} \in 1 + \mathfrak{P}^{n_0}$, as required. This proves part (i) of the lemma.

Looking at the $(1, 2)$ coefficient of A, we see that $z\zeta \in \mathfrak{P}$. Looking at the $(1, 1)$ coefficient of A, we see that $\zeta \in \mathfrak{o}^\times + \mathfrak{P}$.

$$\begin{aligned}
&\chi(\zeta) W^{(1)}\left(\begin{bmatrix} a_1' & b_1' \\ c_1' & d_1' \end{bmatrix}\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right) \\
&\quad = \chi(\zeta)\chi_0(a') W^{(1)}\left(\begin{bmatrix} a_1' & b_1' \\ c_1' & d_1' \end{bmatrix}\begin{bmatrix} 1 & b'/a' \\ c'/a' & d'/a' \end{bmatrix}\right) \\
&\quad = \Lambda(\bar{\zeta}^{-1})\chi_0(\bar{\zeta}^{-1})\chi_0(a') W^{(1)}\left(\begin{bmatrix} a_1' & b_1' \\ c_1' & d_1' \end{bmatrix}\begin{bmatrix} 1 & b'/a' \\ c'/a' & d'/a' \end{bmatrix}\right) = W^{(1)}\left(\begin{bmatrix} a_1' & b_1' \\ c_1' & d_1' \end{bmatrix}\right).
\end{aligned}$$

Here we have used the fact that Λ is trivial on $\mathfrak{o}^\times + \mathfrak{P}$, χ_0 is trivial on $1 + \mathfrak{P}^{n_0}$ and the matrix

$$\begin{bmatrix} 1 & b'/a' \\ c'/a' & d'/a' \end{bmatrix}$$

lies in $K^{(0)}(\mathfrak{p}^{n_0})$. □

Let $n_0 = \max\{1, n\}$, as above. Given a complex number s, define the function $W^\#(\,\cdot\,, s) : G(F) \to \mathbb{C}$ as follows.

i) If $g \notin M(F)N(F)\eta \mathrm{I}\Gamma(\mathfrak{P}^{n_0})$, then $W^\#(g, s) = 0$.

ii) If $g = mn\eta k\gamma$ with $m \in M(F)$, $n \in N(F)$, $k \in \mathrm{I}$, $\gamma \in \Gamma(\mathfrak{P}^{n_0})$, then $W^\#(g, s) = W^\#(m\eta, s)$.

iii) For $\zeta \in L^\times$ and $\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \in M^{(2)}(F)$,

$$W^\#\left(\begin{bmatrix} \zeta & & & \\ & 1 & & \\ & & \bar{\zeta}^{-1} & \\ & & & 1 \end{bmatrix}\begin{bmatrix} 1 & & & \\ & a' & & b' \\ & & \mu & \\ & c' & & d' \end{bmatrix}\eta, s\right) = \left|N(\zeta)\cdot\mu^{-1}\right|^{3(s+1/2)}\chi(\zeta)\, W^{(1)}\left(\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right), \tag{4-8}$$

where $\mu = \bar{a}'d' - b'\bar{c}'$.

By Lemma 4.1, we see that $W^\#$ is well-defined. It is an element of $I(s, \chi, \chi_0, \tau)$.

4.4. *Support of $W^\#$.* We choose $W^\#$ as above and B as in Proposition 3.8, with $B(1) = 1$. Note that $B(1) \neq 0$ by the comments in the beginning of Section 4.3.

Then the integral (4-4) becomes

$$Z(s)=\sum_{\substack{l\in\mathbb{Z}\\ m\geq 0}}\sum_{t} W^{\#}(\eta h(l,m)t,s)\,B(h(l,m)t)\,V_t^{l,m}, \tag{4-9}$$

where t runs through the double coset representatives from Proposition 3.3 and

$$V_t^{l,m}=\mathrm{vol}\big(R(F)\setminus R(F)\,h(l,m)\,t\,\mathrm{I}\big).$$

To compute (4-9), we need to find out for what values of l, m, t is $\eta h(l,m)t$ in the support of $W^{\#}$. Write $\eta h(l,m)=h(l,m)\,\eta_m$, where

$$\eta_m=\begin{bmatrix}1&&&\\ \varpi^m\alpha&1&&\\ &&1&-\varpi^m\bar{\alpha}\\ &&&1\end{bmatrix}.$$

Since $h(l,m)\in M(F)$, we need to know for which values of m, t is $\eta_m t$ in the support of $W^{\#}$. This is done in the following lemma.

Lemma 4.2. *Let t be any double coset representative from* Proposition 3.3. *Then $\eta_m t$ lies in the support, $MN\eta\,\mathrm{I}\Gamma(\mathfrak{P}^{n_0})$, of $W^{\#}$ if and only if $m=0$ and $t=1$.*

Proof. We first consider the case $m>0$. Note that it is enough to show that $\eta_m t\notin MN\eta\,\mathrm{I}\Gamma(\mathfrak{P})$. For any double coset representative t, we have $t^{-1}\eta_m t\equiv 1 \pmod{\mathfrak{P}}$ and hence $t^{-1}\eta_m t\in\Gamma(\mathfrak{P})$. So it is enough to show that $t\notin MN\eta\,\mathrm{I}\Gamma(\mathfrak{P})$ for any t. Suppose there are $\hat{m}\in M$, $\hat{n}\in N$ such that $A=\eta^{-1}\hat{m}\hat{n}t\in\mathrm{I}\Gamma(\mathfrak{P})$. Using $\hat{m},\hat{n}\in K^G$ and

$$\mathrm{I}\Gamma(\mathfrak{P})\subset\begin{bmatrix}\mathfrak{o}+\mathfrak{P}&\mathfrak{P}&\mathfrak{o}+\mathfrak{P}&\mathfrak{o}+\mathfrak{P}\\ \mathfrak{o}+\mathfrak{P}&\mathfrak{o}+\mathfrak{P}&\mathfrak{o}+\mathfrak{P}&\mathfrak{o}+\mathfrak{P}\\ \mathfrak{P}&\mathfrak{P}&\mathfrak{o}+\mathfrak{P}&\mathfrak{o}+\mathfrak{P}\\ \mathfrak{P}&\mathfrak{P}&\mathfrak{P}&\mathfrak{o}+\mathfrak{P}\end{bmatrix} \tag{4-10}$$

we get a contradiction for every $t\in W$. We now consider the case $m=0$. First let $t=1$. Taking $\hat{m}=\hat{n}=1$, we easily see that $\eta\in MN\eta\,\mathrm{I}\Gamma(\mathfrak{P}^{n_0})$, as required. Now assume that $t\neq 1$. Suppose, there are $\hat{m}\in M$, $\hat{n}\in N$ such that $A=\eta^{-1}\hat{m}\hat{n}\eta t\in\mathrm{I}\Gamma(\mathfrak{P})$. Again, using $\hat{m},\hat{n}\in K^G$ and (4-10) we get a contradiction for $t\neq 1$. This completes the proof of the lemma. □

4.5. *Integral computation.* From Lemma 4.2, we see that the integral (4-9) is equal to

$$Z(s)=\sum_{l\geq 0} W^{\#}(\eta h(l,0),s)\,B(h(l,0))\,V_1^{l,0}. \tag{4-11}$$

Arguing as in [Furusawa 1993, §3.5], we get

$$V_1^{l,0} = \frac{\left(1-\left(\frac{L}{\mathfrak{p}}\right)q^{-1}\right)q}{(1+q)^2(1+q^2)}q^{3l}.$$

From Proposition 3.8 and (4-8), we get $B(h(l,0)) = (-\omega q^{-3})^l$ and

$$W^{\#}(\eta h(l,0), s) = q^{-3(s+1/2)l}\omega_\tau(\varpi^{-l})\, W^{(1)}\left(\begin{bmatrix}\varpi^l & \\ & 1\end{bmatrix}\right).$$

We set

$$C = \frac{\left(1-\left(\frac{L}{\mathfrak{p}}\right)q^{-1}\right)q}{(1+q)^2(1+q^2)}.$$

We have

$$Z(s) = C\sum_{l\geq 0}(-\omega)^l q^{-3(s+1/2)l}\omega_\tau(\varpi^{-l})\, W^{(1)}\left(\begin{bmatrix}\varpi^l & \\ & 1\end{bmatrix}\right). \tag{4-12}$$

We will now substitute the value of $W^{(1)}$, from the table obtained in Section 4.2, into (4-12) for all possible GL_2 representations τ.

(4-13)

$$Z(s) = \begin{cases} C\left(1+\omega\alpha(\varpi^{-1})q^{-3s-2}\right)^{-1}\left(1+\omega\beta(\varpi^{-1})q^{-3s-2}\right)^{-1} \\ \qquad\qquad \text{if } \tau = \alpha\times\beta,\ \alpha,\beta \text{ unramified, } \alpha\beta^{-1}\neq|\cdot|^{\pm 1}; \\ C\left(1+\omega\alpha(\varpi^{-1})q^{-3s-2}\right)^{-1} \\ \qquad\qquad \text{if } \tau = \alpha\times\beta,\ \alpha \text{ unramified, } \beta \text{ ramified } \alpha\beta^{-1}\neq|\cdot|^{\pm 1}; \\ C\left(1+\omega\Omega'(\varpi^{-1})q^{-3s-5/2}\right)^{-1} \\ \qquad\qquad \text{if } \tau = \Omega'\mathrm{St}_{\mathrm{GL}_2},\ \Omega' \text{ unramified;} \\ C \qquad\qquad \text{otherwise.} \end{cases}$$

Let $\tilde{\tau}$ denote the contragredient of the representation τ. We get the following theorem on the integral representation of L-functions.

Theorem 4.3. *Let*

$$\pi = \Omega\,\mathrm{St}_{\mathrm{GSp}_4}$$

be the Steinberg representation of $\mathrm{GSp}_4(F)$ *twisted by an unramified quadratic character* Ω. *Let* τ *be any irreducible admissible representation of* $\mathrm{GL}_2(F)$. *Let* $Z(s)$ *be the integral defined in* (4-4). *Choose* B *as in Section 3 and* $W^{\#}$ *as in Section 4.3. Then we have*

$$Z(s) = Y'(s)\,L\left(3s+\tfrac{1}{2},\ \pi\times\tilde{\tau}\right), \tag{4-14}$$

where

$$Y'(s) = \begin{cases} C\left(1-\Omega(\varpi)\,\Omega'(\varpi^{-1})q^{-3s-3/2}\right) & \text{if } \tau = \Omega'\,\mathrm{St}_{\mathrm{GL}_2},\ \Omega' \text{ unramified;} \\ C & \text{otherwise.} \end{cases}$$

Here,

$$C=\frac{\left(1-\left(\frac{L}{\mathfrak{p}}\right)q^{-1}\right)q}{(1+q)^2(1+q^2)}.$$

Proof. This follows from (4-13) and from the following definition of L-functions for the representation $\pi=\Omega\,\mathrm{St}_{\mathrm{GSp}_4}$, with Ω unramified and quadratic, twisted by $\tilde{\tau}$:

$$L(s,\pi\times\tilde{\tau})=\begin{cases}\left(1-\Omega(\varpi)\,\alpha(\varpi^{-1})q^{-s-3/2}\right)^{-1}\left(1-\Omega(\varpi)\,\beta(\varpi^{-1})q^{-s-3/2}\right)^{-1}\\ \qquad\qquad \text{if } \tau=\alpha\times\beta,\ \alpha,\beta \text{ unramified, } \alpha\beta^{-1}\neq|\cdot|^{\pm1};\\ \left(1-\Omega(\varpi)\,\alpha(\varpi^{-1})q^{-s-3/2}\right)^{-1}\\ \qquad\qquad \text{if } \tau=\alpha\times\beta,\ \alpha \text{ unramified, } \beta \text{ ramified } \alpha\beta^{-1}\neq|\cdot|^{\pm1};\\ \left(1-\Omega(\varpi)\,\Omega'(\varpi^{-1})q^{-s-1}\right)^{-1}\left(1-\Omega(\varpi)\,\Omega'(\varpi^{-1})q^{-s-2}\right)^{-1}\\ \qquad\qquad \text{if } \tau=\Omega'\mathrm{St}_{\mathrm{GL}_2},\ \Omega' \text{ unramified;}\\ 1 \qquad\qquad \text{otherwise.}\end{cases}$$

□

5. Global theory

In the previous section, we computed the nonarchimedean integral representation of the L-function $L(s,\pi\times\tilde{\tau})$ for the Steinberg representation of GSp_4 twisted by any GL_2 representation. In [Furusawa 1993], the integral has been computed for both π and τ unramified. In [Pitale and Schmidt 2009c], the integral has been calculated for an unramified representation π twisted by any ramified GL_2 representation τ. In the same paper, the archimedean integral was computed for π_∞ a holomorphic (or limit of holomorphic) discrete series representation with scalar minimal K-type, and τ_∞ any representation of $\mathrm{GL}_2(\mathbb{R})$. In this section, we will put together all the local computations and obtain an integral representation of a global L-function. We will start with a Siegel cuspidal newform F of weight l with respect to the Borel congruence subgroup of square-free level. We will obtain an integral representation of the L-function of F twisted by any irreducible cuspidal automorphic representation τ of $\mathrm{GL}_2(\mathbb{A})$. When τ is obtained from a holomorphic cusp form of the same weight l as F, we obtain a special value result for the L-function, in the spirit of Deligne's conjectures.

5.1. ***Siegel modular form and Bessel model.*** Let M be a square-free positive integer and l be any positive integer. Set

$$B(M):=\left\{g\in\mathrm{Sp}_4(\mathbb{Z}):g\equiv\begin{bmatrix}*&0&*&*\\ *&*&*&*\\ 0&0&*&*\\ 0&0&0&*\end{bmatrix}\pmod M\right\}.$$

Let F be a Siegel newform of weight l with respect to $B(M)$. We refer the reader to [Saha 2009, §8] or [Schmidt 2005] for definition and details on newforms with square-free level. The Fourier expansion of F is given by

$$F(Z) = \sum_{T>0} A(T)\, e^{2\pi i \operatorname{tr}(TZ)},$$

where T runs over all semi-integral, symmetric, positive definite 2×2 matrices. We obtain a well-defined function $\Phi = \Phi_F$ on $H(\mathbb{A})$, where $\mathbb{A}$ is the ring of adeles of $\mathbb{Q}$, by

$$\Phi(\gamma h_\infty k_0) = \mu_2(h_\infty)^l \det(J(h_\infty, i1_2))^{-l} F(h_\infty\langle i1_2\rangle),$$

where $\gamma \in H(\mathbb{Q})$, $h_\infty \in H^+(\mathbb{R})$, $k_0 \in \prod_{p\nmid M} H(\mathbb{Z}_p) \prod_{p|M} \mathrm{I}_p$. Let V_F be the space generated by the right translates of Φ_F and let π_F be one of the irreducible components. Then $\pi_F = \otimes \pi_p$, where π_∞ is a holomorphic discrete series representation of $H(\mathbb{R})$ of lowest weight (l, l), for a finite prime $p \nmid M$, π_p is an irreducible, unramified representation of $H(\mathbb{Q}_p)$, and for $p|M$, π_p is a twist $\Omega_p \mathrm{St}_{\mathrm{GSp}_4}$ of the Steinberg representation of $H(\mathbb{Q}_p)$ by an unramified quadratic character Ω_p.

For a positive integer $D \equiv 0, 3 \pmod 4$, set

$$S(-D) = \begin{cases} \begin{bmatrix} \frac{1}{4}D & 0 \\ 0 & 1 \end{bmatrix} & \text{if } D \equiv 0 \pmod 4, \\ \begin{bmatrix} \frac{1}{4}(1+D) & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} & \text{if } D \equiv 3 \pmod 4. \end{cases}$$

Let $L = \mathbb{Q}(\sqrt{-D})$ and $T(\mathbb{A}) \simeq \mathbb{A}_L^\times$ be the adelic points of the group defined in (3-1). Let $R(\mathbb{A}) = T(\mathbb{A})\, U(\mathbb{A})$ be the Bessel subgroup of $H(\mathbb{A})$. Let Λ be a character of

$$T(\mathbb{A})/T(\mathbb{Q})\, T(\mathbb{R}) \prod_{p\nmid M} T(\mathbb{Z}_p) \prod_{p|M} T_p^0, \tag{5-1}$$

where $T(\mathbb{Z}_p) = T(\mathbb{Q}_p) \cap \mathrm{GL}_2(\mathbb{Z}_p)$ and $T_p^0 = T(\mathbb{Z}_p) \cap \Gamma_p^0$. Here

$$\Gamma_p^0 = \left\{ g \in \mathrm{GL}_2(\mathbb{Z}_p) : g \equiv \begin{bmatrix} * & 0 \\ * & * \end{bmatrix} \pmod{p\mathbb{Z}_p} \right\}.$$

Note that, under the isomorphism (3-2), T_p^0 corresponds to $\mathbb{Z}_p^\times + p\mathfrak{o}_{L_p}$, where $\mathfrak{o}_{L_p}$ is the ring of integers of the two dimensional algebra $L \otimes_{\mathbb{Q}} \mathbb{Q}_p$. Let ψ be a character of $\mathbb{Q} \setminus \mathbb{A}$ that is trivial on $\mathbb{Z}_p$ for all primes p and satisfies $\psi(x) = e^{-2\pi i x}$ for all $x \in \mathbb{R}$. We define the global Bessel function of type (Λ, θ) associated to $\bar{\Phi}$ by

$$B_{\bar{\Phi}}(h) = \int_{Z_H(\mathbb{A})\, R(\mathbb{Q}) \setminus R(\mathbb{A})} (\Lambda \otimes \theta)(r)^{-1} \bar{\Phi}(rh)\, dr,$$

where

$$\theta\left(\begin{bmatrix} 1 & X \\ & 1 \end{bmatrix}\right) = \psi(\operatorname{tr}(SX)) \qquad \text{and} \qquad \bar{\Phi}(h) = \overline{\Phi(h)}.$$

If $B_{\bar\Phi}$ is nonzero, then $B_{\bar\varphi}$ is nonzero for any $\varphi \in \pi_F$. We say that π_F has a global Bessel model of type (Λ, θ) if $B_{\bar\Phi} \neq 0$. We shall make the following assumption on the representation π_F.

Assumption. π_F has a global Bessel model of type (Λ, θ) such that

A1. $-D$ is the fundamental discriminant of $\mathbb{Q}(\sqrt{-D})$.

A2. Λ is a character of (5-1).

A3. For $p|M$, if $L \otimes \mathbb{Q}_p$ is split and Λ_p is unramified, then

$$\Omega_p(\varpi_p)\,\Lambda_p((1, \varpi_p)) \neq 1.$$

Remark 5.1. In [Furusawa 1993; Pitale and Schmidt 2009b; 2009c; Saha 2009], nonvanishing of a suitable Fourier coefficient of F is assumed, while in [Pitale and Schmidt 2009a], the existence of a suitable global Bessel model for π_F is assumed. We explain the relation of the assumption above to nonvanishing of certain Fourier coefficients of F. Let $\{t_j\}$ be a set of representatives for (5-1). One can take $t_j \in \mathrm{GL}_2(\mathbb{A}_\mathrm{f})$. Write

$$t_j = \gamma_j m_j \kappa_j,$$

with $\gamma_j \in \mathrm{GL}_2(\mathbb{Q})$, $m_j \in \mathrm{GL}_2^+(\mathbb{R})$ and $\kappa_j \in \prod_{p\nmid M} \mathrm{GL}_2(\mathbb{Z}_p) \prod_{p|M} \Gamma_p^0$. Set

$$S_j := \det(\gamma_j)^{-1}\, {}^t\gamma_j S(-D) \gamma_j.$$

Note that $\{S_j\}_j$ is a *subset* of the set of representatives of $\Gamma^0(M)$ equivalence classes of primitive, semi-integral positive definite 2×2 matrices of discriminant $-D$.

From [Saha 2009] or [Sugano 1985], we have, for $h_\infty \in H^+(\mathbb{R})$,

$$B_{\bar\Phi}(h_\infty) = \mu_2(h_\infty)^l\, \overline{\det(J(h_\infty, I))^{-l}}\, e^{-2\pi i\, \mathrm{tr}(S(-D)\, \overline{h_\infty\langle I\rangle})} \sum_j \Lambda(t_j)^{-1} \overline{A(S_j)}, \tag{5-2}$$

and $B_{\bar\Phi}(h_\infty) = 0$ for $h_\infty \notin H^+(\mathbb{R})$. Suppose that there is a semi-integral, symmetric, positive definite 2×2 matrix T satisfying

i) $-D = \det(2T)$ is the fundamental discriminant of $L = \mathbb{Q}(\sqrt{-D})$.

ii) T is $\Gamma^0(M)$ equivalent to one of the S_j.

iii) The Fourier coefficient $A(T) \neq 0$.

Then it is clear from (5-2) that one can choose a Λ such that parts A1 and A2 of the assumption are satisfied. If $M = 1$ (as in [Furusawa 1993; Pitale and Schmidt 2009b; 2009c]) or, every prime $p|M$ is inert in L (as in [Saha 2009]), then $\{S_j\}_j$ is the *complete* set of representatives of $\Gamma^0(M)$ equivalence classes and hence, condition (i) above implies condition (ii) to give the assumption from [Furusawa 1993; Pitale and Schmidt 2009b; 2009c] and [Saha 2009]. We have to include part

A3 of the assumption to guarantee that the Iwahori spherical vector in π_p, for $p|M$, is a test vector for the Bessel functional.

We abbreviate $a(\Lambda) = \sum \Lambda(t_j)\, A(S_j)$. For $h \in H(\mathbb{A})$, we have

$$B_{\bar{\Phi}}(h) = \overline{a(\Lambda)} \prod_p B_p(h_p),$$

where B_∞ is as defined in [Pitale and Schmidt 2009c], for a finite prime $p \nmid M$, B_p is the spherical vector in the (Λ_p, θ_p)-Bessel model for π_p, and for $p|M$, B_p is the vector in the (Λ_p, θ_p)-Bessel model for π_p defined by Proposition 3.8 and 3.10. For $p < \infty$, we have normalized the B_p so that $B_p(1) = 1$.

5.2. ***Global induced representation and global integral.*** Let $\tau = \bigotimes \tau_p$ be an irreducible cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A})$ with central character ω_τ. For every prime $p < \infty$, let p^{n_p} be the conductor of τ_p. For almost all p, we have $n_p = 0$. Set $N = \prod_p p^{n_p}$. Choose l_1 to be any weight occurring in τ_∞. Let χ_0 be a character of $\mathbb{A}_L^\times$ such that $\chi_0|_{\mathbb{A}^\times} = \omega_\tau$ and $\chi_{0,\infty}(\zeta) = \zeta^{l_2}$ for any $\zeta \in S^1$. Here, l_2 depends on l_1 and l by the formula

$$l_2 = \begin{cases} l_1 - 2l & \text{if } l \le l_1, \\ -l_1 & \text{if } l \ge l_1, \end{cases}$$

as in [Pitale and Schmidt 2009c]. The existence of such a character is guaranteed by Lemma 5.3.1 of that reference. Define another character χ of $\mathbb{A}_L^\times$ by

$$\chi(\zeta) = \chi_0(\bar{\zeta})^{-1} \Lambda(\bar{\zeta})^{-1}.$$

Let $I(s, \chi_0, \chi, \tau)$ be the induced representation of $G(\mathbb{A})$ obtained in an analogous way to the local situation in Sect. 4.1. We will now define a global section $f_\Lambda(g, s)$. We realize the representation τ as a subspace of $L^2\big(\mathrm{GL}_2(\mathbb{Q}) \setminus \mathrm{GL}_2(\mathbb{A})\big)$ and let $\hat{f}$ be the automorphic cusp form such that the space of τ is generated by the right translates of $\hat{f}$. The function $\hat{f}$ corresponds to a cuspidal Hecke newform on the complex upper half plane. Then, $\hat{f}$ is factorizable. Write $\hat{f} = \otimes \hat{f}_p$ such that $\hat{f}_\infty$ is the function of weight l_1 in τ_∞. For $p < \infty$, $\hat{f}_p$ is the unique newform in τ_p with $\hat{f}_p(1) = 1$. Using χ_0, extend $\hat{f}$ to a function of $\mathrm{GU}(1, 1; L)(\mathbb{A})$.

For a finite prime p, set

$$K_p^G := \begin{cases} G(\mathbb{Z}_p) & \text{if } p \nmid MN; \\ \mathrm{I}\Gamma((p\mathfrak{o}_{L_p})^{n_{p,0}}) & \text{if } p|M; \\ H(\mathbb{Z}_p)\Gamma((p\mathfrak{o}_{L_p})^{n_p}) & \text{if } p|N,\ p \nmid M. \end{cases}$$

Here, in the second case, $n_{p,0} = \max(1, n_p)$. Set $K^G(M, N) = \prod_{p<\infty} K_p^G$ and let K_∞ be the maximal compact subgroup of $G(\mathbb{R})$. Let η be the element of $G(\mathbb{Q})$ defined in (4-4). Let $\eta_{M,N}$ be the element of $G(\mathbb{A})$ such that the p-component is given by η for $p|MN$ and by 1 for $p \nmid MN$. For $s \in \mathbb{C}$, define $f_\Lambda(\,\cdot\,, s)$ on $G(\mathbb{A})$ by

i) $f_\Lambda(g, s) = 0$ if $g \notin M(\mathbb{A})\, N(\mathbb{A})\, \eta_{M,N} K_\infty K^G(M, N)$.

ii) If $m = m_1 m_2$, $m_i \in M^{(i)}(\mathbb{A})$, $n \in N(\mathbb{A})$, $k = k_0 k_\infty$, $k_0 \in K^G(M, N)$, $k_\infty \in K_\infty$, then

$$f_\Lambda(mn\eta_{M,N}k, s) = \delta_P^{1/2+s}(m)\, \chi(m_1)\, \hat{f}(m_2)\, f(k_\infty). \tag{5-3}$$

Recall that $\delta_P(m_1 m_2) = \left|N_{L/\mathbb{Q}}(m_1)\mu_1(m_2)^{-1}\right|^3$.

Here, $M^{(1)}(\mathbb{A})$, $M^{(2)}(\mathbb{A})$, $N(\mathbb{A})$ are the adelic points of the algebraic groups defined by (4-1)–(4-3) and f is the function on K_∞ defined in [Pitale and Schmidt 2009c]. As in [Pitale and Schmidt 2009c], it can be checked that f_Λ is well-defined. For $\mathrm{Re}(s)$ large enough we can form the Eisenstein series

$$E(g, s; f_\Lambda) := \sum_{\gamma \in P(\mathbb{Q})\backslash G(\mathbb{Q})} f_\Lambda(\gamma g, s).$$

In fact, $E(g, s; f_\Lambda)$ has a meromorphic continuation to the entire plane. In [Furusawa 1993], Furusawa studied integrals of the form

$$Z(s, f_\Lambda, \varphi) = \int_{H(\mathbb{Q})\, Z_H(\mathbb{A})\backslash H(\mathbb{A})} E(h, s; f_\Lambda)\, \varphi(h)\, dh, \tag{5-4}$$

where $\varphi \in V_\pi$. Theorem 2.4 of [Furusawa 1993], the "basic identity", states that

$$Z(s, f_\Lambda, \varphi) = \int_{R(\mathbb{A})\backslash H(\mathbb{A})} W_{f_\Lambda}(\eta h, s)\, B_\varphi(h)\, dh, \tag{5-5}$$

where B_φ is the Bessel function corresponding to φ and W_{f_Λ} is the function defined by

$$W_{f_\Lambda}(g) = \int_{\mathbb{Q}\backslash\mathbb{A}} f_\Lambda\left(\begin{bmatrix} 1 &&& \\ & 1 && x \\ && 1 & \\ &&& 1 \end{bmatrix} g\right) \psi(cx)\, dx, \qquad g \in G(\mathbb{A}).$$

The function W_{f_Λ} is a pure tensor and we can write

$$W_{f_\Lambda}(g, s) = \prod_p W_p^{\#}(g_p, s).$$

Then we see that $W_\infty^{\#}$ is as defined in [Pitale and Schmidt 2009c]. For a finite prime $p \nmid M$, the $W_p^{\#}$ is the function defined in Section 4.5 of that reference. For $p|M$, the $W_p^{\#}$ is as in Section 4.3. It follows from (5-5) that

$$Z(s, f_\Lambda, \bar{\Phi}) = \prod_{p \le \infty} Z_p(s, W_p^{\#}, B_p),$$

where

$$Z_p(s, W_p^\#, B_p) = \int_{R(\mathbb{Q}_p)\backslash H(\mathbb{Q}_p)} W_p^\#(\eta h, s)\, B_p(h)\, dh.$$

When $p \nmid MN$, $p < \infty$, the integral Z_p is evaluated in [Furusawa 1993]. For $p = \infty$ or $p|N$, $p \nmid M$, the integral Z_p is calculated in [Pitale and Schmidt 2009c, Theorems 3.5.1 and 4.4.1]. For $p|M$, the integral Z_p is calculated in Theorem 4.3. Putting all of this together we get the following global theorem.

Theorem 5.2. *Let F be a Siegel cuspidal newform of weight l with respect to $B(M)$, where l is any positive integer and M is square-free, satisfying the assumption stated in Section 5.1. Let Φ be the adelic function corresponding to F, and let π_F be an irreducible component of the cuspidal automorphic representation generated by Φ. Let τ be any irreducible cuspidal automorphic representation of* $\mathrm{GL}_2(\mathbb{A})$*. Let the global characters χ, χ_0 and Λ, as well as the global section $f_\Lambda \in I(s, \chi, \chi_0, \tau)$, be chosen as above. Then the global integral* (5-4) *is given by*

$$Z(s, f_\Lambda, \bar{\Phi}) = \Big(\prod_{p\le\infty} Y_p(s)\Big) \frac{L(3s+1/2, \pi\times\tilde{\tau})}{L(6s+1, \omega_\tau^{-1})\, L(3s+1, \tilde{\tau}\times \mathcal{AI}(\Lambda))} \tag{5-6}$$

with

$$Y_\infty(s) = \overline{a(\Lambda)}\, i^{l+l_2} \frac{a^+}{2}\pi D^{-3s-l/2} \cdot \frac{(4\pi)^{-3s+3/2-l}}{6s+2l+l_2-1}\, \frac{\Gamma\big(3s+l-1+(ir)/2\big)\,\Gamma\big(3s+l-1-(ir)/2\big)}{\Gamma(3s+l-l_1/2-1/2)}. \tag{5-7}$$

Here, $\mathcal{AI}(\Lambda)$ is the automorphic representation of $\mathrm{GL}_2(\mathbb{A})$ *obtained from Λ via automorphic induction. The factor $Y_p(s)$ is one for $p \nmid MN$. For $p \nmid M$, $p|N$, the factor $Y_p(s)$ is given in* [Pitale and Schmidt 2009c, Theorem 3.5.1]. *For $p|M$, we have $Y_p(s) = L_p(6s+1, \omega_{\tau_p}^{-1})\, L\big(3s+1, \tilde{\tau}_p \times \mathcal{AI}(\Lambda_p)\big)\, Y'_p(s)$, where $Y'_p(s)$ is given in Theorem 4.3. The number r and a^+ are as in the archimedean calculation in* [Pitale and Schmidt 2009c], *and the constant $a(\Lambda)$ is defined in Section 5.1.*

5.3. ***Special values of L-functions.*** In this section, we will use Theorem 5.2 to obtain a special value result for the L-function in the case that τ corresponds to a holomorphic cusp form of the same weight as F. Let $\Psi \in S_l(N, \chi')$, the space of holomorphic cusp forms on the complex upper half plane $\mathfrak{h}_1$ of weight l with respect to $\Gamma_0(N)$ and nebentypus χ'. Here $N = \prod_p p^{n_p}$ is any positive integer and χ' is a Dirichlet character modulo N. We have as a Fourier expansion

$$\Psi(z) = \sum_{n=1}^{\infty} b_n e^{2\pi i n z}.$$

We will assume that Ψ is primitive, which means that Ψ is a newform, a Hecke eigenform and is normalized so that $b_1 = 1$. Let $\omega = \bigotimes \omega_p$ be the character of $\mathbb{A}^\times/\mathbb{Q}^\times$ corresponding to χ'. Let $K^{(0)}(N) := \prod_{p|N} K^{(0)}(\mathfrak{p}^{n_p}) \prod_{p\nmid N} \mathrm{GL}_2(\mathbb{Z}_p)$ with the local congruence subgroups

$$K^{(0)}(\mathfrak{p}^n) = \mathrm{GL}_2(\mathbb{Z}_p) \cap \begin{bmatrix} 1+p^n\mathbb{Z}_p & \mathbb{Z}_p \\ p^n\mathbb{Z}_p & \mathbb{Z}_p \end{bmatrix}$$

as in (4-5). Let $K_0(N) := \prod_{p|N} K_0(\mathfrak{p}^{n_p}) \prod_{p\nmid N} \mathrm{GL}_2(\mathbb{Z}_p)$, where

$$K_0(\mathfrak{p}^n) = \mathrm{GL}_2(\mathbb{Z}_p) \cap \begin{bmatrix} \mathbb{Z}_p & \mathbb{Z}_p \\ p^n\mathbb{Z}_p & \mathbb{Z}_p \end{bmatrix}.$$

Evidently, $K^{(0)}(N) \subset K_0(N)$. Let λ be the character of $K_0(N)$ given by

$$\lambda\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) := \prod_{p|N} \omega_p(a_p).$$

With these notations, we now define the adelic function f_Ψ by

$$f_\Psi(\gamma_0 m k) = \lambda(k)\,\frac{\det(m)^{l/2}}{(\gamma i+\delta)^l}\,\Psi\Big(\frac{\alpha i+\beta}{\gamma i+\delta}\Big),$$

where $\gamma_0 \in \mathrm{GL}_2(\mathbb{Q})$,

$$m = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \in \mathrm{GL}_2^+(\mathbb{R})$$

and $k \in K_0(N)$. Define a character χ_0, as in the previous section, with $l_2 = -l$. Using χ_0, extend f_Ψ to a function on $\mathrm{GU}(1,1;L)(\mathbb{A})$. We can take $\hat{f} = f_\Psi$ in (5-3) and obtain the section f_Λ. Now, [Pitale and Schmidt 2009c, Lemma 5.4.2] gives us that, for $g \in G^+(\mathbb{R})$, the function

$$\mu_2(g)^{-l}\det(J(g, i1_2))^l\,E(g,s;\,f_\Lambda)$$

only depends on $Z = g\langle i1_2\rangle$. We define the function $\mathscr{E}$ on

$$\mathbb{H}_2 := \{Z \in M_2(\mathbb{C}) : i({}^t\bar{Z} - Z) \text{ is positive definite}\}$$

by the formula

$$\mathscr{E}(Z,s) = \mu_2(g)^{-l}\det\big(J(g,i1_2)\big)^l\,E\Big(g, \frac{s}{3}+\frac{l}{6}-\frac{1}{2};\, f_\Lambda\Big),$$

where $g \in G^+(\mathbb{R})$ is such that $g\langle i1_2\rangle = Z$. The series that defines $\mathscr{E}(Z,s)$ is absolutely convergent for $\mathrm{Re}(s) > 3 - l/2$ (see [Klingen 1967]). We assume that $l > 6$. Now, we can set $s = 0$ and obtain a holomorphic Eisenstein series $\mathscr{E}(Z,0)$ on $\mathbb{H}_2$. Let

$$\Gamma^G(M,N) := G(\mathbb{Q}) \cap G^+(\mathbb{R})K^G(M,N).$$

We have

$$\Gamma^G(M,N)\cap H(\mathbb{Q})=B(M).$$

Then $\mathscr{E}(Z,0)$ is a modular form of weight l with respect to $\Gamma^G(M,N)$. Its restriction to $\mathfrak{h}_2$, the Siegel upper half space, is a modular form of weight l with respect to $B(M)$. By [Harris 1984], we know that the Fourier coefficients of $\mathscr{E}(Z,0)$ are algebraic.

Set

$$V(M):=\bigl[\mathrm{Sp}_4(\mathbb{Z}):B(M)\bigr]^{-1}$$

and define, for any two Siegel modular forms F_1, F_2 of weight l with respect to $B(M)$, the Petersson inner product by

$$\langle F_1,F_2\rangle=\tfrac{1}{2}V(M)\int_{B(M)\backslash\mathfrak{h}_2}F(Z)\,\overline{F_2(Z)}\,(\det(Y))^{l-3}dX\,dY.$$

Arguing as in [Pitale and Schmidt 2009c, Lemma 5.6.2] or [Saha 2009, Proposition 9.0.5], we get

$$Z\bigl(\tfrac{1}{6}l-\tfrac{1}{2},f_\Lambda,\bar{\Phi}\bigr)=\bigl\langle\mathscr{E}(Z,0),F\bigr\rangle. \tag{5-8}$$

Let

$$\Gamma^{(2)}(M):=\{g\in\mathrm{Sp}_4(\mathbb{Z}):g\equiv 1\pmod M\}$$

be the principal congruence subgroup of $\mathrm{Sp}_4(\mathbb{Z})$. We denote the space of all Siegel cusp forms of weight l with respect to $\Gamma^{(2)}(M)$ by $S_l(\Gamma^{(2)}(M))$. For a Hecke eigenform $F\in S_l(\Gamma^{(2)}(M))$, let $\mathbb{Q}(F)$ be the subfield of $\mathbb{C}$ generated by all the Hecke eigenvalues of F. From [Garrett 1992, p. 460], we see that $\mathbb{Q}(F)$ is a totally real number field. Let $S_l(\Gamma^{(2)}(M),\mathbb{Q}(F))$ be the subspace of $S_l(\Gamma^{(2)}(M))$ consisting of cusp forms whose Fourier coefficients lie in $\mathbb{Q}(F)$. Again by [Garrett 1992, p. 460], $S_l(\Gamma^{(2)}(M))$ has an orthogonal basis $\{F_i\}$ of Hecke eigenforms $F_i\in S_l(\Gamma^{(2)}(M),\mathbb{Q}(F_i))$. In addition, if F is a Hecke eigenform such that $F\in S_l(\Gamma^{(2)}(M),\mathbb{Q}(F))$, then one can take $F_1=F$ in the above basis. Hence, we assume that the Siegel newform F of weight l with respect to $B(M)$ considered in the previous section satisfies $F\in S_l(\Gamma^{(2)}(M),\mathbb{Q}(F))$. Then, arguing as in [Pitale and Schmidt 2009b, Lemma 5.4.3], we have

$$\frac{\bigl\langle\mathscr{E}(Z,0),F\bigr\rangle}{\langle F,F\rangle}\in\bar{\mathbb{Q}}, \tag{5-9}$$

where $\bar{\mathbb{Q}}$ is the algebraic closure of $\mathbb{Q}$ in $\mathbb{C}$. Let

$$\langle\Psi,\Psi\rangle_1:=\bigl(\mathrm{SL}_2(\mathbb{Z}):\Gamma_1(N)\bigr)^{-1}\int_{\Gamma_1(N)\backslash\mathfrak{h}_1}|\Psi(z)|^2y^{l-2}dx\,dy,$$

where

$$\Gamma_1(N) := \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_0(N) : a, d \equiv 1 \pmod{N} \right\}.$$

We have the following generalization of [Furusawa 1993, Theorem 4.8.3].

Theorem 5.3. *Let l, M be positive integers such that $l > 6$ and M is square-free. Let F be a cuspidal Siegel newform of weight l with respect to $B(M)$ such that $F \in S_l(\Gamma^{(2)}(M), \mathbb{Q}(F))$, satisfying the assumption from Sect. 5.1. Let $\Psi \in S_l(N, \chi')$ be a primitive form, with $N = \prod p^{n_p}$, any positive integer, and χ', any Dirichlet character modulo N. Let π_F and τ_Ψ be the irreducible cuspidal automorphic representations of* $\mathrm{GSp}_4(\mathbb{A})$ *and* $\mathrm{GL}_2(\mathbb{A})$ *corresponding to F and Ψ. Then*

$$\frac{L\left(\frac{l}{2} - 1,\ \pi_F \times \tilde{\tau}_\Psi\right)}{\pi^{5l-8}\langle F, F\rangle \langle \Psi, \Psi\rangle_1} \in \bar{\mathbb{Q}}. \tag{5-10}$$

Proof. Arguing as in the proof of [Pitale and Schmidt 2009c, Theorem 5.7.1], together with (5-8) and (5-9), we get the theorem. □

Special value results like the one above have been obtained in [Böcherer and Heim 2006; Furusawa 1993; Pitale and Schmidt 2009b; 2009c; Saha 2009].

Acknowledgments

We thank Ralf Schmidt for all his help, and in particular for explaining how to obtain a Bessel model from a Whittaker model in the split case. We also thank Abhishek Saha for several fruitful discussions.

References

[Böcherer and Heim 2006] S. Böcherer and B. E. Heim, "Critical values of L-functions on $\mathrm{GSp}_2 \times \mathrm{GL}_2$", *Math. Z.* **254**:3 (2006), 485–503. MR 2007f:11053 Zbl 1197.11059

[Cartier 1979] P. Cartier, "Representations of p-adic groups: a survey", pp. 111–155 in *Automorphic forms, representations and L-functions, I* (Corvallis, OR, 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **33**, Amer. Math. Soc., Providence, R.I., 1979. MR 81e:22029 Zbl 0421.22010

[Deligne 1979] P. Deligne, "Valeurs de fonctions L et périodes d'intégrales", pp. 313–346 in *Automorphic forms, representations and L-functions, Part 2* (Oregon State Univ., Corvallis, Ore., 1977), edited by A. Borel and W. Casselman, Proc. Sympos. Pure Math. **XXXIII**, Amer. Math. Soc., Providence, R.I., 1979. MR 81d:12009 Zbl 0449.10022

[Furusawa 1993] M. Furusawa, "On L-functions for GSp(4) × GL(2) and their special values", *J. Reine Angew. Math.* **438** (1993), 187–218. MR 94e:11057 Zbl 0770.11025

[Garrett 1992] P. B. Garrett, "On the arithmetic of Siegel–Hilbert cuspforms: Petersson inner products and Fourier coefficients", *Invent. Math.* **107**:3 (1992), 453–481. MR 93e:11060

[Harris 1984] M. Harris, "Eisenstein series on Shimura varieties", *Ann. of Math.* (2) **119**:1 (1984), 59–94. MR 85j:11052 Zbl 0589.10030

[Klingen 1967] H. Klingen, “Zum Darstellungssatz für Siegelsche Modulformen”, *Math. Z.* **102** (1967), 30–43. MR 36 #2555

[Novo-dvorsky and Piatetski-Shapiro 1973] M. E. Novo-dvorsky and I. I. Piatetski-Shapiro, “Generalized Bessel models for the symplectic group of rank 2”, *Mat. Sb.* (*N.S.*) **90 (132)** (1973), 246–256, 326. In Russian; translated in *Math. USSR-Sb.* **19** (1972), 243–255. MR 49 #3045

[Pitale and Schmidt 2009a] A. Pitale and R. Schmidt, “Bessel models for lowest weight representations of $\mathrm{GSp}(4, \mathbb{R})$”, *Int. Math. Res. Notices* **2009**:7 (2009), 1159–1212. MR 2010h:11084 Zbl 05553685

[Pitale and Schmidt 2009b] A. Pitale and R. Schmidt, “Integral representation for L-functions for $\mathrm{GSp}_4 \times \mathrm{GL}_2$”, *J. Number Theory* **129**:6 (2009), 1272–1324. MR 2010f:11079 Zbl 05559081 arXiv 0807.3522

[Pitale and Schmidt 2009c] A. Pitale and R. Schmidt, “Integral representation for L-functions for $\mathrm{GSp}_4 \times \mathrm{GL}_2$, II”, preprint, 2009. Zbl 05559081 arXiv 0908.1611

[Prasad and Takloo-Bighash 2007] D. Prasad and R. Takloo-Bighash, “Bessel models for GSp(4)”, preprint, 2007. arXiv 0712.2092

[Roberts and Schmidt 2007] B. Roberts and R. Schmidt, *Local newforms for GSp(4)*, Lecture Notes in Mathematics **1918**, Springer, Berlin, 2007. MR 2008g:11080 Zbl 1126.11027

[Saha 2009] A. Saha, “L-functions for holomorphic forms on $\mathrm{GSp}(4) \times \mathrm{GL}(2)$ and their special values”, *Int. Math. Res. Not.* **2009**:10 (2009), 1773–1837. MR 2010g:11088 Zbl 05565833

[Schmidt 2002] R. Schmidt, “Some remarks on local newforms for GL(2)”, *J. Ramanujan Math. Soc.* **17**:2 (2002), 115–147. MR 2003g:11056 Zbl 0997.11040

[Schmidt 2005] R. Schmidt, “Iwahori-spherical representations of GSp(4) and Siegel modular forms of degree 2 with square-free level”, *J. Math. Soc. Japan* **57**:1 (2005), 259–293. MR 2005i:11065 Zbl 1166.11325

[Sugano 1985] T. Sugano, “On holomorphic cusp forms on quaternion unitary groups of degree 2”, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **31**:3 (1985), 521–568. MR 86f:11040 Zbl 0559.10020

[Takloo-Bighash 2000] R. Takloo-Bighash, “L-functions for the p-adic group GSp(4)”, *Amer. J. Math.* **122**:6 (2000), 1085–1120. MR 2001k:11090 Zbl 0978.22017

Received September 30, 2009.

AMEYA PITALE
AMERICAN INSTITUTE OF MATHEMATICS
360 PORTAGE AVENUE
PALO ALTO, CA 94306-2244
UNITED STATES

pitale@aimath.org

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

AN INTEGRAL EXPRESSION OF THE FIRST NONTRIVIAL ONE-COCYCLE OF THE SPACE OF LONG KNOTS IN $\mathbb{R}^3$

KEIICHI SAKAI

Our main object of study is a certain degree-one cohomology class of the space $\mathcal{K}_3$ of long knots in $\mathbb{R}^3$. We describe this class in terms of graphs and configuration space integrals, showing the vanishing of some anomalous obstructions. To show that this class is not zero, we integrate it over a cycle studied by Gramain. As a corollary, we establish a relation between this class and ($\mathbb{R}$-valued) Casson's knot invariant. These are $\mathbb{R}$-versions of the results which were previously proved by Teiblyum, Turchin and Vassiliev over $\mathbb{Z}/2$ in a different way from ours.

1. Introduction

A *long knot* in $\mathbb{R}^n$ is an embedding $f : \mathbb{R}^1 \hookrightarrow \mathbb{R}^n$ that agrees with the standard inclusion $\iota(t) = (t, 0, \dots, 0)$ outside $[-1, 1]$. We denote by $\mathcal{K}_n$ the space of long knots in $\mathbb{R}^n$ equipped with C^∞-topology.

In [Cattaneo et al. 2002] a cochain map $I : \mathcal{D}^* \to \Omega^*_{DR}(\mathcal{K}_n)$ from a certain *graph complex* $\mathcal{D}^*$ was constructed for $n > 3$. The cocycles of $\mathcal{K}_n$ corresponding to *trivalent graph cocycles* via I generalize an integral expression of finite type invariants for (long) knots in $\mathbb{R}^3$ [Altschuler and Freidel 1997; Bott and Taubes 1994; Kohno 1994; Volić 2007]. In [Sakai 2008] the author found a *nontrivalent* graph cocycle $\Gamma \in \mathcal{D}^*$ and proved that, when $n > 3$ is odd, it gives a nonzero cohomology class $[I(\Gamma)] \in H^{3n-8}_{DR}(\mathcal{K}_n)$. On the other hand, when $n = 3$, some obstructions to I being a cochain map (called *anomalous obstructions*; see for example [Volić 2007, Section 4.6]) may survive, so even the closedness of $I(\Gamma)$ was not clear. However, the obstructions for trivalent graph cocycles X (of "even orders") in fact vanish [Altschuler and Freidel 1997], hence the map I still yields closed zero-forms $I(X)$ of $\mathcal{K}_3$ (they are finite type invariants). This raises our hope

The author is partially supported by Grant-in-Aid for Young Scientists (B) 21740038, the Sumitomo Foundation, the Iwanami Fujukai Foundation, and a Research Fellowship for Young Scientists (228006) from the Japan Society for the Promotion of Science Research.
MSC2000: primary 58D10; secondary 55P48, 57M25, 57M27, 81Q30.
Keywords: the space of long knots, configuration space integrals, nontrivalent graphs, an action of little cubes, Gramain cycles, Casson's knot invariant.

that all obstructions for any graphs may vanish and hence the map I could be a cochain map even when $n = 3$.

In this paper we will show (in Theorem 2.4) that the obstructions for the non-trivalent graph cocycle Γ mentioned above also vanish, hence the map I yields the first example of a closed one-form $I(\Gamma)$ of $\mathcal{K}_3$. To show that $[I(\Gamma)] \in H^1_{DR}(\mathcal{K}_3)$ is not zero, we will study in part how $I(\Gamma)$ fits into a description of the homotopy type of $\mathcal{K}_3$ given in [Budney 2010; 2007; Budney and Cohen 2009]. It is known that on each component $\mathcal{K}_3(f)$ that contains $f \in \mathcal{K}_3$, there exists a one-cycle G_f called the *Gramain cycle* [Gramain 1977; Budney 2010; Turchin 2006; Vassiliev 2001]. The Kronecker pairing gives an isotopy invariant $V : f \mapsto \langle I(\Gamma),\ G_f\rangle$. We show in Theorem 3.1 that V coincides with *Casson's knot invariant* v_2, which is characterized as the coefficient of z^2 in the Alexander–Conway polynomial. This result will be generalized in Theorem 3.6 for one-cycles obtained by using an action of *little two-cubes operad* on the space $\tilde{\mathcal{K}}_3$ of *framed* long knots [Budney 2007].

Closely related results have appeared in [Turchin 2006; Vassiliev 2001], where the $\mathbb{Z}/2$-reduction of a cocycle v_3^1 of $\mathcal{K}_n$ ($n \geq 3$), appearing in the E_1-term of Vassiliev's spectral sequence [Vassiliev 1992], was studied. A natural quasi-isomorphism $\mathcal{D}^* \to E_0 \otimes \mathbb{R}$ maps our cocycle Γ to v_3^1. In this sense, our results can be seen as "lifts" of those in [Turchin 2006; Vassiliev 2001] to $\mathbb{R}$.

The invariant v_2 can also be interpreted as the linking number of colinearity manifolds [Budney et al. 2005]. Notice that in each formulation (including the one in this paper) the value of v_2 is computed by counting some colinearity pairs on the knot.

2. Construction of a close differential form

Configuration space integral. We review briefly how we can construct (closed) forms of $\mathcal{K}_n$ from graphs. For full details see [Cattaneo et al. 2002; Volić 2007].

Let X be a *graph* in the sense of those references (see Figure 1 for examples). Let v_{i} and v_{f} be the numbers of the *interval vertices* (or *i-vertices* for short; those on the specified oriented line) and the *free* vertices (or *f-vertices*; those which are not interval vertices) of X, respectively. With X we associate a configuration space

$$C_X := \left\{ \begin{array}{l} (f; x_1, \dots, x_{v_{\mathrm{i}}}; x_{v_{\mathrm{i}}+1}, \dots, x_{v_{\mathrm{i}}+v_{\mathrm{f}}}) \\ \in \mathcal{K}_n \times \mathrm{Conf}\,(\mathbb{R}^1, v_{\mathrm{i}}) \times \mathrm{Conf}\,(\mathbb{R}^n, v_{\mathrm{f}}) \end{array} \;\middle|\; \begin{array}{l} f(x_i) \neq x_j \text{ for any} \\ 1 \leq i \leq v_{\mathrm{i}} < j \leq v_{\mathrm{i}} + v_{\mathrm{f}} \end{array} \right\},$$

where $\mathrm{Conf}\,(M, k) := M^{\times k} \setminus \bigcup_{1\leq i<j\leq k}\{x_i = x_j\}$ for a space M.

Let e be the number of the edges of X. Define $\omega_X \in \Omega^{(n-1)e}_{DR}(C_X)$ as the wedge of closed $(n-1)$-forms $\varphi^*_\alpha \mathrm{vol}_{S^{n-1}}$, where $\varphi_\alpha : C_X \to S^{n-1}$ is the *Gauss map*, which assigns a unit vector determined by two points in $\mathbb{R}^n$ corresponding to the vertices adjacent to an edge α of X (for an i-vertex corresponding to $x_i \in \mathbb{R}^1$, we consider the point $f(x_i) \in \mathbb{R}^n$). Here we assume that $\mathrm{vol}_{S^{n-1}}$ is "(anti)symmetric", namely $i^*\mathrm{vol}_{S^{n-1}} = (-1)^n \mathrm{vol}_{S^{n-1}}$ for the antipodal map $i : S^{n-1} \to S^{n-1}$. Then $I(X) \in \Omega_{DR}^{(n-1)e-v_{\mathrm{i}}-nv_{\mathrm{f}}}(\mathcal{K}_n)$ is defined by

$$I(X) := (\pi_X)_* \omega_X,$$

the integration along the fiber of the natural fibration $\pi_X : C_X \to \mathcal{K}_n$. This fiber is a subspace of $\mathrm{Conf}(\mathbb{R}^1, v_{\mathrm{i}}) \times \mathrm{Conf}(\mathbb{R}^n, v_{\mathrm{f}})$. Such integrals converge, since the fiber can be compactified in such a way that the forms $\varphi_\alpha^* \mathrm{vol}_{S^{n-1}}$ are still well-defined on the compactification [Bott and Taubes 1994, Proposition 1.1]. We extend I linearly onto $\mathcal{D}^*$, a cochain complex spanned by graphs. The differential δ of $\mathcal{D}^*$ is defined as a signed sum of graphs obtained by "contracting" the edges one at a time.

One of the results of [Cattaneo et al. 2002] states that $I : \mathcal{D}^* \to \Omega_{DR}^*(\mathcal{K}_n)$ is a cochain map if $n > 3$. The proof is outlined as follows. By the generalized Stokes theorem, $dI(X) = \pm(\pi_X^\partial)_* \omega_X$, where π_X^∂ is the restriction of π_X to the codimension one strata of the boundary of the (compactified) fiber of π_X. Each codimension one stratum corresponds to a collision of subconfigurations in C_X, or equivalently to $A \subset V(X) \cup \{\infty\}$ (here $V(X)$ is the set of vertices of X) with a consecutiveness property: if two i-vertices p, q are in A, then all the other i-vertices between p and q are in A. Here "$\infty \in A$" means that the points x_l ($l \in A$) escape to infinity. When $\infty \notin A$, the interior $\mathrm{Int}\,\Sigma_A$ of the corresponding stratum Σ_A to A is described by the pullback square

(2-1)
$$\begin{tikzcd} & \mathrm{Int}\,\Sigma_A \arrow[r] \arrow[dl, "\pi_X^{\partial_A}"'] \arrow[d] & \hat{B}_A \arrow[d, "\rho_A"] \\ \mathcal{K}_n & C_{X/X_A} \arrow[l, "\pi_{X/X_A}"] \arrow[r, "D_A"'] & B_A \end{tikzcd}$$

Here

- X_A is the maximal subgraph of X with $V(X_A) = A$, and X/X_A is a graph obtained by collapsing the subgraph X_A to a single vertex v_A;
- $B_A = S^{n-1}$ if A contains at least one i-vertex, and $B_A = \{*\}$ otherwise;
- if A consists of i-vertices $i_1, \dots, i_s$ ($s > 0$) and f-vertices $i_{s+1}, \dots, i_{s+t}$, then

$$\hat{B}_A := \left\{ \begin{array}{l} (v; (x_{i_1}, \dots, x_{i_s}; x_{i_{s+1}}, \dots, x_{i_{s+t}})) \\ \in S^{n-1} \times \mathrm{Conf}(\mathbb{R}^1, s) \times \mathrm{Conf}(\mathbb{R}^n, t) \end{array} \middle| \begin{array}{l} x_{i_p} v \neq x_{i_q} \text{ for any} \\ 1 \le p \le s < q \le s+t \end{array} \right\} \Big/ \sim,$$

where $\sim$ is defined by

$$\begin{aligned} &\big(v; (x_{i_1}, \dots, x_{i_s}; x_{i_{s+1}}, \dots, x_{i_{s+t}})\big) \sim \\ &\qquad \big(v; (a(x_{i_1} + r), \dots, a(x_{i_s} + r); a(x_{i_{s+1}} + rv), \dots, a(x_{i_{s+t}} + rv))\big), \end{aligned}$$

for any $a \in \mathbb{R}_{>0}$ and $r \in \mathbb{R}$ (if A consists only of t f-vertices, then

$$\hat{B}_A := \mathrm{Conf}\,(\mathbb{R}^n, t)/(\mathbb{R}^1_{>0} \rtimes \mathbb{R}^n),$$

where $\mathbb{R}^1_{>0} \rtimes \mathbb{R}^n$ acts on $\mathrm{Conf}\,(\mathbb{R}^n, t)$ by scaling and translation);

- ρ_A is the natural projection;
- when A contains at least one i-vertex, $D_A : C_{X/X_A} \to S^{n-1}$ maps $(f; (x_i))$ to $f'(x_{v_A})/|f'(x_{v_A})|$.

We omit the case $\infty \in A$; see [Cattaneo et al. 2002, Appendix].

By properties of fiber integrations and pullbacks, the integration of ω_X along $\mathrm{Int}\,\Sigma_A$ can be written as $(\pi_{X/X_A})_*(\omega_{X/X_A} \wedge D_A^*(\rho_A)_*\hat{\omega}_{X_A})$, where $\hat{\omega}_{X_A} \in \Omega^*_{DR}(\hat{B}_A)$ is defined similarly to $\omega_X \in \Omega^*_{DR}(C_X)$.

The stratum Σ_A is called *principal* if $|A| = 2$, *hidden* if $|A| \geq 3$, and *infinity* if $\infty \in A$. Since two-point collisions correspond to contractions of edges, we have $dI(X) = I(\delta X)$ modulo the integrations along hidden and infinity faces. When $n > 3$, the hidden/infinity contributions turn out to be zero; in fact $(\rho_A)_*\hat{\omega}_{X_A} = 0$ if $n > 3$ and if A is not principal; see [Cattaneo et al. 2002, Appendix] or the next example. This proves that the map I is a cochain map if $n > 3$.

Example 2.1. Here we show one example of vanishing of an integration along a hidden face Σ_A. Let X be the seventh graph in Figure 1 and $A := \{1, 4, 5\}$. Then in (2-1), $B_A = S^{n-1}$ since A contains an i-vertex 1, and

$$\hat{B}_A = \{(v; x_1; x_4, x_5) \in S^{n-1} \times \mathbb{R}^1 \times \mathrm{Conf}\,(\mathbb{R}^n, 2) \mid x_1 v \neq x_4, x_5\}/\sim,$$

where $(v; x_1; x_4, x_5) \sim (v; a(x_1 + r); a(x_4 + rv), a(x_5 + rv))$ for any $a > 0$ and $r \in \mathbb{R}^1$. The subgraph X_A consists of three vertices $1, 4, 5$ and three edges 14, 15 and 45. The open face $\mathrm{Int}\,\Sigma_A$, where three points $f(x_1)$, x_4 and x_5 collide with each other, is a hidden face and is described by the square (2-1). Then the integration of ω_X along $\mathrm{Int}\,\Sigma_A$ is $(\pi_{X/X_A})_*(\omega_{X/X_A} \wedge D_A^*(\rho_A)_*\hat{\omega}_{X_A})$, where

$$\hat{\omega}_{X_A} = \varphi_{14}^*\mathrm{vol}_{S^{n-1}} \wedge \varphi_{15}^*\mathrm{vol}_{S^{n-1}} \wedge \varphi_{45}^*\mathrm{vol}_{S^{n-1}} \in \Omega^{3(n-1)}_{DR}(\hat{B}_A),$$

$$\varphi_{1j} := \frac{x_j - x_1 v}{|x_j - x_1 v|} \quad (j = 4, 5), \qquad \varphi_{45} := \frac{x_5 - x_4}{|x_5 - x_4|}.$$

In this case we can prove that $(\rho_A)_*\hat{\omega}_{X_A} = 0$, hence the integration of ω_X along $\mathrm{Int}\,\Sigma_A$ vanishes. Indeed a fiberwise involution $\chi : \hat{B}_A \to \hat{B}_A$ defined by

$$\chi(v; x_1; x_4, x_5) := (v; x_1; 2x_1 v - x_4, 2x_1 v - x_5)$$

preserves the orientation of the fiber but $\chi^*\hat{\omega}_{X_A} = -\hat{\omega}_{X_A}$ (here we use that $\mathrm{vol}_{S^{n-1}}$ is antisymmetric), hence we have $(\rho_A)_*\hat{\omega}_{X_A} = -(\rho_A)_*\hat{\omega}_{X_A}$.

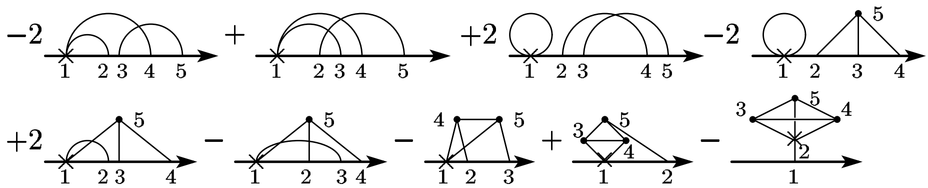


Figure 1. A graph cocycle Γ.

Nontrivalent cocycle. It is shown in [Cattaneo et al. 2002] that, when $n > 3$, the induced map I on cohomology restricted to the space of trivalent graph cocycles is injective. In [Sakai 2008], the author gave the first example of a nontrivalent graph cocycle Γ (Figure 1) which also gives a nonzero class $[I(\Gamma)] \in H^{3n-8}_{DR}(\mathcal{K}_n)$ when $n > 3$ is odd.

In Figure 1, nontrivalent vertices and trivalent f-vertices are marked by $\times$ and $\bullet$, respectively, and other crossings are not vertices. Here we say an i-vertex v is trivalent if there is exactly one edge emanating from v other than the specified oriented line. Each edge ij $(i < j)$ is oriented so that i is the initial vertex.

Remark 2.2. An analogous nontrivalent graph cocycle for the space of embeddings $S^1 \hookrightarrow \mathbb{R}^n$ for even $n \geq 4$ can be found in [Longoni 2004].

If $n = 3$, integrations along some hidden faces (called *anomalous contributions*) might survive, so the map I might fail to be a cochain map. However, nonzero anomalous contributions arise from limited hidden faces.

Theorem 2.3. *Let X be a graph and $A \subset V(X) \cup \{\infty\}$ be such that Σ_A is not principal. When $n = 3$, the integration of ω_X along Σ_A can be nonzero only if the subgraph X_A is trivalent.*

Our main theorem is proved by using Theorem 2.3.

Theorem 2.4. *$I(\Gamma) \in \Omega^1_{DR}(\mathcal{K}_3)$ is a closed form.*

Proof. We call the nine graphs in Figure 1 $\Gamma_1, \dots, \Gamma_9$, respectively. The graphs Γ_i, $i \neq 3, 4, 9$, do not contain trivalent subgraphs X_A satisfying the *consecutive property*; see the paragraph just before (2-1). So $dI(\Gamma_i) = I(d\Gamma_i)$ for $i \neq 3, 4, 9$ by Theorem 2.3.

Possibly the integration of ω_{Γ_i} $(i = 3, 4, 9)$ along Σ_A $(A := \{2, \dots, 5\})$ might survive, since the corresponding subgraph X_A is trivalent. However, we can prove $(\rho_A)_* \hat{\omega}_{X_A} = 0$ (and hence $dI(\Gamma_i) = I(d\Gamma_i)$) as follows: $(\rho_A)_* \hat{\omega}_{X_A} = 0$ for Γ_3, because there is a fiberwise free action of $\mathbb{R}_{>0}$ on $\hat{B}_A$ given by translations of x_2 and x_4 [Volić 2007, Proposition 4.1] which preserves $\hat{\omega}_{X_A}$. Thus $(\rho_A)_* \hat{\omega}_{X_A} = 0$ by dimensional reason. The proof for Γ_4 has appeared in [Bott and Taubes 1994, page 5271]; $\hat{\omega}_{X_A} = 0$ on $\hat{B}_A$ since the image of the Gauss map $\varphi : B_A \to (S^2)^3$ corresponding to three edges of X_A is of positive codimension. As for Γ_9, $(\rho_A)_*\hat{\omega}_{X_A} = 0$ follows from $\deg(\rho_A)_*\hat{\omega}_{X_A} = 4$ which exceeds $\dim B_A$ (in fact $B_A = \{*\}$ in this case). □

Proof of Theorem 2.3. Let A be a subset of $V(X)$ with $|A| \geq 3$ or $\infty \in A$, and X_A is nontrivalent. We must show the vanishing of the integrations along the nonprincipal face Σ_A of the fiber of $C_X \to \mathcal{K}_3$. To do this it is enough to show $(\rho_A)_*\hat{\omega}_{X_A} = 0$. By dimensional arguments [Cattaneo et al. 2002, (A.2)] the contributions of infinite faces vanish. So below we consider the hidden faces Σ_A with $|A| \geq 3$.

If X_A has a vertex of valence ≤ 2, then $(\rho_A)_*\hat{\omega}_{X_A} = 0$ is proved by dimensional arguments or existence of a fiberwise symmetry of B_A which reverses the orientation of the fiber of $\rho_A : \hat{B}_A \to B_A$ but preserves the integrand $\hat{\omega}_{X_A}$ (like χ from Example 2.1, see also [Cattaneo et al. 2002, Lemmas A.7–A.9]).

Next, consider the case that there is a vertex of X_A of valence ≥ 4. Let e, s and t be the numbers of the edges, the i-vertices and the f-vertices of X_A, respectively. Then $\deg \hat{\omega}_{X_A} = 2e$ and the dimension of the fiber of ρ_A is $s + 3t - k$, where $k = 2$ or 4 according to whether $s > 0$ or $s = 0$ [Cattaneo et al. 2002, (A.1)]. Thus $(\rho_A)_*\hat{\omega}_{X_A} \in \Omega^*_{DR}(B_A)$ is of degree $2e - s - 3t + k$. It is not difficult to see $2e - s - 3t > 0$ because at least one vertex of X_A is of valence ≥ 4. Hence $\deg(\rho_A)_*\hat{\omega}_{X_A}$ exceeds $\dim B_A$ ($= 0$ or 2) and hence $(\rho_A)_*\hat{\omega}_{X_A} = 0$.

Thus only the integrations along Σ_A with X_A trivalent can survive. □

Remark 2.5. Every finite type invariant v for long knots in $\mathbb{R}^3$ can be written as a sum of $I(\Gamma_v)$ (Γ_v is a trivalent graph cocycle) and some "correction terms" which kill the contributions of hidden faces corresponding to trivalent subgraphs [Altschuler and Freidel 1997; Bott and Taubes 1994; Kohno 1994; Volić 2007]. So by Theorem 2.3 the problem whether $I : \mathcal{D}^* \to \Omega^*_{DR}(\mathcal{K}_3)$ is a cochain map or not is equivalent to the problem whether one can eliminate all the correction terms from integral expressions of finite type invariants.

3. Evaluation on some cycles

Here we will show that $[I(\Gamma)] \in H^1_{DR}(\mathcal{K}_3)$ restricted to some components of $\mathcal{K}_3$ is not zero.

We introduce two assumptions to simplify computations.

Assumption 1. The support of (antisymmetric) vol_{S^2} is contained in a sufficiently small neighborhood of the poles $(0, 0, \pm 1)$ as in [Sakai 2008]. So only the configurations with the images of the Gauss maps lying in a neighborhood of $(0, 0, \pm 1)$ can nontrivially contribute to various integrals below. Presumably $[I(\Gamma)] \in H^1_{DR}(\mathcal{K}_3)$ may be independent of choices of vol_{S^2} [Cattaneo et al. 2002, Proposition 4.5].

Assumption 2. Every long knot in $\mathbb{R}^3$ is contained in xy-plane except for over-arc of each crossing, and each over-arc is in $\{0 \le z \le h\}$ for a sufficiently small $h > 0$ so that the projection onto xy-plane is a regular diagram of the long knot.

The Gramain cycle. For any $f \in \mathcal{K}_3$, we denote by $\mathcal{K}_3(f)$ the component of $\mathcal{K}_3$ which contains f. Regarding $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ and fixing f, we define the map $G_f : S^1 \to \mathcal{K}_3(f)$, called the *Gramain cycle*, by $G_f(s)(t) := R(s)f(t)$, where $R(s) \in \mathrm{SO}(3)$ is the rotation by the angle s fixing the "long axis" (the x-axis). G_f generates an infinite cyclic subgroup of $\pi_1(\mathcal{K}_3(f))$ if f is nontrivial [Gramain 1977]. The homology class $[G_f] \in H_1(\mathcal{K}_3(f))$ is independent of the choice of f in the connected component; if $f_t \in \mathcal{K}_3$ $(0 \le t \le 1)$ is an isotopy connecting f_0 and f_1, then $G_{f_t} : [0, 1] \times S^1 \to \mathcal{K}_3$ gives a homotopy between G_{f_0} and G_{f_1}. Therefore the Kronecker pairing gives an isotopy invariant $V(f) := \langle I(\Gamma), G_f \rangle$ for long knots.

Theorem 3.1. *The invariant V is equal to Casson's knot invariant v_2.*

Corollary 3.2. $[I(\Gamma)|_{\mathcal{K}_3(f)}] \in H^1_{DR}(\mathcal{K}_3(f))$ *is not zero if* $v_2(f) \neq 0$. □

We will prove two statements that characterize Casson's knot invariant: V is of finite type of order two and $V(3_1) = 1$, where 3_1 is the long trefoil knot. To do this, we will represent G_f using a *Browder operation*, as in [Sakai 2008].

Little cubes action. Let $\tilde{\mathcal{K}}_n$ be the space of *framed* long knots in $\mathbb{R}^n$ (embeddings $\tilde{f} : \mathbb{R}^1 \times D^{n-1} \hookrightarrow \mathbb{R}^n$ that are standard outside $[-1, 1] \times D^{n-1}$). There is a homotopy equivalence $\Phi : \tilde{\mathcal{K}}_3 \simeq \mathcal{K}_3 \times \mathbb{Z}$ [Budney 2007] that maps $\tilde{f}$ to the pair $(\tilde{f}|_{\mathbb{R}^1\times\{(0,0)\}}, \mathrm{fr}\,\tilde{f})$, where the *framing number* $\mathrm{fr}\,\tilde{f}$ is defined as the linking number of $\tilde{f}|_{\mathbb{R}^1\times\{(0,0)\}}$ with $\tilde{f}|_{\mathbb{R}^1\times\{(1,0)\}}$. Since $\mathrm{fr}\,\tilde{f}$ is additive under the connected sum, Φ is a homotopy equivalence of H-spaces. In general, $\tilde{\mathcal{K}}_n \simeq \mathcal{K}_n \times \Omega\,\mathrm{SO}(n-1)$ as H-spaces, where Ω stands for the based loop space functor.

In [Budney 2007] an action of the *little two-cubes operad* on the space $\tilde{\mathcal{K}}_n$ was defined. Its second stage gives a map $S^1 \times (\tilde{\mathcal{K}}_n)^2 \to \tilde{\mathcal{K}}_n$ up to homotopy, which is given as "shrinking one knot f and sliding it along another knot g by using the framing, and repeating the same procedure with f and g exchanged" [Budney 2007, Figure 2]. Fixing a generator of $H_1(S^1)$, we obtain the *Browder operation* $\lambda : H_p(\tilde{\mathcal{K}}_n) \otimes H_q(\tilde{\mathcal{K}}_n) \to H_{p+q+1}(\tilde{\mathcal{K}}_n)$, which is a graded Lie bracket satisfying the Leibniz rule with respect to the product induced by the connected sum. The author proved in [Sakai 2008] that $\langle I(\Gamma), r_*\lambda(e, v)\rangle = 1$ when $n > 3$ is odd, where $r : \tilde{\mathcal{K}}_n \to \mathcal{K}_n$ is the forgetting map, $e \in H_{n-3}(\tilde{\mathcal{K}}_n)$ comes from the space of framings, and $v \in H_{2(n-3)}(\tilde{\mathcal{K}}_n)$ is the first nonzero class of $\mathcal{K}_n$ represented by a map $(S^{n-3})^{\times 2} \to \mathcal{K}_n$ (see below).

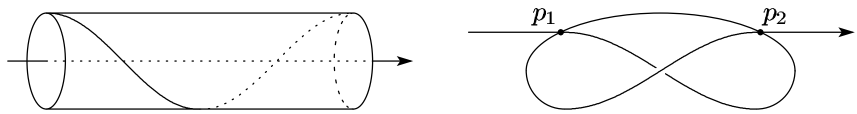


Figure 2. The cycles e and $v = v(T)$.

The case $n = 3$. In [Sakai 2008] the assumption $n > 3$ was used only to deduce the closedness of $I(\Gamma)$ from the results of Cattaneo et al. [2002]. The cycles e and v are defined even when $n = 3$:

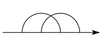
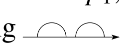
- Under the homotopy equivalence $\tilde{\mathcal{K}}_3 \simeq \mathcal{K}_3 \times \mathbb{Z}$, the zero-cycle e is given by $(\iota, 1)$ where ι is the trivial long knot ($\iota(t) = (t, 0, 0)$ for any $t \in \mathbb{R}^1$).
- The zero-cycle $v = v(T)$ is given by $\sum_{\varepsilon_i=\pm 1} \varepsilon_1\varepsilon_2 T_{\varepsilon_1,\varepsilon_2}$, where $T = 3_1$ and $T_{\varepsilon_1,\varepsilon_2}$ is T with its crossing p_i, for $i = 1, 2$ changed to be positive if $\varepsilon_i = +1$ and negative if $\varepsilon_i = -1$ (see Figure 2).

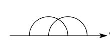
Notice that, for any $f \in \mathcal{K}_3$ and any pair (p_1, p_2) of its crossings, an analogous zero-cycle $v = v(f; p_1, p_2)$ can be defined.

Regard $f \in \mathcal{K}_3$ as a zero-cycle of $\tilde{\mathcal{K}}_3$ (with $\operatorname{fr} f = 0$) and consider $r_*\lambda(e, f)$. During a knot f "going through" e, f rotates once around the x-axis. Thus the one-cycle $r_*\lambda(e, f)$ is homologous to the Gramain cycle G_f. This leads us to the fact that, for $v = v(f; p_1, p_2)$, the one-cycle $r_*\lambda(e, v)$ is homologous to the sum $\sum_{\varepsilon_i=\pm 1} \varepsilon_1\varepsilon_2 G_{f_{\varepsilon_1,\varepsilon_2}}$. This is why we can apply the method in [Sakai 2008] to compute

$$D^2V(f) := \sum_{\varepsilon_j=\pm 1} \varepsilon_1\varepsilon_2 V(f_{\varepsilon_1,\varepsilon_2}) = \sum_{\varepsilon_j=\pm 1} \varepsilon_1\varepsilon_2 \langle I(\Gamma), G_{f_{\varepsilon_1,\varepsilon_2}}\rangle = \langle I(\Gamma), r_*\lambda(e, v(f))\rangle.$$

Recall that our graph cocycle Γ is a sum of nine graphs $\Gamma_1, \dots, \Gamma_9$ (see Figure 1). By Assumption 1, the integration $\langle I(\Gamma_i), G_f\rangle$ can be computed by "counting" the configurations with all the images of the Gauss maps corresponding to edges of Γ_i being around the poles of S^2. Lemma 3.4 below was proved in such a way in [Sakai 2008] when $n > 3$. Since $[v(f)] \in H_0(\mathcal{K}_3(f))$ is independent of small $h > 0$ (see Assumption 2), we may compute $D^2V(f)$ in the limit $h \to 0$.

Definition 3.3. We say that a pair (p_1, p_2) of crossings of f *respects the diagram*  if there exist $t_1 < t_2 < t_3 < t_4$ where $f(t_1)$ and $f(t_3)$ correspond to p_1, while $f(t_2)$ and $f(t_4)$ correspond to p_2. The notion of (p_1, p_2) respecting  or  is defined analogously.

Lemma 3.4 [Sakai 2008]. *Suppose that (p_1, p_2) respects* . *Then, in the limit $h \to 0$, $P_i(f) := \sum_{\varepsilon_j=\pm 1}\varepsilon_1\varepsilon_2\langle I(\Gamma_i), G_{f_{\varepsilon_1,\varepsilon_2}}\rangle$ converges to zero for $i \neq 2$, and $P_2(f)$ converges to* 1. *Thus $D^2V(f) = 1$.*

Outline of proof. Let $\hat{C}_{\Gamma_i} \to S^1$ be the pullback of $C_{\Gamma_i} \to \mathcal{K}_3$ via G_f, and let $\hat{G}_f : \hat{C}_{\Gamma_i} \to C_{\Gamma_i}$ be the lift of G_f. By the properties of pullbacks and fiber integrations,

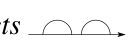
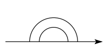
$$P_i(f) = \sum_{\varepsilon_i=\pm 1} \varepsilon_1 \varepsilon_2 \int_{\hat{C}_{\Gamma_i}} \hat{G}^*_{f_{\varepsilon_1,\varepsilon_2}} \omega_{\Gamma_i}. \tag{3-1}$$

Let $t_1 < \cdots < t_4$ be such that $f(t_1)$ and $f(t_3)$ correspond to p_1, while $f(t_2)$ and $f(t_4)$ correspond to p_2. Define the subspace $C'_{\Gamma_i} \subset \hat{C}_{\Gamma_i}$ as consisting of $(G_f(s); (x_j))$ $(s \in S^1)$ such that, for each $j = 1, 2$, there is a pair (l, m) of i-vertices of Γ_i such that x_l is on the over-arc of p_j, x_m is on the under-arc of p_j, and there is a sequence of edges in Γ_i from l to m.

First observation: The integration over $\hat{C}_{\Gamma_i} \setminus C'_{\Gamma_i}$ does not essentially contribute to $P_i(f)$ in the limit $h \to 0$. This is because, over $\hat{C}_{\Gamma_i} \setminus C'_{\Gamma_i}$, the integrals in (3-1) are well defined and continuous even when $h = 0$ (p_j becomes a double point), so two terms in $P_i(f)$ corresponding to $\varepsilon_j = \pm 1$ cancel each other. This implies $\lim_{h\to 0} P_i(f) = 0$ for $i = 7, 8, 9$, since $C'_{\Gamma_i} = \varnothing$ if $\sharp\{\text{i-vertices}\} \le 3$.

Second observation: Consider the configurations $(x_i) \in C'_{\Gamma_i}$ such that, for any pair (l, m) of i-vertices of Γ_i with x_l on the over-arc of p_j and x_m on the under-arc of p_j, all the points x_k (k is in a sequence in Γ_i from l to m) are not near p_j. Such configurations also do not essentially contribute to $P_i(f)$ in the limit $h \to 0$, by the same reason as above. This implies $\lim_{h\to 0} P_i(f) = 0$ for $i = 4, 5, 6$; the configurations $(x_l) \in C'_{\Gamma_i}$ $(4 \le i \le 6)$ must be such that the point $x_l \in \mathbb{R}^1$ $(1 \le l \le 4)$ is near t_l. By the second observation, the "free point" x_5 must be near p_1 or p_2. But then $\omega_{\Gamma_i} = 0$, since at least one Gauss map φ_{l5} has its image outside the support of vol_{S^2} (see Assumption 1). Thus $\lim_{h\to 0} P_i(f) = 0$.

Finally consider the $P_i(f)$, for $i = 1, 2, 3$. For $i = 1$ we have $\omega_{\Gamma_1} = 0$ over C'_{Γ_1}, since the Gauss map corresponding to the edge 12 has its image outside of the support of vol_{S^2}. The same reasoning, using the loop edge 11, shows that $\omega_{\Gamma_3} = 0$ over C'_{Γ_3}. Only $P_2(f)$ survives, since the configurations with x_1 near t_1, x_2 near t_2, x_3 and x_4 near t_3, and x_5 near t_4, contribute nontrivially to the integral [Sakai 2008, Lemma 4.6]. □

Lemma 3.5. *If (p_1, p_2) respects*  *or* *, then $D^2 V(f) = 0$.*

Proof. For $i = 4, \ldots, 9$, we see in the same way as in Lemma 3.4 that $P_i(f)$ approaches 0 as $h \to 0$. That $\lim_{h\to 0} P_i(f)$ for $i = 2, 3$ and the -case for $i = 1$ is proved by the first observation in the proof of Lemma 3.4.

In the -case for $P_1(f)$ over C'_{Γ_1} only the configurations with x_j near t_j, with $j = 1, 2, 3$, and x_5 near t_4 may essentially contribute to $P_1(f)$; in this case the edges 12 and 35 join the over/under arcs of p_1 and p_2 respectively. However, the Gauss map φ_{14} cannot have its image in the support of vol_{S^2}, so ω_{Γ_1} vanishes. □

Proof of Theorem 3.1. For three crossings (p_1, p_2, p_3) of $f \in \mathcal{K}_3$, consider the third difference

$$D^3V(f) := \sum_{\varepsilon_j=\pm 1} \varepsilon_1\varepsilon_2\varepsilon_3 V(f_{\varepsilon_1,\varepsilon_2,\varepsilon_3}) = D^2V(g_{+1}) - D^2V(g_{-1}),$$

where $g_{\pm 1} := f_{+1,+1,\pm 1}$ and $D^2V(g_{\pm 1})$ are taken with respect to (p_1, p_2). Since the pair (p_1, p_2) of g_{+1} respects the same diagram as (p_1, p_2) of g_{-1}, we have $D^2V(g_{+1}) = D^2V(g_{-1})$ by the above Lemmas 3.4, 3.5. Thus $D^3V = 0$ and hence V is finite type of order two. Moreover $V(\iota) = 0$ for the trivial long knot ι since $\mathcal{K}_3(\iota)$ is contractible [Hatcher 1983]; therefore $G_\iota \sim 0$, and $V(3_1) = 1$ by Lemma 3.4 and $V(\iota) = 0$. These properties uniquely characterize Casson's knot invariant v_2. □

The Browder operations. We denote a framed long knot corresponding to (f, k) under the equivalence $\tilde{\mathcal{K}}_3 \simeq \mathcal{K}_3 \times \mathbb{Z}$ by $f^k \in \tilde{\mathcal{K}}_3$ (unique up to homotopy). As mentioned above, the Gramain cycle can be written as $[G_f] = [r_*\lambda(f^k, \iota^1)]$ (k may be arbitrary). Below we will evaluate $I(\Gamma)$ on more general cycles $r_*\lambda(f^k, g^l)$ of $\mathcal{K}_3$ for any nontrivial $f, g \in \mathcal{K}_3$ and $k, l \in \mathbb{Z}$. This generalizes Theorem 3.1.

Theorem 3.6. *We have $\langle I(\Gamma),\ r_*\lambda(f^k, g^l)\rangle = lv_2(f) + kv_2(g)$ for any $f, g \in \mathcal{K}_3$ and $k, l \in \mathbb{Z}$.*

Corollary 3.7. *If at least one of $v_2(f)$ and $v_2(g)$ is not zero, then*

$$[I(\Gamma)|_{\mathcal{K}_3(f\sharp g)}] \in H^1_{DR}(\mathcal{K}_3(f\sharp g)) \neq 0,$$

where $\sharp$ stands for the connected sum.

Proof. This is because $r_*\lambda(f^k, g^l)$ is a one-cycle of $\mathcal{K}_3(f\sharp g)$ for any $k, l \in \mathbb{Z}$. Since $v_2(f)$ or $v_2(g)$ is not zero, there exist some k, l such that $lv_2(f) + kv_2(g) \neq 0$, so $\langle I(\Gamma),\ r_*\lambda(f^k, g^l)\rangle \neq 0$ by Theorem 3.6. □

Remark 3.8. If $v_2(f) = -v_2(g)$, then $v_2(f\sharp g) = 0$ since it is known that v_2 is additive under $\sharp$. Hence we cannot deduce $[I(\Gamma)|_{\mathcal{K}_3(f\sharp g)}] \neq 0$ from Corollary 3.2. Moreover if $v_2(f) = -v_2(g) \neq 0$, then Corollary 3.7 implies $[I(\Gamma)|_{\mathcal{K}_3(f\sharp g)}] \neq 0$.

To prove Theorem 3.6, first we remark that $f^m \sim f^0\sharp\iota^m$. Since λ satisfies the Leibniz rule, $\lambda(f^k, g^l)$ is homologous to

$$\lambda(f^0, g^0)\sharp\iota^{k+l} + \lambda(f^0, \iota^l)\sharp g^k + \lambda(\iota^k, g^0)\sharp f^l + \lambda(\iota^k, \iota^l)\sharp f^0\sharp g^0.$$

Since by definition $r_*\lambda(f^k, \iota^m) \sim mG_f$ ($k, m \in \mathbb{Z}$) and $G_\iota \sim 0$,

$$r_*\lambda(f^k, g^l) \sim r_*\lambda(f^0, g^0) + lG_f\sharp g + kf\sharp G_g. \tag{3-2}$$

Notice that $\sharp$ makes $\mathcal{K}_3$ an H-space and induces a coproduct Δ on $H^*_{DR}(\mathcal{K}_3)$.

Lemma 3.9. $\Delta([I(\Gamma)]) = 1 \otimes [I(\Gamma)] + [I(\Gamma)] \otimes 1 \in H^*_{DR}(\mathcal{K}_3)^{\otimes 2}$.

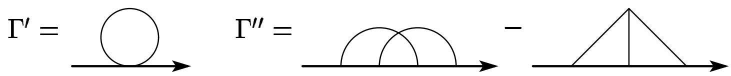


Figure 3. Graph cocycles Γ' and Γ''.

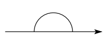
Proof. $\mathfrak{D}$ also admits Δ defined as a "separation" of the graphs by removing a point from the specified oriented line [Cattaneo et al. 2005, Section 3.2]. Theorem 6.3 of [Cattaneo et al. 2005] shows, without using $n > 3$, that $(I \otimes I)\Delta(X) = \Delta I(X)$ if X satisfies $dI(X) = I(\delta X)$.

As for our graphs in Figure 1, $\Delta\Gamma_i = 1 \otimes \Gamma_i + \Gamma_i \otimes 1$ $(i \neq 3, 4)$ and

$$\Delta(\Gamma_3 - \Gamma_4) = 1 \otimes (\Gamma_3 - \Gamma_4) + (\Gamma_3 - \Gamma_4) \otimes 1 + \Gamma' \otimes \Gamma'' + \Gamma'' \otimes \Gamma',$$

where Γ' and Γ'' are as shown in Figure 3. Thus

$$\Delta I(\Gamma) = 1 \otimes I(\Gamma) + I(\Gamma) \otimes 1 + I(\Gamma') \otimes I(\Gamma'') + I(\Gamma'') \otimes I(\Gamma').$$

But in fact $\Gamma' = \delta\Gamma_0$ where $\Gamma_0 =$ , and $I(\Gamma') = dI(\Gamma_0)$ since there is no hidden face in the boundary of the fiber of π_{Γ_0}. □

By (3-2), Lemma 3.9 and Theorem 3.1,

$$\langle I(\Gamma),\ r_*\lambda(f^k, g^l)\rangle = \langle I(\Gamma),\ r_*\lambda(f^0, g^0)\rangle + l v_2(f) + k v_2(g).$$

Thus it suffices to prove Theorem 3.6 in the case $k = l = 0$.

Proof of Theorem 3.6. Fix g and regard $\langle I(\Gamma),\ r_*\lambda(f^0, g^0)\rangle$ as an invariant $V_g(f)$ of f. We choose two crossings p_1 and p_2 from the diagram of f in xy-plane, and compute $D^2 V_g(f) := \sum_{\varepsilon_1,\varepsilon_2} \varepsilon_1\varepsilon_2 \langle I(\Gamma), r_*\lambda(f^0_{\varepsilon_1,\varepsilon_2}, g^0)\rangle$ in the limit $h \to 0$ as on page 414. If this is zero for any (p_1, p_2), then the arguments similar to that in the proof of Theorem 3.1 show that V_g is of order two and takes the value zero for the trefoil knot, thus identically $V_g = 0$ for any g. This will complete the proof.

We will compute each $P'_i := \sum_{\varepsilon=\pm 1} \langle I(\Gamma_i), r_*\lambda(f^0_{\varepsilon_1,\varepsilon_2}, g^0)\rangle$ $(1 \le i \le 9)$ in the limit $h \to 0$. The two observations appearing in the proof of Lemma 3.4 allow us to conclude $P'_i \to 0$ for $4 \le i \le 9$ in the same way as before, so we compute P'_i for $i = 1, 2, 3$ below. We may concentrate on the integration over C'_{Γ_i} by the first observation. Recall $C'_{\Gamma_i} \subset S^1 \times \mathrm{Conf}(\mathbb{R}^1, s) \times \mathrm{Conf}(\mathbb{R}^3, t)$ by definition. We take the S^1-parameter $\alpha \in S^1 = \mathbb{R}^1/2\pi\mathbb{Z}$ so that g goes through f during $0 \le \alpha \le \pi$, and f goes through g during $\pi \le \alpha \le 2\pi$.

First consider the integration over $0 \le \alpha \le \pi$. We may shrink g sufficiently small. Then the sliding of g through f does not affect the integration, so almost all the integrations converge to zero for the same reasons as in Lemmas 3.4 and 3.5. Only the configurations $(x_i) \in C'_{\Gamma_1}$ with x_1 and x_2 near p_1 may essentially contribute to P'_1 when g comes around p_1; the form $\varphi^*_{12}\mathrm{vol}_{S^2}$ may detect the knotting of g. However, the two terms for $\varepsilon_1 = \pm 1$ cancel each other.

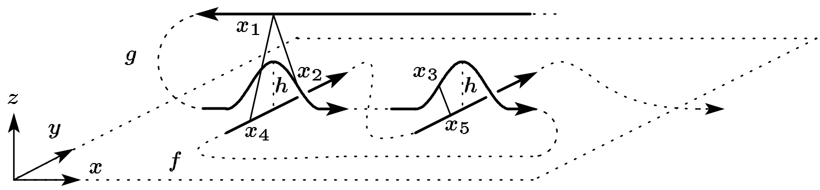
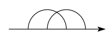


Figure 4. When f comes near an under-arc of g.

Next consider the integration over $\pi \le \alpha \le 2\pi$. There may be two types of contributions to P_i'. One type comes from the configurations in which all the points on the knot concentrate in a neighborhood of f. Such a contribution depends only on the framing number $\mathrm{fr}\, g$ of g, not on the global knotting of g. Since $\mathrm{fr}\, g^0 = 0$ here, such configurations do not essentially contribute to P_i'.

The other possible contributions arise when f comes near the crossings of g. For example, consider the case that (p_1, p_2) respects . When f comes near a crossing of g, a configuration $(x_1, \dots, x_5) \in C_{\Gamma_1}$ as in Figure 4 is certainly in C'_{Γ_1}, so it may contribute to P_1'.

However, such contributions converge to zero in the limit $h \to 0$, because x_1 cannot be near p_1 (see the second observation in the proof of Lemma 3.4). For Γ_3, we should take the configuration $(x_1, \dots, x_5)$ with x_j $(2 \le j \le 5)$ near t_{j-1} into account; but in this case the Gauss map φ_{11} cannot have the image in the support of vol_{S^2}. In such ways we can check that all such contributions of Γ_i $(i = 1, 2, 3)$ can be arbitrarily small. □

References

[Altschuler and Freidel 1997] D. Altschuler and L. Freidel, "Vassiliev knot invariants and Chern–Simons perturbation theory to all orders", *Comm. Math. Phys.* **187**:2 (1997), 261–287. MR 99c:57008 Zbl 0949.57012

[Bott and Taubes 1994] R. Bott and C. Taubes, "On the self-linking of knots", *J. Math. Phys.* **35**:10 (1994), 5247–5287. MR 95g:57008 Zbl 0863.57004

[Budney 2007] R. Budney, "Little cubes and long knots", *Topology* **46**:1 (2007), 1–27. MR 2008c:55015 Zbl 1114.57003

[Budney 2010] R. Budney, "Topology of knot spaces in dimension 3", *Proc. Lond. Math. Soc.* (3) **101**:2 (2010), 477–496. MR 2679699 Zbl 1201.57021

[Budney and Cohen 2009] R. Budney and F. Cohen, "On the homology of the space of knots", *Geom. Topol.* **13**:1 (2009), 99–139. MR 2009k:55029 Zbl 1163.57027

[Budney et al. 2005] R. Budney, J. Conant, K. P. Scannell, and D. Sinha, "New perspectives on self-linking", *Adv. Math.* **191**:1 (2005), 78–113. MR 2005h:57013 Zbl 1078.57011

[Cattaneo et al. 2002] A. S. Cattaneo, P. Cotta-Ramusino, and R. Longoni, "Configuration spaces and Vassiliev classes in any dimension", *Algebr. Geom. Topol.* **2** (2002), 949–1000. MR 2004a:57014 Zbl 1029.57009

[Cattaneo et al. 2005] A. S. Cattaneo, P. Cotta-Ramusino, and R. Longoni, "Algebraic structures on graph cohomology", *J. Knot Theory Ramifications* **14**:5 (2005), 627–640. MR 2006g:58021 Zbl 1091.57008

[Gramain 1977] A. Gramain, "Sur le groupe fundamental de l'espace des noeuds", *Ann. Inst. Fourier* (*Grenoble*) **27**:3 (1977), 29–44. MR 57 #1552

[Hatcher 1983] A. E. Hatcher, "A proof of the Smale conjecture, $\mathrm{Diff}(S^3) \simeq \mathrm{O}(4)$", *Ann. of Math.* (2) **117**:3 (1983), 553–607. MR 85c:57008 Zbl 0531.57028

[Kohno 1994] T. Kohno, "Vassiliev invariants and de Rham complex on the space of knots", pp. 123–138 in *Symplectic geometry and quantization* (Sanda and Yokohama, 1993), edited by Y. Maeda et al., Contemp. Math. **179**, Amer. Math. Soc., Providence, RI, 1994. MR 96g:57010 Zbl 0876.57009

[Longoni 2004] R. Longoni, "Nontrivial classes in $H^*(\mathrm{Imb}(S^1, \mathbb{R}^n))$ from nontrivalent graph cocycles", *Int. J. Geom. Methods Mod. Phys.* **1**:5 (2004), 639–650. MR 2005k:57026 Zbl 1069.81037

[Sakai 2008] K. Sakai, "Nontrivalent graph cocycle and cohomology of the long knot space", *Algebr. Geom. Topol.* **8**:3 (2008), 1499–1522. MR 2009h:58023 Zbl 1151.57012

[Turchin 2006] V. Turchin, "Calculating the first nontrivial 1-cocycle in the space of long knots", *Mathematical Notes* **80** (2006), 101–108. Zbl 1143.57007

[Vassiliev 1992] V. A. Vassiliev, *Complements of discriminants of smooth maps: topology and applications*, Transl. Math. Monographs **98**, Amer. Math. Soc., Providence, RI, 1992. MR 94i:57020

[Vassiliev 2001] V. A. Vassiliev, "Combinatorial formulas for cohomology of knot spaces", *Mosc. Math. J.* **1**:1 (2001), 91–123. MR 2002g:55028 Zbl 1015.57003

[Volić 2007] I. Volić, "A survey of Bott–Taubes integration", *J. Knot Theory Ramifications* **16**:1 (2007), 1–42. MR 2008f:57022 Zbl 1128.57013

Received April 29, 2010.

KEIICHI SAKAI
DEPARTMENT OF MATHEMATICAL SCIENCES
SHINSHU UNIVERSITY
3-1-1 ASAHI, MATSUMOTO 390-8621
JAPAN

ksakai@math.shinshu-u.ac.jp
http://math.shinshu-u.ac.jp/~ksakai/index.html

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

BURGHELEA–HALLER ANALYTIC TORSION FOR TWISTED DE RHAM COMPLEXES

GUANGXIANG SU

We extend the Burghelea–Haller analytic torsion to the twisted de Rham complexes, and compare it with the twisted refined analytic torsion defined by Huang. Finally, we briefly discuss the Cappell–Miller analytic torsion.

1. Introduction

Let E be a unitary flat vector bundle on a closed Riemannian manifold M. Ray and Singer [1971] defined an analytic torsion associated to (M, E) and proved that it does not depend on the Riemannian metric on M. Moreover, they conjectured that this analytic torsion coincides with the classical Reidemeister torsion defined using a triangulation on M (see [Milnor 1966]). This conjecture was later proved in two celebrated papers [Cheeger 1979; Müller 1978]. Müller [1993] generalized this result to the case when E is a unimodular flat vector bundle on M. Inspired by the considerations of Quillen [1985], Bismut and Zhang [1992] reformulated the above Cheeger–Müller theorem as an equality between the Reidemeister and Ray–Singer metrics defined on the determinant of cohomology, and proved an extension of it to the case of general flat vector bundle over M. The method used by Bismut and Zhang is different from that of Cheeger and Müller in that it makes use of a deformation by Morse functions introduced by Witten [1982] on the de Rham complex.

Braverman and Kappeler [2007b; 2007c; 2008] defined the refined analytic torsion for a flat vector bundle over an odd dimensional manifold and showed that it equals the Turaev torsion [1989] (see also [Farber and Turaev 2000]) up to multiplication by a complex number of absolute value one. Burghelea and Haller [2007; 2008], following a suggestion of Müller, defined a generalized analytic torsion associated to a nondegenerate symmetric bilinear form on a flat vector bundle over an arbitrary dimensional manifold and make an explicit conjecture between this generalized analytic torsion and the Turaev torsion. This conjecture was proved up to sign in [Burghelea and Haller 2010] and in full generality in [Su and Zhang 2008]. Cappell and Miller [2010] used non-self-adjoint Laplace operators to define

MSC2000: primary 58J52; secondary 19K56.
Keywords: analytic torsion, symmetric bilinear form, de Rham complex.

another complex-valued analytic torsion and used the method in [Su and Zhang 2008] to prove an extension of the Cheeger–Müller theorem.

Mathai and Wu [2008; 2010b] generalized the classical Ray–Singer analytic torsion to the twisted de Rham complex with an odd degree closed differential form H. In [Mathai and Wu 2010a], they defined and studied analytic torsion of $\mathbb{Z}_2$-graded elliptic complexes. Huang [2010a] generalized Braverman and Kappeler's refined analytic torsion to the twisted de Rham complex, proved a duality theorem and compared it with the twisted Ray–Singer metric.

In this paper, supposing there exists a nondegenerate symmetric bilinear form on the flat vector bundle E, we generalize the Burghelea–Haller analytic torsion to the twisted de Rham complex. For the odd dimensional manifold, we also compare it with the twisted refined analytic torsion and the twisted Ray–Singer metric.

The rest of this paper is organized as follows. In Section 2, supposing there exists a $\mathbb{Z}_2$-graded nondegenerate symmetric bilinear form on a $\mathbb{Z}_2$-graded finite dimensional complex, we define a symmetric bilinear torsion on it. In Section 3, we generalize the Burghelea–Haller analytic torsion to the twisted de Rham complex. In Section 4, when the dimension of the manifold is odd, we show that the twisted Burghelea–Haller analytic torsion is independent of the Riemannian metric g, the symmetric bilinear form b and the representative H in the cohomology class $[H]$. In Section 5, we compare this new torsion with the twisted refined analytic torsion. In Section 6, we briefly discuss the Cappell–Miller analytic torsion on the twisted de Rham complex of an odd dimensional manifold.

2. Symmetric bilinear torsion on a finite dimensional $\mathbb{Z}_2$-graded complex

Consider a cochain complex

$$0 \longrightarrow C^0 \xrightarrow{d_0} C^1 \xrightarrow{d_1} \cdots \xrightarrow{d_{n-1}} C^n \longrightarrow 0$$

of finite dimensional complex vector space. Set

$$C^{\bar{k}} = \bigoplus_{i = k \bmod 2} C^i, \quad k = 0, 1.$$

Let

$$(C^\bullet, d) : \cdots \xrightarrow{d_{\bar{1}}} C^{\bar{0}} \xrightarrow{d_{\bar{0}}} C^{\bar{1}} \xrightarrow{d_{\bar{1}}} C^{\bar{0}} \xrightarrow{d_{\bar{0}}} \cdots \tag{2-1}$$

be a $\mathbb{Z}_2$-graded cochain complex. Denote its cohomology by $H^{\bar{k}}$, $k = 0, 1$. Set

$$\det(C^\bullet, d) = \det C^{\bar{0}} \otimes (\det C^{\bar{1}})^{-1}, \quad \det(H^\bullet, d) = \det H^{\bar{0}} \otimes (\det H^{\bar{1}})^{-1}.$$

Then we have a canonical isomorphism between the determinant lines

$$\phi : \det(C^\bullet, d) \to \det(H^\bullet, d). \tag{2-2}$$

Suppose that there is a nondegenerate symmetric bilinear form on $C^{\bar{k}}$, $k=0,1$. Then it induces a nondegenerate symmetric bilinear form $b_{\det H^\bullet(C^\bullet,d)}$ on the determinant line $\det(H^\bullet, d)$ via the isomorphism (2-2). Let $d_{\bar{k}}^{\#}$ be the adjoint of $d_{\bar{k}}$ with respect to the nondegenerate symmetric bilinear form and define

$$\Delta_{b,\bar{k}} = d_{\bar{k}}^{\#} d_{\bar{k}} + d_{\overline{k+1}} d_{\overline{k+1}}^{\#}.$$

Let λ be the generalized eigenvalue of $\Delta_{b,\bar{k}}$ and let $C_b^{\bar{k}}(\lambda)$ be the generalized λ-eigenspace of $\Delta_{b,\bar{k}}$. Then we have a b-orthogonal decomposition

$$C^{\bar{k}} = \bigoplus_{\lambda} C_b^{\bar{k}}(\lambda) \tag{2-3}$$

and the inclusion $C_b^{\bar{k}}(0) \to C^{\bar{k}}$ induces an isomorphism in cohomology. Particularly, we obtain a canonical isomorphism

$$\det H^\bullet(C_b^\bullet(0)) \cong \det H^\bullet(C^\bullet). \tag{2-4}$$

Proposition 2.1. *The following identity holds*:

$$\begin{aligned} &b_{\det H^\bullet(C^\bullet,d)} \\ &\quad = b_{\det H^\bullet(C_b^\bullet(0),d)} \cdot \det\big(d_{\bar{0}}^{\#} d_{\bar{0}}|_{C_b^{\bar{0},\perp}(0)\cap \operatorname{im} d_{\bar{0}}^{\#}}\big)^{-1} \cdot \det\big(d_{\bar{1}}^{\#} d_{\bar{1}}|_{C_b^{\bar{1},\perp}(0)\cap \operatorname{im} d_{\bar{1}}^{\#}}\big), \end{aligned} \tag{2-5}$$

where $C_b^{\bar{k},\perp}(0) = \bigoplus_{\lambda\neq 0} C_b^{\bar{k}}(\lambda)$, $k=0,1$.

Proof. Same as [Burghelea and Haller 2007, Lemma 3.3]. Suppose $(C_1^\bullet, b_1)$ and $(C_2^\bullet, b_2)$ are finite-dimensional $\mathbb{Z}_2$-graded complexes equipped with $\mathbb{Z}_2$-graded nondegenerate symmetric bilinear forms. Clearly, $H^\bullet(C_1^\bullet \oplus C_2^\bullet) = H^\bullet(C_1^\bullet) \oplus H^\bullet(C_2^\bullet)$ and we obtain a canonical isomorphism of determinant lines

$$\det H^\bullet(C_1^\bullet \oplus C_2^\bullet) = \det H^\bullet(C_1^\bullet) \otimes \det H^\bullet(C_2^\bullet).$$

Then we have

$$b_{\det H^\bullet(C_1^\bullet\oplus C_2^\bullet)} = b_{\det H^\bullet(C_1^\bullet)} \otimes b_{\det H^\bullet(C_2^\bullet)}.$$

In view of the b-orthogonal decomposition (2-3) we may therefore without loss of generality assume $\ker \Delta_{b,\bar{k}} = 0$, $k=0,1$. Then by the lemma just cited we have

$$C^{\bar{k}} = \operatorname{im} d_{\overline{k+1}} \oplus \operatorname{im} d_{\bar{k}}^{\#}.$$

This decomposition is b-orthogonal and invariant under Δ_b. Thus we have the exact complexes

$$0 \longrightarrow C^{\bar{0}} \cap \operatorname{im} d_{\bar{0}}^{\#} \xrightarrow{d_{\bar{0}}} C^{\bar{1}} \cap \operatorname{im} d_{\bar{0}} \longrightarrow 0,$$
$$0 \longrightarrow C^{\bar{1}} \cap \operatorname{im} d_{\bar{1}}^{\#} \xrightarrow{d_{\bar{1}}} C^{\bar{0}} \cap \operatorname{im} d_{\bar{1}} \longrightarrow 0.$$

Then from [Burghelea and Haller 2007, Example 3.2], we get the proposition. □

3. Symmetric bilinear torsion on the twisted de Rham complexes

In this section, we suppose that there is a fiberwise nondegenerate symmetric bilinear form on E. Then we define a symmetric bilinear torsion on the determinant line of the twisted de Rham complex.

Twisted de Rham complexes. In this section, we review the twisted de Rham complexes from [Mathai and Wu 2008].

Let M be a closed Riemannian manifold and $E \to M$ be a complex flat vector bundle with flat connection ∇. Let H be an odd-degree closed differential form on M. We set $\Omega^{\bar{0}} = \Omega^{\mathrm{even}}(M, E)$, $\Omega^{\bar{1}} = \Omega^{\mathrm{odd}}(M, E)$ and $\nabla^H = \nabla + H\wedge$. We define the twisted de Rham cohomology groups as

$$H^{\bar{k}}(M, E, H) = \frac{\ker\big(\nabla^H : \Omega^{\bar{k}}(M, E) \to \Omega^{\overline{k+1}}(M, E)\big)}{\mathrm{im}\big(\nabla^H : \Omega^{\overline{k+1}}(M, E) \to \Omega^{\bar{k}}(M, E)\big)}, \quad k = 0, 1.$$

Suppose H is replaced by $H' = H - dB$ for some $B \in \Omega^{\bar{0}}(M)$, then there is an isomorphism $\varepsilon_B = e^B \wedge \cdot : \Omega^\bullet(M, E) \to \Omega^\bullet(M, E)$ satisfying

$$\varepsilon_B \circ \nabla^H = \nabla^{H'} \circ \varepsilon_B.$$

Therefore ε_B induces an isomorphism

$$\varepsilon_B : H^\bullet(M, E, H) \to H^\bullet(M, E, H')$$

on the twisted de Rham cohomology.

The construction of the symmetric bilinear torsion. Suppose that there exists a nondegenerate symmetric bilinear form on E. To simplify notation, let $C^{\bar{k}} = \Omega^{\bar{k}}(X, E)$ and let $d_{\bar{k}} = d_{\bar{k}}^{E,H}$ be the operator ∇^H acting on $C^{\bar{k}}$ ($k = 0, 1$). Then $d_{\bar{1}} d_{\bar{0}} = d_{\bar{0}} d_{\bar{1}} = 0$ and we have a complex

$$\cdots \xrightarrow{d_{\bar{1}}} C^{\bar{0}} \xrightarrow{d_{\bar{0}}} C^{\bar{1}} \xrightarrow{d_{\bar{1}}} C^{\bar{0}} \xrightarrow{d_{\bar{0}}} \cdots. \tag{3-1}$$

The metric g^M and the symmetric bilinear form b determine together a symmetric bilinear form on $\Omega^\bullet(M, E)$ such that if $u = \alpha f$, $v = \beta g \in \Omega^\bullet(M, E)$ such that $\alpha, \beta \in \Omega^\bullet(M)$, $f, g \in \Gamma(E)$, then

$$\beta_{g,b}(u, v) = \int_M (\alpha \wedge *\beta) b(f, g), \tag{3-2}$$

where $*$ is the Hodge star operator. Denote by $d_{\bar{k}}^{\#}$ the adjoint of $d_{\bar{k}}$ with respect to the nondegenerate symmetric bilinear form (3-2). Then we define the Laplacians

$$\Delta_{b,\bar{k}} = d_{\bar{k}}^{\#} d_{\bar{k}} + d_{\overline{k+1}} d_{\overline{k+1}}^{\#}, \quad k = 0, 1.$$

If λ is in the spectrum of $\Delta_{b,\bar{k}}$, then the image of the associated spectral projection is finite dimensional and contains smooth forms only. Referring to this image as the (generalized) λ-eigenspace of $\Delta_{b,\bar{k}}$ and denoting it by $\Omega^{\bar{k}}_{\{\lambda\}}(M, E)$, there exists $N_\lambda \in \mathbb{N}$ such that

$$(\Delta_{b,\bar{k}} - \lambda)^{N_\lambda}|_{\Omega^{\bar{k}}_{\{\lambda\}}(M,E)} = 0.$$

Therefore for different generalized eigenvalues λ, μ, the spaces $\Omega^{\bar{k}}_{\{\lambda\}}(M, E)$ and $\Omega^{\bar{k}}_{\{\mu\}}(M, E)$ are $\beta_{g,b}$-orthogonal.

For any $a \geq 0$, set

$$\Omega^{\bar{k}}_{[0,a]}(M, E) = \bigoplus_{0\leq|\lambda|\leq a} \Omega^{\bar{k}}_{\{\lambda\}}(M, E).$$

Then $\Omega^{\bar{k}}_{[0,a]}(M, E)$ is finite dimensional and one gets a nondegenerate symmetric bilinear form

$$b_{\det H^\bullet(\Omega^\bullet_{[0,a]},d)} \quad \text{on} \quad \det H^\bullet(\Omega^\bullet_{[0,a]}, d).$$

Let $\Omega^{\bar{k}}_{(a,+\infty)}(M, E)$ denote the $\beta_{g,b}$-orthogonal complement to $\Omega^{\bar{k}}_{[0,a]}(M, E)$.

For the subcomplexes $(\Omega^{\overline{k+1}}_{(a,+\infty)}(M, E), d)$, since the operators $d_{\bar{k}} d^{\#}_{\bar{k}}$ and $\Delta_{b,\overline{k+1}}$ are equal and invertible on $\mathrm{im}(d_{\bar{k}}) \cap \Omega^{\overline{k+1}}_{(a,+\infty)}(M, E)$, we have

$$P_{\bar{k}} := d^{\#}_{\bar{k}}(d_{\bar{k}} d^{\#}_{\bar{k}})^{-1} d_{\bar{k}} = d^{\#}_{\bar{k}}(\Delta_{b,\overline{k+1}})^{-1} d_{\bar{k}} \tag{3-3}$$

is a pseudodifferential operator of order 0 and satisfies

$$P^2_{\bar{k}} = P_{\bar{k}}.$$

By definition we have

$$\begin{aligned} \zeta(s, d^{\#}_{\bar{k}} d_{\bar{k}}|_{\mathrm{im}\, d^{\#}_{\bar{k}} \cap \Omega^{\bar{k}}_{(a,+\infty)}(M,E)}) &= \mathrm{Tr}(\Delta^{-s}_{b,\bar{k}} P_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)}) \\ &= \mathrm{Tr}(P_{\bar{k}} \Delta^{-s}_{b,\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)}). \end{aligned} \tag{3-4}$$

Then $\zeta(s, d^{\#}_{\bar{k}} d_{\bar{k}}|_{\mathrm{im}\, d^{\#}_{\bar{k}} \cap \Omega^{\bar{k}}_{(a,+\infty)}(M,E)})$ has a meromorphic extension to the whole complex plane and, by [Wodzicki 1984, Section 7], it is regular at 0. So by [Wodzicki 1984; Grubb and Seeley 1995], we have the following analogue of [Mathai and Wu 2008, Theorem 2.1].

Theorem 3.1. *For $k = 0, 1$, $\zeta(s, d^{\#}_{\bar{k}} d_{\bar{k}}|_{\mathrm{im}\, d^{\#}_{\bar{k}} \cap \Omega^\bullet_{(a,+\infty)}(M,E)})$ is holomorphic in the half plane for $\mathrm{Re}(s) > n/2$ and extends meromorphically to $\mathbb{C}$ with possible poles at $\{(n-l)/2,\ l = 0, 1, 2, \dots\}$ only, and is holomorphic at $s = 0$.*

Then for $k = 0, 1$ and any $a \geq 0$, the regularized zeta determinant

$$\det{}'(d^{\#}_{\bar{k}} d_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)}) := \exp\bigl(-\zeta'(0, d^{\#}_{\bar{k}} d_{\bar{k}}|_{\mathrm{im}\, d^{\#}_{\bar{k}} \cap \Omega^{\bar{k}}_{(a,+\infty)}(M,E)})\bigr). \tag{3-5}$$

is well defined.

Proposition 3.2. *The symmetric bilinear form on* $\det H^\bullet(\Omega^\bullet(M,E,H),d)$ *given by*

$$b_{\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E),d)}\cdot\det{}'(d_{\bar 0}^{\#}d_{\bar 0}|_{\Omega^{\bar 0}_{(a,+\infty)}(M,E)})^{-1}\cdot\big(\det{}'(d_{\bar 1}^{\#}d_{\bar 1}|_{\Omega^{\bar 1}_{(a,+\infty)}(M,E)})\big)\tag{3-6}$$

is independent of the choice of $a\geq 0$.

Proof. Let $0\leq a<c<\infty$. We have

$$(\Omega^{\bar k}_{[0,c]}(M,E),d_{\bar k})=(\Omega^{\bar k}_{[0,a]}(M,E),d_{\bar k})\oplus(\Omega^{\bar k}_{(a,c]}(M,E),d_{\bar k}),\tag{3-7}$$

$$(\Omega^{\bar k}_{(a,+\infty)}(M,E),d_{\bar k})=(\Omega^{\bar k}_{(a,c]}(M,E),d_{\bar k})\oplus(\Omega^{\bar k}_{(c,+\infty)}(M,E),d_{\bar k}).\tag{3-8}$$

By the definition of the determinant,

$$\det{}'(d_{\bar k}^{\#}d_{\bar k}|_{\Omega^{\bar k}_{(a,+\infty)}(M,E)})=\det{}'(d_{\bar k}^{\#}d_{\bar k}|_{\Omega^{\bar k}_{(a,c]}(M,E)})\cdot\det{}'(d_{\bar k}^{\#}d_{\bar k}|_{\Omega^{\bar k}_{(c,+\infty)}(M,E)}).\tag{3-9}$$

Applying Proposition 2.1 to (3-7),

$$b_{\det H^\bullet(\Omega^\bullet_{[0,c]})}=b_{\det H^\bullet(\Omega^\bullet_{[0,a]})}\cdot\det{}'(d_{\bar 0}^{\#}d_{\bar 0}|_{\Omega^{\bar 0}_{(a,c]}(M,E)})^{-1}\cdot\big(\det{}'(d_{\bar 1}^{\#}d_{\bar 1}|_{\Omega^{\bar 1}_{(a,c]}(M,E)})\big).$$

Then we get the proposition. □

Definition 3.3. The symmetric bilinear form defined by (3-6) is called the Ray–Singer symmetric bilinear torsion on $\det H^\bullet(\Omega^\bullet(M,E,H),d)$ and is denoted by $\tau_{b,\nabla,H}$.

4. Symmetric bilinear torsion under metric and flux deformations

In this section, we will use the methods in [Mathai and Wu 2008] to study the dependence of the torsion on the metric g, the symmetric bilinear form b and the flux H.

Variation of the torsion with respect to the metric and symmetric bilinear form. We assume that M is a closed compact oriented manifold of odd dimension. Suppose the pair (g_u,b_u) is deformed smoothly along a one-parameter family with parameter $u\in\mathbb{R}$. Let $Q_{\bar k}$ be the spectral projection onto $\Omega^{\bar k}_{[0,a]}(M,E)$ and $\Pi_{\bar k}=1-Q_{\bar k}$ be the spectral projection onto $\Omega^{\bar k}_{(a,+\infty)}(M,E)$. Let

$$\alpha=*_u^{-1}\frac{\partial *_u}{\partial u}+b_u^{-1}\frac{\partial b_u}{\partial u}.$$

Lemma 4.1. *Under the assumptions above*,

$$\frac{\partial}{\partial u}\log\Big(\det{}'(d_{\bar 0}^{\#}d_{\bar 0}|_{\Omega^{\bar 0}_{(a,+\infty)}(M,E)})^{-1}\cdot\big(\det{}'(d_{\bar 1}^{\#}d_{\bar 1}|_{\Omega^{\bar 1}_{(a,+\infty)}(M,E)})\big)\Big)=-\sum_{k=0,1}(-1)^k\operatorname{Tr}(\alpha Q_{\bar k}).\tag{4-1}$$

Proof. While $d_{\bar{k}}$ is independent of u, we have

$$\frac{\partial d^{\#}_{\bar{k}}}{\partial u} = -[\alpha, d^{\#}_{\bar{k}}].$$

Using $P_{\bar{k}} d^{\#}_{\bar{k}} = d^{\#}_{\bar{k}}$, $d_{\bar{k}} P_{\bar{k}} = d_{\bar{k}}$ and $P^2_{\bar{k}} = P_{\bar{k}}$, we get $d^{\#}_{\bar{k}} d_{\bar{k}} P_{\bar{k}} = P_{\bar{k}} d^{\#}_{\bar{k}} d_{\bar{k}} = d^{\#}_{\bar{k}} d_{\bar{k}}$ and

$$P_{\bar{k}} \frac{\partial P_{\bar{k}}}{\partial u} P_{\bar{k}} = 0.$$

Following the $\mathbb{Z}$-graded case, we set

$$\begin{aligned} f(s,u) &= \sum_{k=0,1} (-1)^k \int_0^{+\infty} t^{s-1} \operatorname{Tr}(e^{-td^{\#}_{\bar{k}} d_{\bar{k}}} P_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)})\, dt \\ &= \Gamma(s) \sum_{k=0,1} (-1)^k \zeta(s, d^{\#}_{\bar{k}} d_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)}). \end{aligned} \tag{4-2}$$

Using the above identities on $P_{\bar{k}}$, the trace property and by an application of Duhamel's principal, we get

$$\begin{aligned} \frac{\partial f}{\partial u} &= \sum_{k=0,1} (-1)^k \int_0^{+\infty} t^{s-1} \operatorname{Tr}\Big(t[\alpha, d^{\#}_{\bar{k}}] d_{\bar{k}} e^{-td^{\#}_{\bar{k}} d_{\bar{k}}} \Pi_{\bar{k}} + e^{-td^{\#}_{\bar{k}} d_{\bar{k}}} \frac{\partial P_{\bar{k}}}{\partial u} P_{\bar{k}} \Pi_{\bar{k}} \Big)\, dt \\ &= \sum_{k=0,1} (-1)^k \int_0^{+\infty} t^{s-1} \operatorname{Tr}\Big(t\alpha[d^{\#}_{\bar{k}}, d_{\bar{k}} e^{-td^{\#}_{\bar{k}} d_{\bar{k}}}] \Pi_{\bar{k}} + P_{\bar{k}} e^{-td^{\#}_{\bar{k}} d_{\bar{k}}} \frac{\partial P_{\bar{k}}}{\partial u} \Pi_{\bar{k}} \Big)\, dt \\ &= \sum_{k=0,1} (-1)^k \int_0^{+\infty} t^{s-1} \operatorname{Tr}\Big(t\alpha (e^{-td^{\#}_{\bar{k}} d_{\bar{k}}} d^{\#}_{\bar{k}} d_{\bar{k}} - e^{-td_{\bar{k}} d^{\#}_{\bar{k}}} d_{\bar{k}} d^{\#}_{\bar{k}}) \Pi_{\bar{k}} \\ &\qquad\qquad + e^{-td^{\#}_{\bar{k}} d_{\bar{k}}} P_{\bar{k}} \frac{\partial P_{\bar{k}}}{\partial u} \Pi_{\bar{k}} \Big)\, dt \\ &= \sum_{k=0,1} (-1)^k \int_0^{+\infty} t^{s} \operatorname{Tr}\big(\alpha e^{-t\Delta_{b,\bar{k}}} \Delta_{b,\bar{k}} \Pi_{\bar{k}} \big)\, dt \\ &= -\sum_{k=0,1} (-1)^k \int_0^{+\infty} t^{s} \frac{\partial}{\partial t} \operatorname{Tr}\big(\alpha (e^{-t\Delta_{b,\bar{k}}} \Pi_{\bar{k}}) \big)\, dt. \end{aligned} \tag{4-3}$$

Integrating by parts, we have

$$\begin{aligned} \frac{\partial f}{\partial u} &= s \sum_{k=0,1} (-1)^k \int_0^{+\infty} t^{s-1} \operatorname{Tr}\big(\alpha (e^{-t\Delta_{b,\bar{k}}} \Pi_{\bar{k}}) \big)\, dt \\ &= s \sum_{k=0,1} (-1)^k \left(\int_0^1 + \int_1^{+\infty} \right) t^{s-1} \operatorname{Tr}\left(\alpha e^{-t\Delta_{b,\bar{k}}} (1 - Q_{\bar{k}}) \right) dt. \end{aligned} \tag{4-4}$$

Since α is a smooth tensor and n is odd, the asymptotic expansion as $t \downarrow 0$ for $\operatorname{Tr}(\alpha e^{-t\Delta_{b,\bar{k}}})$ does not contain a constant term. Therefore $\int_0^1 t^{s-1} \operatorname{Tr}(\alpha e^{-t\Delta_{b,\bar{k}}})\, dt$

does not have a pole at $s = 0$. On the other hand, because of the exponential decay of $\mathrm{Tr}(\alpha e^{-t\Delta_{b,\bar{k}}}\Pi_{\bar{k}})$ for large t,

$$\int_1^{+\infty} t^{s-1}\,\mathrm{Tr}(\alpha e^{-t\Delta_{b,\bar{k}}}\Pi_{\bar{k}})$$

is an entire function in s. So

$$\text{(4-5)}\quad \frac{\partial f}{\partial u}\bigg|_{s=0} = -s\sum_{k=0,1}(-1)^k\int_0^1 t^{s-1}\,\mathrm{Tr}(\alpha Q_{\bar{k}})\,dt\big|_{s=0} = -\sum_{k=0,1}(-1)^k\,\mathrm{Tr}(\alpha Q_{\bar{k}})$$

and hence

$$\text{(4-6)}\qquad \frac{\partial}{\partial u}\sum_{k=0,1}(-1)^k\zeta(0, d_{\bar{k}}^{\#}d_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)}) = 0.$$

Finally, from (4-5), (4-6), we have

$$\text{(4-7)}\quad \det{}'(d_{\bar{0}}^{\#}d_{\bar{0}}|_{\Omega^{\bar{0}}_{(a,+\infty)}(M,E)})^{-1}\cdot\big(\det{}'(d_{\bar{1}}^{\#}d_{\bar{1}}|_{\Omega^{\bar{1}}_{(a,+\infty)}(M,E)})\big)$$
$$= \exp\bigg(\lim_{s\to 0}\bigg(f(s,u) - \frac{1}{s}\sum_{k=0,1}(-1)^k\zeta\big(0, d_{\bar{k}}^{\#}d_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)}\big)\bigg)\bigg),$$

and the result follows. □

Lemma 4.2. *Under the same assumptions, along any one-parameter deformation of* (g_u, b_u), *we have*

$$\text{(4-8)}\qquad \frac{\partial}{\partial w}\bigg|_u\left(\frac{b_{w,\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E),d)}}{b_{u,\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E),d)}}\right) = \sum_{k=0,1}(-1)^k\,\mathrm{Tr}(\alpha Q_{\bar{k}}).$$

Proof. For sufficiently small $w - u$, the restriction of the spectral projection

$$Q_{\bar{k}}|_{\Omega^{\bar{k}}_{u,[0,a]}(M,E)} : \Omega^{\bar{k}}_{u,[0,a]}(M,E) \to \Omega^{\bar{k}}_{w,[0,a]}(M,E)$$

is an isomorphism of complexes. Then for sufficiently small $w - u$, we have

$$\text{(4-9)}\quad \frac{b_{w,\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E),d)}}{b_{u,\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E),d)}}$$
$$= \det\big((\beta_{g_u,b_u}|_{\Omega^{\bar{0}}_{u,[0,a]}(M,E)})^{-1}(Q_{\bar{0}}|_{\Omega^{\bar{0}}_{u,[0,a]}(M,E)})^*$$
$$\cdot(\beta_{g_w,b_w}|_{\Omega^{\bar{0}}_{w,[0,a]}(M,E)})\big)\cdot\det\big((\beta_{g_u,b_u}|_{\Omega^{\bar{1}}_{u,[0,a]}(M,E)})^{-1}(Q_{\bar{1}}|_{\Omega^{\bar{1}}_{u,[0,a]}(M,E)})^*$$
$$\cdot(\beta_{g_w,b_w}|_{\Omega^{\bar{1}}_{w,[0,a]}(M,E)})\big)^{-1}.$$

Then similarly to [Burghelea and Haller 2007], we get (4-8). □

Combining Lemma 4.1 and Lemma 4.2, we have:

Theorem 4.3. *Let M be a closed, compact manifold of odd dimension, E be a flat vector bundle over M, and H be a closed differential form on M of odd degree. Then the symmetric bilinear torsion $\tau_{b,\nabla,H}$ on the twisted de Rham complex does not depend on the choices of the Riemannian metric on M and the symmetric bilinear form b in a same homotopy class of nondegenerate symmetric bilinear forms on E.*

Variation of analytic torsion with respect to the flux in a cohomology class. We continue to assume that $\dim M$ is odd and use the same notation as above. Suppose the (real) flux form H is deformed smoothly along a one-parameter family with parameter $v \in \mathbb{R}$ in such a way that the cohomology class $[H] \in H^{\bar{1}}(M, \mathbb{R})$ is fixed. Then $\partial H/\partial v = -dB$ for some form $B \in \Omega^{\bar{0}}(M)$ that depends smoothly on v; let

$$\beta = B \wedge \cdot\,.$$

Lemma 4.4. *Under the above assumptions,*

$$\frac{\partial}{\partial v}\log\Big(\det{}'(d_{\bar{0}}^{\#}d_{\bar{0}}|_{\Omega^{\bar{0}}_{(a,+\infty)}(M,E)})^{-1}\cdot\big(\det{}'(d_{\bar{1}}^{\#}d_{\bar{1}}|_{\Omega^{\bar{1}}_{(a,+\infty)}(M,E)})\big)\Big) = -2\sum_{k=0,1}(-1)^k\operatorname{Tr}(\beta Q_{\bar{k}}). \tag{4-10}$$

Proof. As in the proof of Lemma 4.1, we set

$$f(s,v) = \sum_{k=0,1}(-1)^k\int_0^{+\infty} t^{s-1}\operatorname{Tr}(e^{-td_{\bar{k}}^{\#}d_{\bar{k}}}P_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)})\,dt.$$

We note that B, hence β is real. Using

$$\frac{\partial d_{\bar{k}}}{\partial v} = [\beta, d_{\bar{k}}],\qquad \frac{\partial d_{\bar{k}}^{\#}}{\partial v} = -[\beta^{\#}, d_{\bar{k}}^{\#}],\qquad P_{\bar{k}}^2 = P_{\bar{k}} = P_{\bar{k}}^{\#},\qquad P_{\bar{k}}\frac{\partial P_{\bar{k}}}{\partial v}P_{\bar{k}} = 0$$

and Duhamel's principle, similarly to [Mathai and Wu 2008, Lemma 3.5], we get

$$\frac{\partial f}{\partial v} = -2\sum_{k=0,1}(-1)^k\int_0^{+\infty} t^s\frac{\partial}{\partial t}\operatorname{Tr}(\beta e^{-t\Delta_{b,\bar{k}}}\Pi_{\bar{k}})\,dt. \tag{4-11}$$

The rest is similar to the proof of Lemma 4.1. □

Lemma 4.5. *Under the same assumptions, along any one-parameter deformation of H that fixes the cohomology class $[H]$, we have*

$$\frac{\partial}{\partial w}\bigg|_v\left(\frac{b_{\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E,H^w),d)}}{b_{\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E,H^v),d)}}\right) = 2\sum_{k=0,1}(-1)^k\operatorname{Tr}(\beta Q_{\bar{k}}), \tag{4-12}$$

where we identify $\det H^\bullet(M, E, H)$ *along the deformation.*

Proof. For sufficiently small $w - v$, we have

$$Q_{\bar{k}}\varepsilon_B : \Omega^{\bar{k}}_{[0,a]}(M, E, H^v) \to \Omega^{\bar{k}}_{[0,a]}(M, E, H^w)$$

is an isomorphism of complexes and the induced symmetric bilinear form on the determinant line $\det H^\bullet(\Omega^\bullet_{[0,a]}(M, E, H^v), d)$ is

$$\begin{aligned}\text{(4-13)}\quad \Big(\big(\det(Q_{\bar{k}}\varepsilon_B)^* b_{\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E,H^w),d)}\big)\Big)(\,\cdot\,,\,\cdot\,)\\ = b_{\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E,H^w),d)}\big(\det(Q_{\bar{k}}\varepsilon_B)\,\cdot\,,\, \det(Q_{\bar{k}}\varepsilon_B)\,\cdot\,\big),\end{aligned}$$

where

$$\det(Q_{\bar{k}}\varepsilon_B) : \det H^\bullet\left(\Omega^*_{[0,a]}(M, E, H^v)\right) \to \det H^\bullet\left(\Omega^*_{[0,a]}(M, E, H^w)\right)$$

is the induced isomorphism on the determinant lines. Then we can compare it with $b_{\det H^\bullet(\Omega^\bullet_{[0,a]}(M,E,H^u),d)}$, and similarly to [Mathai and Wu 2008, Lemma 3.7], we get (4-12). □

Combining Lemma 4.4 and Lemma 4.5, we have:

Theorem 4.6. *Let M be a closed, compact manifold of odd dimension, E be a flat vector bundle over M. Suppose H and H' are closed differential forms on M of odd degrees representing the same de Rham cohomology class, and let B be an even form so that $H' = H - dB$. Then the symmetric bilinear torsion satisfies $(\det\varepsilon_B)^*\tau_{b,\nabla,H'} = \tau_{b,\nabla,H}$.*

5. Compare with the refined analytic torsion

In this section, we will compare the symmetric bilinear torsion $\tau_{b,\nabla,H}$ with the refined analytic torsion $\rho_{\mathrm{an}}(\nabla^H)$ defined in [Huang 2010a]. The main theorem of this section is the following.

Theorem 5.1. *Let M be a closed odd dimensional manifold, E be a complex vector bundle over M with connection ∇, H be a closed odd-degree differential form on M. Suppose there exists a nondegenerate symmetric bilinear form on E. Then*

$$\text{(5-1)}\qquad \tau_{b,\nabla,H}\left(\rho_{\mathrm{an}}(\nabla^H)\right) = \pm e^{-2\pi i\left(\eta(\nabla^H) - \operatorname{rank} E\cdot\eta_{\mathrm{trivial}}\right)}.$$

(Here $\eta(\nabla^H)$ and η_{trivial} are defined in [Huang 2010a].)

We will use the method in [Braverman and Kappeler 2007a] to prove the theorem and the proof will be given later.

Let h be a Hermitian metric on E. One can construct the Ray–Singer analytic torsion as an inner product on $\det H^\bullet(M, E, H)$, or equivalently as a metric on the determinant line; see [Huang 2010a, (6.13)]. We denote the resulting inner product by $\tau_{h,\nabla,H}$. Then by our Theorem 5.1 and Theorem 6.2 of the same reference, we get:

Corollary 5.2. *If* $\dim M$ *is odd*, $\left|\dfrac{\tau_{b,\nabla,H}}{\tau_{h,\nabla,H}}\right| = 1.$

The dual connection. Let M be an odd dimensional closed manifold and E be a flat vector bundle over M, with flat connection ∇. Assume that there exists a nondegenerate symmetric bilinear form b on E. The dual connection ∇' to ∇ on E with respect to the form b is defined by the formula

$$db(u, v) = b(\nabla u, v) + b(u, \nabla' v), \quad u, v \in \Gamma(M, E).$$

We denote by E' the flat vector bundle (E, ∇').

Choices of the metric and the spectral cut. Until the end of this section we fix a Riemannian metric g on M and set $\mathcal{B}^H = \mathcal{B}(\nabla^H, g) = \Gamma\nabla^H + \nabla^H\Gamma$ and $\mathcal{B}'^H = \mathcal{B}'(\nabla'^H, g) = \Gamma\nabla'^H + \nabla'^H\Gamma$, where $\Gamma : \Omega^\bullet(M, E) \to \Omega^\bullet(M, E)$ is the chirality operator defined by

$$\Gamma\omega = i^{(n+1)/2} * (-1)^{q(q+1)/2}\omega, \quad \omega \in \Omega^q(M, E).$$

We also fix $\theta \in (-\pi/2, 0)$ such that both θ and $\theta + \pi$ are Agmon angles for the odd signature operator $\mathcal{B}^H$. One easily checks that

$$(\nabla^H)^\# = \Gamma\nabla'^H\Gamma, \quad (\nabla'^H)^\# = \Gamma\nabla^H\Gamma, \quad \text{and} \quad (\mathcal{B}^H)^\# = \mathcal{B}'^H. \tag{5-2}$$

As $\mathcal{B}^H$ and $(\mathcal{B}^H)^\#$ have the same spectrum it then follows that

$$\eta(\mathcal{B}'^H) = \eta(\mathcal{B}^H) \quad \text{and} \quad \mathrm{Det}_{\mathrm{gr},\theta}(\mathcal{B}'^H) = \mathrm{Det}_{\mathrm{gr},\theta}(\mathcal{B}^H). \tag{5-3}$$

Proof of Theorem 5.1. The symmetric bilinear form $\beta_{g,b}$ induces a nondegenerate symmetric bilinear form

$$H^j(M, E') \otimes H^{n-j}(M, E) \to \mathbb{C}, \quad j = 0, \dots, n,$$

and, hence, identifies $H^j(M, E')$ with the dual space of $H^{n-j}(M, E)$. Using the construction of [Huang 2010a, Section 5.1] (with $\tau : \mathbb{C} \to \mathbb{C}$ be the identity map) we obtain a linear isomorphism

$$\alpha : \det H^\bullet(M, E, H) \to \det H^\bullet(M, E', H). \tag{5-4}$$

Lemma 5.3. *Let $E \to M$ be a complex vector bundle over a closed oriented odd dimensional manifold M endowed with a nondegenerate bilinear form b and let ∇ be a flat connection on E. Let ∇' denote the connection dual to ∇ with respect to b. Let H be a closed odd-degree differential form on M. Then*

$$\alpha\left(\rho_{\mathrm{an}}(\nabla^H)\right) = \rho_{\mathrm{an}}(\nabla'^H). \tag{5-5}$$

The proof is that of [Huang 2010a, Theorem 5.3] and will be omitted. (Actually, it is simpler, since $\mathscr{B}^H$ and $\mathscr{B}'^H$ have the same spectrum, so there is no complex conjugation involved.)

For simplicity, we set

$$\tau_{b,\nabla,H,(a,+\infty)} = \det{}'(d_{\bar 0}^{\#} d_{\bar 0}|_{\Omega^{\bar 0}_{(a,+\infty)}(M,E)})^{-1}\cdot(\det{}'(d_{\bar 1}^{\#} d_{\bar 1}|_{\Omega^{\bar 1}_{(a,+\infty)}(M,E)})).$$

Setting $\Delta'^H = (\nabla'^H)^{\#}\nabla'^H + \nabla'^H(\nabla'^H)^{\#}$, we then have

$$\Delta'^H = \Gamma\Delta^H\Gamma.$$

Lemma 5.4. $$\tau_{b,\nabla,H,(a,+\infty)} = \tau_{b,\nabla',H,(a,+\infty)}.$$

Proof. Applying (5-2) and using the fact that

$$\nabla'^H : \Omega^{\bar k}_{(a,+\infty)}(M,E,H)\cap \operatorname{im}(\nabla'^H)^{\#} \to \Omega^{\overline{k+1}}_{(a,+\infty)}(M,E,H)\cap \operatorname{im}\nabla'^H$$

is an isomorphism, we get

$$\begin{aligned}
\tau_{b,\nabla,H,(a,+\infty)} &= \prod_{k=0,1}\det{}'\big((\nabla^H)^{\#}\nabla^H|_{\Omega^{\bar k}_{(a,+\infty)}(M,E,H)}\big)^{(-1)^{k+1}} \\
&= \prod_{k=0,1}\det{}'\big(\Gamma\nabla'^H(\nabla'^H)^{\#}\Gamma|_{\Omega^{\bar k}_{(a,+\infty)}(M,E,H)}\big)^{(-1)^{k+1}} \\
&= \prod_{k=0,1}\det{}'\big(\nabla'^H(\nabla'^H)^{\#}|_{\Omega^{\bar k}_{(a,+\infty)}(M,E,H)}\big)^{(-1)^{k}} \\
&= \prod_{k=0,1}\det{}'\big((\nabla'^H)^{\#}\nabla'^H|_{\Omega^{\bar k}_{(a,+\infty)}(M,E,H)}\big)^{(-1)^{k+1}} = \tau_{b,\nabla',H},
\end{aligned}\tag{5-6}$$

which completes the proof. □

Then for any $h\in\det H^{\bullet}(M,E,H)$, we have

$$\tau_{b,\nabla,H}(h) = \tau_{b,\nabla',H}(\alpha(h)).\tag{5-7}$$

Hence, by (5-5) and (5-7),

$$\tau_{b,\nabla,H}\left(\rho_{\mathrm{an}}(\nabla^H)\right) = \tau_{b,\nabla',H}\left(\rho_{\mathrm{an}}\nabla'^H\right).\tag{5-8}$$

Let

$$\tilde\nabla = \begin{pmatrix}\nabla & 0\\ 0 & \nabla'\end{pmatrix},\qquad \tilde\nabla^H = \begin{pmatrix}\nabla^H & 0\\ 0 & \nabla'^H\end{pmatrix}.$$

Then, for any $a\ge 0$,

$$\tau_{b,\tilde\nabla,H,(a,+\infty)} = \tau_{b,\nabla,H,(a,+\infty)}\cdot\tau_{b,\nabla',H,(a,+\infty)},$$
$$\tau_{b,\tilde\nabla,H}\big(\rho_{\mathrm{an}}(\tilde\nabla^H)\big) = \tau_{b,\nabla,H}\big(\rho_{\mathrm{an}}(\nabla^H)\big)\cdot\tau_{b,\nabla',H}\big(\rho_{\mathrm{an}}(\nabla'^H)\big).$$

Combining the latter equality with (5-8) shows that

$$\tau_{b,\tilde{\nabla},H}\big(\rho_{\mathrm{an}}(\tilde{\nabla}^H)\big) = \tau_{b,\nabla,H}\big(\rho_{\mathrm{an}}(\nabla^H)\big)^2.$$

Hence, (5-1) is equivalent to the equality

$$\tau_{b,\tilde{\nabla},H}\big(\rho_{\mathrm{an}}(\tilde{\nabla}^H)\big) = e^{-4\pi i\left(\eta(\nabla^H)-\operatorname{rank} E\cdot \eta_{\mathrm{trivial}}\right)}. \tag{5-9}$$

By a slight modification of the deformation argument in [Braverman and Kappeler 2007a, Section 4.7] where the untwisted case was treated, we obtain (5-9). This concludes the proof of Theorem 5.1. □

6. On the Cappell–Miller analytic torsion

In this section, we briefly discuss the extension of the Cappell–Miller analytic torsion to the twisted de Rham complexes. Let $\dim M$ be odd.

In the notation above, we have the twisted de Rham complex $\nabla^H : \Omega^{\bar{k}}(M,E) \to \Omega^{\overline{k+1}}(M,E)$ and the chirality operator $\Gamma : \Omega^{\bar{k}}(M,E) \to \Omega^{\overline{k+1}}(M,E)$, $k = 0, 1$. Define

$$d^\flat_{\bar{k}} = \Gamma d_{\bar{k}} \Gamma : \Omega^{\bar{k}}(M,E) \to \Omega^{\overline{k+1}}(M,E).$$

Then consider the non-self-adjoint Laplacian

$$\Delta^\flat_{\bar{k}} = (d_{\bar{k}} + d^\flat_{\bar{k}})^2 : \Omega^{\bar{k}}(M,E) \to \Omega^{\bar{k}}(M,E).$$

For any $a \geq 0$, let $\Omega^{\flat,\bar{k}}_{[0,a]}(M,E)$ ($\Omega^{\flat,\bar{k}}_{(a,+\infty)}(M,E)$) denote the span in $\Omega^{\bar{k}}(M,E)$ of the generalized eigensolutions of $\Delta^\flat_{\bar{k}}$ with generalized eigenvalues with absolute value in $[0,a]$ ($(a,+\infty)$). Then we have the decomposition of the complex

$$(\Omega^\bullet(M,E), d) = (\Omega^{\flat,\bullet}_{[0,a]}(M,E), d) \oplus (\Omega^{\flat,\bullet}_{(a,+\infty)}(M,E), d).$$

The subcomplex $(\Omega^{\flat,\bullet}_{[0,a]}(M,E), d)$ is a $\mathbb{Z}_2$-graded finite dimensional complex. Then we can define the torsion element $\rho^\flat_{\Gamma_{[0,a]}} \otimes \rho^\flat_{\Gamma_{[0,a]}} \in \det H^\bullet(\Omega^{\flat,\bullet}_{[0,a]}(M,E), d)^2 \cong \det H^\bullet(M,E,H)^2$, where $\rho^\flat_{\Gamma_{[0,a]}}$ defined by [Huang 2010a, (2.22)]. On the other hand, for the subcomplex $(\Omega^{\flat,\bullet}_{(a,+\infty)}(M,E), d)$, the following zeta-regularized determinant is well defined (see (3-5)):

$$\det{}'(d^\flat_{\bar{k}} d_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)}) := \exp\big(-\zeta'(0, d^\flat_{\bar{k}} d_{\bar{k}}|_{\operatorname{im} d^\flat_{\bar{k}} \cap \Omega^{\flat,\bar{k}}_{(a,+\infty)}(M,E)})\big). \tag{6-1}$$

Considering the square of the graded determinant defined in [Huang 2010a, (2.38)], for the $\mathbb{Z}_2$-graded finite dimensional complex $\Omega^{\flat,\bullet}_{(a,c]}(M,E)$, $0 \leq a < c < \infty$, we find that

$$\det{}'(d^\flat_{\bar{0}} d_{\bar{0}}|_{\Omega^{\flat,\bar{0}}_{(a,c]}(M,E)}) \cdot \det{}'(d^\flat_{\bar{1}} d_{\bar{1}}|_{\Omega^{\flat,\bar{1}}_{(a,c]}(M,E)})^{-1} = \big(\mathrm{Det}_{\mathrm{gr}}(\mathcal{B}_{\bar{0}}|_{\Omega^{\flat,\bullet}_{(a,c]}(M,E)})\big)^2.$$

Then by [Huang 2010a, Proposition 2.7], we easily get:

Proposition 6.1. *The torsion element defined by*

$$\rho^{\flat}_{\Gamma_{[0,a]}} \otimes \rho^{\flat}_{\Gamma_{[0,a]}} \cdot \prod_{k=0,1} \left(\det{}'(d^{\flat}_{\bar{k}} d_{\bar{k}}|_{\Omega^{\bar{k}}_{(a,+\infty)}(M,E)})\right)^{(-1)^k} \in \det H^{\bullet}(M, E, H)^2 \tag{6-2}$$

is independent of the choice of $a \geq 0$.

Definition 6.2. The torsion element in $\det H^{\bullet}(M, E, H)^2$ defined by (6-2) is called the twisted Cappell–Miller analytic torsion for the twisted de Rham complex and is denoted by $\tau_{\nabla,H}$.

Next we study the torsion $\tau_{\nabla,H}$ under metric and flux deformations. Since the methods are the same as the cases in the twisted refined analytic torsion [Huang 2010a] and the twisted Burghelea–Haller analytic torsion above, we only briefly outline the results.

Theorem 6.3 (Metric independence). *Let M be a closed odd dimensional manifold, E be a complex vector bundle over M with flat connection ∇ and H be a closed odd-degree differential form on M. Then the torsion $\tau_{\nabla,H}$ is independent of the choice of the Riemannian metric g.*

Proof. By the definition of $\tau_{\nabla,H}$ and the observation on the determinants, this theorem follows easily from Proposition 2.4 and Equations (3.18) and (4.14) of [Huang 2010a]. □

Theorem 6.4 (Flux representative independence). *Let M be a closed odd dimensional manifold and E be a complex vector bundle over M with flat connection ∇. Suppose H and H' are closed differential forms on M of odd degrees representing the same de Rham cohomology class, and let B be an even form so that $H' = H - dB$. Then we have $\tau_{\nabla,H'} = \det(\varepsilon_B)\tau_{\nabla,H}$.*

Proof. From the above observation, this follows easily from Lemmas 4.6 and 4.7 of [Huang 2010a]. □

From the definition in (6-2), we see that the twisted Cappell–Miller analytic torsion is closely related to the twisted refined analytic torsion $\rho_{\text{an}}(\nabla^H)$. Explicitly:

Theorem 6.5 (compare [Huang 2010b, Theorem 4.5]). *In* $\det H^{\bullet}(M, E, H)^2$,

$$\rho_{\text{an}}(\nabla^H) \otimes \rho_{\text{an}}(\nabla^H) = \tau_{\nabla,H} e^{-2\pi i\left(\eta(\nabla^H) - \operatorname{rank} E \cdot \eta_{\text{trivial}}\right)}. \tag{6-3}$$

Proof. The twisted refined analytic torsion [Huang 2010a, (4.15)] is defined by

$$\rho_{\text{an}}(\nabla^H) = \operatorname{Det}_{\text{gr},\theta}(\mathcal{B}^H_{\bar{0},(\lambda,\infty)}) \cdot \rho_{\Gamma_{[0,\lambda]}} \cdot e^{i\pi(\operatorname{rank} E)\eta_{\text{trivial}}}.$$

By [Huang 2010a, (5.31)], we have

$$
\begin{aligned}
(6\text{-}4)\quad & \rho_{\mathrm{an}}(\nabla^H)\otimes\rho_{\mathrm{an}}(\nabla^H)\\
&= \rho_{\Gamma_{[0,\lambda]}}\otimes\rho_{\Gamma_{[0,\lambda]}}\cdot\exp(2\xi_\lambda(\nabla^H,g^M,\theta))\\
&\qquad\cdot\exp\Bigl(-2i\pi\eta_\lambda(\nabla^H)-i\pi\sum_{k=0,1}(-1)^k d^-_{\bar{k},\lambda}+2i\pi(\operatorname{rank}E)\eta_{\mathrm{trivial}}\Bigr),
\end{aligned}
$$

where $\eta_\lambda(\nabla^H)$, $\xi_\lambda(\nabla^H,g^M,\theta)$, and $d^-_{\bar{k},\lambda}$ are defined in equations (3.17), (3.18), and (3.19) of [Huang 2010a]. By (6-2) and (6-4), we find that

$$
\begin{aligned}
(6\text{-}5)\quad & \rho_{\mathrm{an}}(\nabla^H)\otimes\rho_{\mathrm{an}}(\nabla^H)\\
&= \tau_{\nabla,H}\ \exp(-2i\pi\eta_\lambda(\nabla^H)-i\pi\sum_{k=0,1}(-1)^k d^-_{\bar{k},\lambda}+2i\pi(\operatorname{rank}E)\eta_{\mathrm{trivial}}).
\end{aligned}
$$

From [Huang 2010a, (5.28)], we get

$$
(6\text{-}6)\qquad 2\eta_\lambda(\nabla^H)+\sum_{k=0,1}(-1)^k d^-_{\bar{k},\lambda}\equiv 2\eta(\nabla^H)\ \mathrm{mod}\ 2\mathbb{Z}.
$$

Then (6-5) and (6-6) imply (6-3). □

Theorem 5.1 and Theorem 6.5 give the relation between the twisted Burghelea–Haller analytic torsion $\tau_{b,\nabla,H}$ and the twisted Cappell–Miller analytic torsion $\tau_{\nabla,H}$ if there is a nondegenerate symmetric bilinear form the bundle E.

Corollary 6.6. *If there is a nondegenerate symmetric bilinear form on E and* $\dim M$ *is odd*, *we have*

$$
\tau_{b,\nabla,H}(\tau_{\nabla,H})=\pm 1.
$$

Remark 6.7. Almost at the same time of the preprint [Su 2010] of this paper, Huang [2010b] defined and studied the twisted Cappell–Miller torsion both for holomorphic and analytic cases.

References

[Bismut and Zhang 1992] J.-M. Bismut and W. Zhang, *An extension of a theorem by Cheeger and Müller*, Astérisque **205**, 1992. With an appendix by F. Laudenbach. MR 93j:58138 Zbl 0781.58039

[Braverman and Kappeler 2007a] M. Braverman and T. Kappeler, "Comparison of the refined analytic and the Burghelea–Haller torsions", *Ann. Inst. Fourier* (*Grenoble*) **57**:7 (2007), 2361–2387. Festival Yves Colin de Verdière. MR 2009f:58051 Zbl 1147.58033

[Braverman and Kappeler 2007b] M. Braverman and T. Kappeler, "Ray–Singer type theorem for the refined analytic torsion", *J. Funct. Anal.* **243**:1 (2007), 232–256. MR 2008e:58042 Zbl 1122.58018

[Braverman and Kappeler 2007c] M. Braverman and T. Kappeler, "Refined analytic torsion as an element of the determinant line", *Geom. Topol.* **11** (2007), 139–213. MR 2008a:58031 Zbl 1135.58014

[Braverman and Kappeler 2008] M. Braverman and T. Kappeler, "Refined analytic torsion", *J. Differential Geom.* **78**:2 (2008), 193–267. MR 2009a:58041 Zbl 1147.58034

[Burghelea and Haller 2007] D. Burghelea and S. Haller, “Complex-valued Ray–Singer torsion”, *J. Funct. Anal.* **248**:1 (2007), 27–78. MR 2008b:58035 Zbl 1131.58020

[Burghelea and Haller 2008] D. Burghelea and S. Haller, “Torsion, as a function on the space of representations”, pp. 41–66 in *C^*-algebras and elliptic theory II*, edited by D. Burghelea et al., Trends Math., Birkhäuser, Basel, 2008. MR 2009i:58045 Zbl 1147.58035

[Burghelea and Haller 2010] D. Burghelea and S. Haller, “Complex valued Ray–Singer torsion II”, *Math. Nachr.* **283**:10 (2010), 1372–1402. MR 2744135 Zbl 05815198

[Cappell and Miller 2010] S. E. Cappell and E. Y. Miller, “Complex-valued analytic torsion for flat bundles and for holomorphic bundles with (1, 1) connections”, *Comm. Pure Appl. Math.* **63**:2 (2010), 133–202. MR 2011a:58061 Zbl 1201.58021

[Cheeger 1979] J. Cheeger, “Analytic torsion and the heat equation”, *Ann. of Math.* (2) **109**:2 (1979), 259–322. MR 80j:58065a Zbl 0412.58026

[Farber and Turaev 2000] M. Farber and V. Turaev, “Poincaré–Reidemeister metric, Euler structures, and torsion”, *J. Reine Angew. Math.* **520** (2000), 195–225. MR 2001m:58069 Zbl 0938.57020

[Grubb and Seeley 1995] G. Grubb and R. T. Seeley, “Weakly parametric pseudodifferential operators and Atiyah–Patodi–Singer boundary problems”, *Invent. Math.* **121**:3 (1995), 481–529. MR 96k:58216 Zbl 0851.58043

[Huang 2010a] R.-T. Huang, “Refined analytic torsion for twisted de Rham complexes”, preprint, 2010. arXiv 1001.0654

[Huang 2010b] R.-T. Huang, “Twisted Cappell–Miller holomorphic and analytic torsions”, preprint, 2010. arXiv 1001.3917

[Mathai and Wu 2008] V. Mathai and S. Wu, “Analytic torsion for twisted de Rham complexes”, preprint, 2008. arXiv 0810.4204

[Mathai and Wu 2010a] V. Mathai and S. Wu, “Analytic torsion of $\mathbb{Z}_2$-graded elliptic complexes”, preprint, 2010. arXiv 1001.3212

[Mathai and Wu 2010b] V. Mathai and S. Wu, “Twisted analytic torsion”, *Sci. China Math.* **53**:3 (2010), 555–563. MR 2608312 Zbl 1202.58019

[Milnor 1966] J. Milnor, “Whitehead torsion”, *Bull. Amer. Math. Soc.* **72** (1966), 358–426. MR 33 #4922 Zbl 0147.23104

[Müller 1978] W. Müller, “Analytic torsion and R-torsion of Riemannian manifolds”, *Adv. in Math.* **28**:3 (1978), 233–305. MR 80j:58065b Zbl 0395.57011

[Müller 1993] W. Müller, “Analytic torsion and R-torsion for unimodular representations”, *J. Amer. Math. Soc.* **6**:3 (1993), 721–753. MR 93m:58119 Zbl 0789.58071

[Quillen 1985] D. Quillen, “Determinants of Cauchy–Riemann operators on Riemann surfaces”, *Funct. Anal. Appl.* **14**:1 (1985), 31–34. MR 86g:32035 Zbl 0603.32016

[Ray and Singer 1971] D. B. Ray and I. M. Singer, “R-torsion and the Laplacian on Riemannian manifolds”, *Advances in Math.* **7** (1971), 145–210. MR 45 #4447 Zbl 0239.58014

[Su 2010] G. Su, “Burghelea–Haller analytic torsion for twisted de Rham complexes”, preprint, 2010. arXiv 1001.2728

[Su and Zhang 2008] G. Su and W. Zhang, “A Cheeger–Müller theorem for symmetric bilinear torsions”, *Chin. Ann. Math. Ser. B* **29**:4 (2008), 385–424. MR 2009j:58053 Zbl 1152.58028

[Turaev 1989] V. G. Turaev, “Euler structures, nonsingular vector fields, and Reidemeister-type torsions”, *Izv. Akad. Nauk SSSR Ser. Mat.* **53**:3 (1989), 607–643, 672. In Russian. MR 90m:57021 Zbl 0692.57015

[Witten 1982] E. Witten, "Supersymmetry and Morse theory", *J. Differential Geom.* **17**:4 (1982), 661–692 (1983). MR 84b:58111 Zbl 0499.53056

[Wodzicki 1984] M. Wodzicki, "Local invariants of spectral asymmetry", *Invent. Math.* **75**:1 (1984), 143–177. MR 85g:58089 Zbl 0538.58038

Received January 25, 2010. Revised September 23, 2010.

GUANGXIANG SU
CHERN INSTITUTE OF MATHEMATICS AND LPMC
NANKAI UNIVERSITY
TIANJIN 300071
CHINA

guangxiangsu@nankai.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

$K(n)$-LOCALIZATION OF THE $K(n+1)$-LOCAL E_{n+1}-ADAMS SPECTRAL SEQUENCES

TAKESHI TORII

We construct a spectral sequence converging to the homotopy set of maps from a spectrum to the $K(n)$-localization of the $K(n+1)$-local sphere. We also construct a map of spectral sequences from the $K(n)$-local E_n-Adams spectral sequence to the preceding one. Then we compare the map on E_2-terms with a map induced by the inflation maps of continuous cohomology groups for Morava stabilizer groups. As an application we show that ζ_n in $\pi_{-1}(L_{K(n)}S^0)$ represented by the reduced norm map in the $K(n)$-local E_n-Adams spectral sequence has a nontrivial image under the map $\pi_*(L_{K(n)}S^0) \to \pi_*(L_{K(n)}L_{K(n+1)}S^0)$.

1. Introduction

The motivation of this note is toward understanding the relationship between the $K(n)$-local category and the $K(n+1)$-local category. For each prime number p, the stable homotopy category of p-local spectra has a filtration of full subcategories corresponding to the height filtration of the moduli space of formal groups [Morava 1985]. The n-th associated graded part of the filtration is equivalent to the $K(n)$-local category, that is, the Bousfield localization of the stable homotopy category with respect to the n-th Morava K-theory spectrum $K(n)$ [Hovey and Strickland 1999]. So it can be considered that the stable homotopy category of p-local spectra is built up from the $K(n)$-local categories for various n. In fact, the chromatic convergence theorem [Ravenel 1992] says that a p-local finite spectrum X is homotopy equivalent to the homotopy inverse limit of the chromatic tower $\cdots \to L_{n+1}X \to L_nX \to \cdots \to L_0X$, where L_n is the Bousfield localization functor with respect to the wedge of Morava K-theories $K(0)\vee K(1)\vee\cdots\vee K(n)$. This means that a p-local finite spectrum X can be recovered from $\{L_nX\}_{n\geq 0}$ through the chromatic tower. Furthermore, if the chromatic splitting conjecture is true, then it implies that the p-completion of a finite spectrum X is a direct summand of the product $\prod_n L_{K(n)}X$ [Hovey 1995]. This means that it is not necessary to reconstruct the tower but it is sufficient to know all $L_{K(n)}X$ to obtain

MSC2000: primary 55T25, 55P42; secondary 55Q51, 55N22, 55N20.
Keywords: Adams spectral sequence, $K(n)$-localization, Morava E-theory.

some information of X. Since the chromatic splitting conjecture is concerned with the relationship among various chromatic pieces, it is important to understand the relationship between the $K(n)$-local category and the $K(n+1)$-local category.

Let E_n be the n-th Morava E-theory spectrum. The $K(n)$-local E_n-Adams spectral sequence $L_{K(n)}E_r^{s,t}(W)$ is a natural spectral sequence for any spectrum W,

$$L_{K(n)}E_2^{s,t}(W) = H_c^s(G_n;\, E_n^t(W)) \Longrightarrow [W,\, L_{K(n)}S^0]^{s+t},$$

which converges to $[W,\, L_{K(n)}S^0]^*$ strongly and conditionally; see [Devinatz and Hopkins 2004, Appendix A]. On the E_2-term, G_n is the n-th extended Morava stabilizer group, and $H_c^s(G_n;\, E_n^t(W))$ is a continuous cohomology group for the profinite group G_n with coefficients in the profinite module $E_n^t(W)$.

We construct a natural spectral sequence converging to $[W,\, L_{K(n)}L_{K(n+1)}S^0]^*$ by applying the $K(n)$-localization functor to the $K(n+1)$-local E_{n+1}-Adams resolution of $L_{K(n+1)}S^0$. Let $\mathbb{A} = L_{K(n)}E_{n+1}$ be the $K(n)$-localization of the $(n+1)$-st Morava E-theory E_{n+1}. We identify the E_2-term as a cohomology group based on the continuous cochain complex for G_{n+1} with coefficients in the topological module $\mathbb{A}^*(W)$. We call this spectral sequence the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence for W.

Theorem 4.7. *For any spectrum W, there is a natural spectral sequence*

$$L_{K(n)}L_{K(n+1)}E_2^{s,t}(W) = H_c^s(G_{n+1};\, \mathbb{A}^t(W)) \Longrightarrow [W,\, L_{K(n)}L_{K(n+1)}S^0]^{s+t},$$

which converges strongly and conditionally.

By the $K(n)$-localization of the $K(n+1)$-localization map $S^0 \to L_{K(n+1)}S^0$, we obtain a map $L_{K(n)}S^0 \to L_{K(n)}L_{K(n+1)}S^0$, which induces a map

$$[W,\, L_{K(n)}S^0]^* \to [W,\, L_{K(n)}L_{K(n+1)}S^0]^*$$

for any spectrum W. We construct in Theorem 6.2 a natural map of spectral sequences

$$\varphi_r(W) : L_{K(n)}E_r^{s,t}(W) \longrightarrow L_{K(n)}L_{K(n+1)}E_r^{s,t}(W),$$

which converges to the map $[W,\, L_{K(n)}S^0]^{s+t} \to [W,\, L_{K(n)}L_{K(n+1)}S^0]^{s+t}$. Furthermore, we give an interpretation of the map on E_2-terms. We construct a natural homomorphism

$$\theta(W) : H_c^*(G_n;\, E_n^*(W)) \longrightarrow H_c^*(G_{n+1};\, \mathbb{A}^*(W)),$$

which is obtained from some kind of inflation maps (see (7-1)).

Theorem 7.6. *The map $\varphi_2(W)$ coincides with $\theta(W)$ for any spectrum W.*

By the Hopkins–Miller theorem [Devinatz and Hopkins 2004, Theorem 6], we know that there is a nontrivial element $\zeta_n \in \pi_{-1}(L_{K(n)}S^0)$ which is represented by the reduced norm map of G_n in the E_2-term of the $K(n)$-local E_n-Adams spectral sequence. Let ω_n be the image of ζ_n under the map

$$\pi_*(L_{K(n)}S^0) \to \pi_*(L_{K(n)}L_{K(n+1)}S^0).$$

As an application of our results, we show the following theorem.

Theorem 8.1. *The image ω_n is nontrivial.*

The organization of the remaining sections is as follows: In Section 2 we review the results in [Torii 2010a]. We recall the construction of a commutative ring spectrum $\mathbb{B}$ which is an extension of both of E_n and E_{n+1}, and the action of the group $\mathbb{G} = G_n \times_\Gamma G_{n+1}$ on $\mathbb{B}$. In Section 3 we introduce a topology for $\mathbb{A}^*$-modules of certain type, and study modules of continuous maps from a topological space to such a topological $\mathbb{A}^*$-module. In particular, we show that the functor $\mathrm{Map}_c(T, \mathbb{A}^*(-))$ is a generalized cohomology theory for any compact space T. In Section 4 we construct the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence by applying the $K(n)$-localization functor to the $K(n+1)$-local E_{n+1}-Adams resolution of $L_{K(n+1)}S^0$, and prove Theorem 4.7. In Section 5 we define a cohomology of $\mathbb{G}$ with coefficients in $\mathbb{B}^*(W)$ for the purpose of connecting the cohomology of G_n and that of G_{n+1}. Then we show that the inflation map from the cohomology of G_{n+1} with coefficients in $\mathbb{A}^*(W)$ to the cohomology of $\mathbb{G}$ with coefficients in $\mathbb{B}^*(W)$ is an isomorphism for any spectrum W. In Section 6 we construct a map of spectral sequences from the $K(n)$-local E_n-Adams spectral sequence to the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence. In Section 7 we construct a homomorphism $\theta(W)$ from the cohomology group of G_n with coefficients in $E_n^*(W)$ to the cohomology group of G_{n+1} with coefficients in $\mathbb{A}^*(W)$ by using the cohomology of $\mathbb{G}$ with coefficients in $\mathbb{B}^*(W)$ constructed in Section 5. Then we identify this homomorphism with the map of spectral sequences on E_2-terms, and prove Theorem 7.6. In Section 8 we prove Theorem 8.1 as an application of the results obtained earlier.

2. The ring spectrum $\mathbb{B}$

In this section we review the results in [Torii 2010a]. We recall the construction of a commutative ring spectrum $\mathbb{B}$ and two ring spectrum maps $\Theta : E_{n+1} \to \mathbb{B}$ and $I : E_n \to \mathbb{B}$. Furthermore, we recall that the action of a profinite group $\mathbb{G}$ on $\mathbb{B}$ and the equivariance of Θ and I under the actions of $\mathbb{G}$.

Let p be a prime number, and let n be a positive integer. We fix a finite field $\boldsymbol{F}$ which contains the finite fields $\mathbb{F}_{p^n}$ and $\mathbb{F}_{p^{n+1}}$. Note that the minimal field satisfying the condition is $\mathbb{F}_{p^n} \otimes \mathbb{F}_{p^{n+1}} \cong \mathbb{F}_{p^{n^2+n}}$. We denote by W the ring of Witt vectors with coefficients in $\boldsymbol{F}$. We define variants of the n-th Morava E-theory spectrum E_n and the $(n+1)$-st Morava E-theory spectrum E_{n+1} such that the coefficient rings are given by

$$E_n^* = W[\![w_1, \ldots, w_{n-1}]\!][w^{\pm 1}], \quad E_{n+1}^* = W[\![u_1, \ldots, u_n]\!][u^{\pm 1}].$$

There is an associated degree 0 formal group law F_n over E_n^0 since E_n is complex oriented and even-periodic. The formal group law F_n is a universal deformation of the Honda formal group law H_n of height n over $\boldsymbol{F}$. Note that we can take F_n as a p-typical formal group law. The Morava stabilizer group S_n is defined to be the group of automorphisms of H_n over $\boldsymbol{F}$. Then the extended Morava stabilizer group G_n is defined to be the semi-direct product $G_n = \Gamma \ltimes S_n$, where $\Gamma = \mathrm{Gal}(\boldsymbol{F}/\mathbb{F}_p)$ is the Galois group of $\boldsymbol{F}$ over the prime field $\mathbb{F}_p$. We can identify G_n with the group of automorphisms of the ring spectrum E_n in the stable homotopy category. Then $g = (\gamma, s) \in \Gamma \ltimes S_n = G_n$ induces a ring homomorphism $g^* : E_n^* \to E_n^*$. We denote by F_n^g the formal group law obtained from F_n by the coefficient change along g^*. Then there is a unique isomorphism $t(g) : F_n \to F_n^g$ of formal group laws which is a lifting of the isomorphism $s : H_n \to H_n^\gamma = H_n$. There are projections $G_n \to \Gamma$ and $G_{n+1} \to \Gamma$. We define a profinite group $\mathbb{G}$ to be the fiber product of G_n and G_{n+1} over Γ

$$\mathbb{G} = G_n \times_\Gamma G_{n+1}.$$

Let $K(n)$ be the n-th Morava K-theory spectrum at p. We denote by $\mathbb{A}$ the commutative ring spectrum $L_{K(n)} E_{n+1}$, the Bousfield localization of E_{n+1} with respect to $K(n)$. The coefficient ring of $\mathbb{A}$ is given by the following Lemma.

Lemma 2.1. *The coefficient ring $\mathbb{A}^*$ is isomorphic to $(E_{n+1}^*[u_n^{-1}])^\wedge_{I_n}$, the completion of the localization $E_{n+1}^*[u_n^{-1}]$ at the ideal $I_n = (p, u_1, \ldots, u_{n-1})$. Hence $\mathbb{A}^*$ is a graded complete Noetherian regular local ring isomorphic to*

$$(W((u_n)))^\wedge_p[\![u_1, \ldots, u_{n-1}]\!][u^{\pm 1}]$$

with residue field $\boldsymbol{F}((u_n))[u^{\pm 1}]$.

Proof. There is a tower $\{M(J)\}_J$ of generalized Moore spectra of height n as in [Hovey and Strickland 1999, Proposition 4.2]. If $J = (p^{a_0}, v_1^{a_1}, \ldots, v_{n-1}^{a_{n-1}})$, then $(E_{n+1} \wedge M(J))^* = E_{n+1}^*/(p^{a_0}, u_1^{a_1}, \ldots, u_{n-1}^{a_{n-1}})$ since $v_i = u_i u^{p^i - 1}$ for $i = 1, \ldots, n-1$. We set $X^\wedge_{I_n} = \operatorname{holim}_{\longleftarrow J} X \wedge M(J)$ for a spectrum X. Since E_{n+1} is Landweber exact of height $(n+1)$, it satisfies the telescope conjecture at n in the sense of [Hovey 1997, Definition 1.5.2]. Then $L_{K(n)} E_{n+1} \simeq (E_{n+1}[v^{-1}])^\wedge_{I_n}$ by [Hovey 1997, Theorem 1.5.4], where v is a generalized v_n-element in E_{n+1}^* in the sense of [Hovey 1997, Definition 1.2.2]. We can take $v_n = u_n u^{p^n - 1} \in \pi_{2p^n - 2} E_{n+1}$ as a generalized v_n-element. Since the sequence $p^{a_0}, u_1^{a_1}, \ldots, u_{n-1}^{a_{n-1}}$ is regular in $E_{n+1}^*[v_n^{-1}] = E_{n+1}^*[u_n^{-1}]$, $(E_{n+1}[v_n^{-1}] \wedge M(J))^* = E_{n+1}^*[u_n^{-1}]/(p^{a_0}, u_1^{a_1}, \ldots, u_{n-1}^{a_{n-1}})$ if $J = (p^{a_0}, v_1^{a_1}, \ldots, v_{n-1}^{a_{n-1}})$. Then we see that $\mathbb{A}^* = (L_{K(n)} E_{n+1})^*$ is the completion of $E_{n+1}^*[u_n^{-1}]$ at the ideal $I_n = (p, u_1, \ldots, u_{n-1})$: $\mathbb{A}^* \cong (E_{n+1}^*[u_n^{\pm 1}])^\wedge_{I_n}$. Since the sequence $p, u_1, \dots, u_{n-1}$ is regular in $E^*_{n+1}[u_n^{\pm 1}]$, and it generates a maximal ideal, $\mathbb{A}^*$ is a graded regular local ring with maximal ideal generated by $p, u_1, \dots, u_{n-1}$ and residue field $\boldsymbol{F}((u_n))[u^{\pm 1}]$.

The obvious ring homomorphism $W[\![u_n]\!] \to \mathbb{A}^*$ extends to $(W((u_n)))^\wedge_p \to \mathbb{A}^*$, since u_n is a unit in $\mathbb{A}^*$, and $\mathbb{A}^*$ is p-complete. Furthermore, since $\mathbb{A}^*$ is I_n-adically complete, the obvious ring homomorphism $(W((u_n)))^\wedge_p[u_1, \dots, u_{n-1}][u^{\pm 1}] \to \mathbb{A}^*$ extends to $(W((u_n)))^\wedge_p[\![u_1, \dots, u_{n-1}]\!][u^{\pm 1}] \to \mathbb{A}^*$. The ring

$$(W((u_n)))^\wedge_p[\![u_1, \dots, u_{n-1}]\!][u^{\pm 1}]$$

is a graded complete regular local ring with maximal ideal generated by $p, u_1, \dots, u_{n-1}$ and residue field $\boldsymbol{F}((u_n))[u^{\pm 1}]$. Since the ring homomorphism

$$(W((u_n)))^\wedge_p[\![u_1, \dots, u_{n-1}]\!][u^{\pm 1}] \to \mathbb{A}^*$$

is continuous, and it induces an isomorphism on the associated graded rings, we obtain an isomorphism between $\mathbb{A}^*$ and $(W((u_n)))^\wedge_p[\![u_1, \dots, u_{n-1}]\!][u^{\pm 1}]$. □

Since a complete local ring is Henselian, $\mathbb{A}^*$ is a Henselian ring by Lemma 2.1.

Lemma 2.2 [Milne 1980, Proposition I.4.4]. *Let R be a Henselian ring with residue field k. Then the functor $S \mapsto S \otimes_R k$ induces an equivalence between the category of finite étale R-algebras and the category of finite étale k-algebras.*

Let $\bar{F}_{n+1}$ be the formal group law over $\boldsymbol{F}((u_n))$ obtained from F_{n+1} by the reduction $E^0_{n+1} \to \boldsymbol{F}((u_n))$. Then the height of $\bar{F}_{n+1}$ is n. Since the isomorphism classes of formal group laws over a separably closed field are classified by their height, there is an isomorphism between $\bar{F}_{n+1}$ and the height n Honda formal group law H_n over the separable closure $\boldsymbol{F}((u_n))^{\mathrm{sep}}$. In [Torii 2003, §2.3] we have constructed an extension field L of $\boldsymbol{F}((u_n))$, where L is the minimal extension such that there is an isomorphism between $\bar{F}_{n+1}$ and H_n. The extension L is Galois over $\boldsymbol{F}((u_n))$ with Galois group isomorphic to S_n. There is a sequence of finite Galois extensions of $\boldsymbol{F}((u_n))$

$$\boldsymbol{F}((u_n)) = L(-1) \to L(0) \to L(1) \to \cdots \tag{2-1}$$

such that $L = \bigcup_i L(i)$. We denote by $S_n(i)$ the Galois group for $\boldsymbol{F}((u_n)) \to L(i)$. Then $S_n(i)$ is a finite quotient group of S_n of order $(p^n-1)p^{ni}$, and $S_n = \varprojlim_i S_n(i)$. The action of G_{n+1} on E^0_{n+1} induces an action on the residue field $\boldsymbol{F}((u_n))$ of $\mathbb{A}^0$. By [Torii 2003, §2.4], there is an action of $\mathbb{G}$ on L, which is an extension of the action of G_{n+1} on $\boldsymbol{F}((u_n))$ and the action of S_n on L as Galois group. Note that $L(i)$ is stable under the action of $\mathbb{G}$ for all i.

By Lemma 2.2, the sequence of Galois extensions (2-1) induces a sequence of graded commutative rings

$$\mathbb{A}^* = \mathbb{B}(-1)^* \to \mathbb{B}(0)^* \to \mathbb{B}(1)^* \to \cdots .$$

The ring $\mathbb{B}(i)^*$ is an even-periodic graded complete Noetherian regular local ring with residue field $L(i)[u^{\pm 1}]$. Furthermore, $\mathbb{A}^* \to \mathbb{B}(i)^*$ is a Galois extension of graded commutative rings with Galois group $S_n(i)$ in the sense of [Chase et al. 1965; Greither 1992]. Let $\mathbb{B}(\infty)^*$ be the direct limit of the sequence: $\mathbb{B}(\infty)^* = \mathrm{colim}_{\overrightarrow{i}}\, \mathbb{B}(i)^*$. Then we define a graded commutative ring $\mathbb{B}^*$ to be the completion of $\mathbb{B}(\infty)^*$ at the ideal $I_n = (p, u_1, \ldots, u_{n-1})$

$$\mathbb{B}^* = (\mathbb{B}(\infty)^*)^{\wedge}_{I_n}.$$

By Lemma 2.2, there is a unique lifting of the action of $\mathbb{G}$ on $\mathbb{B}^*$ and $\mathbb{B}(i)^*$ for $0 \le i \le \infty$ compatible with canonical inclusions.

By the $\mathbb{A}^*$-algebra structures, we can regard $\mathbb{B}^*$ and $\mathbb{B}(i)^*$ for $0 \le i \le \infty$ as Landweber exact even-periodic graded commutative rings. We denote the corresponding commutative ring spectra by $\mathbb{B}$ and $\mathbb{B}(i)$ for $0 \le i \le \infty$, respectively. Hence we obtain a sequence of commutative ring spectra

$$\mathbb{A} = \mathbb{B}(-1) \to \mathbb{B}(0) \to \mathbb{B}(1) \to \cdots .$$

Then we have $\mathbb{B}(\infty) = \mathrm{hocolim}_{\overrightarrow{i}}\, \mathbb{B}(i)$ and $\mathbb{B} = L_{K(n)}\mathbb{B}(\infty)$. We define a ring spectrum map $\Theta : E_{n+1} \to \mathbb{B}$ to be the composition

$$\Theta : E_{n+1} \longrightarrow L_{K(n)}E_{n+1} = \mathbb{A} \longrightarrow \mathbb{B}.$$

By [Torii 2003, §2.3], the formal group law induced by the ring homomorphism $E_n^0 \to \boldsymbol{F} \hookrightarrow L$ is isomorphic to the formal group law induced by the ring homomorphism $E_{n+1}^0 \to \boldsymbol{F}((u_n)) \hookrightarrow L$. By the universality of the formal group law F_n associated with E_n, there exists a ring homomorphism $E_n^* \to \mathbb{B}^*$ and an isomorphism Φ between the formal group laws F_n and F_{n+1} over $\mathbb{B}^0$

$$\Phi : F_{n+1} \xrightarrow{\cong} F_n.$$

Note that $\mathbb{B}^0$ is the minimal extension ring of both of E_n^0 and E_{n+1}^0 such that there exists an isomorphism between F_n and F_{n+1}. Since E_n and $\mathbb{B}$ are even-periodic Landweber exact commutative ring spectra, the ring homomorphism $E_n^* \to \mathbb{B}^*$ extends to a ring spectrum map

$$I : E_n \longrightarrow \mathbb{B}.$$

By the projection $\mathbb{G} \to G_n$, we can consider that $\mathbb{G}$ acts on E_n as automorphisms of commutative ring spectrum in the stable homotopy category. Also, by the projection $\mathbb{G} \to G_{n+1}$, we can consider that $\mathbb{G}$ acts on E_{n+1} as automorphisms of commutative ring spectrum.

Proposition 2.3 [Torii 2010a, §4]. *The profinite group $\mathbb{G}$ acts on the commutative ring spectrum $\mathbb{B}$ in the stable homotopy category. The ring spectrum maps I : $E_n \to \mathbb{B}$ and $\Theta : E_{n+1} \to \mathbb{B}$ are equivariant with respect to the actions of $\mathbb{G}$.*

Remark 2.4 [Torii 2010b]. The ring spectrum $\mathbb{B}$ supports a commutative S-algebra structure and the group $\mathbb{G}$ acts on $\mathbb{B}$ in the category of commutative S-algebras. Let $T = L_{K(n)}S^0 \otimes_{\mathbb{Z}_p} W$ be the commutative S-algebra obtained from $L_{K(n)}S^0$ by adjoining a primitive $(p^m - 1)$-st root of unity, where m is the dimension of $\boldsymbol{F}$ over $\mathbb{F}_p$. Then there is an equivalence $\mathbb{B} \simeq L_{K(n)}(E_n \overset{\wedge}{\underset{T}{}} \mathbb{A})$ of commutative S-algebras. In particular, when $\boldsymbol{F} = \mathbb{F}_{p^{n^2+n}}$, there is an equivalence $\mathbb{B} \simeq L_{K(n)}(E'_n \wedge E'_{n+1})$ of commutative S-algebras, where E'_n and E'_{n+1} are the standard Morava E-theory spectra so that $\pi_0 E'_n/I_n = \mathbb{F}_{p^n}$ and $\pi_0 E'_{n+1}/I_{n+1} = \mathbb{F}_{p^{n+1}}$. In this case

$$\mathrm{Gal}(\boldsymbol{F}/\mathbb{F}_p) \cong \mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) \times \mathrm{Gal}(\mathbb{F}_{p^{n+1}}/\mathbb{F}_p) \quad \text{and} \quad \mathbb{G} \cong G'_n \times G'_{n+1},$$

where $G'_n = \mathrm{Gal}(\mathbb{F}_{p^n}/\mathbb{F}_p) \ltimes S_n$ and $G'_{n+1} = \mathrm{Gal}(\mathbb{F}_{p^{n+1}}/\mathbb{F}_p) \ltimes S_{n+1}$ are the standard extended Morava stabilizer groups.

3. Mapping space $\mathrm{Map}_c(T, \mathbb{A}^*(W))$

To interpret the E_2-term of the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence which will be constructed in Section 4 below as a cohomology group of G_{n+1}, we need to give an appropriate topology for $\mathbb{A}^*$-cohomology groups. In this section we introduce a topology for $\mathbb{A}^*$-modules of certain type, and study modules of continuous maps from a topological space to such an $\mathbb{A}^*$-module.

For a topological space T, and a topological module M, denote by $\mathrm{Map}_c(T, M)$ the module of continuous maps from T to M. Recall the fact that a surjection between profinite groups has a continuous section of topological spaces [Serre 1994, Proposition I.1.2.1]. This implies that $\mathrm{Map}_c(T, -)$ gives an exact functor from the category of profinite modules to that of abelian groups. The coefficient ring E^*_{n+1} is a graded complete Noetherian local ring with maximal ideal $I_{n+1} = (p, u_1, \dots, u_n)$. Since E^*_{n+1}/I^r_{n+1} is a graded finite ring for each r, E^*_{n+1} is a graded profinite ring. Let N be a finitely generated E^*_{n+1}-module. Then N is a graded profinite abelian group. In this case there is an easy description for $\mathrm{Map}_c(T, N)$ as follows.

Lemma 3.1. *If N is a finitely generated E^*_{n+1}-module, there is a natural isomorphism*

$$\mathrm{Map}_c(T, N) \cong \mathrm{Map}_c(T, E^*_{n+1}) \otimes_{E^*_{n+1}} N.$$

Proof. Since N is finitely generated, there is an exact sequence of profinite modules $N^1 \to N^0 \to N \to 0$, where N^i is finitely generated free for $i = 0, 1$. This induces two exact sequences $\mathrm{Map}_c(T, N^1) \to \mathrm{Map}_c(T, N^0) \to \mathrm{Map}_c(T, N) \to 0$ and $\mathrm{Map}_c(T, E^*_{n+1}) \otimes N^1 \to \mathrm{Map}_c(T, E^*_{n+1}) \otimes N^0 \to \mathrm{Map}_c(T, E^*_{n+1}) \otimes N \to 0$. Since N^i is finitely generated free, we have $\mathrm{Map}_c(T, N^i) \cong \mathrm{Map}_c(T, E^*_{n+1}) \otimes N^i$ for $i = 0, 1$. Hence we obtain that $\mathrm{Map}_c(T, N) \cong \mathrm{Map}_c(T, E^*_{n+1}) \otimes N$. □

Corollary 3.2. *For an ideal I of E^*_{n+1} and a finitely generated E^*_{n+1}-module N, there is a natural isomorphism*

$$\mathrm{Map}_c(T, N/IN) \cong \mathrm{Map}_c(T, N)/I\mathrm{Map}_c(T, N).$$

By Lemma 3.1, it is fundamental to understand $\mathrm{Map}_c(T, E^*_{n+1})$. Recall that a module over a (graded) regular local ring is called profree if it is isomorphic to the completion at the maximal ideal of some free module (see [Hovey and Strickland 1999, Theorem A.9] for equivalent conditions of profree modules).

Proposition 3.3. *For a topological space T, $\mathrm{Map}_c(T, E^*_{n+1})$ is a profree E^*_{n+1}-module.*

Proof. Put $P = \mathrm{Map}_c(T, E^*_{n+1})$. We have $P \cong \varprojlim_r \mathrm{Map}_c(T, E^*_{n+1}/I^r_{n+1})$, since $E^*_{n+1} \cong \varprojlim_r E^*_{n+1}/I^r_{n+1}$. Then $P \cong \varprojlim_r P/I^r_{n+1}P$ by Corollary 3.2. This shows that P is L-complete by [Hovey and Strickland 1999, Theorem A.6(a)]. Since $p, u_1, \dots, u_n$ is a regular sequence on E^*_{n+1},

$$0 \to E^*_{n+1}/I_k \xrightarrow{u_k} E^*_{n+1}/I_k \to E^*_{n+1}/I_{k+1} \to 0$$

is an exact sequence of profinite modules for $k = 0, 1, \dots, n$. By applying the functor $\mathrm{Map}_c(T, -)$, we obtain an exact sequence

$$0 \to P/I_kP \xrightarrow{u_k} P/I_kP \to P/I_{k+1}P \to 0$$

for $k = 0, 1, \dots, n$ by Corollary 3.2. Hence $p, u_1, \dots, u_n$ is a regular sequence on P, and P is profree by [Hovey and Strickland 1999, Theorem A.9]. □

Recall that $\mathbb{A} = L_{K(n)}E_{n+1}$ and $\mathbb{A}^* \cong E^*_{n+1}[u_n^{-1}]^{\wedge}_{I_n} = \varprojlim_r E^*_{n+1}/I^r_n[u_n^{-1}]$ by Lemma 2.1. We denote by J_n the ideal of $\mathbb{A}^*$ generated by $p, u_1, \dots, u_{n-1}$, that is, $J_n = I_n\mathbb{A}^* \subset \mathbb{A}^*$. Then we have $\mathbb{A}^*/J^r_n = E^*_{n+1}/I^r_n[u_n^{-1}]$. Note that $\mathbb{A}^*/J^r_n$ is a graded ring of formal Laurent series over an Artinian local ring. To introduce a topology for $\mathbb{A}^*$-modules of certain type, we first consider the case of such a ring.

Definition 3.4. Let R be a (graded) Artinian local ring. Then the ring $R[\![a]\!]$ of formal power series is a Noetherian local ring. Note that the topology of $R[\![a]\!]$ coincides with the (a)-adic topology since the maximal ideal of R is nilpotent. We give the ring $R((a)) = R[\![a]\!][a^{-1}]$ of formal Laurent series a $R[\![a]\!]$-linear topology such that $R[\![a]\!]$ is an open submodule. Then $R((a))$ is a union of open submodules

$a^r R[\![a]\!]$ for $r \in \mathbb{Z}$: $R((a)) = \bigcup_{r\in\mathbb{Z}} a^r R[\![a]\!]$. For an $R[\![a]\!]$-module N, we give the (a)-adic topology on N. The localization $N[a^{-1}]$ is an $R((a))$-module. Let N' be the image of the localization map $N \to N[a^{-1}]$. Then N' is an $R[\![a]\!]$-submodule of $N[a^{-1}]$. We give an $R[\![a]\!]$-linear topology on $N[a^{-1}]$ such that N' is an open submodule. Then $N[a^{-1}]$ is a union of open submodules $a^r N'$ for $r \in \mathbb{Z}$: $N[a^{-1}] = \bigcup_{r\in\mathbb{Z}} a^r N'$.

For an $R[\![a]\!]$-module N, the localization map $N \to N[a^{-1}]$ induces a map $\mathrm{Map}_c(T, N)[a^{-1}] \to \mathrm{Map}_c(T, N[a^{-1}])$ of $R((a))$-modules. The following lemma gives a sufficient condition that this map is an isomorphism.

Lemma 3.5. *Let R be a (graded) Artinian local ring with finite residue field, and let T be a compact space. For an $R[\![a]\!]$-module N, there is a natural isomorphism*

$$\mathrm{Map}_c(T, N[a^{-1}]) \cong \mathrm{Map}_c(T, N')[a^{-1}],$$

where N' is the image of the localization map $N \to N[a^{-1}]$. Furthermore, if N is (a)-torsion free or finitely generated, then there is a natural isomorphism

$$\mathrm{Map}_c(T, N[a^{-1}]) \cong \mathrm{Map}_c(T, N)[a^{-1}].$$

Proof. Since $N[a^{-1}]$ is a union of open submodules $a^r N'$ for $r \in \mathbb{Z}$, any continuous map from T to $N[a^{-1}]$ factors through $a^r N'$ for some r. Hence

$$\mathrm{Map}_c(T, N')[a^{-1}] \xrightarrow{\cong} \mathrm{Map}_c(T, N[a^{-1}]).$$

If N is (a)-torsion free, then $N' = N$. Assume that N is finitely generated. Let K be the kernel of the surjection $N \to N'$. Since $N[a^{-1}] \cong N'[a^{-1}]$, $K[a^{-1}] = 0$. Since K is finitely generated, there is a positive integer m such that $a^m K = 0$. Since $R[\![a]\!]$ is profinite, $\mathrm{Map}_c(T, -)$ is an exact functor on the category of finitely generated $R[\![a]\!]$-modules. Then the exact sequence $0 \to K \to N \to N' \to 0$ induces an exact sequence $0 \to \mathrm{Map}_c(T, K) \to \mathrm{Map}_c(T, N) \to \mathrm{Map}_c(T, N') \to 0$. The fact that $a^m K = 0$ implies $a^m \mathrm{Map}_c(T, K) = 0$. Hence $\mathrm{Map}_c(T, K)[a^{-1}] = 0$. So we obtain that $\mathrm{Map}_c(T, N)[a^{-1}] \cong \mathrm{Map}_c(T, N')[a^{-1}]$. □

We define a topology for $\mathbb{A}^*$-modules of the form $\varprojlim_r N/I_n^r[u_n^{-1}]$ for some E_{n+1}^*-module N.

Definition 3.6. For an $\mathbb{A}^*/J_n^r$-module M, since $\mathbb{A}^*/J_n^r$ is a graded ring of formal Laurent series over an Artinian local ring, we give a topology on M as in Definition 3.4. For an E_{n+1}^*-module N, we define an $\mathbb{A}^*$-module $\mathbb{A}^* N$ by

$$\mathbb{A}^* N = N[u_n^{-1}]^{\wedge}_{I_n} = \varprojlim_r N/I_n^r N[u_n^{-1}].$$

Then $N/I_n^r[u_n^{-1}]$ is an $\mathbb{A}^*/J_n^r$-module. We give $\mathbb{A}^* N = \varprojlim_r N/I_n^r N[u_n^{-1}]$ a topology by using the inverse limit topology.

Note that there is an isomorphism $\mathbb{A}^*E^*_{n+1} \cong \mathbb{A}^*$. If N is a finitely generated E^*_{n+1}-module, then $N[u_n^{-1}]$ is finitely generated over the Noetherian ring $E^*_{n+1}[u_n^{-1}]$. Then the completion of $N[u_n^{-1}]$ at the ideal I_n is given by the tensor product with $\mathbb{A}^*$. Hence there is a natural isomorphism $\mathbb{A}^*N \cong \mathbb{A}^* \otimes_{E^*_{n+1}} N$ for any finitely generated E^*_{n+1}-module N, and the functor $\mathbb{A}^*(-)$ is exact on the category of finitely generated E^*_{n+1}-modules.

In the rest of this section we study the functor $\mathrm{Map}_c(T, \mathbb{A}^*(-))$ with T compact.

Lemma 3.7. *If T is a compact space and N is a finitely generated E^*_{n+1}-module, then there is a natural isomorphism of $\mathbb{A}^*$-modules*

$$\mathrm{Map}_c(T, \mathbb{A}^*N) \cong \mathbb{A}^*\mathrm{Map}_c(T, N).$$

Proof. Since $\mathbb{A}^*N = \varprojlim_r N/I_n^r N[u_n^{-1}]$, we have

$$\mathrm{Map}_c(T, \mathbb{A}^*N) \cong \varprojlim_r \mathrm{Map}_c(T, N/I_n^r N[u_n^{-1}]).$$

By Lemma 3.5 and Corollary 3.2,

$$\mathrm{Map}_c(T, N/I_n^r N[u_n^{-1}]) \cong \mathrm{Map}_c(T, N)/I_n^r \mathrm{Map}_c(T, N)[u_n^{-1}].$$

Hence $\mathrm{Map}_c(T, \mathbb{A}^*N)$ is isomorphic to $\varprojlim_r \mathrm{Map}_c(T, N)/I_n^r \mathrm{Map}_c(T, N)[u_n^{-1}] = \mathbb{A}^*\mathrm{Map}_c(T, N)$. □

The basic case is when $N = E^*_{n+1}$:

Proposition 3.8. *For any compact space T, $\mathrm{Map}_c(T, \mathbb{A}^*)$ is a profree $\mathbb{A}^*$-module.*

Proof. By Proposition 3.3, $\mathrm{Map}_c(T, E^*_{n+1})$ is profree over E^*_{n+1}, and is thus a direct summand of some product $\prod_\alpha E^*_{n+1}$ by [Hovey and Strickland 1999, Proposition A.13]. Hence it is sufficient to show that $\mathbb{A}^*(\prod_\alpha E^*_{n+1})$ is profree over $\mathbb{A}^*$. For $k = 0, 1, \dots, n-1$, we put $M = E^*_{n+1}/I_k$ and $N = E^*_{n+1}/I_{k+1}$. Let K_r be the kernel of the map $M/I_n^r M \xrightarrow{u_k} M/I_n^r M$, and let L_r be the kernel of the map $M/I_n^r M \to N/I_n^r N$. Then there are exact sequences $0 \to K_r \to M/I_n^r M \to L_r \to 0$ and $0 \to L_r \to M/I_n^r M \to N/I_n^r N \to 0$. Since E^*_{n+1} is regular, the canonical map $K_{r+1} \to K_r$ is 0. Then

$$\varprojlim_r((\prod_\alpha K_r)[u_n^{-1}]) = \varprojlim_r^1((\prod_\alpha K_r)[u_n^{-1}]) = 0.$$

Hence we obtain $\varprojlim_r((\prod_\alpha M/I_n^r M)[u_n^{-1}]) \xrightarrow{\cong} \varprojlim_r((\prod_\alpha L_r)[u_n^{-1}])$, and

$$0 = \varprojlim_r^1((\prod_\alpha M/I_n^r M)[u_n^{-1}]) \cong \varprojlim_r^1((\prod_\alpha L_r)[u_n^{-1}]).$$

This implies that the sequence

$$\begin{aligned} 0 \to \varprojlim_r((\prod_\alpha M/I_n^r)[u_n^{-1}]) \\ \xrightarrow{u_k} \varprojlim_r((\prod_\alpha M/I_n^r M)[u_n^{-1}]) \longrightarrow \varprojlim_r((\prod_\alpha N/I_n^r N)[u_n^{-1}]) \to 0 \end{aligned}$$

is exact. This shows that $p, u_1, \ldots, u_{n-1}$ is a regular sequence on $\mathbb{A}^*(\prod_\alpha E_{n+1}^*)$. Therefore $\mathbb{A}^*(\prod_\alpha E_{n+1}^*)$ is profree $\mathbb{A}^*$-module by [Hovey and Strickland 1999, Theorem A.9]. □

The map from T to the one point space $*$ induces a ring homomorphism $\mathbb{A}^* = \mathrm{Map}_c(*, \mathbb{A}^*) \to \mathrm{Map}_c(T, \mathbb{A}^*)$. Then the composition with the commutative MU^*-algebra structure map $MU^* \to \mathbb{A}^*$ gives $\mathrm{Map}_c(T, \mathbb{A}^*)$ a commutative MU^*-algebra structure. Since a profree module over $\mathbb{A}^*$ is Landweber exact, we obtain the following corollary

Corollary 3.9. *If T is a compact space, then $\mathrm{Map}_c(T, \mathbb{A}^*)$ is Landweber exact.*

We have a similar description for $\mathrm{Map}_c(T, \mathbb{A}^*N)$ as in Lemma 3.1 when T is a compact space and N is a finitely generated E_{n+1}^*-module as follows.

Proposition 3.10. *If T is a compact space and N is a finitely generated E_{n+1}^*-module, then there is a natural isomorphism of $\mathbb{A}^*$-modules*

$$\mathrm{Map}_c(T, \mathbb{A}^*N) \cong \mathrm{Map}_c(T, \mathbb{A}^*) \underset{\mathbb{A}^*}{\otimes} \mathbb{A}^*N.$$

For the proof of Proposition 3.10, we prepare the following (well-known) lemmas.

Lemma 3.11 ([Lam 1999, Proposition 4.4]). *Let R be a (graded) ring. If M is a finitely presented module over R, then $(\prod_\alpha R) \otimes_R M \cong \prod_\alpha M$.*

Proof. Since M is finitely presented, there is an exact sequence $M^1 \to M^0 \to M \to 0$, where M^i is finitely generated free for $i = 0, 1$. Then there are two exact sequences $(\prod_\alpha R) \otimes M^1 \to (\prod_\alpha R) \otimes M^0 \to (\prod_\alpha R) \otimes M \to 0$ and $\prod_\alpha M^1 \to \prod_\alpha M^0 \to \prod_\alpha M \to 0$. Since M^i is finitely generated free, $(\prod_\alpha R) \otimes M^i \cong \prod_\alpha M^i$ for $i = 0, 1$. Hence we obtain $(\prod_\alpha R) \otimes M \cong \prod_\alpha M$. □

Lemma 3.12. *If F is a profree $\mathbb{A}^*$-module and M is a finitely generated $\mathbb{A}^*$-module, then $F \otimes_{\mathbb{A}^*} M$ is J_n-adically complete.*

Proof. Since F is profree, it is a direct summand of some product $\prod_\alpha \mathbb{A}^*$ by [Hovey and Strickland 1999, Proposition A.13]. Since a direct summand of complete module is complete, it is sufficient to show that $(\prod_\alpha \mathbb{A}^*) \otimes M$ is complete. By Lemma 3.11, $(\prod_\alpha \mathbb{A}^*) \otimes M \cong \prod_\alpha M$, and $\prod_\alpha M$ is complete. □

Proof of Proposition 3.10. By Lemma 3.1, $\mathrm{Map}_c(T, N) \cong \mathrm{Map}_c(T, E_{n+1}^*) \otimes_{E_{n+1}^*} N$. Then we see that $\mathbb{A}^*\mathrm{Map}_c(T, N)$ is the completion of $\mathbb{A}^*\mathrm{Map}_c(T, E_{n+1}^*) \otimes_{\mathbb{A}^*} \mathbb{A}^*N$ at the ideal J_n. By Lemma 3.12, we see that $\mathbb{A}^*\mathrm{Map}_c(T, E_{n+1}^*) \otimes_{\mathbb{A}^*} \mathbb{A}^*N$ is J_n-adically complete. Hence we obtain

$$\mathbb{A}^*\mathrm{Map}_c(T, N) \cong \mathbb{A}^*\mathrm{Map}_c(T, E_{n+1}^*) \otimes_{\mathbb{A}^*} \mathbb{A}^*N. \qquad \square$$

Let $\mathcal{S}$ be the stable homotopy category, and let $\mathcal{K}$ be the $K(n)$-local stable homotopy category. For a $K(n)$-local spectrum $X \in \mathcal{K}$, we define $\Lambda''(X)$ to be the full subcategory of the comma category $(\mathcal{S} \downarrow X)$, whose objects are maps $X'' \to X$ from finite spectra X'' of type at least n. Then $\Lambda''(X)$ is an essentially small filtered category (see [Hovey and Strickland 1999, §9] and [Hovey et al. 1997, §2.3]). For a spectrum $W \in \mathcal{S}$, we set $\Lambda(W) = \Lambda''(L_{K(n)}W)$. The following lemma gives a sufficient condition that we can describe a generalized cohomology group of W in terms of cohomology groups of W_λ for $\lambda \in \Lambda(W)$.

Lemma 3.13. *Let R be a $K(n)$-local commutative ring spectrum. Suppose that the coefficient ring R^* is even-periodic and R^0 is a linearly compact Noetherian ring. Then there is a natural isomorphism*

$$R^*(W) \cong \varprojlim_\lambda R^*(W_\lambda)$$

for any $W \in \mathcal{S}$, where the inverse limit is taken over $\lambda \in \Lambda(W)$.

Proof. For $W \in \mathcal{S}$, we set $F^*(W) = \varprojlim_\lambda R^*(W_\lambda)$. Note that $R^*(W) \cong R^*(L_{K(n)}W)$ for any $W \in \mathcal{S}$ since R is $K(n)$-local. Then it is sufficient to show that $R^*(X) \cong F^*(X)$ for any $X \in \mathcal{K}$. By the assumption of the coefficient ring R^*, the functor $R^*(-)$ on the category of finite spectra takes values in the category of linearly compact R^*-modules and continuous maps. Then $F^*(-)$ is a cohomology theory on $\mathcal{S}$ by [Hovey et al. 1997, Proposition 2.3.16] and [Hovey and Strickland 1999, Proposition 9.2]. There is a natural transformation $R^*(-) \to F^*(-)$ of cohomology theories, which induces an isomorphism

$$R^*(X'') \xrightarrow{\cong} F^*(X'')$$

for any finite spectrum X'' of type at least n. Since $L_{K(n)}F(n)$ is a graded weak generator of $\mathcal{K}$ for any finite spectrum $F(n)$ of type n ([Hovey and Strickland 1999, Theorem 7.3]), we obtain that $R^*(X) \xrightarrow{\cong} F^*(X)$ for any $X \in \mathcal{K}$. □

Definition 3.14. For a finite spectrum X of type at least n, $E^*_{n+1}(X)$ is annihilated by a power of I_n, and $\mathbb{A}^*(X) \cong E^*_{n+1}(X)[u_n^{-1}]$ is a module over $\mathbb{A}^*/J_n^r = E^*_{n+1}/I_n^r[u_n^{-1}]$ for some r. We give a topology on $\mathbb{A}^*(X)$ as in Definition 3.6. For a spectrum W, $\mathbb{A}^*(W) \cong \varprojlim_\lambda \mathbb{A}^*(W_\lambda)$ by Lemma 3.13, where W_λ are finite spectra of type at least n. We give a topology on $\mathbb{A}^*(W)$ by the inverse limit topology.

For a compact space T and a finite spectrum X of type at least n,

$$\mathrm{Map}_c(T, \mathbb{A}^*(X)) \cong \mathrm{Map}_c(T, \mathbb{A}^*) \otimes_{\mathbb{A}^*} \mathbb{A}^*(X)$$

by Proposition 3.10, and $\mathrm{Map}_c(T, \mathbb{A}^*)$ is profree by Proposition 3.8. To study the functor $\mathrm{Map}_c(T, \mathbb{A}^*(-))$ on the stable homotopy category $\mathcal{S}$, we consider the following functor. Let F be a profree $\mathbb{A}^*$-module. We define a functor $H_F(-)$ from the stable homotopy category $\mathcal{S}$ to the category of $\mathbb{A}^*$-modules by

$$H_F(W) = \varprojlim_\lambda F \otimes_{\mathbb{A}^*} \mathbb{A}^*(W_\lambda),$$

where the inverse limit is taken over $\lambda \in \Lambda(W)$.

Lemma 3.15. *The functor $H_F(-)$ is a cohomology theory on $\mathcal{S}$.*

Proof. Since F is a direct summand of some product $\prod_\alpha \mathbb{A}^*$ by [Hovey and Strickland 1999, Proposition A.13], it is sufficient to show that the functor $Z \mapsto \varprojlim_\lambda (\prod_\alpha \mathbb{A}^*) \otimes_{\mathbb{A}^*} \mathbb{A}^*(W_\lambda)$ is a cohomology theory. Since $\mathbb{A}^*(W_\lambda)$ is finitely presented, $(\prod_\alpha \mathbb{A}^*) \otimes_{\mathbb{A}^*} \mathbb{A}^*(W_\lambda) \cong \prod_\alpha \mathbb{A}^*(W_\lambda)$ by Lemma 3.11. Hence

$$\varprojlim_\lambda (\prod_\alpha \mathbb{A}^*) \otimes_{\mathbb{A}^*} \mathbb{A}^*(W_\lambda) \cong \prod_\alpha \mathbb{A}^*(W),$$

and $\prod_\alpha \mathbb{A}^*(W)$ is a cohomology theory. This completes the proof. $\square$

The following theorem will be used to identify the E_2-term of the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence to the continuous cohomology group of G_{n+1} in Section 4 below.

Theorem 3.16. *For any compact space T, the functor $\mathrm{Map}_c(T, \mathbb{A}^*(-))$ is a cohomology theory.*

Proof. By Proposition 3.10, there is a natural isomorphism

$$\mathrm{Map}_c(T, \mathbb{A}^*(W)) \cong \varprojlim_\lambda \mathrm{Map}_c(T, \mathbb{A}^*) \underset{\mathbb{A}^*}{\hat{\otimes}} \mathbb{A}^*(W_\lambda).$$

But $\mathrm{Map}_c(T, \mathbb{A}^*)$ is profree by Proposition 3.8. Therefore the theorem follows from Lemma 3.15. $\square$

4. Construction of the spectral sequence

We set $\widehat{\mathbb{S}} = L_{K(n)} L_{K(n+1)} S^0$. In this section we construct a spectral sequence which converges strongly and conditionally to $[W, \widehat{\mathbb{S}}]^*$ for any spectrum W by applying the $K(n)$-localization functor to the $K(n+1)$-local E_{n+1}-Adams resolution of $L_{K(n+1)} S^0$. Then we describe the E_2-term in terms of the continuous cohomology group of G_{n+1} with coefficients in $\mathbb{A}^*(W)$.

Let $E_n^{\wedge s}$ be the $K(n)$-localization of the smash product of s-copies of E_n

$$E_n^{\wedge s} = L_{K(n)}(\overbrace{E_n \wedge \cdots \wedge E_n}^{s}).$$

The commutative ring spectrum structure on E_n gives $E_n^{\wedge \bullet+1} = \{E_n^{\wedge s+1}\}_{s \geq 0}$ a cosimplicial $K(n)$-local commutative ring spectrum structure with augmentation

$L_{K(n)}S^0 \xrightarrow{\varepsilon} E_n^{\wedge \bullet+1}$. Then the associated cochain complex

$$* \to L_{K(n)}S^0 \xrightarrow{\varepsilon} E_n \xrightarrow{d} E_n^{\wedge 2} \xrightarrow{d} E_n^{\wedge 3} \xrightarrow{d} \cdots \tag{4-1}$$

is a $K(n)$-local E_n-Adams resolution of $L_{K(n)}S^0$ in the sense of [Miller 1981; Devinatz and Hopkins 2004]. We denote the sequence (4-1) by $\mathrm{Res}(E_n; L_{K(n)}S^0)$. There is an associated diagram of exact triangles

$$\begin{tikzcd} L_{K(n)}S^0 = Y^0 \arrow[dr, "j"'] & & Y^1 \arrow[ll, "i"'] \arrow[dr, "j"'] & & Y^2 \arrow[ll, "i"'] \arrow[dr, "j"'] & & Y^3 \arrow[ll, "i"'] & \\ & E_n \arrow[ur, "k"'] & & \Sigma^{-1}E_n^{\wedge 2} \arrow[ur, "k"'] & & \Sigma^{-2}E_n^{\wedge 3} \arrow[ur, "k"'] & & \cdots, \end{tikzcd} \tag{4-2}$$

in the $K(n)$-local stable homotopy category, where k has degree -1 and $jk = d$. We denote by $\mathrm{Ad}(E_n; L_{K(n)}S^0)$ the diagram of exact triangles (4-2).

For any spectrum W, by applying the functor $[W, -]^*$ to $\mathrm{Ad}(E_n; L_{K(n)}S^0)$ we obtain a $K(n)$-local E_n-Adams spectral sequence

$$L_{K(n)}E_r^{s,t}(W) \Longrightarrow [W, L_{K(n)}S^0]^{s+t}$$

with $L_{K(n)}E_2^{s,t}(W) \cong H_c^s(G_n; E_n^t(W))$. This spectral sequence converges strongly and conditionally. Furthermore, since $L_{K(n)}S^0$ is $K(n)$-local E_n-nilpotent [Devinatz and Hopkins 2004, Proposition A.3], the filtration (4-2) has the following property: There exists $N > 0$ such that $Y^{s+N} \to Y^s$ is null for all $s \geq 0$. This property implies that there exist positive integers $r(n)$ and $s(n)$, which do not depend on W, such that $L_{K(n)}E_{r(n)}^{s,*}(W) = 0$ for $s > s(n)$.

By applying the $K(n)$-localization functor to $\mathrm{Ad}(E_{n+1}; L_{K(n+1)}S^0)$, we obtain the following diagram $L_{K(n)}\mathrm{Ad}(E_{n+1}, L_{K(n+1)}S^0)$ of exact triangles

$$\begin{tikzcd} \widehat{\mathbb{S}} = Z^0 \arrow[dr, "j"'] & & Z^1 \arrow[ll, "i"'] \arrow[dr, "j"'] & & Z^2 \arrow[ll, "i"'] \arrow[dr, "j"'] & & Z^3 \arrow[ll, "i"'] & \\ & L_{K(n)}E_{n+1} \arrow[ur, "k"'] & & \Sigma^{-1}L_{K(n)}E_{n+1}^{\wedge 2} \arrow[ur, "k"'] & & \Sigma^{-2}L_{K(n)}E_{n+1}^{\wedge 3} \arrow[ur, "k"'] & & \cdots. \end{tikzcd} \tag{4-3}$$

For any spectrum W, applying the functor $[W, -]^*$ to $L_{K(n)}\mathrm{Ad}(E_{n+1}, L_{K(n+1)}S^0)$, we obtain a spectral sequence

$$L_{K(n)}L_{K(n+1)}E_r^{s,t}(W) \Longrightarrow [W, \widehat{\mathbb{S}}]^{s+t}.$$

We call this spectral sequence the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence.

Lemma 4.1. *The spectral sequence $L_{K(n)}L_{K(n+1)}E_r^{s,t}(W) \Longrightarrow [W, \widehat{\mathbb{S}}]^{s+t}$ converges conditionally and strongly for any spectrum W.*

Proof. There exists $N > 0$ such that $Y^{s+N} \to Y^s$ is null for all $s \geq 0$. Applying the $K(n)$-localization functor, we see that $Z^{s+N} \to Z^s$ is also null for all $s \geq 0$. This implies that the filtration of $[W, \widehat{\mathbb{S}}]^*$ is finite. Hence the spectral sequence converges strongly by [Boardman 1999, Definition 5.2]. Also, we obtain that $\varprojlim_n [W, Z^n]^* = \varprojlim^1_n [W, Z^n]^* = 0$. Hence the spectral sequence converges conditionally by [Boardman 1999, Definition 5.10]. □

Remark 4.2. Note that there exist positive integers r_0 and s_0, which do not depend on W, such that $L_{K(n)}L_{K(n+1)}E_{r_0}^{s,*}(W) = 0$ for $s > s_0$.

In the rest of this section we identify the E_2-term of the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence $L_{K(n)}L_{K(n+1)}E_r^{s,t}(W)$ with the continuous cohomology group of G_{n+1} with coefficients in $\mathbb{A}^*(W)$. Let $C(s) = E_{n+1}^{\wedge s+1}$. The E_1-term of the spectral sequence is given by $E_1^{s,t} = [W, L_{K(n)}C(s)]^t$. There is an isomorphism $C(s)^* \cong \mathrm{Map}_c(G_{n+1}^s, E_{n+1}^*)$ (see [Devinatz and Hopkins 2004, §2]). Then we see that $C(s)^*$ is profree over E_{n+1}^* by Proposition 3.3. The following lemma gives a similar description for $L_{K(n)}C(s)^*$.

Lemma 4.3. *For $s \geq 0$, we have $L_{K(n)}C(s)^* \cong \mathrm{Map}_c(G_{n+1}^s, \mathbb{A}^*)$.*

Proof. There is a tower $\{M(J)\}_J$ of generalized Moore spectra of type n as in [Hovey and Strickland 1999, Proposition 4.2] such that $L_{K(n)}W \simeq \mathrm{holim}_J L_n W \wedge M(J)$ for any spectrum W [Hovey and Strickland 1999, Proposition 7.10(e)]. Since $C(s)$ is Landweber exact of height $(n+1)$, we obtain that $L_{K(n)}C(s)^* \cong \mathbb{A}^*C(s)^*$. Then $\mathbb{A}^*C(s)^* \cong \mathrm{Map}_c(G_{n+1}^s, \mathbb{A}^*)$ by Lemma 3.7, since

$$C(s)^* \cong \mathrm{Map}_c(G_{n+1}^s, E_{n+1}^*).$$

□

Corollary 4.4. *For $s \geq 0$, $L_{K(n)}C(s)^*$ is Landweber exact and profree over $\mathbb{A}^*$.*

Proof. This follows from Proposition 3.8 and Corollary 3.9. □

Then we obtain a description for the E_1-term $[W, L_{K(n)}C(s)]^*$ as a module of continuous maps from G_{n+1}^s to $\mathbb{A}^*(W)$.

Proposition 4.5. *For any spectrum W, there is a natural isomorphism*

$$[W, L_{K(n)}C(s)]^* \cong \mathrm{Map}_c(G_{n+1}^s, \mathbb{A}^*(W)).$$

Proof. By Lemma 4.3 and Corollary 4.4, $L_{K(n)}C(s)^* \cong \mathrm{Map}_c(G_{n+1}^s, \mathbb{A}^*)$ is Landweber exact. Then there is a natural isomorphism

$$[W, L_{K(n)}C(s)]^* \cong \mathrm{Map}_c(G_{n+1}^s, \mathbb{A}^*) \otimes_{\mathbb{A}^*} \mathbb{A}^*(W)$$

for any finite spectrum W. By Proposition 3.10, the right hand side is isomorphic to $\mathrm{Map}_c(G_{n+1}^s, \mathbb{A}^*(W))$. Since $\mathrm{Map}_c(G_{n+1}^s, \mathbb{A}^*)$ is even concentrated, there is a unique extension to a cohomology theory for any spectra by [Hovey and Strickland 1999, Theorem 2.8]. Obviously, $[-, L_{K(n)}C(s)]^*$ is such an extension. On the other hand, $\mathrm{Map}_c(G^s_{n+1}, \mathbb{A}^*(-))$ is also an extension by Theorem 3.16. Therefore $[W, L_{K(n)}C(s)]^* \cong \mathrm{Map}_c(G^s_{n+1}, \mathbb{A}^*(W))$ for any spectrum W. □

For a topological group G and a topological G-module M, denote by $C^*_c(G; M)$ the continuous cochain complex of G with coefficients in M. Define $H^*_c(G; M)$ to be the cohomology group of $C^*_c(G; M)$, and call it the continuous cohomology of G with coefficients in M. Let $[W, C(*)]^t$ be the cochain complex associated with the cosimplicial abelian group $[W, C(\bullet)]^t$. Then there is a natural isomorphism $[W, C(*)]^t \cong C^*_c(G_{n+1}, E^t_{n+1}(W))$ of cochain complexes [Devinatz and Hopkins 2004, §4]. By Proposition 4.5, this implies a natural isomorphism $[W, L_{K(n)}C(*)]^t \cong C^*_c(G_{n+1}, \mathbb{A}^t(W))$ of cochain complexes. Hence we obtain the following corollary.

Corollary 4.6. *For any spectrum W, there is a natural isomorphism*

$$H^s([W, L_{K(n)}C(*)]^t) \cong H^s_c(G_{n+1}; \mathbb{A}^t(W)).$$

As a summary we obtain the following theorem.

Theorem 4.7. *For any spectrum W, there is a natural spectral sequence*

$$L_{K(n)}L_{K(n+1)}E^{s,t}_r(W)$$

which converges strongly and conditionally to $[W, \widehat{\mathbb{S}}]^*$:

$$L_{K(n)}L_{K(n+1)}E^{s,t}_2(W) \Longrightarrow [W, \widehat{\mathbb{S}}]^{s+t}.$$

The E_2-term is given by

$$L_{K(n)}L_{K(n+1)}E^{s,t}_2(W) \cong H^s_c(G_{n+1}; \mathbb{A}^t(W)).$$

Furthermore, there exist positive integers r_0 and s_0 such that

$$L_{K(n)}L_{K(n+1)}E^{s,*}_{r_0}(W) = 0$$

for $s > s_0$, where r_0 and s_0 do not depend on W.

5. The cohomology group $\boldsymbol{H}^*_c(\mathbb{G}; \mathbb{B}^*(W))$

In this section we introduce a cohomology group $\boldsymbol{H}^*_c(\mathbb{G}; \mathbb{B}^*(W))$ of $\mathbb{G}$ with coefficients in $\mathbb{B}^*(W)$ for a spectrum W. Then we show that $\boldsymbol{H}^*_c(\mathbb{G}; \mathbb{B}^*(W))$ is naturally isomorphic to the continuous cohomology group $H^*_c(G_{n+1}; \mathbb{A}^*(W))$ of G_{n+1} with coefficients in $\mathbb{A}^*(W)$. The cohomology group $\boldsymbol{H}^*_c(\mathbb{G}; \mathbb{B}^*(W))$ will be used to connect the E_2-term of the $K(n)$-local E_n-Adams spectral sequence for W and the E_2-term of the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence for W in Section 7 below.

First we introduce a topology for modules of continuous maps from a profinite group to an $\mathbb{A}^*$-module of certain type. Then we study a continuous cohomology group of a profinite group with coefficients in such a topological module of mappings.

Definition 5.1. Let G be a profinite group. Suppose that $M = \varprojlim_\lambda \mathbb{A}^* N_\lambda$ with the inverse limit topology, where $\{N_\lambda\}_{\lambda\in\Lambda}$ is a cofiltered system of finitely generated E^*_{n+1}-modules. By Lemma 3.7, there is an isomorphism

$$\mathrm{Map}_c(G, M) \cong \varprojlim_\lambda \mathbb{A}^* \mathrm{Map}_c(G, N_\lambda).$$

We give a topology on $\mathbb{A}^*\mathrm{Map}_c(G, N_\lambda)$ as in Definition 3.6. Then we give a topology on $\mathrm{Map}_c(G, M)$ by the inverse limit topology. For any spectrum W, $\mathbb{A}^*(W) \cong \varprojlim_\lambda \mathbb{A}^* E^*_{n+1}(W_\lambda)$ by Lemma 3.13, where W_λ are finite spectra of type at least n. We give a topology on $\mathrm{Map}_c(G, \mathbb{A}^*(W))$ as above.

The following lemma shows that the mapping spaces have an expected adjunction property.

Lemma 5.2. *Let G and H be profinite groups. Suppose that $M = \varprojlim_\lambda \mathbb{A}^* N_\lambda$ with the inverse limit topology, where $\{N_\lambda\}_{\lambda\in\Lambda}$ is a cofiltered system of finitely generated E^*_{n+1}-modules. Then there is an isomorphism*

$$\mathrm{Map}_c(G, \mathrm{Map}_c(H, M)) \cong \mathrm{Map}_c(G \times H, M).$$

Proof. We have

$$\mathrm{Map}_c(G, \mathrm{Map}_c(H, M) = \varprojlim_\lambda \mathrm{Map}_c(G, \mathrm{Map}_c(H, \mathbb{A}^* N_\lambda),$$
$$\mathrm{Map}_c(G \times H, M) = \varprojlim_\lambda \mathrm{Map}_c(G \times H, \mathbb{A}^* N_\lambda).$$

Hence it is sufficient to show that the lemma holds when $M = \mathbb{A}^* N$ with finitely generated N. Suppose that N is a finitely generated E^*_{n+1}-module. Let N_r be the image of the localization map $N/I_n^r N \to N/I_n^r N[u_n^{-1}]$, and let $L_r = \mathrm{Map}_c(H, N_r)$. Note that N_r and L_r are (u_n)-torsion free. By Lemma 3.5, $\mathrm{Map}_c(H, \mathbb{A}^* N) = \varprojlim_r L_r[u_n^{-1}]$. Then $\mathrm{Map}_c(G, \mathrm{Map}_c(H, \mathbb{A}^* N)) = \varprojlim_r \mathrm{Map}_c(G, L_r[u_n^{-1}])$. Again by Lemma 3.5, we have $\mathrm{Map}_c(G, L_r[u_n^{-1}]) = \mathrm{Map}_c(G, L_r)[u_n^{-1}]$. The fact that N_r is a profinite module implies that $\mathrm{Map}_c(G, L_r) = \mathrm{Map}_c(G \times H, N_r)$. By Lemma 3.5, we obtain $\varprojlim_r \mathrm{Map}_c(G \times H, N_r)[u_n^{-1}] = \mathrm{Map}_c(G \times H, \mathbb{A}^* N)$. □

Corollary 5.3. *Let G and H be profinite groups. For any spectrum W, there is a natural isomorphism*

$$\mathrm{Map}_c(G, \mathrm{Map}_c(H, \mathbb{A}^*(W)))) \cong \mathrm{Map}_c(G \times H, \mathbb{A}^*(W)).$$

Suppose that a profinite group G continuously acts on a topological module M from the right. For $q > 0$, we define a right G-action on $\mathrm{Map}_c(G, M)$ by

$$\varphi^g(h_1, \ldots, h_q) = \varphi(h_1 g^{-1}, \ldots, h_q g^{-1})^g,$$

where $\varphi \in \mathrm{Map}_c(G^q, M)$ and $g, h_1, \ldots, h_q \in G$. Then $\mathrm{Map}_c(G^q, M)$ is a topological G-module. The following proposition shows that the coinduced module $\mathrm{Map}_c(G^q, M)$ is acyclic with respect to $H_c^*(G; -)$.

Proposition 5.4. *Let G be a profinite group. Suppose that $M = \varprojlim_\lambda \mathbb{A}^* N_\lambda$ with the inverse limit topology, where $\{N_\lambda\}_{\lambda\in\Lambda}$ is a cofiltered system of finitely generated E_{n+1}^*-modules. Furthermore, suppose that G continuously acts on M. For $p > 0$ and $q > 0$, we have $H_c^p(G; \mathrm{Map}_c(G^q, M)) = 0$, and $H_c^0(G; \mathrm{Map}_c(G^q, M)) = \mathrm{Map}_c(G^q, M)^G$.*

Proof. Set

$$C_c^{-1}(G; \mathrm{Map}_c(G^q, M)) = \mathrm{Map}_c(G^q, M)^G, \quad C^{p,q} = C_c^p(G; \mathrm{Map}_c(G^q, M)).$$

Then $C^{p,q} \cong \mathrm{Map}_c(G^q \times G^{p+1}, M)^G$ by Lemma 5.2. The boundary map $d^p : C^{p,q} \to C^{p+1,q}$ is given by

$$d^p f(h_1, \ldots, h_q; g_0, \ldots, g_{p+1}) = \sum_{i=0}^{p+1} (-1)^i f(h_1, \ldots, h_q; g_0, \ldots, g_{i-1}, g_{i+1}, \ldots, g_{p+1}).$$

We define $s^p : C^{p,q} \to C^{p-1,q}$ by

$$s^p f(h_1, \ldots, h_q; g_0, \ldots, g_{p-1}) = f(h_1, \ldots, h_q; h_q, g_0, \ldots, g_{p-1}).$$

Then we can verify that $s^{p+1} d^p(f) + d^{p-1} s^p(f) = f$ for any $f \in C^{p,q}$. □

Corollary 5.5. *Let $p > 0$ and $q > 0$. Then $H_c^p(G_{n+1}; \mathrm{Map}_c(G_{n+1}^q, \mathbb{A}^*(W))) = 0$ and $H_c^0(G_{n+1}; \mathrm{Map}_c(G_{n+1}^q, \mathbb{A}^*(W))) = \mathrm{Map}_c(G_{n+1}^q, \mathbb{A}^*(W))^{G_{n+1}}$ for any spectrum W.*

Next we define a cohomology group $\boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W))$. For this purpose, we introduce a topology on $\mathbb{B}(i)^*(W)$.

Definition 5.6. For a spectrum W, $\mathbb{B}(i)^*(W)$ is a product of finite many copies of $\mathbb{A}^*(W)$ since $\mathbb{B}(i)^*$ is finitely generated free over $\mathbb{A}^*$. We give a topology on $\mathbb{B}(i)^*(W)$ by the product topology.

Recall that the group $\mathbb{G} = G_{n+1} \times_\Gamma G_n$ acts on the cohomology theory $\mathbb{B}^*(-)$ as multiplicative cohomology operations by Proposition 2.3. For $i \geq -1$, we set $\mathbb{G}(i) = G_{n+1} \times_\Gamma G_n(i)$, where $G_n(i) = \Gamma \ltimes S_n(i)$. Then $\mathbb{G}(i)$ acts on $\mathbb{B}(i)^*(W)$ naturally and continuously. Note that we can write $\mathbb{B}(i)^*(W) = \varprojlim_\lambda \mathbb{A}^* N_\lambda$ with finitely generated E_{n+1}^*-modules N_λ since $\mathbb{B}(i)^*$ is finitely generated free over $\mathbb{A}^*$. Then $\mathrm{Map}_c(\mathbb{G}(i)^{p+1}, \mathbb{B}(i)^*(W))$ is a topological module for any $p \geq 0$ as in Definition 5.1.

Definition 5.7. For a spectrum W, we define a cochain complex $\boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ by

$$\boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W)) = \varprojlim_\lambda \varinjlim_i \ C_c^*(\mathbb{G}(i); \mathbb{B}(i)^*(W_\lambda)),$$

where the inverse limit is taken over $\lambda \in \Lambda(W)$. Then we define a cohomology group $\boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ of $\mathbb{G}$ with coefficients in $\mathbb{B}^*(W)$ to be the cohomology group of $\boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))$

$$\boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W)) = H^*(\boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))).$$

Note that both of $\boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ and $\boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ are not functors of $\mathbb{B}^*(W)$ in spite of their notation.

For a continuous cochain complex $C_c^*(G_{n+1}; \mathbb{A}^*(W))$ of G_{n+1} with coefficients in $\mathbb{A}^*(W)$, there is an isomorphism

$$C_c^*(G_{n+1}; \mathbb{A}^*(W)) \cong \varprojlim_\lambda C_c^*(G_{n+1}; \mathbb{A}^*(W_\lambda)).$$

The canonical maps $\mathbb{A}^*(W_\lambda) \to \mathbb{B}(i)^*(W_\lambda)$ and the projections $\mathbb{G}(i) \to G_{n+1}$ define a cochain map

$$C_c^*(G_{n+1}; \mathbb{A}^*(W)) \longrightarrow \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W)).$$

We call the induced map on cohomology groups an inflation map

$$H_c^*(G_{n+1}; \mathbb{A}^*(W)) \longrightarrow \boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W)). \tag{5-1}$$

In the rest of this section we prove the following theorem.

Theorem 5.8. *The inflation map $H_c^*(G_{n+1}; \mathbb{A}^*(W)) \to \boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ is an isomorphism for any spectrum W.*

By definition, $\boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ is the cohomology group of the inverse limit of the cochain complexes $\varinjlim_i C_c^*(\mathbb{G}(i); \mathbb{B}(i)^*(W_\lambda))$. For the cohomology group of the inverse limit of cochain complexes $\{C_\lambda^*\}_{\lambda \in \Lambda}$, we have a spectral sequence to describe it in terms of the cohomology groups of C_λ^* under suitable circumstances.

Lemma 5.9. *Let $\{C_\lambda^*\}_{\lambda \in \Lambda}$ be a system of cochain complexes indexed by a small category Λ. We assume that $\varprojlim_\lambda^j \ C_\lambda^* = 0$ for $j > 0$. Then there is a spectral sequence*

$$E_2^{s,t} = \varprojlim_\lambda^s \ H^t(C_\lambda^*) \Longrightarrow H^{s+t}(\varprojlim_\lambda C_\lambda^*).$$

Proof. Let $\prod^* C^*_\lambda$ be the double complex associated to the cosimplicial replacement [Bousfield and Kan 1972, XI.5] of $\{C^*_\lambda\}$. Then we have two spectral sequences

$$\varprojlim{}^s_\lambda \; H^t(C^*_\lambda) \implies H^{s+t}(\textstyle\prod^* C^*_\lambda),$$
$$H^s(\varprojlim{}^t_\lambda \; C^*_\lambda) \implies H^{s+t}(\textstyle\prod^* C^*_\lambda).$$

By the assumption, the second spectral sequence collapses to give $H^*(\varprojlim_\lambda C^*_\lambda) \cong H^*(\prod^* C^*_\lambda)$. Hence the first spectral sequence gives the desired one. □

The next lemma gives a sufficient condition for all the higher inverse limits to vanish.

Lemma 5.10. *Let F be a profree $\mathbb{A}^*$-module. Then $\varprojlim{}^j_\lambda \; F \otimes_{\mathbb{A}^*} \mathbb{A}^*(W_\lambda) = 0$ for $j > 0$.*

Proof. Since F is a direct summand of some product of (suspensions of) $\mathbb{A}^*$ by [Hovey and Strickland 1999, Proposition A.13], we may assume that $F = \prod_\alpha \mathbb{A}^*$. For a finite spectrum W_λ, $F \otimes \mathbb{A}^*(W_\lambda) \cong \prod_\alpha \mathbb{A}^*(W_\lambda)$ since $\mathbb{A}^*(W_\lambda)$ is a finitely presented $\mathbb{A}^*$-module. Then we have $\varprojlim{}^j_\lambda \; \prod_\alpha \mathbb{A}^*(W_\lambda) \cong \prod_\alpha \varprojlim{}^j_\lambda \; \mathbb{A}^*(W_\lambda)$. The lemma follows from the fact that $\varprojlim{}^j_\lambda \; \mathbb{A}^*(W_\lambda) = 0$ for $j > 0$ since $\mathbb{A}^*(W_\lambda)$ is a linearly compact $\mathbb{A}^*$-module for all λ. □

By Proposition 3.8, $\mathrm{Map}_c(G^{q+1}_{n+1}; \mathbb{A}^*)$ and $\mathrm{Map}_c(\mathbb{G}(i)^{q+1}, \mathbb{B}(i)^*)$ are profree $\mathbb{A}^*$-modules. Then the completion of $\varinjlim_i C^*_c(\mathbb{G}(i); \mathbb{B}(i)^*)$ at I_n is also a profree $\mathbb{A}^*$-module. By Lemma 5.10, we obtain that $\varprojlim{}^j_\lambda \; C^*_c(G_{n+1}; \mathbb{A}^*(W_\lambda)) = 0$ and $\varprojlim{}^j_\lambda \; \varinjlim_i C^*_c(\mathbb{G}(i); \mathbb{B}(i)^*(W_\lambda)) = 0$ for $j > 0$. Hence, by Lemma 5.9, we obtain two spectral sequences

$$
\begin{aligned}
{}_{I}E_2^{s,t} &= \varprojlim{}^s_\lambda \;\; H^t_c(G_{n+1}; \mathbb{A}^*(W_\lambda)) &&\implies H^*_c(G_{n+1}; \mathbb{A}^*(W)),\\
{}_{II}E_2^{s,t} &= \varprojlim{}^s_\lambda \; \varinjlim_i H^t_c(\mathbb{G}(i); \mathbb{B}(i)^*(W_\lambda)) &&\implies \boldsymbol{H}^*_c(\mathbb{G}; \mathbb{B}^*(W)).
\end{aligned}
$$

The system of cochain maps

$$\{C^*_c(G_{n+1}; \mathbb{A}^*(W_\lambda))\}_\lambda \longrightarrow \{\varinjlim_i C^*_c(\mathbb{G}(i); \mathbb{B}(i)^*(W_\lambda))\}_\lambda$$

induces a morphism of spectral sequences

$$f_r : {}_{I}E_r^{*,*} \longrightarrow {}_{II}E_r^{*,*} \tag{5-2}$$

which converges to the inflation map (5-1).

We show that this morphism of spectral sequences is an isomorphism from the E_2-terms onward. For this purpose, it is sufficient to show that the inflation map $H^*_c(G_{n+1}; \mathbb{A}^*(W)) \to H^*_c(\mathbb{G}(i); \mathbb{B}(i)^*(W))$ is an isomorphism for $i \geq 0$. We shall construct two acyclic resolutions $I^*(W)$ and $J^*(i, W)$ of $\mathbb{A}^*(W)$ with respect to $H^*_c(G_{n+1}; -)$ so that

$$I^*(W)^{G_{n+1}} \cong C^*_c(G_{n+1}; \mathbb{A}^*(W)) \quad \text{and} \quad J^*(i, W)^{G_{n+1}} \cong C^*_c(\mathbb{G}(i); \mathbb{B}(i)^*(W)).$$

We shall enlarge the complexes $C_c^*(G_{n+1};\mathbb{A}^*(W))$ and $C_c^*(\mathbb{G}(i);\mathbb{B}(i)^*(W))$ to double complexes $C_c^*(G_{n+1};I^*(W))$ and $C_c^*(G_{n+1};J(i,W))$. We shall construct a map of double complexes $C_c^*(G_{n+1};I^*(W))\to C_c^*(G_{n+1};J(i,W))$, which induces the inflation map $H_c^*(G_{n+1};\mathbb{A}^*(W))\to H_c^*(\mathbb{G}(i);\mathbb{B}(i)^*(W))$. Then we shall show that the map of double complexes induces an isomorphism on cohomology groups.

First, we construct an acyclic resolution $I^*(W)$ of $\mathbb{A}^*(W)$. We set

$$I^q(W)=\mathrm{Map}_c(G_{n+1}^{q+1},\mathbb{A}^*(W))$$

the topological $\mathbb{A}^*$-module of all continuous maps from G_{n+1}^{q+1} to $\mathbb{A}^*(W)$. Define a map $d^q: I^q(W)\to I^{q+1}(W)$ by

$$d^q(f)(g_0,\dots,g_{q+1})=\sum_{j=0}^{q+1}(-1)^j f(g_0,\dots,g_{j-1},g_{j+1},\dots,g_{q+1}).$$

Then $I^*(W)=\{I^q(W),d^q\}_{q\geq -1}$ forms an augmented cochain complex satisfying $I^{-1}(W)=\mathbb{A}^*(W)$. The group G_{n+1} acts on the cochain complex $I^*(W)$ and

$$I^*(W)^{G_{n+1}}\cong C_c^*(G_{n+1};\mathbb{A}^*(W)).$$

Lemma 5.11. *For $p>0$ and $q\geq 0$, we have*

$$H_c^p(G_{n+1};I^q(W))=0 \quad \text{and} \quad H_c^0(G_{n+1};I^q(W))=C_c^q(G_{n+1};\mathbb{A}^*(W)).$$

The sequence $0\to\mathbb{A}^*(W)\xrightarrow{d^{-1}} I^0(W)\xrightarrow{d^1} I^1(W)\xrightarrow{d^2}\cdots$ *is a split exact sequence of topological $\mathbb{A}^*$-modules. Hence $I^*(W)$ is an acyclic resolution of $\mathbb{A}^*(W)$ with respect to $H_c^*(G_{n+1};-)$.*

Proof. Since $I^q(W)=\mathrm{Map}_c(G_{n+1}^{q+1},\mathbb{A}^*(W))$, the first assertion is a consequence of Corollary 5.5. We define $s^q: I^q(W)\to I^{q-1}(W)$ by $s^q(f)(g_0,\dots,g_{q-1})=f(e,g_0,\dots,g_{q-1})$. Then we can verify that $\{s^q\}_{q\geq 0}$ gives a desired splitting. $\square$

Next we construct another acyclic resolution $J^*(i,W)$ of $\mathbb{A}^*(W)$. We set

$$J^q(i,W)=\mathrm{Map}_c(\mathbb{G}(i)^{q+1},\mathbb{B}(i)^*(W))^{S_n(i)}.$$

the topological $\mathbb{A}^*$-module of all $S_n(i)$-equivariant continuous maps from $\mathbb{G}(i)^{q+1}$ to $\mathbb{B}(i)^*(W)$. Define a map $d^q: J^q(i,W)\to J^{q+1}(i,W)$ by

$$d^q f(g_0,\dots,g_{p+1})=\sum_{j=0}^{p+1}(-1)^j f(g_0,\dots,g_{j-1},g_j,\dots,g_{p+1}).$$

Then $J^*(i,W)=\{J^q(i,W),d^q\}_{q\geq -1}$ forms an augmented cochain complex with $J^{-1}(i,W)=\mathbb{A}^*(W)$. The group G_{n+1} acts on $J^*(i,W)$ and

$$J^*(i,W)^{G_{n+1}}\cong C_c^*(\mathbb{G}(i);\mathbb{B}(i)^*(W)).$$

We compare $J^*(i,W)$ with $I^*(W)$. Let $D^* = C^*(S_n(i); \mathbb{B}(i)^*)$ be the cochain complex of $S_n(i)$ with coefficients in $\mathbb{B}(i)^*$. Since $\mathbb{A}^* \to \mathbb{B}(i)^*$ is a Galois extension with Galois group $S_n(i)$, there is an isomorphism $D^q \cong \mathbb{B}(i)^{*\otimes(q+1)}$. Then the differential $d^q : D^q \to D^{q+1}$ corresponds to $d^q : \mathbb{B}(i)^{*(q+1)} \to \mathbb{B}(i)^{*(q+2)}$ given by

$$d^q(b_0 \otimes \cdots \otimes b_q) = \sum_{j=0}^{q} (-1)^j b_0 \otimes \cdots \otimes b_{j-1} \otimes 1 \otimes b_j \otimes \cdots \otimes b_q$$

for $b_0, \dots, b_q \in \mathbb{B}(i)^*$. Since $\mathbb{G}(i) \cong G_{n+1} \times S_n(i)$ as an $S_n(i)$-space, and D^q is a finitely generated free $\mathbb{A}^*$-module, we see that $J^q(i,W) \cong I^q(W) \otimes D^q$. Then the differential $d^q : J^q(i,W) \to J^{q+1}(i,W)$ corresponds to

$$d^q : I^q(i,W) \otimes \mathbb{B}(i)^{*\otimes(q+1)} \to I^{q+1}(i,W) \otimes \mathbb{B}(i)^{*\otimes(q+2)}$$

given by

$$\begin{aligned}
&d^q(f \otimes b_0 \otimes \cdots \otimes b_q)(g_0, \dots, q_{q+1}) \\
&= \sum_{j=0}^{q+1} (-1)^j f(g_0, \dots, g_{j-1}, g_{j+1}, \dots, g_{q+1}) \otimes b_0 \otimes \cdots \otimes b_{j-1} \otimes 1 \otimes b_j \otimes \cdots \otimes b_q.
\end{aligned}$$

Proposition 5.12. *For $p > 0$ and $q \geq 0$, we have*

$$H_c^p(G_{n+1}; J^q(i,W)) = 0 \quad \textit{and} \quad H_c^0(G_{n+1}; J^q(i,W)) = C_c^q(G_{n+1}; \mathbb{A}^*(W)).$$

The sequence $0 \to \mathbb{A}^*(W) \xrightarrow{d^{-1}} J^0(i,W) \xrightarrow{d^0} J^1(i,W) \xrightarrow{d^2} \cdots$ is a split exact sequence of topological $\mathbb{A}^*$-modules. Hence $J^*(i,W)$ is an acyclic resolution of $\mathbb{A}^*(W)$ with respect to $H_c^*(G_{n+1}; -)$.

Proof. Let $M = \mathrm{Map}(S_n(i)^q, \mathbb{B}(i)^*(W))$. We have an isomorphism $J^q(i,W) \cong \mathrm{Map}_c(G_{n+1}^{q+1}, M)$ of topological G_{n+1}-modules. Since M is a product of finite many copies of $\mathbb{A}^*(W)$, we can write $M = \varprojlim_\lambda \mathbb{A}^* N_\lambda$ with finitely generated N_λ. Then the first assertion follows from Proposition 5.4. There is a continuous map $\varepsilon : \mathbb{B}^*(i) \to \mathbb{A}^*$ of topological $\mathbb{A}^*$-modules such that $\varepsilon \circ \eta = 1$, where $\eta : \mathbb{A}^* \to \mathbb{B}^*(i)$ is the unit. Define a map $s^q : I^q(i,W) \otimes \mathbb{B}(i)^{*\otimes(q+1)} \to I^{q-1}(i,W) \otimes \mathbb{B}(i)^{*\otimes q}$ by

$$s^q(f \otimes b_0 \otimes \cdots \otimes b_q)(g_0, \dots, g_{q-1}) = f(e, g_0, \dots, g_{q-1}) \otimes \varepsilon(b_0) b_1 \otimes \cdots \otimes b_q.$$

Then we can verify that $\{s^q\}_{q \geq 0}$ gives a desired splitting. □

We consider the double complexes $C_c^*(G_{n+1}; I^*(W))$ and $C_c^*(G_{n+1}; J^*(i,W))$. The canonical inclusion $\mathbb{A}^*(W) \to \mathbb{B}(i)^*(W)$ and the projection $\mathbb{G}(i) \to G_{n+1}$ induce a cochain map $I^*(W) \to J^*(i,W)$, which is equivariant under the actions of G_{n+1}. Hence we obtain a map of double complexes

$$C_c^*(G_{n+1}; I^*(W)) \longrightarrow C_c^*(G_{n+1}; J^*(i,W)). \tag{5-3}$$

We denote by $\mathrm{Tot}^* C^{*,*}$ the total cochain complex of a double complex $C^{*,*}$.

Lemma 5.13. *The cochain map*

$$\mathrm{Tot}^* C_c^*(G_{n+1};\, I^*(W)) \to \mathrm{Tot}^* C_c^*(G_{n+1};\, J^*(i,W))$$

is a quasi-isomorphism.

Proof. This follows from the fact that the map (5-3) induces an isomorphism on cohomology groups on the second index by Lemma 5.11 and Proposition 5.12. □

Since the invariant subcomplex $I^*(W)^{G_{n+1}}$ is isomorphic to $C_c^*(G_{n+1};\, \mathbb{A}^*(W))$, there is a cochain map

$$C_c^*(G_{n+1};\, \mathbb{A}^*(W)) \longrightarrow \mathrm{Tot}^* C_c^*(G_{n+1};\, I^*(W)).$$

Since the invariant subcomplex $J^*(i,W)^{G_{n+1}}$ is isomorphic to $C_c^*(\mathbb{G}(i);\, \mathbb{B}(i)^*(W))$, there is a cochain map

$$C_c^*(\mathbb{G}(i);\, \mathbb{B}(i)^*(W)) \longrightarrow \mathrm{Tot}^* C_c^*(G_{n+1};\, J^*(i,W)).$$

Then we obtain the commutative diagram of cochain complexes

$$\begin{tikzcd} C_c^*(G_{n+1};\, \mathbb{A}^*(W)) \arrow[r] \arrow[d] & C_c^*(\mathbb{G}(i);\, \mathbb{B}(i)^*(W)) \arrow[d] \\ \mathrm{Tot}^* C_c^*(G_{n+1};\, I^*(W)) \arrow[r] & \mathrm{Tot}^* C_c^*(G_{n+1};\, J^*(i,W)), \end{tikzcd} \tag{5-4}$$

where the top horizontal arrow induces the inflation map

$$H_c^*(G_{n+1};\, \mathbb{A}^*(W)) \longrightarrow H_c^*(\mathbb{G}(i);\, \mathbb{B}(i)^*(W)).$$

Lemma 5.14. *The vertical arrows in the diagram* (5-4) *are quasi-isomorphisms.*

Proof. By Lemma 5.11, the cohomology group of $C_c^*(G_{n+1};\, I^*(W))$ on the first index is isomorphic to $C_c^*(G_{n+1};\, \mathbb{A}^*(W))$. Hence the left vertical arrow is a quasi-isomorphism. By Proposition 5.12, the cohomology group of $C_c^*(G_{n+1};\, J^*(i,W))$ on the first index is isomorphic to $C_c^*(\mathbb{G}(i);\, \mathbb{B}(i)^*(W))$. Hence the right vertical arrow is a quasi-isomorphism. □

Corollary 5.15. *The inflation map $H_c^*(G_{n+1};\, \mathbb{A}^*(W)) \longrightarrow H_c^*(\mathbb{G}(i);\, \mathbb{B}(i)^*(W))$ is an isomorphism for any spectrum W and any $i \geq 0$.*

Proof of Theorem 5.8. Corollary 5.15 implies that the morphism (5-2) of spectral sequences is an isomorphism from the E_2-terms onward. Hence the inflation map (5-1) is an isomorphism. □

Remark 5.16. Let Λ be an essentially small cofiltered category. For a system $\{N_\lambda\}_{\lambda\in\Lambda}$ of finitely generated twisted E_{n+1}^*-G_{n+1}-modules, we set $M = \varprojlim_\lambda \mathbb{A}^* N_\lambda$ and $\mathbb{B}^* M = \varprojlim_\lambda \mathbb{B}^* \otimes_{\mathbb{A}^*} \mathbb{A}^* N_\lambda$. By the same method as above, we can define $\boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^* M)$ and show that there is an isomorphism

$$H_c^*(G_{n+1}; M) \xrightarrow{\cong} \boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^* M).$$

6. Morphism of spectral sequences

In this section we construct a natural morphism of spectral sequences from the $K(n)$-local E_n-Adams spectral sequence to the $K(n)$-localization of the $K(n+1)$-local E_{n+1}-Adams spectral sequence.

Let BP be the Brown–Peterson spectrum at p. We denote by $BP^{\wedge s}$ the smash product of s copies of BP:

$$BP^{\wedge s} = \overbrace{BP \wedge \cdots \wedge BP}^{s}.$$

The commutative ring spectrum structure on BP makes $BP^{\wedge\bullet+1} = \{BP^{\wedge s+1}\}_{s\geq 0}$ a cosimplicial object in the p-local stable homotopy category with augmentation $S^0_{(p)} \xrightarrow{\varepsilon} BP^{\wedge\bullet+1}$. Then the associated cochain complex

$$* \to S^0_{(p)} \xrightarrow{\varepsilon} BP \xrightarrow{d} BP^{\wedge 2} \xrightarrow{d} BP^{\wedge 3} \xrightarrow{d} \cdots \tag{6-1}$$

is a p-local BP-Adams resolution of $S^0_{(p)}$ in the sense of [Miller 1981; Devinatz and Hopkins 2004]. We denote by $\mathrm{Res}(BP; S^0_{(p)})$ the sequence (6-1). Then $\mathrm{Res}(BP; S^0_{(p)})$ gives us a diagram of exact triangles

$$\begin{tikzcd}[column sep=small]
S^0_{(p)} = X^0 \arrow[dr, "j"'] & & X^1 \arrow[ll, "i"'] \arrow[dr, "j"'] & & X^2 \arrow[ll, "i"'] \arrow[dr, "j"'] & & X^3 \arrow[ll, "i"'] & \cdots, \\
& BP \arrow[ur, "k"'] & & \Sigma^{-1} BP^{\wedge 2} \arrow[ur, "k"'] & & \Sigma^{-2} BP^{\wedge 3} \arrow[ur, "k"'] & &
\end{tikzcd} \tag{6-2}$$

where k has degree -1 and $jk = d$. We denote by $\mathrm{Ad}(BP; S^0_{(p)})$ the diagram of exact triangles (6-2).

By applying the $K(n)$-localization functor to the augmented cosimplicial commutative ring spectrum $S^0_{(p)} \xrightarrow{\varepsilon} BP^{\wedge\bullet+1}$, we obtain an augmented cosimplicial $K(n)$-local commutative ring spectrum $L_{K(n)}S^0 \xrightarrow{\varepsilon} L_{K(n)}BP^{\wedge\bullet+1}$, and the associated augmented cochain complex

$$* \to L_{K(n)}S^0 \xrightarrow{\varepsilon} L_{K(n)}BP \xrightarrow{d} L_{K(n)}BP^{\wedge 2} \xrightarrow{d} L_{K(n)}BP^{\wedge 3} \xrightarrow{d} \cdots . \tag{6-3}$$

We denote by $L_{K(n)}\mathrm{Res}(BP;\, S^0_{(p)})$ the sequence (6-3).

Proposition 6.1. *The sequence $L_{K(n)}\mathrm{Res}(BP;\, S^0_{(p)})$ is a $K(n)$-local E_n-Adams resolution of $L_{K(n)}S^0$.*

Proof. To prove the proposition, it suffices to show that $L_{K(n)}BP^{\wedge s}$ is E_n-injective for $s > 0$ and the sequence (6-3) is E_n-exact. By [Hovey and Sadofsky 1999, Theorem B], $L_{K(n)}BP$ is a coproduct of (suspensions of) $L_{K(n)}E(n)$'s in the $K(n)$-local category. Since $L_{K(n)}E(n)$ is a direct summand of E_n, $L_{K(n)}BP$ is E_n-injective. Hence $L_{K(n)}BP^{\wedge s}$ is E_n-injective for $s > 0$. To prove that the sequence (6-3) is E_n-exact, it is sufficient to show that the sequence (6-3) smashing with E_n is a split exact sequence. There is a canonical ring spectrum map $\eta : L_{K(n)}BP \to E_n$. Then the following map

$$L_{K(n)}(E_n \wedge BP^{\wedge s+1}) \xrightarrow{1\wedge\eta\wedge 1^{\wedge s}} L_{K(n)}(E_n \wedge E_n \wedge BP^{\wedge s}) \xrightarrow{m\wedge 1^{\wedge s}} L_{K(n)}(E_n \wedge BP^{\wedge s})$$

for $s \geq 0$ gives a splitting, where m is the multiplication of E_n. □

The $K(n)$-localization functor gives a map of cosimplicial objects $BP^{\bullet+1} \to E_n^{\bullet+1}$ covering the map $S^0_{(p)} \to L_{K(n)}S^0$. This induces a map

$$L_{K(n)}\mathrm{Res}(BP;\, S^0_{(p)}) \to \mathrm{Res}(E_n;\, L_{K(n)}S^0)$$

of cochain complexes and a map $L_{K(n)}\mathrm{Ad}(BP;\, S^0) \to \mathrm{Ad}(E_n;\, L_{K(n)}S^0)$ of diagrams of exact triangles. By Proposition 6.1, the map

$$L_{K(n)}\mathrm{Res}(BP;\, S^0_{(p)}) \to \mathrm{Res}(E_n;\, L_{K(n)}S^0)$$

is a cochain homotopy equivalence. Hence $L_{K(n)}\mathrm{Ad}(BP;\, S^0) \to \mathrm{Ad}(E_n;\, L_{K(n)}S^0)$ is an equivalence of diagram of exact triangles in an appropriate sense.

The canonical ring spectrum map $BP \to E_{n+1}$ induces a map of diagrams of exact triangles

$$\mathrm{Ad}(BP;\, S^0_{(p)}) \longrightarrow L_{K(n+1)}\mathrm{Ad}(BP;\, S^0_{(p)}) \xrightarrow{\simeq} \mathrm{Ad}(E_{n+1};\, L_{K(n+1)}S^0).$$

By applying the $K(n)$-localization functor to this map, we obtain a map of diagrams of exact triangles

$$L_{K(n)}\mathrm{Ad}(BP;\, S^0_{(p)}) \longrightarrow L_{K(n)}\mathrm{Ad}(E_{n+1};\, L_{K(n+1)}S^0).$$

Then this map of exact triangles implies the following theorem.

Theorem 6.2. *For any spectrum W, there is a natural morphism of spectral sequences*

$$\varphi_r(W) : L_{K(n)}E_r^{s,t}(W) \longrightarrow L_{K(n)}L_{K(n+1)}E_r^{s,t}(W),$$

which converges to $[W, L_{K(n)}S^0]^* \to [W, \widehat{\mathbb{S}}]^*$.

7. The inflation map

In Section 6 we constructed a natural morphism

$$\varphi_r(W) : L_{K(n)}E_r^{*,*}(W) \to L_{K(n)}L_{K(n+1)}E_r^{*,*}(W)$$

of spectral sequences for any spectrum W. In this section we construct a natural map $\theta(W) : H_c^*(G_n; E_n^*(W)) \to H_c^*(G_{n+1}; \mathbb{A}^*(W))$ by using the cohomology group $\boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ in Section 5. Then we show that $\theta(W)$ coincides with $\varphi_2(W)$.

For a spectrum W, define cochain complexes $C_{BP}^{*,*}(W)$ and $L_{K(n)}C_{BP}^{*,*}(W)$ by

$$\begin{aligned} C_{BP}^{s,*}(W) &= [W, BP^{\wedge s+1}]^*, \\ L_{K(n)}C_{BP}^{s,*}(W) &= [W, L_{K(n)}(BP^{\wedge s+1})]^*. \end{aligned}$$

The ring spectrum maps $BP \to L_{K(n)}BP \to E_n$ induce cochain maps

$$C_{BP}^{*,*}(W) \to L_{K(n)}C_{BP}^{*,*}(W) \to C_c^*(G_n; E_n^*(W)).$$

We shall describe the cochain map $C_{BP}^{*,*}(W) \to C_c^*(G_n; E_n^*(W))$ in terms of formal group laws. The universal deformation F_n over E_n^0 induces a graded ring homomorphism $BP_* \to E_{n*}$. Recall that, for $g = (\gamma, s) \in \Gamma \ltimes S_n = G_n$, there is a unique isomorphism $t(g) : F_n \to F_n^g$ over E_n^0, which is a lifting of the isomorphism $s : H_n \to H_n^\gamma = H_n$ over $\boldsymbol{F}$. For $g, h \in G_n$, we set $t(g,h) = t(h) \circ t(g)^{-1} : F_n^g \to F_n^h$. For a sequence $\boldsymbol{g} = (g_0, g_1, \ldots, g_s)$ of elements in G_n, we define a graded ring homomorphism

$$t(\boldsymbol{g}) : BP_*(BP)^{\otimes(s+1)} \longrightarrow E_{n*}$$

to be the map representing the following string of isomorphisms of formal group laws

$$F_n \xrightarrow{t(g_0)} F_n^{g_0} \xrightarrow{t(g_0,g_1)} F_n^{g_1} \xrightarrow{t(g_1,g_2)} \cdots \xrightarrow{t(g_{s-1},g_s)} F_n^{g_s}.$$

For a spectrum W, we denote by $\mathrm{ev}(\boldsymbol{g}) : C_c^s(G_n; E_n^*(W)) \to E_n^*(W)$ the evaluation map at $\boldsymbol{g} = (g_0, g_1, \ldots, g_s)$. If W is a finite spectrum, we denote its S-dual by DW. Then there are natural isomorphisms $BP^{-*}(W) \cong BP_*(DW)$ and $E_n^{-*}(W) \cong E_{n*}(DW) \cong BP_*(DW) \otimes_{BP_*} E_{n*}$. In particular, we have

$$C_{BP}^{s,-*}(W) \cong BP_*(DW) \otimes_{BP_*} BP_*(BP)^{\otimes s}.$$

Lemma 7.1. *Let W be a finite spectrum. For a sequence $\boldsymbol{g} = (g_0, g_1, \ldots, g_s)$ of elements in G_n, the composition* $C_{BP}^{s,-*}(W) \longrightarrow C_c^s(G_n; E_n^{-*}(W)) \xrightarrow{\mathrm{ev}(\boldsymbol{g})} E_n^{-*}(W)$

is given by

$$\begin{aligned} BP_*(DW)\otimes_{BP_*} BP_*(BP)^{\otimes s} &\xrightarrow{\psi\otimes 1^{\otimes s}} BP_*(DW)\otimes_{BP_*} BP_*(BP)^{\otimes(s+1)} \\ &\xrightarrow{1\otimes t(\boldsymbol{g})} BP_*(DW)\otimes_{BP_*} E_{n*}, \end{aligned}$$

where ψ is the $BP_*(BP)$-comodule structure map of $BP_*(DW)$.

Proof. For $g \in G_n$, the ring spectrum map $g : E_n \to E_n$ induces a map $g^{-*} : E_n^{-*}(W) \to E_n^{-*}(W)$. This map g^{-*} is given by the composition

$$\begin{aligned} BP_*(DW)\otimes_{BP_*} E_{n*} &\xrightarrow{\psi\otimes 1} BP_*(DW)\otimes_{BP_*} BP_*(BP)\otimes_{BP_*} E_{n*} \\ &\xrightarrow{1\otimes t(g)\otimes g_*} BP_*(DW)\otimes_{BP_*} E_{n*}. \end{aligned}$$

Next we consider the map $g_0\wedge\cdots\wedge g_s : E_n^{\wedge s+1} \to E_n^{\wedge s+1}$. This induces a map $(g_0\wedge\cdots\wedge g_s)^{-*} : (E_n^{\wedge s+1})^{-*}(W) \to (E_n^{\wedge s+1})^{-*}(W)$. Note that there is a natural isomorphism $(E_n^{\wedge s+1})^{-*}(W) \cong BP_*(DW)\otimes_{BP_*} \pi_*E_n^{\wedge s+1}$ since $\pi_*E_n^{\wedge s+1}$ is Landweber exact. Then $(g_0\wedge\cdots\wedge g_s)^{-*}$ is given by

$$\begin{aligned} BP_*(DW)\otimes_{BP_*} \pi_*E_n^{\wedge s+1} &\xrightarrow{\psi\otimes 1} BP_*(DW)\otimes_{BP_*} BP_*(BP)\otimes_{BP_*} \pi_*E_n^{\wedge s+1} \\ &\xrightarrow{1\otimes t(g_0)\otimes \pi_*(g_0\wedge\cdots\wedge g_s)} BP_*(DW)\otimes_{BP_*} E_{n*}\otimes_{E_{n*}} \pi_*E_n^{\wedge s+1} \\ &\cong BP_*(DW)\otimes_{BP_*} \pi_*E_n^{\wedge s+1}. \end{aligned}$$

The lemma follows from the fact that the composition

$$C_{BP}^{s,-*}(W) \longrightarrow C_c^s(G_n; E_n^{-*}(W)) \xrightarrow{\mathrm{ev}(\boldsymbol{g})} E_n^{-*}(W)$$

is induced by the map $BP^{\wedge s+1} \to E_n^{\wedge s+1} \xrightarrow{g_0\wedge\cdots\wedge g_s} E_n^{\wedge s+1} \xrightarrow{m} E_n$, where m is the multiplication map of the ring spectrum E_n. □

Next we construct a cochain map $C_c^*(G_n; E_n^*(W)) \longrightarrow \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))$, which induces a map $H_c^*(G_n; E_n^*(W)) \longrightarrow \boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W))$.

Lemma 7.2. *The ring spectrum map $I : E_n \to \mathbb{B}$ and the projection $\mathbb{G} \to G_n$ induce a cochain map $C_c^*(G_n; E_n^*(W)) \longrightarrow \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ for any spectrum W.*

Proof. There are isomorphisms

$$\begin{aligned} C_c^*(G_n; E_n^*(W)) &\cong \varprojlim_\lambda \varinjlim_i C_c^*(G(i), E_n^*(W_\lambda)), \\ \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W)) &\cong \varprojlim_\lambda \varinjlim_i C_c^*(\mathbb{G}(i), \mathbb{B}(i)^*(W_\lambda)). \end{aligned}$$

Then the canonical maps $E_n^*(W_\lambda) \to \mathbb{B}(i)^*(W_\lambda)$ and the projections $\mathbb{G}(i) \to G_n(i)$ induce the desired cochain map. □

Remark 7.3. Let Λ be an essentially small cofiltered category. For a system $\{N_\lambda\}_{\lambda\in\Lambda}$ of finitely generated twisted E_n^*-G_n-modules annihilated by a power of the ideal I_n, we set $N = \varprojlim_\lambda N_\lambda$ and $\mathbb{B}^*N = \varprojlim_\lambda \mathbb{B}^* \otimes_{E_n^*} N_\lambda$. By the same method as above, we can obtain a cochain map $C_c^*(G_n; N) \to \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*N)$.

Recall that in Section 5 we defined a cochain map $C_c^*(G_{n+1}; \mathbb{A}^*(W)) \longrightarrow \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))$, which induces an isomorphism of cohomology groups

$$H_c^*(G_{n+1}; \mathbb{A}^*(W)) \xrightarrow{\cong} \boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W))$$

by Theorem 5.8. We define a map

$$\theta(W) : H_c^*(G_n; E_n^*(W)) \longrightarrow H_c^*(G_{n+1}; \mathbb{A}^*(W)) \tag{7-1}$$

by the composition

$$H_c^*(G_n; E_n^*(W)) \longrightarrow \boldsymbol{H}_c^*(\mathbb{G}; \mathbb{B}^*(W)) \xleftarrow{\cong} H_c^*(G_{n+1}; \mathbb{A}^*(W)),$$

where the first map is induced by the cochain map in Lemma 7.2.

In the rest of this section we compare $\theta(W)$ to $\varphi_2(W)$. The ring spectrum maps $BP \to L_{K(n)}BP \to L_{K(n)}E_{n+1} = \mathbb{A}$ induce cochain maps

$$C_{BP}^{*,*}(W) \to L_{K(n)}C_{BP}^{*,*}(W) \to C_c^*(G_{n+1}; \mathbb{A}^*(W)).$$

We consider the following diagram of cochain complexes

$$\begin{array}{ccc} C_{BP}^{*,*}(W) & \longrightarrow & C_c^*(G_{n+1}; \mathbb{A}^*(W)) \\ \downarrow & & \downarrow \\ C_c^*(G_n; E_n^*(W)) & \longrightarrow & \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W)). \end{array} \tag{7-2}$$

This diagram is not commutative but we shall show that it is cochain homotopy commutative for finite spectra W by constructing a natural cochain homotopy.

Lemma 7.4. *If W is a finite spectrum, then the diagram* (7-2) *is cochain homotopy commutative.*

Proof. Let $\pi : \mathbb{G} \to G_n$ be the projection. For $g, h \in \mathbb{G}$, we have an isomorphism of formal group laws $t(\pi(g), \pi(h)) : F_n^{\pi(g)} \to F_n^{\pi(h)}$ over E_n^0. If we regard $t(\pi(g), \pi(h))$ as a power series over $\mathbb{B}^0$, then we obtain an isomorphism of formal group laws $t(g, h) : F_n^g \to F_n^h$ over $\mathbb{B}^0$. In the same way we obtain an isomorphism of formal group laws $u(g, h) : F_{n+1}^g \to F_{n+1}^h$ over $\mathbb{B}^0$. Recall that there is an isomorphism of formal group laws $\Phi : F_{n+1} \to F_n$ over $\mathbb{B}^0$. For a sequence $\boldsymbol{g} = (g_0, g_1, \ldots, g_s)$ of elements in $\mathbb{G}$, consider the following diagram of formal groups laws and isomorphisms over $\mathbb{B}^0$

$$\begin{array}{l} F_{n+1} \xrightarrow{u(g_0)} F_{n+1}^{g_0} \xrightarrow{u(g_0,g_1)} F_{n+1}^{g_1} \longrightarrow \cdots \longrightarrow F_{n+1}^{g_i} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \big\downarrow \Phi^{g_i} \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad F_n^{g_i} \xrightarrow{t(g_i,g_{i+1})} F_n^{g_{i+1}} \longrightarrow \cdots \longrightarrow F_n^{g_s}. \end{array}$$

This diagram induces a graded ring homomorphism $T_i(\boldsymbol{g}) : BP_*(BP)^{\otimes(s+2)} \to \mathbb{B}_*$. We fix an isomorphism between $\mathbb{B}^{-*}(W)$ and $BP_*(DW) \otimes_{BP_*} \mathbb{B}_*$, where $\mathbb{B}_*$ is a BP_*-module through the graded ring homomorphism $BP_* \to \mathbb{B}_*$ classifying the p-typical formal group law F_{n+1}. We define a map $C_{BP}^{s+1,-*}(W) \to \mathbb{B}^{-*}(W)$ by

$$\begin{aligned} BP_*(DW) \otimes_{BP_*} BP_*(BP)^{\otimes(s+1)} &\xrightarrow{\psi\otimes 1^{\otimes(s+1)}} BP_*(DW) \otimes_{BP_*} BP_*(BP)^{\otimes(s+2)} \\ &\xrightarrow{1\otimes T_i(\boldsymbol{g})} BP_*(DW) \otimes_{BP_*} \mathbb{B}_*. \end{aligned}$$

This map extends to a map

$$S_i : C_{BP}^{s+1,*}(W) \longrightarrow \varinjlim_i \mathrm{Map}_c(\mathbb{G}(i)^{s+1}, \mathbb{B}(i)^*(W))^{\mathbb{G}(i)} = \boldsymbol{C}_c^s(\mathbb{G}; \mathbb{B}^*(W)).$$

We shall verify that $\sum_{i=0}^s (-1)^i S_i$ is a desired cochain homotopy. First note that the map $E_n^{-*}(W) \to \mathbb{B}^{-*}(W) \cong BP_*(DW) \otimes_{BP_*} \mathbb{B}_*$ is given by

$$\begin{aligned} BP_*(DW) \otimes_{BP_*} E_{n*} &\xrightarrow{\psi\otimes 1} BP_*(DW) \otimes_{BP_*} BP_*(BP) \otimes_{BP_*} E_{n*} \\ &\xrightarrow{1\otimes\Phi\otimes I_*} BP_*(DW) \otimes_{BP_*} \mathbb{B}_*, \end{aligned}$$

where $\Phi : BP_*(BP) \to \mathbb{B}_*$ is the graded ring homomorphism classifying the isomorphism $\Phi : F_{n+1} \to F_n$, and $I_* : E_{n*} \to \mathbb{B}_*$ is the induced map by the ring spectrum map I. Let a^* be the cochain map $C_{BP}^{*,*}(W) \to C_c^*(G_n; E_n^*(W)) \to \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))$ and let b^* be the cochain map $C_{BP}^{*,*}(W) \to C_c^*(G_{n+1}; E_{n+1}^*(W)) \to \boldsymbol{C}_c^*(\mathbb{G}; \mathbb{B}^*(W))$. We see that $\mathrm{ev}(\boldsymbol{g}) \circ a^s$ is given by

$$\begin{aligned} BP_*(DW) \otimes_{BP_*} BP_*(BP)^{\otimes s} &\xrightarrow{\psi\otimes 1^{\otimes s}} BP_*(DW) \otimes_{BP_*} BP_*(BP)^{\otimes(s+1)} \\ &\xrightarrow{1\otimes U(\boldsymbol{g})} BP_*(DW) \otimes_{BP_*} \mathbb{B}_*, \end{aligned}$$

where $U(\boldsymbol{g})$ is the graded ring homomorphism classifying the following string of isomorphisms of formal group laws

$$F_{n+1} \xrightarrow{t(g_0)\circ\Phi} F_n^{g_0} \xrightarrow{t(g_0,g_1)} F_n^{g_1} \xrightarrow{t(g_1,g_2)} \cdots \xrightarrow{t(g_{s-1},g_s)} F_n^{g_s}.$$

In the cosimplicial module $C^{\bullet,*}_{BP}(W)$, the map $d_i : C^{s,-*}_{BP}(W) \to C^{s+1,-*}_{BP}(W)$ is given by

$$d_i = \begin{cases} \psi \otimes 1^{\otimes s} & \text{if } i = 0, \\ 1 \otimes 1^{\otimes(i-1)} \otimes \Delta \otimes 1^{\otimes(s-i)} & \text{if } 1 \le i \le s, \\ 1 \otimes 1^{\otimes s} \otimes \eta_L & \text{if } i = s+1, \end{cases}$$

where $\Delta : BP_*(BP) \to BP_*(BP)^{\otimes 2}$ is the comultiplication, and $\eta_L : BP_* \to BP_*(BP)$ is the left unit. Then we see that

$$\begin{aligned} S_0 \circ d_0 &= a^s, \\ S_i \circ d_j &= d_j \circ S_{i-1} \quad \text{for } 0 \le j < i \le s, \\ S_{i-1} \circ d_i &= S_i \circ d_i \qquad \text{for } 0 < i \le s, \\ S_i \circ d_j &= d_{j-1} \circ S_i \quad \text{for } 0 \le i < j-1 \le s, \\ S_s \circ d_{s+1} &= b^s. \end{aligned}$$

This implies that

$$\sum_{i=0}^{s} (-1)^i S_i \circ \sum_{j=0}^{s+1} (-1)^j d_j + \sum_{j=0}^{s} (-1)^j d_j \circ \sum_{i=0}^{s-1} (-1)^i S_i = a^s - b^s.$$

This completes the proof. □

For a spectrum W, we have a similar diagram of cochain complexes

$$\begin{array}{ccc} L_{K(n)} C^{*,*}_{BP}(W) & \longrightarrow & C^*_c(G_{n+1}; \mathbb{A}^*(W)) \\ \big\downarrow & & \big\downarrow \\ C^*_c(G_n; E^*_n(W)) & \longrightarrow & \boldsymbol{C}^*_c(\mathbb{G}; \mathbb{B}^*(W)). \end{array} \tag{7-3}$$

When W is a finite spectrum, we let $S(W) : C^{*,*}_{BP}(W) \to \boldsymbol{C}^{*-1}_c(\mathbb{G}; \mathbb{B}^*(W))$ be the cochain homotopy constructed in the proof of Lemma 7.4. Then $S(W)$ extends to a cochain homotopy $L_{K(n)} S(W) : L_{K(n)} C^{*,*}_{BP}(W) \to \boldsymbol{C}^{*-1}_c(\mathbb{G}; \mathbb{B}^*(W))$, which makes the diagram (7-3) homotopy commutative.

Proposition 7.5. *For any spectrum W, the diagram* (7-3) *is cochain homotopy commutative.*

Proof. Since the cochain homotopy $L_{K(n)} S(W)$ is natural for finite spectra W, we obtain a cochain homotopy

$$\varprojlim_\lambda L_{K(n)} S(W_\lambda) : \\ \varprojlim_\lambda L_{K(n)} C^{*,*}_{BP}(W_\lambda) \longrightarrow \varprojlim_\lambda \boldsymbol{C}^{*-1}_c(\mathbb{G}; \mathbb{B}^*(W_\lambda)) = \boldsymbol{C}^{*-1}_c(\mathbb{G}; \mathbb{B}^*(W)),$$

where the inverse limits are taken over $\lambda \in \Lambda(W)$. Then the composition with the cochain map $L_{K(n)}C_{BP}^{*,*}(W) \longrightarrow \varprojlim_{\lambda} L_{K(n)}C_{BP}^{*,*}(W_\lambda)$ makes the diagram (7-3) cochain homotopy commutative. □

Theorem 7.6. *The map*

$$\theta(W) : H_c^*(G_n; E_n^*(W)) \to H_c^*(G_{n+1}; E_{n+1}^*(W))$$

coincides with the map $\varphi_2(W)$ for any spectrum W.

Proof. In the diagram (7-3) the left vertical arrow is a quasi-isomorphism by Proposition 6.1. So is the right vertical arrow, by Theorem 5.8. The theorem follows because the top horizontal arrow induces the map $\varphi_2(W)$ and the bottom horizontal arrow induces the map $\theta(W)$. □

8. Nontriviality of the image of ζ_n

In this section we prove Theorem 8.1 as an application of the results in this note. By the Hopkins–Miller theorem [Devinatz and Hopkins 2004, Theorem 6], we know that there exists a nontrivial element $\zeta_n \in \pi_{-1}(L_{K(n)}S^0)$, which is represented by the reduced norm map of G_n in the E_2-term of the $K(n)$-local E_n-Adams spectral sequence. The $K(n)$-localization of the $K(n+1)$-localization map $S^0 \to L_{K(n+1)}S^0$ induces a map $L_{K(n)}S^0 \to L_{K(n)}L_{K(n+1)}S^0$. In this section we show that the image of ζ_n under the map $\pi_*(L_{K(n)}S^0) \to \pi_*(L_{K(n)}L_{K(n+1)}S^0)$ is nontrivial as an application of Theorems 4.7 and 5.8.

By Theorem 6.2, we have a morphism of spectral sequences

$$\varphi_r = \varphi_r(S^0) : L_{K(n)}E_r^{*,*}(S^0) \longrightarrow L_{K(n)}L_{K(n+1)}E_r^{*,*}(S^0),$$

which converges to $\pi_*(L_{K(n)}S^0) \to \pi_*(L_{K(n)}L_{K(n+1)}S^0)$. Then φ_2 is identified with the inflation map

$$\theta = \theta(S^0) : H_c^*(G_n; E_n^*) \longrightarrow H_c^*(G_{n+1}; \mathbb{A}^*)$$

by Theorem 5.8. The reduced norm map of G_n defines an element $z_n \in H_c^1(G_n; E_n^0)$ which represents $\zeta_n \in \pi_{-1}(L_{K(n)}S^0)$. We set $w_n = \theta(z_n) \in H_c^1(G_{n+1}; \mathbb{A}^0)$, and denote by ω_n the image of ζ_n under the map $\pi_*(L_{K(n)}S^0) \to \pi_*(L_{K(n)}L_{K(n+1)}S^0)$. Then w_n is a permanent cycle and it represents ω_n.

Theorem 8.1. $\omega_n \in \pi_{-1}(L_{K(n)}L_{K(n+1)}S^0)$ *is nontrivial.*

Proof. In [Torii 2003] we constructed a map

$$\theta' : H_c^*(G_n; \boldsymbol{F}[w^{\pm 1}]) \longrightarrow H_c^*(G_{n+1}; \boldsymbol{F}((u_n))[u^{\pm 1}]).$$

Then there exists a commutative diagram

$$\begin{tikzcd} H_c^*(G_n; E_n^*) \arrow[r, "\theta"] \arrow[d, "\pi"'] & H_c^*(G_{n+1}; \mathbb{A}^*) \arrow[d, "\pi"] \\ H_c^*(G_n; \boldsymbol{F}[w^{\pm 1}]) \arrow[r, "\theta'"] & H_c^*(G_{n+1}; \boldsymbol{F}((u_n))[u^{\pm 1}]), \end{tikzcd}$$

where the vertical arrows π are canonical reduction maps. In [Torii 2005] we calculated the image of $\theta' : H_c^1(G_n; \boldsymbol{F}[w^{\pm 1}]) \to H_c^1(G_{n+1}; \boldsymbol{F}((u_n))[u^{\pm 1}])$, and we showed that $\theta'(\pi(z_n))$ is nontrivial. This implies that $\theta(z_n) \in H_c^1(G_{n+1}; \mathbb{A}^0)$ is nontrivial. Since $\theta(z_n)$ is a permanent cycle and lies in the 1-line of the spectral sequence, it represents a nontrivial element in $\pi_{-1}(L_{K(n)} L_{K(n+1)} S^0)$. □

References

[Boardman 1999] J. M. Boardman, "Conditionally convergent spectral sequences", pp. 49–84 in *Homotopy invariant algebraic structures* (Baltimore, MD, 1998), edited by J.-P. Meyer et al., Contemporary Mathematics **239**, AMS, Providence, RI, 1999. MR 2000m:55024 Zbl 0947.55020

[Bousfield and Kan 1972] A. K. Bousfield and D. M. Kan, *Homotopy limits, completions and localizations*, Lecture Notes in Mathematics **304**, Springer, Berlin, 1972. MR 51 #1825 Zbl 0259.55004

[Chase et al. 1965] S. U. Chase, D. K. Harrison, and A. Rosenberg, "Galois theory and Galois cohomology of commutative rings", *Mem. Am. Math. Soc.* **52** (1965), 15–33. MR 33 #4118 Zbl 0143.05902

[Devinatz and Hopkins 2004] E. S. Devinatz and M. J. Hopkins, "Homotopy fixed point spectra for closed subgroups of the Morava stabilizer groups", *Topol.* **43**:1 (2004), 1–47. MR 2004i:55012 Zbl 1047.55004

[Greither 1992] C. Greither, *Cyclic Galois extensions of commutative rings*, Lecture Notes in Mathematics **1534**, Springer, Berlin, 1992. MR 94k:11121 Zbl 0788.13003

[Hovey 1995] M. A. Hovey, "Bousfield localization functors and Hopkins' chromatic splitting conjecture", pp. 225–250 in *The Čech centennial* (Boston, 1993), edited by M. Cenkl and H. R. Miller, Contemporary Mathematics **181**, AMS, Providence, RI, 1995. MR 96m:55010 Zbl 0830.55004

[Hovey 1997] M. A. Hovey, "v_n-Elements in ring spectra and applications to bordism theory", *Duke Math. J.* **88**:2 (1997), 327–356. MR 98d:55017 Zbl 0880.55006

[Hovey and Sadofsky 1999] M. A. Hovey and H. Sadofsky, "Invertible spectra in the $E(n)$-local stable homotopy category", *J. Lond. Math. Soc.* (2) **60**:1 (1999), 284–302. MR 2000h:55017 Zbl 0947.55013

[Hovey and Strickland 1999] M. A. Hovey and N. P. Strickland, *Morava K-theories and localisation*, Mem. Am. Math. Soc. **139**:666, 1999. MR 99b:55017 Zbl 0929.55010

[Hovey et al. 1997] M. A. Hovey, J. H. Palmieri, and N. P. Strickland, *Axiomatic stable homotopy theory*, Mem. Am. Math. Soc. **128**:610, 1997. MR 98a:55017 Zbl 0881.55001

[Lam 1999] T. Y. Lam, *Lectures on modules and rings*, Graduate Texts in Mathematics **189**, Springer, New York, 1999. MR 99i:16001 Zbl 0911.16001

[Miller 1981] H. R. Miller, "On relations between Adams spectral sequences, with an application to the stable homotopy of a Moore space", *J. Pure Appl. Algebra* **20**:3 (1981), 287–312. MR 82f:55029 Zbl 0459.55012

[Milne 1980] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series **33**, Princeton University Press, Princeton, NJ, 1980. MR 81j:14002 Zbl 0433.14012

[Morava 1985] J. Morava, "Noetherian localisations of categories of cobordism comodules", *Ann. Math.* (2) **121**:1 (1985), 1–39. MR 86g:55004 Zbl 0572.55005

[Ravenel 1992] D. C. Ravenel, *Nilpotence and periodicity in stable homotopy theory*, Ann. Math. Studies **128**, Princeton University Press, Princeton, NJ, 1992. MR 94b:55015 Zbl 0774.55001

[Serre 1994] J.-P. Serre, *Cohomologie galoisienne*, 5th ed., Lecture Notes in Math. **5**, Springer, Berlin, 1994. Translated in *Galois cohomology*, Springer, Berlin, 2002. MR 96b:12010 Zbl 0812.12002

[Torii 2003] T. Torii, "On degeneration of one-dimensional formal group laws and applications to stable homotopy theory", *Am. J. Math.* **125**:5 (2003), 1037–1077. MR 2004i:55006 Zbl 1046.55003

[Torii 2005] T. Torii, "On relations between 1-lines of Adams–Novikov spectral sequences modulo invariant prime ideals", *Topol. Appl.* **150**:1-3 (2005), 33–57. MR 2005k:55026 Zbl 1069.55002

[Torii 2010a] T. Torii, "Comparison of Morava E-theories", *Math. Zeitschrift* **266**:4 (2010), 933–951. MR 2729298 Zbl 1203.55004

[Torii 2010b] T. Torii, "On E_∞-structure of the generalized Chern character", *Bull. Lond. Math. Soc.* **42**:4 (2010), 680–690. MR 2669689 Zbl 05770958

Received April 3, 2009. Revised December 17, 2009.

TAKESHI TORII
DEPARTMENT OF MATHEMATICS
OKAYAMA UNIVERSITY
OKAYAMA 700-8530
JAPAN

torii@math.okayama-u.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

THOMPSON'S GROUP IS DISTORTED IN THE THOMPSON–STEIN GROUPS

CLAIRE WLADIS

We show that the inclusion map of the generalized Thompson groups $F(n_i)$ is exponentially distorted in the Thompson–Stein groups $F(n_1, \ldots, n_k)$ for $k > 1$. One consequence is that F is exponentially distorted in $F(n_1, \ldots, n_k)$ for $k > 1$ whenever $n_i = 2^m$ for some m (whenever no i, m exist such that $n_i = 2^m$, there is no obviously "natural" inclusion map of F into $F(n_1, \ldots, n_k)$). This is the first known example in which the natural embedding of one of the Thompson-type groups into another is not quasi-isometric.

1. Introduction

In this paper, we use some of the motivating ideas behind the proofs of the metric properties developed in [Wladis 2009] to show that the inclusion map of the generalized Thompson groups $F(n_i)$ into $F(n_1, \ldots, n_k)$ is exponentially distorted for $k > 1$. A quasi-isometric embedding of a subgroup into a larger group induces a metric on the subgroup that is equivalent to subgroup metric. In contrast, when an embedding is not quasi-isometric, the subgroup distortion measures the extent to which this metric is distorted by the embedding map (for formal definitions, see Section 4).

We give here the first known example of the natural embedding of one Thompson-type group being distorted inside another. Burillo, Cleary and Stein [Burillo et al. 2001] showed that $F(n)$ is quasi-isometrically embedded into $F(m)$ for all $n, m \in \mathbb{N} - \{1\}$, and along with Taback, that F is quasi-isometrically embedded in Thompson's group T [Burillo et al. 2009]. Different methods have been used to show that $F^n \times \mathbb{Z}^m$ is quasi-isometrically embedded in F for all $m, n \in \mathbb{N}$ [Burillo 1999; Cleary and Taback 2003; Guba and Sapir 1999; Guba and Sapir 1997]. Since the development of the main theorem of this paper, Burillo and Cleary [2010] have

This work was supported in part by a grant from The City University of New York PSC-CUNY Research Award Program and the CUNY Scholar Incentive Award. The author would also like to thank the Technische Universität Berlin for its hospitality during the writing of this paper.
MSC2000: 20F65.
Keywords: Thompson's group, piecewise linear homeomorphism, Stein group, Higman group, quasi-isometrically embedded subgroup, distorted subgroup.

used similar methods as those described here to prove that the canonical embeddings of Thompson's groups F and V are also distorted in the higher dimensional Thompson's group nV.

Robert Thompson introduced the three groups named after him in the early 1960s (see [McKenzie and Thompson 1973]). Denoted by $F \subset T \subset V$, they have provided many interesting group-theoretic counterexamples: T and V were the first known infinite, simple, finitely presented groups, and F was the first known example of a torsion-free infinite-dimensional FP_∞ group. For more information see [Cannon et al. 1996].

The groups $F(n_1, \dots, n_k)$, generalizing F, were first explored in depth by Melanie Stein [1992]. Related explorations of general classes in this family of groups, each of which can be considered to be a generalization of the Thompson groups, include [Higman 1974; Brown and Geoghegan 1984; Brown 1987; Brin and Guzmán 1998; Brin and Squier 2001; Bieri and Strebel 1985].

Definition 1.1. The *Thompson–Stein group* $F(n_1, \dots, n_k)$, where $k \in \mathbb{N}$ and $n_1, \dots, n_k \in \{2, 3, 4, \dots\}$ are pairwise relatively prime, is the group of piecewise linear orientation-preserving homeomorphisms of the closed unit interval with finitely many breakpoints in $\mathbb{Z}[\frac{1}{n_1 \cdots n_k}]$ and slopes in the group $\langle n_1, n_2, \dots, n_k \rangle$ in each linear piece. We abbreviate $F(2)$ by F.

Stein [1992] explored the homological and simplicity properties of $F(n_1, \dots, n_k)$ and showed that they are of type FP_∞ and finitely presented, and gave a technique for computing infinite and finite presentations. In [Wladis 2009], using Stein's presentations, we developed the theory of tree-pair diagram representation for elements of $F(n_1, \dots, n_k)$, gave a unique normal form, and calculated sharp upper and lower bounds on the metric in terms of the number of leaves in the minimal tree-pair diagram representative. The proofs in this paper use the normal form results and some of the same motivating ideas behind the metric approximations used in our 2009 paper.

The results of this article hold for all groups of the form $F(n_1, \dots, n_k)$ that satisfy the condition $n_1 - 1 \mid n_j - 1$ for all $j \in \{1, \dots, k\}$; throughout this paper, when we refer to the group $F(n_1, \dots, n_k)$, this divisibility criterion will be implied. Groups not satisfying this criterion will have a significantly different group presentation, and therefore require alternate normal form and metric techniques than those presented here or in [Wladis 2009]. Much of the introductory material in this paper is summarized from that paper, where more detail can be found.

2. Representing elements using tree-pair diagrams

The proofs in this paper depend heavily on the representation of elements of $F(m)$ and $F(n_1, \dots, n_k)$ by tree-pair diagrams; see [Wladis 2007; 2009] for more details.

Definition 2.1. An *n-ary caret*, or caret of type n, is a graph which has $n+1$ vertices joined by n edges: one vertex has degree n (the parent) and the rest have degree 1 (the children).

An *$(n_1, \dots, n_k)$-ary tree* is a graph formed by joining a finite number of carets by identifying the child vertex of one caret with the parent vertex of another so that every caret in the tree has a type in $\{n_1, \dots, n_k\}$. An *$(n_1, \dots, n_k)$-ary tree-pair diagram* is an ordered pair of $(n_1, \dots, n_k)$-ary trees with the same number of leaves.

If a vertex in a tree has degree 1, it is referred to as a *leaf*.

An $(n_1, \dots, n_k)$-ary tree represents a subdivision of $[0, 1]$ using the following recursive process, which assigns a subinterval of $[0, 1]$ to each leaf in the tree: the root vertex represents the interval $[0, 1]$; for a given n-ary caret in the tree with parent vertex representing $[a, b]$, the n child vertices represent the subintervals $\left[a, a+\frac{1}{n}\right], \left[a+\frac{1}{n}, a+\frac{2}{n}\right], \dots, \left[b-\frac{1}{n}, b\right]$ respectively.

Every element of $F(n_1, \dots, n_k)$ can be represented by an $(n_1, \dots, n_k)$-ary tree-pair diagram and vice versa. We number the leaves in a tree beginning with zero, in increasing order from left to right; a leaf's placement in this order is determined by the relative position of the subinterval within $[0, 1]$ which it represents. Once the leaves of each tree in a tree-pair diagram are numbered, then the element of $F(n_1, \dots, n_k)$ which it represents is the map which takes the subinterval of $[0, 1]$ represented by the ith leaf in the domain tree to the subinterval of $[0, 1]$ represented by the ith leaf in the range tree. Because every element of $F(n_1, \dots, n_k)$ is a piecewise linear map with fixed endpoints, it can be determined solely by the ordered subintervals in the domain and range. For example, the element given in Figure 1 is just the map $\left\{\left[0, \frac{1}{2}\right], \left[\frac{1}{2}, \frac{3}{4}\right], \left[\frac{3}{4}, 1\right]\right\} \to \left\{\left[0, \frac{1}{3}\right], \left[\frac{1}{3}, \frac{2}{3}\right], \left[\frac{2}{3}, 1\right]\right\}$.

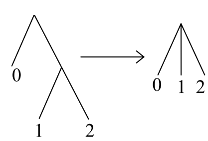


Figure 1. An example element of $F(2, 3)$.

Equivalence and minimality of tree-pair diagrams. We will analyze properties of $F(m)$ and $F(n_1, \dots, n_k)$ by identifying each group element with an equivalence class of tree-pair diagrams, so we must have criteria for equivalence. And because our metric is based on using a minimal tree-pair diagram representative for an element, we also give minimality criteria.

Definition 2.2. Two trees are *equivalent* if they represent the same subdivision of the unit interval; two tree-pair diagrams are *equivalent* if they represent the same element of $F(n_1, \dots, n_k)$.

An *exposed caret pair* in a tree-pair diagram is a pair of carets of the same type, one in each tree, such that all the child vertices of each caret are leaves, and both sets of leaves have identical leaf index numbers. Exposed caret pairs can be canceled in a tree-pair diagram to produce an equivalent tree-pair diagram with fewer leaves. Analogously, we can add a pair of identical carets to a tree-pair diagram to the leaves with the same index number in each tree and obtain an equivalent tree-pair diagram.

Definition 2.3. An $(n_1, \ldots, n_k)$-ary tree-pair diagram is *minimal* if it has the smallest number of leaves of any tree-pair diagram in the equivalence class representing a given element of $F(n_1, \ldots, n_k)$. In $F(m)$, a tree-pair diagram is minimal if and only if it contains no exposed caret pairs.

Definition 2.4. For any given $j \in \{1, \ldots, k\}$, the n_j*-valence* of a leaf $l \in T$ is the number of n_j-ary carets which have an edge on the path from the root vertex to l; it is denoted by $v_{n_j}(l)$. If we refer to just the valence of l, or $\boldsymbol{v}(l)$, this refers to the vector $\langle v_{n_1}(l), \ldots, v_{n_k}(l)\rangle$.

Theorem 2.5 [Wladis 2009]. *The $(n_1, \ldots, n_k)$-ary trees T and S are equivalent if and only if $L(T) = L(S)$ and $\boldsymbol{v}(l_i) = \boldsymbol{v}(k_i)$ for all leaves $l_i \in T$, $k_i \in S$.*

Tree-pair diagram composition. To find ba for $b, a \in F(n_1, \ldots, n_k)$, $b = (T_-, T_+)$ and $a = (S_-, S_+)$, we need to make S_+ equivalent to T_-. This is accomplished by adding carets to T_- and S_+ (and by extension to the leaves with the same index numbers in T_+ and S_- respectively) until the valence of all leaves of both T_- and S_+ are the same. If we let T_-^*, T_+^*, S_-^* S_+^* denote T_-, T_+, S_-, S_+, respectively, after this addition of carets; then the (possibly nonminimal) product is (S_-^*, T_+^*) (see Figure 2). The process of tree-pair diagram composition always terminates; see [Wladis 2009].

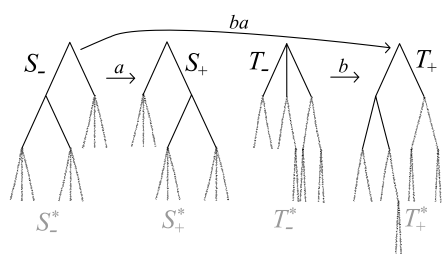


Figure 2. Composition of two elements of $F(2, 3)$. Solid lines indicate the carets present in the original elements a and b, and dotted lines indicate carets that must be added during composition. The tree-pair diagram representative of ba is the pair which contains the domain tree of a and the range tree of b, with both hatched and solid line carets included.

3. The metric in $F(n)$ and $F(n_1, \dots, n_k)$

Standard presentations. Stein [1992] gave a method for finding the finite presentations for the groups $F(n_1, \dots, n_k)$; in [Wladis 2009] we computed the exact finite presentations explicitly. For the sake of simplicity, we give the presentation for $F(2, 3)$ only here. For presentations for $F(n_1, \dots, n_k)$ more generally, see [Wladis 2009].

Theorem 3.1 [Stein 1992; Wladis 2009]. *Thompson's group $F(2, 3)$ admits the infinite presentation with generators $x_0, y_0, z_0, x_1, y_1, z_1, \dots$ and relators*

$$\gamma_j x_i = x_i \gamma_{j+1} \quad \text{and} \quad \gamma_j z_i = z_i \gamma_{j+2} \qquad \text{when } i < j \text{ for } \gamma = x, y, z;$$
$$y_{i+1} z_i = y_i x_{i+1} x_i \quad \text{and} \quad x_i z_{i+1} z_i = z_i x_{i+2} x_{i+1} x_i \quad \text{for all } i.$$

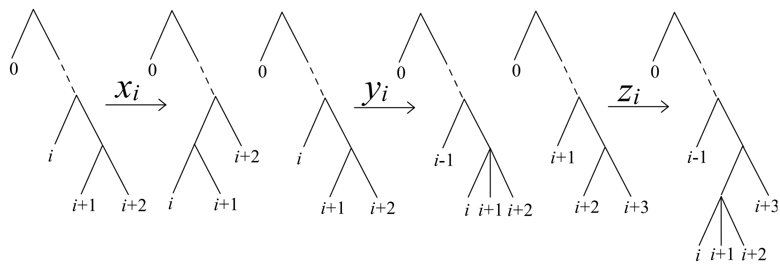


Figure 3. Infinite generators for $F(2, 3)$.

Theorem 3.2 [Stein 1992; Wladis 2009]. *$F(2, 3)$ admits the finite presentation with generators $\{x_0, x_1, y_0, y_1\}$ and relators*

$$x_2x_0 = x_0x_3, \quad y_2x_0 = x_0y_3, \quad x_1z_0 = z_0x_3, \quad y_1z_0 = z_0y_3,$$
$$x_3x_1 = x_1x_4, \quad y_3x_1 = x_1y_4, \quad x_2z_1 = z_1x_4, \quad y_2z_1 = z_1y_4,$$
$$x_0z_1z_0 = z_0x_2x_1x_0, \qquad x_1z_2z_1 = z_1x_3x_2x_1,$$

where

$$x_3 = x_1^{-1}x_2x_1, \quad y_3 = x_1^{-1}y_2x_1, \quad z_0 = y_1^{-1}y_0x_1x_0,$$
$$x_4 = x_2^{-1}x_3x_2, \quad y_4 = x_2^{-1}y_3x_2, \quad z_1 = y_2^{-1}y_1x_2x_1, \quad z_2 = y_3^{-1}y_2x_3x_2.$$

The standard presentations for F (see [Brown 1987]) are:

$$\text{Infinite: } \{x_0, x_1, x_2, \cdots \mid x_j x_i = x_i x_{j+1} \text{ for } i < j\}$$
$$\text{Finite: } \{x_0, x_1 \mid [x_0x_1^{-1}, x_0^{-1}x_1x_0], [x_0x_1^{-1}, x_0^{-2}x_1x_0^2]\}$$

The metric. It is well known that the metric in F and $F(n)$ is quasi-isometric to the number of carets (or equivalently to the number of leaves) in the minimal tree-pair diagram representative of a given group element. However, this does not hold for $F(n_1, \dots, n_k)$ when $k > 1$; it is this fact which will be exploited to show that F is distorted in $F(n_1, \dots, n_k)$.

Notation 3.3. The notation $|x|_{F(n)}$ and $|x|_{F(n_1,\ldots,n_k)}$ will be used to represent the length of the element x in $F(n)$ and $F(n_1, \ldots, n_k)$ respectively, with respect to the standard finite generating set.

Notation 3.4. The notation $L(T)$, $L(T_-, T_+)$, and $L(x)$ denotes the number of leaves in the tree T, in either tree of the tree-pair diagram (T_-, T_+), and in either tree of the minimal tree-pair diagram for x respectively.

We note that both trees in a tree-pair diagram have the same number of leaves.

Theorem 3.5 [Fordham and Cleary 2009; Burillo et al. 2001]. *For $x \in F(n)$, $|x|_{F(n)}$ is quasi-isometric to $L(x)$* (see Definition 4.1 for formal definition).

Theorem 3.6 [Wladis 2009]. *There exist fixed $B, C \in \mathbb{N}$ such that*

$$\log_B L(x) \leq |w|_{F(n_1,\ldots,n_k)} \leq CL(w) \quad \textit{for all } x \in F(n_1, \ldots, n_k).$$

These bounds are sharp.

Normal form. A unique normal form exists for $F(n_1, \ldots, n_k)$ with respect to the standard infinite presentations. This normal form essentially provides an algorithm for converting a tree-pair diagram into an algebraic expression in the normal form and vice versa. For the main proofs of this paper, we will introduce several elements for which we will give both an algebraic expression in the normal form and a tree-pair diagram representative. To understand the proofs that follow, one need only consider the tree-pair diagrams, and one need not see explicitly how the algebraic expression comes from the tree-pair diagram representative, so for the sake of space and simplicity of presentation, we have omitted a full explanation of how to write out the normal form for a given element in $F(n_1, \ldots, n_k)$; however, full details on this algorithm can be found in [Wladis 2009].

4. Quasi-isometry and subgroup distortion

A quasi-isometrically embedded subgroup has a metric that is equivalent to the induced metric within the larger group. In contrast, an embedding which is not quasi-isometric can be said to be distorted, and the type of this distortion measures the extent to which the metric is distorted by the embedding map.

Definition 4.1. The groups X and Y are *quasi-isometric* if there exist fixed $c_1, c_2 > 0$ and an embedding $f : X \to Y$ such that

$$\frac{1}{c_1}|x|_X - c_2 \leq |f(x)|_Y \leq c_1|x|_X + c_2,$$

where $|x|_X$ and $|x|_Y$ are the lengths of $x \in X$ and $x \in Y$ respectively, with respect to a fixed finite generating set. When $X \subset Y$, the embedding f will be assumed to be the inclusion map, so we often omit explicit mention of the embedding itself.

Let $x \in X \subset Y$. The *distortion function* is defined by

$$D(r) = \frac{1}{r} \max \left\{ |x|_X, |x|_Y \;\middle|\; |x|_Y < r \right\}.$$

For finitely generated groups, the distortion function is bounded if and only if the inclusion map of X into Y is a quasi-isometric embedding. When $D(r)$ is a function that grows without bound as $r \to \infty$, then we say that X is distorted in Y; the function type of $D(r)$ determines the type of the distortion (i.e. we say that a subgroup with exponential $D(r)$ is exponentially distorted). We use the notation $\sim$ to denote quasi-isometry. The property of quasi-isometry is transitive: whenever $X \sim Y$ and $Y \sim Z$, $X \sim Z$.

F is exponentially distorted in $F(n_1, \ldots, n_k)$. We begin by proving that the inclusion map of $F(n_i)$ is exponentially distorted in $F(n_1, \ldots, n_k)$ whenever there exists $j \in \{1, \ldots, k\}$ such that $n_i - 1 \mid n_j - 1$ by constructing a distorted element in $F(n_i)$ explicitly. In the next section, we generalize this result to all $i \in \{1, \ldots, l\}$.

Definition 4.2. We say that a tree is *balanced* if $\boldsymbol{v}(l_i) = \boldsymbol{v}(l_j)$ for all leaves $l_i, l_j \in T$.

Theorem 4.3. *$F(n_i)$ is exponentially distorted in $F(n_1, \ldots, n_k)$ for $k > 1$ whenever there exists n_j such that $j \in \{1, \ldots, k\}$, $i \neq j$, and $n_i - 1 \mid n_j - 1$.*

Proof. For the sake of readability, we will restrict all the explicit details of this proof to the canonical embedding of F into $F(2, 3)$ since this is the simplest case. However, this proof holds for all $F(n_i)$ that meet the stated conditions of the theorem; at key points in this proof, we will indicate what adjustments need to be made to generalize the results to the general case.

We will show that $w = y_0^{-n} x_0 y_0^n$ is such that $|w|_F \geq \frac{1}{A} 3^n$ for some $A \in \mathbb{N}$ by showing that $L(w) \geq \frac{1}{A} 3^n$. We consider the product of the representative tree-pair diagrams given in Figure 4.

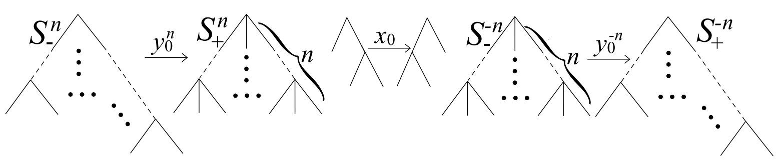


Figure 4. The product $w = y_0^{-n} x_0 y_0^n$.

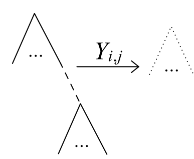
In order to perform this composition, a binary caret must be added to every leaf in S_-^n and S_+^n, to produce $(S_-^n)^1$ and $(S_+^n)^1$ respectively. Then a second binary caret must be added to the leaves with index numbers $3^n, \ldots, 2 \cdot 3^n - 1$ in both $(S_-^n)^1$ and $(S_+^n)^1$ to produce $(S_-^n)^2$ and $(S_+^n)^2$ respectively. Then a balanced n-level ternary tree (identical to S_+^n) must be added to each leaf of T_- and T_+. And finally, a binary caret must be added to each leaf in S_-^{-n} and S_+^{-n} to produce $(S_-^{-n})^1$ and $(S_+^{-n})^1$ respectively, and then another binary caret must be added to the leaves with index numbers $0, \dots, 3^n - 1$ in $(S_-^{-n})^1$ and $(S_+^{-n})^1$ to produce $(S_-^{-n})^2$ and $(S_+^{-n})^2$ respectively. It is clear then that $\big((S_-^n)^2, (S_+^{-n})^2\big)$ is a tree-pair diagram for w whose number of leaves is $2 \cdot 3^n - 1$. However, $\big((S_-^n)^2, (S_+^{-n})^2\big)$ may not be minimal. In fact, there exist exposed caret pairs in $\big((S_-^n)^2, (S_+^{-n})^2\big)$, but not enough to significantly reduce the number of leaves in the tree-pair diagram; to see this, we list the leftmost leaf index number of every exposed caret in $\big((S_-^n)^2, (S_+^{-n})^2\big)$:

$$\begin{aligned}
(S_-^n)^2 &: 0, 2, 4, \dots, 3^n - 3, \text{ (even)}\\
&\qquad \mathbf{3^n}, \mathbf{3^n+2}, \mathbf{3^n+4}, \dots, \mathbf{2 \cdot 3^n - 1}, 2 \cdot 3^n + 1, 2 \cdot 3^n + 3, \dots, 3 \cdot 3^n - 2 \text{ (odd)}\\
(S_+^n)^2 &: 0, 2, 4, \dots, 3^n - 3, \mathbf{3^n-1}, \mathbf{3^n+1}, \mathbf{3^n+3}, \dots, \mathbf{2 \cdot 3^n - 2}, \text{ (even)}\\
&\qquad 2 \cdot 3^n + 1, 2 \cdot 3^n + 3, 2 \cdot 3^n + 5, \dots, 3 \cdot 3^n - 2 \text{ (odd)}
\end{aligned}$$

It is clear that all exposed carets with leftmost leaf number in bold cannot cancel, because these leaves in the domain tree have odd index numbers and these leaves in the range tree have even index numbers. So

$$L(w) \geq (2 \cdot 3^n - 2) - (3^n - 1) = 3^n + 1,$$

and because the metric in F is quasi-isometric to the number of leaves in the minimal tree-pair diagram representative of an element, there exists $A \in \mathbb{N}$ such that $|w|_F \geq \frac{1}{A} 3^n$. However, clearly $|w|_{F(2,3)} \leq 2n + 1$.



To generalize this proof for $F(n_i)$ in $F(n_1, \dots, n_k)$, we begin by defining the element $Y_{i,j}$ as the element with tree-pair diagram of the form given on the right. (In the case $i = 1$, we simply have $Y_{i,j} = (y_j)_0$.)

We define Z_i as the element with the tree-pair diagram given in Figure 5. We consider the product

$$w_{i,j,n} = Y_{i,j}^{-n} Z_i Y_{i,j}^n$$

given in that figure the same way that we considered $y_0^n x_0 y_0^{-n}$ for F in $F(2, 3)$ in Figure 4. After adding all carets to each tree-pair diagram in Figure 5, as necessary in order for composition to take place, the resulting diagram $\big((S_-^n)^2, (S_+^{-n})^2\big)$ for $w_{i,j,n}$ will have exposed carets whose leftmost leaf index numbers are as follows,

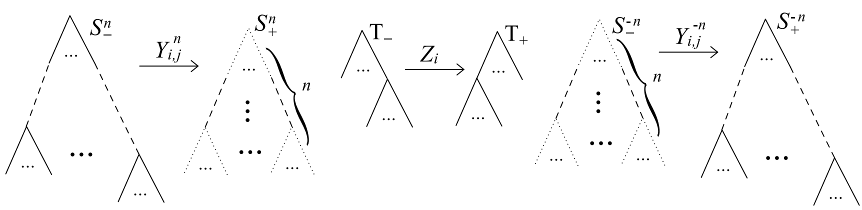


Figure 5. The product $Y_{i,j}^{-n} Z_i Y_{i,j}^n$. Solid carets are n_i-ary and dotted carets are n_j-ary (also in inset above).

where $c = \lfloor n_j^n / n_i \rfloor$ and $*$ denotes "not divisible by n_i":

$$
\begin{array}{lll}
(S_-^n)^2: & 0,\, n_i,\, 2n_i,\, 3n_i,\, \ldots,\, (c-1)n_i, & \text{(divisible by } n_i) \\
& \boldsymbol{(n_i-1)n_j^n,\, (n_i-1)n_j^n+n_i,\, (n_i-1)n_j^n+2n_i},\, \ldots, & (*) \\
& \boldsymbol{(2n_i-1)n_j^n-(c+2)n_i},\, 2(n_i-1)n_j^n-(c+1)n_i,\, \ldots,\, (2n_i-1)n_j^n-n_i & (*) \\
(S_+^n)^2: & 0,\, n_i,\, 2n_i,\, 3n_i,\, \ldots,\, (c-1)n_i,\, \boldsymbol{cn_i},\, \ldots,\, \boldsymbol{(n_j^n-1)n_i}, & \text{(divisible by } n_i) \\
& 2(n_i-1)n_j^n-(c+1)n_i,\, 2(n_i-1)n_j^n-cn_i,\, \ldots,\, (2n_i-1)n_j^n-n_i & (*)
\end{array}
$$

Because n_i and n_j are relatively prime, the carets with leftmost leaf numbers in bold will not cancel. Thus

$$
\begin{aligned}
L(w_{i,j,n}) &\geq [(2n_i-1)n_j^n-(c+2)n_i]-[cn_i] \\
&= (2n_i-1)n_j^n-(2c-2)n_i \\
&> (2n_i-3)n_j^n-2n_i \\
&> n_j^n-2n_i,
\end{aligned}
$$

the last inequality being a consequence of $cn_i < n_j^n$ and $n_i \geq 2$. However, if we let $A = |Y_{i,j}|_{F(n_1,\ldots,n_k)}$ and $B = |Z_i|_{F(n_1,\ldots,n_k)}$, we have

$$
\begin{aligned}
|w_{i,j,n}|_{F(n_1,\ldots,n_k)} &\leq |Y_{i,j}^{-n}|_{F(n_1,\ldots,n_k)} + |Z_i|_{F(n_1,\ldots,n_k)} + |Y_{i,j}^{n}|_{F(n_1,\ldots,n_k)} \\
&\leq An+B+An = 2An+B. \qquad \square
\end{aligned}
$$

$F(n_i)$ is exponentially distorted in $F(n_1,\ldots,n_k)$. We now extend the results of the last two pages to all n_i such that $i \in \{1,\ldots,k\}$. We will again do this by explicitly constructing a product in $F(n_1,\ldots,n_k)$ that produces an element in $F(n_i)$ so that the number of leaves in the product is logarithmic with respect to the number of factors in $F(n_1,\ldots,n_k)$. Without the added condition that $n_i-1 \mid n_j-1$ for some $j \in \{1,\ldots,k\}$, this product will have to be more complex than the one constructed in the last section; however, the underlying structure will be similar. We begin by defining elements of $F(n_1,\ldots,n_k)$ which will occur in our product. As in the previous section, for the sake of clarity we give our detailed proof for the embedding of $F(3)$ into $F(2,3)$, including notes indicating how this can be generalized for any $F(n_i)$ into $F(n_1,\ldots,n_k)$ that meet the conditions of Theorem 4.5.

Notation 4.4. For a fixed $i \in \{1,\ldots,k\}$ we define A_i, Z_i, λ_i as follows:

$i=2$	arbitrary i
$A_2 = x_0 y_0^{-1}$ (see Figure 9)	A_i has the form seen in Figure 7, left
$Z_2 = y_1 z_1 y_3^{-1} y_1^{-1}$ (see Figure 9)	Z_i has the form seen in Figure 7, middle
$\lambda_2 = x_0 y_1^{-1}$ (see Figure 6)	λ_i has the form seen in Figure 7, right

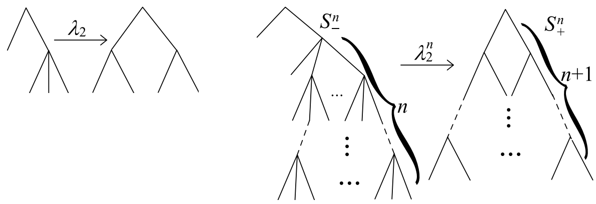


Figure 6. The elements λ_2 and λ_2^n in $F(2,3)$. Level i from the top in S_-^n has 2^{i-2} ternary carets.

For readability, the theorem and proof that follow are restricted to the case $F(2,3)$, which is illustrated in Figure 6. However, the proof can be extended to all cases by using the generalized elements given in Figure 7. Particular examples of more complicated λ_i can be seen in Figure 8.

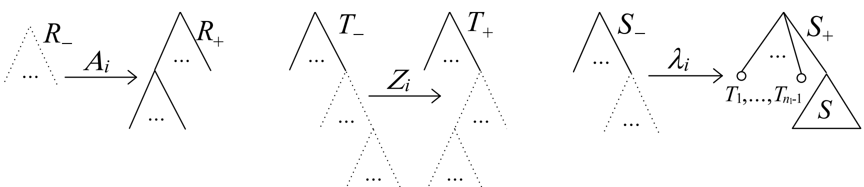


Figure 7. The elements A_i, Z_i, and λ_i in $F(n_1, \dots, n_k)$. Solid carets are n_1-ary and dotted carets are n_i-ary. On the right, S is a balanced n_1-ary tree where $L(S) \le n_i$, while $T_1, \dots, T_{n_1-1}$ are (possibly empty) n_1-ary subtrees of $D(S)$ levels or less, chosen as needed in order to make $L(S_-) = L(S_+)$. For simplicity, we fill in the subtrees $T_1, \dots, T_{n_1-1}$ from left to right, but this is not strictly necessary. For specific examples, see Figure 8.

Theorem 4.5. *The canonical embedding of $F(n_i)$ is exponentially distorted in $F(n_1, \dots, n_k)$ for all $i \in \{1, \dots, k\}$.*

Proof. We will establish this by showing that the product $W_{2,n} = (\lambda_2^n A_2)^{-1} z_2 (\lambda_2^n A_2)$ is an element of $F(3)$, and that it has a minimal tree-pair diagram representative whose number of leaves is of the order B^n for some fixed $B > 1$. All of the following steps generalize in a straightforward way to show the same result for $F(n_i)$ in $F(n_1, \dots, n_k)$ by simply replacing all the elements A_2, λ_2, Z_2 with their general formulations.

It is clear that $|W_{2,n}|_{F(2,3)} < 4n+8$ while $|W_{2,n}|_{F(3)} \sim L(W_{2,n})$. Straightforward computation of the product $W_{2,n}$, illustrated in Figure 9, shows that we must do the following:

(i) Add n levels of binary carets to each leaf in the trees T_- and T_+ of Z_2.

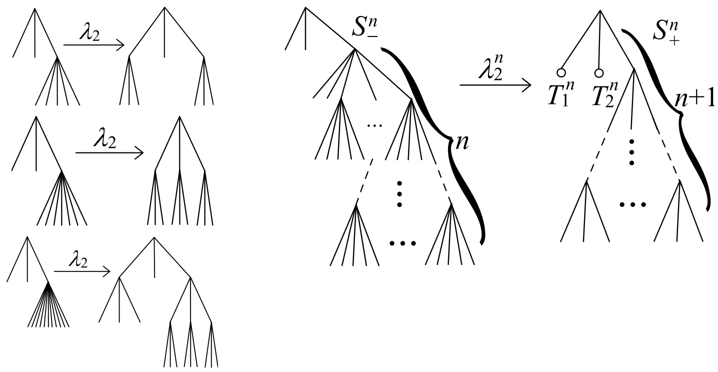


Figure 8. More complex examples of $\lambda_i \in F(n_1, \ldots, n_k)$. Left column: The element λ_2 in $F(3, 5)$, $F(3, 7)$ and $F(3, 11)$. Right: the element λ_2^n in $F(3, 5)$; level i from the top in S_-^n has 3^{i-2} quinary (5-ary) carets, and T_1^n, T_2^n are ternary subtrees.

(ii) Add a ternary caret to the 2^n rightmost leaves of S_+^n and S_-^{-n} (and by extension to the 2^n rightmost leaves of S_-^n and S_+^{-n}), and then add a ternary caret to the rightmost 2^n leaves of these added ternary carets in S_+^n (and S_-^n respectively) and to the leftmost 2^n leaves of these added ternary carets in S_-^{-n} (and S_+^{-n} respectively).

We can then see that the (not necessarily minimal) tree-pair diagram of the resulting product $\lambda_2^{-n} Z_2 \lambda_2^n$ has $3 \cdot 2^{n+1}$ leaves, and the only nonternary carets in each tree are the root carets. Conjugating this product by A_2 then produces a tree-pair diagram for $W_{2,n}$ with $(3 \cdot 2^{n+1} + 1)$ leaves consisting entirely of ternary carets (so clearly $W_{2,n} \in F(3)$).

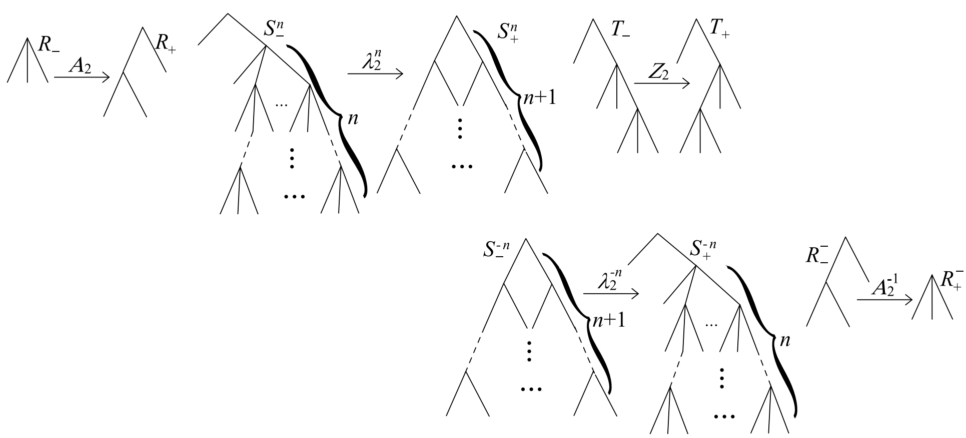


Figure 9. The product $(\lambda_2^n A_2)^{-1} Z_2 (\lambda_2^n A_2)$.

Now we need only show that a significant number of these leaves will not cancel. Using a similar argument to that in the proof of Theorem 4.3 where we tracked the leaf numbers and their divisors, it is easy to show that less than 2^{n+1} leaves will cancel, so we can conclude that $L(W_{2,n}) \geq 2^{n+1}$. □

Acknowledgements. The author would like to thank José Burillo, Sean Cleary, Melanie Stein and Ashot Minasyan for helpful discussion and comments during the preparation of this article.

References

[Bieri and Strebel 1985] R. Bieri and R. Strebel, "On groups of PL homeomorphisms of the real line", Preprint, Math. Sem. der Univ. Frankfurt, Frankfurt am Main, 1985.

[Brin and Guzmán 1998] M. G. Brin and F. Guzmán, "Automorphisms of generalized Thompson groups", *J. Algebra* **203**:1 (1998), 285–348. MR 99d:20056 Zbl 0930.20039

[Brin and Squier 2001] M. G. Brin and C. C. Squier, "Presentations, conjugacy, roots, and centralizers in groups of piecewise linear homeomorphisms of the real line", *Comm. Algebra* **29**:10 (2001), 4557–4596. MR 2002h:57047 Zbl 0986.57025

[Brown 1987] K. S. Brown, "Finiteness properties of groups", *J. Pure Appl. Algebra* **44**:1-3 (1987), 45–75. MR 88m:20110 Zbl 0613.20033

[Brown and Geoghegan 1984] K. S. Brown and R. Geoghegan, "An infinite-dimensional torsion-free FP_∞ group", *Invent. Math.* **77**:2 (1984), 367–381. MR 85m:20073 Zbl 0557.55009

[Burillo 1999] J. Burillo, "Quasi-isometrically embedded subgroups of Thompson's group F", *J. Algebra* **212**:1 (1999), 65–78. MR 99m:20051 Zbl 0951.20028

[Burillo and Cleary 2010] J. Burillo and S. Cleary, "Metric properties of higher-dimensional Thompson's groups", *Pacific J. Math.* **248**:1 (2010), 49–62. MR 2734164 Zbl 05809054

[Burillo et al. 2001] J. Burillo, S. Cleary, and M. I. Stein, "Metrics and embeddings of generalizations of Thompson's group F", *Trans. Amer. Math. Soc.* **353**:4 (2001), 1677–1689. MR 2001k:20087 Zbl 0989.20030

[Burillo et al. 2009] J. Burillo, S. Cleary, M. I. Stein, and J. Taback, "Combinatorial and metric properties of Thompson's group T", *Trans. Amer. Math. Soc.* **361**:2 (2009), 631–652. MR 2010e:20061 Zbl 1196.20048

[Cannon et al. 1996] J. W. Cannon, W. J. Floyd, and W. R. Parry, "Introductory notes on Richard Thompson's groups", *Enseign. Math.* (2) **42**:3-4 (1996), 215–256. MR 98g:20058 Zbl 0880.20027

[Cleary and Taback 2003] S. Cleary and J. Taback, "Geometric quasi-isometric embeddings into Thompson's group F", *New York J. Math.* **9** (2003), 141–148. MR 2004k:20083 Zbl 1038.20026

[Fordham and Cleary 2009] S. B. Fordham and S. Cleary, "Minimal length elements of Thompson's groups $F(p)$", *Geom. Dedicata* **141** (2009), 163–180. MR 2010e:20064 Zbl 1196.20049

[Guba and Sapir 1997] V. S. Guba and M. V. Sapir, *Diagram groups*, Mem. Amer. Math. Soc. **130**:620, Amer. Math. Soc., Providence, RI, 1997. MR 98f:20013 Zbl 0930.20033

[Guba and Sapir 1999] V. S. Guba and M. V. Sapir, "On subgroups of the R. Thompson group F and other diagram groups", *Mat. Sb.* **190**:8 (1999), 3–60. In Russian; translated in *Sb. Math.* **190**:7–8 (1999), 1077–1130. MR 2001m:20045 Zbl 1095.20021

[Higman 1974] G. Higman, *Finitely presented infinite simple groups*, Notes on Pure Mathematics **8**, Department of Pure Mathematics, Department of Mathematics, I.A.S., Australian National University, Canberra, 1974. MR 51 #13049

[McKenzie and Thompson 1973] R. McKenzie and R. J. Thompson, "An elementary construction of unsolvable word problems in group theory", pp. 457–478 in *Word problems: decision problems and the Burnside problem in group theory* (Irvine, CA, 1969), edited by W. W. Boone et al., Studies in Logic and the Foundations of Math. **71**, North-Holland, Amsterdam, 1973. MR 53 #629 Zbl 0286.02047

[Stein 1992] M. I. Stein, "Groups of piecewise linear homeomorphisms", *Trans. Amer. Math. Soc.* **332**:2 (1992), 477–514. MR 92k:20075 Zbl 0798.20025

[Wladis 2007] C. Wladis, "Thompson's group $F(n)$ is not minimally almost convex", *New York J. Math.* **13** (2007), 437–481. MR 2009c:20074 Zbl 1174.20013

[Wladis 2009] C. Wladis, "The word problem and the metric for the Thompson–Stein groups", Preprint, 2009. To appear in the J. London Math. Soc. arXiv 0811.3036

Received February 7, 2010.

CLAIRE WLADIS
DEPARTMENT OF MATHEMATICS
BOROUGH OF MANHATTAN COMMUNITY COLLEGE
CITY UNIVERSITY OF NEW YORK
199 CHAMBERS ST.
NEW YORK, NY 10007
UNITED STATES

cwladis@gmail.com
http://www.cwladis.com/math

PACIFIC JOURNAL OF MATHEMATICS
Vol. 250, No. 2, 2011

PARABOLIC MEROMORPHIC FUNCTIONS

ZHENG JIAN-HUA

Following the definition of parabolic rational functions and in view of the behavior of transcendental meromorphic functions, we give the definition of parabolic transcendental meromorphic functions. We discuss their dynamical behavior and prove the existence of conformal measures and invariant measures over their Julia sets, thus extending Denker and Urbański's work on parabolic rational functions. However, our method for proving the existence of the conformal measures differs in that we use the Perron–Frobenius–Ruelle operator.

1. Introduction and notations

Let $f(z)$ be a meromorphic function that is transcendental or rational with degree at least two. Let $f^n(z)$ be the n-th iterate of $f(z)$, let $\mathcal{F}(f)$ be the Fatou set of $f(z)$, and let $\hat{\mathcal{J}}(f)=\hat{\mathbb{C}}\setminus\mathcal{F}(f)$, which is the Julia set of $f(z)$. If f is transcendental, then $\infty\in\hat{\mathcal{J}}(f)$, and set $\mathcal{J}(f)=\hat{\mathcal{J}}(f)\setminus\{\infty\}$ and $\mathcal{J}_\infty(f)=\bigcup_{n=0}^{\infty}f^{-n}(\infty)$. If $\mathcal{J}_\infty(f)$ contains at least three points, then $\mathcal{J}(f)=\overline{\mathcal{J}_\infty(f)}$ and so f is analytic on $\mathcal{F}(f)$. $\mathcal{F}(f)$ is open and consists of at most a countable number of components, which are called Fatou components. Since $\mathcal{F}(f)$ is completely invariant, the image of every Fatou component under f is contained in a Fatou component. A Fatou component U is called periodic if $f^m(U)\subset U$ for some $m\geq 1$ and the least such m is called its period; U is preperiodic if $f^m(U)$ is periodic for some $m\geq 1$ but U is not periodic; U is wandering if $f^n(U)\cap f^m(U)=\varnothing$ for $m\neq n$. The periodic Fatou components are classified into five types: attracting domain, parabolic domain, Siegel disk, Herman ring and Baker domain. The Baker domain and wandering domain are possible only for transcendental meromorphic functions.

By $\operatorname{sing}(f^{-1})$ we mean the closure of the set of all finite critical and asymptotic values of $f(z)$ in the complex plane $\mathbb{C}$ and by $\widehat{\operatorname{sing}}(f^{-1})$ the closure of the set of all critical and asymptotic values of $f(z)$ in the extended complex plane $\hat{\mathbb{C}}=\mathbb{C}\cup\{\infty\}$. Hence if $f(z)$ has multiple poles, then ∞ is a critical value of $f(z)$

Supported by the NSF of China, grant number 10871108.
MSC2000: primary 30B20, 30D05, 30D10; secondary 34M05, 37F10.
Keywords: parabolic meromorphic functions, expansive map, conformal measures, pressure functions, Poincaré exponents.

and $\infty \in \widehat{\text{sing}}(f^{-1})$. If ∞ is an asymptotic value of $f(z)$, then $\infty \in \widehat{\text{sing}}(f^{-1})$, but in any case, $\infty \notin \text{sing}(f^{-1})$. Then $\infty \notin \widehat{\text{sing}}(f^{-1})$ if and only if $f(z)$ has no multiple poles and no ∞ as an asymptotic value and ∞ is not a limit point of finite singular values of $f(z)$. We denote by $\mathcal{P}(f)$ the postsingular set defined to be the closure in $\hat{\mathbb{C}}$ of

$$\bigcup_{n=0}^{\infty} f^n\Big(\text{sing}(f^{-1}) \setminus \bigcup_{j=0}^{n-1} f^{-j}(\infty)\Big)$$

and set $\hat{\mathcal{P}}(f) = \mathcal{P}(f) \cup \widehat{\text{sing}}(f^{-1})$.

In [Zheng 2008] we proved that for a hyperbolic meromorphic function on the complex plane, the Hausdorff dimension of the radial Julia set $\mathcal{J}_r(f)$ is equal to the Poincaré exponent $s(f)$ of f over $\mathcal{J}(f)$. Actually, the proof showed that

$$\dim_h \mathcal{J}(f) = \dim_H \mathcal{J}_r(f) = s(f),$$

where $\dim_h \mathcal{J}(f)$ is the hyperbolic dimension of $\mathcal{J}(f)$. The first equality above was proved in [Rempe 2009] for the general case. For a hyperbolic meromorphic function on the Riemann sphere, the author proved that

$$\dim_h \mathcal{J}(f) = \dim_H \mathcal{J}(f) = \lambda(f) = s(f),$$

where $\lambda(f)$ is the exponent of conformal measure of f over $\mathcal{J}(f)$, and there exists the invariant Gibbs measure that is equivalent to the $\lambda(f)$-conformal measure which extends the results in [Kotus and Urbański 2002]. Here, we say a probability measure μ over $\mathcal{J}(f)$ is a s-conformal measure for f if $f^{\times}(z)^s$ is the Jacobian of f over $\mathcal{J}(f)$ with respect to μ, that is, for any Borel subset A of $\mathcal{J}(f)$ such that f is injective on A, we have

$$\mu(f(A)) = \int_A f^{\times}(z)^s \mathrm{d}\mu.$$

In this paper, we investigate parabolic meromorphic functions. The papers [Denker and Urbański 1991a; 1991b; Aaronson et al. 1993] are careful investigations of the Hausdorff dimension, conformal measure and invariant measure of parabolic rational functions. The definition of a parabolic rational function is clear: we know that a rational function f with degree at least two is called parabolic if $\hat{\mathcal{J}}(f) \cap \widehat{\text{sing}}(f^{-1}) = \varnothing$ and f has at least one rational indifferent periodic point. However, the transcendental case is more complicated.

Definition 1.1. Let f be a transcendental meromorphic function in $\mathbb{C}$. We say that f is *parabolic on the complex plane* if $\mathcal{P}(f) \cap \mathcal{J}(f)$ is finite and nonempty, each point in $\mathcal{P}(f) \cap \mathcal{J}(f)$ is a rational indifferent periodic point of f, and $\text{sing}(f^{-1})$ is contained in $\mathcal{F}(f)$. We say f is *parabolic on the Riemann sphere* (or *with respect to the spherical metric*) if f is parabolic on the complex plane and $\infty \notin \hat{\mathcal{P}}(f)$.

We denote by $\mathcal{P}(\mathbb{C})$ and $\mathcal{P}(\hat{\mathbb{C}})$ the set of all parabolic transcendental meromorphic functions on $\mathbb{C}$ and $\hat{\mathbb{C}}$, respectively. A rational function has only a finite number of rational indifferent periodic points, while a transcendental meromorphic function may have infinitely many rational indifferent periodic points. Since every rational indifferent periodic point must be in $\mathcal{P}(f)\cap\mathcal{J}(f)$, a parabolic meromorphic function on the complex plane has only finitely many rational indifferent periodic points. In Definition 1.1, we need to stress the condition that $\mathrm{sing}(f^{-1}) \subset \mathcal{F}(f)$: although a rational indifferent periodic point cannot be a critical point, it may be a critical value, and for transcendental case it may be an asymptotic value. This can be explained by considering the functions $z(z-1)^2$ and ze^z. The point 0 is a rational indifferent fixed point of $z(z-1)^2$, which is also a critical value, and of ze^z, which is also an asymptotic value. The functions $z(z-1)^2$ and ze^z satisfy the conditions for parabolicity on the complex plane (Definition 1.1) except for the requirement that $\mathrm{sing}(f^{-1}) \subset \mathcal{F}(f)$. Hence they are not parabolic on the complex plane. If $\infty \notin \hat{\mathcal{P}}(f)$, then f is of bounded type, that is, in class $\mathcal{B}$. Clearly, a parabolic meromorphic function on the Riemann sphere is in class $\mathcal{B}$, that is, $\mathcal{P}(\hat{\mathbb{C}}) \subset \mathcal{B}$.

Let f be a transcendental meromorphic function in class $\mathcal{S}$, so that $\mathrm{sing}(f^{-1})$ is finite. If $\mathrm{sing}(f^{-1}) \subset \mathcal{F}(f)$ (resp. $\widehat{\mathrm{sing}}(f^{-1}) \subset \mathcal{F}(f)$), then f is hyperbolic whenever it has no rational indifferent periodic points; otherwise it is parabolic on the complex plane (resp. on the Riemann sphere). This is because f has only attracting domains and/or parabolic domains. For a general case, see Theorem 3.1 and Theorem 3.2 below. A simple calculation yields that $\tan z$ is in $\mathcal{P}(\hat{\mathbb{C}})$.

In the papers cited above, Denker, Urbański, and Aaronson obtained the existence of a conformal measure and an invariant measure, and showed they are equivalent for parabolic rational functions. Using the results attained in [Zheng 2009] by developing Walters' theory, we extend some of the Denker and Urbański's results to the parabolic transcendental meromorphic function, and establish:

Theorem 1.2. *Let $f(z)$ be a parabolic meromorphic function on the Riemann sphere. Then $f(z)$ has a s-conformal measure μ_s and $P(f, s) = 0$.*

Here $P(f, t)$ is the pressure of f at t, whose definition is given in Lemma 3.8. Applying a result from [Martens 1992] we determine conditions about the existence of μ_s-equivalent, f-invariant measure:

Theorem 1.3. *Let $f(z)$ be a parabolic meromorphic function on the Riemann sphere. Assume that s-conformal measure μ_s is atomless. Then $f(z)$ has a μ_s-equivalent, f-invariant measure if for some $a \in \mathcal{J}(f) \setminus \bigcup_{n=0}^{\infty} f^{-n}(\Omega)$, where $\Omega = \mathcal{P}(f) \cap \mathcal{J}(f)$, we have*

$$\sum_{n=0}^{\infty} \mathcal{L}_s^n(\mathbb{1})(a) = \infty.$$

Here we say a measure m is f-invariant if $m(f^{-1}(A)) = m(A)$ for any Borel subset A of $\mathcal{J}(f)$. Actually, Ω is the set of all rational indifferent periodic points of $f(z)$ and $\mathcal{L}_s(\mathbb{1}) = \mathcal{L}_{-s\log f^\times}(\mathbb{1})$ is the Perron–Frobenius–Ruelle operator for $-s\log f^\times(z)$ over $\mathcal{J}(f)$ and please see the statements before Lemma 3.8 for its definition.

Question 1.4. *For $f \in \mathcal{P}(\hat{\mathbb{C}})$, is $\dim_H \mathcal{J}(f)$ always equal to s?*

We conjecture the answer is affirmative.

2. Conformal measures and expansiveness of covering maps

To discuss the existence of conformal measures of parabolic meromorphic functions, we need some results from [Zheng 2009] on the existence of conformal measures for covering maps. Let $(\hat{X}, d)$ be a compact metric space and X be an open and dense subset of $\hat{X}$ and X_0 an open and dense subset of X. For a point $x \in \hat{X}$, $B(x,\delta)$ is the ball centered at x with radius δ. $\mathscr{C}(\Lambda)$ will denote the set of all real-valued continuous functions on $\Lambda = \hat{X}$, X or X_0. Let $T : X_0 \to X$ be continuous and $\varphi \in \mathscr{C}(X_0)$.

Definition 2.1. An ordered pair (T, φ) is called *admissible* if:

(1a) For each $x \in X$, the set $T^{-1}(x)$ is at most countable.

(1b) T has the uniform covering property: there exists a $\delta > 0$ such that for each $x \in X$, $T^{-1}(B_X(x,\delta))$ can be written uniquely as a disjoint union of a finite or countable number of open subsets $A_i(x)$ ($1 \le i \le N \le \infty$) of X_0 and for each i, T is a homeomorphism of $A_i(x)$ onto $B_X(x,\delta)$, where $B_X(x,\delta) = B(x,\delta) \cap X$. For simplicity, we will call $A_i(x)$ the injective component of T^{-1} over $B_X(x,\delta)$ and δ the injectivity radius.

(1c) The inverse of T is locally uniformly continuous: $\forall \varepsilon > 0$, $\exists \delta_0$ with $0 < \delta_0 < \delta$ such that for each $x \in X$ and each $y \in X_0$ with $T(y) = x$, once $d(x,x') < \delta_0$ for $x' \in X$, we have $d(T_y^{-1}(x), T_y^{-1}(x')) < \varepsilon$, where T_y^{-1} is the branch of the inverse of T which sends x to y. That is to say, every injective component of T^{-1} over $B_X(x,\delta_0)$ has diameter less than ε.

(1d) $\varphi \in \mathscr{C}(X_0)$ is summable on X, that is to say,

$$\sup\left\{\sum_{T(y)=x} \exp\varphi(y) : x \in X\right\} < +\infty.$$

(1e) For all $\varepsilon > 0$, there exists a $0 < \delta_1 < \delta$ such that for any pair $x, x' \in X$, once $d(x,x') < \delta_1$, we have

$$\sum_{T(y)=x} \left|\exp\varphi(T_y^{-1}(x)) - \exp\varphi(T_y^{-1}(x'))\right| < \varepsilon,$$

that is, $\sum_{T(y)=x} \left|\exp\varphi(T_y^{-1}(x)) - \exp\varphi(T_y^{-1}(x'))\right| \to 0$ uniformly as $d(x, x')$ goes to 0.

We now give a condition under which (1b) implies (1c).

Lemma 2.2 [Zheng 2009, Lemma 2.1 and following remark]. *Let T satisfy* (1b) *with $X = \hat{X}$. The inverse of T is locally uniformly continuous, that is, T satisfies* (1c), *if one of following statements holds*:

(1) *For arbitrary $\varepsilon > 0$, we have a $0 < \eta \le \varepsilon$ such that for each $x \in X$, $\partial B(x, \eta) \subset X_0$.*

(2) *all limit points of $T^{-1}(x)$ for each $x \in X$ lie in $X \setminus X_0$ and* (1) *holds only for $x \in X \setminus X_0$.*

We can define for a summable function φ on X_0 the Perron–Frobenius–Ruelle operator by setting

$$\mathcal{L}_\varphi(f)(x) := \sum_{T(y)=x} f(y)\exp\varphi(y) \quad \text{for } x \in X.$$

Obviously, $\mathcal{L}_\varphi(f)(x)$ is a bounded real-valued function on X when f is a bounded real-valued function on X_0. Sometimes, we write $\mathcal{L}_{\varphi,T}$ for $\mathcal{L}_\varphi$ to emphasize T. It is obvious that T^n is a continuous mapping of $T^{-n+1}X_0$ to X. Set

$$S_n\varphi(y) = \sum_{i=0}^{n-1} \varphi(T^i(y)) \quad \text{for } y \in T^{-n+1}X_0.$$

Noting that $T^{-n+1}X_0 \subseteq X_0$, we easily deduce that

$$\mathcal{L}^n_{\varphi,T}(f)(x) = \mathcal{L}_{S_n\varphi,T^n}(f)(x) = \sum_{T^n(y)=x} f(y)\exp(S_n\varphi(y)) \quad \text{for } x \in X. \tag{2-1}$$

(Here and throughout the paper we denote by $\mathcal{L}^n_{\varphi,T}$ the n-th iterate of $\mathcal{L}_{\varphi,T}$.) We want to get the desired probability measure on $\hat{X}$ through the dual operator of the $\mathcal{L}_\varphi$ over $\mathcal{M}(\hat{X})$, here $\mathcal{M}(\hat{X})$ denotes the set of all probability measures over $\hat{X}$.

Theorem 2.3. *Let (T, φ) be admissible.*

(1) *For each fixed positive integer N, $(T^N, S_N\varphi)$ is admissible.*

(2) *$\mathcal{L}_\varphi$ can be extended to a linear operator of $\mathcal{C}(\hat{X})$ to itself, which is still denoted by $\mathcal{L}_\varphi$.*

(3) *There exists a $\mu \in \mathcal{M}(\hat{X})$ such that $\mathcal{L}^*_\varphi(\mu) = \lambda\mu$, $\lambda = \mathcal{L}^*_\varphi(\mu)(1) > 0$, where $\mathcal{L}^*_\varphi$ is the dual operator of $\mathcal{L}_\varphi$, and the following statements hold*:
   - (3a) *$\lambda\exp(-\varphi)$ is the Jacobian of T with respect to μ.*
   - (3b) *μ is positively nonsingular and nonsingular for T, that is, $\mu \circ T \ll \mu$ and $\mu \circ T^{-1} \ll \mu$.*

With the exception of part (1), this theorem is a modification of the main result from [Walters 1978] (in which the expanding property is stressed; compare [Zheng 2009, Theorem 2.1 and following remark]). Obviously, the λ in Theorem 2.3 satisfies

$$\lambda^n = \mathscr{L}^{*n}_{\varphi}(\mu)(\mathbb{1}) = \mu(\mathscr{L}^n_{\varphi}(\mathbb{1}))$$

and therefore

$$\inf_{x\in X}\{\mathscr{L}^n_{\varphi}(\mathbb{1})(x)\} \le \lambda^n \le \sup_{x\in X}\{\mathscr{L}^n_{\varphi}(\mathbb{1})(x)\}. \tag{2-2}$$

Lemma 2.4. *Let T, φ, λ and μ be as in Theorem 2.3. Assume that there exist a sequence of positive number $\{K_n\}$ such that, for any $x, x' \in X$,*

$$e^{-K_n}\mathscr{L}^n_{\varphi}(\mathbb{1})(x') \le \mathscr{L}^n_{\varphi}(\mathbb{1})(x) \le e^{K_n}\mathscr{L}^n_{\varphi}(\mathbb{1})(x') \tag{2-3}$$

and $K_n/n \to 0$ as $n \to \infty$. Then

$$\log\lambda = \lim_{n\to\infty}\frac{1}{n}\log\mathscr{L}^n_{\varphi}(\mathbb{1})(x)$$

and the limit exists uniformly on X.

Indeed, the inequality (2-3) holds if

$$|S_n\varphi(y) - S_n\varphi(y')| \le K_n,$$

whenever y and y' are in a component of $T^{-n}(B_X(x,\delta))$, for any $x \in X$. This is proved by noting that $\hat{X}$ is compact and a finite number of such disks $B_X(x,\delta)$ cover X.

We have pointed out that the existence of a conformal measure does not require expansiveness; however the existence of an equivalent invariant measure seems to depend on this property, from Walters theory. In the second part of this section, we consider the expansiveness of a continuous map of $\hat{X}$ from X_0 preparing for the discussion of a parabolic meromorphic function on its Julia set. Since a transcendental meromorphic function is not a self-mapping of a compact metric space, this forces us to analyze carefully the definition of expansiveness.

Definition 2.5. A continuous map $T : \hat{X} \to \hat{X}$ of a compact metric space $(\hat{X}, d)$ is called *expansive* if there exists $\delta > 0$ such that we have $x = y$ if $d(T^n(x), T^n(y)) < \delta$ for all $n \ge 0$.

This definition of an expansive self-mapping of a compact metric space is not suitable to the case when T is a continuous map from X_0 into $\hat{X}$, where X_0 is a dense open subset of $\hat{X}$. For example, if there exist a point $w \in \hat{X} \setminus X_0$ and two points $x, y \in X_0$ such that $T^n(x) \to w$ and $T^n(y) \to w$ as $n \to \infty$ and $T^n(x) \ne T^n(y)$, then $d(T^n(x), T^n(y)) \to 0$ as $n \to \infty$. Thus such a continuous map can never satisfy Definition 2.5; in particular, according to this definition, no transcendental meromorphic function is expansive over its Julia set, since its escape set to infinity is nonempty.

Neither is Definition 2.5 suitable to the case when T is an infinite-to-one continuous map from X_0 to $\hat{X}$. Indeed, take a point $a \in \hat{X}$ such that $T^{-1}(a)$ contains a countable sequence $x_n \to x \in \hat{X}$ and then $d(x_n, x_{n+1}) \to 0$.

Let us analyze Definition 2.5 a bit further. Assume T is expansive. Given $x \neq y$, there exist two possibilities: either $T^m(x) = T^m(y)$ and $T^{m-1}(x) \neq T^{m-1}(y)$, for some $m \geq 1$; or $T^n(x) \neq T^n(y)$ for each n. In the first case, we have

$$d(T^{m-1}(x), T^{m-1}(y)) > \delta, \tag{2-4}$$

that is, $y \notin T^{-m+1}(B(T^{m-1}(x), \delta))$ and $x \notin T^{-m+1}(B(T^{m-1}(y), \delta))$ with $T^m(x) = T^m(y)$. In the second case, we have $d(T^{n_k}(x), T^{n_k}(y)) > \delta$ for a increasing sequence of natural numbers $\{n_k\}$ with $n_k \to \infty$, that is, $y \notin T^{-n_k}(B(T^{n_k}(x), \delta))$ and $x \notin T^{-n_k}(B(T^{n_k}(y), \delta))$. If T is not homeomorphism, then $T^{-n_k}(B(T^{n_k}(y), \delta))$ may contain two disjoint components $A^j_{n_k}$ $(j = 1, 2)$ such that T^{n_k} maps $A^j_{n_k}$ onto $B(T^{n_k}(y), \delta)$, while the definition above of expansive maps does not allow x being in any component $A^j_{n_k}$. We note that the crucial point of expansiveness is in the component $A^j_{n_k}(y)$ which contains y and that $T^{n_k} : A^j_{n_k}(y) \to B(T^{n_k}(y), \delta)$ expands the distance. From this point of view, we can extend the above definition of expansive maps to the case when T is a continuous map from X_0 to $\hat{X}$, where X_0 is a dense open subset of $\hat{X}$. Generally, the component of the preimage of a set B by a map T containing y will be denoted by $T_y^{-1}(B)$.

Definition 2.6. A continuous map $T : X_0 \to \hat{X}$ is called *precisely expansive* if there exists $\delta > 0$ such that for $x \neq y$ in $\hat{X}$, one of the following statements holds:

(1) For some $s \geq 0$, at least one of $T^s(x)$ and $T^s(y)$ is in $\hat{X} \setminus X_0$ and $T^s(x) \neq T^s(y)$;

(2) For some $m \geq 1$ with $T^m(x) = T^m(y) \in \hat{X}$ but $T^{m-1}(x) \neq T^{m-1}(y)$, we have $y \notin T_x^{-m}(B(T^m(x), \delta))$ and $x \notin T_y^{-m}(B(T^m(y), \delta))$;

(3) For a sequence of natural numbers $\{n_k\}$ with $n_k < n_{k+1} \to \infty$,
$$y \notin T_x^{-n_k}(B(T^{n_k}(x), \delta)) \quad \text{and} \quad x \notin T_y^{-n_k}(B(T^{n_k}(y), \delta)).$$

We call this δ the expansive constant for T. Note that item (2) in Definition 2.6 implies the uniform covering property (1b) of T with injectivity radius at least $\delta/2$. Generally, we cannot require that $T^{m-1}(x)$ and $T^{m-1}(y)$ have a distance with positive infimum, but if $T^{-1}(a)$ is finite for each $a \in \hat{X}$, such a positive infimum for the distance exists; see (2-4).

Obviously, the property of precise expansiveness implies that two points x and y will coincide if for every n, $y \in T_x^{-n}(B(T^n(x), \delta))$ and $x \in T_y^{-n}(B(T^n(y), \delta))$.

A continuous map $T : \hat{X} \to \hat{X}$ is precisely expansive if it is expansive. (For such an expansive map T, the set $T^{-1}(x)$ is finite for each $x \in \hat{X}$.)

When one considers a homeomorphism $T : \hat{X} \to \hat{X}$, there exists an equivalent definition of expansiveness, namely, the existence of a generator. An open cover α of $\hat{X}$ is called a *one-sided generator* for T if $\bigcap_{n=0}^{\infty} T^{-n}\overline{A}_n$ contains at most one point for any choice of $\{A_n\}$ from α. We set $\overline{\alpha} = \{\overline{A} : A \in \alpha\}$. We will consider a similar result for a precisely expansive map.

If α and β are two sets of subsets of $\hat{X}$, we denote by $\alpha \vee \beta$ the set of all subsets with the form $A \cap B$, for all $A \in \alpha$ and $B \in \beta$. Further, we set

$$\operatorname{diam} \alpha = \sup\{\operatorname{diam} A : A \in \alpha\}.$$

Definition 2.7. A finite cover α of $\hat{X}$ is called a *one-sided generator* for a continuous map $T : X_0 \to \hat{X}$, if each element of $\bigvee_{n=0}^{\infty} T^{-n}\overline{\alpha}$ has at most one point. Equivalently, the cover α is a one-sided generator for T if and only if

$$\operatorname{diam} \bigvee_{j=0}^{n} T^{-j}\overline{\alpha} \to 0 \quad \text{as } n \to \infty.$$

Theorem 2.8. *A continuous map $T : X_0 \to \hat{X}$ is precisely expansive if and only if there exists a one-sided generator for T and T has the uniform covering property* (1b) *with a fixed injectivity radius.*

Proof. Suppose that T is precisely expansive with expansive constant δ. The uniform covering property (1b) of T follows from (2) in Definition 2.6. Therefore, we only need to prove the existence of a one-sided generator.

Take a finite cover α of $\hat{X}$ by open balls with radius $\delta/2$. Let

$$E = \bigcap_{n=0}^{\infty} T_0^{-n}(A_n)$$

be an element of $\bigvee_{n=0}^{\infty} T^{-n}\overline{\alpha}$, where each A_n lies in $\overline{\alpha}$ and $T_0^{-n}(A_n)$ is a component of $T^{-n}(A_n)$. Suppose that $x, y \in E$. Then for every n, we have $x, y \in T_0^{-n}(A_n)$ and $T^n(x), T^n(y) \in A_n = \bar{B}(x_n, \delta/2)$ for a point $x_n \in \hat{X}$. Obviously, $A_n \subset B(T^n(x), \delta)$ and $A_n \subset B(T^n(y), \delta)$. From this it follows that $y \in T_x^{-n}(B(T^n(x), \delta))$ and $x \in T_y^{-n}(B(T^n(y), \delta))$. Then $x = y$, which shows that E contains at most one point. We have proved that α is a one-sided generator.

Now suppose that there exists a one-sided generator α for T and T has the uniform covering property (1b) with injective radius δ. Let η be a positive number less than the Lebesgue number of α and δ. Given two distinct points $x, y \in \hat{X}$, we assume that $\forall\, n$, $T^n(x), T^n(y) \in X_0$. Suppose that (3) in Definition 2.6 does not hold for η and therefore, there exists a $m \geq 1$ such that for all $n \geq m$, we have $y \in T_x^{-n}(B(T^n(x), \eta))$ and $x \in T_y^{-n}(B(T^n(y), \eta))$ and $T^m(y) \in T^m T_x^{-n}(B(T^n(x), \eta))$

and $T^m(x) \in T^m T_y^{-n}(B(T^n(y), \eta))$ and clearly, $T^m(y) \in T_{T^m(x)}^{-n+m}(B(T^n(x), \eta))$ and $T^m(x) \in T_{T^m(y)}^{-n+m}(B(T^n(y), \eta))$. Since η is less than the Lebesgue number of α, $B(T^n(x), \eta) \subset A_{n-m}$ for some $A_{n-m} \in \alpha$ and for $n \geq m$. Thus $T^m(x)$ and $T^m(y)$ are in an element of $\bigvee_{n=0}^{\infty} T^{-n}\alpha$. It follows that $T^m(x) = T^m(y)$.

Now we can assume that $T^m(x) = T^m(y) \in \hat{X}$ but $T^{m-1}(x) \neq T^{m-1}(y)$. It follows from the uniform covering property (1b) of T that

$$T^{m-1}(y) \notin T_{T^{m-1}(x)}^{-1}(B(T^m(x), \delta)) \quad \text{and} \quad T^{m-1}(x) \notin T_{T^{m-1}(y)}^{-1}(B(T^m(y), \delta)).$$

Obviously, $y \notin T_x^{-m}(B(T^m(x), \delta))$ and $x \notin T_y^{-m}(B(T^m(y), \delta))$. Therefore, T is precisely expansive. □

3. Dynamical properties of parabolic meromorphic functions

A meromorphic function is a map from the complex plane $\mathbb{C}$ into the extended complex plane $\hat{\mathbb{C}}$. In this section, we consider two metrics: the euclidean metric d on $\mathbb{C}$ and the spherical metric d_∞ on $\hat{\mathbb{C}}$. The metric space $(\mathbb{C}, d)$ is noncompact, but the metric space $(\hat{\mathbb{C}}, d_\infty)$ is compact. And $(\mathbb{C}, d_\infty)$ is a subspace of $(\hat{\mathbb{C}}, d_\infty)$. We are in $(\mathbb{C}, d_\infty)$ and $(\hat{\mathbb{C}}, d_\infty)$ to consider the situation of conformal measures. Set $B(a, \delta) = \{z : d(z, a) < \delta\}$ for $a \in \mathbb{C}$ and $B_\infty(a, \delta) = \{z : d_\infty(z, a) < \delta\}$ for $a \in \hat{\mathbb{C}}$.

We begin with basic dynamical properties of parabolic meromorphic functions.

Theorem 3.1. *Let f be a parabolic meromorphic function on $\mathbb{C}$ and in Class $\mathcal{B}$. Then it has finitely many and at least one parabolic domain and at most finitely many attracting domains without other types of stable domains and furthermore, $\mathcal{P}(f)$ is bounded.*

Proof. Clearly, $f(z)$ has at least one but only finitely many rational indifferent periodic points, and the number of its parabolic domains is finite and positive. Notice that $f(z)$ is in Class $\mathcal{B}$ and if $f(z)$ has a Baker domain U, then $\{f^n\}$ in U has a finite limit point. By Theorem 2.2 of [Zheng 2003], the limit point is in $\mathcal{P}(f) \cap \mathcal{J}(f)$ and so it is a rational indifferent periodic point. A contradiction is derived as every $f^n(z)$ is analytic at it. This implies that $f(z)$ has no Baker domains at all. By Theorem 2.1 of the same reference, all limit points of $\{f^n\}$ in a wandering domain are in $\mathcal{P}(f) \cap \mathcal{J}(f)$ and if a limit point is finite and not prepoles, then there exist infinitely many limit points. Thus $f(z)$ has no wandering domains. Since the boundaries of Siegel disks and Herman rings are contained in $\mathcal{P}(f) \cap \mathcal{J}(f)$, $f(z)$ therefore has no Siegel disks and Herman rings. Obviously, $f(z)$ may have attracting domains. Suppose that $f(z)$ has infinitely many attracting domains. Since every cycle of attracting domains contains at least a singular value, we take a singular value from every cycle of attracting domains to form a sequence of singular values which has a finite limit point, and clearly the limit point is in $\mathcal{J}(f)$. This implies that $\operatorname{sing}(f^{-1}) \cap \mathcal{J}(f) \neq \varnothing$. A contradiction is derived.

It is obvious that $\mathcal{P}(f)$ is bounded. $\square$

Theorem 3.2. *Let f be a transcendental meromorphic function satisfying the parabolic condition on the complex plane in Definition 1.1, except for* $\operatorname{sing}(f^{-1}) \subset \mathcal{F}(f)$. *If $\mathcal{P}(f)$ is bounded, then it has finitely many and at least one parabolic domain and at most finitely many attracting domains without other types of stable domains.*

Proof. From the proof of Theorem 3.1, it is sufficient to prove that the number of attracting domains is finite. Suppose that $f(z)$ has infinitely many attracting domains. Let $\{a_n\}$ be the sequence of all distinct attracting periodic points of f and let E be the set of all limit points of $\{a_n\}$. It is clear that $E \subset \mathcal{J}(f)$. Since every a_n is in the derived set of $\mathcal{P}(f)$, we have $E \subset \mathcal{P}(f)$, and every point in E is a rational indifferent periodic point of $f(z)$. Hence E is finite and we write $E = \{b_1, b_2, \dots, b_q\}$. Obviously, $f(E) \subseteq E$. We choose a $\delta > 0$ and a $\eta > \delta$ such that $f(B(b_j, \delta)) \subset B(f(b_j), \eta)$ $(j = 1, \dots, q)$ and f is univalent on each $B(b_j, \delta)$ and $\{B(b_j, \eta)\}$ are disjoint. For all $n \geq N$, we have $a_n \in \bigcup_{j=1}^{q} B(b_j, \delta)$. We can take a cycle of attracting periodic points $\{a, f(a), \dots, f^{p-1}(a)\}$ in $\bigcup_{j=1}^{q} B(b_j, \delta)$. Assume that $a \in B(b_1, \delta)$ and $f(a) \in B(f(b_1), \eta)$ so that $f(a) \in B(f(b_1), \delta)$. Thus $\{a, f(a), \dots, f^{p-1}(a)\} \subset \bigcup_{j=0}^{m-1} B(f^j(b_1), \delta)$, where m is the period of b_1, and $p = km$ for a positive integer k. This implies that in $B(b_1, \delta)$, $f^{km}(a) = a$. However, it is impossible for sufficiently small δ in view of the expansiveness in a neighborhood of rational indifferent periodic cycles. $\square$

The following describes equivalently the function in $\mathcal{P}(\hat{\mathbb{C}})$.

Theorem 3.3. *A meromorphic function is parabolic on the Riemann sphere if and only if it has finitely many and at least one parabolic domain and at most finitely many attracting domains without other types of stable domains and* $\widehat{\operatorname{sing}}(f^{-1}) \subset \mathcal{F}(f)$.

Proof. We just need to prove the "only if". That $\widehat{\operatorname{sing}}(f^{-1}) \subset \mathcal{F}(f)$ implies that $\infty \notin \widehat{\operatorname{sing}}(f^{-1})$ and $\operatorname{sing}(f^{-1})$ is bounded. Since $f(z)$ has only finitely many attracting and parabolic domains without other types of stable domains, $\bigcup_{n=0}^{\infty} f^n(\operatorname{sing}(f^{-1})) \subset \mathcal{F}(f)$ and the limit points of $\bigcup_{n=0}^{\infty} f^n(\operatorname{sing}(f^{-1}))$ on $\mathcal{J}(f)$ are rational indifferent periodic points of f. Thus f is parabolic on the Riemann sphere. $\square$

Denker and Urbański [1991a] investigated such properties of parabolic rational functions as the convergent speed of backward orbits of points in a small neighborhood of rational indifferent periodic points and expansive property over the Julia set, which we attempt to extend to transcendental case. The local properties of rational indifferent periodic points, for example, the Fatou's flower theorem, can be directly transferred to transcendental case. For convenience, we collect some of them.

Let $f(z)$ be a parabolic meromorphic function on the Riemann sphere and let Ω be the set of all rational indifferent periodic points of $f(z)$. The following result is basic.

Lemma 3.4. *For every $\theta > 0$ there exists $\delta = \delta(\theta) > 0$ such that for every $a \in \hat{\mathcal{J}}(f) \setminus B(\Omega, \theta)$, we have $B_\infty(a, 2\delta) \cap \mathcal{P}(f) = \varnothing$. In particular, all analytic branches of the inverse of f^n are well defined on $B_\infty(a, 2\delta)$ and $B_\infty(f(a), 2\delta)$ for every $n = 1, 2, \ldots$.*

The dynamical behavior in a neighborhood of a rational indifferent periodic point was discussed in [Denker and Urbański 1991a] in view of the Fatou's Flower Theorem. Some of their results are extracted as follows.

Lemma 3.5. *Let ω be a rational indifferent periodic point of a meromorphic function $f(z)$ with period p and $(f^p)'(\omega) = 1$. Then there exists $0 < \eta < 1$ such that*

$$|(f_\omega^{-p})'(z)| < 1 \quad \text{and} \quad |f_\omega^{-p}(z) - \omega| < |z - \omega|$$

for every $z \in B(\omega, \eta) \cap \mathcal{J}(f) \setminus \{\omega\}$, where f_ω^{-p} is the branch of the inverse of f^p sending ω to ω. And the branch f_ω^{-np} of f^{-np} sending ω to ω is well defined and is an analytical homemophism from $B(\omega, \eta) \cap \mathcal{J}(f)$ into $B(\omega, \eta) \cap \mathcal{J}(f)$.

We stress that f_ω^{-np} is not conformal on $B(\omega, \eta) \cap \mathcal{J}(f)$ (the definition of conformality can be found in [Zheng 2009]), as it has no bounded distortions over there. f^{np} is not expanding near ω.

Lemma 3.6. *Let $f(z)$ be a meromorphic function which is precisely expansive from $\mathcal{J}(f)$ to $\hat{\mathcal{J}}(f)$. Then $\widehat{\text{sing}}(f^{-1}) \subset \mathcal{F}(f)$ and f^n is precisely expansive from $\hat{\mathcal{J}}(f) \setminus \bigcup_{j=0}^{n-1} f^{-j}(\infty)$ to $\hat{\mathcal{J}}(f)$.*

Proof. Suppose that $\widehat{\text{sing}}(f^{-1}) \cap \hat{\mathcal{J}}(f) \neq \varnothing$. From this intersecting set, take a point a. Let δ be an arbitrary small fixed positive number. If a is a critical value of $f(z)$, for a $0 < \eta < \delta$ we have a component U of $f^{-1}(B_\infty(a, \eta))$ with $\operatorname{diam}_\infty U < \delta$ such that $f : U \to B_\infty(a, \eta)$ has covering number at least 2. There exist two distinct points z_1 and z_2 in U such that $f(z_1) = f(z_2)$. This contradicts the precisely expansive property of f. Assume that a is an asymptotic value and U is a tract of f over $B_\infty(a, \eta)$. Then there exists a sequence of points $\{z_n\}$ such that $z_n \to \infty$ and $f(z_n) = b \in B_\infty(a, \eta)$. Thus for all sufficiently large n, $d_\infty(z_n, z_{n+1}) < \delta$ and this contradicts the precisely expansive property of f. It is obvious that f^n is precisely expansive. □

We remark that the condition $\widehat{\text{sing}}(f^{-1}) \subset \mathcal{F}(f)$ may not imply that f is parabolic or hyperbolic, but it does when f is rational.

Theorem 3.7. *A parabolic meromorphic function f on the Riemann sphere is precisely expansive over $\hat{\mathcal{J}}(f)$.*

Proof. Take two distinct points x and y in $\mathcal{J}(f)$. Assume without any loss of generality that $f^n(x) \neq \infty$ and $f^n(y) \neq \infty$ for every n, or $f^n(x) = f^n(y) = \infty$ for some n. According to Definition 2.6, we need to treat the two cases, as follows.

(I) For some m, $f^m(x) = f^m(y) = a \in \hat{\mathcal{J}}(f)$ but $f^{m-1}(x) \neq f^{m-1}(y)$. Take a number Θ such that $0 < \Theta < \operatorname{dist}(\operatorname{sing}(f^{-1}), \mathcal{J}(f))$. If $a \notin B(\Omega, \Theta)$, then f^m is univalent from $f_x^{-m}(B_\infty(a, \delta))$ onto $B_\infty(a, \delta)$ with $\delta = \delta(\Theta)$ (by Lemma 3.4), so $y \notin f_x^{-m}(B_\infty(a, \delta))$. If $a \in B(\Omega, \Theta)$, then f is univalent from $f^{-1}_{f^{m-1}(x)}(B_\infty(a, \delta))$ onto $B_\infty(a, \delta)$ so that $f^{m-1}(y) \notin f^{-1}_{f^{m-1}(x)}(B_\infty(a, \delta))$ and furthermore, $y \notin f_x^{-m+1} \circ f^{-1}_{f^{m-1}(x)}(B_\infty(a, \delta)) = f_x^{-m}(B_\infty(a, \delta))$.

(II) for each n, $f^n(x) \neq f^n(y)$. If for a sequence of positive integers $\{n_k\}$ tending to ∞ such that $f^{n_k}(x) \notin B(\Omega, \Theta)$, then $f_x^{-n_k}$ is a single-valued function over $B_\infty(f^{n_k}(x), \delta)$ and therefore $\operatorname{diam} f_x^{-n_k}(B_\infty(f^{n_k}(x), \delta)) \to 0$ as $k \to \infty$. This implies that for all sufficiently large k, $y \notin f_x^{-n_k}(B_\infty(f^{n_k}(x), \delta)$. Now assume that for all $n \geq N$, $f^n(x) \in B(\Omega, \Theta)$ and $f^n(y) \in B(\Omega, \Theta)$. When Θ is sufficiently small, we have for some m, $f^m(x), f^m(y) \in \Omega$. Then $f^m(x)$ and $f^m(y)$ are distinct rational indifferent periodic points of $f(z)$ so that $B_\infty(f^n(x), \delta)$ is disjoint from $B_\infty(f^n(y), \delta)$ for all $n \geq m$. Obviously, $y \notin f_x^{-n}(B_\infty(f^n(x), \delta))$ and $x \notin f_y^{-n}(B_\infty(f^n(y), \delta))$. $\square$

That a rational function is parabolic if and only if it is expansive with at least one rational indifferent periodic point is proved in [Denker and Urbański 1991a]. In view of Theorem 2.8, Theorem 3.7 implies that $f(z)$ has a one-sided generator over $\hat{\mathcal{J}}(f)$. Actually, we can also use the existence of a one-sided generator to show the precisely expansive property of a parabolic meromorphic function on the Riemann sphere, as in view of Lemma 3.5, for each n, $f^n(z)$ has the uniformly covering property (1b) over $\hat{\mathcal{J}}(f)$ with a fixed injectivity radius.

In what follows, let us discuss the existence of conformal measures of parabolic meromorphic functions on the Riemann sphere. We shall use the results in Section 2 to attain our purpose. Let $f(z)$ be a parabolic meromorphic function on the Riemann sphere and let d_∞ be the Riemann spherical metric. Hence $(\hat{\mathcal{J}}(f), d_\infty)$ is a compact metric space. Consider the continuous map $f : \mathcal{J}(f) \to \hat{\mathcal{J}}(f)$ under the Riemann spherical metric. This map is not conformal, so we cannot use Theorem 3.1 of [Zheng 2009] to achieve our purpose. We take a different approach.

Define the *pressure* $P(f, t)$ for a parabolic meromorphic function f over $\hat{\mathcal{J}}(f)$ as follows. Define $\varphi_t : \mathcal{J}(f) \to \mathbb{R}$ by $\varphi_t(z) = -t \log f^\times(z)$, and set $\mathcal{L}_t = \mathcal{L}_{\varphi_t}$.

Thus, for a fixed value $a \in \hat{\mathcal{J}}(f)$ and $g \in \mathcal{C}(\mathcal{J}(f))$, we have

$$\mathcal{L}_t(g)(a) = \sum_{f(z)=a} \frac{g(z)}{f^\times(z)^t}.$$

Obviously, $\mathcal{L}_t^n(\mathbb{1})(a) = \sum_{f^n(z)=a} (f^n)^\times(z)^{-t}$. Set

$$P_a(f,t) = \limsup_{n\to\infty} \frac{1}{n} \log \mathcal{L}_t^n(\mathbb{1})(a);$$

then the pressure is

$$\hat{P}(f,t) = \sup\{P_a(f,t) : a \in \hat{\mathcal{J}}(f)\}.$$

Lemma 3.8. *Let f be a parabolic meromorphic function on the Riemann sphere.*

(1) $\hat{P}(f,t) \geq 0$.

(2) $P_a(f,t) = P_b(f,t)$ *whenever* $a, b \in \hat{\mathcal{J}}(f) \setminus \Omega$.

We write $P(f,t)$ for $P_a(f,t)$ for $a \in \hat{\mathcal{J}}(f) \setminus \Omega$.

Proof. (1) Take a point $a \in \Omega$ with period p. For each n, $(f^{np})^\times(a) = 1$ and $\mathcal{L}_t^{np}(\mathbb{1})(a) > 1$. This implies that $\hat{P}(f,t) \geq 0$.

(2) Assume that $P_a(f,t) < \infty$. For arbitrarily small $\varepsilon > 0$ and for all sufficiently large n, we have

$$\begin{aligned} e^{n(P_a(f,t)+\varepsilon)} &\geq \mathcal{L}_t^n(\mathbb{1})(a) \\ &= \sum_{f^n(z)=a} (f^n)^\times(z)^{-t} \\ &= \sum_{f^p(w)=a} \sum_{f^{n-p}(z)=w} (f^p)^\times(w)^{-t} (f^{n-p})^\times(z)^{-t} \\ &\geq (f^p)^\times(w)^{-t} \sum_{f^{n-p}(z)=w} (f^{n-p})^\times(z)^{-t}, \end{aligned}$$

where $w \in f^{-p}(a)$. Since $a, b \notin \Omega$, we have a $\delta > 0$ such that $B_\infty(a, 2\delta) \cap \mathcal{P}(f) = \varnothing$ and $B_\infty(b, 2\delta) \cap \mathcal{P}(f) = \varnothing$. We can choose a p such that $f^{-p}(a) \cap B_\infty(b,\delta) \neq \varnothing$ and therefore by the Koebe distortion theorem for the Riemann spherical metric, for an absolute constant K we have

$$e^{n(P_a(f,t)+\varepsilon)} \geq (f^p)^\times(w)^{-t} \sum_{f^{n-p}(z)=b} K^{-t} (f^{n-p})^\times(z)^{-t}.$$

This yields that $P_a(f,t) + \varepsilon \geq P_b(f,t)$ and so $P_a(f,t) \geq P_b(f,t)$. The same argument implies that $P_b(f,t) \geq P_a(f,t)$. Hence $P_b(f,t) = P_a(f,t)$. □

Set $\tau(f) = \inf\{t \geq 0 : P(f,t) < \infty\}$ and $s(f) = \inf\{t \geq 0 : P(f,t) \leq 0\}$. We call $s(f)$ the *Poincaré exponent*.

Lemma 3.9. *Let f be a parabolic meromorphic function on the Riemann sphere.*

(I) $\tau(f) \le s(f) \le 2$.

(II) *$P(f,t)$ is strictly decreasing and convex in $t \in (\tau(f), +\infty)$.*

(III) *If $t \ge s(f)$, then $\varphi_t(z) = -t \log f^\times(z)$ is summable, and $P(f,s) = 0 = \hat{P}(f,t)$.*

Proof. (I) Take a point $a \in \mathcal{J}(f)$ and $r > 0$ such that $B_\infty(a, 2r) \cap \hat{\mathcal{P}}(f) = \varnothing$. Let f_z^{-n} be the analytic branch of f^{-n} over $B_\infty(a,r)$ sending a to z with $f^n(z) = a$. Set $U(z) = f_z^{-n}(B_\infty(a,r))$. By the Koebe covering theorem for the spherical metric, we have

$$U(z) \supseteq B_\infty(z, Kr(f_z^{-n})^\times(a))$$

for an absolute constant K. Thus, noting that $U(z)$ is disjoint for distinct $z \in f^{-n}(a)$, we have

$$\sum_{f^n(z)=a} \pi(Kr(f_z^{-n})^\times(a))^2 \le \sum_{f^n(z)=a} \text{spherical area of}(U(z)) \le \pi,$$

and furthermore, using $(f_z^{-n})^\times(a) = ((f^n)^\times(z))^{-1}$, we obtain

$$\sum_{f^n(z)=a} \frac{1}{((f^n)^\times(z))^2} \le (Kr)^{-2}.$$

This implies that $P(f,2) = P_a(f,2) \le 0$ and hence $s(f) \le 2$.

(II) Take a point $a \in \mathcal{J}(f)$ and a $\delta > 0$ such that $B_\infty(a,\delta) \cap \mathcal{P}(f) = \varnothing$. Then there exists an integer N such that for $n \ge N$

$$d_\infty(f^n(x), f^n(y)) \ge \lambda C^n d_\infty(x,y),$$

with $C > 1$ and $\lambda > 0$, whenever x and y are in an injective component of $f^{-n}(B_\infty(a,\delta))$ and $(f^n)^\times(w) > \lambda C^n, \forall w \in f^{-n}(a)$. This easily implies that $P(f,t) = P_a(f,t)$ is strictly decreasing and convex in t. (See the proof of Theorem 2.3 of [Zheng 2008]).

(III) For arbitrary $t > s(f)$, $P(f,t) < 0$. For a fixed $a \in \hat{\mathcal{J}}(f) \setminus \Omega$, $\mathcal{L}_t^n(\mathbb{1})(a) \to 0$ as $n \to \infty$ and hence for $n \ge m$, $\mathcal{L}_t^n(a) < 1$. Take $z_j (1 \le j \le q)$ such that $\hat{\mathcal{J}}(f) \subset \bigcup_{j=1}^q B_\infty(z_j, \delta/2)$, where δ is chosen such that for each j, $B_\infty(z_j, 2\delta) \cap \operatorname{sing}(f^{-1}) = \varnothing$. Take a positive integer N such that for arbitrary pair of j and i, $B_\infty(z_j,\delta) \cap f^{-N+1}(z_i) \ne \varnothing$. For a P, we have $\mathcal{L}_t^{PN}(\mathbb{1})(a) < 1$. This implies that $\mathcal{L}_t^N(\mathbb{1})(b) < 1$ for some $b \in \hat{\mathcal{J}}(f) \setminus \Omega$. Then $b \in B_\infty(z_{j_0}, \delta/2)$ for some j_0. We find $M = M(j_0)$ disks $B_\infty(b_i, \eta)$ $(1 \le i \le M)$ covering the $B_\infty(z_{j_0}, \delta/2)$ such that each disk $B_\infty(b_i, 2\eta)$ does not intersect $\operatorname{sing}(f^{-N})$. In view of the Koebe distortion theorem, we have

$$\mathcal{L}_t^N(\mathbb{1})(z_{j_0}) \le K^{Mt} \mathcal{L}_t^N(\mathbb{1})(b) < K^{Mt},$$

where K is an absolute constant.

For each $j \in \{1, 2, \ldots, q\}$, $f^{-N+1}(z_{j_0}) \cap B_\infty(z_j, \delta/2) \neq \varnothing$, from which we take a point $w^j_{j_0}$. We have

$$\begin{aligned}\mathcal{L}_t^N(\mathbb{1})(z_{j_0}) &= \sum_{f^N(z)=z_{j_0}} (f^N)^\times(z)^{-t} = \sum_{f^{N-1}(w)=z_{j_0}} (f^{N-1})^\times(w)^{-t} \sum_{f(z)=w} f^\times(z)^{-t} \\ &\geq (f^{N-1})^\times(w^j_{j_0})^{-t} \sum_{f(z)=w^j_{j_0}} f^\times(z)^{-t} \\ &= (f^{N-1})^\times(w^j_{j_0})^{-t} \mathcal{L}_t(\mathbb{1})(w^j_{j_0});\end{aligned}$$

equivalently,

$$\begin{aligned}\mathcal{L}_t(\mathbb{1})(w^j_{j_0}) &\leq (f^{N-1})^\times(w^j_{j_0})^t \mathcal{L}_t^N(\mathbb{1})(z_{j_0}) \\ &< (f^{N-1})^\times(w^j_{j_0})^t K^{Mt}.\end{aligned}$$

Set

$$C = \max\{(f^{N-1})^\times(w^v_j) K^{M(j)} : 1 \leq j, v \leq q\}.$$

For each $w \in \hat{\mathcal{J}}(f)$ we have $w \in B_\infty(z_j, \delta/2)$, so $w \in B_\infty(w^j_{j_0}, \delta)$ for some j. By the Koebe distortion theorem,

$$\mathcal{L}_t(\mathbb{1})(w) \leq L^t \mathcal{L}_t(\mathbb{1})(w^j_{j_0}) < L^t C^t$$

for an absolute constant $L > 0$. This yields that φ_t is summable. Letting t approach $s(f)$ from above, we have

$$\mathcal{L}_s(\mathbb{1})(w) \leq L^s C^s.$$

We have proved that $\varphi_s = -s \log f^\times(z)$ with $s = s(f)$ is summable on $\hat{\mathcal{J}}(f)$ so that $P(f, s) \leq 0$. This immediately implies that $P(f, s) = 0$.

Now we prove that $\hat{P}(f, t) = 0$. For $t > s(f)$, we know that $P(f, t) < 0$. Therefore, we want to calculate $P_a(f, t) = 0$ for $a \in \Omega$. It suffices to prove that $\mathcal{L}_t^n(\mathbb{1})(a)$ is uniformly bounded in n and t for $a \in \Omega$. Assume without loss that the period of a is 1. We take $\eta > 0$ such that $B_\infty(w, \eta) \cap \mathcal{P}(f) = \varnothing$ for $w \in f^{-1}(a) \setminus \{a\}$ and $B_\infty(\infty, \eta) \cap \mathcal{P}(f) = \varnothing$. We can take finitely many w_j, for $1 \leq j \leq q$, such that $w_j \in f^{-1}(a) \setminus \{a\}$ and $\{B_\infty(w_j, \eta/2)\}$ together with $B_\infty(\infty, \eta/2)$ form a covering of $f^{-1}(a) \setminus \{a\}$. By the Koebe distortion theorem, for some c with $P(f, t) < c < 0$, we have for $n \geq N$

$$\mathcal{L}_t^n(\mathbb{1})(w) \leq e^{nc} \quad \text{for } w \in f^{-1}(a) \setminus \{a\}.$$

Set

$$\sum_{\substack{f(w)=a \\ w \neq a}} f^\times(w)^{-t} = K_t.$$

We have

$$
\begin{aligned}
\text{(3-1)}\qquad \mathcal{L}_t^n(\mathbb{1})(a) &= \sum_{f^n(z)=a} (f^n)^{\times}(z)^{-t} \\
&= (f^n)^{\times}(a)^{-t} + \sum_{\substack{f^n(z)=a\\ z\neq a}} (f^n)^{\times}(z)^{-t} \\
&= 1 + \sum_{f(w)=a} \sum_{\substack{f^{n-1}(z)=w\\ z\neq a}} f^{\times}(w)^{-t}(f^{n-1})^{\times}(z)^{-t} \\
&\le 1 + e^{(n-1)c} \sum_{\substack{f(w)=a\\ w\neq a}} f^{\times}(w)^{-t} + \sum_{\substack{f^{n-1}(z)=a\\ z\neq a}} (f^{n-1})^{\times}(z)^{-t} \\
&\le 1 + K_t e^{(n-1)c} + \sum_{f^{n-1}(z)=a,z\neq a} (f^{n-1})^{\times}(z)^{-t} \\
&\le K_t(1+e^c+\cdots+e^{(n-1)c}) + \sum_{\substack{f^N(z)=a\\ z\neq a}} (f^N)^{\times}(z)^{-t} \\
&< K_t \frac{1}{1-e^c} + \sum_{\substack{f^N(z)=a\\ z\neq a}} (f^N)^{\times}(z)^{-t}.
\end{aligned}
$$

This implies that $\hat{P}(f,t) = \max_{a\in\Omega} P_a(f,t) = 0$.

For the case when $s = s(f)$, for arbitrarily small $\varepsilon > 0$ it follows from the above implication that there exists $N = N(\varepsilon)$ such that

$$
\mathcal{L}_s^n(\mathbb{1})(a) \le K_s \frac{e^{n\varepsilon}-1}{e^{\varepsilon}-1} + \sum_{\substack{f^N(z)=a\\ z\neq a}} (f^N)^{\times}(z)^{-s}.
$$

This implies that $\hat{P}(f,s) \le \varepsilon$ and hence $\hat{P}(f,s) = 0$. $\square$

The next result reflects the expansiveness of a parabolic meromorphic function over $\hat{\mathcal{J}}(f)$. Its idea comes from [Rippon and Stallard 1999].

Lemma 3.10. *Let $f(z)$ be a parabolic meromorphic function on $\mathbb{C}$ and in class $\mathcal{B}$. There exists $c > 0$ such that for each n and $z \in \mathcal{J}(f) \setminus \bigcup_{j=0}^{n-1} f^{-j}(\infty)$, we have*

$$
\text{(3-2)}\qquad |(f^n)'(z)| > c\frac{|f^n(z)|+1}{|z|+1}.
$$

Let M_m be the set of all points $z \in \mathcal{J}(f) \setminus \mathcal{J}_\infty(f)$ for which there exists a sequence $\{s_k\}$ with $s_k \in [km,(k+1)m]$ and $f^{s_k}(z) \notin B(\Omega,\theta)$ for some constant $\theta > 0$. There *exist constants $c > 0$ and $\lambda > 1$ such that*

$$|(f^n)'(z)| > c\lambda^n \frac{|f^n(z)|+1}{|z|+1} \quad \textit{for } z \in M_m. \tag{3-3}$$

Proof. Assume without loss of generality that $\{z : |z| < 1\} \subset \mathcal{F}(f)$. In view of Theorem 3.1, take a $R > 1$ such that $\mathcal{P}(f) \subset B(0, R)$ and $|f^n(0)| < R$ for all $n \in \mathbb{N}$. In view of the result in [Rippon and Stallard 1999] (compare [Zheng 2003]), we have

$$|(f^n)'(z)| > \frac{|f^n(z)|(\log|f^n(z)| - \log R)}{4|z|} \tag{3-4}$$

for $z \in \mathcal{J}(f) \setminus \mathcal{J}_\infty(f)$; furthermore, for $z \in \mathcal{J}(f) \setminus \mathcal{J}_\infty(f)$ with $|f^n(z)| \geq e^2 R$, we have

$$|(f^n)'(z)| > \frac{|f^n(z)|+1}{4(|z|+1)}.$$

We first prove (3-2) for $n = 1$. Since $d(\mathcal{J}(f), \operatorname{sing}(f^{-1})) > 0$, we can take a positive number $A \geq 1$ such that

$$B\left(z, \frac{|z|+1}{A}\right) \cap \operatorname{sing}(f^{-1}) = \varnothing$$

for any $z \in \mathcal{J}(f)$ and

$$B(0, 1) \not\subseteq f^{-1}\left(B\left(f(z), \frac{|f(z)|+1}{A}\right)\right)$$

for $z \in \mathcal{J}(f) \setminus f^{-1}(\infty)$ with $|f(z)| < e^2 R$. Then for a fixed $z \in \mathcal{J}(f) \setminus f^{-1}(\infty)$, we have

$$B\left(f(z), \frac{|f(z)|+1}{A}\right) \cap \operatorname{sing}(f^{-1}) = \varnothing$$

and f_z^{-1} is a single-valued analytic branch on $B(f(z), (|f(z)|+1)/A)$ tending $f(z)$ to z. Let U be the component of $f^{-1}(B(f(z), (|f(z)|+1)/A))$ containing z. Then $f : U \to B(f(z), (|f(z)|+1)/A) = B$ (say) is univalent and U is simply connected. In view of the hyperbolic metric principle, we have

$$\lambda_U(z) = \lambda_B(f(z))|f'(z)| = \frac{2A|f'(z)|}{|f(z)|+1}.$$

For $z \in \mathcal{J}(f) \setminus f^{-1}(\infty)$ with $|f(z)| < e^2 R$, $B(0, 1) \not\subseteq U$. If $0 \notin U$, then $|z|\lambda_U(z) \geq \frac{1}{4}$; If $0 \in U$, then for a with $|a| \leq 1$, $|z-a|\lambda_U(z) \geq \frac{1}{4}$. Therefore, we always have $(|z|+1)\lambda_U(z) \geq \frac{1}{4}$. These imply that

$$|f'(z)| \geq \frac{1}{8A} \frac{|f(z)|+1}{|z|+1}. \tag{3-5}$$

This proves (3-2) for $n = 1$ with $c = \frac{1}{8A}$.

Suppose (3-2) is not true. Then there exist a sequence of positive integers $\{m_k\}$ and a sequence of points $z_k \in \mathcal{J}(f) \setminus \bigcup_{j=0}^{m_k-1} f^{-j}(\infty)$ such that

$$\varepsilon_k = \frac{|(f^{m_k})'(z_k)|(|z_k|+1)}{|(f^{m_k})(z_k)|+1} \to 0 \quad \text{as } k \to \infty.$$

We can take a positive number C such that for $z \in \mathcal{J}(f) \setminus B(\Omega, \theta)$,

$$B\left(z, 2\frac{|z|+1}{C}\right) \cap \hat{\mathcal{P}}(f) = \varnothing. \tag{3-6}$$

If $f^{m_k}(z_k) \notin B(\Omega, \theta)$, a single-valued analytic branch g_k of f^{-m_k} sending $f^{m_k}(z_k)$ to z_k exists on $B(f^{m_k}(z_k), 2(|f^{m_k}(z_k)|+1)/C)$. By the Koebe covering theorem, we have

$$\begin{aligned} g_k\left(B\left(f^{m_k}(z_k), \frac{|f^{m_k}(z_k)|+1}{C}\right)\right) &\supseteq B\left(z_k, \frac{|f^{m_k}(z_k)|+1}{4C}|g_k'(f^{m_k}(z_k))|\right) \\ &= B\left(z_k, \frac{|z_k|+1}{4C\varepsilon_k}\right) \\ &\supseteq B\left(0, \frac{|z_k|+1}{4C\varepsilon_k} - |z_k|\right). \end{aligned} \tag{3-7}$$

Now assume that $f^j(z_k) \in B(\Omega, \theta)$, $p_k \le j \le m_k$ and $f^{p_k-1}(z_k) \notin B(\Omega, \theta)$. By Lemma 3.5, we have $|(f^{m_k-p_k})'(f^{p_k}(z_k))| \ge 1$ and so for a positive constant a,

$$|(f^{m_k-p_k})'(f^{p_k}(z_k))| \ge a\frac{|f^{m_k}(z_k)|+1}{|f^{p_k}(z_k)|+1}.$$

Thus

$$\begin{aligned} |f'(f^{p_k-1}(z_k))||(f^{p_k-1})'(z_k)| &= |(f^{p_k})'(z_k)| \\ &= \frac{|(f^{m_k})'(z_k)|}{|(f^{m_k-p_k})'(f^{p_k}(z_k))|} \\ &\le \frac{|(f^{m_k})'(z_k)|}{|f^{m_k}(z_k)|+1}\frac{|f^{p_k}(z_k)|+1}{a} \\ &= \frac{\varepsilon_k(|f^{p_k}(z_k)|+1)}{a(|z_k|+1)} \end{aligned}$$

and in view of (3-5), we have

$$|f'(f^{p_k-1}(z_k))| \ge \frac{1}{8A}\frac{|f^{p_k}(z_k)|+1}{|f^{p_k-1}(z_k)|+1}.$$

Combining the above two inequalities implies that

$$|(f^{p_k-1})'(z_k)| \le \frac{\varepsilon_k 8A(|f^{p_k-1}(z_k)|+1)}{a(|z_k|+1)}. \tag{3-8}$$

Since $f^{p_k-1}(z_k) \notin B(\Omega, \theta)$, from (3-8) we have

$$h_k\left(B\left(f^{p_k-1}(z_k), \frac{|f^{p_k-1}(z_k)|+1}{C}\right)\right) \supseteq B\left(0, \frac{a(|z_k|+1)}{32AC\varepsilon_k} - |z_k|\right),$$

where h_k is the analytic branch of f^{-p_k+1} which sends $f^{p_k-1}(z_k)$ to z_k. This together with (3-7) shows the existence of a sequence of positive integers $\{n_k\}$ such that

$$f_{z_k}^{-n_k}\left(B\left(f^{n_k}(z_k), \frac{|f^{n_k}(z_k)|+1}{C}\right)\right) \supseteq B\left(0, \frac{a(|z_k|+1)}{32AC\varepsilon_k} - |z_k|\right).$$

But this gives

$$\frac{a(|z_k|+1)}{32AC\varepsilon_k} - |z_k| \to +\infty$$

as $k \to \infty$, and a contradiction is derived. We have proved (3-2).

Now we prove (3-3). Let $z \in M_m$. In view of (3-4), there exists an $R_0 > R$ such that

$$|(f^n)'(z)| > 2(1+c^{-1})\frac{|f^n(z)|+1}{|z|+1}, \text{ for } n \in \mathbb{N}, |f^n(z)| > R_0 \tag{3-9}$$

where c is the constant in (3-2). Using the same argument as in the proof of (3-2), we can also attain (3-9) for $n \geq N \geq m$, $z \in (\mathcal{J}(f) \setminus \bigcup_{j=0}^{n-1} f^{-j}(\infty)) \cap B(0, R_0)$ with $f^n(z) \notin B(\Omega, \theta)$.

For any $0 \leq p < 2N$, we treat two cases. If $|f^{2N+p}(z)| > R_0$, from (3-9) we have

$$|(f^{2N+p})'(z)| > 2\frac{|f^{2N+p}(z)|+1}{|z|+1};$$

If $|f^{2N+p}(z)| \leq R_0$, for some $N \leq N_1 \leq 2N+p$ we have either $|f^{N_1}(z)| \leq R_0$ and $f^{N_1}(z) \notin B(\Omega, \theta)$ or $|f^{N_1}(z)| > R_0$. Therefore from (3-2) and (3-9) we have

$$\begin{aligned}|(f^{2N+p})'(z)| &= |(f^{2N+p-N_1})'(f^{N_1}(z))||(f^{N_1})'(z)| \\ &> c\frac{|f^{2N+p}(z)|+1}{|f^{N_1}(z)|+1}2c^{-1}\frac{|f^{N_1}(z)|+1}{|z|+1} \\ &= 2\frac{|f^{2N+p}(z)|+1}{|z|+1}.\end{aligned}$$

For $n \geq 2N$, we write $n = 2qN + p$ with $0 \leq p < 2N$ and thus

$$|(f^n)'(z)| > 2^q\frac{|f^n(z)|+1}{|z|+1} > \frac{1}{2}(2^{\frac{1}{2N}})^n\frac{|f^n(z)|+1}{|z|+1}.$$

For $1 \leq n < 2N$, we use (3-2). □

The next result confirms the existence of a measure that becomes s-conformal.

Lemma 3.11. *Let $f(z)$ be a parabolic meromorphic function on the Riemann sphere. Then (f, φ_s) is admissible over $\hat{\mathcal{J}}(f)$.*

Proof. We check the conditions in Definition 2.1. Obviously, for f, (1a) and (1b) hold, and (1d) holds by virtue of Lemma 3.9(III). In view of Lemma 2.2, (1c) is true for f. We state (1e) for (f, φ_s) as follows: for all $\varepsilon > 0$, there exists $\delta_1 \in (0, \delta)$ such that for any pair $a, b \in \hat{\mathcal{J}}(f)$, the condition $d_\infty(a, b) < \delta_1$ implies

$$\sum_{f(z)=a} \left|\exp \varphi_s(f_z^{-1}(a)) - \exp \varphi_s(f_z^{-1}(b))\right| < \varepsilon;$$

that is,

$$\sum_{f(z)=a} \left|\frac{1}{f^\times(z)^s} - \frac{1}{f^\times(z')^s}\right| < \varepsilon, \tag{3-10}$$

where $z' = f_z^{-1}(b)$. From Lemma 3.2 of [Zheng 2009], noting that φ_s is summable, (3-10) follows from

$$\left|1 - \frac{f^\times(z)^s}{f^\times(z')^s}\right| \leq C_s d_\infty(a, b), \tag{3-11}$$

whenever $d_\infty(a, b) < \delta$. And (3-11) can be proved via the same argument used in the proof of Lemma 3.1 of [Zheng 2008] and the inequality (3-2). □

Now we are in the position to prove Theorem 1.2, which, as we recall, states that any f in $\mathcal{P}(\hat{\mathbb{C}})$ has a s-conformal measure with $P(f, s) = 0$.

Proof of Theorem 1.2. In view of Lemma 3.11 and Theorem 2.3, there exists a probability measure μ with $\mathcal{L}_s^*(\mu) = \lambda\mu$, $\lambda = \mathcal{L}_s^*(\mu)(\mathbb{1})$, satisfying the conditions in Theorem 2.3(3). We calculate λ using (2-2), and obtain

$$\lambda^n = \mu(\mathcal{L}_s^n(\mathbb{1})) \leq \sup\{\mathcal{L}_s^n(\mathbb{1})(x) : x \in \hat{\mathcal{J}}(f)\}.$$

Using the same argument as in the proof of (3-1), for arbitrarily small $\varepsilon > 0$, we have for $n \geq N$

$$\sup\{\mathcal{L}_s^n(\mathbb{1})(x) : x \in \hat{\mathcal{J}}(f)\} \leq Kne^{n\varepsilon},$$

so $\log \lambda \leq 0$.

If $\mu(\{a\}) > 0$ for a point $a \in \hat{\mathcal{J}}(f)$, then $\lambda^n \geq \mu(\{a\})\mathcal{L}_s^n(\mathbb{1})(a)$ and so $\log \lambda \geq P(f, s) = 0$. Now assume that μ is atomless and we can find a disk $B_\infty(a, \eta)$ with $\mu(B_\infty(a, \eta)) > 0$ which does not intersect $\mathcal{P}(f)$. Thus

$$\begin{aligned}\lambda^n &\geq \mu(B_\infty(a, \eta)) \inf\{\mathcal{L}_s^n(\mathbb{1})(x) : x \in B_\infty(a, \eta) \cap \hat{\mathcal{J}}(f)\} \\ &\geq \mu(B_\infty(a, \eta)) K^{-s} \mathcal{L}_s^n(\mathbb{1})(a)\end{aligned}$$

so that $\log \lambda \geq P(f, s) = 0$. Therefore, we have proved that $\lambda = 1$ and μ is a s-conformal measure of $f(z)$ over $\hat{\mathcal{J}}(f)$. □

In what follows, we consider the existence of a f-invariant measure equivalent to the s-conformal measure μ_s. We cannot get such an invariant measure from Walters' result. Therefore, we will complete our discussion in light of the results of Martens.

Lemma 3.12 [Martens 1992, Proposition 2.6]. *Let μ be a σ-finite Borel measure on a σ-compact space X and $f : X \to X$ a measurable map. Then f has a μ-equivalent, σ-finite invariant measure $\mathfrak{m}$, if the following statements hold*:

(1) *There exist a countable collection of pairwise disjoint Borel sets $G = \{I_j : j \in \mathbb{N}\}$ of X such that each I_j is σ-compact, $0 < \mu(I_j) < \infty$, $\mu(X \setminus \bigcup_{j=1}^{\infty} I_j) = 0$ and for all pair I_i and I_j, for some $n \geq 0$, $\mu(f^{-n}(I_i) \cap I_j) > 0$.*

(2) *There exists a σ-finite measure ν having properties that for each $I \in G$ there exists a $K > 0$ such that $K^{-1} \leq \nu(I) \leq K$, $\sup_{n\geq 0} \nu(f^{-n}(I)) < \infty$, and*
$$\frac{1}{K}\frac{\mu(A)}{\mu(I)} \leq \frac{\nu(f^{-n}(A))}{\nu(f^{-n}(I))} \leq K\frac{\mu(A)}{\mu(I)}.$$
for all measurable sets $A \subset I$ and all $n \in \mathbb{N}$.

(3) $\sum_{n=0}^{\infty} \nu(f^{-n}(\hat{I})) = \infty$ *for some $\hat{I} \in G$.*

Actually, $\mathfrak{m}$ is a weakly convergent limit of $\{Q_n(\nu)\}$ on each $I \in G$, where
$$Q_n(\nu) = \frac{\sum_{j=0}^{n-1} f_*^i \nu}{\sum_{j=0}^{n-1} f_*^i \nu(\hat{I})}$$
and for a Borel measurable map g, $g_*\nu = \nu \circ g^{-1}$.

Let $f(z)$ be a parabolic meromorphic function in $\mathcal{P}(\hat{\mathbb{C}})$ and let μ_s be the s-conformal measure determined in Theorem 1.2. Assume that μ_s is atomless. Set $X_0 = \mathcal{J}(f) \setminus \bigcup_{n=0}^{\infty} f^{-n}(\Omega)$ and $X = \hat{\mathcal{J}}(f) \setminus \bigcup_{n=0}^{\infty} f^{-n}(\Omega)$. Then $\mu_s(X_0) = \mu_s(X) = 1$ and we can construct a countable collection of disjoint Borel sets $G = \{I_j : j \in \mathbb{N}\}$ of X such that for each j, $I_j \subset B_\infty(a_j, \delta_j)$ and $B_\infty(a_j, 2\delta_j) \cap \hat{\mathcal{P}}(f) = \varnothing$ for some $a_j \in I_j$ and which satisfies (1) in Lemma 3.12. In view of the Koebe distortion theorem for the spherical metric and the definition of s-conformal measure, we easily prove (2) in Lemma 3.12 for f and G with respect to μ_s and $\nu = \mu_s$. Therefore, the crucial point is in (3) in Lemma 3.12. We have
$$\begin{aligned}\mu_s(f^{-n}(I_j)) &= \sum_{f^n(z)=a_j} \mu_s(f_z^{-n}(I_j)) = \sum_{f^n(z)=a_j} \int_{I_j} (f_z^{-n})^\times(w)^s \mu_s(w) \\ &\geq \sum_{f^n(z)=a_j} K^{-s} (f_z^{-n})^\times(a_j)^s \mu_s(I_j) = K^{-s} \mu_s(I_j) \mathcal{L}_s^n(\mathbb{1})(a_j)\end{aligned}$$
and
$$\mu_s(f^{-n}(I_j)) \leq K^s \mu_s(I_j) \mathcal{L}_s^n(\mathbb{1})(a_j),$$

where K is the Koebe distortion constant. Thus we have

$$K^{-s}\mu_s(I_j)\sum_{n=0}^{\infty}\mathcal{L}_s^n(\mathbb{1})(a_j) \le \sum_{n=0}^{\infty}\mu_s(f^{-n}(I_j)) \le K^s\mu_s(I_j)\sum_{n=0}^{\infty}\mathcal{L}_s^n(\mathbb{1})(a_j).$$

In view of Lemma 3.12, $f(z)$ has an f-invariant, σ-finite measure m which is equivalent to μ_s if $\sum_{n=0}^{\infty}\mathcal{L}_s^n(\mathbb{1})(a) = \infty$ for some $a \in \mathcal{J}(f) \setminus \bigcup_{n=0}^{\infty} f^{-n}(\Omega)$.

In view of the statements above, we have actually proved Theorem 1.3.

On the other hand, assume that $f(z)$ has an f-invariant, σ-finite measure m_s which is equivalent to μ_s. Take an $a \in \mathcal{J}(f) \setminus \bigcup_{n=0}^{\infty} f^{-n}(\Omega)$, and $B_\infty(a, 2\delta) \cap \mathcal{P}(f) = \varnothing$ for some $\delta > 0$. Set $I = B_\infty(a, \delta) \cap \mathcal{J}(f)$. Then $\mu_s(I) > 0$, and $m_s(I) > 0$ and for each n, $m_s(f^{-n}(I)) = m_s(I)$. This implies that

$$\sum_{n=0}^{\infty} m_s(f^{-n}(I)) = \infty.$$

Then if the Radon–Nikodym derivative $dm_s/d\mu_s$ of m_s with respect to μ_s is bounded, we have $\sum_{n=0}^{\infty}\mathcal{L}_s^n(\mathbb{1})(a) = \infty$.

References

[Aaronson et al. 1993] J. Aaronson, M. Denker, and M. Urbański, "Ergodic theory for Markov fibred systems and parabolic rational maps", *Trans. Amer. Math. Soc.* **337**:2 (1993), 495–548. MR 94g:58116 Zbl 0789.28010

[Denker and Urbański 1991a] M. Denker and M. Urbański, "Hausdorff and conformal measures on Julia sets with a rationally indifferent periodic point", *J. London Math. Soc.* (2) **43**:1 (1991), 107–118. MR 92k:58153 Zbl 0734.28007

[Denker and Urbański 1991b] M. Denker and M. Urbański, "Absolutely continuous invariant measures for expansive rational maps with rationally indifferent periodic points", *Forum Math.* **3**:6 (1991), 561–579. MR 92k:58154 Zbl 0745.28008

[Kotus and Urbański 2002] J. Kotus and M. Urbański, "Conformal, geometric and invariant measures for transcendental expanding functions", *Math. Ann.* **324**:3 (2002), 619–656. MR 2003j:37072 Zbl 1009.37032

[Martens 1992] M. Martens, "The existence of σ-finite invariant measures, applications to real one-dimensional dynamics", preprint, 1992. arXiv math/9201300v1

[Rempe 2009] L. Rempe, "Hyperbolic dimension and radial Julia sets of transcendental functions", *Proc. Amer. Math. Soc.* **137**:4 (2009), 1411–1420. MR 2010h:37100 Zbl 05544554

[Rippon and Stallard 1999] P. J. Rippon and G. M. Stallard, "Iteration of a class of hyperbolic meromorphic functions", *Proc. Amer. Math. Soc.* **127**:11 (1999), 3251–3258. MR 2000b:30033 Zbl 0931.30017

[Walters 1978] P. Walters, "Invariant measures and equilibrium states for some mappings which expand distances", *Trans. Amer. Math. Soc.* **236** (1978), 121–153. MR 57 #6371

[Zheng 2003] J.-H. Zheng, "Singularities and limit functions in iteration of meromorphic functions", *J. London Math. Soc.* (2) **67**:1 (2003), 195–207. MR 2003j:37071 Zbl 1075.37013

[Zheng 2008] J. H. Zheng, "Dynamics of hyperbolic meromorphic functions", preprint, 2008.

[Zheng 2009] J. H. Zheng, “Conformal measures, Poincaré exponents and dimensions for a class of maps having covering property”, preprint, 2009.

Received January 11, 2010. Revised November 5, 2010.

ZHENG JIAN-HUA
DEPARTMENT OF MATHEMATICAL SCIENCES
TSINGHUA UNIVERSITY
BEIJING 100084
CHINA
jzheng@math.tsinghua.edu.cn

CONTENTS

Volume 250, no. 1 and no. 2

Guidelines for Authors

Authors may submit manuscripts at pjm.math.berkeley.edu/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a 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 250 No. 2 April 2011

0030-8730(201104)250:2;1-F