Pacific Journal of Mathematics

Volume 290 No. 1 September 2017

PACIFIC JOURNAL OF MATHEMATICS

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

msp.org/pjm

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

The subscription price for 2017 is US $450/year for the electronic version, and $625/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.1

THE VIETORIS–RIPS COMPLEXES OF A CIRCLE

MICHAŁ ADAMASZEK AND HENRY ADAMS

Given a metric space X and a distance threshold $r > 0$, the Vietoris–Rips simplicial complex has as its simplices the finite subsets of X of diameter less than r. A theorem of Jean-Claude Hausmann states that if X is a Riemannian manifold and r is sufficiently small, then the Vietoris–Rips complex is homotopy equivalent to the original manifold. Little is known about the behavior of Vietoris–Rips complexes for larger values of r, even though these complexes arise naturally in applications using persistent homology. We show that as r increases, the Vietoris–Rips complex of the circle obtains the homotopy types of the circle, the 3-sphere, the 5-sphere, the 7-sphere, etc., until finally it is contractible. As our main tool we introduce a directed graph invariant, the *winding fraction*, which in some sense is dual to the circular chromatic number. Using the winding fraction we classify the homotopy types of the Vietoris–Rips complex of an arbitrary (possibly infinite) subset of the circle, and we study the expected homotopy type of the Vietoris–Rips complex of a uniformly random sample from the circle. Moreover, we show that as the distance parameter increases, the ambient Čech complex of the circle (i.e., the nerve complex of the covering of a circle by all arcs of a fixed length) also obtains the homotopy types of the circle, the 3-sphere, the 5-sphere, the 7-sphere, etc., until finally it is contractible.

1. Introduction

Given a metric space X and a distance threshold $r > 0$, the *Vietoris–Rips simplicial complex* $\mathbf{VR}_<(X; r)$ has as its simplices the finite subsets of X of diameter less than r. As the maximal simplicial complex determined by its 1-skeleton, it is an example of a clique (or flag) complex. Vietoris–Rips complexes were used by

Research of Adams was supported by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation. Adamaszek is supported by VILLUM FONDEN through the network for Experimental Mathematics in Number Theory, Operator Algebras, and Topology. Part of this research was carried out when Adamaszek was at the Max Planck Institute for Informatics in Saarbrücken.
MSC2010: primary 05E45, 55U10; secondary 68R05.
Keywords: Vietoris–Rips complex, Čech complex, homotopy, clique complex, circular arc, persistent homology.

Vietoris [1927] to define a (co)homology theory for metric spaces, and by Rips [Gromov 1987] to study hyperbolic groups.

More recently, Vietoris–Rips complexes are used in computational algebraic topology and in topological data analysis. In this context the metric space X is often a finite sample from some unknown subset $M \subseteq \mathbb{R}^d$, and one would like to use X to recover topological features of M. The idea behind *topological persistence* is to reconstruct $\mathbf{VR}_<(X; r)$ as the distance threshold r varies from small to large, to disregard short-lived topological features as the result of sampling noise, and to trust topological features which persist as being representative of the shape of M. For example, with *persistent homology* one attempts to reconstruct the homology groups of M from the homology of $\mathbf{VR}_<(X; r)$ as r varies [Edelsbrunner and Harer 2010; Carlsson 2009; Carlsson et al. 2008].

Part of the motivation for using Vietoris–Rips complexes in applied contexts comes from the work of Hausmann and Latschev. Hausmann [1995] proves that if M is a closed Riemannian manifold and if r is sufficiently small compared to the injectivity radius of M, then $\mathbf{VR}_<(M; r)$ is homotopy equivalent to M. Latschev [2001] furthermore shows that if X is Gromov–Hausdorff close to M (for example a sufficiently dense finite sample) and r is sufficiently small, then $\mathbf{VR}_<(X; r)$ recovers the homotopy type of M. As the main idea of persistence is to allow r to vary, we would like to understand what happens when r is beyond the "sufficiently small" range.

As the main result of this paper, we show that as the distance threshold increases, the Vietoris–Rips complex $\mathbf{VR}_<(S^1; r)$ of the circle obtains the homotopy types of the odd-dimensional spheres S^1, S^3, S^5, S^7, …, until finally it is contractible. To our knowledge, this is the first computation for a noncontractible connected manifold M of the homotopy types of $\mathbf{VR}_<(M; r)$ for arbitrary r (and also a first computation of the persistent homology of $\mathbf{VR}_<(M; r)$). Our main result confirms, for the case $M = S^1$, a conjecture of Hausmann [1995, (3.12)] that for M a compact Riemannian manifold, the connectivity of $\mathbf{VR}_<(M; r)$ is a nondecreasing function of the distance threshold r.

As our main tools we introduce *cyclic graphs*, a combinatorial abstraction of Vietoris–Rips graphs for subsets of the circle, and their invariant called the *winding fraction*. In a sense which we make precise, the winding fraction is a directed dual of the *circular chromatic number* of a graph [Hell and Nešetřil 2004, Chapter 6]. In [Adamaszek et al. 2016] we proved that for $X \subseteq S^1$ finite, $\mathbf{VR}_<(X; r)$ is homotopy equivalent to either an odd-dimensional sphere or a wedge sum of spheres of the same even dimension; the theory of winding fractions gives us quantitative control over which homotopy type occurs, and also over the behavior of induced maps between complexes. As applications we classify the homotopy types of $\mathbf{VR}_<(X; r)$ for arbitrary (possibly infinite) subsets $X \subseteq S^1$, and we analyze the evolution of the homotopy type of $\mathbf{VR}_<(X; r)$ when $X \subseteq S^1$ is chosen uniformly at random.

Čech complexes are a second geometric construction producing a simplicial complex from a metric space. The *Čech complex* $\check{\mathbf{C}}_<(S^1; r)$ is defined as the nerve of the collection of all open arcs of length $2r$ in the circle of circumference 1. For r small the nerve theorem [Hatcher 2002, Corollary 4G.3] implies that $\check{\mathbf{C}}_<(S^1; r)$ is homotopy equivalent to S^1. Just as we study $\mathbf{VR}_<(S^1; r)$ in the regime where r is too large for Hausmann's result to apply, we also study $\check{\mathbf{C}}_<(S^1; r)$ in the regime where r is too large for the nerve theorem to apply. We show that as r increases, the ambient Čech complex $\check{\mathbf{C}}_<(S^1; r)$ also obtains the homotopy types of S^1, S^3, S^5, S^7, ..., until finally it is contractible.

All of this has analogues for the complexes $\mathbf{VR}_\le(S^1; r)$ and $\check{\mathbf{C}}_\le(S^1; r)$, defined by sets of diameter at most r, respectively closed arcs of length $2r$. We have:

Main result (Theorems 7.4, 7.6, 9.7, 9.8). *Let* $0 < r < \frac{1}{2}$. *There are homotopy equivalences* (*for* $l = 0, 1, \ldots$):

$$\mathbf{VR}_<(S^1; r) \simeq S^{2l+1} \qquad \text{if } \tfrac{l}{2l+1} < r \le \tfrac{l+1}{2l+3},$$

$$\check{\mathbf{C}}_<(S^1; r) \simeq S^{2l+1} \qquad \text{if } \tfrac{l}{2(l+1)} < r \le \tfrac{l+1}{2(l+2)},$$

$$\mathbf{VR}_\le(S^1; r) \simeq \begin{cases} S^{2l+1} & \text{if } \frac{l}{2l+1} < r < \frac{l+1}{2l+3}, \\ \bigvee^{\mathfrak{c}} S^{2l} & \text{if } r = \frac{l}{2l+1}, \end{cases}$$

$$\check{\mathbf{C}}_\le(S^1; r) \simeq \begin{cases} S^{2l+1} & \text{if } \frac{l}{2(l+1)} < r < \frac{l+1}{2(l+2)}, \\ \bigvee^{\mathfrak{c}} S^{2l} & \text{if } r = \frac{l}{2(l+1)}, \end{cases}$$

Contents of the paper. In Section 2 we introduce preliminary concepts and notation, including Vietoris–Rips complexes. We introduce cyclic graphs and develop their invariant called the winding fraction in Section 3. In Section 4 we show how this invariant affects the homotopy type of the clique complex of a cyclic graph. In Section 5 we show that the homotopy type of the Vietoris–Rips complex stabilizes for sufficiently dense samples of S^1. We apply the winding fraction to study the evolution of Vietoris–Rips complexes for random subsets of S^1 in Section 6. The main result appears in Sections 7 and 8, where we show how to compute the homotopy types of Vietoris–Rips complexes of arbitrary (possibly infinite) subsets of S^1; in particular we describe $\mathbf{VR}_<(S^1; r)$. In Section 9 we transfer these results to the Čech complexes of the circle. The appendices contain proofs of auxiliary results in linear algebra and probability.

2. Preliminaries

We refer the reader to [Hatcher 2002; Kozlov 2008] for basic concepts in topology and combinatorial topology.

Simplicial complexes. For K a simplicial complex, let $V(K)$ be its vertex set. The link of a vertex $v \in V(K)$ is $\mathrm{lk}_K(v) = \{\sigma \in K \mid v \notin \sigma \text{and } \sigma \cup \{v\} \in K\}$. We will identify an abstract simplicial complex with its geometric realization.

Definition 2.1. For an undirected graph G the *clique complex* $\mathrm{Cl}(G)$ is the simplicial complex with vertex set $V(G)$ and with faces determined by all cliques (complete subgraphs) of G.

Vietoris–Rips complexes. The Vietoris–Rips complex is used to capture a notion of proximity in a metric space.

Definition 2.2. Suppose X is a metric space and $r > 0$ is a real number. The *Vietoris–Rips* complex $\mathbf{VR}_{\leq}(X; r)$ is the simplicial complex with vertex set X, where a finite subset $\sigma \subseteq X$ is a face if and only if the diameter of σ is at most r. Analogously, the complex $\mathbf{VR}_{<}(X; r)$ is characterized by finite subsets whose diameter is strictly less than r.

Every Vietoris–Rips complex is the clique complex of its 1-skeleton. We will write $\mathbf{VR}(X; r)$, omitting the subscripts $<$ and $\leq$, in statements which remain true whenever either inequality is applied consequently throughout.

Conventions regarding the circle. We give the circle S^1 the arc-length metric scaled so that the circumference of S^1 is 1. For $x, y \in S^1$ we denote by $[x, y]_{S^1}$ the closed clockwise arc from x to y and by $\vec{d}(x, y)$ its length — the clockwise distance from x to y. For a fixed choice of $0 \in S^1$ each point $x \in S^1$ can be identified with the real number $\vec{d}(0, x) \in [0, 1)$, and this will be our coordinate system on S^1. For any two numbers $x, y \in \mathbb{R}$ we define $[x, y]_{S^1} = [x \bmod 1, y \bmod 1]_{S^1}$ and $\vec{d}(x, y) = \vec{d}(x \bmod 1, y \bmod 1)$. Open and half-open arcs are defined similarly. If $x_1, x_2, \ldots, x_s \in S^1$ then we will write

$$x_1 \prec x_2 \prec \cdots \prec x_s$$

if $x_1, \ldots, x_s$ appear on S^1 in this clockwise order, or equivalently if they are pairwise distinct and $\sum_{i=1}^{s} \vec{d}(x_i, x_{i+1}) = 1$, where $x_{s+1} = x_1$. We define

$$\vec{d}_n(i, j) = n \cdot \vec{d}\left(\tfrac{i}{n}, \tfrac{j}{n}\right)$$

to be the "forward" distance from i to j in $\mathbb{Z}/n$.

Directed graphs. Throughout this work a *directed graph* is a pair $\vec{G} = (V, E)$ with V the set of vertices and $E \subseteq V \times V$ the set of directed edges, subject to the conditions $(v, v) \notin E$ (no loops) and $(v, w) \in E \Longrightarrow (w, v) \notin E$ (no edges oriented in both directions). The edge (v, w) will also be denoted by $v \to w$. For a vertex $v \in V(\vec{G})$ we define the out- and in-neighborhoods

$$N^+(\vec{G}, v) = \{w : v \to w\} \quad \text{and} \quad N^-(\vec{G}, v) = \{w : w \to v\},$$

as well as their closed versions

$$N^+[\vec{G}, v] = N^+(\vec{G}, v) \cup \{v\} \quad \text{and} \quad N^-[\vec{G}, v] = N^-(\vec{G}, v) \cup \{v\}.$$

A *directed cycle* of length s in $\vec{G}$ is a sequence of vertices $v_1, \ldots, v_s$ such that there is an edge $v_i \to v_{i+1}$ for all $i = 1, \ldots, s$, where $v_{s+1} = v_1$.

For a directed graph $\vec{G}$ we will denote by G the *underlying undirected graph* obtained by forgetting the orientations. If G is an undirected graph we write $v \sim w$ when v and w are adjacent, and we define

$$N(G, v) = \{w : w \sim v\} \quad \text{and} \quad N[G, v] = N(G, v) \cup \{v\}.$$

For $v \in V$ let $\vec{G} \setminus v$ be the directed graph obtained by removing vertex v and all of its incident edges, and for $e \in E$ let $\vec{G} \setminus e$ be obtained by removing edge e. The undirected versions $G \setminus v$ and $G \setminus e$ are defined similarly.

All graphs considered in this paper are finite.

3. Cyclic graphs, winding fractions, and dismantling

In this section we develop the combinatorial theory of cyclic graphs, dismantling, and winding fractions.

We are going to work with the notion of a cyclic order. While there exist definitions of a cyclic order based on the abstract ternary relation of betweenness [Huntington 1916], the following definition will be sufficient for our purpose. A *cyclic order* on a finite set S of cardinality n is a bijection $h : S \to \{0, \ldots, n-1\}$. Denoting $x_i = h^{-1}(i)$ we write this simply as

$$x_0 \prec x_1 \prec \cdots \prec x_{n-1}.$$

If $n \geq 3$ this gives rise to a betweenness relation: we write $x_i \prec x_j \prec x_k$ if $i < j < k$ or $k < i < j$ or $j < k < i$. If $S' \subseteq S$ then any cyclic order on S restricts in an obvious way to a cyclic order on S'.

A subinterval in such a cyclic ordering of S is either (1) the empty set, (2) a set of the form $\{x_i, \ldots, x_j\}$ for $0 \leq i \leq j \leq n-1$, or (3) a set of the form $\{x_j, \ldots, x_{n-1}, x_0, \ldots, x_i\}$ for $0 \leq i < j \leq n-1$. In particular, S itself is also a subinterval.

A function $f : S \to T$ between cyclic orders is *cyclic monotone* if (1) for every $t \in T$ the set $f^{-1}(t)$ is a subinterval of S, and (2) $f(s) \prec f(s') \prec f(s'')$ in T implies $s \prec s' \prec s''$ in S for any $s, s', s'' \in S$.

One easily shows that if f is cyclic monotone, $s \prec s' \prec s''$, and $f(s), f(s'), f(s'')$ are pairwise distinct, then $f(s) \prec f(s') \prec f(s'')$. Moreover, if $f : S \to T$ is cyclic monotone then the preimage of any subinterval of T is a subinterval of S.

We will concentrate on the following class of directed graphs.

Definition 3.1. A directed graph $\vec{G}$ is called *cyclic* if its vertices can be arranged in a cyclic order $v_0 \prec v_1 \prec \cdots \prec v_{n-1}$ subject to the following condition: if there is a directed edge $v_i \to v_j$, then either $j \equiv i+1 \mod n$ or there are directed edges

$$v_i \to v_{j-1 \bmod n} \quad \text{and} \quad v_{i+1 \bmod n} \to v_j.$$

In the future all arithmetic operations on the vertex indices are understood to be reduced modulo n; for instance we will write simply v_{i+k} for $v_{i+k \bmod n}$.

Two examples of cyclic graphs are shown in Figure 1. Cyclic graphs are a special case of directed graphs with a *round enumeration*; the latter are defined by the above definition when edges with double (opposite) orientations are allowed. For a comprehensive survey of related graph classes, see [Lin and Szwarcfiter 2009], especially Theorem 10.

We begin with some basic properties of cyclic graphs.

Lemma 3.2. *Let $\vec{G}$ be a cyclic graph with n vertices in cyclic order $v_0 \prec \cdots \prec v_{n-1}$. Then*:

(a) *For every $i = 0, \ldots, n-1$ there exist $s, s' \geq 0$ such that*

$$N^+[\vec{G}, v_i] = \{v_i, v_{i+1}, \ldots, v_{i+s}\} \quad \textit{and} \quad N^-[\vec{G}, v_i] = \{v_{i-s'}, \ldots, v_{i-1}, v_i\}.$$

(b) *For every $i = 0, \ldots, n-1$ we have inclusions*

$$N^+(\vec{G}, v_i) \subseteq N^+[\vec{G}, v_{i+1}] \quad \textit{and} \quad N^-(\vec{G}, v_{i+1}) \subseteq N^-[\vec{G}, v_i].$$

(c) *Every induced subgraph of $\vec{G}$ is a cyclic graph.*

(d) *If $\vec{G}$ contains a directed cycle then $v_i \to v_{i+1}$ for all $i = 0, \ldots, n-1$.*

Proof. Parts (a) and (b) follow directly from the definition. The cyclic orientation inherited from $\vec{G}$ is a cyclic orientation of any induced subgraph, which gives (c). To prove (d), take a directed cycle and replace any edge $v_i \to v_j$ with $j \neq i+1$ by a path $v_i \to v_{i+1} \to v_j$. After finitely many steps of this kind we get a directed cycle in which every edge is of the form $v_i \to v_{i+1}$. □

The main examples of cyclic graphs of interest in this paper are provided in the next two definitions.

Definition 3.3. For a finite subset $X \subseteq S^1$ and real number $0 < r < \frac{1}{2}$, the directed Vietoris–Rips graphs $\overrightarrow{\mathrm{VR}}_{\leq}(X; r)$ and $\overrightarrow{\mathrm{VR}}_{<}(X; r)$ are defined as

$$\overrightarrow{\mathrm{VR}}_{\leq}(X; r) = \big(X,\ \{x_1 \to x_2 : 0 < \vec{d}(x_1, x_2) \leq r\}\big)$$

and

$$\overrightarrow{\mathrm{VR}}_{<}(X; r) = \big(X,\ \{x_1 \to x_2 : 0 < \vec{d}(x_1, x_2) < r\}\big).$$

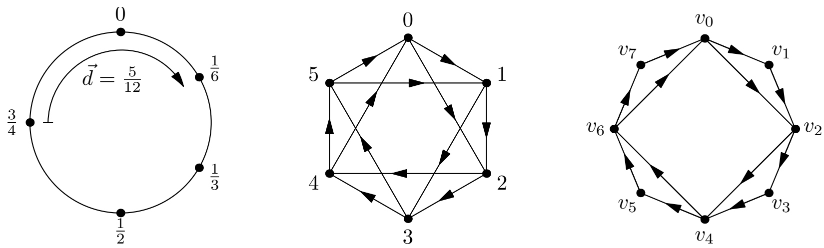


Figure 1. Left: the coordinate system on S^1. Middle: the cyclic graph $\vec{C}_6^2$. Right: a cyclic graph which is not a Vietoris–Rips graph. Its odd-numbered vertices are dominated; see Definition 3.9.

It is clear that the Vietoris–Rips graph is cyclic with respect to the clockwise ordering of X; in particular the two meanings of the symbol $\prec$ denoting clockwise order in S^1 and cyclic order of the vertices of $\overrightarrow{\mathrm{VR}}(X; r)$ agree.

As before, we will omit the subscript and write $\overrightarrow{\mathrm{VR}}(X; r)$ in statements which apply to both $<$ and $\leq$. Not every cyclic graph is a Vietoris–Rips graph of a subset of S^1: an example is in Figure 1. Our interest in Vietoris–Rips graphs stems from the fact that a Vietoris–Rips complex is the clique complex of the corresponding undirected Vietoris–Rips graph, namely $\mathbf{VR}(X; r) = \mathrm{Cl}(\mathrm{VR}(X; r))$.

Definition 3.4. For integers n and k with $0 \leq k < \frac{1}{2}n$, the directed graph $\vec{C}_n^k$ has vertex set $\{0, \dots, n-1\}$ and edges $i \to (i+s) \bmod n$ for all $i = 0, \dots, n-1$ and $s = 1, \dots, k$. Equivalently

$$i \to j \quad \text{if and only if} \quad 0 < \vec{d}_n(i, j) \leq k.$$

The directed graphs $\vec{C}_n^k$ are cyclic with respect to the natural cyclic order of the vertices. Note that $\vec{C}_n^k$ is a Vietoris–Rips graph of the vertex set of a regular n-gon, or in our notation:

$$\vec{C}_n^k = \overrightarrow{\mathrm{VR}}_{\leq}\left(\left\{0, \tfrac{1}{n}, \dots, \tfrac{n-1}{n}\right\}; \tfrac{k}{n}\right). \tag{1}$$

The cyclic graphs $\vec{C}_n^k$ will play a prominent role in our analysis of the Vietoris–Rips graphs.

A *homomorphism of directed graphs* $f\colon \vec{G} \to \vec{H}$ is a vertex map such that for every edge $v \to w$ in $\vec{G}$ either $f(v) = f(w)$ or there is an edge $f(v) \to f(w)$ in $\vec{H}$. Directed graphs with homomorphisms form a category. We now define a class of homomorphisms for the subcategory of cyclic graphs.

Definition 3.5. Suppose that $\vec{G}$ and $\vec{H}$ are cyclic graphs, with vertex ordering $v_0 \prec \cdots \prec v_{n-1}$ in $\vec{G}$. A vertex map $f\colon \vec{G} \to \vec{H}$ is a *cyclic homomorphism* if f is cyclic monotone, if f is a homomorphism of directed graphs, and if f is not constant whenever $\vec{G}$ has a directed cycle.

Note that if $\vec{H} = \overrightarrow{\mathrm{VR}}(X;r)$ then the condition "f is cyclic monotone and not constant" is equivalent to the equation

$$\sum_i \vec{d}(f(v_i), f(v_{i+1})) = 1. \tag{2}$$

Lemma 3.6. *Cyclic homomorphisms have the following properties.*

(a) *If $f\colon \vec{G} \to \vec{H}$ is a cyclic homomorphism and $\vec{G}$ has a directed cycle then so does $\vec{H}$.*

(b) *The composition of two cyclic homomorphisms is a cyclic homomorphism.*

(c) *The inclusion of a cyclic subgraph* (*with inherited cyclic orientation*) *is a cyclic homomorphism.*

Proof. For (a), note that if $\vec{G}$ has a directed cycle, then by Lemma 3.2(d) it has a directed cycle $\cdots \to v_i \to v_{i+1} \to \cdots$ through all its vertices. The image of that cycle under f is not constant, and by removing adjacent repetitions one gets a directed cycle in $\vec{H}$.

For (b) suppose $f\colon \vec{G} \to \vec{H}$ and $g : \vec{H} \to \vec{K}$ are cyclic homomorphisms of cyclic graphs. We first check that gf is cyclic monotone. Indeed, for a vertex w in $\vec{K}$ the set $(gf)^{-1}(w)$ is the preimage under f of the subinterval $g^{-1}(w)$, hence a subinterval. If $gf(v) \prec gf(v') \prec gf(v'')$, then using that f and then g are cyclic monotone, we get $v \prec v' \prec v''$.

Now we only have to check that if $\vec{G}$ has a directed cycle then gf is not constant. By part (a), all three graphs have a directed cycle. Suppose, on the contrary, that $g(f(v_i)) = w$ for all $v_i \in V(\vec{G})$. Since g is not constant there is a vertex u of $\vec{H}$ not in the image of f with $g(u) \neq w$, and since f is not constant there is an index i such that $f(v_i) \prec u \prec f(v_{i+1})$ is cyclically ordered in $\vec{H}$. Since $v_i \to v_{i+1}$ in $\vec{G}$ we have $f(v_i) \to f(v_{i+1})$ in $\vec{H}$. That in turn implies $f(v_i) \to u \to f(v_{i+1})$ in $\vec{H}$ and therefore $w \to g(u) \to w$ in $\vec{K}$. This contradicts our definition of a directed graph (no edges oriented in both directions), and hence gf is not constant.

Part (c) is clear. □

We can now define the main numerical invariant of cyclic graphs.

Definition 3.7. The *winding fraction* of a cyclic graph $\vec{G}$ is

$$\mathrm{wf}(\vec{G}) = \sup\left\{\tfrac{k}{n} : \text{there exists a cyclic homomorphism } \vec{C}_n^k \to \vec{G}\right\}.$$

For a finite subset $X \subseteq S^1$ we also introduce the shorthand notation

$$\mathrm{wf}_{\le}(X;r) = \mathrm{wf}(\overrightarrow{\mathrm{VR}}_{\le}(X;r)) \quad \text{and} \quad \mathrm{wf}_{<}(X;r) = \mathrm{wf}(\overrightarrow{\mathrm{VR}}_{<}(X;r)).$$

The next proposition records the basic properties of the winding fraction.

Proposition 3.8. *The winding fraction satisfies the following properties.*

(a) $\mathrm{wf}(\vec{G}) > 0$ *if and only if* $\vec{G}$ *has a directed cycle.*

(b) *If* $\vec{G} \to \vec{H}$ *is a cyclic homomorphism then* $\mathrm{wf}(\vec{G}) \leq \mathrm{wf}(\vec{H})$.

(c) *If* $X \subseteq S^1$ *is a finite set and* $0 < r < \frac{1}{2}$ *then*

$$\mathrm{wf}_{\leq}(X;r) \leq r, \quad \mathrm{wf}_{<}(X;r) < r, \quad \textit{and} \quad \mathrm{wf}_{<}(X;r) \leq \mathrm{wf}_{\leq}(X;r).$$

(d) $\mathrm{wf}(\vec{C}_n^k) = \frac{k}{n}$.

Proof. For (a) note that if $\vec{G}$ has a directed cycle, then by Lemma 3.2(d) the map $i \mapsto v_i$ defines a cyclic homomorphism $\vec{C}_n^1 \to \vec{G}$ with $n = |V(\vec{G})|$. Conversely, if $\vec{G}$ has no directed cycle then by Lemma 3.6(a) it admits cyclic homomorphisms only from the graphs $\vec{C}_n^0$.

Part (b) follows from the definition of the winding fraction and the fact that a composition of cyclic homomorphisms is a cyclic homomorphism.

Now we prove the first inequality of (c). Suppose that $f \colon \vec{C}_n^k \to \vec{\mathrm{VR}}_{\leq}(X;r)$ is a cyclic homomorphism with $k \geq 1$, which means that for every $i = 0, \dots, n-1$ we have $\vec{d}(f(i), f(i+k)) \leq r$. Since every arc of the form $[f(j), f(j+1)]_{S^1}$ is covered by exactly k arcs $[f(i), f(i+k)]_{S^1}$,

$$nr \geq \sum_i \vec{d}(f(i), f(i+k)) = k \sum_j \vec{d}(f(j), f(j+1)) = k,$$

where in the last equality we used (2). It follows that $\frac{k}{n} \leq r$ and $\mathrm{wf}_{\leq}(X;r) \leq r$.

The second inequality has similar proof with strict inequalities and the third one follows from (b) since we have a subgraph inclusion $\vec{\mathrm{VR}}_{<}(X;r) \hookrightarrow \vec{\mathrm{VR}}_{\leq}(X;r)$.

For (d), the identity automorphism of $\vec{C}_n^k$ shows $\mathrm{wf}(\vec{C}_n^k) \geq \frac{k}{n}$. Conversely, applying part (c) with $X = \left\{0, \frac{1}{n}, \dots, \frac{n-1}{n}\right\}$ and $r = \frac{k}{n}$ gives, by (1), that $\mathrm{wf}(\vec{C}_n^k) \leq \frac{k}{n}$. □

We now describe a practical way of computing the winding fraction. The method uses graph reductions modeled on the notion of *dismantlings* of undirected graphs (called *folds* in [Hell and Nešetřil 2004, Section 2.11] or *LC reductions* in [Matoušek 2008]), and hence we use the same terminology.

Definition 3.9. Suppose $\vec{G}$ is a cyclic graph with vertex ordering $v_0 \prec \cdots \prec v_{n-1}$. A vertex v_i is *dominated by* v_{i+1} (or just *dominated*) if $N^-(\vec{G}, v_{i+1}) = N^-[\vec{G}, v_i]$.

Lemma 3.10. *If* $\vec{G}$ *is a cyclic graph and* v_i *is dominated by* v_{i+1}, *then the map* $f \colon \vec{G} \to \vec{G} \setminus v_i$ *given by*

$$f(v_j) = \begin{cases} v_j & \text{if } j \neq i, \\ v_{i+1} & \text{if } j = i \end{cases}$$

is a cyclic homomorphism. The composition $\vec{G} \setminus v_i \hookrightarrow \vec{G} \xrightarrow{f} \vec{G} \setminus v_i$ *is the identity.*

Proof. We first check that f is a homomorphism of directed graphs. First note the map f preserves all edges avoiding v_i. If $v_k \to v_i$ then $v_k \to v_{i+1}$ because v_i is dominated by v_{i+1}. If $v_i \to v_k$ then either $k = i + 1$, and then $f(v_i) = f(v_k)$, or there is also an edge $v_{i+1} \to v_k$ because $\vec{G}$ is cyclic.

The map f is a cyclic homomorphism because it clearly preserves the cyclic ordering, and if $\vec{G}$ has a directed cycle then it has at least three vertices, in which case f is not constant.

The last claim is obvious. $\square$

The removal of a dominated vertex can be repeated as long as the new graph has a dominated vertex.

Definition 3.11. We say a cyclic graph $\vec{G}$ *dismantles* to an induced subgraph $\vec{H}$ if there is a sequence of graphs $\vec{G} = \vec{G}_0, \vec{G}_1, \dots, \vec{G}_s = \vec{H}$ such that $\vec{G}_i$ is obtained from $\vec{G}_{i-1}$ by removing a dominated vertex for $i = 1, \dots, s$.

If $\vec{G}$ dismantles to $\vec{H}$ then the composition of cyclic homomorphisms $\vec{G}_i \to \vec{G}_{i+1}$ provided by Lemma 3.10 gives a cyclic homomorphism $\vec{G} \to \vec{H}$. Moreover the composition

$$\vec{H} \hookrightarrow \vec{G} \to \vec{H}$$

is the identity of $\vec{H}$. The next proposition answers the question of when the dismantling process of a cyclic graph must stop.

Proposition 3.12. *A cyclic graph without a dominated vertex is isomorphic to $\vec{C}_n^k$ for some $0 \le k < \frac{1}{2}n$. As a consequence every cyclic graph dismantles to an induced subgraph of the form $\vec{C}_n^k$.*

Proof. Let $\vec{G}$ be a cyclic graph with vertex ordering $v_0 \prec \cdots \prec v_{n-1}$ and with no dominated vertex. By Lemma 3.2(a) for every $j = 0, \dots, n-1$ there is an $e(j)$ such that $N^+[\vec{G}, v_j] = \{v_j, \dots, v_{e(j)}\}$. For every $i = 0, \dots, n-1$,

$$N^-[\vec{G}, v_i] \setminus N^-(\vec{G}, v_{i+1}) = \{v_j : e(j) = i\},$$

where $N^-(\vec{G}, v_{i+1}) \subseteq N^-[\vec{G}, v_i]$ by Lemma 3.2(b). It follows that

$$\sum_i |N^-[\vec{G}, v_i] \setminus N^-(\vec{G}, v_{i+1})| = n.$$

Since $\vec{G}$ has no dominated vertices, all n summands above are positive and therefore all are equal to 1. We have

$$|N^-(\vec{G}, v_{i+1})| = |N^-[\vec{G}, v_i]| - 1 = |N^-(\vec{G}, v_i) \cup \{v_i\}| - 1 = |N^-(\vec{G}, v_i)|.$$

Denote the common value of $|N^-(\vec{G}, v_i)|$ by k. Using Lemma 3.2(a) again we see that $N^-[\vec{G}, v_i] = \{v_{i-k}, \dots, v_i\}$ for all i, and so $\vec{G}$ is isomorphic to $\vec{C}_n^k$. $\square$

Remark 3.13. In [Adamaszek et al. 2017] we prove that, regardless of the choices of a dominated vertex made in the process, every dismantling of a cyclic graph ends up with the same subgraph. Such strong uniqueness is not needed in this paper.

Our notion is modeled on the more classical dismantling of undirected graphs; see [Hell and Nešetřil 2004]. In that setting the end result of the dismantling process is unique only up to isomorphism; see [Matoušek 2008] or [Hell and Nešetřil 2004, Theorem 2.60].

We can now give a recipe for computing the winding fraction.

Proposition 3.14. *If a cyclic graph $\vec{G}$ dismantles to $\vec{C}_n^k$ then* $\mathrm{wf}(\vec{G}) = \frac{k}{n}$.

Proof. The graph $\vec{G}$ has both cyclic homomorphisms $\vec{C}_n^k \hookrightarrow \vec{G}$ and $\vec{G} \to \vec{C}_n^k$, so the claim follows from Proposition 3.8 parts (b) and (d). □

The following result gives the converse of Proposition 3.8(b).

Proposition 3.15. *There is a cyclic homomorphism $f\colon \vec{G} \to \vec{H}$ if and only if* $\mathrm{wf}(\vec{G}) \le \mathrm{wf}(\vec{H})$.

Proof. The "only if" part is handled by Proposition 3.8(b).

For any $0 \le k < \frac{1}{2}n$ and $d \ge 1$ consider two maps $\iota : \vec{C}_n^k \to \vec{C}_{nd}^{kd}$ and $\tau : \vec{C}_{nd}^{kd} \to \vec{C}_n^k$ given by

$$\iota(i) = di, \quad \text{and} \quad \tau(j) = \left\lfloor \tfrac{j}{d} \right\rfloor.$$

It is easy to see that ι and τ are cyclic homomorphisms.

To prove the "if" part, suppose that $\vec{G}$ dismantles to $\vec{C}_n^k$ and $\vec{H}$ dismantles to $\vec{C}_{n'}^{k'}$. Proposition 3.14 and the assumption $\mathrm{wf}(\vec{G}) \le \mathrm{wf}(\vec{H})$ imply $\frac{k}{n} \le \frac{k'}{n'}$. Then we have a cyclic homomorphism

$$\vec{G} \to \vec{C}_n^k \xrightarrow{\iota} \vec{C}_{nn'}^{kn'} \hookrightarrow \vec{C}_{nn'}^{nk'} \xrightarrow{\tau} \vec{C}_{n'}^{k'} \hookrightarrow \vec{H},$$

where the first and last map come from dismantling, and the middle map is a subgraph inclusion since $kn' \le nk'$. □

The winding fraction is in a sense dual to the well-studied concept of *circular chromatic number*; see [Hell and Nešetřil 2004, Chapter 6]. For an arbitrary undirected graph G the circular chromatic number $\chi_c(G)$ is defined as the infimum over numbers $\frac{n}{k}$ such that there is a map $V(G) \to \mathbb{Z}/n$ which maps every edge to a pair of numbers *at least* k apart. By Proposition 3.15 we have

$$\mathrm{wf}(\vec{G}) = \inf\left\{ \tfrac{k}{n} : \text{there exists a cyclic homomorphism } \vec{G} \to \vec{C}_n^k \right\}$$

which leads to the following description: $\mathrm{wf}(\vec{G})$ is the infimum over numbers $\frac{k}{n}$ such that there is an order-preserving map $V(G) \to \mathbb{Z}/n$ which maps every edge to a pair of numbers *at most* k apart.

4. Winding fractions determine homotopy types

We now analyze the influence of the winding fraction $\mathrm{wf}(\vec{G})$ on the topology of the clique complex $\mathrm{Cl}(G)$.

A homomorphism $f\colon G \to H$ of undirected graphs is a vertex map such that $v \sim w$ implies $f(v) = f(w)$ or $f(v) \sim f(w)$. Every homomorphism of directed graphs $\vec{G} \to \vec{H}$ determines a homomorphism of the underlying undirected graphs $G \to H$, and in turn also a simplicial map $\mathrm{Cl}(G) \to \mathrm{Cl}(H)$. The assignment $\vec{G} \mapsto \mathrm{Cl}(G)$ is a functor from the category of directed graphs to topological spaces, and also a functor from the subcategory of cyclic graphs to topological spaces.

Lemma 4.1. *If $\vec{G}$ is a cyclic graph and v_i is a dominated vertex, then the cyclic homomorphisms $\vec{G} \setminus v_i \hookrightarrow \vec{G}$ and $\vec{G} \to \vec{G} \setminus v_i$ from Lemma 3.10 induce homotopy equivalences of clique complexes.*

Proof. Using the conditions listed in Lemma 3.2(b) and Definition 3.9 we get

$$N[G, v_i] = N^-[\vec{G}, v_i] \cup N^+(\vec{G}, v_i) \subseteq N^-(\vec{G}, v_{i+1}) \cup N^+[\vec{G}, v_{i+1}] = N[G, v_{i+1}].$$

Hence the link $\mathrm{lk}_{\mathrm{Cl}(G)}(v_i)$ is a cone with apex v_{i+1}, or in other words, $\mathrm{Cl}(G)$ is obtained from $\mathrm{Cl}(G \setminus v_i)$ by attaching a cone over a cone. It follows that the inclusion $\mathrm{Cl}(G \setminus v_i) \hookrightarrow \mathrm{Cl}(G)$ is a homotopy equivalence. Since the composition $\mathrm{Cl}(G \setminus v_i) \hookrightarrow \mathrm{Cl}(G) \to \mathrm{Cl}(G \setminus v_i)$ is the identity, also $\vec{G} \to \vec{G} \setminus v_i$ induces a homotopy equivalence. □

Corollary 4.2. *If a cyclic graph $\vec{G}$ dismantles to $\vec{H}$ then the maps of clique complexes induced by $\vec{H} \hookrightarrow \vec{G}$ and $\vec{G} \to \vec{H}$ are homotopy equivalences.*

To determine the homotopy types of $\mathrm{Cl}(G)$ for arbitrary cyclic graphs $\vec{G}$ we recall the following result, proved with different methods in [Adamaszek 2013] and [Adamaszek et al. 2016].

Theorem 4.3. *For $0 \le k < \frac{1}{2}n$ there are homotopy equivalences*

$$\mathrm{Cl}(C_n^k) \simeq \begin{cases} S^{2l+1} & \text{if } \frac{l}{2l+1} < \frac{k}{n} < \frac{l+1}{2l+3} \text{ for some } l = 0, 1, \ldots, \\ \bigvee^{n-2k-1} S^{2l} & \text{if } \frac{k}{n} = \frac{l}{2l+1} \text{ for some } l = 0, 1, \ldots. \end{cases}$$

By convention an empty wedge sum is a point. We immediately obtain the following result.

Theorem 4.4. *If $\vec{G}$ is a cyclic graph then*

$$\mathrm{Cl}(G) \simeq \begin{cases} S^{2l+1} & \text{if } \frac{l}{2l+1} < \mathrm{wf}(\vec{G}) < \frac{l+1}{2l+3} \text{ for some } l = 0, 1, \ldots, \\ \bigvee^{n-2k-1} S^{2l} & \text{if } \mathrm{wf}(\vec{G}) = \frac{l}{2l+1} \text{ and } \vec{G} \text{ dismantles to } \vec{C}_n^k. \end{cases}$$

Proof. Graph $\vec{G}$ dismantles to some $\vec{C}_n^k$ for $0 \le k < \frac{1}{2}n$ by Proposition 3.12, and then we have $\mathrm{Cl}(G) \simeq \mathrm{Cl}(C_n^k)$ by Corollary 4.2. From Proposition 3.14 we get $\mathrm{wf}(\vec{G}) = \frac{k}{n}$, and plugging this into Theorem 4.3 gives the result. □

Corollary 4.5. *If $X \subseteq S^1$ is a finite set and $0 \le r < \frac{1}{2}$ then*

$$\mathbf{VR}(X;r) \simeq \begin{cases} S^{2l+1} & \text{if } \frac{l}{2l+1} < \mathrm{wf}(X;r) < \frac{l+1}{2l+3} \text{ for some } l = 0, 1, \ldots, \\ \bigvee^{n-2k-1} S^{2l} & \text{if } \mathrm{wf}(X;r) = \frac{l}{2l+1} \text{ and } \vec{\mathrm{VR}}(X;r) \text{ dismantles to } \vec{C}_n^k. \end{cases}$$

Proof. For the cyclic graph $\vec{\mathrm{VR}}(X;r)$ we have $\mathbf{VR}(X;r) = \mathrm{Cl}(\mathrm{VR}(X;r))$. □

Remark 4.6. A *circular-arc graph* (CA) is an intersection graph of a collection of arcs in S^1. A circular-arc graph is *proper* (PCA) if no arc contains another and *unit* (UCA) if all arcs have the same length. We have inclusions of graph classes UCA $\subsetneq$ PCA $\subsetneq$ CA. If $\vec{G}$ is a cyclic graph then one can show G is a PCA graph, and if $X \subseteq S^1$ is finite and $0 \le r < \frac{1}{2}$ then the Vietoris–Rips graph $\mathrm{VR}(X;r)$ is a UCA graph. In [Adamaszek et al. 2016] we proved that the clique complex of any CA graph has the homotopy type of S^{2l+1} or a wedge of copies of S^{2l} for some $l \ge 0$. The theory of winding fractions refines the result by providing quantitative control over which homotopy type occurs, and by allowing us to understand induced maps. These features will be crucial for the applications we present in the following sections.

There is a clear difference in the behavior of $\mathrm{Cl}(G)$ when $\mathrm{wf}(\vec{G})$ is one of the *singular* values $\frac{l}{2l+1}, l = 0, 1, \ldots$ as opposed to a *generic* value $\frac{l}{2l+1} < \mathrm{wf}(\vec{G}) < \frac{l+1}{2l+3}$. We now discuss additional properties of $\mathrm{Cl}(G)$ in the generic situation. The next lemmas describe the effect of a vertex or edge removal on the homotopy type of $\mathrm{Cl}(G)$.

Lemma 4.7. *Suppose that $\vec{G}$ is a cyclic graph and $v \in V(\vec{G})$. If*

$$\frac{l}{2l+1} < \mathrm{wf}(\vec{G} \setminus v) \le \mathrm{wf}(\vec{G}) < \frac{l+1}{2l+3},$$

then the inclusion $\vec{G} \setminus v \hookrightarrow \vec{G}$ induces a homotopy equivalence of clique complexes.

Proof. By Theorem 4.4 the complexes $\mathrm{Cl}(G \setminus v)$ and $\mathrm{Cl}(G)$ are both homotopy equivalent to S^{2l+1}. Let $\vec{G}_v$ denote the cyclic subgraph of $\vec{G}$ induced by $N(G, v)$, so $\mathrm{lk}_{\mathrm{Cl}(G)}(v) = \mathrm{Cl}(G_v)$. The decomposition $\mathrm{Cl}(G) = \mathrm{Cl}(G \setminus v) \cup_{\mathrm{Cl}(G_v)} (\mathrm{Cl}(G_v) * v)$ yields a Mayer–Vietoris long exact sequence of homology groups whose only nontrivial part is

$$0 \to \widetilde{H}_{2l+1}(\mathrm{Cl}(G_v)) \to \underset{\substack{\| \\ \mathbb{Z}}}{\widetilde{H}_{2l+1}(\mathrm{Cl}(G \setminus v))} \to \underset{\substack{\| \\ \mathbb{Z}}}{\widetilde{H}_{2l+1}(\mathrm{Cl}(G))} \to \widetilde{H}_{2l}(\mathrm{Cl}(G_v)) \to 0 \tag{3}$$

Since $\vec{G}_v$ is cyclic, by Theorem 4.4 the homology $\widetilde{H}_*(\mathrm{Cl}(G_v))$ is free and concentrated in at most one dimension. In view of (3) this is possible only if $\widetilde{H}_*(\mathrm{Cl}(G_v))=0$ and the middle map in (3) is an isomorphism. So $\vec{G}\setminus v \hookrightarrow \vec{G}$ induces a homology isomorphism between spaces homotopy equivalent to S^{2l+1}, and hence is a homotopy equivalence by the Hurewicz and Whitehead theorems. □

Lemma 4.8. *Suppose that $\vec{G}$ is a cyclic graph and $e \in E(\vec{G})$ is an edge such that $\vec{G}\setminus e$ is also a cyclic graph. If*

$$\tfrac{l}{2l+1} < \mathrm{wf}(\vec{G}\setminus e) \le \mathrm{wf}(\vec{G}) < \tfrac{l+1}{2l+3},$$

then the inclusion $\vec{G}\setminus e \hookrightarrow \vec{G}$ induces a homotopy equivalence of clique complexes.

Proof. Let $e=(a,b)$ and denote by $\vec{G}_e$ the cyclic subgraph of $\vec{G}$ induced by $N(G,a)\cap N(G,b)$. Then we have a decomposition

$$\mathrm{Cl}(G)=\mathrm{Cl}(G\setminus e)\cup_{\mathrm{Cl}(G_e)*\{a,b\}}(\mathrm{Cl}(G_e)*e)=\mathrm{Cl}(G\setminus e)\cup_{\Sigma\ \mathrm{Cl}(G_e)}(\mathrm{Cl}(G_e)*e).$$

By Mayer–Vietoris this yields the exact sequence

$$\begin{array}{ccccccccc}0 \to \widetilde{H}_{2l}(\mathrm{Cl}(G_e)) & \to & \widetilde{H}_{2l+1}(\mathrm{Cl}(G\setminus e)) & \to & \widetilde{H}_{2l+1}(\mathrm{Cl}(G)) & \to & \widetilde{H}_{2l-1}(\mathrm{Cl}(G_e)) \to 0 \\ & & \| & & \| & & \\ & & \mathbb{Z} & & \mathbb{Z} & & \end{array}$$

where $\widetilde{H}_k(\mathrm{Cl}(G_e))=\widetilde{H}_{k+1}(\Sigma\,\mathrm{Cl}(G_e))$. The proof can now be completed as in Lemma 4.7. □

Proposition 4.9. *Suppose $f\colon \vec{G}\to\vec{H}$ is a cyclic homomorphism and*

$$\tfrac{l}{2l+1} < \mathrm{wf}(\vec{G}) \le \mathrm{wf}(\vec{H}) < \tfrac{l+1}{2l+3}.$$

Then f induces a homotopy equivalence of clique complexes.

Proof. We proceed in three stages. First, suppose that $f\colon\vec{G}\to\vec{H}$ is injective on the vertices, i.e., it is an inclusion of a subgraph (not necessarily induced). In that case f can be factored as a composition of cyclic homomorphisms

$$\vec{G}=\vec{G}_0\hookrightarrow\vec{G}_1\hookrightarrow\cdots\hookrightarrow\vec{G}_s=\vec{H}$$

where each inclusion $\vec{G}_i\hookrightarrow\overrightarrow{G_{i+1}}$ is an extension by a single vertex or by a single edge. Since $\frac{l}{2l+1}<\mathrm{wf}(\vec{G})\le\mathrm{wf}(\vec{G}_i)\le\mathrm{wf}(\vec{H})<\frac{l+1}{2l+3}$, the result follows from Lemmas 4.7 and 4.8.

Next, we prove the statement for an arbitrary cyclic homomorphism $f\colon\vec{C}_n^k\to\vec{C}_{n'}^{k'}$ with

$$\tfrac{l}{2l+1}<\tfrac{k}{n}\le\tfrac{k'}{n'}<\tfrac{l+1}{2l+3}.$$

Our first goal is to find a factorization $f=\tau\circ f_d$ where $f_d:\vec{C}_n^k\to\vec{C}_{dn'}^{dk'}$ is injective and $\tau:\vec{C}_{dn'}^{dk'}\to\vec{C}_{n'}^{k'}$ is given by $\tau(j)=\lfloor\frac{j}{d}\rfloor$.

Let $j_0 \prec \cdots \prec j_s$, with $1 \leq s \leq n-1$, be the cyclically ordered vertices of the image of f in $\vec{C}_{n'}^{k'}$. Since f is a cyclic homomorphism, each preimage $f^{-1}(j_q)$ is an interval modulo n. Now define the cyclically ordered vertices $i_0 \prec \cdots \prec i_s$ in $\vec{C}_n^k$ by $f^{-1}(j_q) = \{i_q, \ldots, i_{q+1}-1\}$. Choose $d \geq \max\{|f^{-1}(j_q)| : q = 0, \ldots, s\}$ and define a map $f_d : \vec{C}_n^k \to \vec{C}_{dn'}^{dk'}$ by

$$f_d(i) = dj_q + \vec{d}_n(i_q, i) \quad \text{for } i \in \{i_q, \ldots, i_{q+1}-1\}.$$

Note that $0 \leq i - i_q < |f^{-1}(j_q)| \leq d$; therefore f_d preserves the cyclic ordering and hence is a cyclic homomorphism so long as it is a homomorphism of directed graphs. It suffices to check that for every $i = 0, \ldots, n-1$ we have

$$\vec{d}_{dn'}(f_d(i), f_d(i+k)) \leq dk'.$$

Suppose that $i \in f^{-1}(j_q)$ and $i + k \in f^{-1}(j_{q'})$; necessarily $\vec{d}_{n'}(j_q, j_{q'}) \leq k'$. If $\vec{d}_{n'}(j_q, j_{q'}) \leq k' - 1$ then

$$\vec{d}_{dn'}(f_d(i), f_d(i+k)) \leq \vec{d}_{dn'}(dj_q, dj_{q'} + d) \leq dk'.$$

If $j_{q'} = j_q + k'$ then

$$\begin{aligned}\vec{d}_{dn'}(f_d(i), f_d(i+k)) &= dk' + \vec{d}_n(i_{q'}, i+k) - \vec{d}_n(i_q, i) \\ &= dk' + \vec{d}_n(i, i+k) - \vec{d}_n(i_q, i_{q'}) = dk' + k - \vec{d}_n(i_q, i_{q'}).\end{aligned}$$

We have $\vec{d}_n(i_q, i_{q'}) \geq k$, for otherwise $\vec{d}_n(i_q - 1, i_{q'}) \leq k$ and $\vec{d}_{n'}(f(i_q - 1), f(i_{q'})) = \vec{d}_{n'}(j_{q-1}, j_{q'}) \geq k' + 1$ would contradict the fact that f is a homomorphism. This ends the proof that f_d is a cyclic homomorphism.

Consider the cyclic homomorphisms $\iota : \vec{C}_{n'}^{k'} \to \vec{C}_{dn'}^{dk'}$ and $\tau : \vec{C}_{dn'}^{dk'} \to \vec{C}_{n'}^{k'}$ given by

$$\iota(i) = di \quad \text{and} \quad \tau(j) = \left\lfloor \tfrac{j}{d} \right\rfloor.$$

We have a commutative diagram

$$\begin{tikzcd} \vec{C}_n^k \arrow[r, "f_d", "\simeq"'] \arrow[dr, "f"'] & \vec{C}_{dn'}^{dk'} \arrow[d, "\tau"] & \vec{C}_{n'}^{k'} \arrow[l, "\iota"', "\simeq"] \arrow[dl, "\simeq"', "\mathrm{id}"] \\ & \vec{C}_{n'}^{k'} & \end{tikzcd}$$

where $\simeq$ indicates the map induces a homotopy equivalence of clique complexes; for the inclusions f_d and ι this follows from the first part of the proof. From the diagram we conclude that f induces a homotopy equivalence.

Finally, to prove the general case, suppose that $\vec{G}$ dismantles to $\vec{C}_n^k$ and $\vec{H}$ dismantles to $\vec{C}_{n'}^{k'}$ with $\mathrm{wf}(\vec{G}) = \frac{k}{n} \leq \frac{k'}{n'} = \mathrm{wf}(\vec{H})$. The composition

$$\vec{C}_n^k \xhookrightarrow{\simeq} \vec{G} \xrightarrow{f} \vec{H} \xrightarrow{\simeq} \vec{C}_{n'}^{k'}$$

induces a homotopy equivalence of clique complexes, and therefore so does f. $\square$

We defer until Section 8 a further study of the combinatorics of $\mathrm{Cl}(G)$ when $\mathrm{wf}(\vec{G}) = \frac{l}{2l+1}$ is a singular value.

5. Density implies stability

In this section we make precise the heuristic observation that the winding fraction $\mathrm{wf}(X;r)$ increases with the density of X in S^1. For this we recall the notion of covering from metric geometry.

Definition 5.1. A subset X of a metric space M is an *ε-covering* if every point of M is within distance less than ε from some point in X.

A finite subset $X \subseteq S^1$ is an ε-covering of S^1 if and only if every two cyclically consecutive points in X are less than 2ε apart.

As motivation for this section, we note that if $0 < r < \frac{1}{3}$ and $X \subseteq S^1$ is a finite subset, then $\mathbf{VR}_<(X;r) \simeq S^1$ if and only if X is an $\left(\frac{r}{2}\right)$-covering of S^1. The next proposition is an analogue of this observation for bigger winding fractions and therefore for higher-dimensional homotopy types of $\mathbf{VR}_<(X;r)$.

Proposition 5.2. *Suppose that $0 < r < \frac{1}{2}$ and $X \subseteq S^1$ is a finite subset. If X is an ε-covering of S^1 for some $\varepsilon > 0$ then $\mathrm{wf}_<(X;r) > r - 2\varepsilon$.*

Proof. We can assume that $r - 2\varepsilon > 0$. There exists an $\varepsilon' < \varepsilon$ such that X is also an ε'-covering. It suffices to show that whenever $0 < \frac{k}{n} < r - 2\varepsilon'$ then there is a cyclic homomorphism $\vec{C}_n^k \to \vec{\mathrm{VR}}_<(X;r)$, since then we get

$$\mathrm{wf}_<(X;r) \geq r - 2\varepsilon' > r - 2\varepsilon.$$

Fix $0 < \frac{k}{n} < r - 2\varepsilon'$. For every $i = 0, \ldots, n-1$ let $x_i \in X$ be the point closest to $\frac{i}{n}$. (The uniqueness of each x_i can be assured by an infinitesimal rotation, if necessary) Then $x_0, \ldots x_{n-1}$ appear on S^1 in this clockwise order (possibly with repetitions) and, since $\varepsilon' < \frac{1}{2}r < \frac{1}{4}$, not all of the x_i are the same. By the triangle inequality

$$\begin{aligned} d(x_i, x_{i+k}) &\leq d\left(x_i, \tfrac{i}{n}\right) + d\left(\tfrac{i}{n}, \tfrac{i+k}{n}\right) + d\left(\tfrac{i+k}{n}, x_{i+k}\right) \\ &< \varepsilon' + \tfrac{k}{n} + \varepsilon' = \tfrac{k}{n} + 2\varepsilon' < r. \end{aligned}$$

It follows that the map $i \mapsto x_i$ determines a cyclic homomorphism $\vec{C}_n^k \to \vec{\mathrm{VR}}_<(X;r)$, and the proof is complete. $\square$

This leads to the following conclusion.

Proposition 5.3. *Suppose that* $\frac{l}{2l+1} < r \le r' < \frac{l+1}{2l+3}$ *and* $\delta = r - \frac{l}{2l+1}$. *If* X *and* Y *are finite subsets of* S^1, $X \subseteq Y$, *and* X *is a* $\frac{\delta}{2}$*-covering of* S^1, *then in the diagram*

$$\begin{tikzcd}
\mathbf{VR}_{\le}(X;r) \arrow[r, hook] & \mathbf{VR}_{\le}(Y;r') \\
\mathbf{VR}_{<}(X;r) \arrow[r, hook] \arrow[u, hook] & \mathbf{VR}_{<}(Y;r') \arrow[u, hook]
\end{tikzcd}$$

all spaces are homotopy equivalent to S^{2l+1} *and all maps are homotopy equivalences.*

For the spaces in the bottom row and the bottom map the same conclusion holds under the weaker assumption $\frac{l}{2l+1} < r \le r' \le \frac{l+1}{2l+3}$.

Proof. Proposition 5.2 gives $\mathrm{wf}_<(X;r) > r - \delta = \frac{l}{2l+1}$ and by Lemma 3.2(b) we have $\mathrm{wf}_\le(Y;r') \le r' < \frac{l+1}{2l+3}$. Hence the four cyclic graphs underlying the diagram have their winding fractions in the open interval $\left(\frac{l}{2l+1}, \frac{l+1}{2l+3}\right)$. The statement now follows from Proposition 4.9.

If $r' = \frac{l+1}{2l+3}$ then by Proposition 3.8(c) we still have $\mathrm{wf}_<(Y;r') < r' = \frac{l+1}{2l+3}$ and Proposition 4.9 applies in the bottom row. $\square$

We end this section with a partial converse of Proposition 5.2.

Proposition 5.4. *Suppose that* $\frac{l}{2l+1} < r$ *and* $\delta = r - \frac{l}{2l+1}$. *If* $X \subseteq S^1$ *is a finite subset with* $\mathrm{wf}_<(X;r) > \frac{l}{2l+1}$ *then* X *is a* $\left((l+\frac{1}{2})\delta\right)$*-covering of* S^1.

Proof. Suppose that $\overrightarrow{\mathrm{VR}}_<(X;r)$ dismantles to $\vec{C}_n^k$ with $\frac{k}{n} > \frac{l}{2l+1}$. Let $x_0 \prec \cdots \prec x_{n-1}$ be the points of X which induce the subgraph $\vec{C}_n^k \hookrightarrow \overrightarrow{\mathrm{VR}}_<(X;r)$. The proof will be complete if we show the following claim: for every i there exists a $j \neq i$ such that $\vec{d}(x_i, x_j) < (2l+1)\delta$. Without loss of generality it suffices to prove this for $i = 0$. We can assume that $(2l+1)\delta < 1$, for otherwise the claim is trivial.

Consider the directed path in $\overrightarrow{\mathrm{VR}}_<(X;r)$:

$$x_0 \to x_k \to x_{2k} \to \cdots \to x_{(2l+1)k}.$$

Since $(2l+1)k > nl$ this path makes at least l revolutions around the circle, hence

$$\sum_{i=0}^{2l} \vec{d}(x_{ik}, x_{(i+1)k}) > l.$$

On the other hand

$$\sum_{i=0}^{2l} \vec{d}(x_{ik}, x_{(i+1)k}) < (2l+1)r = l + (2l+1)\delta < l+1.$$

It follows that the directed path covers exactly l full circle lengths plus the arc $[x_0, x_{(2l+1)k}]_{S^1}$ whose length, by the last inequality, is less than $(2l+1)\delta$. That proves the claim. $\square$

The results of this section can be summarized as follows. Suppose that $\frac{l}{2l+1} < r < \frac{l+1}{2l+3}$ and $\delta = r - \frac{l}{2l+1}$. Then we know by Proposition 3.8(c) that for any finite subset $X \subseteq S^1$ we have $\mathrm{wf}_<(X; r) < r < \frac{l+1}{2l+3}$. If we think of X as an evolving (increasing) set, then the homotopy type of $\mathbf{VR}_<(X; r)$ stabilizes at S^{2l+1} at the same time when X becomes an ε-covering for some $\varepsilon \in \left[\frac{1}{2}\delta, \left(l + \frac{1}{2}\right)\delta\right]$. If l is constant this is a very tight window as $\delta \to 0$.

6. Evolution of random samples

We now apply the winding fraction to study the evolution of Vietoris–Rips complexes of random subsets of S^1. Let $\mathcal{X}_n \subseteq S^1$ be a subset obtained by sampling n points uniformly and independently from S^1. The connectivity of the graph $\mathrm{VR}(\mathcal{X}_n; r)$ when $r = r(n) \to 0$ as $n \to \infty$ has been extensively studied by many authors; see [Imany-Nabiyyi 2008] and the references therein. We obtain asymptotic thresholds for the higher-dimensional connectivity of $\mathbf{VR}(\mathcal{X}_n; r)$ when r is large. In particular, we analyze how many random samples are required until the homotopy type of $\mathbf{VR}(\mathcal{X}_n; r)$ matches that of $\mathbf{VR}(S^1; r)$, extending Latschev's approximation result [2001] for S^1 to r values that are no longer sufficiently small.

In this section we always assume that $l \geq 0$ is fixed and $\frac{l}{2l+1} < r < \frac{l+1}{2l+3}$. We define $\delta = r - \frac{l}{2l+1}$. The probability that two points of $\mathcal{X}_n$ are in distance exactly r for any fixed r is 0, and therefore all results hold for $\mathbf{VR}_<$ as well as $\mathbf{VR}_\leq$. Just as nontrivial asymptotic results about the connectedness of the graph $\mathrm{VR}(\mathcal{X}_n; r)$ can be obtained for $r \to 0$ as $n \to \infty$, in our higher-dimensional regime it makes sense to assume that $r \to \frac{l}{2l+1}$, that is $\delta \to 0$, as $n \to \infty$. We use the standard asymptotic notation $f(\delta) = \Theta(g(\delta))$ as $\delta \to 0$ when there are constants $C_1, C_2 > 0$ (which can depend on l) such that $C_1 g(\delta) \leq f(\delta) \leq C_2 g(\delta)$.

Let $M(r)$ and $N(r)$ be the random variables counting the number n of random points in S^1 until $\mathrm{wf}(\mathcal{X}_n; r)$ reaches, resp. exceeds, the value $\frac{l}{2l+1}$. Formally, consider the random process $(\mathcal{X}_1, \mathcal{X}_2, \dots)$ where $\mathcal{X}_{i+1}$ is obtained from $\mathcal{X}_i$ by adding a single uniformly random point. Define

$$M(r) = \min\left\{n : \mathrm{wf}(\mathcal{X}_n; r) \geq \tfrac{l}{2l+1}\right\} \quad \text{and} \quad N(r) = \min\left\{n : \mathrm{wf}(\mathcal{X}_n; r) > \tfrac{l}{2l+1}\right\},$$

where $\min \varnothing = \infty$. The random variables $M(r)$ and $N(r)$ describe the last two transition points in the evolution of $\mathbf{VR}(\mathcal{X}_n; r)$, since $M(r) \leq n < N(r)$ means $\mathbf{VR}(\mathcal{X}_n; r)$ is homotopy equivalent to a wedge of copies of S^{2l}, and $n \geq N(r)$ gives $\mathbf{VR}(\mathcal{X}_n; r) \simeq S^{2l+1}$. We will determine the asymptotic expectations $\mathbf{E}[M(r)]$ and $\mathbf{E}[N(r)]$.

Theorem 6.1. *Let $\frac{l}{2l+1} < r < \frac{l+1}{2l+3}$ for some fixed $l \geq 0$ and let $\delta = r - \frac{l}{2l+1}$. Then*

$$\mathbf{E}[M(r)] = \Theta\left(\left(\tfrac{1}{\delta}\right)^{\frac{2l}{2l+1}}\right) \quad \textit{and} \quad \mathbf{E}[N(r)] = \Theta\left(\tfrac{1}{\delta}\log\tfrac{1}{\delta}\right) \quad \textit{as } \delta \to 0.$$

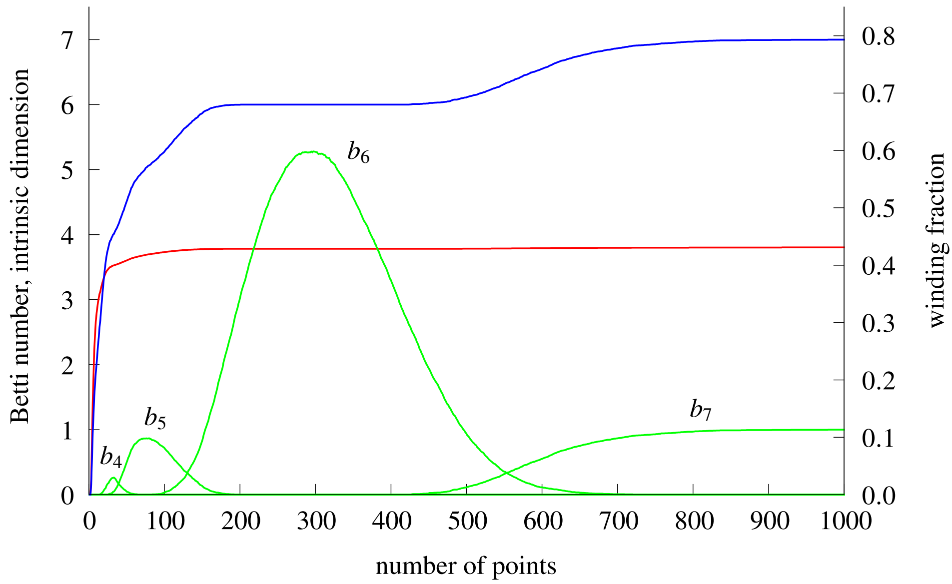


Figure 2. The evolution of $\mathbf{VR}(\mathcal{X}_n; r)$ with $r = 0.432$; see Example 6.2. The red curve is the average winding fraction, the blue curve is the average intrinsic dimension, and the green curves are the average Betti numbers b_4, b_5, b_6, b_7 (from left to right). Note the support of b_6 (for example) mostly coincides with the average intrinsic dimension being close to 6.

In particular, the expected number of random points n until $\mathbf{VR}(\mathcal{X}_n; r) \simeq S^{2l+1}$ *is* $\Theta\big(\frac{1}{\delta}\log\frac{1}{\delta}\big)$ *as* $\delta \to 0$.

Note that the winding fraction of $\frac{l}{2l+1}$ is achieved much sooner than it is exceeded (in fact $\mathbf{E}[M(r)]$ is sublinear in $1/\delta$). It means that we are expecting a long interval of n for which $\mathbf{VR}(\mathcal{X}_n; r)$ is a wedge of $2l$-spheres, before reaching the final homotopy type of S^{2l+1}.

Example 6.2. Suppose $\frac{3}{7} < r = 0.432 < \frac{4}{9}$ with $l = 3$, $\delta \approx 0.00343$, and $1/\delta \approx 291$. Figure 2 shows the average evolution of $\mathbf{VR}(\mathcal{X}_n; r)$ for $1 \le n \le 1000$. The red curve plots the average winding fraction, which rapidly approaches $\frac{3}{7}$ and then exceeds it around $n = 600$ to approach r. The homotopy type then stabilizes at S^7. For clarity of the presentation the blue curve depicts the average *intrinsic dimension*, which we define as $2l$ when $\mathrm{wf}(\,\cdot\,) = \frac{l}{2l+1}$ and as $2l+1$ when $\frac{l}{2l+1} < \mathrm{wf}(\,\cdot\,) < \frac{l+1}{2l+3}$.

We first prove the second claim of Theorem 6.1. For $\varepsilon > 0$ let $C(\varepsilon)$ be the random variable which counts the number of steps until $\mathcal{X}_n$ becomes a $\big(\frac{1}{2}\varepsilon\big)$-covering of S^1.

By Propositions 5.2 and 5.4 we have

$$C((2l+1)\delta) \le N(r) \le C(\delta). \tag{4}$$

It is well known that

$$\mathbf{E}[C(\varepsilon)] = \Theta(\varepsilon^{-1}\log\varepsilon^{-1}) \tag{5}$$

as $\varepsilon \to 0$; see [Solomon 1978, Equation (4.16)], which gives a more precise answer. The asymptotics of (5) can also be seen heuristically as follows. Divide S^1 into $K = \Theta(\varepsilon^{-1})$ arcs of length $\Theta(\varepsilon)$. Think of the random process $\mathcal{X}_n$ as throwing balls into K urns (arcs) independently. Then the event that $\mathcal{X}_n$ is a ε-covering coincides with the event that each urn contains a ball. By the classical coupon collector's problem this happens, in expectation, after $n = \Theta(K \log K)$ balls as $K \to \infty$. Combining (5) with (4) gives $\mathbf{E}[N(r)] = \Theta(\delta^{-1}\log\delta^{-1})$ as $\delta \to 0$.

To prove the first statement of Theorem 6.1 we need some auxiliary results. A subset $Y \subseteq S^1$ will be called (ε, m)-*regular* if $|Y| = m$ and there is a bijection from Y to the vertices of some regular inscribed m-gon which moves each point by distance less than ε. We previously showed that achieving $\mathrm{wf}(X; r) > \frac{l}{2l+1}$ coincides with X being a $\Theta(\delta)$-covering, and the next lemma shows that $\mathrm{wf}(X; r) \ge \frac{l}{2l+1}$ is achieved when X contains a $(\Theta(\delta), 2l+1)$-regular subset.

Lemma 6.3. *Let $\frac{l}{2l+1} < r < \frac{l+1}{2l+3}$ and $\delta = r - \frac{l}{2l+1}$. For a finite subset $X \subseteq S^1$:*

(a) *If X has a $\left(\frac{1}{2}\delta, 2l+1\right)$-regular subset then* $\mathrm{wf}(X; r) \ge \frac{l}{2l+1}$.

(b) *If* $\mathrm{wf}(X; r) \ge \frac{l}{2l+1}$ *then X has a $(4l\delta, 2l+1)$-regular subset.*

Proof. For (a) let $\{x_0, \ldots, x_{2l}\} \subseteq X$ be the $\left(\frac{1}{2}\delta, 2l+1\right)$-regular subset. We can assume $x_i \in \left(\frac{i}{2l+1} - \frac{1}{2}\delta, \frac{i}{2l+1} + \frac{1}{2}\delta\right)_{S^1}$. Since $\delta < \frac{1}{2l+1}$ we have $x_0 \prec x_1 \prec \cdots \prec x_{2l}$ cyclically ordered in S^1 as well as in $\overrightarrow{\mathrm{VR}}(X; r)$. We have

$$\vec{d}(x_i, x_{i+l}) < \frac{l}{2l+1} + 2 \cdot \frac{1}{2}\delta = r$$

and hence a cyclic homomorphism $\vec{C}^l_{2l+1} \hookrightarrow \overrightarrow{\mathrm{VR}}(X; r)$.

To prove (b) suppose $\mathrm{wf}(X; r) \ge \frac{l}{2l+1}$. By Proposition 3.15 there is a directed homomorphism $f : \vec{C}^l_{2l+1} \to \overrightarrow{\mathrm{VR}}(X; r)$. Denote $x_i = f(i)$. For every $i = 0, \ldots, 2l$,

$$\vec{d}(x_i, x_{i+1}) = 1 - \vec{d}(x_{i+1}, x_{i+l+1}) - \vec{d}(x_{i+l+1}, x_i) > 1 - 2r = \frac{1}{2l+1} - 2\delta.$$

It follows that for $j = 1, \ldots, 2l$,

$$\vec{d}(x_i, x_{i+j}) > \frac{j}{2l+1} - 2j\delta \ge \frac{j}{2l+1} - 4l\delta$$

and in turn

$$\vec{d}(x_i, x_{i+j}) = 1 - \vec{d}(x_{i+j}, x_i) < 1 - \frac{2l+1-j}{2l+1} + 4l\delta = \frac{j}{2l+1} + 4l\delta.$$

It follows that each x_j lies in distance less than $4l\delta$ from the j-th vertex of the regular $(2l+1)$-gon with x_0 as a vertex. □

Let $R_m(\varepsilon)$ be the random variable which counts the number of steps until $\mathcal{X}_n$ contains a (ε, m)-regular subset. In Section B we show that for every fixed m

$$\mathbf{E}[R_m(\varepsilon)] = \Theta\big(\varepsilon^{-\frac{m-1}{m}}\big) \qquad \text{as } \varepsilon \to 0. \tag{6}$$

Here we only give a heuristic explanation using our previous urn model with $K = \Theta(\varepsilon^{-1})$ urns identified with arcs of length $\Theta(\varepsilon)$. Divide the urns into K/m groups of size m, each group consisting of arcs centered approximately around the vertices of a regular m-gon. Then the event that $\mathcal{X}_n$ has an (ε, m)-regular subset coincides with the event that every urn in some group contains a ball. This can be correlated with the generalized birthday paradox, where we require one urn to contain m balls (the case $m = 2$ is the classical birthday paradox). The expected waiting time for this to happen is $\Theta\big(K^{\frac{m-1}{m}}\big)$ as $K \to \infty$; see [Klamkin and Newman 1967, Theorem 2].

This proves the first claim of Theorem 6.1 since by Lemma 6.3 we have

$$R_{2l+1}(4l\delta) \le N(r) \le R_{2l+1}\big(\tfrac{1}{2}\delta\big).$$

The proof of Theorem 6.1 is now complete.

7. Vietoris–Rips complexes for subsets of S^1

The definition of the Vietoris–Rips complex $\mathbf{VR}(X; r)$ makes sense for an arbitrary metric space X, not necessarily finite nor discrete. Hausmann [1995] studied the case when X is a closed Riemannian manifold. In this section we show that for an arbitrary subset $X \subseteq S^1$ and $r > 0$, the complex $\mathbf{VR}(X; r)$ has the homotopy type of an odd-dimensional sphere or a wedge of even-dimensional spheres. We will also study the complexes $\mathbf{VR}_<(S^1; r)$ and $\mathbf{VR}_\le(S^1; r)$ in more detail.

For an arbitrary metric space X the geometric realization of $\mathbf{VR}(X; r)$ is given the topology of a CW-complex, that is the weak topology with respect to finite-dimensional skeleta, or equivalently, the weak topology with respect to subcomplexes induced by finite subsets of X. Formally, let $F(X)$ be the poset of all finite subsets of X ordered by inclusion. Then for each r we have a functor $\mathbf{VR}(-; r) : F(X) \to \mathrm{Top}$ and

$$\mathbf{VR}(X; r) = \mathrm{colim}_{Y \in F(X)} \mathbf{VR}(Y; r) \simeq \mathrm{hocolim}_{Y \in F(X)} \mathbf{VR}(Y; r),$$

where the last equivalence is a consequence of the fact that all maps $\mathbf{VR}(Y; r) \hookrightarrow \mathbf{VR}(Y'; r)$ for $Y \subseteq Y'$ are inclusions of closed subcomplexes, hence cofibrations. See [Welker et al. 1999, Section 3] for the statements of all diagram comparison theorems used in this section.

For a finite subset $Y_0 \subseteq X$ let $F(X; Y_0)$ be the subposet of $F(X)$ consisting of all sets which contain Y_0. Since this poset is cofinal in $F(X)$, we also have

$$\mathbf{VR}(X; r) = \operatorname{colim}_{Y \in F(X;Y_0)} \mathbf{VR}(Y; r) \simeq \operatorname{hocolim}_{Y \in F(X;Y_0)} \mathbf{VR}(Y; r).$$

For an arbitrary subset $\varnothing \neq X \subseteq S^1$ and $0 < r < \frac{1}{2}$ we define

$$\begin{aligned} \mathrm{wf}_<(X; r) &= \sup\{\mathrm{wf}_<(Y; r) : Y \subseteq X, |Y| < \infty\}, \\ \mathrm{wf}_\leq(X; r) &= \sup\{\mathrm{wf}_\leq(Y; r) : Y \subseteq X, |Y| < \infty\}. \end{aligned} \tag{7}$$

The supremum in (7) need not be attained when X is infinite. When X is finite then $\mathrm{wf}(X; r)$ agrees with our previous definition of this symbol since the supremum is attained by $Y = X$. The following proposition shows that in the generic case, $\mathbf{VR}(X; r)$ has the homotopy type of an odd-dimensional sphere.

Proposition 7.1. *Suppose that $\varnothing \neq X \subseteq S^1$ and $0 < r < \frac{1}{2}$. Either of the conditions*

(1) $\frac{l}{2l+1} < \mathrm{wf}(X; r) < \frac{l+1}{2l+3}$, *or*
(2) $\mathrm{wf}(X; r) = \frac{l+1}{2l+3}$ *and the supremum is not attained*,

for some $l = 0, 1, \ldots$, implies that $\mathbf{VR}(X; r) \simeq S^{2l+1}$.

Moreover, if $r' \geq r$ is another value of the distance parameter for which (1) *or* (2) *hold with the same l, then the inclusion $\mathbf{VR}(X; r) \hookrightarrow \mathbf{VR}(X; r')$ is a homotopy equivalence.*

Proof. Either of the two conditions (1) or (2) implies there is a finite subset $Y_0 \subseteq X$ such that for every finite subset Y with $Y_0 \subseteq Y \subseteq X$, we have $\frac{l}{2l+1} < \mathrm{wf}(Y; r) < \frac{l+1}{2l+3}$. By Proposition 4.9 all maps in the diagram $\mathbf{VR}(-; r) : F(X; Y_0) \to \mathrm{Top}$ are homotopy equivalences between spaces homotopy equivalent to S^{2l+1}, and therefore

$$\mathbf{VR}(X; r) \simeq \operatorname{hocolim}_{Y \in F(X, Y_0)} \mathbf{VR}(Y; r) \simeq S^{2l+1}.$$

Furthermore, $\frac{l}{2l+1} < \mathrm{wf}(Y; r) \leq \mathrm{wf}(Y; r') < \frac{l+1}{2l+3}$, hence the same is true for $\mathbf{VR}(X; r')$. The maps $\mathbf{VR}(Y; r) \to \mathbf{VR}(Y; r')$ now define a natural transformation of diagrams $\mathbf{VR}(-; r) \to \mathbf{VR}(-; r')$ which is a levelwise homotopy equivalence by Proposition 4.9. It follows that the induced map of (homotopy) colimits is a homotopy equivalence. $\square$

Remark 7.2. The same argument shows that under the assumptions of the last proposition the map $\mathbf{VR}(Y_0; r) \hookrightarrow \mathbf{VR}(X; r)$ is a homotopy equivalence whenever $Y_0 \subseteq X$ is a finite set with $\mathrm{wf}(Y_0; r) > \frac{l}{2l+1}$.

As the next lemma shows, the winding fractions behave in the expected way for dense subsets of the circle.

Lemma 7.3. *If X is dense in S^1 and $0 < r < \frac{1}{2}$ then* $\mathrm{wf}_<(X; r) = \mathrm{wf}_\leq(X; r) = r$. *In the case of* $\mathrm{wf}_<$ *the supremum is not attained.*

Proof. For every $\varepsilon > 0$ the set X contains a finite ε-covering of S^1. Proposition 5.2 now gives $\mathrm{wf}(X; r) \geq r$. The reverse inequality and the second statement of the lemma follow from Proposition 3.8(c). $\square$

We can now give a complete description of the homotopy types of $\mathbf{VR}_<(S^1; r)$ for arbitrary r.

Theorem 7.4. *If X is dense in S^1 (in particular when $X = S^1$) and $0 < r < \frac{1}{2}$, then*

$$\mathbf{VR}_<(X; r) \simeq S^{2l+1} \quad \textit{for } \tfrac{l}{2l+1} < r \leq \tfrac{l+1}{2l+3},\ l = 0, 1, \ldots.$$

Moreover, if $\frac{l}{2l+1} < r \leq r' \leq \frac{l+1}{2l+3}$ then the inclusion $\mathbf{VR}_<(X; r) \hookrightarrow \mathbf{VR}_<(X; r')$ is a homotopy equivalence.

Proof. By Lemma 7.3 we have $\mathrm{wf}_<(X; r) = r$ and the supremum is not attained, meaning that either (1) or (2) in Proposition 7.1 is satisfied. $\square$

Proposition 7.1 describes the homotopy types of the complex $\mathbf{VR}(X; r)$ in all generic situations. The only singular cases it does not cover occur when $\mathrm{wf}(X; r)$ is of the form $\frac{l}{2l+1}$ and this value is in fact attained by some finite subset $Y_0 \subseteq X$. We deal with this in the next two statements.

Proposition 7.5. *Suppose that $\varnothing \neq X \subseteq S^1$ and $0 < r < \frac{1}{2}$. If $\mathrm{wf}(X; r) = \frac{l}{2l+1}$ for some $l = 0, 1, \ldots$ and the supremum in the definition of $\mathrm{wf}(X; r)$ is attained, then $\mathbf{VR}(X; r)$ is homotopy equivalent to a wedge sum of spheres of dimension $2l$.*

Theorem 7.6. *For $0 \leq r < \frac{1}{2}$ we have a homotopy equivalence*

$$\mathbf{VR}_\leq(S^1; r) \simeq \begin{cases} S^{2l+1} & \textit{if } \frac{l}{2l+1} < r < \frac{l+1}{2l+3},\ l = 0, 1, \ldots, \\ \bigvee^{\mathfrak{c}} S^{2l} & \textit{if } r = \frac{l}{2l+1}, \end{cases}$$

where $\mathfrak{c}$ is the cardinality of the continuum. Moreover, if $\frac{l}{2l+1} < r \leq r' < \frac{l+1}{2l+3}$ then the inclusion $\mathbf{VR}_\leq(S^1; r) \hookrightarrow \mathbf{VR}_\leq(S^1; r')$ is a homotopy equivalence.

We delay the proofs of Proposition 7.5 and Theorem 7.6 until Section 8. Note that Theorems 7.4 and 7.6 together provide a complete description of the homotopy types of $\mathbf{VR}(S^1; r)$ for arbitrary r. They also give the persistent homology of $\mathbf{VR}(S^1; r)$, where we refer the reader to [Chazal et al. 2014] for information on the persistent homology of Vietoris–Rips complexes.

Corollary 7.7. *The persistent homology of $\mathbf{VR}_<(S^1; r)$ contains a single interval $\left(\frac{l}{2l+1}, \frac{l+1}{2l+3}\right]$ in each homological dimension $2l + 1$, and the persistent homology of $\mathbf{VR}_\leq(S^1; r)$ contains a single interval $\left(\frac{l}{2l+1}, \frac{l+1}{2l+3}\right)$ in each homological dimension* $2l + 1$.

Remark 7.8. Hausmann [1995, (3.12)] conjectured that if M is a compact Riemannian manifold then the connectivity $\mathrm{conn}(\mathbf{VR}_<(M; r))$ is a nondecreasing function of r, and our results confirm this conjecture for $M = S^1$. Hausmann [op. cit., (3.11)] furthermore conjectured that for r sufficiently small, $\mathbf{VR}_<(M;r)$ is homotopy equivalent to $\mathbf{VR}_<(Y_0;r)$ for some finite subset $Y_0 \subseteq M$. For $M = S^1$ we confirm this conjecture for all r, sufficiently small or otherwise.

8. Singular winding fractions

In this section we return to study cyclic graphs for which $\mathrm{wf}(\vec{G}) = \frac{l}{2l+1}$ is a singular value. Our aim is to describe a convenient structure in the homology group $\widetilde{H}_{2l}(\mathrm{Cl}(G))$, which we then use to prove Proposition 7.5 and Theorem 7.6.

We consider first a cyclic graph $\vec{C}_n^k$ with $\frac{k}{n} = \frac{l}{2l+1}$. Since l and $2l+1$ are coprime we have $(k,n) = (dl, d(2l+1))$ for some integer $d \geq 1$. We have $d = n - 2k$ and so by Theorem 4.4 we can write

$$\mathrm{Cl}(C_{d(2l+1)}^{dl}) \simeq \bigvee^{d-1} S^{2l}.$$

When $(k,n) = (l, 2l+1)$ the graph C_{2l+1}^l is a clique and $\mathrm{Cl}(C_{2l+1}^l)$ is the full simplex with $2l+1$ vertices.

The next case, $d = 2$ and $(k,n) = (2l, 2(2l+1))$, is particularly interesting for our purposes. The nonedges of the graph $C_{2(2l+1)}^{2l}$ are pairs of the form $\{i, i+2l+1\}$, which are the antipodal pairs in the evenly-spaced model

$$C_{2(2l+1)}^{2l} = \mathrm{VR}_{\leq}\big(\big\{\tfrac{i}{2(2l+1)} : i = 0, \ldots, 4l+1\big\}; \tfrac{l}{2l+1}\big).$$

It follows that the clique complex $\mathrm{Cl}(C_{2(2l+1)}^{2l})$ is isomorphic to the standard triangulation of S^{2l} as the boundary of the cross-polytope of dimension $2l+1$. We fix the $2l$-dimensional cycle in $\mathrm{Cl}(C_{2(2l+1)}^{2l})$:

$$\begin{aligned} \iota_{2l} &= (-1)^{l(l+3)/2} \cdot ([0] - [2l+1]) \wedge ([1] - [2l+2]) \wedge \cdots \wedge ([2l] - [4l+1]) \\ &= [0, 2, \ldots, 4l] - [1, 3, \ldots, 4l+1] \pm \cdots, \end{aligned} \tag{8}$$

which is (up to sign) the fundamental cycle of the boundary of the cross-polytope. Here $[x_0] \wedge \cdots \wedge [x_k]$ denotes the oriented simplex $[x_0, \ldots, x_k]$, and we have chosen the sign so that the oriented simplices $[0, 2, \ldots, 4l]$ and $[1, 3, \ldots, 4l+1]$ appear with coefficients $+1$ and -1 respectively. Indeed, in the oriented cycle $([0] - [2l+1]) \wedge \cdots \wedge ([2l] - [4l+1])$ the sign on $[0, 2l+2, 2, \ldots, 2l-2, 4l, 2l]$ is $(-1)^l$, and then after $\frac{l(l+1)}{2}$ transpositions this gives the sign $(-1)^{l(l+3)/2}$ on $[0, 2, \ldots, 4l]$. The argument for $[1, 3, \ldots, 4l+1]$ is similar. The corresponding homology class $\iota_{2l} \in \widetilde{H}_{2l}(\mathrm{Cl}(C_{2(2l+1)}^{2l})) = \mathbb{Z}$ is a generator. (Here and in the following we will use the same symbol to denote a (co)cycle and its (co)homology class, and sometimes also the map which induces the given class.)

Definition 8.1. Suppose that $\vec{G}$ is a cyclic graph. A nonzero homology class $\alpha \in \widetilde{H}_{2l}(\mathrm{Cl}(G))$ is called *cross-polytopal* if there is a cyclic homomorphism $f: \vec{C}^{2l}_{2(2l+1)} \to \vec{G}$ such that $\alpha = f_*(\iota_{2l})$.

An immediate consequence of the definition is that the image of a cross-polytopal class under a cyclic homomorphism $\vec{G} \to \vec{H}$ is again cross-polytopal, unless it is zero. Note that if f is not injective on the vertices then $f_*(\iota_{2l}) = 0$ because a homology class of degree $2l$ in a clique complex must be supported on at least $4l+2$ vertices; see for instance [Kahle 2009, Lemma 5.3].

Our aim is to classify all cross-polytopal homology classes for cyclic graphs. We begin with the description of a class of cyclic homomorphisms.

Lemma 8.2. *Let $d \geq 1$ and $(k, n) = (dl, d(2l+1))$.*

(a) *Every cyclic homomorphism $\vec{C}^{l}_{2l+1} \to \vec{C}^{k}_{n}$ is of the form θ_a for some $a = 0, \dots, n-1$, where*

$$\theta_a(i) \equiv a + di \mod n.$$

(b) *Every injective cyclic homomorphism $\vec{C}^{2l}_{2(2l+1)} \to \vec{C}^{k}_{n}$ is of the form $\alpha_{a,b}$ for some $a = 0, \dots, n-1$ and $b = a+1, \dots, a+d-1$, where*

$$\alpha_{a,b}(i) = \begin{cases} a + d \cdot \frac{i}{2} \mod n & \text{if } i \text{ is even}, \\ b + d \cdot \frac{i-1}{2} \mod n & \text{if } i \text{ is odd}. \end{cases}$$

Remark 8.3. Every cyclic homomorphism θ in part (a) is determined by the choice of $a = \theta(0)$ and the condition $\theta(i+1) = \theta(i) + d \mod n$. Similarly, in part (b) every cyclic homomorphism is determined by the two initial values $a = \alpha(0)$ and $b = \alpha(1)$, together with the requirement that $\alpha(i+2) \equiv \alpha(i) + d \mod n$.

Proof. To prove (a) let $\theta : \vec{C}^{l}_{2l+1} \to \vec{C}^{k}_{n}$ be a cyclic homomorphism with $l > 0$, since the case $l = 0$ is clear. Then

$$(2l+1)k \geq \sum_{i=0}^{2l} \vec{d}_n(\theta(i), \theta(i+l)) = l \cdot \sum_{i=0}^{2l} \vec{d}_n(\theta(i), \theta(i+1)) = ln,$$

where the last equality follows from (2). Since the two extremes are in fact equal, we must have $\vec{d}_n(\theta(i), \theta(i+l)) = k$ for all i, which implies

$$\vec{d}_n(\theta(i), \theta(i+1)) = n - \vec{d}_n(\theta(i+1), \theta(i+l+1)) - \vec{d}_n(\theta(i+l+1), \theta(i)) = n - 2k = d,$$

as required. Clearly every θ_a is a cyclic homomorphism, hence (a) is proved.

Part (b) follows immediately, since $\vec{C}^{2l}_{2(2l+1)}$ contains two induced copies of $\vec{C}^{l}_{2l+1}$ with vertex sets $\{0, 2, \dots, 4l\}$ and $\{1, 3, \dots, 4l+1\}$. For an injective cyclic homomorphism α we must have $\alpha(0) \prec \alpha(1) \prec \alpha(2)$ cyclically ordered in $\vec{C}^{k}_{n}$, i.e., $a \prec b \prec a + d$, which yields the restrictions on b. $\square$

The cyclic homomorphism $\alpha_{a,b}$ evaluated on the fundamental cycle ι_{2l} determines a cycle as well as a homology class in $\widetilde{H}_{2l}(\mathrm{Cl}(C^{dl}_{d(2l+1)}))$. We will continue to denote both with $\alpha_{a,b}$. The chain representation of the cycle $\alpha_{a,b}$ starts with

$$\alpha_{a,b} = [a, a+d, \ldots, a+2l\cdot d] - [b, b+d, \ldots, b+2l\cdot d] \pm \cdots \tag{9}$$

Compare this to (8). The homology classes $\alpha_{a,b}$ for various pairs (a, b) satisfy a number of relations worked out in the proof of the next proposition, which is the main result concerning cyclic graphs with $\mathrm{wf}(\vec{G}) = \frac{l}{2l+1}$.

Proposition 8.4. *Suppose $\vec{G}$ is a cyclic graph which dismantles to $\vec{C}^{dl}_{d(2l+1)}$. Then the homology group $\widetilde{H}_{2l}(\mathrm{Cl}(G)) = \mathbb{Z}^{d-1}$ has a basis $\{e_1, \ldots, e_{d-1}\}$ such that all the cross-polytopal elements in $\widetilde{H}_{2l}(\mathrm{Cl}(G))$ are*

$$\pm e_1, \ldots, \pm e_{d-1}$$

and

$$e_i - e_j, \quad 1 \le i, j \le d-1,\ i \neq j.$$

In particular, there are exactly $d(d-1)$ cross-polytopal elements in $\widetilde{H}_{2l}(\mathrm{Cl}(G))$.

Proof. In the first step we will prove the result for $\vec{G} = \vec{C}^{dl}_{d(2l+1)}$. Denote $(k, n) = (dl, d(2l+1))$.

For an oriented simplex σ in a simplicial complex K let $\sigma^\vee$ denote the cochain which assigns 1 to σ, -1 to σ with opposite orientation, and 0 to all other oriented simplices of K. For every $a = 0, \ldots, n-1$ define a cochain β_a in $\mathrm{Cl}(C^k_n)$ by

$$\beta_a = [a, a+d, \ldots, a+2l\cdot d]^\vee.$$

Since the face $[a, a+d, \ldots, a+2l\cdot d]$ is maximal in $\mathrm{Cl}(C^k_n)$, the cochain β_a is in fact a cocycle, and it determines a cohomology class which we denote with the same symbol. Using (9) we verify that for $1 \le i, j \le d-1$,

$$\beta_i(\alpha_{j,d}) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Since the groups $\widetilde{H}_{2l}(\mathrm{Cl}(C^k_n))$ and $\widetilde{H}^{2l}(\mathrm{Cl}(C^k_n))$ are both free abelian of rank $d-1$, the above implies that $\{\alpha_{1,d}, \ldots, \alpha_{d-1,d}\}$ is a basis of homology and $\{\beta_1, \ldots, \beta_{d-1}\}$ is its dual basis of cohomology. In particular, every element $v \in \widetilde{H}_{2l}(\mathrm{Cl}(C^k_n))$ has a decomposition

$$v = \sum_{i=1}^{d-1} \beta_i(v)\cdot \alpha_{i,d}. \tag{10}$$

Note that $\beta_i(\alpha_{a,b})$ depends only on the evaluation of β_i on the two leading terms in (9), since β_i evaluates to 0 on all the omitted terms. The oriented simplices appearing in $\alpha_{a,b}$ and $\alpha_{a+d,b+d}$ differ by a cyclic shift, hence by an even number of $2l$ transpositions, and are therefore equal. That means we have the identity

$$\alpha_{a+d,b+d} = \alpha_{a,b}.$$

It follows that all cross-polytopal classes can be written as $\alpha_{a,b}$ with $0 \le a \le d-1$ and $a+1 \le b \le a+d-1$.

If $a = 0$ then $1 \le b \le d-1$ and the only nonzero pairing in (10) is $\beta_b(\alpha_{0,b}) = -1$, and hence $\alpha_{0,b} = -\alpha_{b,d}$.

If $1 \le a < b \le d-1$ then $\beta_a(\alpha_{a,b}) = 1$ and $\beta_b(\alpha_{a,b}) = -1$; hence $\alpha_{a,b} = \alpha_{a,d} - \alpha_{b,d}$.

If $1 \le a \le d-1$ and $b = d$ then $\alpha_{a,b} = \alpha_{a,d}$ is itself one of the generators.

If $1 \le a \le d-1$ and $d+1 \le b \le a+d-1$ then $1 \le b-d < a \le d-1$. Using the cyclic shift argument we obtain $\beta_a(\alpha_{a,b}) = 1$ and $\beta_{b-d}(\alpha_{a,b}) = -1$, hence $\alpha_{a,b} = \alpha_{a,d} - \alpha_{b-d,d}$.

It follows that the proposition is true with $e_i = \alpha_{i,d}$ for $i = 1, \dots, d-1$.

Now suppose $\vec{G}$ is an arbitrary cyclic graph which dismantles to $\vec{C}_n^k$. By Corollary 4.2 the cyclic homomorphisms $\vec{C}_n^k \xrightarrow{\iota} \vec{G} \xrightarrow{\pi} \vec{C}_n^k$ induce isomorphisms

$$\widetilde{H}_{2l}(\mathrm{Cl}(C_n^k)) \xrightarrow{\cong} \widetilde{H}_{2l}(\mathrm{Cl}(G)) \xrightarrow{\cong} \widetilde{H}_{2l}(\mathrm{Cl}(C_n^k))$$

with the composition being the identity. It follows that the cross-polytopal classes $\iota_*(e_1), \dots, \iota_*(e_{d-1})$ form a basis of $\widetilde{H}_{2l}(\mathrm{Cl}(G))$ and that $\pm\iota_*(e_i)$ and $\iota_*(e_i) - \iota_*(e_j)$, $i \neq j$, are cross-polytopal. Moreover, if $\alpha \in \widetilde{H}_{2l}(\mathrm{Cl}(G))$ is cross-polytopal then $\pi_*(\alpha)$ is one of $\pm e_i$ or $e_i - e_j$, $i \neq j$, and therefore α must be one of $\pm\iota_*(e_i)$ or $\iota_*(e_i) - \iota_*(e_j)$, $i \neq j$. That completes the proof. $\square$

We are now prepared to prove Proposition 7.5, using the algebraic fact in Proposition A.1 of the appendix.

Proof of Proposition 7.5. Let $Y_0 \subseteq X$ be a finite subset that achieves $\mathrm{wf}(Y_0; r) = \frac{l}{2l+1}$. Then we have $\mathrm{wf}(Y; r) = \frac{l}{2l+1}$ for any finite subset Y with $Y_0 \subseteq Y \subseteq X$. By Corollary 4.5 every space in the diagram

$$\mathbf{VR}(X; r) = \mathrm{colim}_{Y \in F(X; Y_0)} \mathbf{VR}(Y; r)$$

is homotopy equivalent to a finite wedge sum of $2l$-spheres. It follows immediately that $\mathbf{VR}(X; r)$ is simply-connected and its homology is torsion-free and concentrated in degree $2l$. It remains to show that the group $\widetilde{H}_{2l}(\mathbf{VR}(X; r))$ is free abelian. Indeed, if this is the case then $\mathbf{VR}(X; r)$ is a model of the Moore space $M(\bigoplus^{\kappa} \mathbb{Z}, 2l)$, unique up to homotopy and equivalent to $\bigvee^{\kappa} S^{2l}$, for some cardinal number κ.

A nonzero homology class in $\widetilde{H}_{2l}(\mathbf{VR}(X; r))$ will be called *cross-polytopal* if it is the image under the inclusion $\mathbf{VR}(Y; r) \hookrightarrow \mathbf{VR}(X; r)$ of a cross-polytopal class in $\widetilde{H}_{2l}(\mathbf{VR}(Y; r))$ for some finite $Y \in F(X; Y_0)$. Since the groups $\widetilde{H}_{2l}(\mathbf{VR}(Y; r))$

are generated by cross-polytopal classes, the same is true about their colimit, $\widetilde{H}_{2l}(\mathbf{VR}(X;r))$.

A subset B of an abelian group G is called *independent* if for every finite subset $\{b_1,\dots,b_s\}\subseteq B$ the identity $\sum_{i=1}^{s} a_i b_i = 0$, with $a_1,\dots,a_s\in\mathbb{Z}$, implies $a_1=\cdots=a_s=0$. An independent set B generates a free abelian subgroup of G with basis B. Now let $\mathcal{B}$ be the family of all subsets $B\subseteq\widetilde{H}_{2l}(\mathbf{VR}(X;r))$ such that

(a) all elements of B are cross-polytopal,

(b) B is independent.

The family $\mathcal{B}$ is nonempty and closed under increasing unions. Using Zorn's lemma pick an inclusionwise maximal set B satisfying (a) and (b). If B generates $\widetilde{H}_{2l}(\mathbf{VR}(X;r))$ then we are done, since the group $\langle B\rangle$ generated by B is free abelian.

We suppose for a contradiction that B does not generate $\widetilde{H}_{2l}(\mathbf{VR}(X;r))$, and hence there exists a cross-polytopal class $v\notin\langle B\rangle$ since $\widetilde{H}_{2l}(\mathbf{VR}(X;r))$ is generated by cross-polytopal classes. By maximality of B the set $B\cup\{v\}$ violates (b), and hence there exists a nontrivial linear relation involving v and a finite number of elements $b_1,\dots,b_s\in B$. In other words, some nontrivial multiple of v lies in the subgroup of $\widetilde{H}_{2l}(\mathbf{VR}(X;r))$ generated by $b_1,\dots,b_s$. The same relation holds for the cross-polytopal representatives $v, b_1,\dots,b_s$ at some finite stage $\widetilde{H}_{2l}(\mathbf{VR}(Y;r))$ of the colimit, where $\mathbf{VR}(Y;r)$ dismantles to $\vec{C}^{dl}_{d(2l+1)}$. Changing signs if necessary we may assume that each of the elements $v, b_1,\dots,b_s\in\widetilde{H}_{2l}(\mathbf{VR}(Y;r))=\mathbb{Z}^{d-1}$ is of the form e_i or e_i-e_j, $i<j$, for the basis $\{e_1,\dots,e_{d-1}\}$ from Proposition 8.4. Now Proposition A.1 implies that v itself lies in the subgroup of $\widetilde{H}_{2l}(\mathbf{VR}(Y;r))$ generated by $b_1,\dots,b_s$, and hence in the subgroup of $\widetilde{H}_{2l}(\mathbf{VR}(X;r))$ generated by $b_1,\dots,b_s$. This contradiction shows that in fact $\widetilde{H}_{2l}(\mathbf{VR}(X;r))=\langle B\rangle$ is free abelian. $\square$

The last item in this section is the proof of Theorem 7.6.

Proof of Theorem 7.6. By Lemma 7.3 we have $\mathrm{wf}_{\le}(S^1;r)=r$, and so all statements concerning the generic values of r and r' follow from part (1) of Proposition 7.1.

If $r=\frac{l}{2l+1}$ then the value of $\mathrm{wf}_{\le}\big(S^1;\frac{l}{2l+1}\big)=\frac{l}{2l+1}$ is attained by the vertex set of any regular $(2l+1)$-gon. Proposition 7.5 implies that $\mathbf{VR}_{\le}\big(S^1;\frac{l}{2l+1}\big)$ is homotopy equivalent to a wedge of copies of S^{2l}, and so it remains to count the number of wedge summands. For $t\in\big(0,\frac{1}{2l+1}\big)_{S^1}$ let

$$Y_t=\Big\{\tfrac{i}{2l+1},\, t+\tfrac{i}{2l+1} : i=0,\dots,2l\Big\}.$$

We have an isomorphism $\vec{\mathrm{VR}}_{\le}\big(Y_t;\frac{l}{2l+1}\big)=\vec{C}^{2l}_{2(2l+1)}$, hence each inclusion

$$j_t:\mathbf{VR}_{\le}\big(Y_t;\tfrac{l}{2l+1}\big)\hookrightarrow\mathbf{VR}_{\le}\big(S^1;\tfrac{l}{2l+1}\big)$$

determines a homology class $\alpha_t=j_{t*}(\iota_{2l})$ in the complex $\mathbf{VR}_{\le}\big(S^1;\frac{l}{2l+1}\big)$. Each simplex $\beta_t=\big[t+\frac{i}{2l+1} : i=0,\dots,2l\big]$ is a maximal face of $\mathbf{VR}_{\le}\big(S^1;\frac{l}{2l+1}\big)$, which appears in the support of α_t but not in any other α_s for $s \neq t$. This implies that the classes α_t are independent, and hence $\widetilde{H}_{2l}\big(\mathbf{VR}_{\leq}\big(S^1; \frac{l}{2l+1}\big)\big)$ contains a free abelian group of rank $\mathfrak{c}$. We get a corresponding upper bound by noting that the cardinality of the set of $2l$-simplices in $\mathbf{VR}_{\leq}\big(S^1; \frac{l}{2l+1}\big)$ is also $\mathfrak{c}$, and hence the cardinality of the wedge sum is $\mathfrak{c}$. □

Remark 8.5. Chambers et al. [2010, Section 6(1)] asked if for all $k \geq 2$ and any finite subset $X \subseteq \mathbb{R}^2$ the homology group $\widetilde{H}_k(\mathbf{VR}(X; r))$ is generated by induced k-dimensional cross–polytopal spheres (for all k these are complexes of the form $\mathrm{Cl}(C^k_{2k+2})$, where we considered $k = 2l$ in this section). Proposition 8.4 confirms this when $k = 2l$ for subsets $X \subseteq S^1 \subseteq \mathbb{R}^2$. When $k = 2l + 1$ is odd the claim fails already for $X \subseteq S^1$. For example, one can check that for $\frac{1}{3} < \frac{k}{n} < \frac{3}{8}$ the graph C^k_n does not contain an induced subgraph isomorphic to C^3_8, yet $\widetilde{H}_3(\mathrm{Cl}(C^k_n)) = \widetilde{H}_3(S^3) \neq 0$ by Theorem 4.3.

9. Čech complexes

The Čech complex is another simplicial complex commonly associated with a metric space. For a point x in a metric space M, let $B_<(x; r)$ and $B_\leq(x; r)$ denote the open and closed balls in M with center x and radius r.

Definition 9.1. For a subset $X \subseteq M$ of an ambient metric space M and $r > 0$, the *Čech complex* $\check{\mathbf{C}}_<(X, M; r)$ is the simplicial complex with vertex set X, where a finite subset $\sigma \subseteq X$ is a face if and only if $\bigcap_{x\in\sigma} B_<(x; r) \neq \varnothing$. Analogously, the faces of the complex $\check{\mathbf{C}}_\leq(X, M; r)$ satisfy $\bigcap_{x\in\sigma} B_\leq(x; r) \neq \varnothing$.

As before, we will omit the subscript in statements which apply to both $<$ and $\leq$. An equivalent definition of $\check{\mathbf{C}}(X, M; r)$ is as the nerve of the family of balls $\{B(x; r) : x \in X\}$. Chazal, de Silva, and Oudot [Chazal et al. 2014] refer to these complexes as *ambient Čech complexes* with landmark set X and witness set M. We have the inclusion $\check{\mathbf{C}}(X, M; r/2) \subseteq \mathbf{VR}(X; r)$, and if M is a geodesic space then $\mathrm{VR}(X; r)$ is the 1-skeleton not only of $\mathbf{VR}(X; r)$ but also of $\check{\mathbf{C}}(X, M; r/2)$.

Notation 9.2. If $X \subseteq S^1$ then we write $\check{\mathbf{C}}(X; r)$ for $\check{\mathbf{C}}(X, S^1; r)$.

If $M = S^1$ then the balls are open or closed arcs, and one can see that finite $\sigma \subseteq X$ is a face of $\check{\mathbf{C}}(X; r)$ if and only if σ is contained in some arc of length $2r$.

One can develop a parallel theory of dismantling, winding fractions, and homotopy types for the complexes $\check{\mathbf{C}}(X; r)$ with $X \subseteq S^1$, leading to straightforward analogues of all the results from this paper. We note that the sequence of critical values $\big(0, \frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \dots, \frac{l}{2l+1}, \dots\big)$ determining the transitions of homotopy types will be replaced in the case of Čech complexes with the sequence $\big(0, \frac{1}{4}, \frac{2}{6}, \frac{3}{8}, \dots, \frac{l}{2(l+1)}, \dots\big)$, and we refer the reader to [Adamaszek et al. 2016] for some results in the case of finite X. Instead of pursuing the parallel theory of winding fractions for Čech

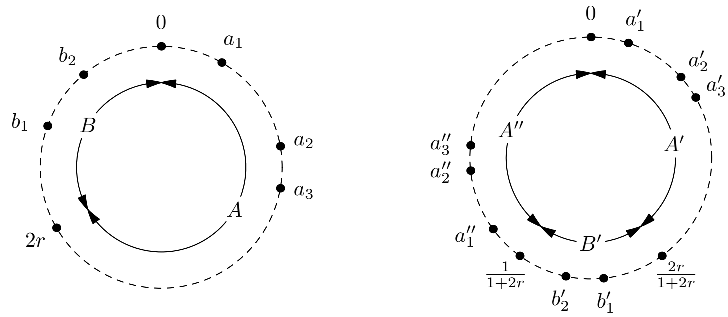


Figure 3. The action of the operator T_r from Theorem 9.3. Left: a set X split as $X = A \sqcup B$, where $A = X \cap [0, 2r)_{S^1}$ and $B = X \cap [2r, 1)_{S^1}$. Right: $T_r(X) = A' \sqcup B' \sqcup A''$, where A', A'', and B' are suitably rescaled and shifted copies of A and B. The map $\pi_r : T_r(X) \to X$ sends back A' to A, B' to B, and A'' to A.

complexes, we provide a direct transformation from Vietoris–Rips complexes to Čech complexes. This transformation recovers most but not all of the results that could be obtained with the parallel theory, and we believe it is of independent interest. To our knowledge, Theorems 9.7 and 9.8 are the first computation for a noncontractible connected manifold M of the homotopy types of $\check{\mathbf{C}}(M, M; r)$ for arbitrary r.

Let $\mathcal{P}(S^1)$ denote the power set of S^1. If $X \subseteq S^1$ and $a, b \in \mathbb{R}$ then we write $aX + b = \{(ax + b) \bmod 1 : x \in X\}$, where it is understood that each point x is represented by a real number in $[0, 1)$.

Theorem 9.3. *For each* $0 < r < \frac{1}{2}$ *let* $T_r : \mathcal{P}(S^1) \to \mathcal{P}(S^1)$ *be given by*

$$T_r(X) = \tfrac{1}{1+2r} X \cup \left(\tfrac{1}{1+2r} \cdot (X \cap [0, 2r)_{S^1}) + \tfrac{1}{1+2r}\right).$$

Then the (*noncontinuous*) *map* $\pi_r : S^1 \to S^1$ *defined by*

$$\pi_r(y) = (1 + 2r)y \bmod 1 \quad \textit{for} \ \ y \in [0, 1)$$

induces a simplicial homotopy equivalence

$$\pi_r : \mathbf{VR}_{\leq}\left(T_r(X); \tfrac{2r}{1+2r}\right) \xrightarrow{\simeq} \check{\mathbf{C}}_{\leq}(X; r).$$

Proof. We first verify that $\pi_r(T_r(X)) = X$. Take any $y \in T_r(X)$. If $y = \frac{1}{1+2r}x$ for $x \in X$ then $\pi_r(y) = x$. If $y = \frac{1}{1+2r}x + \frac{1}{1+2r}$ for some $x \in X \cap [0, 2r)_{S_1}$ then $\pi_r(y) \equiv x + 1 \equiv x \bmod 1$. It means that π_r restricts to a surjection $\pi_r : T_r(X) \to X$.

Next we check that π_r induces a map of simplicial complexes. Let σ be any face of the complex $\mathbf{VR}_{\leq}\left(T_r(X); \frac{2r}{1+2r}\right)$ and let $x_0 = \min(\sigma)$, so that $\sigma \cap [0, x_0)_{S^1} = \varnothing$.

To prove that the subset $\pi_r(\sigma)$ is a face of $\check{\mathbf{C}}_\le(X;r)$ we need to show that it is contained in a closed arc of length $2r$. There are three cases:

- $x_0 \in \big[\frac{1}{1+2r}, 1\big)_{S^1}$. Then $\sigma \subseteq \big[\frac{1}{1+2r}, 1\big)_{S^1}$ and $\pi_r(\sigma) \subseteq [0, 2r)_{S^1}$.
- $x_0 \in \big[\frac{2r}{1+2r}, \frac{1}{1+2r}\big)_{S^1}$. Then the only way x_0 can be in distance at most $\frac{2r}{1+2r}$ from the other points in σ is if that distance is measured clockwise from x_0. It means that $\sigma \subseteq \big[x_0, x_0 + \frac{2r}{1+2r}\big]_{S^1}$ with $x_0 + \frac{2r}{1+2r} < 1$ as well as $\pi_r(\sigma) \subseteq [(1+2r)x_0, (1+2r)x_0 + 2r \mod 1]_{S^1}$.
- $x_0 \in \big[0, \frac{2r}{1+2r}\big)_{S^1}$. Note that $\big(x_0 - \frac{2r}{1+2r}\big) \bmod 1 = x_0 + \frac{1}{1+2r}$, hence we can write $\sigma \subseteq \big[x_0, x_0 + \frac{2r}{1+2r}\big]_{S^1} \cup \big[x_0 + \frac{1}{1+2r}, 1\big)_{S^1}$. An application of π_r gives

$$\begin{aligned}\pi_r(\sigma) &\subseteq [(1+2r)x_0, (1+2r)x_0 + 2r]_{S^1} \cup [(1+2r)x_0, 2r)_{S^1} \\ &\subseteq [(1+2r)x_0, (1+2r)x_0 + 2r]_{S^1}.\end{aligned}$$

To prove that π_r is a homotopy equivalence it suffices to check that the preimage $\pi_r^{-1}(\tau)$ of every face $\tau \in \check{\mathbf{C}}_\le(X;r)$ is contractible. The conclusion is then provided by the simplicial version of Quillen's Theorem A due to Barmak [2011, Theorem 4.2]. Suppose that

$$\tau = \{a_1, \dots, a_s\} \cup \{b_1, \dots, b_t\},$$

where possibly $s = 0$ or $t = 0$, and

$$0 \le a_1 < \cdots < a_s < 2r \le b_1 < \cdots < b_t < 1.$$

The preimage $\pi_r^{-1}(\tau)$ is the subcomplex of $\mathbf{VR}_\le\big(T_r(X); \frac{2r}{1+2r}\big)$ induced by the vertex set

$$V(\pi_r^{-1}(\tau)) = \{a'_1, \dots, a'_s\} \cup \{b'_1, \dots, b'_t\} \cup \{a''_1, \dots, a''_s\}$$

where $a'_i = \frac{1}{1+2r}a_i$, $b'_i = \frac{1}{1+2r}b_i$, and $a''_i = \frac{1}{1+2r}a_i + \frac{1}{1+2r}$; see Figure 3.

Note that

$$\begin{cases} \vec{d}(a'_i, b'_j) = \frac{1}{1+2r}\vec{d}(a_i, b_j) & \text{for all } i, j, \\ \vec{d}(b'_i, a''_j) = \frac{1}{1+2r}\vec{d}(b_i, a_j) & \text{for all } i, j. \end{cases}$$

Moreover,

$$\begin{cases} \vec{d}(b'_i, b'_j) = \frac{1}{1+2r}\vec{d}(b_i, b_j) & \text{for } i < j, \\ \vec{d}(a'_i, a''_j) = \frac{1}{1+2r}\vec{d}(a_i, a_j) & \text{for } i > j, \\ \vec{d}(a''_i, a'_i) = \frac{2r}{1+2r}. \end{cases}$$

Since τ is contained in an arc of length $2r$, there is a vertex $x_0 \in \tau$ such that $\tau \subseteq [x_0, x_0 + 2r]_{S^1}$. We find a vertex $y_0 \in V(\pi_r^{-1}(\tau))$ such that y_0 is at most $\frac{2r}{1+2r}$ away from every other vertex of $V(\pi_r^{-1}(\tau))$. This will end the proof, since then $\pi_r^{-1}(\tau)$ is a cone with apex y_0. We need to consider four cases.

- If $x_0 = a_q$ for some $1 \le q \le s$ then $y_0 = a'_q$. The arc $[a_q, a_q + 2r]_{S^1}$ contains all points of τ, therefore
$$\big[a'_q, a'_q + \tfrac{2r}{1+2r}\big]_{S_1}$$
contains all points of $V(\pi_r^{-1}(\tau))$ up to the point preceding a''_q. Moreover $\vec{d}(a''_q, a'_q) = \frac{2r}{1+2r}$ and $[a''_q, a'_q]_{S^1}$ covers the remaining points.
- If $x_0 = b_1$ and $s = 0$ then take $y_0 = b'_1$. Since $\vec{d}(b_1, b_t) \le 2r$, we have $\vec{d}(b'_1, b'_t) \le \frac{2r}{1+2r}$.
- If $x_0 = b_1$ and $s > 0$ then take $y_0 = a''_s$. From $\vec{d}(b_1, a_s) \le 2r$ we get the inequality $\vec{d}(b'_1, a''_s) \le \frac{2r}{1+2r}$; also $\vec{d}(a''_s, a_s) = \frac{2r}{1+2r}$. This covers the distances from a''_s to all points of $V(\pi_r^{-1}(\tau))$.
- The last case, $x_0 = b_q$ with $q \ge 2$, is impossible since $\vec{d}(b_q, b_{q-1}) > 2r$. $\square$

Remark 9.4. Theorem 9.3 has no variant for $\mathbf{VR}_<$ and $\check{\mathbf{C}}_<$, since the set $T_r(X)$ contains pairs of points in distance exactly $\frac{2r}{1+2r}$ whose existence is essential for the proof.

In many natural circumstances maps of Čech complexes can be lifted to maps of Vietoris–Rips complexes via π_r. Below we describe the case of inclusions.

Proposition 9.5. *Suppose $X \subseteq S^1$ and $0 < r \le r' < \frac{1}{2}$. Then the (noncontinuous) map $\eta : S^1 \to S^1$ given by*
$$\eta(y) = \tfrac{1+2r}{1+2r'} \cdot y \quad \textit{for } y \in [0, 1)$$
determines a map of Vietoris–Rips complexes which makes the following diagram commute:
$$\begin{tikzcd}
\mathbf{VR}_\le\big(T_r(X); \tfrac{2r}{1+2r}\big) \arrow[r, "\eta"] \arrow[d, "\pi_r"', "\simeq"] & \mathbf{VR}_\le\big(T_{r'}(X); \tfrac{2r'}{1+2r'}\big) \arrow[d, "\pi_{r'}"', "\simeq"] \\
\check{\mathbf{C}}_\le(X; r) \arrow[r, hook, "\subseteq"] & \check{\mathbf{C}}_\le(X; r')
\end{tikzcd}$$

Proof. We first verify that η gives a well-defined map of Vietoris–Rips complexes in the top row of the diagram; it suffices to check this map on vertices and edges. For vertices, pick any $y \in T_r(X)$. If $y = \frac{1}{1+2r}x$ for $x \in X$ then $\eta(y) = \frac{1}{1+2r'}x \in T_{r'}(X)$. If $y = \frac{1}{1+2r}x + \frac{1}{1+2r}$ for $x \in X \cap [0, 2r)$, then $\eta(y) = \frac{1}{1+2r'}x + \frac{1}{1+2r'}$ with $x \in X \cap [0, 2r) \subseteq X \cap [0, 2r')$, and hence also in this case $\eta(y) \in T_{r'}(X)$. For edges, we suppose that $0 \le y < y' < 1$. If $\vec{d}(y, y') \le \frac{2r}{1+2r}$ then
$$\vec{d}(\eta(y), \eta(y')) \le \tfrac{2r}{1+2r} \cdot \tfrac{1+2r}{1+2r'} \le \tfrac{2r'}{1+2r'}.$$
If $\vec{d}(y', y) \le \frac{2r}{1+2r}$ then $\vec{d}(y, y') \ge \frac{1}{1+2r}$, hence we get
$$\vec{d}(\eta(y), \eta(y')) \ge \tfrac{1}{1+2r} \cdot \tfrac{1+2r}{1+2r'} = \tfrac{1}{1+2r'},$$

and therefore $\vec{d}(\eta(y'), \eta(y)) \leq \frac{2r'}{1+2r'}$.

Commutativity of the diagram follows from a direct calculation:

$$\pi_{r'}(\eta(y)) \equiv (1+2r') \cdot \tfrac{1+2r}{1+2r'} \cdot y \equiv (1+2r)y \equiv \pi_r(y) \bmod 1. \qquad \square$$

For arbitrary $X \subseteq S^1$ Theorem 9.3 allows one to determine the homotopy type of $\check{\mathbf{C}}_{\leq}(X; r)$ from an efficiently constructible instance of the Vietoris–Rips complex. Here are some examples.

Example 9.6. Let $X_n = \{0, \frac{1}{n}, \dots, \frac{n-1}{n}\} \subseteq S^1$. Then $\check{\mathbf{C}}_{\leq}(X_n; \frac{k}{2n})$ is the complex whose maximal faces are generated from $\{0, \frac{1}{n}, \dots, \frac{k}{n}\}$ via rotations by $\frac{1}{n}$. If $r = \frac{k}{2n}$ then $\frac{2r}{1+2r} = \frac{k}{n+k}$ and $T_r(X_n) = X_{n+k}$. We obtain a homotopy equivalence

$$\mathrm{Cl}(C^k_{n+k}) = \mathbf{VR}_{\leq}\big(X_{n+k}; \tfrac{k}{n+k}\big) \xrightarrow{\simeq} \check{\mathbf{C}}_{\leq}\big(X_n; \tfrac{k}{2n}\big).$$

This special case was proved in [Adamaszek et al. 2016, Theorem 8.5].

Theorem 9.7. *For* $0 < r < \frac{1}{2}$ *we have a homotopy equivalence*

$$\check{\mathbf{C}}_{\leq}(S^1; r) \simeq \begin{cases} S^{2l+1} & \text{if } \frac{l}{2(l+1)} < r < \frac{l+1}{2(l+2)},\ l = 0, 1, \dots, \\ \bigvee^{\mathfrak{c}} S^{2l} & \text{if } r = \frac{l}{2(l+1)}. \end{cases}$$

Moreover, if $\frac{l}{2(l+1)} < r \leq r' < \frac{l+1}{2(l+2)}$ *then the inclusion* $\check{\mathbf{C}}_{\leq}(S^1; r) \hookrightarrow \check{\mathbf{C}}_{\leq}(S^1; r')$ *is a homotopy equivalence.*

Proof. Note that $r \mapsto \frac{2r}{1+2r}$ is a monotone map which takes the interval $\big[\frac{l}{2(l+1)}, \frac{l+1}{2(l+2)}\big)$ to $\big[\frac{l}{2l+1}, \frac{l+1}{2l+3}\big)$. Since $T_r(S^1) = S^1$ we get a homotopy equivalence

$$\mathbf{VR}_{\leq}\big(S^1; \tfrac{2r}{1+2r}\big) \xrightarrow{\simeq} \check{\mathbf{C}}_{\leq}(S^1; r),$$

and the statement of homotopy types now follows from Theorem 7.6.

The statement about inclusions will follow from Proposition 9.5 if we show that the map $\mathbf{VR}_{\leq}\big(S^1; \frac{2r}{1+2r}\big) \xrightarrow{\eta} \mathbf{VR}_{\leq}\big(S^1; \frac{2r'}{1+2r'}\big)$ is a homotopy equivalence. Pick a finite set $Y_0 \subseteq S^1$ with $\mathrm{wf}_{\leq}\big(Y_0; \frac{2r}{1+2r}\big) > \frac{l}{2l+1}$. We have a commutative diagram

$$\begin{tikzcd} \mathbf{VR}_{\leq}\big(Y_0; \frac{2r}{1+2r}\big) \arrow[r, "\eta"] \arrow[d, hook] & \mathbf{VR}_{\leq}\big(Y_0 \cup \eta(Y_0); \frac{2r'}{1+2r'}\big) \arrow[d, hook] \\ \mathbf{VR}_{\leq}\big(S^1; \frac{2r}{1+2r}\big) \arrow[r, "\eta"] & \mathbf{VR}_{\leq}\big(S^1; \frac{2r'}{1+2r'}\big). \end{tikzcd}$$

The vertical inclusions are homotopy equivalences by Remark 7.2. The map η is injective and preserves the clockwise order of points on S^1, hence it is a cyclic homomorphism of the cyclic graphs underlying the top row of the diagram. As $\frac{l}{2l+1} < \mathrm{wf}_{\leq}\big(Y_0; \frac{2r}{1+2r}\big) \leq \mathrm{wf}_{\leq}\big(Y_0 \cup \eta(Y_0); \frac{2r'}{1+2r'}\big) < \frac{l+1}{2l+3}$, the top row is a homotopy equivalence by Proposition 4.9. $\square$

The following is an analogue of Theorem 7.4 for Čech complexes in the case when $X = S^1$.

Theorem 9.8. *For $0 < r < \frac{1}{2}$ we have a homotopy equivalence*

$$\check{\mathbf{C}}_<(S^1; r) \simeq S^{2l+1} \quad \text{for } \tfrac{l}{2(l+1)} < r \le \tfrac{l+1}{2(l+2)},\ l = 0, 1, \ldots.$$

Moreover, if $\frac{l}{2(l+1)} < r \le r' \le \frac{l+1}{2(l+2)}$ then the inclusion $\check{\mathbf{C}}_<(S^1; r) \hookrightarrow \check{\mathbf{C}}_<(S^1; r')$ is a homotopy equivalence.

Proof. Fix $\frac{l}{2(l+1)} < r \le \frac{l+1}{2(l+2)}$ and note that $\check{\mathbf{C}}_<\big(S^1, S^1; r\big) = \operatorname{colim}_n \check{\mathbf{C}}_\le\big(S^1; r - \frac{1}{n}\big)$. All inclusions

$$\check{\mathbf{C}}_\le\big(S^1; r - \tfrac{1}{n}\big) \hookrightarrow \check{\mathbf{C}}_\le\big(S^1; r - \tfrac{1}{n+1}\big)$$

are cofibrations and by Theorem 9.7 they are self-homotopy equivalences of S^{2l+1} for sufficiently large n. That proves the statement of homotopy types.

For the statement about inclusions, note the inclusions

$$\check{\mathbf{C}}_\le\big(S^1; r - \tfrac{1}{n}\big) \hookrightarrow \check{\mathbf{C}}_\le\big(S^1; r' - \tfrac{1}{n}\big)$$

define a natural transformation of diagrams which is a levelwise homotopy equivalence for sufficiently large n by Theorem 9.7. It follows that the induced map of (homotopy) colimits $\check{\mathbf{C}}_<(S^1; r) \hookrightarrow \check{\mathbf{C}}_<(S^1; r')$ is a homotopy equivalence. □

10. Concluding remarks

A natural generalization of our results would be to investigate the complexes $\mathbf{VR}(M; r)$ and $\check{\mathbf{C}}(M, M; r)$ for Riemannian manifolds M other than S^1, though very little is known along these lines. Intriguing examples include the spheres S^n and tori $(S^1)^n$ for $n \ge 2$. One difficulty is that it is not known whether the homotopy type of $\mathbf{VR}(M; r)$ can be approximated by those of complexes $\mathbf{VR}(X; r)$ for sufficiently dense subsets $X \subseteq M$. Furthermore, already for $M = S^2$ the complete list of homotopy types of complexes $\mathbf{VR}(X; r)$ for finite subsets $X \subseteq S^2$ is not known.

Towards the goal of understanding Vietoris–Rips complexes of more spaces, we briefly describe two results, the homotopy types of annuli and of tori equipped with the ℓ_∞ metric, which can be derived from our computation of $\mathbf{VR}(S^1; r)$ using known tools.

Proposition 10.1. *Consider the annulus $D(\rho, \tilde{\rho}) = \{(x, y) \in \mathbb{R}^2 : \rho^2 \le x^2 + y^2 \le \tilde{\rho}^2\}$ with the Euclidean metric. Then for any $r > 0$ the space $\mathbf{VR}_<(D(\rho, \tilde{\rho}); r)$ is homotopy equivalent to an odd-dimensional sphere or to a point.*

Proof. The homotopy which radially deforms the annulus onto its inner boundary does not increase distances, and so it is a *crushing* map in the sense of Hausmann [1995]. By Proposition (2.2) of that reference, the inclusion of the Vietoris–Rips complex of S^1 into that of $D(\rho, \tilde{\rho})$ is a homotopy equivalence, and so the result follows from Theorem 7.4. $\square$

We include a proof of a result for which we were unable to find a published reference.

Proposition 10.2. *Suppose $(M_1, d_1), \dots, (M_n, d_n)$ are metric spaces and $M = M_1 \times \cdots \times M_n$ is their product equipped with the supremum metric*

$$\ell_\infty((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max\{d_i(x_i, y_i) : i = 1, \dots, n\}.$$

Then for any $r > 0$ we have a homotopy equivalence

$$\mathbf{VR}(M; r) \simeq \mathbf{VR}(M_1; r) \times \cdots \times \mathbf{VR}(M_n; r).$$

Proof. For simplicial complexes $K_1, \dots, K_n$ the categorical product [Kozlov 2008, Definition 4.25] (in the category of abstract simplicial complexes) is the complex $\prod_i K_i$ with vertex set $V(K_1) \times \cdots \times V(K_n)$ and with faces given by the condition: $\sigma \in \prod_i K_i$ if and only if $\sigma \subseteq \sigma_1 \times \cdots \times \sigma_n$ for some $\sigma_i \in K_i,\ i = 1, \dots, n$. Since a subset of M has diameter equal to the maximum of the diameters of its coordinate projections, we get an isomorphism of simplicial complexes

$$\mathbf{VR}(M; r) = \prod_{i=1}^{n} \mathbf{VR}(M_i; r).$$

There is a homotopy equivalence $\prod_i K_i \simeq K_1 \times \cdots \times K_n$ when each K_i is a finite simplicial complex [Kozlov 2008, Proposition 15.23], and one can see that the finiteness assumption is not necessary by combining the same proof with a version of the nerve lemma for infinite simplicial complexes [Björner 1995, Theorem 10.6]. That ends the proof. $\square$

Applied to the torus $\mathbb{T}^n = (S^1)^n$ the last proposition yields the homotopy types of $\mathbf{VR}(\mathbb{T}^n; r)$ for the ℓ_∞ metric on $\mathbb{T}^n$. It would be interesting to investigate the homotopy types of $\mathbf{VR}(\mathbb{T}^n; r)$ for other ℓ_p metrics on $\mathbb{T}^n$, especially for the ℓ_2 metric.

Appendix A

We prove the following algebraic fact, which is used in the proof of Proposition 7.5.

Proposition A.1. *Suppose that $V = \{e_1, \dots, e_n\}$ is a basis of the free abelian group $\mathbb{Z}^n$. Consider the set of $n + \binom{n}{2}$ vectors*

$$\widetilde{V} = \{e_1, \dots, e_n\} \cup \{e_i - e_j : 1 \le i < j \le n\}.$$

For an arbitrary choice $v, v_1, \dots, v_k \in \widetilde{V}$, if the subgroup of $\mathbb{Z}^n$ generated by $\{v_1, \dots, v_k\}$ contains some nonzero multiple of v, then it also contains v.

Proof. We can assume $v, v_1, \dots, v_k$ are pairwise distinct. Let $A = \{v, v_1, \dots, v_k\}$. Note that when expressed in the basis $\{e_1, \dots, e_n\}$, any two vectors in $\widetilde{V}$ have at most one nonzero coordinate in common. By symmetry it suffices to consider the cases $v = e_1$ and $v = e_1 - e_2$.

First suppose $v = e_1$. We have the identity

$$pe_1 = \sum_{v_i \in A \setminus \{e_1\}} a_i v_i$$

for some $p, a_1, \dots, a_k \in \mathbb{Z}$ with $p \neq 0$. Consider a labeled graph G with vertex set A, where two vectors are connected by an edge with label i $(1 \le i \le n)$ if they both have nonzero i-th coordinate. Let A_1 be the vertex set of the connected component of G containing e_1. Then we still have the identity

$$pe_1 = \sum_{v_i \in A_1 \setminus \{e_1\}} a_i v_i$$

because the vectors in A_1 and $A \setminus A_1$ contribute to two nonoverlapping sets of coordinates.

It is not possible that all the vectors in $A_1 \setminus \{e_1\}$ are of the form $e_i - e_j$. Indeed, any linear combination of such vectors has the sum of its coordinates equal to 0, whereas pe_1 does not. Hence the connected component A_1 contains some vector e_l with $l \neq 1$. Consider the shortest path in G from e_l to e_1. It is easy to see that no edge label appears along this path more than once and that all the intermediate vertices are vectors of the form $e_i - e_j$. The shortest path has the form

$$e_l = e_{l_0} \to \pm(e_{l_0} - e_{l_1}) \to \pm(e_{l_1} - e_{l_2}) \to \cdots \to \pm(e_{l_{s-1}} - e_{l_s}) \to e_{l_s} = e_1$$

for some $s \ge 1$, where $l_0 = l$ and $l_s = 1$, all l_i are pairwise distinct, and $\pm(e_i - e_j)$ stands for $e_{\min(i,j)} - e_{\max(i,j)}$. Now we obtain a presentation

$$e_1 = e_{l_0} + (e_{l_1} - e_{l_0}) + (e_{l_2} - e_{l_1}) + \cdots + (e_{l_s} - e_{l_{s-1}})$$

of e_1 as a linear combination of elements of $A_1 \setminus \{e_1\}$ (with coefficients ± 1). That ends the proof of the proposition for $v = e_1$.

The other case, $v = e_1 - e_2$, can be reduced to the previous one as follows. Set $e_1' = e_1 - e_2$, $e_2' = -e_2$, $e_3' = e_3 - e_2$, $\dots$, $e_n' = e_n - e_2$. The set $V' = \{e_1', \dots, e_n'\}$ is a basis of $\mathbb{Z}^n$. Moreover, up to signs, the sets $\widetilde{V}'$ and $\widetilde{V}$ coincide. The assumption that $p(e_1 - e_2)$ is a combination of $v_1, \dots, v_k \in \widetilde{V}$ is therefore equivalent to the assumption that pe_1' is a combination of $\pm v_1, \dots, \pm v_k \in \widetilde{V}'$. From the previous case we get that $e_1' = e_1 - e_2$ is also a linear combination of $v_1, \dots, v_k$. $\square$

Appendix B

In this section we prove (6) which gives the expected waiting time for the appearance of an (ε, m)-regular subset in a random sampling of S^1.

We will first determine the waiting times for some occupancy problems in the "balls into bins" model. Let $K \geq 1$ be the number of bins and fix a constant $m \geq 1$. Consider the following random experiments.

(a) We throw balls independently and uniformly at random into K bins until one of the bins contains m balls. Let $A_m(K)$ be the random variable denoting the number of balls thrown. By [Klamkin and Newman 1967, Theorem 2],

$$\mathbf{E}[A_m(K)] = \Theta(K^{\frac{m-1}{m}}) \quad \text{as } K \to \infty.$$

This is known as the generalized birthday paradox, the case $m = 2$ (and $K = 365$ in the folklore formulation) being the classical birthday paradox.

(b) We throw balls as before, but each time a ball is thrown we assign it, uniformly at random, with one of m colors. When a bin with m balls appears, we call the sequence of colors in that bin, in the order in which they were thrown, the *outcome* of the experiment. The outcome is *good* if all of the m balls have different colors. The number of balls thrown is still given by the random variable $A_m(K)$, since the colors do not influence the stopping condition. Since the balls were colored independently and uniformly, each outcome is equally likely. In particular, the probability of a good outcome is $m!/m^m$.

(c) We repeat the experiment of (b) until we obtain a good outcome, each time starting with a fresh set of empty bins. Let τ be the random variable counting the number of repetitions and let $B_m(K)$ be the total number of balls thrown. We have

$$B_m(K) = A_m(K)_1 + \cdots + A_m(K)_\tau,$$

where the $A_m(K)_i$ are independent random variables with the distribution of $A_m(K)$. Clearly τ is a stopping time with respect to these variables, and by the discussion in (b) we have $\mathbf{E}[\tau] = m^m/m!$. Now Wald's equation gives

$$\mathbf{E}[B_m(K)] = \mathbf{E}[A_m(K)] \cdot \mathbf{E}[\tau] = \mathbf{E}[A_m(K)] \cdot \frac{m^m}{m!} = \Theta(K^{\frac{m-1}{m}}) \quad \text{as } K \to \infty.$$

(d) We throw balls independently and uniformly at random into K bins and we color each ball uniformly with one of m colors, until some bin contains at least one ball of each color. If $C_m(K)$ is the random variable counting the number of balls thrown then $A_m(K) \leq C_m(K) \leq B_m(K)$ and

$$\mathbf{E}[C_m(K)] = \Theta(K^{\frac{m-1}{m}}) \quad \text{as } K \to \infty.$$

In the classical case $m = 2$ this is known as the birthday paradox with two types (a boy sharing a birthday with a girl); we were unable to find a literature reference for this result with arbitrary m.

Recall that $R_m(\varepsilon)$ is the number of points chosen uniformly at random from S^1 until an (ε, m)-regular subset appears. We claim that for any integer $K \geq 1$,

$$R_m\Big(\frac{1}{Km}\Big) \leq C_m(K). \tag{11}$$

To see this divide S^1 into arcs of length $\frac{1}{Km}$. Each union of m arcs whose centers form a regular m-gon represents one of our K bins. A uniformly random point $x \in S^1$ can be chosen by picking a uniformly random point $y \in \big[0, \frac{1}{m}\big)_{S^1}$ and a random number $i \in \{0, \ldots, m-1\}$ and setting $x = y + \frac{i}{m}$; note y determines the bin and i determines the color of the ball. When some bin contains a ball of each color, the corresponding points are within $\frac{1}{Km}$ $\big($in fact even $\frac{1}{2Km}\big)$ from the vertices of a regular m-gon.

Next, we claim that

$$\mathbf{E}[A_m(K)] \leq 2\mathbf{E}\big[R_m\big(\tfrac{1}{4Km}\big)\big]. \tag{12}$$

Let P_1 be the collection of arcs and bins as above, and let P_2 be the same collection rotated by $\frac{1}{2Km}$. If Y is $\big(\frac{1}{4Km}, m\big)$-regular then all points of Y belong to the same bin with respect to P_1 or with respect to P_2 (or both). Let p_1 (resp., p_2) be the probability that the first time a $\big(\frac{1}{4Km}\big)$-regular set emerges in the random process $(\mathcal{X}_1, \mathcal{X}_2, \ldots)$, it is contained in one bin with respect to P_1 (resp., P_2). By symmetry $p_1 = p_2$, therefore $p_1 \geq \frac{1}{2}$. If we repeat the whole process until it ends with a set in P_1 then the expected number of repetitions is $\mathbf{E}[\tau] = \frac{1}{p_1} \leq 2$. An argument similar to that in (c) above proves (12).

Letting $K \to \infty$ in (11) and (12) and using the asymptotics of $\mathbf{E}[A_m(K)]$ and $\mathbf{E}[C_m(K)]$ we obtain

$$\mathbf{E}[R_m(\varepsilon)] = \Theta\Big(\big(\tfrac{1}{\varepsilon}\big)^{\frac{m-1}{m}}\Big) \quad \text{as } \varepsilon \to 0,$$

which proves (6).

Acknowledgements

We thank Florian Frick, Matthew Kahle, Vin de Silva, and Francis Motta for helpful conversations, and the anonymous referee for helpful remarks concerning the presentation.

References

[Adamaszek 2013] M. Adamaszek, "Clique complexes and graph powers", *Israel J. Math.* **196**:1 (2013), 295–319. MR Zbl

[Adamaszek et al. 2016] M. Adamaszek, H. Adams, F. Frick, C. Peterson, and C. Previte-Johnson, "Nerve complexes of circular arcs", *Discrete Comput. Geom.* **56**:2 (2016), 251–273. MR Zbl

[Adamaszek et al. 2017] M. Adamaszek, H. Adams, and F. Motta, "Random cyclic dynamical systems", *Adv. in Appl. Math.* **83** (2017), 1–23. MR Zbl

[Barmak 2011] J. A. Barmak, "On Quillen's Theorem A for posets", *J. Combin. Theory Ser. A* **118**:8 (2011), 2445–2453. MR Zbl

[Björner 1995] A. Björner, "Topological methods", pp. 1819–1872 in *Handbook of combinatorics, II*, edited by R. L. Graham et al., Elsevier, Amsterdam, 1995. MR Zbl

[Carlsson 2009] G. Carlsson, "Topology and data", *Bull. Amer. Math. Soc.* (*N.S.*) **46**:2 (2009), 255–308. MR Zbl

[Carlsson et al. 2008] G. Carlsson, T. Ishkhanov, V. de Silva, and A. Zomorodian, "On the local behavior of spaces of natural images", *Int. J. Comp. Vision* **76**:1 (2008), 1–12.

[Chambers et al. 2010] E. W. Chambers, V. de Silva, J. Erickson, and R. Ghrist, "Vietoris–Rips complexes of planar point sets", *Discrete Comput. Geom.* **44**:1 (2010), 75–90. MR Zbl

[Chazal et al. 2014] F. Chazal, V. de Silva, and S. Oudot, "Persistence stability for geometric complexes", *Geom. Dedicata* **173** (2014), 193–214. MR Zbl

[Edelsbrunner and Harer 2010] H. Edelsbrunner and J. L. Harer, *Computational topology: an introduction*, American Mathematical Society, Providence, RI, 2010. MR Zbl

[Gromov 1987] M. Gromov, "Hyperbolic groups", pp. 75–263 in *Essays in group theory*, edited by S. M. Gersten, Math. Sci. Res. Inst. Publ. **8**, Springer, New York, 1987. MR Zbl

[Hatcher 2002] A. Hatcher, *Algebraic topology*, Cambridge Univ. Press, 2002. MR Zbl

[Hausmann 1995] J.-C. Hausmann, "On the Vietoris–Rips complexes and a cohomology theory for metric spaces", pp. 175–188 in *Prospects in topology* (Princeton, 1994), edited by F. Quinn, Ann. of Math. Stud. **138**, Princeton Univ. Press, 1995. MR Zbl

[Hell and Nešetřil 2004] P. Hell and J. Nešetřil, *Graphs and homomorphisms*, Oxford Lecture Series in Mathematics and its Applications **28**, Oxford Univ. Press, 2004. MR Zbl

[Huntington 1916] E. V. Huntington, "A set of independent postulates for cyclic order", *Proc. Nat. Acad. Sci. USA* **2**:11 (1916), 630–631. JFM

[Imany-Nabiyyi 2008] R. Imany-Nabiyyi, "The sizes of components in random circle graphs", *Discuss. Math. Graph Theory* **28**:3 (2008), 511–533. MR Zbl

[Kahle 2009] M. Kahle, "Topology of random clique complexes", *Discrete Math.* **309**:6 (2009), 1658–1671. MR Zbl

[Klamkin and Newman 1967] M. S. Klamkin and D. J. Newman, "Extensions of the birthday surprise", *J. Combinatorial Theory* **3**:3 (1967), 279–282. MR

[Kozlov 2008] D. Kozlov, *Combinatorial algebraic topology*, Algorithms and Computation in Mathematics **21**, Springer, Berlin, 2008. MR Zbl

[Latschev 2001] J. Latschev, "Vietoris–Rips complexes of metric spaces near a closed Riemannian manifold", *Arch. Math.* (*Basel*) **77**:6 (2001), 522–528. MR Zbl

[Lin and Szwarcfiter 2009] M. C. Lin and J. L. Szwarcfiter, "Characterizations and recognition of circular-arc graphs and subclasses: a survey", *Discrete Math.* **309**:18 (2009), 5618–5635. MR Zbl

[Matoušek 2008] J. Matoušek, "LC reductions yield isomorphic simplicial complexes", *Contrib. Discrete Math.* **3**:2 (2008), 37–39. MR Zbl

[Solomon 1978] H. Solomon, *Geometric probability*, Regional Conf. Series in Appl. Math. **28**, Society for Industrial and Applied Mathematics, Philadelphia, 1978. MR Zbl

[Vietoris 1927] L. Vietoris, "Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen", *Math. Ann.* **97**:1 (1927), 454–472. MR JFM

[Welker et al. 1999] V. Welker, G. M. Ziegler, and R. T. Živaljević, "Homotopy colimits: comparison lemmas for combinatorial applications", *J. Reine Angew. Math.* **509** (1999), 117–149. MR Zbl

Received May 22, 2015. Revised October 13, 2016.

MICHAŁ ADAMASZEK
DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF COPENHAGEN
UNIVERSITETSPARKEN 5
DK-2100 COPENHAGEN
DENMARK

aszek@mimuw.edu.pl

HENRY ADAMS
DEPARTMENT OF MATHEMATICS
COLORADO STATE UNIVERSITY
1874 CAMPUS DELIVERY
FORT COLLINS, CO 80524
UNITED STATES

adams@math.colostate.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.41

A TALE OF TWO LIOUVILLE CLOSURES

ALLEN GEHRET

An H-field is a type of ordered valued differential field with a natural interaction between ordering, valuation, and derivation. The main examples are Hardy fields and fields of transseries. Aschenbrenner and van den Dries (2002) proved that every H-field K has either exactly one or exactly two Liouville closures up to isomorphism over K, but the precise dividing line between these two cases was unknown. We prove here that this dividing line is determined by λ-freeness, a property of H-fields that prevents certain deviant behavior. In particular, we show that under certain types of extensions related to adjoining integrals and exponential integrals, the property of λ-freeness is preserved. In the proofs we introduce a new technique for studying H-fields, the *yardstick argument* which involves the rate of growth of pseudoconvergence.

MSC2010: primary 12H05, 12J10; secondary 06F20, 12J15, 26A12.
Keywords: H-fields, asymptotic fields, asymptotic couples, differential-valued fields, Liouville extensions, Liouville closures, λ-freeness.

1. Introduction

Consider the classical ordinary differential equation

$$(*)\qquad y' + fy = g$$

where f and g are sufficiently nice real-valued functions. To solve $(*)$, we first perform an *exponential integration* to obtain the so-called *integrating factor*

$$\mu = \exp \textstyle\int f.$$

Then we perform an *integration* to obtain a solution $y = \mu^{-1}\int(g\mu)$. In this paper, we wish to consider integration and exponential integration in the context of *H-fields*. H-fields and all other terms used in this introduction will be properly defined in the body of this paper.

H-fields are a certain kind of ordered valued differential field introduced in [Aschenbrenner and van den Dries 2002] and include all *Hardy fields* containing $\mathbb{R}$; Hardy fields are ordered differential fields of germs of real-valued functions defined on half-lines $(a, +\infty)$, (e.g., see [Bourbaki 1951, Chapitre V] or [Rosenlicht 1983a; 1983b]). Other examples include fields of *transseries* such as the *field of logarithmic-exponential transseries* $\mathbb{T}$ and the *field of logarithmic transseries* $\mathbb{T}_{\log}$ (e.g., see [Écalle 1992; van der Hoeven 2006; ADH 2017]). Our primary reference for the theory of H-fields, and all other things considered in this paper, is the work “Asymptotic differential algebra and model theory of transseries”, by Matthias Aschenbrenner, Lou van den Dries and Joris van der Hoeven, which we refer to as [ADH 2017].

A real closed H-field in which every equation of the form $(*)$ has a nonzero solution, with f and g ranging over K, is said to be *Liouville closed*. If K is an H-field, then a minimal Liouville closed H-field extension of K is called a *Liouville closure* of K. The main result of [Aschenbrenner and van den Dries 2002] is that for any H-field K, exactly one of the following occurs:

(I) K has exactly one Liouville closure up to isomorphism over K.

(II) K has exactly two Liouville closures up to isomorphism over K.

There are three distinct types of H-fields: an H-field K either is *grounded*, *has a gap*, or has *asymptotic integration*. According to that work, grounded H-fields fall into case (I) and H-fields with a gap fall into case (II). If an H-field has asymptotic integration, then it is either in case (I) or (II). However, the precise dividing line between (I) and (II) for asymptotic integration was not known.

The main result of this paper (Theorem 12.1) shows that this dividing line is exactly the property of *λ-freeness*. We prove that if an H-field is λ-free, then it is in case (I), and if an H-field has asymptotic integration and is not λ-free, then it is in case (II). This follows by combining known facts about λ-freeness from [ADH 2017] with our new technical results which show that λ-freeness is preserved under certain adjunctions of integrals and exponential integrals. In order to "defend" the λ-freeness of an H-field in these types of extensions, we introduce the *yardstick argument*, which concerns the "rate of pseudoconvergence" when adjoining integrals and exponential integrals.

We use many definitions and cite many results from [ADH 2017]. As a general rule, any result taken directly from that reference is titled ADH instead of Lemma, Theorem, etc. In citing results in this way we do not imply that they are originally due to the authors of [ADH 2017]; for instance, ADH 4.1 is actually a classical fact of valuation theory due to Kaplansky. Furthermore, in citations we omit qualifiers when no confusion should arise, writing, for example, [Gehret 2017a, 3.2] instead of [Gehret 2017a, Lemma 3.2].

In Section 2, we introduce the notion of a subset S of an ordered abelian group Γ being *jammed*. A set S being jammed corresponds to the elements near the top of S becoming closer and closer together at an unreasonably fast rate. Being jammed is an exotic property which we later wish to avoid.

In Section 3, we recall the basic theory of asymptotic couples and introduce and study the *yardstick property* of subsets of asymptotic couples. Asymptotic couples are pairs (Γ, ψ) where Γ is an ordered abelian group and $\psi : \Gamma \setminus \{0\} \to \Gamma$ is a map which satisfies, among other things, a valuation-theoretic version of l'Hôpital's rule. Asymptotic couples often arise as the value groups of H-fields, where the map ψ comes from the logarithmic derivative operation $f \mapsto f'/f$ for $f \neq 0$. Roughly speaking, a set S has the yardstick property if for any element $\gamma \in S$, there is a larger element $\gamma + \varepsilon(\gamma) \in S$ for a certain "yardstick" $\varepsilon(\gamma) > 0$ which depends on γ and which we can explicitly describe. In contrast to the notion of being jammed, the yardstick property is a desirable tame property. In Section 3 we show, among other things, that the two properties are incompatible, except in a single degenerate case. Asymptotic couples were introduced by Rosenlicht [1979; 1980; 1981] in order to study the value group of a differential field with a so-called *differential valuation*, what we call here a *differential-valued field*. For more on asymptotic couples, including the extension theory of asymptotic couples and some model-theoretic results concerning the asymptotic couples of $\mathbb{T}$ and $\mathbb{T}_{\log}$, see [Aschenbrenner and van den Dries 2000; Aschenbrenner 2003; Gehret 2017b; 2017a] and [ADH 2017, §6.5, §9.2, §9.8 and §9.9].

In Section 4 we recall definitions concerning pseudocauchy sequences in valued fields and some of the elementary facts concerning pseudocauchy sequences. The main result of Section 4 is Lemma 4.4 which is a rational version of Kaplansky's lemma (ADH 4.1). We assume the reader is familiar with basic valuation theory, including notions such as henselianity. As a general reference, see [ADH 2017, Chapters 2 and 3] or [Engler and Prestel 2005].

In Section 5 we give the definitions and relevant properties of *differential fields*, *valued differential fields*, *asymptotic fields*, *pre-differential-valued fields*, *differential-valued fields*, *pre-H-fields* and *H-fields*. These are the types of fields we will be concerned with in the later sections. Nearly everything from this section is from [ADH 2017] except for Lemmas 5.1 and 5.4 which are needed in our proof of Theorem 12.1.

In Section 6 we give a survey of the property of λ-freeness, citing many definitions and results from [ADH 2017, §11.5 and §11.6]. Many of these results we cite, and later use, involve situations where λ-freeness is preserved in certain valued differential field extensions. The main result of this section is Proposition 6.19 which shows that a rather general type of field extension preserves λ-freeness. Proposition 6.19 is related to the yardstick property of Section 3.

In Section 7, Section 8, and Section 9, we show that under various circumstances, if a pre-differential-valued field or a pre-H-field K is λ-free, and we adjoin an integral or an exponential integral to K for an element in K that does not already have an integral or exponential integral, then the resulting field extension will also be λ-free. The arguments in all three sections mirror one another and the main results, Propositions 7.2, 8.3, and 9.3 are all instances of Proposition 6.19.

In Sections 10 and 11 we give two minor applications of the results of Sections 7, 8, and 9. In Section 10 we show that λ-freeness is preserved when passing to the *differential-valued hull* of a λ-free pre-differential-field K (Theorem 10.2). In Section 11 we show that for λ-free differential-valued fields K, the minimum henselian, integration-closed extension $K\left(\int\right)$ of K is also λ-free (Theorem 11.2).

In Section 12 we prove the main result of this paper, Theorem 12.1. Combining it with the results in Section 10, we also obtain a generalization to the setting of pre-H-fields (Corollary 12.3). Finally, we provide proofs of claims made in [Aschenbrenner and van den Dries 2002; 2005] (Corollary 12.6 and Remark 12.7).

Conventions. Throughout, m and n range over the set $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ of natural numbers. By “ordered set” we mean “totally ordered set”.

Let S be an ordered set. Below, the ordering on S will be denoted by $\leq$, and a subset of S is viewed as ordered by the induced ordering. We put $S_\infty := S \cup \{\infty\}$, $\infty \notin S$, with the ordering on S extended to a (total) ordering on S_∞ by $S < \infty$. Suppose that B is a subset of S. We put $S^{>B} := \{s \in S : s > b \text{ for every } b \in B\}$ and we denote $S^{>\{a\}}$ as just $S^{>a}$; similarly for $\geq$, $<$, and $\leq$ instead of $>$. For $a, b \in S$ we put

$$[a, b] := \{x \in S : a \leq x \leq b\}.$$

A subset C of S is said to be *convex* in S if for all $a, b \in C$ we have $[a, b] \subseteq C$. A subset A of S is said to be a *downward closed* in S, if for all $a \in A$ and $s \in S$ we have $s < a \implies s \in A$. For $A \subseteq S$ we put

$$A^{\downarrow} := \{s \in S : s \leq a \text{ for some } a \in A\},$$

which is the smallest downward closed subset of S containing A.

A *well-indexed sequence* is a sequence (a_ρ) whose terms a_ρ are indexed by the elements ρ of an infinite well-ordered set without a greatest element.

Suppose that G is an ordered abelian group. Then we set $G^{\neq} := G \setminus \{0\}$. Also, $G^{<} := G^{<0}$; similarly for $\geq$, $\leq$, and $>$ instead of $<$. We define $|g| := \max\{g, -g\}$ for $g \in G$. For $a \in G$, the *archimedean class* of a is defined by

$$[a] := \{g \in G : |a| \leq n|g| \text{ and } |g| \leq n|a| \text{ for some } n \geq 1\}.$$

The archimedean classes partition G. Each archimedean class $[a]$ with $a \neq 0$ is the disjoint union of the two convex sets $[a] \cap G^{<}$ and $[a] \cap G^{>}$. We order the set $[G] := \{[a] : a \in G\}$ of archimedean classes by

$$[a] < [b] \ :\Longleftrightarrow\ n|a| < |b| \text{ for all } n \geq 1.$$

We have $[0] < [a]$ for all $a \in G^{\neq}$, and

$$[a] \leq [b] \iff |a| \leq n|b| \text{ for some } n \geq 1.$$

We shall consider G to be an ordered subgroup of its *divisible hull* $\mathbb{Q}G$. The divisible hull of G is the divisible abelian group $\mathbb{Q}G := \mathbb{Q} \otimes_{\mathbb{Z}} G$ equipped with the unique ordering which makes it an ordered abelian group containing G as an ordered subgroup.

2. Ordered abelian groups

In this section Γ *is an ordered abelian group,* $S \subseteq \Gamma$*,* $\alpha \in \Gamma$ *and* $n \geq 1$. We define:

$$\alpha + nS := \{\alpha + n\gamma : \gamma \in S\}.$$

A set of the form $\alpha + nS$ is called an *affine transform* of S. Many qualitative properties of a set $S \subseteq \Gamma$ are preserved when passing to an affine transform, for instance:

Lemma 2.1. *S has a supremum in $\mathbb{Q}\Gamma$ if and only if $\alpha + nS$ does.*

Definition 2.2. We say that S is *jammed* (*in* Γ) if $S \neq \varnothing$ does not have a greatest element and for every nontrivial convex subgroup $\Delta \neq \{0\}$ of Γ, there is $\gamma_0 \in S$ such that for every $\gamma_1 \in S^{>\gamma_0}$, $\gamma_1 - \gamma_0 \in \Delta$.

Example 2.3. Suppose $\Gamma \neq \{0\}$ is such that $\Gamma^{>}$ does not have a least element. Then $S := \Gamma^{<\beta}$ is jammed for every $\beta \in \Gamma$. In particular, $\Gamma^{<}$ is jammed.

Most $\Gamma \neq \{0\}$ we will deal with are either divisible or else $[\Gamma^{\neq}]$ does not have a least element and so Example 2.3 will provide a large collection of jammed subsets for such Γ. Of course, not all jammed sets are of the form $S^{\downarrow} = \Gamma^{<\beta}$.

Whether or not S is jammed in Γ depends on the archimedean classes of Γ in the following way:

Lemma 2.4. *Let Γ_1 be an ordered abelian group extension of Γ such that $[\Gamma^{\neq}]$ is coinitial in $[\Gamma_1^{\neq}]$. Then S is jammed in Γ if and only if S is jammed in Γ_1.*

Being jammed is also preserved by affine transforms:

Lemma 2.5. *S is jammed if and only if $\alpha + nS$ is jammed.*

Proof. ($\Longrightarrow$) Let Δ be a nontrivial convex subgroup of Γ. Let $\gamma_0 \in S$ be such that for every $\gamma_1 \in S^{>\gamma_0}$, $\gamma_1 - \gamma_0 \in \Delta$. Consider the element $\delta_0 := \alpha + n\gamma_0 \in \alpha + nS$. Let $\delta_1 \in (\alpha+nS)^{>\delta_0}$. Then $\delta_1 = \alpha + n\gamma_1$ for some $\gamma_1 \in S^{>\gamma_0}$ and $\delta_1 - \delta_0 = n(\gamma_1 - \gamma_0) \in \Delta$. We conclude that $\alpha + nS$ is jammed.

($\Longleftarrow$) Let Δ be a nontrivial convex subgroup of Γ. Let $\delta_0 = \alpha + n\gamma_0 \in \alpha + nS$ be such that $\delta_1 - \delta_0 \in \Delta$ for all $\delta_1 \in (\alpha + nS)^{>\delta_0}$. Then for $\gamma_1 \in S^{>\gamma_0}$ we have $\delta_1 := \alpha + n\gamma_1 \in (\alpha + nS)^{>\delta_0}$ and so $\delta_1 - \delta_0 = n(\gamma_1 - \gamma_0) \in \Delta$. As Δ is convex, it follows that $\gamma_1 - \gamma_0 \in \Delta$. We conclude that S is jammed. $\square$

Whether or not S is jammed depends only on the downward closure $S^{\downarrow}$ of S:

Lemma 2.6. *S is jammed if and only if $S^{\downarrow}$ is jammed.*

3. Asymptotic couples

An *asymptotic couple* is a pair (Γ, ψ) where Γ is an ordered abelian group and $\psi : \Gamma^{\neq} \to \Gamma$ satisfies for all $\alpha, \beta \in \Gamma^{\neq}$,

(AC1) $\alpha + \beta \neq 0 \implies \psi(\alpha + \beta) \geq \min(\psi(\alpha), \psi(\beta))$;

(AC2) $\psi(k\alpha) = \psi(\alpha)$ for all $k \in \mathbb{Z}^{\neq}$, in particular, $\psi(-\alpha) = \psi(\alpha)$;

(AC3) $\alpha > 0 \implies \alpha + \psi(\alpha) > \psi(\beta)$.

If in addition for all $\alpha, \beta \in \Gamma$,

$$0 < \alpha \leq \beta \implies \psi(\alpha) \geq \psi(\beta), \tag{HC}$$

then (Γ, ψ) is said to be of *H-type*, or to be an *H-asymptotic couple*.

By convention we extend ψ to all of Γ by setting $\psi(0) := \infty$. Then

$$\psi(\alpha + \beta) \geq \min(\psi(\alpha), \psi(\beta))$$

holds for all $\alpha, \beta \in \Gamma$, and we construe $\psi : \Gamma \to \Gamma_\infty$ as a (non-surjective) valuation on the abelian group Γ. If (Γ, ψ) is of H-type, then this valuation is convex in the sense of [ADH 2017, §2.4].

For $\alpha \in \Gamma^{\neq}$ we shall also use the following notation:

$$\alpha^{\dagger} := \psi(\alpha), \quad \alpha' := \alpha + \psi(\alpha).$$

The following subsets of Γ play special roles:

$$(\Gamma^{\neq})' := \{\gamma' : \gamma \in \Gamma^{\neq}\}, \qquad (\Gamma^{>})' := \{\gamma' : \gamma \in \Gamma^{>}\},$$
$$\Psi := \psi(\Gamma^{\neq}) = \{\gamma^{\dagger} : \gamma \in \Gamma^{\neq}\} = \{\gamma^{\dagger} : \gamma \in \Gamma^{>}\}.$$

Note that by (AC3) we have $\Psi < (\Gamma^{>})'$. It is also the case that $(\Gamma^{<})' < (\Gamma^{>})'$:

ADH 3.1. *The map $\gamma \mapsto \gamma' = \gamma + \psi(\gamma) : \Gamma^{\neq} \to \Gamma$ is strictly increasing. In particular*:

(1) $(\Gamma^{<})' < (\Gamma^{>})'$, *and*

(2) *for $\beta \in \Gamma$ there is at most one $\alpha \in \Gamma^{\neq}$ such that $\alpha' = \beta$.*

Proof. This follows from [ADH 2017, 6.5.4(iii)]. □

ADH 3.2 [ADH 2017, 9.2.4]. *There is at most one β such that*

$$\Psi < \beta < (\Gamma^{>})'.$$

If Ψ has a largest element, there is no such β.

Definition 3.3. Let (Γ, ψ) be an asymptotic couple. If $\Gamma = (\Gamma^{\neq})'$, then we say that (Γ, ψ) has *asymptotic integration*. If there is $\beta \in \Gamma$ as in ADH 3.2, then we say that β is a *gap* in (Γ, ψ) and that (Γ, ψ) *has a gap*. Finally, we call (Γ, ψ) *grounded* if Ψ has a largest element, and *ungrounded* otherwise.

The notions of asymptotic integration, gaps and being grounded form an important trichotomy for H-asymptotic couples:

ADH 3.4 [ADH 2017, 9.2.16]. *Let (Γ, ψ) be an H-asymptotic couple. Then exactly one of the following is true*:

(1) *(Γ, ψ) has a gap, in particular, $\Gamma \setminus (\Gamma^{\neq})' = \{\beta\}$ where β is a gap in Γ;*

(2) *(Γ, ψ) is grounded, in particular, $\Gamma \setminus (\Gamma^{\neq})' = \{\max \Psi\}$;*

(3) *(Γ, ψ) has asymptotic integration.*

Remark 3.5. Gaps in H-asymptotic couples are the fundamental source of deviant behavior we wish to avoid. If β is a gap in an H-asymptotic couple (Γ, ψ), then there is no $\alpha \in \Gamma$ such that $\alpha' = \beta$, or in other words, β cannot be asymptotically integrated. This presents us with an *irreversible choice*: if we wish to adjoin to (Γ, ψ) an asymptotic integral for β, then we have to choose once and for all if that asymptotic integral will be positive or negative. This phenomenon is referred to as the *fork in the road* and is the primary cause of H-fields having two nonisomorphic Liouville closures, as we shall see in Section 12 below. Gaps also prove to be a main obstruction in the model theory of asymptotic couples. For more on this, see [Aschenbrenner and van den Dries 2000] and [Gehret 2017a].

Definition 3.6 (The Divisible Hull). Given an asymptotic couple (Γ, ψ), ψ extends uniquely to a map $(\mathbb{Q}\Gamma)^{\neq} \to \mathbb{Q}\Gamma$, also denoted by ψ, such that $(\mathbb{Q}\Gamma, \psi)$ is an asymptotic couple. We call $(\mathbb{Q}\Gamma, \psi)$ the *divisible hull* of (Γ, ψ). Here are some basic facts about the divisible hull:

(1) $\psi((\mathbb{Q}\Gamma)^{\neq}) = \Psi = \psi(\Gamma^{\neq})$;

(2) if (Γ, ψ) is of H-type, then so is $(\mathbb{Q}\Gamma, \psi)$;

(3) if (Γ, ψ) is grounded, then so is $(\mathbb{Q}\Gamma, \psi)$;

(4) if $\beta \in \Gamma$ is a gap in (Γ, ψ), then it is a gap in $(\mathbb{Q}\Gamma, \psi)$;

(5) $(\Gamma^{\neq})' = ((\mathbb{Q}\Gamma)^{\neq})' \cap \Gamma$.

For proofs of these facts, see [ADH 2017, §6.5 and 9.2.8]. We say (Γ, ψ) has *rational asymptotic integration* if $(\mathbb{Q}\Gamma, \psi)$ has asymptotic integration.

In the rest of this section (Γ, ψ) *is an* H*-asymptotic couple with asymptotic integration and we let* α, β, γ *range over* Γ.

Definition 3.7. For $\alpha \in \Gamma$ we let $\int\alpha$ denote the unique element $\beta \in \Gamma^{\neq}$ such that $\beta' = \alpha$ and we call $\beta = \int\alpha$ the *integral* of α. This gives us a function $\int : \Gamma \to \Gamma^{\neq}$ which is the inverse of $\gamma \mapsto \gamma' : \Gamma^{\neq} \to \Gamma$. We define the *successor function* $s : \Gamma \to \Psi$ by $\alpha \mapsto \psi\big(\int\alpha\big)$. Finally, we define the *contraction map* $\chi : \Gamma^{\neq} \to \Gamma^{<}$ by $\alpha \mapsto \int \psi(\alpha)$. We extend χ to a function $\Gamma \to \Gamma^{\leq}$ by setting $\chi(0) := 0$.

The successor function gets its name from its behavior on $\psi(\Gamma_{\log}^{\neq})$ in Example 3.8 below (see [Gehret 2017a]). The contraction map gets its name from the way it contracts archimedean classes in the sense of Lemma 3.9(5) below. Contraction maps originate from the study of precontraction groups and ordered exponential fields (see [Kuhlmann 1994; 1995; 2000]).

Example 3.8 (The asymptotic couple of $\mathbb{T}_{\log}$). Define the abelian group $\Gamma_{\log} := \bigoplus_n \mathbb{R}e_n$, equipped with the unique ordering such that $e_n > 0$ for all n, and $[e_m] > [e_n]$ whenever $m < n$. It is convenient to think of an element $\sum r_i e_i$ of $\Gamma_{\log}$ as the vector $(r_0, r_1, r_2, \ldots)$. Next, we define the map $\psi : \Gamma_{\log}^{\neq} \to \Gamma_{\log}$ by

$$(\underbrace{0, \ldots, 0}_{n}, \underbrace{r_n}_{\neq 0}, r_{n+1}, \ldots) \mapsto (\underbrace{1, \ldots, 1}_{n+1}, 0, 0, \ldots).$$

It is easy to verify that $(\Gamma_{\log}, \psi)$ is an H-asymptotic couple with rational asymptotic integration. Furthermore, the functions $\int$, s, and χ are given by the following formulas:

(1) (***integral***) For $\alpha = (r_0, r_1, r_2, \ldots) \in \Gamma_{\log}$, take the unique n such that $r_n \neq 1$ and $r_m = 1$ for $m < n$. Then the formula for $\alpha \mapsto \int\alpha$ is given as follows:

$$\alpha = (\underbrace{1,\ldots,1}_{n}, \underbrace{r_n}_{\neq 1}, r_{n+1}, r_{n+2}, \ldots) \mapsto \textstyle\int\alpha = (\underbrace{0,\ldots,0}_{n}, r_n - 1, r_{n+1}, r_{n+2}, \ldots) : \Gamma_{\log} \to \Gamma_{\log}^{\neq}.$$

(2) (***successor***) For $\alpha = (r_0, r_1, r_2, \ldots) \in \Gamma_{\log}$, take the unique n such that $r_n \neq 1$ and $r_m = 1$ for $m < n$. Then the formula for $\alpha \mapsto s(\alpha)$ is given as follows:

$$\alpha = (\underbrace{1,\ldots,1}_{n}, \underbrace{r_n}_{\neq 1}, r_{n+1}, r_{n+1}, \ldots) \mapsto s(\alpha) = (\underbrace{1,\ldots,1}_{n+1}, 0, 0, \ldots) : \Gamma_{\log} \to \Psi_{\log} \subseteq \Gamma_{\log}.$$

(3) (***contraction***) If $\alpha = 0$, then $\chi(\alpha) = 0$. Otherwise, for $\alpha = (r_0, r_1, r_2, \ldots) \in \Gamma_{\log}^{\neq}$, take the unique n such that $r_n \neq 0$ and $r_m = 0$ for $m < n$. Then the formula for $\alpha \mapsto \chi(\alpha)$ is given as follows:

$$\alpha = (\underbrace{0, \ldots, 0}_{n}, \underbrace{r_n}_{\neq 0}, r_{n+1}, \ldots) \mapsto \chi(\alpha) = (\underbrace{0, \ldots, 0}_{n+1}, -1, 0, 0, \ldots) : \Gamma_{\log} \to \Gamma_{\log}^{\leq}.$$

For more on this example, see [Gehret 2017b; 2017a].

Lemma 3.9. *For all $\alpha, \beta \in \Gamma$ and $\gamma \in \Gamma^{\neq}$:*

(1) (***integral identity***) $\int\alpha = \alpha - s\alpha$.

(2) (***successor identity***) *If $s\alpha < s\beta$, then $\psi(\beta - \alpha) = s\alpha$.*

(3) (***fixed point identity***) *$\beta = \psi(\alpha - \beta)$ if and only if $\beta = s\alpha$.*

(4) $s\alpha < s^2\alpha$.

(5) $[\chi(\gamma)] < [\gamma]$.

(6) $\alpha \neq \beta \implies [\chi(\alpha) - \chi(\beta)] < [\alpha - \beta]$.

(7) $\alpha < \beta \implies \alpha - \chi(\alpha) < \beta - \chi(\beta)$.

Proof. For (1)–(4) we direct the reader to [Gehret 2017a]. (1) is Lemma 3.2 there, (2) is Lemma 3.4 there, (3) is Lemma 3.7 there, (4) is Lemma 3.3 there, and (5) and (6) follow easily from [ADH 2017, 9.2.18 (iii,iv)]. (7) follows from (6). □

Lemma 3.10. *Suppose $\alpha \in (\Gamma^{<})'$ and $n \geq 1$. Then $\alpha + (n+1)(s\alpha - \alpha) \in (\Gamma^{>})'$.*

Proof. Suppose $\alpha \in (\Gamma^{<})'$. Then we have

$$\begin{aligned}
\alpha + (n+1)(s\alpha - \alpha) &= s\alpha + ns\alpha - n\alpha \\
&= \psi\big(\textstyle\int\alpha\big) + n\psi\big(\int\alpha\big) - n\big(\int\alpha\big)' \\
&= \psi\big(\textstyle\int\alpha\big) + n\psi\big(\int\alpha\big) - n\big(\int\alpha\big) - n\psi\big(\int\alpha\big) \\
&= \psi\big(\textstyle\int\alpha\big) - n\int\alpha \\
&= \big(-n\textstyle\int\alpha\big)' \in (\Gamma^{>})'.
\end{aligned}$$

The last part follows because $\alpha \in (\Gamma^{<})'$ if and only if $\int\alpha \in \Gamma^{<}$ if and only if $-n\int\alpha \in \Gamma^{>}$ if and only if $\left(-n\int\alpha\right)' \in (\Gamma^{>})'$. $\square$

Lemma 3.11. *The sets Ψ and $\Psi^{\downarrow}$ are jammed.*

Proof. By Lemma 2.6, it suffices to show that $\Psi^{\downarrow} = (\Gamma^{<})'$ is jammed. By asymptotic integration and ADH 3.4, $(\Gamma^{<})'$ is nonempty and does not have a largest element. Let Δ be a nontrivial convex subgroup of Γ. Take $\delta \in \Delta^{>}$ and set $\gamma_0 := (-\delta)' \in (\Gamma^{<})'$. Then

$$\gamma_0 + 2\delta = \gamma_0 + 2\left(-\int(-\delta)'\right) = \gamma_0 + 2\left(-\int\gamma_0\right) = \gamma_0 + 2(s\gamma_0 - \gamma_0),$$

where the last equality follows from Lemma 3.9(1). Thus $\gamma_0 + 2\delta \in (\Gamma^{>})'$ by Lemma 3.10. In particular, for every $\gamma_1 \in ((\Gamma^{<})')^{>\gamma_0}$, $\gamma_1 - \gamma_0 < 2\delta \in \Delta$. We conclude that $(\Gamma^{<})'$ is jammed. $\square$

Calculation 3.12. Suppose $\gamma \neq 0$. Then

$$\int\left(\gamma' - \int s\gamma'\right) = \gamma + (s\gamma^{\dagger} - \gamma^{\dagger}) = \gamma - \chi(\gamma).$$

Proof. We begin by showing that

$$s(\gamma + s\gamma^{\dagger}) = \gamma^{\dagger}. \tag{A}$$

By (2) and (4) of Lemma 3.9 we have

$$\psi(-\gamma) = \gamma^{\dagger} < s\gamma^{\dagger} = \psi(\gamma^{\dagger} - s\gamma^{\dagger}),$$

which implies $\psi(\gamma^{\dagger} - \gamma - s\gamma^{\dagger}) = \gamma^{\dagger}$. Now (A) follows by Lemma 3.9(3).

We now proceed with our main calculation. The first and second equalities below come from Lemma 3.9(1); the third from the definitions of s and $'$, and the last from (A).

$$\begin{aligned}
\int\left(\gamma' - \int s\gamma'\right) &= \left(\gamma' - \int s\gamma'\right) - s\left(\gamma' - \int s\gamma'\right) \\
&= (\gamma' - s\gamma' + s^2\gamma') - s(\gamma' - s\gamma' + s^2\gamma') \\
&= (\gamma + \gamma^{\dagger} - \gamma^{\dagger} + s\gamma^{\dagger}) - s(\gamma + \gamma^{\dagger} - \gamma^{\dagger} + s\gamma^{\dagger}) \\
&= \gamma + s\gamma^{\dagger} - s(\gamma + s\gamma^{\dagger}) \\
&= \gamma + (s\gamma^{\dagger} - \gamma^{\dagger})
\end{aligned}$$

Finally, note that $-\chi(\gamma) = s\gamma^{\dagger} - \gamma^{\dagger}$ follows from applying Lemma 3.9(1) to $\gamma^{\dagger}$ and the definition of χ. $\square$

Lemma 3.13. *Let $\gamma \in (\Gamma^{>})'$. Then*

$$\int\gamma > -\int s\gamma = -\chi\int\gamma > 0.$$

Furthermore, if $\gamma_0, \gamma_1 \in (\Gamma^{>})'$, then $\gamma_0 \leq \gamma_1$ implies $-\int s\gamma_0 \leq -\int s\gamma_1$.

Proof. We have $s\gamma \in (\Gamma^{<})'$ which implies that $-\int s\gamma > 0$, which gives the second part of the first inequality. For the first part we note that

$$\int\gamma > -\int s\gamma \iff \int\gamma + \int s\gamma > 0 \iff \int\gamma + \chi\int\gamma > 0,$$

this last equivalence being true because $\int\gamma > 0$ and $\left[\chi\int\gamma\right] < \left[\int\gamma\right]$ by Lemma 3.9(5).

For the second inequality, we have

$$\begin{aligned} \gamma_0 \leq \gamma_1 &\implies s\gamma_0 \geq s\gamma_1 && \text{since } \gamma_0, \gamma_1 \in (\Gamma^{>})' \\ &\iff \int s\gamma_0 \geq \int s\gamma_1 && \text{by ADH 3.1} \\ &\iff -\int s\gamma_0 \leq -\int s\gamma_1. \end{aligned}$$

□

Definition 3.14. Let S be a nonempty convex subset of Γ without a greatest element. We say that S has the *yardstick property* if there is $\beta \in S$ such that for every $\gamma \in S^{>\beta}$, $\gamma - \chi(\gamma) \in S$.

Note that if S is a nonempty convex subset of Γ without a greatest element, then S has the yardstick property if and only if $S^{\downarrow}$ has the yardstick property. The following is immediate from Lemma 3.9(7):

Lemma 3.15. *Suppose S is a nonempty convex subset of Γ without a greatest element with the yardstick property. Then for every $\gamma \in S$, $\gamma - \chi(\gamma) \in S$.*

Remark 3.16. The yardstick property says that if you have an element $\gamma \in S$, then you can travel up the set S to a larger element $\gamma - \chi(\gamma)$ in a "measurable" way, i.e., you can increase upwards at least a distance of $-\chi(\gamma)$ and still remain in S. Similar to the property *jammed* from Section 2, this is a qualitative property concerning the top of the set S. Unlike *jammed*, the yardstick property requires the asymptotic couple structure of (Γ, ψ), and the contraction map χ in particular.

The yardstick property and being jammed are incompatible properties, except in the following case:

Lemma 3.17. *Let S be a nonempty convex subset of Γ without a greatest element with the yardstick property. Then S is jammed if and only if $S^{\downarrow} = \Gamma^{<}$.*

Proof. If $S = \Gamma^{<}$, then S is jammed. Now suppose that $S \neq \Gamma^{<}$. We must show that S is not jammed. In the first case, suppose $S \cap \Gamma^{>} \neq \varnothing$ and take $\gamma \in S \cap \Gamma^{>}$. Let Δ be a nontrivial convex subgroup of Γ such that $[\Delta] < [\chi(\gamma)]$. Now let $\gamma_0, \gamma_1 \in S$ be such that $\gamma < \gamma_0 < \gamma_0 - \chi(\gamma_0) < \gamma_1$. Note that

$$\gamma_1 - \gamma_0 > -\chi(\gamma_0) \geq -\chi(\gamma) > \Delta,$$

and we conclude that S is not jammed since $\gamma_0 > \gamma$ was arbitrary.

Next, suppose there is β such that $S < \beta < 0$. Let Δ be a nontrivial convex subgroup of Γ such that $[\beta] > [\chi(\beta)] > [\Delta]$. Let $\gamma \in S$ be arbitrary. Then $\gamma - \chi(\gamma) \in S$. Note that

$$(\gamma - \chi(\gamma)) - \gamma = -\chi(\gamma) \geq -\chi(\beta) > \Delta.$$

We conclude that S is not jammed since γ was arbitrary. $\square$

The following technical variant of the yardstick property will come in handy in Sections 7, 8, and 9:

Definition 3.18. Let $S \subseteq \Gamma$ be a nonempty convex set without a greatest element such that either $S \subseteq (\Gamma^{>})'$ or $S \subseteq (\Gamma^{<})'$. We say that S has the *derived yardstick property* if there is $\beta \in S$ such that for every $\gamma \in S^{>\beta}$,

$$\gamma - {\textstyle\int} s\gamma \in S^{>\beta}.$$

Proposition 3.19. *Suppose $S \subseteq \Gamma$ is a nonempty convex set without a greatest element such that either $S \subseteq (\Gamma^{>})'$ or $S \subseteq (\Gamma^{<})'$ and S has the derived yardstick property. Then $\int S := \{\int s : s \in S\} \subseteq \Gamma$ is nonempty*, *convex*, *does not have a greatest element*, *and has the yardstick property.*

Proof. By ADH 3.1, $\int S$ is nonempty, convex, and does not have a greatest element. Let $\beta \in S$ be such that for every $\gamma \in S^{>\beta}$, $\gamma - \int s\gamma \in S$. Now take $\gamma \in (\int S)^{>\int\beta}$. Then $\gamma' \in S^{>\beta}$, so $\gamma' - \int s\gamma' \in S^{>\beta}$. Thus

$$\textstyle\int\left(\gamma' - \int s\gamma'\right) \in \left(\int S\right)^{>\int\beta}.$$

By Calculation 3.12,

$$\gamma - \chi(\gamma) \in \left({\textstyle\int} S\right)^{>\int\beta}.$$

We conclude that $\int S$ has the yardstick property. $\square$

Example 3.20. (The yardstick property in $(\Gamma_{\log}, \psi)$) To get a feel for what the yardstick property says, suppose $S \subseteq \Gamma_{\log}$ is nonempty, downward closed, and has the yardstick property. Then, given an element $\alpha \neq 0$ in S we may write

$$\alpha = (\underbrace{0, \ldots, 0}_{n}, \underbrace{r_n}_{\neq 0}, r_{n+1}, \ldots),$$

and then the yardstick property says that the following larger element is also in S:

$$\alpha - \chi(\alpha) = (\underbrace{0,\ldots,0}_{n}, r_n, r_{n+1}) - (\underbrace{0,\ldots,0}_{n+1}, -1, 0, 0, \ldots) = (\underbrace{0,\ldots,0}_{n}, r_n, r_{n+1}+1, \ldots) \in S.$$

In fact, by iterating the yardstick property, we find that for *any* m, the following element is in S:

$$(\underbrace{0, \ldots, 0}_{n}, r_n, r_{n+1} + m, \ldots) \in S$$

Thus if Δ is the convex subgroup generated by $-\chi(\alpha)$, it follows that $\alpha + \Delta \subseteq S$.

4. Valued fields

In this section K is a valued field. Let $\mathcal{O}_K$ denote its valuation ring, $\mathcal{o}_K$ the maximal ideal of $\mathcal{O}_K$, $v : K^\times \to \Gamma_K := v(K^\times)$ its valuation with value group Γ_K, and $\operatorname{res} : \mathcal{O}_K \to \boldsymbol{k}_K := \mathcal{O}_K/\mathcal{o}_K$ its residue map with residue field $\boldsymbol{k}_K$, which we may also denote as $\operatorname{res}(K)$. We will suppress the subscript K when the valued field K is clear from context. By convention we extend v to a map $v : K \to \Gamma_\infty$ by setting $v(0) := \infty$.

Given $f, g \in K$ we have the following relations:

$$f \preccurlyeq g :\Longleftrightarrow vf \geq vg \quad (f \text{ is } \textit{dominated} \text{ by } g)$$
$$f \prec g :\Longleftrightarrow vf > vg \quad (f \text{ is } \textit{strictly dominated} \text{ by } g)$$
$$f \asymp g :\Longleftrightarrow vf = vg \quad (f \text{ is } \textit{asymptotic} \text{ to } g)$$

For $f, g \in K^\times$, we have the additional relation:

$$f \sim g :\Longleftrightarrow v(f-g) > vf \quad (f \text{ and } g \text{ are } \textit{equivalent})$$

Both $\asymp$ and $\sim$ are equivalence relations on K and $K^\times$, respectively. We shall also use the following notation:

$$K^{\prec 1} := \{f \in K : f \prec 1\} = \mathcal{o}_K$$
$$K^{\preccurlyeq 1} := \{f \in K : f \preccurlyeq 1\} = \mathcal{O}_K$$
$$K^{\succ 1} := \{f \in K : f \succ 1\} = K \setminus \mathcal{O}_K$$

Pseudocauchy sequences and a Kaplansky lemma. Let (a_ρ) be a well-indexed sequence in K, and $a \in K$. Then (a_ρ) is said to *pseudoconverge to a* (written $a_\rho \rightsquigarrow a$) if for some index ρ_0 we have $a - a_\sigma \prec a - a_\rho$ whenever $\sigma > \rho > \rho_0$. In this case we also say that a is a *pseudolimit of* (a_ρ). We say that (a_ρ) is a *pseudocauchy sequence in K* (or *pc-sequence in K*) if for some index ρ_0 we have

$$\tau > \sigma > \rho > \rho_0 \implies a_\tau - a_\sigma \prec a_\sigma - a_\rho.$$

If $a_\rho \rightsquigarrow a$, then (a_ρ) is necessarily a pc-sequence in K. A pc-sequence (a_ρ) is *divergent in K* if (a_ρ) does not have a pseudolimit in K.

Suppose that (a_ρ) is a pc-sequence in K and $a \in K$ is such that $a_\rho \rightsquigarrow a$. Also let $\gamma_\rho := v(a - a_\rho) \in \Gamma_\infty$, which is eventually in Γ and strictly increasing as a function of ρ. Recall *Kaplansky's Lemma*:

ADH 4.1 [ADH 2017, Prop. 3.2.1]. *Suppose $P \in K[X] \setminus K$. Then $P(a_\rho) \rightsquigarrow P(a)$. Furthermore, there are $\alpha \in \Gamma$ and $i \geq 1$ such that eventually $v(P(a_\rho) - P(a)) = \alpha + i\gamma_\rho$.*

Note that ADH 4.1 concerns *polynomials* $P \in K[X]$. Below we give a version for rational functions, but first a few remarks.

Roughly speaking, we think of the eventual nature of the sequence (γ_ρ) as a "rate of convergence" for the pseudoconvergence $a_\rho \rightsquigarrow a$. ADH 4.1 tells us that the rate of convergence for $P(a_\rho) \rightsquigarrow P(a)$ is very similar to that of $a_\rho \rightsquigarrow a$. Indeed, $(\alpha + i\gamma_\rho)$ is just an affine transform of (γ_ρ) in Γ. We want to show that applying rational functions to (a_ρ) will also have this property. Before we can do this, we need to recall a few more facts from valuation theory.

Suppose that (a_ρ) is a pc-sequence in K. A main consequence of ADH 4.1 is that (a_ρ) falls into one of two categories:

(1) (a_ρ) is of *algebraic type over* K if for *some* nonconstant $P \in K[X]$, $v(P(a_\rho))$ is eventually strictly increasing (equivalently, $P(a_\rho) \rightsquigarrow 0$).

(2) (a_ρ) is of *transcendental type over* K if for *all* nonconstant $P \in K[X]$, $v(P(a_\rho))$ is eventually constant (equivalently, $P(a_\rho) \not\rightsquigarrow 0$).

Suppose (a_ρ) is a pc-sequence of transcendental type over K. Then (a_ρ) is divergent in K. Moreover, if $a_\rho \rightsquigarrow b$ with b in a valued field extension of K, then b will necessarily be transcendental over K.

Now suppose that (a_ρ) is a pc-sequence in K. Take ρ_0 as in the definition of "pseudocauchy sequence" and define $\gamma_\rho := v(a_{\rho'} - a_\rho) \in \Gamma$ for $\rho' > \rho > \rho_0$; this depends only on ρ and the sequence $(\gamma_\rho)_{\rho>\rho_0}$ is strictly increasing. We define the *width* of (a_ρ) to be the following upward closed subset of Γ_∞:

$$\operatorname{width}(a_\rho) = \{\gamma \in \Gamma_\infty : \gamma > \gamma_\rho \text{ for all } \rho > \rho_0\}.$$

The width of (a_ρ) is independent of the choice of ρ_0. The following follows from various results in [ADH 2017, Chapters 2 and 3]:

ADH 4.2. *Let (a_ρ) be a divergent pc-sequence in K and let b be an element of a valued field extension of K such that $a_\rho \rightsquigarrow b$. Then for $\sigma_\rho := v(b - a_\rho) \in \Gamma_\infty$, eventually $\sigma_\rho = \gamma_\rho$ and*

$$\operatorname{width}(a_\rho) = \Gamma_\infty^{>v(b-K)} \quad \textit{and} \quad v(b-K) = \Gamma_\infty^{<\operatorname{width}(a_\rho)}$$

where $v(b-K) = \{v(b-a) : a \in K\} \subseteq \Gamma$.

Remark 4.3. Let b be an element of an immediate valued field extension of K. If $b \notin K$, then $v(b-K) \subseteq \Gamma$ is a nonempty downward closed subset of Γ without a greatest element. We think of $v(b-K)$ as encoding how well elements from K can approximate b. Below we will consider various qualitative properties of such a set $v(b-K)$ and consider what these properties say about the element b itself.

We say that pc-sequences (a_ρ) and (b_σ) in K are *equivalent* if they satisfy any of the following equivalent conditions:

(1) (a_ρ) and (b_σ) have the same pseudolimits in every valued field extension of K;

(2) (a_ρ) and (b_σ) have the same width, and have a common pseudolimit in some valued field extension of K;

(3) there are arbitrarily large ρ and σ such that for all $\rho' > \rho$ and $\sigma' > \sigma$ we have $a_{\rho'} - b_{\sigma'} \prec a_{\rho'} - a_\rho$, and there are arbitrarily large ρ and σ such that for all $\rho' > \rho$ and $\sigma' > \sigma$ we have $a_{\rho'} - b_{\sigma'} \prec b_{\sigma'} - b_\sigma$.

See [ADH 2017, 2.2.17] for details of this equivalence.

Now we assume that L is an immediate extension of K, $a \in L \setminus K$, and (a_ρ) is a pc-sequence in K of transcendental type over K such that $a_\rho \rightsquigarrow a$.

Lemma 4.4. *Let $R(X) \in K(X) \setminus K$. Then there is an index ρ_0 such that, for $\rho > \rho_0$,*

(1) *$R(a_\rho) \in K$ (that is, $R(a_\rho) \neq \infty$);*

(2) *$R(a_\rho) \rightsquigarrow R(a)$;*

(3) *$v(R(a_\rho) - R(a)) = \alpha + i\gamma_\rho$, eventually, for some $\alpha \in \Gamma$ and $i \geq 1$;*

(4) *$(\alpha + i\gamma_\rho)$ is eventually cofinal in $v(R(a) - K)$, with α and i as in* (3);

(5) *$(R(a_\rho))$ is a divergent pc-sequence in K; and*

(6) *$v(R(a) - K) = (\alpha + i v(a - K))^{\downarrow}$, with α and i as in* (3).

Proof. Let $R(X) = P(X)/Q(X)$ with $P, Q \in K[X]^{\neq}$. It is clear there exists ρ_0 such that $R(a_\rho) \in K$ for all $\rho > \rho_0$. Fix such a ρ_0 and assume $\rho > \rho_0$ for the rest of this proof.

We first consider the case that $R(X) = P(X) \in K[X] \setminus K$ is a polynomial. Then (2) and (3) follow from ADH 4.1. We will prove (5) and then (4) and (6) will follow. Assume towards a contradiction that there is $b \in K$ such that $R(a_\rho) \rightsquigarrow b$. Then $R(a_\rho) - b \rightsquigarrow 0$, so (a_ρ) is of algebraic type in view of $R(X) - b \in K[X] \setminus K$. This contradicts the assumption that (a_ρ) is a pc-sequence of transcendental type.

Next consider the case that $R(X) \in K(X) \setminus K[X]$. In particular, $Q(X) \in K[X] \setminus K$ and $Q \nmid P$. Then note that

$$
\begin{aligned}
v\left(\frac{P(a_\rho)}{Q(a_\rho)} - \frac{P(a)}{Q(a)}\right) &= v\left(\frac{P(a_\rho)Q(a) - P(a)Q(a_\rho)}{Q(a_\rho)Q(a)}\right) \\
&= v(P(a_\rho)Q(a) - P(a)Q(a_\rho)) - v(Q(a_\rho)) - v(Q(a)).
\end{aligned}
$$

The quantity $v(Q(a_\rho))$ is eventually constant since (a_ρ) is of transcendental type. Next, set $S(X) := P(X)Q(a) - P(a)Q(X) \in K(a)[X]$. Note that eventually $S(a_\rho) \neq 0$ and thus $S \neq 0$ (otherwise, the polynomial $Q(X) - (Q/P)(a)P(X)$ would be identically zero since it would have infinitely many distinct zeros, which would imply $Q \mid P$). Furthermore, $S(a) = 0$, which shows that $S \in K(a)[X] \setminus K(a)$. By ADH 4.1, it follows that $S(a_\rho) \rightsquigarrow S(a) = 0$. In particular, $v(S(a_\rho))$ is eventually strictly increasing and there are $\alpha \in \Gamma$ and $i \geq 1$ such that eventually $v(S(a_\rho)) = \alpha + i\gamma_\rho$. This shows (2) and (3).

Finally, we will prove (5) and then (4) and (6) will follow. Assume towards a contradiction that $R(a_\rho) \rightsquigarrow b$ with $b \in K$. Then

$$v\left(\frac{P(a_\rho)}{Q(a_\rho)} - b\right) = v(P(a_\rho) - bQ(a_\rho)) - v(Q(a_\rho))$$

is eventually strictly increasing. Therefore so is $v(P(a_\rho) - bQ(a_\rho))$, since $v(Q(a_\rho))$ is eventually constant. This implies that (a_ρ) is of algebraic type, a contradiction. $\square$

5. Differential fields, differential-valued fields and H-fields

Differential fields. A *differential field* is a field K of characteristic zero equipped with a derivation ∂ on K, i.e., an additive map $\partial : K \to K$ which satisfies the Leibniz identity: $\partial(ab) = \partial(a)b + a\partial(b)$ for all $a, b \in K$. For such K we identify $\mathbb{Q}$ with a subfield of K in the usual way.

Let K be a differential field. For $a \in K$, we will often denote $a' := \partial(a)$, and for $a \in K^\times$ we will denote the *logarithmic derivative* of a as $a^\dagger := a'/a = \partial(a)/a$. For $a, b \in K^\times$, note that $(ab)^\dagger = a^\dagger + b^\dagger$, in particular, $(a^k)^\dagger = ka^\dagger$ for $k \in \mathbb{Z}$. The set $\{a \in K : a' = 0\} \subseteq K$ is a subfield of K and is called the *field of constants* of K, and denoted by C_K (or just C if K is clear from the context). If $c \in C$, then $(ca)' = ca'$ for $a \in K$. If $a, b \in K^\times$, then $a^\dagger = b^\dagger$ if and only if $a = bc$ for some $c \in C^\times$.

The following is routine:

Lemma 5.1. *Let K be a differential field. Suppose that $y_0, y_1, \ell \in K$ are such that $y_0, y_1 \notin C$ and $y_i'' = \ell y_i'$ for $i = 0, 1$. Then there are $c_0, c_1 \in C$ such that $c_0 \neq 0$ and $y_1 = c_0 y_0 + c_1$.*

In this paper we are primarily concerned with algebraic extensions and simple transcendental extensions of differential fields. In these cases, we have:

ADH 5.2 [ADH 2017, 1.9.2]. *Suppose K is a differential field and L is an algebraic field extension of K. Then ∂ extends uniquely to a derivation on L.*

ADH 5.3 [ADH 2017, 1.9.4]. *Suppose K is a differential field with field extension $L = K(x)$ where $x = (x_i)_{i \in I}$ is a family in L that is algebraically independent over K. Then there is for each family $(y_i)_{i \in I}$ in L a unique extension of ∂ to a derivation on L with $\partial(x_i) = y_i$ for all $i \in I$.*

If K is a differential field and $s \in K \setminus \partial(K)$, then ADH 5.3 allows us to *adjoin an integral for s*: let $K(x)$ be a field extension of K such that x is transcendental over K. Then by ADH 5.3 there is a unique derivation on $K(x)$ extending ∂ such that $x' = s$. Likewise, if $s \in K \setminus (K^\times)^\dagger$, then we can *adjoin an exponential integral for s*: take $K(x)$ as before and by ADH 5.3 there is a unique derivation on $K(x)$

extending ∂ such that $x' = sx$, and thus $x^\dagger = s$, i.e., "$x = \exp(\int s)$". Adjoining integrals and exponential integrals are basic examples of *Liouville extensions*:

A *Liouville extension* of K is a differential field extension L of K such that C_L is algebraic over C and for each $a \in L$ there are $t_1, \dots, t_n \in L$ with $a \in K(t_1, \dots, t_n)$ and for $i = 1, \dots, n$,

(1) t_i is algebraic over $K(t_1, \dots, t_{i-1})$, or

(2) $t_i' \in K(t_1, \dots, t_{i-1})$, or

(3) $t_i \neq 0$ and $t_i^\dagger \in K(t_1, \dots, t_{i-1})$.

Valued differential fields. A *valued differential field* is a differential field K equipped with a valuation ring $\mathcal{O} \supseteq \mathbb{Q}$ of K. In particular, all valued differential fields have $\operatorname{char} \boldsymbol{k} = 0$.

An *asymptotic differential field*, or just *asymptotic field*, is a valued differential field K such that for all $f, g \in K^\times$ with $f, g \prec 1$,

$$f \prec g \iff f' \prec g'. \tag{A}$$

If K is an asymptotic field, then $C \subseteq \mathcal{O}$ and thus $v(C^\times) = \{0\}$. The following consequence of Lemma 5.1 will be used in Section 12 to obtain our main result:

Lemma 5.4. *Let K be an asymptotic field. Suppose that $y_0, y_1, \ell \in K$ are such that $y_0, y_1 \notin C$ and $y_i'' = \ell y_i'$ for $i = 0, 1$. Then $y_0 \succ 1$ if and only if $y_1 \succ 1$.*

The value group of an asymptotic field always has a natural asymptotic couple structure associated to it:

ADH 5.5 [ADH 2017, 9.1.3]. *Let K be a valued differential field. The following are equivalent*:

(1) *K is an asymptotic field*;

(2) *there is an asymptotic couple (Γ, ψ) with underlying ordered abelian group $\Gamma = v(K^\times)$ such that for all $g \in K^\times$ with $g \not\asymp 1$ we have $\psi(vg) = v(g^\dagger)$.*

If K is an asymptotic field, we call (Γ, ψ) as defined in ADH 5.5(2), the *asymptotic couple of K*.

Convention 5.6. Let L be an expansion of an asymptotic field, and P a property that an asymptotic couple may or may not have. Then "L has property P" means "the asymptotic couple of L has property P". For instance, when we say L is "of *H-type*", equivalently "is *H-asymptotic*", we mean that the asymptotic couple (Γ_L, ψ_L) of L is H-type. Likewise for the properties "asymptotic integration", "grounded", etc.

We say that an asymptotic field K is *pre-differential-valued*, or *pre*-d-*valued*, if the following holds:

$$\text{(PDV)} \qquad \text{for all } f, g \in K^\times, \qquad f \preccurlyeq 1,\ g \prec 1 \implies f' \prec g^\dagger.$$

Every ungrounded asymptotic field is pre-d-valued by [ADH 2017, 10.1.3].

Finally, we say that a pre-d-valued field K is *differential-valued*, or d-*valued*, if it satisfies one of the following three equivalent conditions:

(1) $\mathcal{O} = C + \mathcal{o}$.

(2) $\{\mathrm{res}(a) : a \in C\} = \boldsymbol{k}$.

(3) for all $f \asymp 1$ in K there exists $c \in C$ with $f \sim c$.

Suppose K is a pre-d-valued field of H-type. Define the $\mathcal{O}$-submodule

$$\mathrm{I}(K) := \{y \in K : y \preccurlyeq f' \text{ for some } f \in \mathcal{O}\}$$

of K. We say that K has *integration* if $K = \partial K$, has *exponential integration* if $K = (K^\times)^\dagger$, has *small integration* if $\mathrm{I}(K) = \partial\mathcal{o}$, and has *small exponential integration* if $\mathrm{I}(K) = (1+\mathcal{o})^\dagger$.

Lemma 5.7. *Let K be a pre*-d-*valued field of H-type with small integration. Then K is* d-*valued.*

Proof. Take $f \in K$ such that $f \asymp 1$. Then $f' \in \mathrm{I}(K) = \partial\mathcal{o}$, so we have $\varepsilon \in \mathcal{o}$ such that $f' = \varepsilon'$. Hence $f - \varepsilon = c$ with $c \in C^\times$ and thus $f \sim c$. □

Ordered valued differential fields. A *pre-H-field* is an ordered pre-d-valued field K of H-type whose ordering, valuation, and derivation interact as follows:

(PH1) the valuation ring $\mathcal{O}$ is convex with respect to the ordering;

(PH2) for all $f \in K$, if $f > \mathcal{O}$, then $f' > 0$.

An *H-field* is a pre-H-field K that is also d-valued. Any ordered differential field with the trivial valuation is a pre-H-field.

Example 5.8. Consider the field $L = \mathbb{R}(x)$ with x transcendental over $\mathbb{R}$, equipped with the unique derivation which has constant field $\mathbb{R}$ and $x' = 1$. Furthermore, equip L with the trivial valuation and the unique field ordering determined by requiring $x > \mathbb{R}$. It follows that L is a pre-H-field with residue field isomorphic to $\mathbb{R}(x)$. However, L is not an H-field. Indeed, the residue field is not even algebraic over the image of the constant field $\mathbb{R}$ under the residue map.

Example 5.9. Consider the Hardy field $\mathbb{Q}$. Using [Rosenlicht 1983a, Theorem 2] twice, we can extend to the Hardy field $\mathbb{Q}(x)$ where $x' = 1$, and further extend to the Hardy field $K = \mathbb{Q}(x, \arctan(x))$ where $(\arctan(x))' = 1/(1+x^2)$. Each of these three Hardy fields are pre-H-fields (see [ADH 2017, §10.5]); however, $\mathbb{Q}$

Hardy field		value group		residue field		constant field	H-field?
$\mathbb{Q}$		$\{0\}$		$\mathbb{Q}$		$\mathbb{Q}$	Yes
$\mathbb{Q}(x)$		$\mathbb{Z}v(x)$		$\mathbb{Q}$		$\mathbb{Q}$	Yes
$K = \mathbb{Q}(x, \arctan(x))$	(I)	$\mathbb{Z}v(x)$	(I)	$\mathbb{Q}(\pi)$	(II)	$\mathbb{Q}$	No

and $\mathbb{Q}(x)$ are H-fields, whereas K is *not* an H-field: the constant field of K is $\mathbb{Q}$ whereas the residue field of K is $\mathbb{Q}(\pi)$. Note that in this example the residue field $\mathbb{Q}(\pi)$ is also not algebraic over the image of the constant field $\mathbb{Q}$. For details of these Hardy field extensions and justification of the claims about K, see the table and the following discussion:

(I) Note that $\lim_{x\to\infty} \arctan(x) = \pi/2$, hence $\arctan(x) \preccurlyeq 1$ and the residue field $\operatorname{res}(K)$ of K contains $\mathbb{Q}(\pi)$. Recall that by the Lindemann–Weierstrass theorem [Lindemann 1882], π is transcendental over $\mathbb{Q}$, so $\operatorname{res}(\arctan(x)) = \pi/2$ is transcendental over $\operatorname{res}(\mathbb{Q}(x)) = \mathbb{Q}$. It follows that $\arctan(x)$ is transcendental over $\mathbb{Q}(x)$ (otherwise $\operatorname{res}(K)$ would be algebraic over $\operatorname{res}(\mathbb{Q}(x)) = \mathbb{Q}$). By [ADH 2017, 3.1.31], it follows that $\Gamma_K = \Gamma_{\mathbb{Q}(x)} = \mathbb{Z}v(x)$, and

$$\operatorname{res}(K) = \operatorname{res}(\mathbb{Q}(x))(\operatorname{res}(\arctan(x))) = \mathbb{Q}(\pi/2) = \mathbb{Q}(\pi).$$

(II) As K is a pre-H-field, it follows that the constant field is necessarily a subfield of the residue field $\mathbb{Q}(\pi)$. A routine brute force verification shows that $1/(1+x^2) \notin \partial(\mathbb{Q}(x))$. Thus the differential ring $\mathbb{Q}(x)[\arctan(x)]$ is simple by [ADH 2017, 4.6.10] (see the same work for definitions of *differential ring* and *simple differential ring*). Furthermore, as $\mathbb{Q}(x)[\arctan(x)]$ is finitely generated as a $\mathbb{Q}(x)$-algebra, it follows that C_K is algebraic over $\mathbb{Q}$ by [ADH 2017, 4.6.12]. However, $\mathbb{Q}$ is algebraically closed in $\mathbb{Q}(\pi)$ (because π is transcendental over $\mathbb{Q}$) and so $C_K = \mathbb{Q}$.

Algebraic extensions. *In this subsection K is an asymptotic field.* We fix an algebraic field extension L of K. By ADH 5.2 we equip L with the unique derivation extending the derivation ∂ of K. By *Chevalley's Extension Theorem* [ADH 2017, 3.1.15] we equip L with a valuation extending the valuation of K. Thus L is a valued differential field extension of K. We record here several properties that are preserved in this algebraic extension:

ADH 5.10. *The valued differential field L is an asymptotic field* [ADH 2017, 9.5.3]. *Also*:

(1) *If K is of H-type, then so is L.*

(2) *If K is pre-*d*-valued, then so is L* [ADH 2017, 10.1.22].

(3) *K is grounded if and only if L is grounded.*

(1) and (3) of ADH 5.10 follow from the corresponding facts about the divisible hull of an asymptotic couple; see Definition 3.6.

Furthermore, assume that K is equipped with an ordering making it a pre-H-field, and $L|K$ is an algebraic extension of ordered differential fields.

ADH 5.11. *There is a unique convex valuation ring of L extending the valuation ring of K* [ADH 2017, 3.5.18]. *Equipped with this valuation ring, L is a pre-H-field extension of K* [ADH 2017, 10.5.4]. *Furthermore, if K is an H-field and $L = K^{\mathrm{rc}}$, a real closure of K, then L is also an H-field* [ADH 2017, 10.5.6].

6. λ-freeness

In this section K is an ungrounded H-asymptotic field with $\Gamma \neq \{0\}$.

Logarithmic sequences and λ-sequences.

Definition 6.1. A *logarithmic sequence* (*in K*) is a well-indexed sequence (ℓ_ρ) in $K^{\succ 1}$ such that

(1) $\ell'_{\rho+1} \asymp \ell^\dagger_\rho$, i.e., $v(\ell_{\rho+1}) = \chi(v\ell_\rho)$, for all ρ;

(2) $\ell_{\rho'} \prec \ell_\rho$ whenever $\rho' > \rho$;

(3) (ℓ_ρ) is coinitial in $K^{\succ 1}$: for each $f \in K^{\succ 1}$ there is an index ρ with $\ell_\rho \preccurlyeq f$.

Such sequences exist and can be constructed by transfinite recursion.

Definition 6.2. A λ*-sequence* (*in K*) is a sequence of the form $(\lambda_\rho) = (-(\ell_\rho^{\dagger\dagger}))$ where (ℓ_ρ) is a logarithmic sequence in K.

ADH 6.3 [ADH 2017, 11.5.2]. *Every λ-sequence is a pc-sequence of width* $\{\gamma \in \Gamma_\infty : \gamma > \Psi\}$.

ADH 6.4 [ADH 2017, 11.5.3]. *All λ-sequences are equivalent as pc-sequences.*

For the rest of this section we will fix in K a distinguished logarithmic sequence (ℓ_ρ) along with its corresponding λ-sequence (λ_ρ). Nothing that we will discuss depends on the choice of this λ-sequence.

λ-freeness.

ADH 6.5 [ADH 2017, 11.6.1]. *The following conditions on K are equivalent*:

(1) *(λ_ρ) has no pseudolimit in K*;

(2) *for all $s \in K$ there is $g \in K^{\succ 1}$ such that $s - g^{\dagger\dagger} \succcurlyeq g^\dagger$.*

Definition 6.6. If L is an H-asymptotic field, we say that L is λ*-free* (or has λ*-freeness*) if it is ungrounded with $\Gamma_L \neq \{0\}$, and it satisfies condition (2) in ADH 6.5.

The following is immediate from the definition of λ-freeness and is a remark made after [ADH 2017, 11.6.4]:

ADH 6.7. *Suppose L is an H-asymptotic extension of K such that Ψ is cofinal in Ψ_L. If L is λ-free, then so is K.*

ADH 6.8 [ADH 2017, 11.6.4]. *If K is a directed union of grounded asymptotic subfields, then K is λ-free.*

Lemma 6.9. *If K is a directed union of λ-free asymptotic subfields, then K is λ-free.*

Proof. This follows easily from the (2) characterization of λ-freeness. □

Algebraic extensions. Ultimately, we will show that λ-freeness is preserved under arbitrary Liouville extensions of H-fields. For the time being, we have the following results concerning λ-freeness for algebraic extensions:

ADH 6.10 [ADH 2017, 11.6.7]. *If K is λ-free, then so is its henselization K^{h}.*

ADH 6.11 [ADH 2017, 11.6.8]. *K is λ-free if and only if the algebraic closure K^{a} of K is λ-free.*

Lemma 6.12. *Suppose K is equipped with an ordering making it a pre-H-field. If K is λ-free, then so is its real closure K^{rc}.*

Proof. This follows from ADH 6.11 and ADH 6.7, using the fact that $\Psi_{K^{\mathrm{rc}}} = \Psi$. □

Big exponential integration. The "big" exponential integral extensions considered here complement the Liouville extensions considered in Section 7, Section 8, and Section 9 below. In particular, we fix an element $s \in K$ that does not have an exponential integral in K, i.e., $s \notin (K^\times)^\dagger$, and we assume that s is *bounded away* from the logarithmic derivatives in K in the sense that

$$S := \{v(s - a^\dagger) : a \in K^\times\} \subseteq \Psi^\downarrow.$$

Then under the following circumstances, λ-freeness is preserved when adjoining an exponential integral for such an s:

ADH 6.13 [ADH 2017, 11.6.12]. *Let K be λ-free and Γ be divisible, and let $f^\dagger = s$, where $f \neq 0$ lies in an H-asymptotic field extension of K. Suppose*

(1) *S does not have a largest element, or*

(2) *S has a largest element and $[\gamma + vf] \notin [\Gamma]$ for some $\gamma \in \Gamma$.*

Then $K(f)$ is λ-free.

ADH 6.14 [ADH 2017, 10.5.20 and 11.6.13]. *Suppose K is equipped with an ordering making it a real closed H-field such that $s < 0$. Let $L = K(f)$ be a field extension of K such that f is transcendental over K, equipped with the unique derivation extending the derivation of K such that $f^\dagger = s$. Then there is a unique pair consisting of a valuation of $L = K(f)$ and a field ordering on L making it a* *pre-H-field extension of K with $f > 0$. With this valuation and ordering L is an H-field and Ψ is cofinal in Ψ_L. Furthermore, if K is λ-free, then so is L.*

Gap creators. Let $s \in K$. We say that s *creates a gap over* K if vf is a gap in $K(f)$, for some element $f \neq 0$ in some H-asymptotic field extension of K with $f^\dagger = s$.

ADH 6.15 [ADH 2017, 11.6.1 and 11.6.8]. *If K is λ-free, then K has rational asymptotic integration, and no element of K creates a gap over K.*

Remark 6.16. ADH 6.15 suggests that one way to view λ-freeness is as a *gap prevention property*. How good is λ-freeness as a gap prevention property? Already the above results show that it is impossible to create a gap from algebraic extensions and certain exponential integral extensions of a λ-free field. However, we can do a little bit better than that: by our results Propositions 7.2, 8.3, and 9.3 below, it follows that λ-freeness is also safely preserved (and so gaps are prevented) when passing to much more general Liouville extensions of a λ-free field.

On the other hand, *not* being λ-free does not bode well for preventing a gap:

ADH 6.17. *Suppose K has asymptotic integration, Γ is divisible, and $\lambda_\rho \rightsquigarrow \lambda \in K$. Then $s = -\lambda$ creates a gap over K. Furthermore, for every H-asymptotic extension $K(f)$ of K such that $f^\dagger = s$, vf is a gap in $K(f)$.*

Proof. The first claim is [ADH 2017, 11.5.14] and the second claim is a remark after that. $\square$

The following will be our main method of producing gaps in Liouville extensions of H-fields in Section 12 below:

ADH 6.18. *Suppose that K is equipped with an ordering making it a real closed H-field with asymptotic integration, and $\lambda_\rho \rightsquigarrow \lambda \in K$. Let $L = K(f)$ be a field extension of K with f transcendental over K equipped with the unique derivation extending the derivation of K such that $f^\dagger = -\lambda$. Then there is a unique pair consisting of a valuation of L and a field ordering on L making it an H-field extension of K with $f > 0$. With this valuation and ordering, vf is a gap in L.*

Proof. By [ADH 2017, 11.5.13] we can apply 10.5.20 of the same work with either $-\lambda$ or λ playing the role of s, whichever one is negative. Either way, a positive exponential integral f of $-\lambda$ will be adjoined, as it is the reciprocal of a positive exponential integral of λ. Also $L = K(f)$. By ADH 6.17, vf is a gap in L. $\square$

The yardstick argument. Assume that $L = K(y)$ is an immediate H-asymptotic extension of K where y is transcendental over K. In particular, $v(y - K)$ is a nonempty downward closed subset of Γ without a greatest element.

Proposition 6.19. *Assume K is henselian and λ-free, and $v(y-K) \subseteq \Gamma$ has the yardstick property. Then $L = K(y)$ is λ-free.*

Proof. Assume towards a contradiction that L is not λ-free. Take $\lambda \in L \setminus K$ such that $\lambda_\rho \rightsquigarrow \lambda$. By ADH 6.3, ADH 4.2, and Lemma 3.11, $v(\lambda - K) = \Psi^{\downarrow}$ is jammed. Furthermore, $v(\lambda - K)$ does not have a supremum in $\mathbb{Q}\Gamma$ because K is λ-free and hence has rational asymptotic integration. By the henselian assumption and Lemma 4.4, there are $\alpha \in \Gamma$ and $n \geq 1$ such that $v(\lambda - K) = (\alpha + nv(y-K))^{\downarrow}$. Thus by Lemmas 2.6 and 2.5, $v(y-K)$ is jammed as well. Since $v(y-K)$ also has the yardstick property, by Lemma 3.17 it follows that $v(y-K) = \Gamma^{<}$. However, since $v(\lambda - K)$ does not have a supremum in $\mathbb{Q}\Gamma$, by Lemma 2.1, neither does $v(y-K)$, a contradiction. □

7. Small exponential integration

*In this section K is a henselian pre-*d*-valued field of H-type and we fix an element $s \in K \setminus (K^{\times})^{\dagger}$ such that $v(s) \in (\Gamma^{>})'$.* In particular, *K does not* have small exponential integration. Take a field extension $L = K(y)$ with y transcendental over K, equipped with the unique derivation extending the derivation of K such that $(1+y)^{\dagger} = y'/(1+y) = s$.

ADH 7.1 [ADH 2017, 10.4.3 and 10.5.18]. *There is a unique valuation of L that makes it an H-asymptotic extension of K with $y \not\asymp 1$. With this valuation L is pre-*d*-valued, and is an immediate extension of K with $y \prec 1$. Furthermore, if K is equipped with an ordering making it a pre-H-field, then there is a unique ordering on L making it a pre-H-field extension of K.*

For the rest of this section equip L with this valuation. The main result of this section is the following:

Proposition 7.2. *If K is λ-free, then so is $L = K(y)$.*

The proof of Proposition 7.2 is delayed until the end of the section. The following nonempty set will be of importance in our analysis:

$$S := \left\{ v\left(s - \frac{\varepsilon'}{1+\varepsilon}\right) : \varepsilon \in K^{\prec 1} \right\} \subseteq (\Gamma^{>})' \subseteq \Gamma_{\infty}.$$

ADH 7.3. *The set S does not have a largest element.*

Proof. This is Claim 1 in the proof of [ADH 2017, 10.4.3]. □

Lemma 7.4. *S is a downward closed subset of $(\Gamma^{>})'$; in particular, S is convex.*

Proof. Let $\varepsilon_1 \prec 1$ in K and $\alpha, \beta \in (\Gamma^{>})'$ be such that

$$\alpha < v\left(s - \frac{\varepsilon_1'}{1+\varepsilon_1}\right) = \beta.$$

Let $\delta \prec 1$ in K be such that $v(\delta') = \alpha$ and set $\varepsilon_0 := \delta + \varepsilon_1 + \delta\varepsilon_1$. Note that

$$\begin{aligned}\frac{\varepsilon_1'}{1+\varepsilon_1} - \frac{\varepsilon_0'}{1+\varepsilon_0} &= \frac{\varepsilon_1'}{1+\varepsilon_1} - (1+\delta+\varepsilon_1+\delta\varepsilon_1)^\dagger \\ &= \frac{\varepsilon_1'}{1+\varepsilon_1} - ((1+\delta)(1+\varepsilon_1))^\dagger = \frac{\varepsilon_1'}{1+\varepsilon_1} - \frac{\delta'}{1+\delta} - \frac{\varepsilon_1'}{1+\varepsilon_1} = -\frac{\delta'}{1+\delta}\end{aligned}$$

and thus

$$v\left(\frac{\varepsilon_1'}{1+\varepsilon_1} - \frac{\varepsilon_0'}{1+\varepsilon_0}\right) = v\left(\frac{\delta'}{1+\delta}\right) = \alpha.$$

Finally,

$$v\left(s - \frac{\varepsilon_0'}{1+\varepsilon_0}\right) = v\left(\left(s - \frac{\varepsilon_1'}{1+\varepsilon_1}\right) + \left(\frac{\varepsilon_1'}{1+\varepsilon_1} - \frac{\varepsilon_0'}{1+\varepsilon_0}\right)\right) = \min(\beta,\alpha) = \alpha \in S. \ \square$$

The next lemma shows that S is a transform of the positive portion of $v(y-K)$.

Lemma 7.5. $(v(y-K)^{>0})' = S$, *and equivalently* $v(y-K)^{>0} = \int S$.

Proof. ($\subseteq$) Let $\varepsilon \in K$ be such that $v(y-\varepsilon) > 0$. Then necessarily $\varepsilon \prec 1$ since $y \prec 1$ and so it suffices to prove that $(v(y-\varepsilon))' = v(y'-\varepsilon') \in S$. By (PDV) it follows that $(y-\varepsilon)' \succ \varepsilon'(y-\varepsilon)$. Thus

$$\begin{aligned} s - \frac{\varepsilon'}{1+\varepsilon} &= \frac{y'}{1+y} - \frac{\varepsilon'}{1+\varepsilon} = \frac{y'(1+\varepsilon) - \varepsilon'(1+y)}{(1+y)(1+\varepsilon)} = \frac{(1+\varepsilon)(y-\varepsilon)' - \varepsilon'(y-\varepsilon)}{(1+y)(1+\varepsilon)} \\ &\asymp (1+\varepsilon)(y-\varepsilon)' - \varepsilon'(y-\varepsilon) \asymp y' - \varepsilon'.\end{aligned}$$

We conclude that $v(y'-\varepsilon') = (v(y-\varepsilon))' \in S$.

For the ($\supseteq$) direction, suppose that $\alpha = v(s-\varepsilon'/(1+\varepsilon)) \in S$ where $\varepsilon \in K^{\prec 1}$. Then the calculation in reverse shows that $\alpha = v(y'-\varepsilon') = (v(y-\varepsilon))' \in (v(y-K)^{>0})'$. $\square$

The next lemma gives us a "definable yardstick" that we can use for going up the set S. If K has small integration, then we can obtain a longer yardstick in the sense of Lemma 3.13, however the shorter yardstick will be good enough for our purposes.

Lemma 7.6. *Suppose* $\gamma \in S$. *Then* $\gamma < \gamma - \int s\gamma \in S$. *If* $\mathrm{I}(K) = \partial\mathcal{O}$, *then* $\gamma < \gamma + \int\gamma \in S$. *Thus* S *has the derived yardstick property and so* $v(y-K)^{>0}$ *and* $v(y-K)$ *both have the yardstick property.*

Proof. Let $\gamma \in S$ and take $\varepsilon \prec 1$ in K such that $\gamma = v(s - \varepsilon'/(1+\varepsilon))$. Next take $b \prec 1$ in K such that $v(b') = (v(b))' = \gamma$ (and so $v(b) = \int\gamma$). Take $u \in K$ with $s - \varepsilon'/(1+\varepsilon) = ub'$, so $u \asymp 1$. Next let $\delta \prec 1$ be such that $(1+\varepsilon)(1+ub) = 1+\delta$.

Now note that

$$\begin{aligned} s - \frac{\delta'}{1+\delta} &= s - ((1+\varepsilon)(1+ub))^{\dagger} = s - \frac{\varepsilon'}{1+\varepsilon} - \frac{(ub)'}{1+ub} \\ &= ub' - \frac{(ub)'}{1+ub} = \frac{u^2bb' - u'b}{1+ub}. \end{aligned}$$

However, since $\Psi \ni s^2\gamma < v(u') \in \Gamma^{>\Psi}$, we have

$$v(u'b) = v(u'b'(b^{\dagger})^{-1}) = v(u') - \psi \textstyle\int \gamma + \gamma > s^2\gamma - s\gamma + \gamma = -\int s\gamma + \gamma,$$

the last step following from Lemma 3.9(1). Thus by Lemma 3.13, we have

$$v\Big(s - \frac{\delta'}{1+\delta}\Big) \geq \min(v(u^2bb'), v(u'b)) \geq \min\big(\gamma + \textstyle\int\gamma, -\int s\gamma + \gamma\big) = \gamma - \int s\gamma > \gamma.$$

Finally, by Lemma 7.4, it follows that $\gamma - \int s\gamma \in S$.

If $\mathrm{I}(K) = \partial\mathcal{O}$, then we can arrange $u = 1$ above and thus

$$s - \frac{\delta'}{1+\delta} = \frac{bb'}{1+b} \asymp bb',$$

and so $v(bb') = \gamma + \int\gamma$. The claim about $v(y-K)^{>0}$ now follows from Lemma 7.5 and Proposition 3.19. □

Proposition 7.2 now follows immediately from Lemma 7.6 and Proposition 6.19.

8. Small integration

In this section K is a henselian pre-d-*valued field of H-type and we fix an element $s \in K$ such that $v(s) \in (\Gamma^{>})'$ and $s \notin \partial\mathcal{O}$.* In particular, *$K$ does not* have small integration. Define the following nonempty set:

$$S := \{v(s - \varepsilon') : \varepsilon \in K^{\prec 1}\} \subseteq (\Gamma^{>})' \subseteq \Gamma_{\infty}.$$

As K is pre-d-valued, we have the following, which elaborates on [ADH 2017, 10.2.5(iii)]:

Lemma 8.1. *S has no largest element and is a downward closed subset of $(\Gamma^{>})'$; in particular, S is convex*

Proof. First note that $v(s) \in S$. Next take $\gamma \in S$ with $\gamma \geq v(s)$, and write $\gamma = v(s - \varepsilon')$ for some $\varepsilon \prec 1$ in K. As $\gamma \in (\Gamma^{>})'$, we take $b \prec 1$ in K such that $v(b') = \gamma$. Thus for some $u \asymp 1$ in K we have $v(s - \varepsilon' - ub') > \gamma$. By (PDV), $v(u'b) > v(b') = \gamma$ and so $v(s - \varepsilon' - (ub)') > \gamma$. This shows that S has no largest element. The claim that S is a downward closed subset of $(\Gamma^{>})'$ follows similarly from $S \subseteq (\Gamma^{>})'$. □

Take a field extension $L = K(y)$ with y transcendental over K, equipped with the unique derivation extending the derivation of K such that $y' = s$.

ADH 8.2 [ADH 2017, 10.2.4 and 10.5.8]. *There is a unique valuation of L that makes it an H-asymptotic extension of K with $y \not\asymp 1$. With this valuation L is an immediate extension of K with $y \prec 1$ and L is pre-*d*-valued. Furthermore, if K is equipped with an ordering making it a pre-H-field, then there is a unique ordering on L making it a pre-H-field extension of K.*

For the rest of this section equip L with this valuation. The main result of this section is the following:

Proposition 8.3. *If K is λ-free, then so is $L = K(y)$.*

We will delay the proof of Proposition 8.3 until the end of the section.

Lemma 8.4. $(v(y-K)^{>0})' = S$, *and equivalently* $v(y-K)^{>0} = \int S$.

Proof. ($\subseteq$) Let $\varepsilon \in K$ be such that $y-\varepsilon \prec 1$. Then necessarily $\varepsilon \prec 1$ because $y \prec 1$. Let $\alpha = v(y-\varepsilon)$. We want to show that $\alpha' \in S$. Note that because $y-\varepsilon \not\asymp 1$, we get

$$\alpha' = (v(y-\varepsilon))' = v(y'-\varepsilon') = v(s-\varepsilon') \in S.$$

For the ($\supseteq$) direction, let $\varepsilon \prec 1$ be such that $\alpha = v(s-\varepsilon')$ is an arbitrary element of S. Then by arguing as above, $v(y-\varepsilon) > 0$ and $(v(y-\varepsilon))' = \alpha$. □

Lemma 8.5. *Suppose $\gamma \in S$. Then $\gamma < \gamma - \int s\gamma \in S$. If $\mathrm{I}(K) = (1+\smallo)^\dagger$, then $\gamma < \gamma + \int \gamma \in S$. Thus S has the derived yardstick property and so $v(y-K)^{>0}$ and $v(y-K)$ both have the yardstick property.*

Proof. Suppose $\gamma \in S$ and take $\varepsilon \prec 1$ in K such that $\gamma = v(s-\varepsilon')$. As $\gamma \in (\Gamma^{>})'$, we may take $b \prec 1$ in K such that $b' \asymp s - \varepsilon'$. Thus there is $u \asymp 1$ in K such that $ub' = s - \varepsilon'$. By (PDV), it follows that $v(u') > \Psi$. Thus

$$\begin{aligned} v(s-(\varepsilon-ub)') &= v(s-\varepsilon'-ub'-ub') = v(u'b) \\ &= v(u'b'(b^\dagger)^{-1}) = v(u') - \psi \textstyle\int \gamma + \gamma \\ &> s^2\gamma - s\gamma + \gamma = -\textstyle\int s\gamma + \gamma. \end{aligned}$$

Next, assume that $(1+\smallo)^\dagger = \mathrm{I}(K)$. Since $s-\varepsilon' \in \mathrm{I}(K)$, there is $\delta \prec 1$ such that $s - \varepsilon' = (1+\delta)^\dagger$, i.e.,

$$s-\varepsilon' = \frac{\delta'}{1+\delta}.$$

Now note that

$$s-(\varepsilon+\delta)' = s-\varepsilon'-\delta' = \frac{\delta'}{1+\delta} - \delta' = \frac{-\delta'\delta}{1+\delta} \asymp \delta'\delta,$$

and so

$$S \ni v(s-(\varepsilon+\delta)') = v(\delta'\delta) = \gamma + \textstyle\int \gamma.$$

The claim about $v(y-K)^{>0}$ now follows from Lemma 8.4 and Proposition 3.19. □

Proposition 8.3 now follows immediately from Lemma 8.5 and Proposition 6.19.

9. Big integration

*In this section K is a henselian pre-*d*-valued field of H-type and we fix an element $s \in K$ such that*

$$S := \{v(s - a') : a \in K\} \subseteq (\Gamma^{<})' \subseteq \Gamma_{\infty}.$$

It will necessarily be the case that $s \notin \partial K$ and $v(s) \in (\Gamma^{<})'$.

Lemma 9.1. *S is downward closed and does not have a largest element.*

Proof. Let $\gamma = v(s - a') \in S$ for some $a \in K$. Suppose $\delta < \gamma$ in Γ. Then there is $f \in K$ such that $v(f') = \delta$ and so $\delta = v(s - (a + f)') \in S$. Next, by $S \subseteq (\Gamma^{<})'$, take $b \in K$ such that $b' \asymp s - a'$, and then take $u \asymp 1$ in K with $ub' = s - a'$. By (PDV), $u'b \prec b'$ and thus $\gamma < v(s - a' - (ub)') \in S$. □

Take a field extension $L = K(y)$ with y transcendental over K, equipped with the unique derivation extending the derivation of K such that $y' = s$.

ADH 9.2 [ADH 2017, 10.2.6 and 10.5.8]. *There is a unique valuation of L making it an H-asymptotic extension of K. With this valuation L is an immediate extension of K with $y \succ 1$ and L is pre-*d*-valued. Furthermore, if K is equipped with an ordering making it a pre-H-field, then there is a unique ordering on L making it a pre-H-field extension of K.*

For the rest of this section equip L with this valuation. The main result of this section is the following:

Proposition 9.3. *If K is λ-free, then so is $L = K(y)$.*

We will delay the proof of Proposition 9.3 until the end of the section.

Lemma 9.4. *$v(y - K)' = S$, and equivalently $v(y - K) = \int S$.*

Proof. Let $\gamma = v(y - x)$ with $x \in K$. Then $v(y' - x') = v(s - x') \in S \subseteq (\Gamma^{<})'$ and so $y - x \succ 1$. Thus $\gamma' = (v(y - x))' = v(y' - x') = v(s - x') \in S$. Conversely, if $\gamma = v(s - x') \in S$, then $\gamma = v(y' - x') = (v(y - x))'$. □

By Lemma 9.1, we fix $g \in K^{\succ 1}$ such that $g' \sim s$.

Lemma 9.5. *$S^{>v(s)}$ is cofinal in S, and*

$$S^{>v(s)} = \{v((g(1 + \varepsilon))' - s) : \varepsilon \prec 1\}.$$

Proof. $S^{>v(s)}$ is cofinal in S since $v(s) \in S$ and S does not have a largest element. Suppose $\varepsilon \prec 1$. Then by (PDV), $(g(1 + \varepsilon))' = g' + \varepsilon' g + \varepsilon g' \sim g' \sim s$ and so $(g(1 + \varepsilon))' - s \prec s$. Conversely, suppose $\gamma = v(x' - s) > vs$. Then $x' \sim s$ and so $x' \sim g'$, i.e., $x' - g' \prec g'$. As $g \succ 1$, we get $x - g \prec g$ and so $x = g(1 + \varepsilon)$ for some $\varepsilon \prec 1$. □

Lemma 9.6. *If $\gamma \in S^{>v(s)}$, then $\gamma < \gamma - \int s\gamma \in S$. Thus S has the derived yardstick property and so $v(y-K)$ has the yardstick property.*

Proof. Let $\gamma = v((g(1+\varepsilon))' - s)$ for some $\varepsilon \prec 1$. Note that

$$(g(1+\varepsilon))' - s = g' + g\varepsilon' + g'\varepsilon - s.$$

Next take $\delta \succ 1$ such that

$$\delta' \asymp g' + g\varepsilon' + g'\varepsilon - s,$$

so $v(\delta') = \gamma$, and take $u \asymp 1$ such that

$$u\delta' = g' + g\varepsilon' + g'\varepsilon - s.$$

Then $\delta' \prec g' \asymp s$ and so $\delta \prec g$, i.e., $\delta/g \prec 1$. Furthermore, $u^\dagger \prec \delta^\dagger$ implies that $u'\delta \prec u\delta'$. Now consider the following element of $S^{>v(s)}$:

$$\beta = v\Big(\Big(g\Big(1+\varepsilon - \frac{u\delta}{g}\Big)\Big)' - s\Big).$$

Note that

$$\begin{aligned}\Big(g\Big(1+\varepsilon - \frac{u\delta}{g}\Big)\Big)' - s &= (g + g\varepsilon - u\delta)' - s \\ &= g' + g\varepsilon' + g'\varepsilon - u'\delta - u\delta' - s \\ &= (g' + g\varepsilon' + g'\varepsilon - s - u\delta') - u'\delta \\ &= -u'\delta.\end{aligned}$$

Thus we can use that $v(u') > \Psi$ and $\gamma = v(\delta) + v(\delta^\dagger)$ to get the yardstick:

$$\begin{aligned}v(-u'\delta) &= v(u'(\delta^\dagger)^{-1}\delta') = v(u'(\delta^\dagger)^{-1}) + \gamma \\ &= v(u') - \psi \textstyle\int \gamma + \gamma = v(u') - s\gamma + \gamma \\ &> s^2\gamma - s\gamma + \gamma = -\textstyle\int s\gamma + \gamma.\end{aligned}$$

The claim about $v(y-K)$ now follows from Lemma 9.4 and Proposition 3.19. □

Proposition 9.3 now follows immediately from Lemma 9.6 and Proposition 6.19.

10. The differential-valued hull and H-field hull

*In this section K is a pre-*d*-valued field of H-type.*

ADH 10.1 [ADH 2017, 10.3.1]. *K has a* d*-valued extension* $\mathrm{dv}(K)$ *of H-type such that any embedding of K into any* d*-valued field L of H-type extends uniquely to an embedding of* $\mathrm{dv}(K)$ *into L.*

The d-valued field $\mathrm{dv}(K)$ as in ADH 10.1 above is called the *differential-valued hull of K*.

Theorem 10.2. *If K is λ-free, then* $\mathrm{dv}(K)$ *is λ-free.*

Proof. By iterating applications of ADH 6.10, Proposition 8.3, and Lemma 6.9, we get an immediate henselian λ-free H-asymptotic extension L of K which has small integration. By Lemma 5.7, L will also be d-valued. Thus by ADH 10.1, $\mathrm{dv}(K)$ can be identified with a subfield of L which contains K. Finally, by ADH 6.7 it follows that $\mathrm{dv}(K)$ is λ-free. □

Definition 10.3. A gap β in K is said to be a *true gap* if no $b \asymp 1$ in K satisfies $v(b') = \beta$, and is said to be a *fake gap* otherwise (that is, if there is $b \asymp 1$ in K such that $v(b') = \beta$).

Remark 10.4. Suppose K has a gap β. Then the asymptotic couple (Γ, ψ) "believes" it can make a choice about β, in the sense of Remark 3.5. However, if β is a fake gap, then this choice is completely predetermined by K itself. Indeed, if L is a d-valued extension of K of H-type and β is a fake gap, then there is $\varepsilon \in \mathcal{O}_L$ such that $v(\varepsilon') = \beta$. However, if β is a true gap, then both options of this choice are still available to K, see [ADH 2017, 10.3.2(ii), 10.2.1, and 10.2.2].

Lemma 10.5. *If K is* d*-valued and has a gap β, then β is a true gap.*

Proof. Let K be a d-valued field and consider $\beta \in \Gamma$. Suppose that there is $b \asymp 1$ in K such that $v(b') = \beta$. Then there are $c \in C^\times$ and $\varepsilon \prec 1$ in $K^\times$ such that $b = c + \varepsilon$ and thus $v(b') = v(\varepsilon') = \beta \in (\Gamma^>)'$. In particular, β is not a gap. □

Corollary 10.6. *The differential-valued hull of K has the following properties*:

(1) *If K is grounded, then* $\mathrm{dv}(K)$ *is grounded.*

(2) *If K has a fake gap, then* $\mathrm{dv}(K)$ *is grounded.*

(3) *If K has a true gap, then* $\mathrm{dv}(K)$ *has a true gap.*

(4) *If K has asymptotic integration and is not λ-free, then* $\mathrm{dv}(K)$ *has asymptotic integration and is not λ-free.*

(5) *If K is λ-free, then* $\mathrm{dv}(K)$ *is λ-free.*

Proof. (1)–(4) are a restatement of [ADH 2017, 10.3.2]. (5) is Theorem 10.2. □

The H-field hull of a pre-H-field. *In this subsection we further assume that K is equipped with an ordering making it a pre-H-field.*

ADH 10.7 [ADH 2017, 10.5.13]. *A unique field ordering on* $\mathrm{dv}(K)$ *makes* $\mathrm{dv}(K)$ *a pre-H-field extension of K. Let $H(K)$ be* $\mathrm{dv}(K)$ *equipped with this ordering. Then $H(K)$ is an H-field and embeds uniquely over K into any H-field extension of K.*

The H-field $H(K)$ in ADH 10.7 above is called the *H-field hull of K*. We have the following H-field analogues of Theorem 10.2 and Corollary 10.6:

Corollary 10.8. *If K is λ-free, then $H(K)$ is λ-free.*

Corollary 10.9. *The H-field hull of K has the following properties*:

(1) *If K is grounded, then $H(K)$ is grounded.*

(2) *If K has a fake gap, then $H(K)$ is grounded.*

(3) *If K has a true gap, then $H(K)$ has a true gap.*

(4) *If K has asymptotic integration and is not λ-free, then $H(K)$ has asymptotic integration and is not λ-free.*

(5) *If K is λ-free, then $H(K)$ is λ-free.*

11. The integration closure

In this section K is a d*-valued field of H-type with asymptotic integration.*

ADH 11.1 [ADH 2017, 10.2.7]. *K has an immediate asymptotic extension $K(\int)$ that is henselian, has integration, and embeds over K into any henselian* d*-valued H-asymptotic extension of K that has integration.*

Given any $K(\int)$ with these properties, the only henselian asymptotic subfield of $K(\int)$ containing K and having integration is $K(\int)$.

Theorem 11.2. *If K is λ-free, then so is $K(\int)$.*

Proof. By iterating Lemma 6.9, ADH 6.10, and Propositions 8.3 and 9.3, we obtain a λ-free d-valued immediate H-asymptotic extension L of K that is henselian and has integration. By ADH 11.1, $K(\int)$ can be identified with a subfield of L which contains K. Finally, by ADH 6.7, $K(\int)$ is also λ-free. □

12. The number of Liouville closures

In this section K is a pre-H-field. K is said to be *Liouville closed* if it is a real closed H-field with integration and exponential integration. A *Liouville closure* of K is a Liouville closed H-field extension of K which is also a Liouville extension of K.

Theorem 12.1. *Suppose K is an H-field. Then K has at least one and at most two Liouville closures up to isomorphism over K. In particular*,

(1) *K has exactly one Liouville closure up to isomorphism over K if and only if*
   (a) *K is grounded, or*
   (b) *K is λ-free.*

(2) *K has exactly two Liouville closures up to isomorphism over K if and only if*
   (c) *K has a gap, or*
   (d) *K has asymptotic integration and is not λ-free.*

Theorem 12.1 will follow from the following Proposition, whose proof we delay until later in the section:

Proposition 12.2. *Suppose K is an H-field.*

(1) *If K is λ-free, then K has exactly one Liouville closure up to isomorphism over K.*

(2) *If K has asymptotic integration and is not λ-free, then K has at least two Liouville closures up to isomorphism over K.*

Proof of Theorem 12.1 assuming Proposition 12.2. It is clear that K will be in case (a), (b), (c) or (d), and all four cases are mutually exclusive. If K is in case (a), then K has exactly one Liouville closure up to isomorphism over K, by [ADH 2017, 10.6.23]. If K is in case (c), then K has exactly two Liouville closures up to isomorphism over K, by [ADH 2017, 10.6.25]. Cases (b) and (d) are taken care of by Proposition 12.2 and [ADH 2017, 10.6.12]. □

In general, a pre-H-field which is not also an H-field might not have any Liouville closures at all. For instance, the pre-H-field L from Example 5.8 cannot have any Liouville closures: a Liouville closure of L would necessarily contain $H(L)$, but $H(L)$ cannot be contained inside any Liouville extension of L because $C_{H(L)}$ is not an algebraic extension of $C_L = \mathbb{R}$. In such a situation, the next best thing is to consider Liouville closures of the H-field hull:

Corollary 12.3. *$H(K)$ has at least one and at most two Liouville closures up to isomorphism over K. In particular,*

(1) *$H(K)$ has exactly one Liouville closure up to isomorphism over K if and only if*
   (a) *K is grounded, or*
   (b) *K has a fake gap, or*
   (c) *K is λ-free.*

(2) *$H(K)$ has exactly two Liouville closures up to isomorphism over K if and only if*
   (d) *K has a true gap, or*
   (e) *K has asymptotic integration and is not λ-free.*

Proof. If we replace in the statement of Corollary 12.3 all instances of "up to isomorphism over K" with "up to isomorphism over $H(K)$", then this would follow from Corollary 10.9 and Theorem 12.1. Now, to strengthen the statements to "up to isomorphism over K", use that $H(K)$ is determined up to unique isomorphism in ADH 10.7. □

Liouville towers. *In this subsection K is an H-field.* The primary method of constructing Liouville closures of an H-field is with a *Liouville tower*. A *Liouville tower on K* is a strictly increasing chain $(K_\lambda)_{\lambda \le \mu}$ of H-fields, indexed by the ordinals less than or equal to some ordinal μ, such that

(1) $K_0 = K$;

(2) if λ is a limit ordinal, $0 < \lambda \leq \mu$, then $K_\lambda = \bigcup_{\iota<\lambda} K_\iota$;

(3) for $\lambda < \lambda + 1 \leq \mu$, *either*

 (a) K_λ is not real closed and $K_{\lambda+1}$ is a real closure of K_λ,

 or K_λ is real closed, $K_{\lambda+1} = K_\lambda(y_\lambda)$ with $y_\lambda \notin K_\lambda$ (so y_λ is transcendental over K_λ), and one of the following holds, with $(\Gamma_\lambda, \psi_\lambda)$ the asymptotic couple of K_λ and $\Psi_\lambda := \psi_\lambda(\Gamma_\lambda^{\neq})$:

 (b) $y_\lambda' = s_\lambda \in K_\lambda$ with $y_\lambda \prec 1$ and $v(s_\lambda)$ is a gap in K_λ,
 (c) $y_\lambda' = s_\lambda \in K_\lambda$ with $y_\lambda \succ 1$ and $v(s_\lambda)$ is a gap in K_λ,
 (d) $y_\lambda' = s_\lambda \in K_\lambda$ with $v(s_\lambda) = \max \Psi_\lambda$,
 (e) $y_\lambda' = s_\lambda \in K_\lambda$ with $y_\lambda \prec 1$, $v(s_\lambda) \in (\Gamma_\lambda^>)'$, and $s_\lambda \neq \varepsilon'$ for all $\varepsilon \in K_\lambda^{\prec 1}$,
 (f) $y_\lambda' = s_\lambda \in K_\lambda$ such that $S_\lambda := \{v(s_\lambda - a') : a \in K_\lambda\} < (\Gamma_\lambda^>)'$, and S_λ has no largest element,
 (g) $y_\lambda^\dagger = s_\lambda \in K_\lambda$ with $y_\lambda \sim 1$, $v(s_\lambda) \in (\Gamma_\lambda^>)'$, and $s_\lambda \neq a^\dagger$ for all $a \in K_\lambda^\times$,
 (h) $y_\lambda^\dagger = s_\lambda \in K_\lambda^<$ with $y_\lambda > 0$, and $v(s_\lambda - a^\dagger) \in \Psi_\lambda^\downarrow$ for all $a \in K_\lambda^\times$.

The H-field K_μ is called the *top* of the tower $(K_\lambda)_{\lambda\leq\mu}$. We say that a Liouville tower $(K_\lambda)_{\lambda\leq\mu}$ is *maximal* if it cannot be extended to a Liouville tower $(K_\lambda)_{\lambda\leq\mu+1}$ on K. Given a Liouville tower $(K_\lambda)_{\lambda\leq\mu}$ on K, $0 \leq \lambda < \lambda+1 \leq \mu$, we say $K_{\lambda+1}$ is an *extension of type* $(*)$ for $(*) \in \{(a), (b), \ldots, (h)\}$ if $K_{\lambda+1}$ and K_λ satisfy the properties of item $(*)$ as in the definition of Liouville tower.

ADH 12.4. (1) *Let $(K_\lambda)_{\lambda\leq\mu}$ be a Liouville tower on K. Then*:

 (a) *K_μ is a Liouville extension of K.*
 (b) *The constant field C_μ of K_μ is a real closure of C if $\mu > 0$.*
 (c) $|K_\mu| = |K|$, *hence* $\mu < |K|^+$.

(2) *There is a maximal Liouville tower on K.*

(3) *The top of a maximal Liouville tower on K is Liouville closed, and hence a Liouville closure of K.*

(4) *If $(K_\lambda)_{\lambda\leq\mu}$ is a Liouville tower on K such that no K_λ with $\lambda < \mu$ has a gap, and if K_μ is Liouville closed, then K_μ is the unique Liouville closure of K up to isomorphism over K.*

Proof. (1) is [ADH 2017, 10.6.13], (2) follows from (1)(c), (3) is [ADH 2017, 10.6.14], and (4) is [ADH 2017, 10.6.17]. □

For a set $\Lambda \subseteq \{(a), (b), \ldots, (h)\}$ with $(a) \in \Lambda$, the definition of a Λ*-tower on K* is identical to that of a Liouville tower on K, except that in clause (3) of the above definition only the items from Λ occur. Thus every Λ-tower on K is also a Liouville tower on K. Maximal Λ-towers exist on K by Zorn's Lemma and ADH 12.4(1)(c).

Proof of Proposition 12.2. (1) Assume K is λ-free. By ADH 12.4(4), it suffices to find a Liouville tower $(K_\lambda)_{\lambda\le\mu}$ on K such that K_μ is Liouville closed and no K_λ with $\lambda<\mu$ has a gap. Take a maximal {(a),(e),(f),(g),(h)}-tower $(K_\lambda)_{\lambda\le\mu}$ on K. By Lemmas 6.9, 6.12, Propositions 7.2, 8.3, 9.3 and ADH 6.14, K_λ is λ-free for every $\lambda\le\mu$. In particular, no K_λ with $\lambda<\mu$ has a gap. Finally, by maximality, it follows that K_μ is Liouville closed.

(2) Assume that K has asymptotic integration and is not λ-free. First consider the case that K does not have rational asymptotic integration. Then $K_1=K^{\mathrm{rc}}$ has a gap. By [ADH 2017, 10.6.25] K_1 has two Liouville closures which are not isomorphic over K_1. As K_1 is a real closure of K, they are not isomorphic over K either because the real closure is unique up to unique isomorphism. Thus K has at least two Liouville closures which are not isomorphic over K.

Next, consider the case that K is real closed. In this case, if L is a Liouville closure of K, then $C_L=C$ since C is necessarily real closed. As K is not λ-free, there is $\lambda\in K$ such that $\lambda_\rho\rightsquigarrow\lambda$. Next, let $K_1=K(f)$ be the H-field extension from ADH 6.18 such that $f^\dagger=-\lambda$ and $v(f)$ is a gap in K_1. Again by [ADH 2017, 10.6.25], K_1 has two Liouville closures L_1 and L_2 which are not isomorphic over K_1. There is $\tilde{y}\in L_1^{\prec 1}$ such that $\tilde{y}'=f$ whereas every $y\in L_2$ such that $y'=f$ has the property that $y\succ 1$. Furthermore, as both L_1 and L_2 are Liouville closed, they both contain nonconstant elements y such that $y''=-\lambda y'$.

Claim. *If $y\in L_1\setminus C$ is such that $y''=-\lambda y'$, then $y\preccurlyeq 1$. If $y\in L_2\setminus C$ is such that $y''=-\lambda y'$, then $y\succ 1$.*

Proof of Claim. Suppose $y\in L_1\setminus C$ is such that $y''=-\lambda y'$. Let $\tilde{y}\in L_1^{\prec 1}$ be such that $\tilde{y}'=f$. Then $\tilde{y}\in L_1\setminus C$ since $f\neq 0$. Furthermore $\tilde{y}''=-\lambda\tilde{y}'$ so there are $c_0\in C^\times$ and $c_1\in C$ such that $y=c_0\tilde{y}+c_1$, by Lemma 5.4. It follows that $y\preccurlyeq 1$.

Next, let $y\in L_2\setminus C$ and let $\tilde{y}\in L_2$ be such that $\tilde{y}'=f$. Then $\tilde{y}\notin C$ because $\tilde{y}\succ 1$ and $\tilde{y}''=-\lambda\tilde{y}'$. As in the first case, it will follow from Lemma 5.4 that $y\succ 1$. □

It follows from the claim that L_1 and L_2 are not isomorphic over K.

Finally, consider the case that K is not real closed, and has rational asymptotic integration. By the above case, the real closure K^{rc} has two Liouville closures L_1 and L_2 which are not isomorphic over K^{rc}. These two Liouville closures will also not be isomorphic over K, as real closures are unique up to unique isomorphism. □

The next lemma concerns the appearances of gaps in arbitrary Liouville H-field extensions, not necessarily extensions occurring as the tops of Liouville towers.

Lemma 12.5. *Suppose K is grounded or is λ-free and L is a Liouville H-field extension of K. Then L does not have a gap.*

Proof. We first consider the case that K is λ-free. Let M be the Liouville closure of K which was constructed in the proof of Proposition 12.2. We claim that Ψ

is cofinal in Ψ_M. This follows from the fact that M is constructed as the top of an {(a),(e),(f),(g),(h)}-tower on K: the Ψ-set remains unchanged when passing to extensions of type (a), (e), (f) or (g) and for extensions of type (h), the original Ψ-set is cofinal in the larger Ψ-set by ADH 6.14. Finally, as M is the unique Liouville closure of K up to isomorphism over K, we may identify L with a subfield of M which contains K. Thus Ψ_L is cofinal in Ψ_M. As M is λ-free, so is L by ADH 6.7. In particular, L has rational asymptotic integration and so it does not have a gap.

We next consider the case that K is grounded. Let M be the Liouville closure of K as constructed in the proof of [ADH 2017, 10.6.24] and the remarks following it. In particular, using the notation from the remarks following that proof, we have $M = \bigcup_{n<\omega} \ell^n(K)$ where $\ell^0(K) = K$ and $\ell^{n+1}(K)$, for each n, is a grounded Liouville H-field extension of K such that $\max \Psi_{\ell^{n+1}(K)} = s(\max \Psi_{\ell^n(K)})$. Thus the set $\{s^n(\max \Psi) : n < \omega\}$ is a cofinal subset of Ψ_M. We now identify L with a subfield of M that contains K and consider two cases:

Case 1: $\{s^n(\max \Psi) : n < \omega\} \not\subseteq \Psi_L$

In this case there is a least $N < \omega$ such that $s^N(\max \Psi) \in \Psi_L$ but $s(s^N(\max \Psi)) \in \Psi_M \setminus \Psi_L$. This implies that the element $s^N(\max \Psi) \in \Psi_L$ cannot be asymptotically integrated. The only way this can happen is if $s^N(\max \Psi) = \max \Psi_L$. Thus L is grounded and does not have a gap.

Case 2: $\{s^n(\max \Psi) : n < \omega\} \subseteq \Psi_L$

In this case Ψ_L is cofinal in Ψ_M and so L is λ-free by ADH 6.7. This implies that L has rational asymptotic integration and therefore does not have a gap. $\square$

We also give a characterization of the dichotomy of Theorem 12.1 entirely in terms of gaps appearing in Liouville towers and arbitrary Liouville extensions:

Corollary 12.6. *The following are equivalent*:

(1) *K has exactly two Liouville closures up to isomorphism over K.*

(2) *There is a Liouville tower $(K_\lambda)_{\lambda\leq\mu}$ on K such that some K_λ has a gap.*

(3) *Every maximal Liouville tower $(K_\lambda)_{\lambda\leq\mu}$ on K has some K_λ with a gap.*

(4) *There is a Liouville tower $(K_\lambda)_{\lambda\leq\mu}$ on K with $\mu \geq \omega$ such that either K_0, K_1 or K_2 has a gap.*

(5) *There is an H-field L which has a gap and is a Liouville extension of K.*

Proof. (4) $\Longrightarrow$ (2) and (3) $\Longrightarrow$ (2) are clear. (1) $\Longrightarrow$ (3) and (1) $\Longrightarrow$ (5) follow from ADH 12.4(4).

(1) $\Longrightarrow$ (4): If K has exactly two Liouville closures up to isomorphism over K, then in particular K itself is not Liouville closed. A routine argument shows that every maximal Liouville tower $(K_\lambda)_{\lambda\leq\mu}$ has $\mu \geq \omega$. By Theorem 12.1 either K has a gap or K has asymptotic integration and is not λ-free. If K has a gap, then

for any maximal Liouville tower $(K_\lambda)_{\lambda\le\mu}$, K_0 has a gap. Otherwise, the proof of Proposition 12.2 shows how we can arrange either K_1 or K_2 to have a gap.

(2) $\Longrightarrow$ (1): We will prove the contrapositive. Suppose that K has exactly one Liouville closure up to isomorphism over K and let $(K_\lambda)_{\lambda\le\mu}$ be a Liouville tower on K. We will prove by induction on λ that K_λ is either grounded or λ-free, and thus no K_λ has a gap. The case $\lambda=0$ is clear and the limit ordinal case is taken care of by ADH 6.8 and Lemma 6.9. Suppose $\lambda=\nu+1$ for some ordinal $0\le\nu<\mu$. If K_λ is a real closure of K_ν, then K_λ will be grounded if K_ν is by Definition 3.6 (1) and K_λ will be λ-free if K_ν is by Lemma 6.12. By the inductive hypothesis, K_λ will never be an extension of type (b) or (c). If K_λ is an extension of type (d), then K_λ will also be grounded by [ADH 2017, 10.2.3]. Extensions of type (e), (f) and (g) are necessarily immediate extensions, so if K_ν is grounded, then so is K_λ and if K_ν is λ-free, then so is K_λ by Propositions 7.2, 8.3, and 9.3. Finally, if K_λ is an extension of type (h), and if K_ν is grounded, then so is K_λ by [ADH 2017, 10.5.20], and if K_ν is λ-free then so is K_λ by ADH 6.14.

(5) $\Longrightarrow$ (1): Suppose K has a Liouville H-field extension with a gap. Then by Lemma 12.5, K has a gap or K has asymptotic integration and is not λ-free. By Theorem 12.1, it follows that K has exactly two Liouville closures up to isomorphism over K. □

Remark 12.7. The implication (2) $\Longrightarrow$ (1) of our Corollary 12.6 above occurs without proof in [Aschenbrenner and van den Dries 2002] (see item (II) before their 6.11). Also, (1) $\Longleftrightarrow$ (5) of our Corollary 12.6 is stated without proof in [Aschenbrenner and van den Dries 2005] (see the paragraph after their 4.3).

Acknowledgements

The author thanks the referee for the very careful reading of the manuscript and many helpful suggestions, Chris Miller for suggesting Example 5.9, Santiago Camacho and Elliot Kaplan for their helpful comments, Matthias Aschenbrenner for his encouragement, and especially Lou van den Dries for his gentle guidance and endless patience.

References

[ADH 2017] M. Aschenbrenner, L. van den Dries, and J. van der Hoeven, *Asymptotic differential algebra and model theory of transseries*, vol. 195, Annals of Mathematics Studies, Princeton University Press, Princeton, NJ, 2017. Zbl

[Aschenbrenner 2003] M. Aschenbrenner, “Some remarks about asymptotic couples”, pp. 7–18 in *Valuation theory and its applications, II* (Saskatoon, 1999), edited by F.-V. Kuhlmann et al., Fields Inst. Commun. **33**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[Aschenbrenner and van den Dries 2000] M. Aschenbrenner and L. van den Dries, “Closed asymptotic couples”, *J. Algebra* **225**:1 (2000), 309–358. MR Zbl

[Aschenbrenner and van den Dries 2002] M. Aschenbrenner and L. van den Dries, "H-fields and their Liouville extensions", *Math. Z.* **242**:3 (2002), 543–588. MR Zbl

[Aschenbrenner and van den Dries 2005] M. Aschenbrenner and L. van den Dries, "Asymptotic differential algebra", pp. 49–85 in *Analyzable functions and applications*, edited by O. Costin et al., Contemp. Math. **373**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Bourbaki 1951] N. Bourbaki, *Fonctions d'une variable réelle, Chapitres IV–VII*, Actualités Sci. Ind. **1132**, Hermann et Cie., Paris, 1951. Definitive edition (chapters I-VII) published by Masson, Paris, 1976; translated as *Functions of a real variable*, Springer, 2004. MR Zbl

[Écalle 1992] J. Écalle, *Introduction aux fonctions analysables et preuve constructive de la conjecture de Dulac*, Hermann, Paris, 1992. MR Zbl

[Engler and Prestel 2005] A. J. Engler and A. Prestel, *Valued fields*, Springer, Berlin, 2005. MR Zbl

[Gehret 2017a] A. Gehret, "The asymptotic couple of the field of logarithmic transseries", *J. Algebra* **470** (2017), 1–36. MR Zbl

[Gehret 2017b] A. Gehret, "NIP for the asymptotic couple of the field of logarithmic transseries", *J. Symb. Log.* **82**:1 (2017), 35–61. MR

[van der Hoeven 2006] J. van der Hoeven, *Transseries and real differential algebra*, Lecture Notes in Mathematics **1888**, Springer, Berlin, 2006. MR Zbl

[Kuhlmann 1994] F.-V. Kuhlmann, "Abelian groups with contractions, I", pp. 217–241 in *Abelian group theory and related topics* (Oberwolfach, 1993), edited by R. Göbel et al., Contemp. Math. **171**, Amer. Math. Soc., Providence, RI, 1994. MR Zbl

[Kuhlmann 1995] F.-V. Kuhlmann, "Abelian groups with contractions, II: Weak o-minimality", pp. 323–342 in *Abelian groups and modules* (Padova, 1994), edited by A. Facchini and C. Menini, Math. Appl. **343**, Kluwer Acad. Publ., Dordrecht, 1995. MR Zbl

[Kuhlmann 2000] S. Kuhlmann, *Ordered exponential fields*, Fields Institute Monographs **12**, Amer. Math. Soc., Providence, RI, 2000. MR Zbl

[Lindemann 1882] F. Lindemann, "Ueber die Zahl π", *Math. Ann.* **20**:2 (1882), 213–225. MR JFM

[Rosenlicht 1979] M. Rosenlicht, "On the value group of a differential valuation", *Amer. J. Math.* **101**:1 (1979), 258–266. MR Zbl

[Rosenlicht 1980] M. Rosenlicht, "Differential valuations", *Pacific J. Math.* **86**:1 (1980), 301–319. MR Zbl

[Rosenlicht 1981] M. Rosenlicht, "On the value group of a differential valuation, II", *Amer. J. Math.* **103**:5 (1981), 977–996. MR Zbl

[Rosenlicht 1983a] M. Rosenlicht, "Hardy fields", *J. Math. Anal. Appl.* **93**:2 (1983), 297–311. MR Zbl

[Rosenlicht 1983b] M. Rosenlicht, "The rank of a Hardy field", *Trans. Amer. Math. Soc.* **280**:2 (1983), 659–671. MR Zbl

Received December 12, 2016. Revised February 14, 2017.

ALLEN GEHRET
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
1409 W. GREEN STREET
URBANA, IL 61801
UNITED STATES
agehret2@illinois.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.77

BRAID GROUPS AND QUIVER MUTATION

JOSEPH GRANT AND BETHANY ROSE MARSH

We describe presentations of braid groups of type *ADE* and show how these presentations are compatible with mutation of quivers. In types *A* and *D* these presentations can be understood geometrically using triangulated surfaces. We then give a categorical interpretation of the presentations, with the new generators acting as spherical twists at simple modules on derived categories of Ginzburg dg-algebras of quivers with potential.

1. Introduction

Braid groups are fundamental objects in mathematics. Although they are of a topological and geometric nature, they have an algebraic interpretation: a simple presentation by generators and relations which is just based on adjacency of integers [Artin 1925]. This can be encoded in a line graph, and from there one can generalize to define a group from any finite graph, known as the Artin braid group.

The most well known groups defined from graphs are the Coxeter groups (we restrict to the simply laced cases, for simplicity). These are closely related to Artin braid groups: each Coxeter group is a quotient of a corresponding Artin braid group in a natural way. In particular, the symmetric group on n letters is a quotient of the classical braid group on n strands. Coxeter groups naturally split into two distinct classes: those of finite type, corresponding to the Dynkin diagrams of type *ADE*, and those of infinite type. Although all Artin braid groups are infinite, the Artin braid groups of Dynkin type have a different character to those not of Dynkin type, and are known as Artin groups of "finite type".

This work was supported by the Engineering and Physical Sciences Research Council (grant number EP/G007497/1) and the Mathematical Sciences Research Institute, Berkeley.
MSC2010: primary 13F60, 16G20, 20F36; secondary 16E35, 16E45, 18E30.
Keywords: mutation, braid groups, cluster algebras, Ginzburg dg algebra, spherical twist.

This dichotomy also arises in another area of mathematics which has generated a lot of interest in the recent years: cluster algebras. In this theory, there is a notion of finite-type cluster algebras, which again correspond to the Dynkin diagrams [Fomin and Zelevinsky 2003]. Cluster algebras are specified by a directed graph, known as a quiver, together with other information. A key ingredient in the definition is the notion of mutation, which changes the arrows in a quiver in a nonobvious manner which generalizes reflection at a source or sink. Barot and Marsh [2015] have given new presentations of Coxeter groups of finite type based on quivers obtained from Dynkin diagrams under finite sequences of mutations. Our first result generalizes this to braid groups:

Theorem A. (Theorem 2.12) Let Q be a quiver, with vertices $1, 2, \dots, n$, obtained from a Dynkin quiver by a finite sequence of mutations. Let B_Q be the group with generators $s_1, s_2, \dots, s_n$, subject to the relations

(a) $s_i s_j = s_j s_i$ if there is no arrow between i and j (in either direction);

(b) $s_i s_j s_i = s_j s_i s_j$ if there is an arrow between i and j (in either direction);

(c) $s_{i_1} s_{i_2} \cdots s_{i_n} s_{i_1} \cdots s_{i_{n-2}} = s_{i_2} s_{i_3} \cdots s_{i_n} s_{i_1} \cdots s_{i_{n-1}} = \cdots = s_{i_n} s_{i_1} s_{i_2} \cdots s_{i_n} s_{i_1} s_{i_2} \cdots s_{i_{n-3}}$, whenever $i_1 \to i_2 \to \cdots \to i_n \to i_1$ is a chordless cycle in Q.

Then B_Q is isomorphic to the Artin braid group of the same Dynkin type as Q.

We prove our result via isomorphisms between abstractly defined groups, which can be thought of as mutations of groups, even though the resulting groups are isomorphic. The Artin group presentations we obtain induce presentations of the corresponding Coxeter groups which are distinct from those in [Barot and Marsh 2015]; we also give a compatibility result which shows the relationship between the two presentations. Why have we chosen to use presentations which don't agree with the earlier work? This is explained in the following two sections of the paper, as we now detail.

Certain cluster algebras can be understood using pictures. A (tagged) triangulation of a Riemann surface with marked points on its boundary defines a quiver [Fomin and Zelevinsky 2002; Fomin et al. 2008]; see also [Caldero et al. 2006]. Then mutation of the quiver has a natural interpretation in terms of swapping one diagonal of a given quadrilateral for the other. So these cluster algebras have a topological interpretation. In particular, such descriptions are available for the infinite families of Dynkin type. It is natural to ask whether we can understand the generators above, and the isomorphisms corresponding to mutations, in terms of the geometry of the surface. The answer is yes, as shown by the following theorem:

Theorem B. (Theorem 3.6) Let Δ be a Dynkin diagram of type A_n or type D_n. In the former case, let (X, M) be a disk with $n+3$ marked points on its boundary. In the latter case, let (X, M) be a disk with n marked points on its boundary and one marked point in its interior, taken to be a cone point of order two (so X is an orbifold in this case).

Let $\mathcal{T}$ be a tagged triangulation of (X, M). Let $G_{\mathcal{T}}$ be the graph in (X, M) dual to $\mathcal{T}$. For each vertex i of the quiver $Q_{\mathcal{T}}$ associated to $\mathcal{T}$ as in [Fomin and Zelevinsky 2002; Fomin et al. 2008], let σ_i be the braid of (X, M) associated to the edge of $G_{\mathcal{T}}$ crossing the tagged arc in $\mathcal{T}$ corresponding to i (see Definition 3.3). Then there is an isomorphism between the subgroup $H_{\mathcal{T}}$ of the braid group generated by the σ_i and the group B_Q defined above, taking σ_i to s_i.

Furthermore, in type A_n, the subgroup $H_{\mathcal{T}}$ coincides with the braid group of (X, M), while in type D_n, it is of index two in the braid group of (X, M).

As well as the original combinatorial and commutative algebraic approach to cluster algebras, and the geometric approach described above, there is a third approach which has proved very powerful: the representation theoretic approach [Buan et al. 2006; Caldero et al. 2006]. This approach uses finite-dimensional (noncommutative) algebras and ideas from categorification to better understand cluster algebras, and has received intense study. Braid groups also appear in representation theory and categorification [Rouquier and Zimmermann 2003; Seidel and Thomas 2001]: in many important situations there are actions of braid groups on derived categories via spherical twists. One example of this is given by certain derived categories of differential graded algebras [Ginzburg 2007; Keller and Yang 2011] which are known to cover the categories appearing in the representation theoretic approach to cluster algebras [Amiot 2009]. One might hope that these categorical braid group actions are related to our presentations of braid groups, and we show that this is indeed the case.

First, we make a connection between the categorical and the geometric situations. The relevant differential graded algebras are defined by use of a quiver together with a formal sum of cycles in that quiver known as a potential [Ginzburg 2007]. Mutation of quivers of potential has been defined [Derksen et al. 2008] and, in the situations where our cluster algebra comes from a Riemann surface, the mutation of potentials also has a geometric interpretation [Labardini-Fragoso 2009]. Relying heavily on results of Labardini-Fragoso [2009; 2016], we observe that the potential defined on mutation-Dynkin quivers according to the geometric procedure is equivalent to the "obvious" potential that one might guess (Proposition 4.4). So, while the potential is important, it is in fact entirely determined by the quiver in types A and D. Note that this result could also be proved relatively easily via a direct calculation.

Next we show that we do indeed obtain an action of the groups B_Q (defined using mutation-Dynkin quivers) on derived categories of Ginzburg differential graded algebras in which the generators act via spherical twists. After setting up all the technical machinery correctly, the main difficulty in proving this is to check that the mutation procedure for the groups B_Q, which relates the group associated to a quiver to the group associated to a mutated quiver, actually lifts to the categorical setting as a natural isomorphism of functors. We do this, using important results of Keller and Yang [2011]. From here, we can use the earlier theory developed here to show that the generators s_i of finite type Artin braid groups from Theorem A can be viewed as derived autoequivalences:

Theorem C. (Theorem 4.16) Let (Q, W) be a mutation-Dynkin quiver with potential of type ADE, and let $\Gamma_{Q,W}$ be the corresponding Ginzburg differential graded algebra. Let $\mathrm{D_{fd}}(\Gamma_{Q,W})$ denote the full subcategory of the derived category $\mathrm{D}(\Gamma_{Q,W})$ on objects with finite-dimensional total homology. Then there is a group homomorphism

$$B_Q \to \operatorname{Aut} \mathrm{D_{fd}}(\Gamma_{Q,W}), \qquad s_i \mapsto F_i$$

sending the group generator associated to the vertex i of Q to the spherical twist F_i at the simple $\Gamma_{Q,W}$-module S_i.

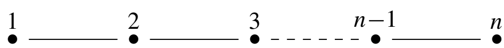
Since we started work on this project, we have become aware of independent work by other authors. A. King and Y. Qiu have a related project, and were aware of the new relations between spherical twists and a topological interpretation of the spherical twist group; see [Qiu 2016], especially Section 10.1. In particular, an independent proof of a version of Theorem 2.10 in types A and D was announced in [Qiu 2016]. A key difference in our approach is the use of an orbifold with cone point of degree two in type D. In [Nagao 2010, §2.2], K. Nagao refers to an action of the mapping class group of a marked surface on the derived category of a Ginzburg dg-algebra associated to a triangulation.

Since we released the first draft of this article, the preprint [Haley et al. 2014] has appeared, where the authors give a presentation (different from the one given here) of the Artin braid group for each diagram of finite type (in the cluster-theoretic sense). This includes the non-simply-laced cases (not considered here) but does not include a topological or categorical interpretation.

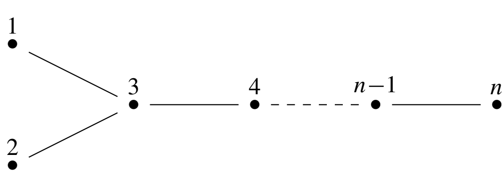
2. Presentations of braid groups

Braid groups. Let Δ be a graph of ADE Dynkin type, i.e., Δ is a graph of type A_n for $n \geq 1$, of type D_n for $n \geq 4$, or of type E_6, E_7 or E_8.

Type A_n:



Type D_n:

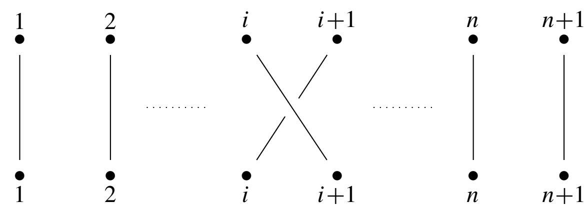


In particular, Δ has no double edges or cycles. Let I be the set of vertices of Δ. We can associate a group B_Δ to Δ, which we call *the braid group of* Δ. It has a distinguished set of generators $S_\Delta = \{s_i\}_{i \in I}$, and the relations depend on whether or not two vertices are connected by an edge. They are

(i) $s_i s_j = s_j s_i$ if i and j are not connected by an edge;

(ii) $s_i s_j s_i = s_j s_i s_j$ if i and j are connected by an edge.

If Δ is of type A_n then we recover the "usual" braid group, sometimes denoted B_{n+1}. Its generators can be visualized as

$s_i =$ 

and the relations of type (i) record the fact that crossings of adjacent pairs of strings which are far apart commute, while relations of type (ii) record a Reidemeister 3 move.

If we also impose the relation that $s_i^2 = 1$ for all $i \in I$ then we recover the Coxeter group of type Δ. More information on Coxeter groups and braid groups can be found in [Humphreys 1990; Kassel and Turaev 2008].

Mutation of quivers. A quiver is just a directed graph. Throughout this article we will only work with quivers with finitely many vertices and finitely many arrows that have no loops or oriented 2-cycles. For a given quiver Q, we again denote its set of vertices by I.

There is a procedure to obtain one quiver from another, called *quiver mutation*, due to Fomin and Zelevinsky [2002, §4]. Fix Q and let $k \in I$. Then we obtain the mutated quiver $\mu_k(Q)$ as follows:

(i) for each pair of arrows $i \to k \to j$ through k, add a formal composite $i \to j$;

(ii) reverse the orientation of all arrows incident with k;

(iii) remove a maximal set of 2-cycles (we may have created 2-cycles in the previous two steps).

It is a basic but important observation that quiver mutation does not change the set of vertices. One can also check that mutation is an involution.

We call a cycle in an unoriented graph (or in the underlying unoriented graph of a quiver) *chordless* if the full subgraph on the vertices of the cycle contains no edges which are not part of the cycle. We call a quiver *Dynkin* if its underlying unoriented graph is a Dynkin graph of type *ADE*, and *mutation-Dynkin* if it can be obtained by mutating a Dynkin quiver finitely many times. By a theorem of Fomin and Zelevinsky [2003, Theorem 1.4], there are only finitely many quivers that can be obtained by mutating a given Dynkin quiver.

The following fact, due to Fomin and Zelevinsky, will be useful to us:

Proposition 2.1. *In any mutation-Dynkin quiver, there are no double arrows and all chordless cycles are oriented.*

Proof. By [Fomin and Zelevinsky 2003, Theorem 1.8], the entries in the corresponding exchange matrix B satisfy $|B_{xy}B_{yx}| \le 3$ for all x, y (known as being 2-finite). Hence there cannot be any double arrows in the quiver.

Now let Q be a mutation-Dynkin quiver and C a chordless cycle in Q. Then, since Q is 2-finite, so is C. By Proposition 9.7 of the same paper, C must be an oriented cycle. □

Groups from quivers. Let Q be a mutation-Dynkin quiver.

Definition 2.2. Let B_Q be the group with generators $S_Q = \{s_i\}_{i \in I}$ subject to the following relations:

(i) $s_i s_j = s_j s_i$ whenever i and j are vertices with no arrow between them,

(ii) $s_i s_j s_i = s_j s_i s_j$ whenever i and j are vertices of Q and there is an arrow between them (in either direction);

(iii)
$$\begin{aligned} s_1 s_2 \cdots s_n s_1 s_2 \cdots s_{n-2} &= s_2 s_3 \cdots s_n s_1 \cdots s_{n-1} \\ &\;\;\vdots \\ &= s_n s_1 s_2 \cdots s_n s_1 s_2 \cdots s_{n-3} \end{aligned}$$

whenever Q contains an oriented chordless n-cycle

$$\begin{tikzcd} 1 \arrow[r] & 2 \arrow[d] \\ n \arrow[u] & \vdots \arrow[l] \end{tikzcd}$$

Remark 2.3. If Q is a Dynkin quiver, then B_Q is (isomorphic to) the Artin braid group of the corresponding Dynkin type.

This presentation is symmetric but not minimal:

Lemma 2.4. *For each single chordless n-cycle, in the presence of the relations of type* (i) *and* (ii), *any one of the relations of type* (iii) *implies all the others.*

Proof. It is enough to show that if the relation

(1) $$s_1 s_2 \cdots s_n s_1 s_2 \cdots s_{n-2} = s_2 s_3 \cdots s_n s_1 s_2 \cdots s_{n-1}$$

holds then

$$s_1 s_2 \cdots s_n s_1 s_2 \cdots s_{n-2} = s_3 s_4 \cdots s_n s_1 s_2 \cdots s_n.$$

So, we assume that (1) holds. Then we have

$$s_2^{-1}s_1s_2\cdots s_ns_1s_2\cdots s_{n-2}s_n = s_3\cdots s_ns_1s_2\cdots s_{n-1}s_n.$$

The left-hand side can be rewritten, using relations of type (i) and (ii), as

$$\begin{aligned}s_2^{-1}s_1s_2\cdots s_ns_1s_2\cdots s_{n-2}s_n &= s_1s_2s_1^{-1}s_3\cdots s_ns_1s_2\cdots s_{n-2}s_n\\ &= s_1s_2s_3\cdots s_{n-1}s_1^{-1}s_ns_1s_2\cdots s_{n-2}s_n\\ &= s_1s_2\cdots s_{n-1}s_ns_1s_n^{-1}s_2\cdots s_{n-2}s_n = s_1s_2\cdots s_ns_1s_2\cdots s_{n-2},\end{aligned}$$

and the result follows. □

Though the relations look different, by taking an appropriate quotient we can obtain the groups defined by Barot and Marsh [2015] directly:

Lemma 2.5. *If, along with the relations of types* (i) *and* (ii), *we also impose the relations* $s_i^2 = 1$ *for all* $i \in I$, *then the group* B_Q *becomes isomorphic to the group* $\Gamma_{U(Q)}$ *defined in* [*Barot and Marsh 2015, Section* 3], *where* $U(Q)$ *is the underlying graph of* Q.

Proof. As our definition is the usual definition of the braid group for a Dynkin quiver, this follows from results in [Barot and Marsh 2015] and the results below on how our groups change with quiver mutation, but since it is straightforward to give a direct proof, we do so.

We need to show that, in the presence of relations (i), (ii), and $s_i^2 = 1$ for all $i \in I$, our extra relation (iii) holds if and only if the relation

$$(s_1s_2\cdots s_{n-1}s_ns_{n-1}\cdots s_2)^2 = 1$$

and its rotations hold for each n-cycle $1 \to 2 \to \cdots \to n \to 1$. By symmetry, it is enough to check that the relation above is equivalent to $s_1s_2\cdots s_ns_1s_2\cdots s_{n-2} = s_2s_3\cdots s_ns_1\cdots s_{n-1}$.

Using that $s_i = s_i^{-1}$, we see our relation is equivalent to

$$s_1s_2\cdots s_ns_1s_2\cdots s_{n-2}s_{n-1}s_{n-2}\cdots s_1s_n\cdots s_3s_2 = 1.$$

Multiplying out the relation from Barot and Marsh, we see that it is equivalent to

$$s_1s_2\cdots s_{n-1}s_ns_{n-1}\cdots s_2s_1s_2\cdots s_{n-1}s_ns_{n-1}\cdots s_2 = 1.$$

Cancelling out n terms on the left and $n-1$ terms on the right of these two expressions, it just remains to show

$$s_1s_2\cdots s_{n-2}s_{n-1}s_{n-2}\cdots s_2s_1 = s_{n-1}s_{n-2}\cdots s_2s_1s_2\cdots s_{n-2}s_{n-1}.$$

As there is an arrow $i \to i+1$ for each i and the cycle is chordless, the symmetric group on n letters maps onto the subgroup generated by $s_1, \ldots, s_{n-1}$ with the

Figure 1. Mutation of a quiver of mutation Dynkin type.

transposition which swaps i and $i+1$ being sent to s_i. It is easy to see that the corresponding relation holds in the symmetric group, with both sides of the equation representing the transposition which swaps 1 and n. □

We will justify our choice of relations in Remark 4.19.

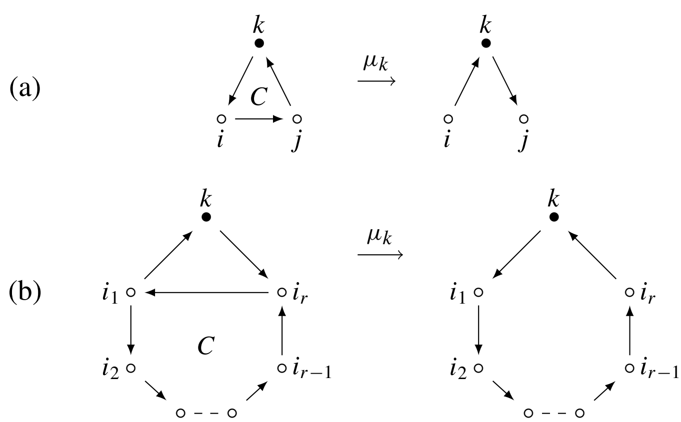
Mutation of groups. Let B_Q be the group associated to the mutation-Dynkin quiver Q, as above, and let k be a vertex of Q. Denote $\mu_k(Q)$ by Q'. Our aim in this section is to show that B_Q is isomorphic to $B_{Q'}$. We will do this by using a group homomorphism $\varphi_k : B_Q \to B_{Q'}$ defined using a formula which lifts the formula used in [Barot and Marsh 2015, §5].

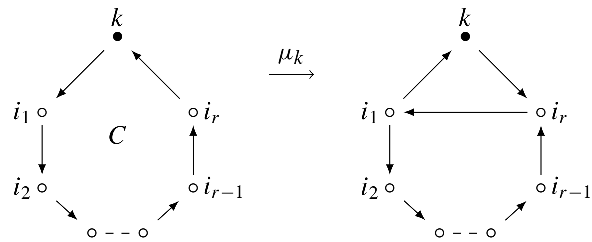
The following lemma follows from results in [Fomin and Zelevinsky 2003] (see [Barot and Marsh 2015, §2]).

Lemma 2.6. *Let Q be a quiver of mutation-Dynkin type, and fix a vertex k of Q. Suppose that k has two neighbouring vertices. Then the possibilities for the induced subquiver of Q containing vertex k and its neighbours are shown in Figure 1. The effect of mutation is shown in each case.*

The following lemma follows from [Barot and Marsh 2015, Lemma 2.5].

Lemma 2.7. *Let Q be a quiver of mutation-Dynkin type, and fix a vertex k of Q. Let C be an oriented cycle in Q. Then C is one of the following. In each case we indicate what happens locally under mutation at k.*



(c) 

(d) *An oriented cycle containing exactly one neighbour of k. Mutation at k reverses the arrow between k and its neighbour in C.*

(e) *An oriented cycle containing no neighbours of k. Mutation at k does not affect C.*

Recall that B_Q is defined using generators s_i for $i \in I$. We denote the corresponding generating set for $B_{Q'}$ by t_i, $i \in I$. Let F_Q be the free group on the generators s_i for $i \in I$.

Definition 2.8. Let $\varphi_k : F_Q \to B_{Q'}$ be the group homomorphism defined by

$$\varphi_k(s_i) = \begin{cases} t_k t_i t_k^{-1} & \text{if } i \to k \text{ in } Q; \\ t_i & \text{otherwise.} \end{cases}$$

Proposition 2.9. *The group homomorphism φ_k induces a group homomorphism (which we also denote by φ_k) from B_Q to $B_{Q'}$.*

Proof. Let us write $\tilde{s}_i = \varphi_k(s_i)$. We must show that the elements $\tilde{s}_i$ in $B_{Q'}$ satisfy the defining relations of B_Q. Note that the t_i satisfy the defining relations for $B_{Q'}$.

Firstly, we check the relations (ii) for an arrow incident with k. Suppose that there is an arrow $i \to k$. Using the fact that $t_i t_k t_i = t_k t_i t_k$,

$$\begin{aligned} \tilde{s}_i \tilde{s}_k \tilde{s}_i &= t_k t_i t_k t_i t_k^{-1} = t_k^2 t_i t_k t_k^{-1} \\ &= t_k^2 t_i. \end{aligned}$$

Also,

$$\tilde{s}_k \tilde{s}_i \tilde{s}_k = t_k^2 t_k t_i t_k^{-1} t_k = t_k^2 t_i.$$

So

$$\tilde{s}_i \tilde{s}_k \tilde{s}_i = \tilde{s}_k \tilde{s}_i \tilde{s}_k,$$

as required.

If there is an arrow $i \leftarrow k$, then

$$\tilde{s}_i \tilde{s}_k \tilde{s}_i = t_i t_k t_i = t_k t_i t_k = \tilde{s}_k \tilde{s}_i \tilde{s}_k.$$

Next, we consider relations (i) and (ii) for all other arrows in Q. Relations of this kind involving pairs of vertices which are not neighbours of k follow immediately from the corresponding relations in B_Q. If only one of the vertices in the relation is a neighbour of k, the relation again follows immediately since t_k commutes with any generator corresponding to a vertex not incident with k in Q' (or equivalently, in Q). So we only need to consider the case where both of the vertices in the pair are incident with k and we can use Lemma 2.6.

Going in either direction in part (a) of Lemma 2.6, the relation $\tilde{s}_i\tilde{s}_j = \tilde{s}_j\tilde{s}_i$ follows from the relation $t_it_j = t_jt_i$ in $B_{Q'}$, so we consider part (b), firstly from left to right. The cycle in Q' gives the relation $t_kt_i = t_jt_kt_it_jt_k^{-1}t_j^{-1}$. Also applying the relation $t_k^{-1}t_j^{-1}t_k^{-1} = t_j^{-1}t_k^{-1}t_j^{-1}$, we obtain

$$\begin{aligned}\tilde{s}_i\tilde{s}_j &= t_kt_it_k^{-1}t_j = t_jt_kt_it_jt_k^{-1}t_j^{-1}t_k^{-1}t_j\\ &= t_jt_kt_it_k^{-1} = \tilde{s}_j\tilde{s}_i.\end{aligned}$$

Going from right to left in part (b), we have, using $t_jt_kt_j = t_kt_jt_k$, $t_it_j = t_jt_i$ and $t_it_kt_i = t_kt_it_k$,

$$\begin{aligned}\tilde{s}_j\tilde{s}_i\tilde{s}_j &= t_kt_jt_k^{-1}t_it_kt_jt_k^{-1} = t_j^{-1}t_kt_jt_it_j^{-1}t_kt_j\\ &= t_j^{-1}t_kt_it_kt_j = t_j^{-1}t_it_kt_it_j\\ &= t_it_j^{-1}t_kt_jt_i = t_it_kt_jt_k^{-1}t_i\\ &= \tilde{s}_i\tilde{s}_j\tilde{s}_i.\end{aligned}$$

Next, we have to check that the $\tilde{s}_i$ satisfy the relations of type (iii) for Q, so we need to consider each type of cycle described in Lemma 2.7. By Lemma 2.4, it is enough to check that, for any given cycle in Q, one of the relations in (iii) holds.

For part (a),

$$\tilde{s}_k\tilde{s}_i\tilde{s}_j\tilde{s}_k = t_kt_it_kt_jt_k^{-1}t_k = t_kt_it_kt_j,$$

while

$$\begin{aligned}\tilde{s}_i\tilde{s}_j\tilde{s}_k\tilde{s}_i &= t_it_kt_jt_k^{-1}t_kt_i = t_it_kt_jt_i\\ &= t_it_kt_it_j,\end{aligned}$$

which is equal to $\tilde{s}_k\tilde{s}_i\tilde{s}_j\tilde{s}_k$ as required.

For part (b), applying a relation for the cycle in Q' in the fourth step,

$$\begin{aligned}\tilde{s}_{i_1}\tilde{s}_{i_2}\cdots\tilde{s}_{i_r}\tilde{s}_{i_1}\tilde{s}_{i_2}\cdots\tilde{s}_{i_{r-2}} &= t_kt_{i_1}t_k^{-1}t_{i_2}\cdots t_{i_r}t_kt_{i_1}t_k^{-1}t_{i_2}\cdots t_{i_{r-2}}\\ &= t_{i_1}^{-1}t_kt_{i_1}t_{i_2}\cdots t_{i_r}t_kt_{i_1}t_k^{-1}t_{i_2}\cdots t_{i_{r-2}}\\ &= t_{i_1}^{-1}t_kt_{i_1}t_{i_2}\cdots t_{i_r}t_kt_{i_1}t_{i_2}\cdots t_{i_{r-2}}t_k^{-1}\\ &= t_{i_1}^{-1}t_{i_1}t_{i_2}\cdots t_{i_r}t_kt_{i_1}t_{i_2}\cdots t_{i_{r-2}}t_{i_{r-1}}t_k^{-1}\\ &= t_{i_2}\cdots t_{i_r}t_kt_{i_1}t_k^{-1}t_{i_2}\cdots t_{i_{r-2}}t_{i_{r-1}}\\ &= \tilde{s}_{i_2}\cdots\tilde{s}_{i_r}\tilde{s}_{i_1}\tilde{s}_{i_2}\cdots\tilde{s}_{i_{r-2}}\tilde{s}_{i_{r-1}}.\end{aligned}$$

For part (c), applying a relation for the cycle in Q' in the fourth step,

$$\begin{aligned}\tilde{s}_{i_1}\tilde{s}_{i_2}\cdots\tilde{s}_{i_{r-1}}\tilde{s}_{i_r}\tilde{s}_k\tilde{s}_{i_1}\cdots\tilde{s}_{i_{r-1}} &= t_{i_1}t_{i_2}\cdots t_{i_{r-1}}t_kt_{i_r}t_k^{-1}t_kt_{i_1}\cdots t_{i_{r-1}}\\ &= t_{i_1}t_{i_2}\cdots t_{i_{r-1}}t_kt_{i_r}t_{i_1}\cdots t_{i_{r-1}}\\ &= t_{i_1}t_kt_{i_2}\cdots t_{i_{r-1}}t_{i_r}t_{i_1}\cdots t_{i_{r-1}}\\ &= t_{i_1}t_kt_{i_1}t_{i_2}\cdots t_{i_{r-1}}t_{i_r}t_{i_1}\cdots t_{i_{r-2}}\\ &= t_kt_{i_1}t_kt_{i_2}\cdots t_{i_{r-1}}t_{i_r}t_{i_1}\cdots t_{i_{r-2}}\\ &= t_kt_{i_1}t_{i_2}\cdots t_{i_{r-1}}t_kt_{i_r}t_k^{-1}t_kt_{i_1}\cdots t_{i_{r-2}}\\ &= \tilde{s}_k\tilde{s}_{i_1}\tilde{s}_{i_2}\cdots\tilde{s}_{i_{r-1}}\tilde{s}_{i_r}\tilde{s}_k\tilde{s}_{i_1}\cdots\tilde{s}_{i_{r-2}},\end{aligned}$$

and we are done. □

Theorem 2.10. *The map $\varphi_k : B_Q \to B_{Q'}$ is a group isomorphism.*

Proof. As mutation is an involution, we can consider the composition

$$\varphi_k : B_Q \xrightarrow{\varphi_k} B_{Q'} \xrightarrow{\varphi_k} B_Q.$$

Fix some $i \in I$. Note that mutation at k does not change whether i and k are connected in the quiver; it just swaps the direction of any arrow that may exist between i and k. So if we have $i \to k$, then $s_i \mapsto t_kt_it_k^{-1} \mapsto s_ks_is_k^{-1}$. If we have $i \leftarrow k$, then $s_i \mapsto t_i \mapsto s_ks_is_k^{-1}$. And if there is no arrow between i and k then $s_i \mapsto t_i \mapsto s_i$. But in this case s_i and s_k commute, so $s_i = s_ks_is_k^{-1}$. Hence in every case $\varphi_k(s_i) = s_ks_is_k^{-1}$, so φ_k is just a conjugation map and therefore $\varphi_k : B_Q \to B_{Q'}$ is an isomorphism. □

Remark 2.11. The inverse of φ_k is the group isomorphism $\varphi_k^{-1} : B_{Q'} \xrightarrow{\sim} B_Q$ defined by

$$\varphi_k^{-1}(t_i) = \begin{cases} s_k^{-1}s_is_k & \text{if } i \to k \text{ in } Q;\\ s_i & \text{otherwise.}\end{cases}$$

Noting Remark 2.3, we have the following:

Theorem 2.12. *If Q is a mutation-Dynkin quiver of type Δ then $B_Q \cong B_\Delta$.*

3. Topological interpretation of the generators

Braid groups. In this section we consider quivers Q which are mutation-equivalent to an orientation of the Dynkin diagram of type Δ, where $\Delta = A_n$ or D_n. By Theorem 2.12, B_Q is isomorphic to the Artin braid group B_Δ of the same Dynkin type. In other words, B_Q gives a presentation of B_Δ. In this section we give a geometric interpretation of this presentation.

We associate an oriented Riemann surface S (with boundary), together with marked points M, to Δ as follows. If $\Delta = A_n$, we take S to be a disk with $n-3$

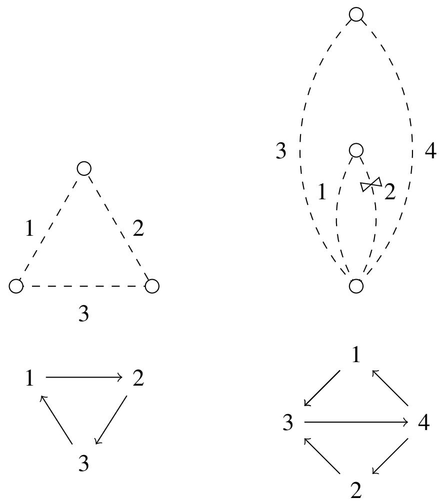


Figure 2. Type I (left) and type II (right) puzzle pieces for tagged triangulations in types A_n and D_n and the corresponding quivers.

marked points on its boundary, as in [Fomin and Zelevinsky 2002; 2003]. If $\Delta = D_n$, we take S to be a disk with one marked point in its interior and n marked points on its boundary, as in [Fomin et al. 2008; Schiffler 2008]. In each case, it was shown that every quiver of the corresponding mutation type arises from a triangulation of (S, M) (tagged, in the type D_n case) in the following way. We follow [Fomin et al. 2008], in a generality great enough to cover both cases (noting that there is at most one interior marked point).

A (simple) *arc* in (S, M) is a curve in S (considered up to isotopy) whose endpoints are marked points in M and which does not have any self-crossings, except possibly at its endpoints. Apart from these endpoints, it must be disjoint from M and the boundary of S, and it must not cut out an unpunctured one- or two-sided polygon.

Two arcs are said to be *compatible* if they are noncrossing in the interior of S. A maximal set of compatible arcs is a *triangulation*.

A *tagged arc* in (S, M) is an arc which does not cut out a once-punctured monogon; each of its ends is tagged, either plain or notched. Plain tags are omitted, while notched tags are displayed using the bow-tie symbol $\bowtie$. An end incident with a boundary marked point is always tagged plain. Two tagged arcs α, β are *compatible* if

(i) the untagged arcs underlying α and β are compatible, and

(ii) if the untagged versions of α and β are different but share an endpoint, then the corresponding ends of α and β are tagged in the same way.

A *tagged triangulation* $\mathcal{T}$ of (S, M) is a maximal collection of tagged arcs in (S, M). Note that if none of the marked points in M lies in the interior of S, every end of an arc in a tagged triangulation must be tagged plain, and tagged triangulations of S can be identified with triangulations of S.

The set M of marked points divides the boundary components of (S, M) into connected components, which we call *boundary arcs*. Note that the boundary arcs do not lie in a triangulation or tagged triangulation of (S, M), by definition.

The tagged triangulation $\mathcal{T}$ can be built up by gluing together puzzle pieces of the two types shown in Figure 2 (see [Fomin et al. 2008, Remark 4.2]) by gluing together along boundary arcs. Note that the puzzle piece of type II can only occur in the type D_n case, and then it occurs exactly once.

If α is an arc in a tagged triangulation $\mathcal{T}$, then the *flip* of $\mathcal{T}$ at α is the unique tagged triangulation containing $\mathcal{T} \setminus \{\alpha\}$ but not containing α. By [Fomin et al. 2008], the set of tagged triangulations of (S, M) is connected under flips.

The *quiver* $Q_{\mathcal{T}}$ of a tagged triangulation $\mathcal{T}$ has vertices corresponding to the arcs in $\mathcal{T}$. The quiver is built up by associating a quiver to each puzzle piece; see Figure 2. If a boundary arc in the puzzle piece is also a boundary arc of (S, M), then the corresponding vertex in the quiver is omitted, together with all incident arrows. The quivers are then glued together by identifying vertices whenever the corresponding edges are glued together in the puzzle pieces.

In order to discuss braid groups, we need to consider more general curves in (S, M). We define a *path* in (S, M) to be a (possibly nonsimple) curve whose endpoints lie in S (not necessarily in M).

Definition 3.1. Let $\mathcal{T}$ be a tagged triangulation of (S, M). We associate a graph to $\mathcal{T}$, which we call the *braid graph* $G_{\mathcal{T}}$ of $\mathcal{T}$, as follows. The vertices $V_{\mathcal{T}}$ of $G_{\mathcal{T}}$ are in bijection with the connected components of the complement of $\mathcal{T}$ in (S, M) and, whenever two such connected components have a common tagged arc on their boundaries, there is an edge in $G_{\mathcal{T}}$ between the corresponding vertices. Thus the edges in $G_{\mathcal{T}}$ are in bijection with the arcs in $\mathcal{T}$.

We choose an embedding ι of $G_{\mathcal{T}}$ into (S, M), mapping each vertex to an interior point of the corresponding connected component of the complement of $\mathcal{T}$ in (S, M) and each edge to a path between the images of its endpoints transverse to the corresponding edge in $\mathcal{T}$. We identify $G_{\mathcal{T}}$ with its image under ι.

Note that in the type A case the braid graph is the tree from Section 3.1 of [Caldero et al. 2006].

We associate an orbifold X to S as follows. In the type A_n case, we just take $X = S$, and in the type D_n case we take X to be S with the interior marked point of S interpreted as a cone point of order two. In each case, the set M of marked points induces a corresponding set of marked points in X, which we also denote

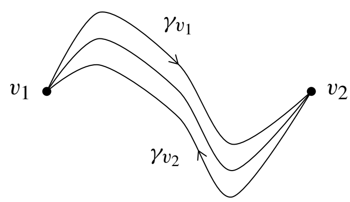


Figure 3. Thickening of the path π (the middle path).

by M. Each arc or tagged arc α in (S, M) induces a corresponding arc or tagged arc in (X, M) which we also denote by α. Thus each (tagged) triangulation $\mathcal{T}$ of (S, M) induces a corresponding set $\mathcal{T}$ of (tagged) arcs in (X, M).

Note also that orbifolds have been used to model cluster algebras in [Felikson et al. 2012]. In this approach, the model for B_n is an orbifold with a cone point of order two, regarded as a folding of D_n, where D_n is modelled by a disk with a single interior marked point (see also Lecture 15 of [Thurston 2012], which was given by A. Felikson).

We denote by X° the orbifold X with the cone point (if there is one) removed (so $X^\circ = X$ in type A_n). Given any set V of vertices in X°, we may consider the corresponding *braid group*, $\Gamma(X, V)$ following [Allcock 2002]. Each element of $\Gamma(X, V)$ (or *braid*) can be regarded as a permutation g of V together with a tuple $\gamma = (\gamma_v)_{v \in V}$ of paths $\gamma_v : [0, 1] \to X^\circ$ such that $\gamma_v(0) = v$ and $\gamma_v(1) = g(v)$ for each $v \in V$. In addition, for each $t \in [0, 1]$, the points $\gamma_v(t)$ for $v \in V$ must all be distinct for all $v \in V$. Braids are considered up to isotopy, and two braids can be multiplied by composing the paths in a natural way; we compose braids from right to left, as for functions.

Remark 3.2. Suppose V and V' are two sets of points in X° and there is a bijection $\rho : V \to V'$. Suppose also that there is a set of paths $\delta_v : [0, 1] \to X^\circ$, for $v \in V$, with $\delta_v(0) = v$ and $\delta_v(1) = \rho(v)$ for all $v \in V$. Suppose furthermore that the points $\gamma_v(t)$ for $v \in V$ and $t \in [0, 1]$ are all distinct. Then the maps δ_v induce a natural isomorphism between $\Gamma(X, V)$ and $\Gamma(X, V')$.

Definition 3.3. Each path π in X° with endpoints v_1, v_2 in V determines a braid σ_π in $\Gamma(X, V)$ as follows (see [Fox and Neuwirth 1962, §7]). We thicken the path π along its length (avoiding the other vertices), closing it off at the end points to form a (topological) disk. We give the boundary of the disk the clockwise orientation. The vertices v_1 and v_2 divide the boundary of the disk into two paths, one from v_1 to v_2 and the other from v_2 to v_1. We define γ_{v_1} to be the former and γ_{v_2} to be the latter. See Figure 3. For $v \in V$ such that $v \neq v_1, v_2$, we define $\gamma_v(t)$ to be v for all $t \in [0, 1]$. Then σ_π is the braid $(\gamma_v)_{v \in V}$. Note that σ_π only depends on the isotopy class of the image of π in (X, V). In particular, it is unchanged if π is reversed.

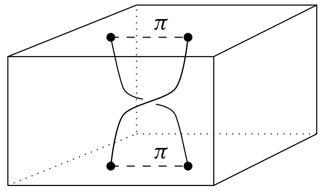


Figure 4. The braid σ_π.

An example of a braid σ_π is displayed as a picture (in the same way as in [Allcock 2002]) in Figure 4. In this figure only, we display π as a dashed line to distinguish it from the braid σ_π.

Interpretation of the generators. Let $\mathcal{T}$ be a triangulation of (S, M). Let $Q_{\mathcal{T}}$ be the quiver of $\mathcal{T}$. Then $Q_{\mathcal{T}}$ has vertices I corresponding to the arcs in $\mathcal{T}$. We denote the arc in $\mathcal{T}$ associated to $i \in I$ by α_i. The corresponding edge in $G_{\mathcal{T}}$ is denoted π_i. Let $\sigma_i = \sigma_{\pi_i} \in \Gamma(X, V_{\mathcal{T}})$ be the corresponding braid. We define $H_{\mathcal{T}}$ to be the subgroup of $\Gamma(X, V_{\mathcal{T}_0})$ generated by the braids σ_i for $i \in I$.

Let $\mathcal{T}_0$ be an initial triangulation of (S, M) defined as follows. In the type A_n case, we choose a marked point P in M and take noncrossing arcs between P and each of the other marked points in M not incident with a boundary arc incident with P. In the type D_n case, we choose two marked points P, Q on the boundary of S. We take two arcs between the interior marked point and Q, one tagged plain at the interior marked point and the other one tagged notched, and an arc between P and Q (not homotopic to a boundary arc). We then take (noncrossing) arcs between P and every other marked point in M on the boundary of S not incident with a boundary arc incident with P. See Figure 5. Then the quiver $Q_{\mathcal{T}_0}$ associated to $Q_{\mathcal{T}_0}$ is a Dynkin quiver of type Δ. By Remark 2.3, $B_{Q_{\mathcal{T}_0}}$ is isomorphic to the Artin braid group of type Δ.

Proposition 3.4. *Let $\mathcal{T}_0$ be the triangulation of (X, M) defined as above. Then there is an isomorphism from $H_{\mathcal{T}_0}$ to $B_{Q_{\mathcal{T}_0}}$ taking the braid σ_i to the generator s_i of $B_{Q_{\mathcal{T}_0}}$. Furthermore, in type A_n, the subgroup $H_{\mathcal{T}_0}$ coincides with $\Gamma(X, V_{\mathcal{T}_0})$, while in type D_n, it is a subgroup of $\Gamma(X, V_{\mathcal{T}_0})$ of index two.*

Proof. For type A_n, see [Fox and Neuwirth 1962] and the explanation in [Allcock 2002, §4]. For type D_n, note that the elements σ_i for $i \in I$ coincide with the generators h_i defined in [Allcock 2002, §1] (via an isomorphism of the kind in Remark 3.2). The result then follows from [Allcock 2002, Theorem 1]. □

The following lemma appears in [Sergiescu 1993, Théorème, part (iv)].

Lemma 3.5. *Let A, B, C be three distinct points in X° and suppose there is a topological disk in X°, with A, B and C lying in order clockwise around its boundary.*

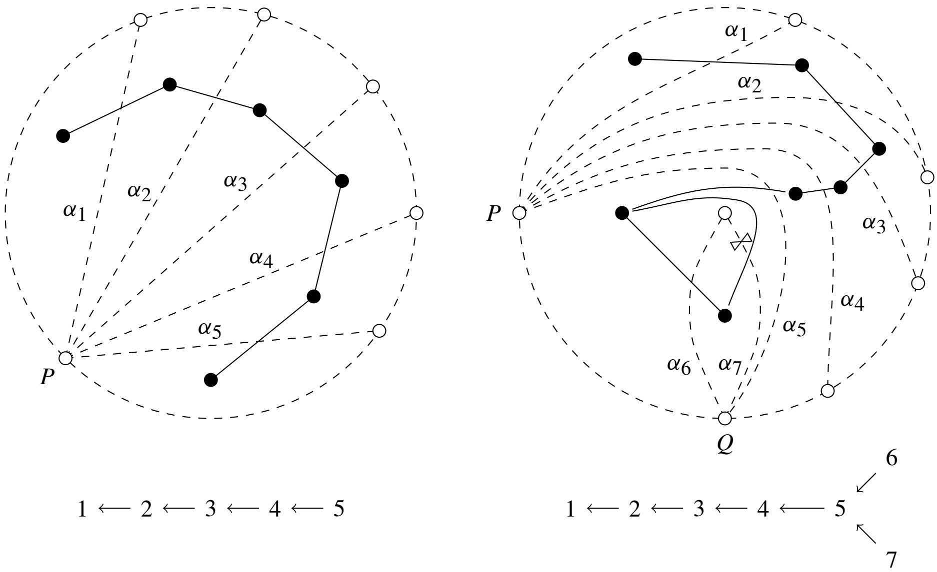


Figure 5. Initial triangulations and the corresponding braid graphs and quivers.

Let AB denote the arc on this boundary between A and B. We define BC and CA similarly. Then $\sigma_{AB}\sigma_{BC} = \sigma_{BC}\sigma_{CA}$.

Theorem 3.6. *Let $\mathcal{T}$ be an arbitrary tagged triangulation of (X, M). Then there is an isomorphism from $H_{\mathcal{T}}$ to $B_{Q_{\mathcal{T}}}$ taking the braid σ_i to the generator s_i of $B_{Q_{\mathcal{T}}}$. Furthermore, in type A_n, the subgroup $H_{\mathcal{T}}$ coincides with $\Gamma(X, V)$, while in type D_n, it is a subgroup of $\Gamma(X, V)$ of index two.*

Proof. The result holds for $\mathcal{T} = \mathcal{T}_0$ by Proposition 3.4. Note that any triangulation can be obtained from $\mathcal{T}_0$ by applying a finite number of flips of tagged triangulations. We show that the theorem is true for an arbitrary tagged triangulation $\mathcal{T}$ by induction on the number of flips required to obtain $\mathcal{T}$ from $\mathcal{T}_0$. To do this, it is enough to show that if the theorem holds for a tagged triangulation $\mathcal{T}$ and α_i is a tagged arc in $\mathcal{T}$ then the theorem also holds for the flip of $\mathcal{T}$ at α_i.

So we assume the result holds for a tagged triangulation $\mathcal{T}$. Thus there is an isomorphism $\psi_{\mathcal{T}} : H_{\mathcal{T}} \to B_{Q_{\mathcal{T}}}$ sending σ_i to s_i. We denote the corresponding elements of $H_{\mathcal{T}'}$ by τ_i and t_i. The tagged arcs in $\mathcal{T}$ are denoted by α_i, for $i \in I$, and we denote the corresponding tagged arcs in $\mathcal{T}'$ by β_i, for $i \in I$. The edges of $G_{\mathcal{T}}$ are denoted π_i, and we denote the edges of $G_{\mathcal{T}'}$ by ρ_i.

We define:

$$\tilde{\tau}_i = \begin{cases} \sigma_k^{-1}\sigma_i\sigma_k, & \text{if } i \to k \text{ in } Q; \\ \sigma_i, & \text{otherwise.} \end{cases}$$

Then it is easy to see that $H_{\mathcal{T}}$ is generated by the $\tilde{\tau}_i$ for $i \in I$.

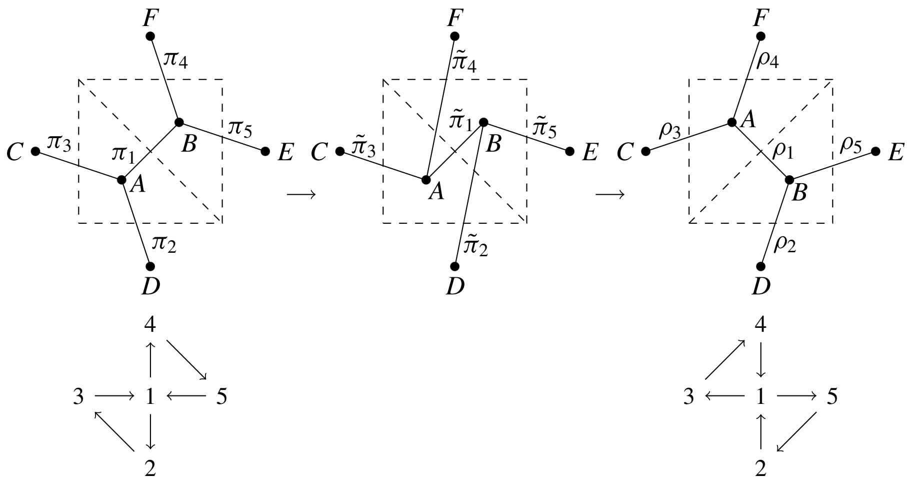


Figure 6. Flip involving an arc (α_1) where two puzzle pieces of type I are glued.

We consider the possible types of flip that can occur, which are determined by the fact that $\mathcal{T}$ can be constructed out of puzzle pieces. Suppose first that $\mathcal{T}'$ is the flip of $\mathcal{T}$ at an arc α where two puzzle pieces of type I are glued together. We label the corresponding vertices in I by $1, 2, 3, 4, 5$, for simplicity, and suppose we are flipping at the edge in $\mathcal{T}$ dual to α_1. The braid graph local to the flip is shown in the left-hand diagram in Figure 6. Applying Lemma 3.5, we see that the middle figure shows paths $\tilde{\pi}_i$ with the property that $\tilde{\tau}_i = \sigma_{\tilde{\pi}_i}$ for $i = 1, 2, 3, 4, 5$.

Rotating vertices A and B clockwise, to get the right-hand diagram in Figure 6, we obtain, via Remark 3.2, an isomorphism from $H_{\mathcal{T}}$ to $H_{\mathcal{T}'}$ taking $\tilde{\tau}_i$ to τ_i for all $i \in I$. The inverse is an isomorphism from $H_{\mathcal{T}'}$ to $H_{\mathcal{T}}$ taking τ_i to $\sigma_k^{-1}\sigma_i\sigma_k$ if there is an arrow $i \to k$ in Q and to σ_i otherwise. Composing with the isomorphism $\varphi_k \circ \psi_{\mathcal{T}}$, where φ_k is the isomorphism in Proposition 2.9, we obtain an isomorphism from $H_{\mathcal{T}'}$ to $B_{Q_{\mathcal{T}'}}$ taking τ_i to t_i as required. This proves the required result in type A, so we are left with the type D case, where puzzle pieces of type II may occur.

We next consider a flip inside a puzzle piece of type II. We can apply essentially the same argument; see Figures 7 and 8. Here we draw the puzzle piece together with the two adjacent triangles, necessarily of type I (since there is only one cone point). To go from the middle diagram to the right-hand diagram in Figure 8, the vertex D should be moved anticlockwise around the cone point. We use the fact that in the right-hand diagram of Figure 8, the resulting path $\tilde{\pi}_1$ is isotopic to the path ρ_1 in $G_{\mathcal{T}'}$, using the fact that the cone point has order two.

Note that the adjoining type I puzzle pieces (in Figures 7 and 8) may not occur, but the argument is easily modified to cover these cases. We also need to consider

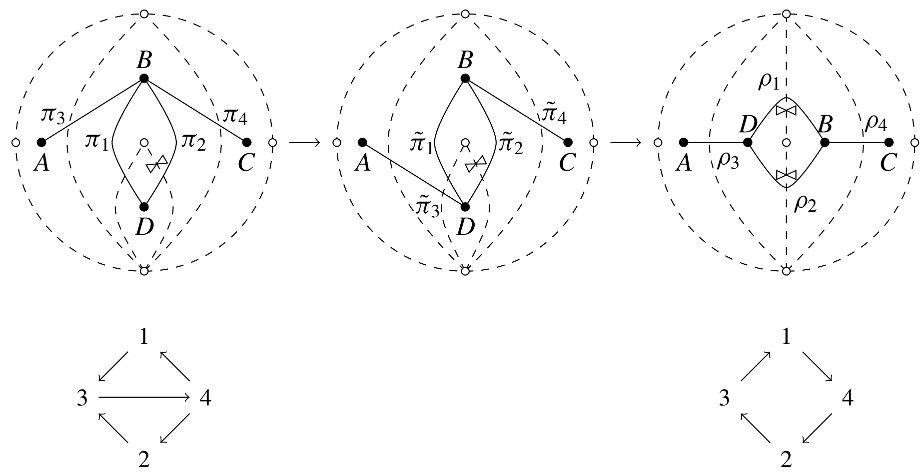


Figure 7. Flip (at α_1) inside a puzzle piece of type II, first case.

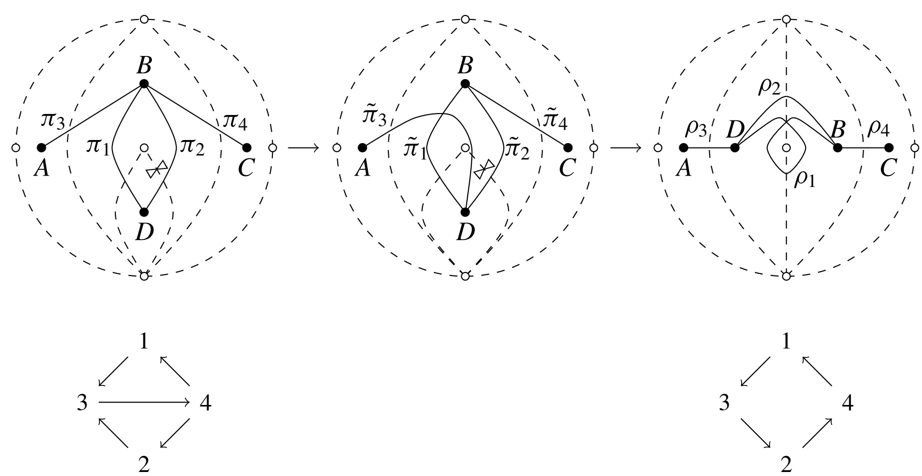


Figure 8. Flip (at α_2) inside a puzzle piece of type II, second case.

the flips from the right-hand diagram in each case to the corresponding left-hand one. We omit the details; a similar argument can be applied in these cases.

Finally, we need to consider a flip involving an arc where a puzzle piece of type I and a puzzle piece of type II have been glued together. These cases are shown in Figures 9 and 10: Figure 9 illustrates the case where the puzzle piece of type I is on the left of the puzzle piece of type II (when it is drawn as shown), while Figure 10 illustrates the case where it is on the right. Again, a similar argument applies in the case of flips from the right-hand diagram to the left-hand one in these cases. □

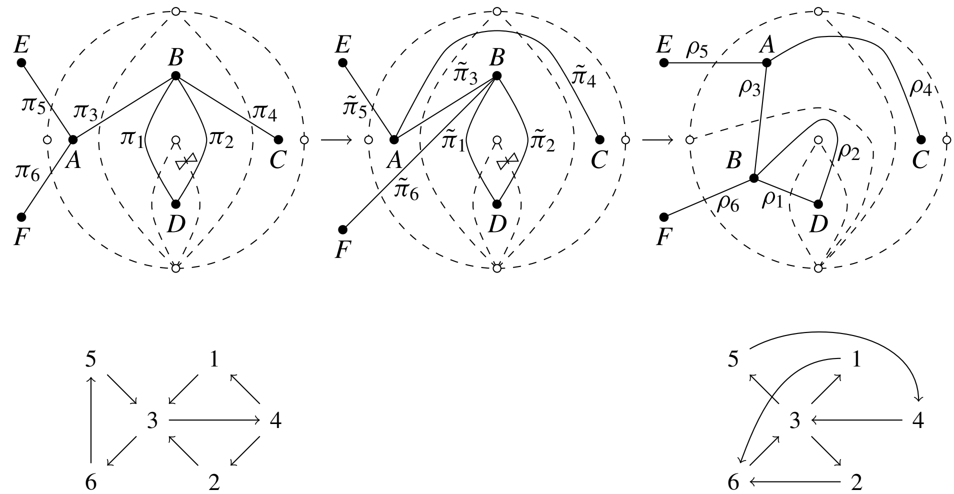


Figure 9. Flip involving an arc (α_3) where puzzle pieces of type I and II are glued, first case.

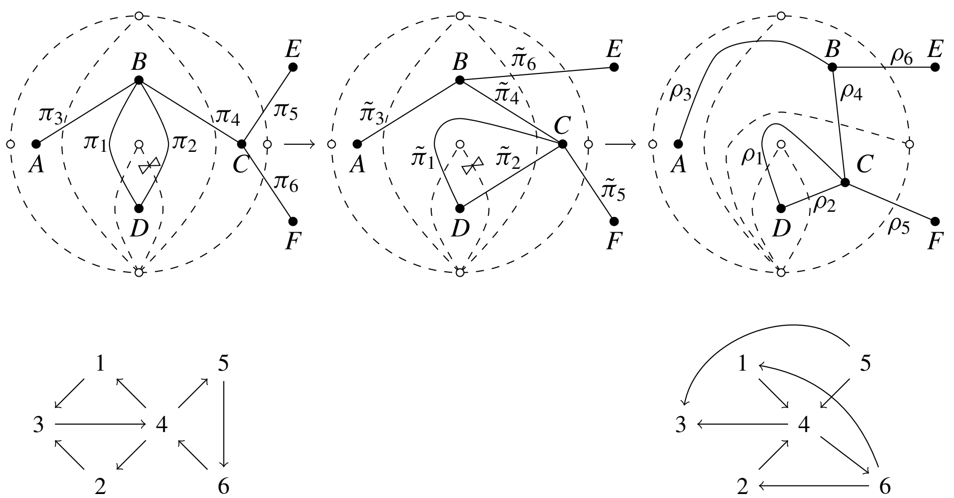


Figure 10. Flip involving an arc (α_4) where puzzle pieces of type I and II are glued, second case.

4. Actions on categories

Quivers with potential. Fix an algebraically closed field $\mathbb{F}$. To any quiver Q we can associate the path algebra $\mathbb{F}Q$, which, as an $\mathbb{F}$-vector space, has basis given by all paths in Q of length ≥ 0, and the multiplication of two paths p_1 and p_2 is their concatenation $p_1 p_2$ if p_1 ends and p_2 starts at the same vertex, and is zero otherwise.

Let $\mathbb{F}Q_{\geq n}$ be the ideal of $\mathbb{F}Q$ generated by the paths in Q of length at least n. We can take the completion $\widehat{\mathbb{F}Q}$ of $\mathbb{F}Q$ with respect to $\mathbb{F}Q_{\geq 1}$, which is defined as:

$$\widehat{\mathbb{F}Q} = \varprojlim_n \frac{\mathbb{F}Q}{\mathbb{F}Q_{\geq n}} = \big\{(a_n + \mathbb{F}Q_{\geq n})_{n=1}^{\infty} \mid a_n \in \mathbb{F}Q,\, \varphi_n(a_n + \mathbb{F}Q_{\geq n}) = a_{n-1} + \mathbb{F}Q_{\geq n-1}\big\},$$

where the limit is taken along the chain of epimorphisms

$$\frac{\mathbb{F}Q}{\mathbb{F}Q_{\geq 1}} \overset{\varphi_2}{\twoheadleftarrow} \frac{\mathbb{F}Q}{\mathbb{F}Q_{\geq 2}} \overset{\varphi_3}{\twoheadleftarrow} \frac{\mathbb{F}Q}{\mathbb{F}Q_{\geq 3}} \twoheadleftarrow \cdots .$$

Let $\widehat{\mathbb{F}Q}_{\mathrm{cyc}}$ denote the subspace of (possibly infinite) linear combinations of cycles in Q. Recall that a *potential* for a quiver Q is an element W of $\widehat{\mathbb{F}Q}_{\mathrm{cyc}}$, regarded up to cyclic equivalence (and for which no two cyclically equivalent paths in Q occur in the decomposition of W). The pair (Q, W) is called a *quiver with potential* [Derksen et al. 2008], which we occasionally abbreviate to QP. The following definition is due to Derksen, Weyman and Zelevinsky:

Definition 4.1 [Derksen et al. 2008, Definition 4.2]. Let Q_1 and Q_2 be two quivers with the same vertex set I, and (Q_1, W_1) and (Q_2, W_2) be two QPs. A *right equivalence* between (Q_1, W_1) and (Q_2, W_2) is an algebra isomorphism $\varphi : \widehat{\mathbb{F}Q}_1 \to \widehat{\mathbb{F}Q}_2$ such that $\varphi(W_1)$ is cyclically equivalent to W_2 and φ is the identity when restricted to the semisimple subalgebra $\mathbb{F}I$ of $\widehat{\mathbb{F}Q}_1$.

A quiver with potential (Q, W) with W containing paths of length two or more is *trivial* if Q is a disjoint union of 2-cycles and there is an algebra automorphism of $\widehat{kQ}$ preserving the span of the arrows of Q (a *change of arrows*) which takes W to the sum of the 2-cycles in Q. A quiver with potential (Q, W) is said to be *reduced* if W is a linear combination of cycles in Q of length 3 or more.

The *splitting theorem* [Derksen et al. 2008, Theorem 4.6] states that every quiver with potential can be written as a direct sum of a reduced quiver with potential and a trivial quiver with potential which are unique up to right equivalence.

Let (Q, W) be a quiver with potential, and let k be a vertex of Q not involved in any 2-cycles. By replacing W with a cyclically equivalent potential on Q if necessary, we can assume that none of the cycles in the decomposition of W start or end at k. We denote by $\tilde{\mu}_k(Q, W)$ the nonreduced mutation of (Q, W) at k in Q, as defined in [Derksen et al. 2008, §5]. Then, by Theorem 5.2 of the same paper, the right equivalence class of $\tilde{\mu}_k(Q, W)$ is determined by the right equivalence class of (Q, W). The *mutation* $\mu_k(Q, W)$ of (Q, W) at k is then defined to be the reduced component of $\tilde{\mu}_k(Q, W)$, and is uniquely determined up to right equivalence, given the right equivalence class of (Q, W).

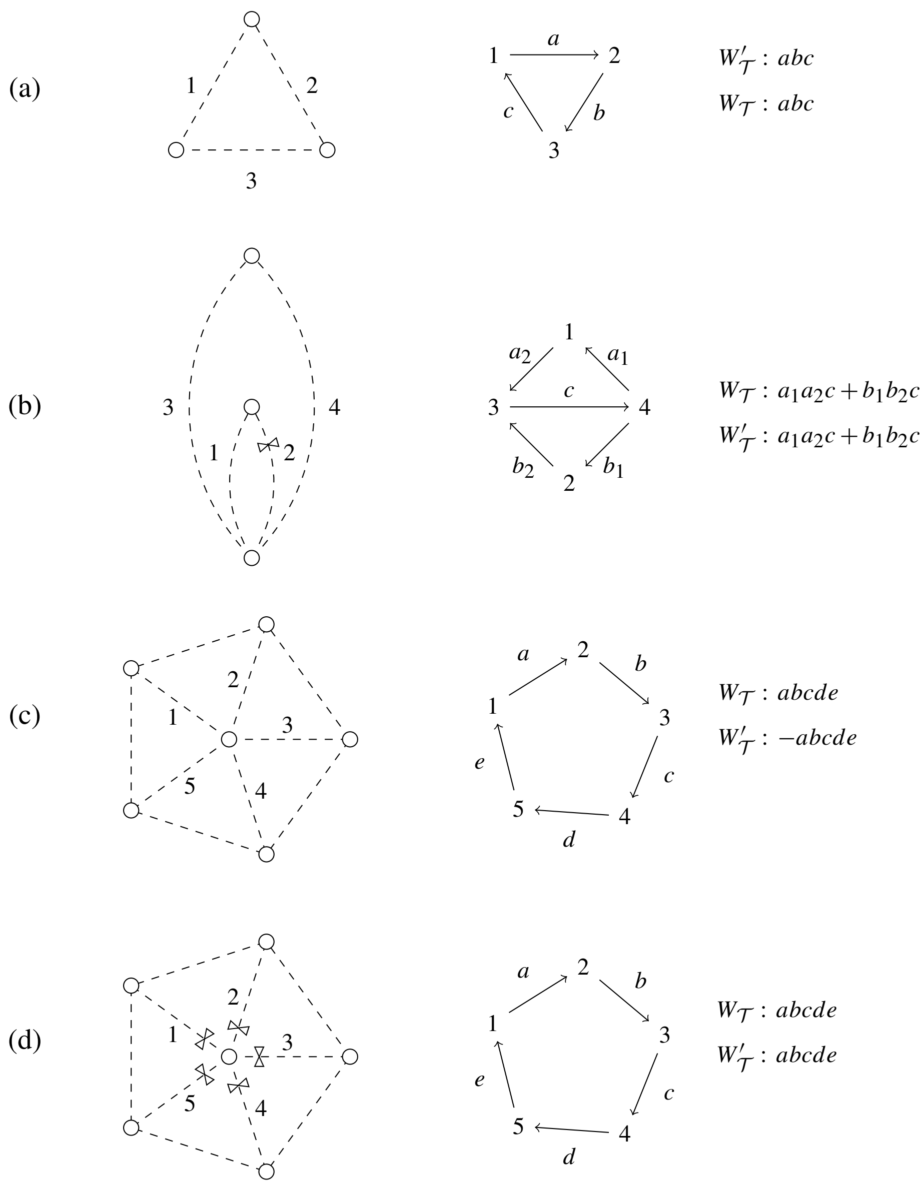


Figure 11. Terms in the potential $W_{\mathcal{T}}$.

As before, we will say a quiver with potential (Q, W) is *Dynkin* if the underlying unoriented graph of Q is an orientation of a Dynkin quiver (and hence $W = 0$). We shall say that a quiver with potential (Q', W') is *mutation-Dynkin* if it can be obtained by repeatedly mutating a Dynkin quiver with potential in the above sense. *Note:* For the rest of this subsection we will restrict to Dynkin types A and D.

Let (S, M) be the Riemann surface with marked points associated to Δ as in Section 3. So, if $\Delta = A_n$, we take S to be a disk with $n-3$ points on its boundary, and if $\Delta = D_n$, we take S to be a disk with one marked point in its interior and n marked points on its boundary.

Let Q be a mutation-Dynkin quiver. By [Fomin et al. 2008], $Q = Q_{\mathcal{T}}$ for some tagged triangulation $\mathcal{T}$ of (S, M). Let W, W' be the sum of the terms coming from local configurations in $\mathcal{T}$ as shown in Figure 11 (where in (c) and (d) there are at least three arcs incident with the interior marked point).

Then $W_{\mathcal{T}}$ is the potential given by taking the sum of the induced cycles in $Q_{\mathcal{T}}$ (i.e., induced subgraphs of $Q_{\mathcal{T}}$ which are cycles), and $W'_{\mathcal{T}}$ is the potential associated to $\mathcal{T}$ in [Labardini-Fragoso 2016, §3], taking the parameter associated to the internal marked point (if there is one) to be equal to -1. Then we have the following:

Lemma 4.2. *The potentials $W_{\mathcal{T}}$ and $W'_{\mathcal{T}}$ are right equivalent.*

Proof. We assume we are in case D_n, since the two potentials coincide in case A_n. If the interior marked point is as in case (c) of Figure 11 (with at least 3 arcs incident with it), then there is a unique triangle in $\mathcal{T}$ with sides 1 and 2. We label the arrows in the corresponding 3-cycle in $W_{\mathcal{T}}$ or $W'_{\mathcal{T}}$ by a, x, y, in order around the cycle. Then the automorphism φ of $\widehat{kQ_{\mathcal{T}}}$ negating a and x and taking each other arrow to itself gives a right equivalence between $W_{\mathcal{T}}$ and $W'_{\mathcal{T}}$, since a and x are not involved in any other terms in any of these potentials.

If the interior marked point is as in case (d), then $W_{\mathcal{T}}$ and $W'_{\mathcal{T}}$ coincide. □

We recall the following special case of [Labardini-Fragoso 2016, Theorem 8.1].

Theorem 4.3 [Labardini-Fragoso 2016]. *Let $\mathcal{T}$, $\mathcal{T}'$ be triangulations of (S, M). If $\mathcal{T}'$ is obtained from $\mathcal{T}$ by flipping at an arc α_k then $\mu_k(Q_{\mathcal{T}}, W'_{\mathcal{T}})$ is right equivalent to $(Q_{\mathcal{T}'}, W'_{\mathcal{T}'})$.*

By [Derksen et al. 2008, Theorem 7.1], it follows from this that the quiver of $\mu_k(Q_{\mathcal{T}}, W_{\mathcal{T}})$ coincides with the quiver obtained from $Q_{\mathcal{T}}$ by Fomin–Zelevinsky quiver mutation at k.

Hence we can effectively ignore potentials:

Proposition 4.4. *Any mutation-Dynkin quiver with potential $(\widetilde{Q}, \widetilde{W})$ of type A or D is right equivalent to $(\widetilde{Q}, W_{\widetilde{Q}})$, where $W_{\widetilde{Q}}$ is the sum of all chordless cycles in $\widetilde{Q}$.*

Proof. Note that a Dynkin quiver with zero potential is of the form $(Q_{\mathcal{T}}, W_{\mathcal{T}})$ for some triangulation $\mathcal{T}$; see [Fomin et al. 2008]. Suppose that $(\widetilde{Q}, \widetilde{W})$ is obtained from a Dynkin quiver with zero potential by iterated mutation in the sense of [Derksen et al. 2008]. Then, by Theorem 4.3 and Lemma 4.2, $(\widetilde{Q}, \widetilde{W})$ is right equivalent to $(Q_{\mathcal{T}}, W_{\mathcal{T}})$ for some triangulation $\mathcal{T}$ of (S, M). □

Note that an alternative proof of Proposition 4.4 would be to compute the mutation of a quiver with potential $(Q_{\mathcal{T}}, W'_{\mathcal{T}})$ directly, and show that it is right equivalent to $(Q_{T'}, W'_{T'})$. This is not too difficult to do, but requires consideration of several cases and still requires arguments dealing with changes of sign as in Lemma 4.2, so we instead refer to [Labardini-Fragoso 2016] above.

Differential graded algebras and modules. Let $\mathbb{F}$ be an algebraically closed field. We think of $\mathbb{F}$ as a graded $\mathbb{F}$-algebra concentrated in degree 0. If $V = \bigoplus V_i$ is a graded $\mathbb{F}$-module then let $V[j]$ be the graded $\mathbb{F}$-module with $(V[j])_i = V_{i+j}$. If $f : V \to W$ is a map of graded vector spaces with homogeneous components $f_i : V_i \to W_i$ then let $f[j] : V[j] \to W[j]$ be the map of graded vector spaces with homogeneous components $f[j]_i : V[j]_i \to W[j]_i$ defined by $f[j]_i(v) = (-1)^j f_{i+j}(v)$ for $v \in V[j]_i = V_{i+j}$. Thus [1] is an endofunctor of the category of graded $\mathbb{F}$-modules, called the *shift functor*.

If V and W are graded vector spaces, we will refer to a map $f : V \to W[j]$ of graded vector spaces as a map from V to W of degree j. We use the Koszul sign rule for graded $\mathbb{F}$-algebras, so if $f : V \to V'$ and $g : W \to W'$ are maps of graded $\mathbb{F}$-modules of degree m and n then

$$(f \otimes g)(v \otimes w) = (-1)^{in} f(v) \otimes g(w)$$

for $v \in V_i$ and $w \in W$.

A unital differential graded algebra (or dg-algebra, or dga) over $\mathbb{F}$ is a graded $\mathbb{F}$-algebra $A = \bigoplus_{i \in Z} A_i$ with multiplication $m : A \otimes_{\mathbb{F}} A \to A$ of degree 0 together with a unit $\iota : \mathbb{F} \hookrightarrow A$ and an $\mathbb{F}$-linear differential $d : A \to A$ of degree $+1$. These should satisfy

- the associativity relation $m \circ (1 \otimes m) = m \circ (m \otimes 1)$;
- the boundary relation $d^2 = 0$;
- the Leibniz relation $d \circ m = m \circ (1 \otimes d + d \otimes 1)$;
- the unital relation $m \circ (\mathrm{id}_A \otimes \iota) = m \circ (\iota \otimes \mathrm{id}_A)$, which should agree with the $\mathbb{F}$-algebra structure of A.

We often denote our dga by (A, d), or simply by A. Each dga (A, d) has an underlying unital graded algebra, obtained by simply forgetting the differential, which we denote $u(A)$.

A left module M for A is a graded left $\mathbb{F}$-module M which has a left action $m_M : A \otimes M \to M$ of $u(A)$ together with a map $d_M : M \to M$ of degree $+1$, called a *differential*, such that

$$d_M \circ m_M = m_M \circ (1 \otimes d_M + d \otimes 1).$$

We always have the *regular* module $M = A$ with $d_M = d$ and $m_M = m$. Similarly, a right module M for A is a graded right $\mathbb{F}$-module M which has a right action $m_M : M \otimes A \to M$ of $u(A)$ together with a differential d_M such that $d_M \circ m_M = m_M \circ (1 \otimes d + d_M \otimes 1)$. If (M, d_M) is an A-module, then $(M[1], d_M[1])$ is also an

A-module, which we sometimes just write as $M[1]$. Modules for A are modules for $u(A)$, simply by forgetting the differential.

A map $f : M \to N$ of left A-modules is a degree 0 map of $u(A)$-modules such that f commutes with the differentials: $d_N \circ f = f \circ d_M$. We thus obtain a category A-Mod of left A-modules, and we write the morphism spaces in this category as $\mathrm{Hom}_{A\text{-Mod}}(M, N)$. A-Mod is an $\mathbb{F}$-category: each morphism space is an $\mathbb{F}$-module.

Given two differential algebras (A, d_A) and (B, d_B), an A-B-bimodule (M, d_M) is a graded $\mathbb{F}$-module which is a left (A, d_A)-module with left action m^ℓ and a right (B, d_B)-module with right action m^r where the two actions commute:

$$m^r \circ (m^\ell \otimes \mathrm{id}_B) = m^\ell \circ (\mathrm{id}_A \otimes m^r).$$

We will always assume that $\mathbb{F}$ acts centrally. Under this assumption we can, and will, identify left A-modules with A-k-bimodules and A-B-bimodules with left $A \otimes_{\mathbb{F}} B^{\mathrm{op}}$-modules, where B^{op} denotes the algebra B with the order of multiplication reversed. A map of bimodules should commute with the differential on both the left and the right, and we obtain an $\mathbb{F}$-category A-Mod-B of A-B-bimodules.

Given a map $f : M \to N$ of left A-modules, we can construct a new left A-module called the *cone* of f, denoted $\mathrm{cone}(f)$. As a left module for $u(A)$, we have $\mathrm{cone}(f) = N \oplus M[1]$. The differential is given by

$$\begin{pmatrix} d_N & 0 \\ f[1] & d_{M[1]} \end{pmatrix}.$$

If L is isomorphic to $\mathrm{cone}(f)$ for some map $f : M \to N$, we say that L is an *extension* of M by $L[-1]$.

We will use the following lemma, whose proof follows immediately from the definitions, repeatedly.

Lemma 4.5. *Let $f : M \to N$ be a map in A-*Mod*.*

(i) *Let $F : A\text{-Mod} \to B\text{-Mod}$ be an additive functor which commutes with the shift functor. Then we have an isomorphism $\mathrm{cone}(Ff) \cong F\,\mathrm{cone}(f)$ in B-*Mod*.*

(ii) *For any commutative diagram*

$$\begin{tikzcd} M \arrow[r, "f"] \arrow[d, "\varphi_M", "\sim"'] & N \arrow[d, "\varphi_N", "\sim"'] \\ M' \arrow[r, "f'"] & N' \end{tikzcd}$$

*in A-*Mod* where both φ_M and φ_N are isomorphisms, we have an isomorphism $\varphi_N \oplus \varphi_M[1] : \mathrm{cone}(f) \to \mathrm{cone}(f')$ of A-modules.*

Let (A, d_A), (B, d_B), and (C, d_C) be dgas. If (M, d_M) is an A-B-bimodule and (N, d_N) is an A-C-bimodule then let $\mathrm{Hom}^i_A(M, N)$ be the space of all graded left

$u(A)$-module maps $f : M \to N$ of degree i. We do not require that these maps commute with the differential. We define

$$\mathrm{Hom}_A(M, N) = \bigoplus_{i \in \mathbb{Z}} \mathrm{Hom}^i_A(M, N),$$

and this is a graded $u(B)$-$u(C)$-bimodule. We also have a version for right modules, which we write as $\mathrm{Hom}_{A^{\mathrm{op}}}(M, N)$.

Note the distinction between $\mathrm{Hom}_A(M, N)$ and the hom spaces in the category A-Mod. With the differential $d(f) = d_N \circ f - (-1)^i f \circ d_M$ for $f \in \mathrm{Hom}^i_A(M, N)$, $\mathrm{Hom}_A(M, N)$ becomes a B-C-bimodule. Similarly, if (M, d_M) is a B-A-bimodule and (N, d_N) is a C-A-bimodule, $\mathrm{Hom}_{A^{\mathrm{op}}}(M, N)$ is a C-B-bimodule. $\mathrm{Hom}_A(-, -)$ is the internal hom in the bimodule category, and we can recover the hom spaces in A-Mod as the 0-cycles of $\mathrm{Hom}_A(M, N)$.

If (M, d_M) is an A-B-bimodule and (N, d_N) is a B-C-bimodule then let $M \otimes_B N$ denote the space $M \otimes_{u(B)} N$. It is a graded $u(A)$-$u(C)$-bimodule: if $m \in M_i$ and $n \in N_j$ then $m \otimes n$ has degree $i + j$. With the differential $d_M \otimes \mathrm{id}_N + \mathrm{id}_M \otimes d_N$, it becomes an A-C-bimodule.

For an A-B-bimodule (M, d_M), we thus have functors

$$M \otimes_B - : B\text{-Mod} \to A\text{-Mod} \quad \text{and} \quad \mathrm{Hom}_A(M, -) : A\text{-Mod} \to B\text{-Mod}.$$

The functor $M \otimes_B -$ is left adjoint to $\mathrm{Hom}_A(M, -)$. For (N, d_N) a left A-module, the counit $\mathrm{ev}_N : M \otimes_B \mathrm{Hom}_A(M, N) \to N$ of the adjunction is the evaluation map, which acts as $x \otimes f \mapsto (-1)^{ij} f(m)$ for $x \in M_i$ and $f \in \mathrm{Hom}^j_A(M, N)$.

Derived categories. Our references are [Keller 1994; 2006].

If A is a graded vector space and d is a differential, i.e., a degree $+1$ endomorphism of A which satisfies $d^2 = 0$, then the i-th homology of A, denoted $H_i(A)$, is the subquotient $\ker d_i / \operatorname{im} d_{i-1}$, where $d_i : A_i \to A_{i+1}$ denotes the restriction of d to A_i. If (A, d) is a dga then the homology $H(A) = \bigoplus H_i(A)$ is a graded algebra, and if M is a left A-module then $H(M) = \bigoplus H_i(M)$ is a left $H(A)$-module. In fact, taking homology is a functor from the category of A-modules to the category of graded $H(A)$-modules. We say that a left A-module M is *acyclic* if $H(M) = 0$, and that a map $f : M \to N$ of A-modules is a *quasi-isomorphism* if $H(f)$ is an isomorphism.

The *category up to homotopy* of A-Mod, denoted $\mathrm{K}(A)$, is the $\mathbb{F}$-category whose objects are all left A-modules and whose morphism spaces, for $M, N \in A$-Mod, are $\mathrm{Hom}_{K(A)}(M, N) = H_0 \mathrm{Hom}_A(M, N)$. The *derived category* of A, denoted $\mathrm{D}(A)$, is the $\mathbb{F}$-category obtained by localizing $\mathrm{K}(A)$ at the full subcategory of acyclic A-modules. As a map of modules is a quasi-isomorphism if and only if its cone is acyclic, this is equivalent to localizing $\mathrm{K}(A)$ at the class of all quasi-isomorphisms. So we have a canonical functor $\mathrm{K}(A) \to \mathrm{D}(A)$, which we call the projection functor.

The finite-dimensional derived category, denoted $\mathrm{D}_{\mathrm{fd}}(A)$, is the full subcategory of $\mathrm{D}(A)$ on objects with finite-dimensional total homology, i.e., on A-modules M such that $H(M)$ is a finite-dimensional $\mathbb{F}$-vector space.

Let (A, d_A) be a dga. We say that

- $P \in A$-Mod is *indecomposable projective* if it is an indecomposable direct summand of the regular module,
- $P \in A$-Mod is *relatively projective* if it is a direct sum of shifts of indecomposable projective modules, and
- $P \in A$-Mod is *cofibrant* if, for each surjective quasi-isomorphism $f : M \to N$, the map $\mathrm{Hom}_{A\text{-Mod}}(P, f) : \mathrm{Hom}_{A\text{-Mod}}(P, M) \to \mathrm{Hom}_{A\text{-Mod}}(P, N)$ is surjective.

The following result characterizes cofibrant modules.

Proposition 4.6 [Keller 1994, Section 3; Keller and Yang 2011, Proposition 2.13]. *An A-module P is cofibrant if and only if it is an iterated extension of a relatively projective module by other relatively projective modules, possibly infinitely many times.*

Let A -cofib denote the full subcategory of $\mathrm{K}(A)$ on the cofibrant objects. The projection functor $\mathrm{K}(A) \to \mathrm{D}(A)$ induces an equivalence $A\text{-cofib} \xrightarrow{\sim} \mathrm{D}(A)$. Each A-module M has a *cofibrant replacement*, defined up to quasi-isomorphism and denoted $\boldsymbol{p}\, M$, which can be realized as the image of M under the left adjoint $\mathrm{D}(A) \to \mathrm{K}(A)$ to the canonical projection functor [Keller 2006, Proposition 3.1].

Let (B, d_B) be another dga and let $F : A\text{-Mod} \to B\text{-Mod}$ be an additive functor. Then F preserves chain homotopies, and so induces a functor $\mathrm{K}(F) : \mathrm{K}(A) \to \mathrm{K}(B)$. If $\mathrm{K}(F)$ preserves quasi-isomorphisms then, by the universal property of localization, it induces a functor $\mathrm{D}(F) : \mathrm{D}(A) \to \mathrm{D}(B)$. If $P \in A\text{-Mod-}B$ is cofibrant as a left A-module then, by [Keller 1994, Theorem 3.1(a)] and [Keller and Yang 2011, Proposition 2.13], $\mathrm{Hom}_A(P, -)$ preserves acyclic modules, and so preserves quasi-isomorphisms. By imitating the proof of [Keller 1994, Theorem 3.1(a)] we see that if $P \in A\text{-Mod-}B$ is cofibrant as a right B-module then $P \otimes_B -$ also preserves acyclic modules. We often write $P \otimes_B -$ and $\mathrm{Hom}_A(P, -)$, instead of $\mathrm{D}(P \otimes_B -)$ and $\mathrm{D}(\mathrm{Hom}_A(P, -))$, for the induced functors $\mathrm{D}(B) \to \mathrm{D}(A)$ and $\mathrm{D}(A) \to \mathrm{D}(B)$.

For an arbitrary $M \in A\text{-Mod-}B$, we get a functor $M \otimes_B^{\boldsymbol{L}} - : \mathrm{D}(B) \to \mathrm{D}(A)$, known as the *left derived functor* of $M \otimes_B -$, by composing the cofibrant replacement functor $\mathrm{D}(B) \to \mathrm{K}(B)$, the tensor functor $\mathrm{K}(M \otimes_B -) : \mathrm{K}(B) \to \mathrm{K}(A)$, and the projection functor $\mathrm{K}(A) \to \mathrm{D}(B)$. By [Keller 1994, Lemma 6.3(a)], we have an isomorphism $M \otimes_B^{\boldsymbol{L}} N \cong \boldsymbol{p}\, M \otimes_B N$ for all $N \in \mathrm{D}(B)$. The following basic, but useful, lemma says that this isomorphism is natural.

Lemma 4.7. *Let $M \in$* Mod-B.

(i) *We have a natural isomorphism of functors $\boldsymbol{p}\, M \otimes_B - \cong M \otimes_B^{\boldsymbol{L}} -$.*

(ii) *If M is cofibrant we have a natural isomorphism of functors $M \otimes_B - \cong M \otimes^{\boldsymbol{L}}_B -$.*

Proof. (i) We need to show that for each $N \in B$-Mod there is a quasi-isomorphism $\varphi_N : \boldsymbol{p}\, M \otimes_B N \to M \otimes_B \boldsymbol{p}\, N$ such that, for all maps $f : N \to N'$, the diagram

$$\begin{tikzcd}
\boldsymbol{p}\, M \otimes_B N \arrow[r, "\varphi_N"] \arrow[d, "\boldsymbol{p}\, M \otimes f"] & M \otimes_B \boldsymbol{p}\, N \arrow[d, "M \otimes \boldsymbol{p}\, f"] \\
\boldsymbol{p}\, M \otimes_B N' \arrow[r, "\varphi_{N'}"] & M \otimes_B \boldsymbol{p}\, N'
\end{tikzcd}$$

commutes. Consider the following diagram:

$$\begin{tikzcd}
& \boldsymbol{p}\, M \otimes_B \boldsymbol{p}\, N \arrow[dl, bend right, "\boldsymbol{p}\, M \otimes \pi_N"'] \arrow[dr, bend left, "\pi_M \otimes \boldsymbol{p}\, N"] & & & & \\
\boldsymbol{p}\, M \otimes_B N \arrow[rr, dashed, "\varphi_N"] \arrow[drrr, "\boldsymbol{p}\, M \otimes f"'] & & M \otimes_B \boldsymbol{p}\, N \arrow[drrr, "M \otimes \boldsymbol{p}\, f"] & & \boldsymbol{p}\, M \otimes_B \boldsymbol{p}\, N' \arrow[dl, bend right, "\boldsymbol{p}\, M \otimes \pi_{N'}"'] \arrow[dr, bend left, "\pi_M \otimes \boldsymbol{p}\, N'"] & \\
& & & \boldsymbol{p}\, M \otimes_B N' \arrow[rr, dashed, "\varphi_{N'}"] & & M \otimes_B \boldsymbol{p}\, N'
\end{tikzcd}$$

As $\boldsymbol{p}\, M$ and $\boldsymbol{p}\, N$ are cofibrant and π_M and π_N are quasi-isomorphisms, both $\boldsymbol{p}\, M \otimes \pi_N$ and $\pi_M \otimes \boldsymbol{p}\, N$ are quasi-isomorphisms, therefore we can define $\varphi_N = (\pi_M \otimes \boldsymbol{p}\, N) \circ (\boldsymbol{p}\, M \otimes \pi_N)^{-1}$ and it is a quasi-isomorphism. Then to check naturality we need to show that

$$(M \otimes \boldsymbol{p}\, f) \circ (\pi_M \otimes \boldsymbol{p}\, N) \circ (\boldsymbol{p}\, M \otimes \pi_N)^{-1} = (\pi_M \otimes \boldsymbol{p}\, N') \circ (\boldsymbol{p}\, M \otimes \pi_{N'})^{-1} \circ (\boldsymbol{p}\, M \otimes f).$$

By the bifunctoriality of the tensor product, the left-hand side is equal to

$$(\pi_M \otimes \boldsymbol{p}\, N') \circ (\boldsymbol{p}\, M \otimes \boldsymbol{p}\, f) \circ (\boldsymbol{p}\, M \otimes \pi_N)^{-1}$$

so we just need to show that

$$\boldsymbol{p}\, f \circ \pi_N^{-1} = \pi_{N'}^{-1} \circ f$$

but this follows from the functoriality $f \circ \pi_N = \pi_{N'} \circ \boldsymbol{p}\, f$ of the cofibrant replacement functor $\boldsymbol{p}$.

(ii) We just need to show that, for M cofibrant, there is a natural isomorphism $M \otimes_B - \cong \boldsymbol{p}\, M \otimes_B -$, and then the result will follow by part (i) of the lemma. This follows because $\pi_M : \boldsymbol{p}\, M \to M$ is a quasi-isomorphism and by the bifunctoriality of the tensor product. □

If the functor $M \otimes^{\boldsymbol{L}}_B -$ is an equivalence $\mathrm{D}(B) \xrightarrow{\sim} \mathrm{D}(A)$, we say that M is a *tilting module*.

We say that a module M is of *finite projective dimension* if its cofibrant replacement is an iterated extension of finitely many shifted indecomposable projective modules.

The following basic lemma will also be useful. It can be found as [Keller 1994, Lemma 6.2(a)]. We include a proof for the convenience of the reader.

Lemma 4.8. *Let M be a left A-module of finite projective dimension and let P be its cofibrant replacement. Then we have a natural isomorphism of functors*

$$\mathrm{Hom}_A(P, A) \otimes_A - \xrightarrow{\sim} \mathrm{Hom}_A(P, -) : A\text{-Mod} \to \mathbb{F}\text{-Mod}.$$

Proof. First note that, for any $P \in A$-Mod, we always have a natural transformation

$$\mathrm{Hom}_A(P, A) \otimes_A - \to \mathrm{Hom}_A(P, -)$$

obtained by starting with the map

$$\mathrm{ev} \otimes 1 : (P \otimes_{\mathbb{F}} \mathrm{Hom}_A(P, A)) \otimes_A M \to A \otimes_A M,$$

using the associativity isomorphism to obtain a map

$$P \otimes_{\mathbb{F}} (\mathrm{Hom}_A(P, A) \otimes_A M) \to A \otimes_A M,$$

then using the adjunction

$$\begin{aligned}&\mathrm{Hom}_{\mathbb{F}}(\mathrm{Hom}_A(P, A) \otimes_A M, \mathrm{Hom}_A(P, A \otimes_A M))\\ &\qquad\qquad \cong \mathrm{Hom}_A(P \otimes_{\mathbb{F}} (\mathrm{Hom}_A(P, A) \otimes_A M), A \otimes_A M),\end{aligned}$$

and finally using the natural isomorphism $A \otimes_A M \cong M$.

To show that our natural transformation is an isomorphism, we use induction on the number of times we need to extend a summand of the regular module to obtain P. We handle the base case as follows: the natural transformation is certainly an isomorphism when P is the regular module and so, as hom functors commute with finite direct sums, it is an isomorphism for all summands of the regular module. For our inductive step, suppose the lemma holds for P_1 and P_2, and let $P = \mathrm{cone}(f)$ for some map $f : P_1 \to P_2$. Then, for $M \in A$-Mod, one can check that the map $\mathrm{Hom}_A(P, A) \otimes_A M \to \mathrm{Hom}_A(P, M)$ comes from the commutative diagram

$$\begin{array}{ccc}\mathrm{Hom}_A(P_1, A) \otimes_A M & \xleftarrow{\mathrm{Hom}(f,A)\otimes M} & \mathrm{Hom}_A(P_2, A) \otimes_A M\\ \downarrow & & \downarrow\\ \mathrm{Hom}_A(P_1, M) & \xleftarrow{\mathrm{Hom}(f,M)} & \mathrm{Hom}_A(P_2, M)\end{array}$$

as in the construction from the second half of Lemma 4.5, where the vertical maps come from the natural transformation described above. Therefore, as both vertical maps are isomorphisms by induction, $\mathrm{Hom}_A(P, A) \otimes_A M \to \mathrm{Hom}_A(P, M)$ is an isomorphism. $\square$

Spherical twists. Our references are [Seidel and Thomas 2001; Rouquier and Zimmermann 2003; Grant 2013].

Let (A, d) be a dga and M be a left A-module with finite-dimensional total homology. Let $d \in \mathbb{Z}$. Following [Seidel and Thomas 2001], we say that M is d-spherical if

- M is a d-Calabi–Yau object, i.e., we have an isomorphism

  $$\mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(M, N) \cong \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(N, M[d])^*$$

  which is functorial in N, and
- $\bigoplus_{i \in \mathbb{Z}} \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(M, M[i])$ is isomorphic as a graded algebra to $\mathbb{F}[x]/\langle x^2 \rangle$, with x in degree d.

Associated to any spherical object M, we have a spherical twist functor $F_M : \mathrm{D}_{\mathrm{fd}}(A) \to \mathrm{D}_{\mathrm{fd}}(A)$ which is defined as follows. First, let $P = \boldsymbol{p}\, M$ be a cofibrant replacement of M. Then let X_M be the cone of the map of A-A-bimodules

$$P \otimes_{\mathbb{F}} \mathrm{Hom}_A(P, A) \xrightarrow{\mathrm{ev}} A,$$

where the nonzero map is the obvious evaluation map. As both $\mathrm{Hom}_A(P, A)$ and A are cofibrant, X_M is cofibrant as a right A-module. Then we define the spherical twist at M by

$$F_M = X_M \otimes_A - : \mathrm{D}_{\mathrm{fd}}(A) \to \mathrm{D}_{\mathrm{fd}}(A).$$

The spherical twist is an autoequivalence of $\mathrm{D}_{\mathrm{fd}}(A)$ (so X_M is a tilting module).

Note that, by Lemmas 4.5 and 4.8, if M has finite projective dimension then

$$F_M(N) \cong P \otimes_{\mathbb{F}} \mathrm{Hom}_A(P, N) \xrightarrow{\mathrm{ev}} N.$$

We next need the fact that spherical twists are intertwined by derived equivalences. The following is a generalization of [Seidel and Thomas 2001, Lemma 2.11].

Proposition 4.9. *Let A, B be dgas. Let $T \in B$-*Mod*-A be a tilting module and*

$$\Phi = T \otimes^{\boldsymbol{L}}_A - : \mathrm{D}_{\mathrm{fd}}(A) \to \mathrm{D}_{\mathrm{fd}}(B)$$

*be the associated derived equivalence. Let $M \in A$-*Mod *have finite-dimensional total homology and suppose it is d-spherical, for some $d \in \mathbb{Z}$. Suppose that $\Phi(M) \in B$-*Mod *has finite-dimensional total homology. Then $\Phi(M)$ is also d-spherical and we have an isomorphism of functors*

$$\Phi \circ F_M \cong F_{\Phi(M)} \circ \Phi : \mathrm{D}_{\mathrm{fd}}(A) \xrightarrow{\sim} \mathrm{D}_{\mathrm{fd}}(B).$$

In particular, we have an isomorphism

$$F_{\Phi(M)} \cong \Phi \circ F_M \circ \Phi^{-1} : \mathrm{D}_{\mathrm{fd}}(B) \xrightarrow{\sim} \mathrm{D}_{\mathrm{fd}}(B)$$

where Φ^{-1} is the quasi-inverse functor of Φ.

Proof. As $\Phi : \mathrm{D}_{\mathrm{fd}}(A) \to \mathrm{D}_{\mathrm{fd}}(B)$ is a derived equivalence it has quasi-inverse $\Phi^{-1} : \mathrm{D}_{\mathrm{fd}}(B) \to \mathrm{D}_{\mathrm{fd}}(A)$, and so we have isomorphisms

$$\mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(B)}(\Phi(M), \Phi(M)[i]) \cong \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(M, M[i])$$

and

$$\begin{aligned} \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(B)}(\Phi(M), N) &\cong \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(M, \Phi^{-1}(N)) \\ &\cong \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(\Phi^{-1}(N), M[d])^* \\ &\cong \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(B)}(N, \Phi(M)[d])^*, \end{aligned}$$

the second natural in $N \in \mathrm{D}_{\mathrm{fd}}(B)$, using the facts that M is a d-Calabi–Yau object and the shift functor $[d]$ commutes with all triangulated functors. Thus $\Phi(M)$ is d-spherical.

By Lemma 4.7 we may assume that T is cofibrant as a right B-module and that $\Phi = T \otimes_A -$. We want to show that

$$T \otimes_A X_M \otimes_A - \cong X_{\Phi(M)} \otimes_B T \otimes_A -,$$

so it is enough to check that we have an isomorphism

$$T \otimes_A X_M \cong X_{\Phi(M)} \otimes_B T$$

in $\mathrm{D}_{\mathrm{fd}}(B \otimes_{\mathbb{F}} A^{\mathrm{op}})$. To construct this isomorphism, we use the following extension of Lemma 4.5, which follows from the triangulated Five Lemma: for any commutative diagram

$$\begin{tikzcd} M \arrow[r, "f"] \arrow[d, "\varphi_M", "\sim"'] & N \arrow[d, "\varphi_N", "\sim"'] \\ M' \arrow[r, "f'"] & N' \end{tikzcd}$$

in B-Mod-A where φ_M and φ_N are both quasi-isomorphisms, we have a quasi-isomorphism $\varphi_N \oplus \varphi_{M[1]} : \mathrm{cone}(f) \to \mathrm{cone}(f')$.

As above, write $P = \boldsymbol{p}\, M$. Then, by the first part of Lemma 4.5, we just need to find two vertical maps which are quasi-isomorphisms and make the following diagram commute:

$$\begin{tikzcd} T \otimes_A P \otimes_{\mathbb{F}} \mathrm{Hom}_B(T \otimes_A P, B) \otimes_B T \arrow[r, "\mathrm{ev} \otimes 1_T"] \arrow[d, dashed, "\sim"'] & B \otimes_B T \arrow[d, dashed, "\sim"'] \\ T \otimes_A P \otimes_{\mathbb{F}} \mathrm{Hom}_A(P, A) \arrow[r, "1_T \otimes \mathrm{ev}"] & T \otimes_A A \end{tikzcd}$$

Our plan is to do this in stages: we will show that the vertical maps in the following diagram exist, and are quasi-isomorphisms, and that the diagram commutes.

$$\begin{tikzcd}
T \otimes_A P \otimes_{\mathbb{F}} \mathrm{Hom}_B(T \otimes_A P, B) \otimes_B T \arrow[r, "\mathrm{ev} \otimes 1_T"] \arrow[d, "\sim"'] & B \otimes_B T \arrow[d, "="'] \\
T \otimes_A P \otimes_{\mathbb{F}} \mathrm{Hom}_B(T \otimes_A P, B \otimes_B T) \arrow[r, "\mathrm{ev}"] \arrow[d, "\sim"'] & B \otimes_B T \arrow[d, "\sim"'] \\
T \otimes_A P \otimes_{\mathbb{F}} \mathrm{Hom}_B(T \otimes_A P, T \otimes_A A) \arrow[r, "\mathrm{ev}"] \arrow[d, "\sim"'] & T \otimes_A A \arrow[d, "="'] \\
T \otimes_A P \otimes_{\mathbb{F}} \mathrm{Hom}_A(P, A) \arrow[r, "1_T \otimes \mathrm{ev}"] & T \otimes_A A
\end{tikzcd}$$

Let us show that the first square commutes. We introduce some temporary notation for the rest of this proof. Let F and G denote the functors $F = T \otimes_A P \otimes_{\mathbb{F}} -$ and $G = \mathrm{Hom}_B(T \otimes_A P, -)$, so F is left adjoint to G, and let H denote the functor $- \otimes_B T$. Then we have unit and counit natural transformations $\varepsilon : FG \to 1$ and $\eta : 1 \to GF$, and a natural isomorphism $\zeta : HF \xrightarrow{\sim} FH$ coming from the associativity isomorphism for tensor products. We first need to define a map

$$\begin{aligned} HFGB = T \otimes_A P \otimes_F \mathrm{Hom}_B(T \otimes_A P, B) \otimes_B T \to T \otimes_A P \otimes_F \mathrm{Hom}_B(T \otimes_A P, B \otimes_B T) \\ = FGHB. \end{aligned}$$

We define this as the composite

$$HFGB \xrightarrow{\zeta GB} FHGB \xrightarrow{F\eta_{HGB}} FGFHGB \xrightarrow{FG\zeta^{-1}GB} FGHFGB \xrightarrow{FGH\varepsilon_B} FGHB.$$

One checks that this is an isomorphism using the same argument as in Lemma 4.8. To see that the diagram commutes, we break it up into smaller diagrams as follows:

$$\begin{tikzcd}
HFGB \arrow[rrrr, "H\varepsilon_B"] \arrow[d, "\zeta GB"', "\sim"] & & & & HB \\
FHGB \arrow[r, "1"] \arrow[d, "F\eta_{HGB}"'] & FHGB \arrow[drr, "\sim", "\zeta^{-1}GB"'] & & & \\
FGFHGB \arrow[ur, "\varepsilon_{FHGB}"'] \arrow[d, "FG\zeta^{-1}GB"', "\sim"] & & & HFGB \arrow[uur, "H\varepsilon_B"] & \\
FGHFGB \arrow[urrr, "\varepsilon_{HFGB}"] \arrow[rrrr, "FGH\varepsilon_B"] & & & & FGHB \arrow[uuu, "\varepsilon_{HB}"']
\end{tikzcd}$$

Now we see that both squares commute by the naturality of ε, the triangle commutes by the triangle identity $\varepsilon_F \circ F\eta = 1_F$, and the pentagon commutes because the isomorphisms are defined by ζ and its inverse.

We use the obvious composite isomorphism $B \otimes_B T \xrightarrow{\sim} T \xrightarrow{\sim} T \otimes_A A$ to define the second square. This commutes because the evaluation map is a counit, and therefore a natural transformation.

To show that the third square commutes, we introduce some more notation. Let F' and G' denote the functors $F' = P \otimes_{\mathbb{F}} -$ and $G' = \mathrm{Hom}_A(P, -)$, so F' is left adjoint to G', and let H' and I' denote the functors $H' = T \otimes_A -$ and $I' = \mathrm{Hom}_B(T, -)$, so H' is left adjoint (in fact, quasi-inverse) to I'. We denote the counit and unit maps of the first adjunction by $\varepsilon' : F'G' \to 1$ and $\eta' : 1 \to G'F'$, and of the second adjunction by $\varepsilon'' : H'I' \to 1$ and $\eta'' : 1 \to I'H'$. Note that, because H' induces an equivalence of derived categories, ε'' and η'' give quasi-isomorphisms when applied to any object.

Using the associativity isomorphism for tensor products we have a natural isomorphism of functors $F \cong H'F'$, and by the uniqueness of right adjoints (or by using the tensor-hom adjunctions directly) this gives another natural isomorphism $G \cong G'I'$.

We now redraw our final square, breaking it up into smaller diagrams:

$$\begin{tikzcd}
FGH'A \arrow[rr, "\varepsilon'_{H'A}"] \arrow[d, "\sim"] & & H'A \\
H'F'G'I'H'A \arrow[r, "H'\varepsilon'_{I'H'A}"] & H'I'H'A \arrow[ur, "\varepsilon'' H'A"] & \\
H'F'G'A \arrow[u, "H'F'G'\eta''_A"] \arrow[rr, "H'\varepsilon'_A"] & & H'A \arrow[ul, "H'\eta''_A"'] \arrow[uu, "1"']
\end{tikzcd}$$

Here, the top square commutes by definition of the isomorphisms $F \cong H'F'$ and $G \cong G'I'$, the triangle commutes by the triangle identity $\varepsilon'' H' \circ H'\eta'' = 1_{H'}$, and the bottom square commutes by the naturality of ε'. $\square$

We now describe the braid relations for spherical twists, as in Propositions 2.12 and 2.13 of [Seidel and Thomas 2001]; see also [Rouquier and Zimmermann 2003; Grant 2015].

Proposition 4.10. *Suppose that M and N are spherical objects of $\mathrm{D_{fd}}(A)$ and let*

$$(M, N) = \dim_{\mathbb{F}} \bigoplus_{n \in \mathbb{Z}} \mathrm{Hom}_{\mathrm{D_{fd}}(A)}(M, N[n]).$$

Let $F_M, F_N : \mathrm{D_{fd}}(A) \xrightarrow{\sim} \mathrm{D_{fd}}(A)$ be the associated spherical twists.

- *If $(M, N) = 0$ then $F_M \circ F_N \cong F_N \circ F_M$.*
- *If $(M, N) = 1$ then $F_M \circ F_N \circ F_M \cong F_N \circ F_M \circ F_N$.*

Ginzburg dg-algebras. There is a well-known method to associate a differential graded algebra to a quiver with potential [Ginzburg 2007, Section 5; Keller and Yang 2011, Section 2.6].

Let (Q, W) be a quiver with potential. Construct a new quiver $\bar{Q}$ by adding arrows to Q: for each arrow $a : i \to j$ in Q we add a new arrow $a^* : j \to i$, and for each vertex i in Q we add a new arrow $t_i : i \to i$. We view $\bar{Q}$ as a graded quiver with the arrows of Q in degree 0, the arrows a^* in degree -1, and the arrows t_i in degree -2. This induces a grading on the path algebra $\mathbb{F}\bar{Q}$ of $\bar{Q}$ such that the degree 0 part $(\mathbb{F}\bar{Q})_0$ is just the path algebra of $\mathbb{F}Q$ of Q. Let $\bar{J}$ denote the ideal of $\mathbb{F}\bar{Q}$ generated by the arrows of $\bar{Q}$, and let $\widehat{\mathbb{F}\bar{Q}}$ denote the completion of the graded algebra $\mathbb{F}\bar{Q}$ with respect to $\bar{J}$, as on page 96.

We define a differential d on $\widehat{\mathbb{F}\bar{Q}}$ by requiring that d be zero on each idempotent e_i associated to a vertex i of Q, specifying how d acts on arrows of $\bar{Q}$, and then extending to the rest of $\widehat{\mathbb{F}\bar{Q}}$ using the Leibniz rule and continuity. For degree reasons, we must have $d(a) = 0$ for each arrow a of Q. For arrows a^*, we set $d(a^*) = \partial_a W$, where ∂_a denotes the cyclic derivative, and for arrows t_i we set $d(t_i) = e_i\big(\sum aa^* - a^*a\big)e_i$, where we sum over all arrows a of Q. Then $\Gamma_{Q,W} = (\widehat{\mathbb{F}\bar{Q}}, d)$ is called the *Ginzburg dga* of (Q, W).

Note that if (Q_1, W_1) and (Q_2, W_2) are right equivalent, then we have an isomorphism of dgas $\Gamma_{Q_1,W_1} \cong \Gamma_{Q_2,W_2}$ [Keller and Yang 2011, Lemma 2.9]. Hence, if we are working with quivers with potential of mutation type A or D, by Proposition 4.4 we only need to consider the Ginzburg dgas Γ_{Q,W_Q}, and so can denote them Γ_Q.

Keller and Yang showed that QP-mutation lifts to equivalences of derived categories of Ginzburg dgas:

Theorem 4.11 [Keller and Yang 2011, Theorem 3.2]. *Suppose that (Q, W) is a QP and that $(Q', W') = \mu_k(Q, W)$ for some $k \in I$. There is a tilting complex T which gives an equivalence of triangulated categories*

$$\mu_k = \mathrm{Hom}_{\Gamma_{Q',W'}}(T, -) : \mathrm{D}(\Gamma_{Q',W'}) \to \mathrm{D}(\Gamma_{Q,W}),$$

and it restricts to an equivalence of triangulated categories,

$$\mu_k = \mathrm{Hom}_{\Gamma_{Q',W'}}(T, -) : \mathrm{D}_{\mathrm{fd}}(\Gamma_{Q',W'}) \to \mathrm{D}_{\mathrm{fd}}(\Gamma_{Q,W}).$$

Recall that, for a dga A, the finite-dimensional derived category $\mathrm{D}_{\mathrm{fd}}(A)$ is d-Calabi–Yau if there exists a bifunctorial isomorphism,

$$\mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(M, N) \cong \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(N, M[d])^*,$$

where $(-)^*$ denotes the k-linear dual. We will need the following important result of Keller and Van den Bergh:

Theorem 4.12 [Keller 2011, Theorem 6.3 and Theorem A.12]. *The category $\mathrm{D}_{\mathrm{fd}}(\Gamma_{Q,W})$ is 3-Calabi–Yau.*

Let (Q, W) be a QP and $\Gamma = \Gamma_{Q,W}$. Associated to each vertex i of Q, we have the 1-dimensional simple left Γ-module, which we denote S_i. Keller and Yang

[2011, Section 2.14] explained how to construct the cofibrant replacement of S_i: as long as we remember the differential, we can proceed as if the Ginzburg dga were an ordinary hereditary algebra, and the underlying $u(\Gamma)$-module of $\boldsymbol{p}\, S_i$ is the direct sum of one copy of the projective P_i and one copy of the shifted projective $P_j[1]$ for each arrow $j \to i$ in Q. Using this, they showed:

Lemma 4.13 [Keller and Yang 2011, Lemma 2.15]. *Let $i, j \in I$ and $n \in Z$ and $\Gamma = \Gamma_{Q,W}$. Then* $\mathrm{Hom}_{\mathrm{D_{fd}}(\Gamma)}(S_i, S_j[n]) = 0$ *if $n \neq 0, 1, 2, 3$, and*

$$\dim_{\mathbb{F}} \mathrm{Hom}_{\mathrm{D_{fd}}(\Gamma)}\big(S_i, S_j[n]\big) = \begin{cases} \delta_{ij} & \text{if } n = 0, \\ \#\{\text{arrows } i \to j \text{ in } Q\} & \text{if } n = 1, \\ \#\{\text{arrows } j \to i \text{ in } Q\} & \text{if } n = 2, \\ \delta_{ij} & \text{if } n = 3, \end{cases}$$

where δ_{ij} is the Kronecker delta.

Relations between functors. By Theorem 4.12, every object of $\mathrm{D_{fd}}(\Gamma_{Q,W})$ is a 3-Calabi–Yau object. By Lemma 4.13,

$$\bigoplus_{j \in \mathbb{Z}} \mathrm{Hom}_{\mathrm{D_{fd}}(\Gamma_{Q,W})}(S_i, S_i[j]) \cong \mathbb{F}[x]/\langle x^2 \rangle$$

with x in degree 3. Hence S_i is 3-spherical, and we have a spherical twist F_{S_i} associated to S_i. We will sometimes write F_i instead of F_{S_i}.

Let k be a vertex of Q, and write $(Q', W') = \mu_k(Q, W)$. Then write $\Gamma = \Gamma_{Q,W}$ and $\Gamma' = \Gamma_{Q',W'}$ for the associated Ginzburg dgas. Write T_i for the left Γ'-module associated to the vertex i of Q' and G_i for the associated autoequivalence F_{T_i} of $\mathrm{D_{fd}}(\Gamma')$. In this section we will study how the spherical twists $F_i : \mathrm{D_{fd}}(\Gamma) \xrightarrow{\sim} \mathrm{D_{fd}}(\Gamma)$ interact with the mutation functors $\mu_k : \mathrm{D_{fd}}(\Gamma') \xrightarrow{\sim} \mathrm{D_{fd}}(\Gamma)$. Our key tools will be Proposition 4.9 and the results on the images of the simple modules under the mutation functors [Keller and Yang 2011, Lemma 3.12(a)], which we will describe below.

If A is a dga and $M, N \in \mathrm{D_{fd}}(A)$, we have a natural map,

$$M \otimes_{\mathbb{F}} \mathrm{Hom}_{\mathrm{D_{fd}}(A)}(M, N) \to N,$$

in $\mathrm{D_{fd}}(A)$ given by evaluation. For any graded vector space V, we have biadjoint functors $- \otimes_F V$ and $- \otimes_F V^*$, and these respect the left A-module structure, so we also obtain a natural map,

$$M \to N \otimes_{\mathbb{F}} \mathrm{Hom}_{\mathrm{D_{fd}}(A)}(M, N)^*,$$

in $\mathrm{D_{fd}}(A)$. Now let $L, N \in A$-Mod. The *universal extension of N by L* is the cone of the natural map

$$N[-1] \to L \otimes_{\mathbb{F}} \mathrm{Hom}_{\mathrm{D_{fd}}(A)}(N[-1], L)^*$$

and the *universal coextension of L by N* is the cone of the natural map

$$N[-1] \otimes_{\mathbb{F}} \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(A)}(N[-1], L) \to L.$$

The following result is due to Keller and Yang:

Lemma 4.14 [Keller and Yang 2011, Lemma 3.12(a)]. *We have $\mu_k(T_k) \cong S_k[1]$ and $\mu_k(T_i)$ is the universal extension of S_i by S_k, for $i \neq k$.*

The following result should be compared to Definition 2.8.

Proposition 4.15. *If Q has no double arrows then we have a natural isomorphism of functors*

$$F_{\mu_k^{-1}(S_i)} \cong \begin{cases} G_k G_i G_k^{-1} & \text{if } i \to k \text{ in } Q, \\ G_i & \text{otherwise.} \end{cases}$$

Proof. We first use Lemma 4.14 to calculate the images of the simple Γ'-modules under the inverse mutation functor μ^{-1}, where $\mu = \mu_k$. We know that $\mu(T_k) \cong S_k[1]$, so $\mu^{-1}(S_k) \cong T_k[-1]$. By assumption, there is at most one arrow between any two vertices in Q. If $i \neq k$ and there is no arrow $i \to k$ in Q then, by Lemma 4.14, $\mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(\Gamma)}(S_i[-1], S_k) = 0$ and so $\mu(T_i) \cong \mathrm{cone}(S_i[-1] \to 0) \cong S_i$, so $\mu^{-1}(S_i) \cong T_i$.

If $i \neq k$ and there is an arrow $i \to k$ in Q then $\mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(\Gamma)}(S_i[-1], S_k)$ is 1-dimensional and so $\mu(T_i) \cong \mathrm{cone}(S_i[-1] \to S_k)$, with the nonzero map determined up to a scalar. We can then use Lemma 4.5 to calculate $\mu(\mathrm{cone}(T_k[-1] \to T_i)$: this is $\mathrm{cone}(\mu(T_k)[-1] \to \mu(T_i))$ where, as μ is an equivalence, the map must again be nonzero and determined up to scalar. We know that $\mu(T_k)[-1] \cong S_k$ and $\mu(T_i)$ is $S_i \oplus S_k$ with appropriate differential. One can check that the injection $S_k \hookrightarrow S_i \oplus S_k$ respects the differentials, and so this must be our nonzero map. This is quasi-isomorphic to the map $0 \to S_i$, and so $\mu(\mathrm{cone}(T_k[-1] \to T_i)) \cong S_i$ and hence $\mu^{-1}(S_i) \cong \mathrm{cone}(T_k[-1] \to T_i)$. Note that this is the universal coextension of T_i by T_k.

Now we check that the formula holds. If $i = k$ then $F_{\mu^{-1}(S_i)} = F_{T_i[1]}$, and as the shift functor on $\mathrm{D}_{\mathrm{fd}}(\Gamma')$ is naturally isomorphic to $\Gamma'[1] \otimes_{\Gamma'} -$ we see that, by Proposition 4.9, $F_{T_i[1]} \cong [1] \circ G_i \circ [-1] \cong G_i$. If $i \neq k$ and there is no arrow $i \to k$ in Q then $\mu^{-1}(S_i) \cong T_i$ so $F_{\mu^{-1}(T_i)} = G_i$.

Finally, suppose $i \neq k$ and there is an arrow $i \to k$ in Q. As mutation at k reverses all arrows incident with k, and can never change the number of arrows incident with k, there must be exactly one arrow $k \to i$ in Q'. We first calculate $G_k(T_i)$: this is

$$\mathrm{cone}(\boldsymbol{p}\, T_k \otimes_{\mathbb{F}} \mathrm{Hom}_{\Gamma'}(\boldsymbol{p}\, T_k, T_i) \to T_i).$$

As $\mathrm{Hom}_{\Gamma'}(\boldsymbol{p}\, T_k, T_i)$ is a differential graded $\mathbb{F}$-module, it is quasi-isomorphic to its homology, which is the direct sum $\bigoplus \mathrm{Hom}_{K(\Gamma')}(T_k, T_i[n])$ with

$$\mathrm{Hom}_{K(\Gamma')}(T_k, T_i[n]) \cong \mathrm{Hom}_{\mathrm{D}_{\mathrm{fd}}(\Gamma')}(T_k, T_i[n])$$

in degree n. So by Lemma 4.14 the homology is only nonzero in degree 1, where it is 1-dimensional, and thus

$$G_k(T_i) \cong \operatorname{cone}(T_k \otimes_{\mathbb{F}} \mathbb{F}[-1] \to T_i).$$

So we see that $\mu^{-1}(S_i) \cong G_k(T_i)$ and therefore, using Proposition 4.9 again, $F_{\mu^{-1}(S_i)} \cong G_k G_i G_k^{-1}$. □

We are now able to show that our braid groups B_Q act via spherical twists on the category $\mathrm{D}_{\mathrm{fd}}(\Gamma)$.

Theorem 4.16. *Let (Q, W) be a mutation-Dynkin quiver with potential of type ADE. Then we have a group homomorphism*

$$B_Q \to \operatorname{Aut} \mathrm{D}_{\mathrm{fd}}(\Gamma_{Q,W}), \qquad s_i \mapsto F_i$$

sending the group generator associated to the vertex $i \in I$ to the spherical twist at the simple $\Gamma_{Q,W}$-module S_i.

Proof. As (Q, W) is mutation-Dynkin, it is obtained by mutating a quiver with potential $(Q'', 0)$ finitely many times, where Q'' is a Dynkin quiver. Then we have a group homomorphism $B_{Q''} \to \operatorname{Aut} \mathrm{D}_{\mathrm{fd}}(\Gamma_{Q'',0})$ by Remark 2.3, Proposition 4.10, and Lemma 4.13. This gives the base case of an inductive argument, so we need to show that if the spherical twists F_i on $\Gamma = \Gamma_{Q,W}$ satisfy the relations of B_Q for a mutation-Dynkin quiver with potential (Q, W) then the spherical twists G_i on $\Gamma' = \Gamma_{Q',W'}$ satisfy the relations of $B_{Q'}$.

Assume the functors $F_i : \mathrm{D}_{\mathrm{fd}}(\Gamma) \to \mathrm{D}_{\mathrm{fd}}(\Gamma)$ satisfy the relations of B_Q and let $\mu = \mu_k : \mathrm{D}_{\mathrm{fd}}(\Gamma') \to \mathrm{D}_{\mathrm{fd}}(\Gamma)$ be the Keller–Yang derived equivalence. Then the functors $\mu^{-1} \circ F_i \circ \mu : \mathrm{D}_{\mathrm{fd}}(\Gamma') \to \mathrm{D}_{\mathrm{fd}}(\Gamma')$ also satisfy the relations of B_Q. By Proposition 4.9 we have

$$\mu^{-1} \circ F_i \circ \mu \cong F_{\mu^{-1}(S_i)},$$

i.e., the following diagram commutes:

$$\begin{tikzcd} \mathrm{D}_{\mathrm{fd}}(\Gamma') \arrow[r, "\mu"] \arrow[d, "F_{\mu^{-1}(S_i)}"] & \mathrm{D}_{\mathrm{fd}}(\Gamma) \arrow[d, "F_{S_i}"] \\ \mathrm{D}_{\mathrm{fd}}(\Gamma') \arrow[r, "\mu"] & \mathrm{D}_{\mathrm{fd}}(\Gamma) \end{tikzcd}$$

So we have a group homomorphism $\rho : B_Q \xrightarrow{\rho} \operatorname{Aut} \mathrm{D}_{\mathrm{fd}}(\Gamma')$ sending s_i to $F_{\mu^{-1}(S_i)}$. By Proposition 2.1, Q has no double arrows, so we can use Proposition 4.15 to write ρ as

$$\begin{aligned} B_Q &\xrightarrow{\rho} \operatorname{Aut} \mathrm{D}_{\mathrm{fd}}(\Gamma'), \\ s_i &\mapsto \begin{cases} G_k G_i G_k^{-1} & \text{if } i \to k \text{ in } Q; \\ G_i & \text{otherwise.} \end{cases} \end{aligned}$$

Precomposing this with the group isomorphism $\varphi_k^{-1} : B'_Q \xrightarrow{\sim} B_Q$ of Remark 2.11, we obtain the group homomorphism

$$\begin{array}{ccccc} B_{Q'} & \xrightarrow{\varphi_k^{-1}} & B_Q & \xrightarrow{\qquad\qquad\rho\qquad\qquad} & \operatorname{Aut} \mathrm{D}_{\mathrm{fd}}(\Gamma') \end{array}$$

$$t_i \longmapsto \begin{Bmatrix} s_k^{-1} s_i s_k & \text{if } i \to k \text{ in } Q, \\ s_i & \text{otherwise} \end{Bmatrix} \mapsto \begin{Bmatrix} G_k^{-1} G_k G_i G_k^{-1} G_k & \text{if } i \to k \text{ in } Q, \\ G_i & \text{otherwise} \end{Bmatrix} \cong G_i$$

as required. □

Remark 4.17. Known results on the faithfulness of braid group actions can be transferred to our setting. Suppose Q'' is an orientation of an *ADE* graph and the usual action $B_{Q''} \to \operatorname{Aut} \mathrm{D}_{\mathrm{fd}}(\Gamma_{Q'',0})$ is faithful. From the proof of Theorem 4.16 we see that our actions of B_Q where Q is of mutation type *ADE* are just built by precomposing group isomorphisms with the action of $B_{Q''}$, and so these are also faithful under this assumption.

It was shown by Seidel and Thomas [2001, Theorem 2.18], building on work of Khovanov and Seidel [2002], that given a collection of d-spherical objects, with $d \geq 2$, in a type A_n-configuration the action of the braid group by spherical twists is faithful. Thus the actions of Theorem 4.16 are faithful in mutation type A. The faithfulness result of Seidel and Thomas was extended to all collections of 2-spherical objects in type *ADE* configurations by Brav and Thomas [2011], using the Garside structure of the braid monoid, but it is not immediately clear how to generalize their argument to the 3-Calabi–Yau situation.

Remark 4.18. Although we have shown that our braid groups of mutation-Dynkin quivers can be realized categorically, this is not a categorification of our earlier results because we cannot decategorify (see, for example, [Baez and Dolan 1998]): we cannot recover Theorem 2.10 from Theorem 4.16 because we use Theorem 2.10 to prove Theorem 4.16. The problem is that, for an arbitrary mutation-Dynkin quiver with potential (Q, W), we do not in advance know the relations satisfied by the spherical twist functors F_i. This question will be addressed in a forthcoming paper.

Remark 4.19. The arguments of [Grant 2015] generalize to show that, if vertices i and j of Q are joined by an arrow, then $F_i F_j F_i \cong F_j F_i F_j$ can be realized as a single periodic twist. Similarly, one can show that if $i \to j \to k \to i$ is a 3-cycle in

Q then

$$F_1F_2F_3F_1 \cong F_2F_3F_1F_2 \cong F_3F_1F_2F_3$$

can be realized as a single periodic twist. This exhausts the possibilities in type A; we will study type D further in the future.

Acknowledgements

Marsh would like to thank K. Baur and A. King for useful discussions concerning orbifold interpretations of cluster algebras. Both authors would like to thank the referee, and I. Webster, for helpful comments on an earlier version of the paper. Some of the work for this paper was done while Marsh was Guest Professor in 2014 at the Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, N-7491, Norway.

References

[Allcock 2002] D. Allcock, "Braid pictures for Artin groups", *Trans. Amer. Math. Soc.* **354**:9 (2002), 3455–3474. MR Zbl

[Amiot 2009] C. Amiot, "Cluster categories for algebras of global dimension 2 and quivers with potential", *Ann. Inst. Fourier (Grenoble)* **59**:6 (2009), 2525–2590. MR Zbl

[Artin 1925] E. Artin, "Theorie der Zöpfe", *Abh. Math. Sem. Univ. Hamburg* **4**:1 (1925), 47–72. MR JFM

[Baez and Dolan 1998] J. C. Baez and J. Dolan, "Categorification", pp. 1–36 in *Higher category theory* (Evanston, IL, 1997), edited by E. Getzler and M. Kapranov, Contemp. Math. **230**, American Mathematical Society, Providence, RI, 1998. MR Zbl

[Barot and Marsh 2015] M. Barot and B. R. Marsh, "Reflection group presentations arising from cluster algebras", *Trans. Amer. Math. Soc.* **367**:3 (2015), 1945–1967. MR Zbl

[Brav and Thomas 2011] C. Brav and H. Thomas, "Braid groups and Kleinian singularities", *Math. Ann.* **351**:4 (2011), 1005–1017. MR Zbl

[Buan et al. 2006] A. B. Buan, B. R. Marsh, M. Reineke, I. Reiten, and G. Todorov, "Tilting theory and cluster combinatorics", *Adv. Math.* **204**:2 (2006), 572–618. MR Zbl

[Caldero et al. 2006] P. Caldero, F. Chapoton, and R. Schiffler, "Quivers with relations arising from clusters (A_n case)", *Trans. Amer. Math. Soc.* **358**:3 (2006), 1347–1364. MR Zbl

[Derksen et al. 2008] H. Derksen, J. Weyman, and A. Zelevinsky, "Quivers with potentials and their representations, I: Mutations", *Selecta Math.* (*N.S.*) **14**:1 (2008), 59–119. MR Zbl

[Felikson et al. 2012] A. Felikson, M. Shapiro, and P. Tumarkin, "Cluster algebras and triangulated orbifolds", *Adv. Math.* **231**:5 (2012), 2953–3002. MR Zbl

[Fomin and Zelevinsky 2002] S. Fomin and A. Zelevinsky, "Cluster algebras, I: Foundations", *J. Amer. Math. Soc.* **15**:2 (2002), 497–529. MR Zbl

[Fomin and Zelevinsky 2003] S. Fomin and A. Zelevinsky, "Cluster algebras, II: Finite type classification", *Invent. Math.* **154**:1 (2003), 63–121. MR Zbl

[Fomin et al. 2008] S. Fomin, M. Shapiro, and D. Thurston, "Cluster algebras and triangulated surfaces, I: Cluster complexes", *Acta Math.* **201**:1 (2008), 83–146. MR Zbl

[Fox and Neuwirth 1962] R. Fox and L. Neuwirth, "The braid groups", *Math. Scand.* **10** (1962), 119–126. MR Zbl

[Ginzburg 2007] V. Ginzburg, "Calabi–Yau algebras", preprint, 2007. arXiv:math/0612139v3

[Grant 2013] J. Grant, "Derived autoequivalences from periodic algebras", *Proc. Lond. Math. Soc.* (3) **106**:2 (2013), 375–409. MR Zbl

[Grant 2015] J. Grant, "Lifts of longest elements to braid groups acting on derived categories", *Trans. Amer. Math. Soc.* **367**:3 (2015), 1631–1669. MR Zbl

[Haley et al. 2014] J. Haley, D. Hemminger, A. Landesman, and H. Peck, "Artin group presentations arising from cluster algebras", preprint, 2014. arXiv

[Humphreys 1990] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge Studies in Advanced Mathematics **29**, Cambridge Univ. Press, 1990. MR Zbl

[Kassel and Turaev 2008] C. Kassel and V. Turaev, *Braid groups*, Graduate Texts in Mathematics **247**, Springer, New York, 2008. MR Zbl

[Keller 1994] B. Keller, "Deriving DG categories", *Ann. Sci. École Norm. Sup.* (4) **27**:1 (1994), 63–102. MR Zbl

[Keller 2006] B. Keller, "On differential graded categories", pp. 151–190 in *International Congress of Mathematicians*, vol. II, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl

[Keller 2011] B. Keller, "Deformed Calabi–Yau completions", *J. Reine Angew. Math.* **654** (2011), 125–180. With an appendix by Michel Van den Bergh. MR Zbl

[Keller and Yang 2011] B. Keller and D. Yang, "Derived equivalences from mutations of quivers with potential", *Adv. Math.* **226**:3 (2011), 2118–2168. MR Zbl

[Khovanov and Seidel 2002] M. Khovanov and P. Seidel, "Quivers, Floer cohomology, and braid group actions", *J. Amer. Math. Soc.* **15**:1 (2002), 203–271. MR Zbl

[Labardini-Fragoso 2009] D. Labardini-Fragoso, "Quivers with potentials associated to triangulated surfaces", *Proc. Lond. Math. Soc.* (3) **98**:3 (2009), 797–839. MR Zbl

[Labardini-Fragoso 2016] D. Labardini-Fragoso, "Quivers with potentials associated to triangulated surfaces, IV: Removing boundary assumptions", *Selecta Math.* (*N.S.*) **22**:1 (2016), 145–189. MR Zbl

[Nagao 2010] K. Nagao, "Triangulated surface, mapping class group and Donaldson–Thomas theory", lecture notes, Graduate School of Mathematics, Nagoya University, 2010, available at http://www.ms.u-tokyo.ac.jp/~kazushi/kinosaki2010/nagao.pdf.

[Qiu 2016] Y. Qiu, "Decorated marked surfaces: spherical twists versus braid twists", *Math. Ann.* **365**:1-2 (2016), 595–633. MR Zbl

[Rouquier and Zimmermann 2003] R. Rouquier and A. Zimmermann, "Picard groups for derived module categories", *Proc. London Math. Soc.* (3) **87**:1 (2003), 197–225. MR Zbl

[Schiffler 2008] R. Schiffler, "A geometric model for cluster categories of type D_n", *J. Algebraic Combin.* **27**:1 (2008), 1–21. MR Zbl

[Seidel and Thomas 2001] P. Seidel and R. Thomas, "Braid group actions on derived categories of coherent sheaves", *Duke Math. J.* **108**:1 (2001), 37–108. MR Zbl

[Sergiescu 1993] V. Sergiescu, "Graphes planaires et présentations des groupes de tresses", *Math. Z.* **214**:3 (1993), 477–490. MR

[Thurston 2012] D. Thurston, "The geometry and algebra of curves on surfaces", course notes, Department of Mathematics, University of California, Berkeley, 2012, available at http://tinyurl.com/qiaochu. Notes taken by Qiaochu Yuan.

Received March 28, 2016. Revised January 13, 2017.

JOSEPH GRANT
SCHOOL OF MATHEMATICS
UNIVERSITY OF EAST ANGLIA
NORWICH RESEARCH PARK
NORWICH
NR4 7TJ
UNITED KINGDOM
j.grant@uea.ac.uk

BETHANY ROSE MARSH
SCHOOL OF MATHEMATICS
UNIVERSITY OF LEEDS
LEEDS
LS2 9JT
UNITED KINGDOM
marsh@maths.leeds.ac.uk

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.117

A PALEY–WIENER THEOREM FOR THE SPECTRAL PROJECTION OF SYMMETRIC GRAPHS

SHIN KOIZUMI

We prove a Paley–Wiener theorem for the spectral projection of symmetric graphs and, as a corollary, derive a Paley–Wiener theorem for the Helgason–Fourier transform. The proof is based on contour integration arguments similar to those used to prove the Paley–Wiener theorem for Euclidean spaces and symmetric spaces.

1. Introduction

The theory of representations of free groups has been studied by many authors in analogy with the semisimple theory. This arises from the realization of a free group as a homogeneous tree and relies upon the use of the Poisson boundary and spherical function. Mantero and Zappa [1983] characterized the image of the Poisson transform of free groups and studied the uniform boundedness of the spherical representation. In [Cowling et al. 1998], Cowling, Meda and Setti studied the images of the Abel transform for various function spaces on homogeneous trees. Cowling and Setti [1999] gave the characterizations of the images of the spaces of compactly supported functions and rapidly decreasing functions.

The concept of tree has been extended in several aspects. For instance, Iozzi and Picardello [1983a; 1983b] extended the context of tree to symmetric graphs and gave an explicit expression of the spherical function. Later, the Plancherel measure on symmetric graphs was explicitly computed in [Kuhn and Soardi 1983; Faraut and Picardello 1984]. Recently Eddine [2013; 2015] investigated the characterization of the Abel transform for symmetric graphs and, as an application, solved the shifted wave equations on it.

In [Koizumi 2013], we studied the spectral projection on homogeneous trees and proved the Paley–Wiener theorem of the spectral projection, which is an analogue of that given by Bray [1996]. In this paper, we shall extend the works in [Koizumi 2013] to the case of symmetric graphs. Unlike the works of Cowling and Setti

MSC2010: primary 43A80; secondary 43A85.
Keywords: spectral projection, symmetric graph, Paley–Wiener theorem, generalized spherical function.

[1999], our proof is based on contour integration arguments, which are usually used to prove the Paley–Wiener theorem for the cases of the Euclidean spaces and the symmetric spaces [Johnson 1979; Campoli 1980].

A brief outline of this paper is as follows: Section 2 is devoted to the overview of the notation of symmetric graphs. In Section 3, we concretely write down the expressions of the Poisson transform on symmetric graphs. In Section 4, we construct the intertwining operators between the spherical representations and give the explicit expressions of the intertwining operators. In Section 5, we study the properties of the spectral projection for symmetric graphs. Finally in Section 6, we show the Paley–Wiener theorem of the spectral projection and prove the Paley–Wiener theorem of the Helgason–Fourier transform.

2. Notation and preliminaries

The standard symbols $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$ are used for the integers, the real numbers and the complex numbers, respectively. Let us set $\mathbb{Z}_k = \mathbb{Z}/k\mathbb{Z}$. Throughout this paper, the imaginary unit is denoted as i. If $x \in \mathbb{C}$, $\Re x$ and $\Im x$ denote its real part and its imaginary part, respectively.

A graph $\mathfrak{X}$ is symmetric of type $k \geq 2$ and order $r \geq 2$ if every vertex v belongs exactly to r polygons with k sides each with no sides and no vertex in common except v, and if every nontrivial loop in $\mathfrak{X}$ runs through all edges of at least one polygon. If $k = 2$, $\mathfrak{X}$ reduces to a homogeneous tree of degree r. In what follows, we write $q = (k-1)(r-1)$, $\tau = 2\pi/\log q$ and $\mathbb{T} = \mathbb{R}/\tau\mathbb{Z}$. Different notions of length on a symmetric graph were introduced in [Iozzi and Picardello 1983a]. Here we use the definition of the length $d(x, y)$ between two vertices $x, y \in \mathfrak{X}$ to denote the minimal number of polygons crossed by a path connecting x and y. We fix a reference point o in $\mathfrak{X}$ and write $|x| = d(x, o)$.

By the same arguments as in [Betori and Pagliacci 1984, Theorem 1], if $k > 2$, it is easy to see that every group acting simply transitively on $\mathfrak{X}$ and isometrically with respect to the metric induced by this length is isometric to the free group $G = \oplus_{i=1}^{r} \mathbb{Z}_k$, while, for $k = 2$, G is isometric to the free product of t copies of $\mathbb{Z}$ and s copies of $\mathbb{Z}_2$, where $2t + s = r$. Hence every vertex of $\mathfrak{X}$ is identified with an element of G and, under this identification, every polygon corresponds to an orbit under right translations by one of the factors $\mathbb{Z}_k$. For $x \in \mathfrak{X}$ and $n \leq |x|$, we write $x^{(n)}$ for the word of length n consisting of the first n blocks of x and simply write x' for $x^{(|x|-1)}$.

Let $\mathfrak{S}_n$ be the set of words of length n in $\mathfrak{X}$. We write Ω for the Poisson boundary of $\mathfrak{X}$. For $\omega \in \Omega$ and $n \in \mathbb{Z}_{\geq 0}$, we denote by ω_n the word of length n consisting the first n blocks. Let $E(x)$ denote the subset of Ω of words that begin with the reduced word $x \in \mathfrak{X}$. We write $\mathcal{M}$ and $\mathcal{M}_n$ for the σ-algebra generated by $\{E(x) : x \in \mathfrak{X}\}$ and σ-subalgebra generated by $\{E(x) : |x| \leqq n\}$ respectively. Then $\mathcal{M}$ makes

Ω into a compact topological space and there exists a natural G-quasi-invariant probability measure ν on $(\Omega, \mathcal{M})$. We write $\mathcal{F}(\Omega)$ for the space comprised of the $\mathcal{M}_n$-measurable functions on Ω. We denote by $\mathcal{F}(\Omega)_c$ the linear span of the characteristic functions of $E(x)$ for $x \in \mathfrak{X}$. The dual space $\mathcal{F}'(\Omega)$ is identified with the space of the martingales on Ω with respect to $\{\mathcal{M}_n\}$.

We write $s_0 = \bigl(\frac{1}{2} - \log_q(k-1)\bigr)i + \frac{1}{2}\tau$ and set

$$\Upsilon = \bigl\{s \in \mathbb{C} : s = \tfrac{1}{2}i + h\tau,\ s = s_0 + h\tau\ (h \in \mathbb{Z})\bigr\}.$$

We define the subsets b_x, c_x, d_x of $\mathfrak{X}$ by the following: for $x \in \mathfrak{X} \setminus \{o\}$

$$\begin{aligned} b_x &= \{y \in \mathfrak{X} : d(y,x) = 1, |y| = |x|\},\\ c_x &= \{y \in \mathfrak{X} : d(y,x) = 2, |y| = |x|\},\\ d_x &= \{y \in \mathfrak{X} : d(y,x) \geq 3, |y| = |x|\},\end{aligned}$$

and $b_o = c_o = d_o = \varnothing$. The subsets $B(x)$ and $C(x)$ of Ω are defined by

$$B(x) = \bigcup_{y \in b_x} E(y), \qquad C(x) = \bigcup_{y \in c_x} E(y).$$

For a function η on Ω and $n \in \mathbb{Z}_{\geq 0}$, we define the averages $E_n\eta$ and $B_n\eta$ as follows:

$$E_n\eta(\omega) = \frac{1}{\nu(E(\omega_n))}\int_{E(\omega_n)} \eta(\omega')\,d\nu(\omega'), \quad B_n\eta(\omega) = \frac{1}{\nu(B(\omega_n))}\int_{B(\omega_n)} \eta(\omega')\,d\nu(\omega').$$

Then, as shown in [Mantero and Zappa 1983, p. 375], the set $\{E_n\eta\}$ is a martingale associated to $\eta \in L^1(\Omega)$ and the n-th martingale difference of η is given by $D_n\eta = E_n\eta - E_{n-1}\eta$. Here we set $E_{-1} = 0$. For $x \in \mathfrak{X}$ and $\omega \in \Omega$, the Poisson kernel $p(x,\omega)$ is defined to be the Radon–Nikodym derivative $d\nu(x^{-1}\omega)/d\nu(\omega)$ and is computed as

$$p(x,\omega) = q^{\zeta(x,\omega)},$$

where $\zeta(x,\omega) = \lim_{m\to\infty}(m - d(x,\omega_m))$ is the Busemann function. As shown in [Iozzi and Picardello 1983b, Proposition 2], for $x \in \mathfrak{S}_n$, we have

$$p(x,\omega) = q^n\chi_{E(x)}(\omega) + \sum_{j=1}^{n} q^{2j-n-1}\chi_{B(x^{(j)})}(\omega) + \sum_{j=1}^{n} q^{2j-n-2}\chi_{C(x^{(j)})}(\omega). \tag{2-1}$$

For $\eta \in L^1(\Omega)$ and $s \in \mathbb{T}$, we define the Poisson transform $P^s\eta$ by

$$P^s\eta(x) = \int_\Omega p(x,\omega)^{1/2+is}\eta(\omega)\,d\nu(\omega). \tag{2-2}$$

By duality, the Poisson transform is naturally extended to $\mathcal{F}'(\Omega)$ and is denoted by the same symbol P^s.

Following [Mantero and Zappa 1983], we define the operators ε and Δ on $\mathfrak{X}$, which are essentially the analogue of E_n and D_n. We set for $n \in \mathbb{Z}_{\geq 0}$

$$S(n,x) = \begin{cases} \{x\}, & |x| \leq n, \\ \{y \in \mathfrak{X} : |y| = |x|,\, y^{(n)} = x^{(n)}\}, & |x| > n. \end{cases}$$

For a function ϕ on $\mathfrak{X}$ and $n \in \mathbb{Z}_{\geq 0}$, we define its average $\varepsilon_n\phi$ by

$$\varepsilon_n\phi(x) = \frac{1}{\operatorname{Card} S(n,x)} \sum_{y \in S(n,x)} \phi(y). \tag{2-3}$$

We also define $\Delta_n\phi$ by

$$\Delta_n\phi(x) = \varepsilon_n\phi(x) - \varepsilon_{n-1}\phi(x).$$

Here we set $\varepsilon_{-1}\phi = 0$. We write μ_1 for the probability measure equidistributed on words of length 1. We also use the notation κ_1 to denote the following:

$$(\phi * \kappa_1)(x) = \frac{1}{k-2} \sum_{y \in b_x} \phi(y).$$

We write $C_c(\mathfrak{X})$ for the set of all compactly supported functions on $\mathfrak{X}$. For $N \in \mathbb{Z}_{\geq 0}$, we denote by $C_N(\mathfrak{X})$ the subset of $C_c(\mathfrak{X})$ consisting of all $f \in C_c(\mathfrak{X})$ such that $\operatorname{supp} f \subseteq \mathfrak{B}_N$. A function ϕ on $\mathfrak{X}$ is said to be radial if $\varepsilon_0\phi = \phi$ and cylindrical if $\varepsilon_N\phi(x) = \phi(x)$ for some $N \in \mathbb{Z}_{\geq 0}$. For any function space $E(\mathfrak{X})$, we denote by $E(\mathfrak{X})^{\#}$ and $E(\mathfrak{X})_c$ the subspaces of $E(\mathfrak{X})$ consisting of all radial functions and cylindrical functions, respectively. A function f on $\mathbb{T}$ is said to be Weyl-invariant if $f(s+\tau) = f(s)$ and $f(-s) = f(s)$.

Finally we pointed out that it is meaningful to study harmonic analysis for symmetric graphs using methods similar to that in symmetric spaces. For example, the explicit expressions of the intertwining operators obtained in Section 4 can be used to construct the composition series of the spherical representations and determine which parts of the subquotients are unitarizable. In Section 6, using this information, we can concretely characterize the image of the compactly supported functions under the Helgason–Fourier transform.

3. The Poisson transform on symmetric graphs

Iozzi and Picardello [1983b] studied the Poisson transform for symmetric graphs. They showed in their paper that the Poisson transform P^s is injective on $\mathcal{F}(\Omega)_c$ if and only if $s \notin \Upsilon$. In this section, by carrying out similar arguments to that in [Mantero and Zappa 1983], we show that P^s is also surjective on $\mathcal{F}'(\Omega)$ when $s \notin \Upsilon$.

As shown in [Iozzi and Picardello 1983b, Theorem 1], we have

$$(P^s\eta * \mu_1)(x) = \gamma(s) P^s\eta(x),$$

where

$$\gamma(s)=\frac{1}{r(k-1)}(q^{1/2+is}+q^{1/2-is}+k-2).$$

By using the equation

$$\chi_{E(x^{(j)})}(\omega)-\chi_{E(x^{(j+1)})}(\omega)=\chi_{C(x^{(j+1)})}(\omega)+\chi_{B(x^{(j+1)})}(\omega),$$

(2-1) is expressed as

$$p(x,\omega)=q^{-|x|}\chi_{E(x^{(0)})}(\omega)+(1-q^{-2})\sum_{j=1}^{|x|}q^{2j-|x|}\chi_{E(x^{(j)})}(\omega) + (1-q^{-1})\sum_{j=1}^{|x|}q^{2j-|x|-1}\chi_{B(x^{(j)})}(\omega). \tag{3-1}$$

Therefore for $\omega\in\Omega$, substituting (3-1) into (2-2), we have

$$P^s\eta(\omega_n)=\sum_{j=0}^{n}b_{j,n}(s)E_j\eta(\omega)+\frac{k-2}{q^{1/2+is}+1}\sum_{j=1}^{n}b_{j,n}(s)B_j\eta(\omega),$$

where $b_{0,n}(s)=q^{-n(1/2+is)}$ and

$$b_{j,n}(s)=\frac{q}{r(k-1)}(1-q^{-1-i2s})q^{-n(1/2+is)+i2js}.$$

By the definitions of B_n and E_n, it is easy to verify that

$$E_mB_n=\begin{cases}B_n, & m\ge n,\\ E_m, & m<n,\end{cases}\qquad B_mE_n=\begin{cases}E_n, & m>n,\\ B_m, & m\le n.\end{cases}$$

And hence we obtain that

$$B_mD_n\eta=\begin{cases}D_n\eta, & m>n,\\ B_n\eta-E_{n-1}\eta, & m=n,\\ 0, & m<n.\end{cases}$$

Hereafter we suppose that $D_M\eta=\eta$ for some $M\ge 0$. We first consider the case when $M>0$. Since $E_j\eta=0$ and $B_j\eta=0$ for $j<M$, we have

$$P^s\eta(\omega_n)=0\quad\text{for } n<M,$$

and

$$P^s\eta(\omega_{M+\ell})=\sum_{j=M}^{M+\ell}b_{j,M+\ell}(s)\eta(\omega)+\frac{k-2}{q^{1/2+is}+1}\sum_{j=M+1}^{M+\ell}b_{j,M+\ell}(s)\eta(\omega) + \frac{k-2}{q^{1/2+is}+1}b_{M,M+\ell}(s)B_M\eta(\omega). \tag{3-2}$$

We define the function $Q_M^0(s)$ on $\mathbb{T}$ by $Q_0^0(s)=1$ and

$$Q_M^0(s)=\frac{\sqrt{q}}{r(k-1)}q^{-M/2}q^{i(M-1)s}(q^{1/2+is}-q^{-1/2-is})$$

for $M>0$. By using this, (3-2) can be written

$$\begin{aligned}P^s\eta(\omega_{M+\ell})&\\ &=q^{-\ell(1/2+is)}Q_M^0(s)\Bigg\{\sum_{j=M}^{M+\ell}q^{i2s(j-M)}\eta(\omega)+\frac{k-2}{q^{1/2+is}+1}\sum_{j=M+1}^{M+\ell}q^{i2s(j-M)}\eta(\omega)\\ &\qquad\qquad+\frac{k-2}{q^{1/2+is}+1}B_M\eta(\omega)\Bigg\}.\end{aligned}$$

Here we set $Q_0(s)=R_0(s)=1$ and

$$\begin{aligned}Q_M(s)&=\frac{\sqrt{q}}{r(k-1)}q^{-M/2}q^{i(M-1)s}(q^{1/2+is}-(k-1)q^{-1/2-is}+k-2),\\ R_M(s)&=\frac{\sqrt{q}}{r(k-1)}q^{-M/2}q^{i(M-1)s}(1-q^{-1/2-is})\end{aligned}$$

for $M>0$. Then a direct computation yields that

$$\begin{aligned}P^s\eta(\omega_{M+\ell})&=q^{-\ell/2}\psi(\ell+1,s)Q_M(s)\eta(\omega)\\ &\quad+(k-2)R_M(s)q^{-\ell(1/2+is)}(B_M\eta(\omega)-\eta(\omega)),\end{aligned}\tag{3-3}$$

where

$$\psi(n,s)=\frac{\sin(ns\log q)}{\sin(s\log q)}.$$

As pointed out in [Cowling and Setti 1999, p. 242], $D_M\eta=\eta$ if and only if η is constant on $E(x)$ for every $x\in\mathfrak{S}_M$ and the average of η with respect to $E(y)$ for $|y|<M$ is equal to 0. Therefore we can regard η as a function on $\mathfrak{X}$ by setting $\eta(x)=E_{|x|}\eta(\omega)$ for $\omega\in E(x)$. Under this identification, we have that $\eta(x)=0$ when $|x|<M$ and $\eta(x)=\eta(x^{(M)})$ when $|x|\geq M$. Moreover, for $x\in\mathfrak{X}$ such that $|x|>M$ and $y\in b_x$, it holds that $\eta(x)=\eta(y)$ because $|x'|=|y'|\geq M$. We also remark that $B_M\eta(\omega)$ corresponds to $\eta*\kappa_1(x^{(M)})$. For these reasons, (3-3) can be rewritten as follows:

$$\begin{aligned}P^s\eta(x)&=q^{-(|x|-M)/2}\psi(|x|-M+1,s)Q_M(s)\eta(x^{(M)})\\ &\quad+(k-2)R_M(s)q^{-(|x|-M)(1/2+is)}(\eta*\kappa_1(x^{(M)})-\eta(x^{(M)})).\end{aligned}\tag{3-4}$$

In the case $M=0$, η is a constant function on Ω and so $P^s\eta$ is expressed in terms of the spherical function ϕ_s given in [Iozzi and Picardello 1983b, Theorem 2] as:

$$P^s\eta(x)=\phi_s(x)\eta(o).$$

We summarize these results in the following proposition.

Proposition 3.1. *Let $\eta \in L^1(\Omega)$ be such that $D_M\eta = \eta$. Then the Poisson transform $P^s\eta$ has the following forms*:

(1) *If $M > 0$,*

$$P^s\eta(x) = 0, \qquad |x| < M,$$

$$P^s\eta(x) = q^{-(|x|-M)/2}\psi(|x|-M+1, s)Q_M(s)\eta(x^{(M)}) + (k-2)R_M(s)q^{-(|x|-M)(1/2+is)}(\eta * \kappa_1(x^{(M)}) - \eta(x^{(M)})), \quad |x| \geq M.$$

(2) *If $M = 0$,*

$$P^s\eta(x) = \phi_s(x)\eta(o).$$

Let $\mathscr{L}_{\gamma(s)}(\mathfrak{X})$ denote the space consisting of the functions ϕ on $\mathfrak{X}$ satisfying the condition $\phi * \mu_1 = \gamma(s)\phi$. For $M \in \mathbb{Z}_{\geq 0}$, we denote by $\mathscr{L}^M_{\gamma(s)}(\mathfrak{X})$ the subspace of $\mathscr{L}_{\gamma(s)}(\mathfrak{X})$ consisting of the functions ϕ which satisfy the following conditions:

(1) $\Delta_M\phi = \phi$,

(2) for $x \in \mathfrak{X}$ such that $|x| > M$ and $y \in b_x$, $\phi(x) = \phi(y)$.

Then $\mathscr{L}^0_{\gamma(s)}(\mathfrak{X})$ is just the space of radial harmonic functions on $\mathfrak{X}$. We prove the following lemma, which is an analogue of [Mantero and Zappa 1983, Lemma 3.2].

Lemma 3.2. *Let $\phi \in \mathscr{L}^M_{\gamma(s)}(\mathfrak{X})$ and $\omega \in \Omega$. Then we have the following*:

(1) *If $M > 0$,*

$$\begin{aligned} &\phi(\omega_n) = 0 \qquad (n < M), \\ &\phi(\omega_{M+\ell}) = q^{-\ell/2}\psi(\ell+1, s)\phi(\omega_M) \\ &\qquad\qquad + (k-2)q^{-(\ell+1)/2}\psi(\ell, s) \times (\phi(\omega_M) - \phi * \kappa_1(\omega_M)). \end{aligned} \tag{3-5}$$

(2) *If $M = 0$,*

$$\phi(\omega_\ell) = \phi_s(\omega_\ell)\phi(o).$$

Proof. Since for $\omega \in \Omega$ and $n \geq M$,

$$\phi*\mu_1(\omega_n) = \frac{1}{r(k-1)}\Bigl(\phi(\omega_{n-1}) + (k-2)\phi*\kappa_1(\omega_n) + \phi(\omega_{n+1}) + \sum_{y \in b_{\omega_{n+1}}}\phi(y) + \sum_{y \in c_{\omega_{n+1}}}\phi(y)\Bigr),$$

and

$$\Delta_{n+1}\phi(\omega_{n+1}) = \phi(\omega_{n+1}) - \frac{1}{q}\Bigl(\phi(\omega_{n+1}) + \sum_{y \in b_{\omega_{n+1}}}\phi(y) + \sum_{y \in c_{\omega_{n+1}}}\phi(y)\Bigr) = 0,$$

we have

$$\phi * \mu_1(\omega_n) = \frac{1}{r(k-1)}\{\phi(\omega_{n-1}) + (k-2)\phi * \kappa_1(\omega_n) + q\phi(\omega_{n+1})\}.$$

Hence the condition $\phi * \mu_1 = \gamma(s)\phi$ implies that

$$(q\beta(s)+k-2)\phi(\omega_n) = \phi(\omega_{n-1}) + (k-2)\phi * \kappa_1(\omega_n) + q\phi(\omega_{n+1}),$$

where $\beta(s) = q^{-1/2+is} + q^{-1/2-is}$. When $n > M$, we have that $\phi * \kappa_1(\omega_n) = \phi(\omega_n)$ and therefore we have the recursion formulae:

$$\phi(\omega_n) = 0, \qquad n < M, \tag{3-6}$$

$$\phi(\omega_{M+1}) = \beta(s)\phi(\omega_M) + (k-2)q^{-1}(\phi(\omega_M) - \phi * \kappa_1(\omega_M)), \tag{3-7}$$

$$\phi(\omega_{M+\ell}) = \beta(s)\phi(\omega_{M+\ell-1}) - q^{-1}\phi(\omega_{M+\ell-2}), \qquad \ell \geq 2. \tag{3-8}$$

In the case $s \notin (\tau/2)\mathbb{Z}$, the difference equation (3-8) has the fundamental solutions $q^{-1/2+is}$ and $q^{-1/2-is}$. So using the initial condition (3-7) to determine the coefficients of the fundamental solutions, we obtain the following expression:

$$\phi(\omega_{M+\ell}) = C_1 q^{\ell(-1/2+is)} + C_2 q^{\ell(-1/2-is)}, \tag{3-9}$$

where

$$C_1 = \frac{q^{is}\phi(\omega_M) + (k-2)q^{-1/2}\{\phi(\omega_M) - \phi * \kappa_1(\omega_M)\}}{q^{is} - q^{-is}},$$

$$C_2 = \frac{-q^{-is}\phi(\omega_M) - (k-2)q^{-1/2}\{\phi(\omega_M) - \phi * \kappa_1(\omega_M)\}}{q^{is} - q^{-is}}.$$

Similarly, when $s = \frac{1}{2}m\tau$ $(m \in \mathbb{Z})$, we also have

$$\phi(\omega_{M+\ell}) = (C'_1 + C'_2\ell)(-1)^{m\ell}q^{-\ell/2}, \tag{3-10}$$

where

$$C'_1 = \phi(\omega_M), \quad C'_2 = \phi(\omega_M) + (k-2)(-1)^m q^{-1/2}(\phi(\omega_M) - \phi * \kappa_1(\omega_M)).$$

Obviously both the expressions (3-9) and (3-10) agree with the equation (3-5). The case $M = 0$ is analogous. This concludes the proof. □

Let $x \in \mathfrak{S}_M$ and $s \notin \Upsilon$. Then we have from (3-4) that

$$P^s\eta(x) = Q_M(s)\eta(x) + (k-2)R_M(s)(\eta * \kappa_1(x) - \eta(x)). \tag{3-11}$$

We put $\phi(x) = P^s\eta(x)$ and write down η in terms of ϕ and $\phi * \kappa_1$. Since

$$\phi * \kappa_1 * \kappa_1(x) = \frac{1}{k-2}\phi(x) + \frac{k-3}{k-2}\phi * \kappa_1(x), \tag{3-12}$$

we have from (3-11) that

$$\phi * \kappa_1(x) = Q_M(s)(\eta * \kappa_1)(x) + R_M(s)(\eta(x) - \eta * \kappa_1(x)). \tag{3-13}$$

Thus, solving the simultaneous equations (3-11) and (3-13), we get the expressions

$$\eta(x)=\frac{(q^{1/2+is}+k-2)\phi(x)-(k-2)\phi*\kappa_1(x)}{q^{1/2+is}Q_M(s)},$$
$$\eta*\kappa_1(x)=\frac{(q^{1/2+is}+1)\phi*\kappa_1(x)-\phi(x)}{q^{1/2+is}Q_M(s)}.$$

The above expressions suggest the following proposition.

Proposition 3.3 (cf., [Mantero and Zappa 1983, Proposition 3.4]). *Let $s\notin\Upsilon$. For $\phi\in\mathscr{L}^M_{\gamma(s)}(\mathfrak{X})$, there exists a function η on Ω such that $D_M\eta=\eta$ and $P^s\eta=\phi$.*

Proof. Suppose that $M>0$. Indeed, define $\eta(\omega)$ by

$$\eta(\omega)=\frac{(q^{1/2+is}+k-2)\phi(\omega_M)-(k-2)\phi*\kappa_1(\omega_M)}{q^{1/2+is}Q_M(s)}.$$

Then $\phi(\omega_M)=P^s\eta(\omega_M)$ is trivial. Applying Lemma 3.2 to our case together with (3-11) and (3-13), we see that

$$\begin{aligned}\phi(\omega_{M+\ell})&=q^{-\ell/2}\psi(\ell+1,s)\phi(\omega_M)+(k-2)q^{-(\ell+1)/2}\psi(\ell,s)\\&\qquad\qquad\times(\phi(\omega_M)-\phi*\kappa_1(\omega_M))\\&=q^{-\ell/2}\psi(\ell+1,s)\{Q_M(s)\eta(\omega_M)+(k-2)R_M(s)(\eta*\kappa_1(\omega_M)-\eta(\omega_M))\}\\&\qquad+(k-2)q^{-(\ell+1)/2}\psi(\ell,s)q^{1/2+is}R_M(s)(\eta(\omega_M)-\eta*\kappa_1(\omega_M))\\&=q^{-\ell/2}\psi(\ell+1,s)Q_M(s)\eta(\omega_M)+(k-2)q^{-\ell/2}R_M(s)\\&\qquad\times\{\psi(\ell+1,s)-q^{is}\psi(\ell,s)\}(\eta*\kappa_1(\omega_M)-\eta(\omega_M))\\&=P^s\eta(\omega_{M+\ell}).\end{aligned}$$

The case $M=0$ is analogous. This concludes the proof. □

The following proposition is proved in the same way as in [Mantero and Zappa 1983, Corollary 3.5] and hence we omit its proof.

Proposition 3.4. *Suppose that $s\notin\Upsilon$. Then the Poisson transform P^s is a bijective operator from $\mathcal{F}'(\Omega)$ onto $\mathscr{L}_{\gamma(s)}(\mathfrak{X})$.*

4. The construction of the intertwining operator

Mantero and Zappa [1983] defined the intertwining operator between the spherical representations for free groups and gave an explicit expression of the intertwining operator. In this section, we extend their results to the case of symmetric graphs.

Let $s\in\mathbb{C}$ and define the action π_s of G on $L^2(\Omega)$ by

$$(\pi_s(g)\eta)(\omega)=p(g\cdot o,\omega)^{1/2+is}\eta(g^{-1}\omega).$$

The representation $(\pi_s, L^2(\Omega))$ is called the spherical representation. We denote by λ the left regular representation of G on $\mathscr{L}_{\gamma(s)}(\mathfrak{X})$. Then as indicated in [Iozzi and Picardello 1983b, p. 372], the Poisson transform P^{-s} intertwines π_s and λ. Therefore, in the case $\pm s \notin \Upsilon$, we see from Proposition 3.4 that the operator on $\mathcal{F}'(\Omega)$ defined by $I_s = (P^s)^{-1}P^{-s}$ is bijective and satisfies the following relation:

$$I_s \pi_s(g) = \pi_{-s}(g) I_s. \tag{4-1}$$

Let $\eta \in L^1(\Omega)$ be such that $D_M\eta = \eta$ for some $M > 0$. Under this assumption, $I_s(B_M\eta) = B_M(I_s\eta)$ because

$$P^s(B_M I_s\eta)(x) = P^s I_s\eta * \kappa_1(x) = P^{-s}\eta * \kappa_1(x) = P^{-s}(B_M\eta)(x).$$

Since

$$P^s I_s\eta(\omega_{M+\ell}) = P^{-s}\eta(\omega_{M+\ell}),$$

we have from (3-3) that

$$\begin{aligned} &q^{-\ell/2}\psi(\ell+1,s)Q_M(s)I_s\eta(\omega) + (k-2)R_M(s)q^{-\ell(1/2+is)}(I_sB_M\eta(\omega) - I_s\eta(\omega)) \\ &= q^{-\ell/2}\psi(\ell+1,s)Q_M(-s)\eta(\omega) + (k-2)R_M(-s)q^{-\ell(1/2-is)}(B_M\eta(\omega) - \eta(\omega)). \end{aligned}$$

Taking $\ell = 0$ and $\ell = 1$ respectively, we obtain from the above equation that

$$\begin{aligned} Q_M^0(s)I_s\eta(\omega) + (k-2)R_M(s)I_sB_M\eta(\omega) \\ = Q_M^0(-s)\eta(\omega) + (k-2)R_M(-s)B_M\eta(\omega) \end{aligned} \tag{4-2}$$

and

$$\begin{aligned} \beta(s)Q_M(s)I_s\eta(\omega) + (k-2)q^{-(1/2+is)}R_M(s)(I_sB_M\eta(\omega) - I_s\eta(\omega)) \\ = \beta(s)Q_M(-s)\eta(\omega) + (k-2)q^{-(1/2-is)}R_M(-s)(B_M\eta(\omega) - \eta(\omega)). \end{aligned} \tag{4-3}$$

Solving the simultaneous equations (4-2) and (4-3), we have

$$\begin{aligned} I_s\eta(\omega) &= \frac{q^{-is}Q_M(-s)\eta(\omega) + (q^{is} - q^{-is})Q_M^0(-s)\eta(\omega)}{q^{is}Q_M(s)} \\ &\quad + \frac{(k-2)(q^{is} - q^{-is})R_M(-s)B_M\eta(\omega)}{q^{is}Q_M(s)}, \end{aligned} \tag{4-4}$$

$$\begin{aligned} I_sB_M\eta(\omega) &= \frac{(q^{is} - q^{-is})R_M(-s)\{\eta(\omega) - B_M\eta(\omega)\}}{q^{is}Q_M(s)} \\ &\quad + \frac{q^{is}Q_M(-s)B_M\eta(\omega)}{q^{is}Q_M(s)}. \end{aligned} \tag{4-5}$$

For $\eta \in L^1(\Omega)$ satisfying $D_M\eta = \eta$, we define η^+ and η^- as

$$\eta^+(\omega) = \eta(\omega) + (k-2)B_M\eta(\omega), \tag{4-6}$$

$$\eta^-(\omega) = \eta(\omega) - B_M\eta(\omega). \tag{4-7}$$

Then we have from (4-4) and (4-5) that

$$I_s\eta^+(\omega) = \frac{Q_M(-s)}{Q_M(s)}\eta^+(\omega), \tag{4-8}$$

$$I_s\eta^-(\omega) = \frac{q^{-is}R_M(-s)}{q^{is}R_M(s)}\eta^-(\omega). \tag{4-9}$$

For $n \in \mathbb{Z}_{\geq 0}$, we denote by $\mathcal{H}_n$ the subspace of $\mathcal{F}(\Omega)_c$ consisting of η such that $D_n\eta = \eta$. We write $\mathcal{H}_n^+$ and $\mathcal{H}_n^-$ for the subspaces of $\mathcal{H}_n$ generated by $\{\eta^+ : \eta \in \mathcal{H}_n\}$ and $\{\eta^- : \eta \in \mathcal{H}_n\}$, respectively. Then it holds that $\mathcal{H}_0 = \mathcal{H}_0^+$ and $\mathcal{H}_n = \mathcal{H}_n^+ \oplus \mathcal{H}_n^-$ for $n > 0$. The expressions (4-8) and (4-9) give the explicit forms of the intertwining operator I_s when restricted to $\mathcal{H}_n^+$ and $\mathcal{H}_n^-$, respectively.

Finally in this section, we list some properties of the Poisson transform. Analogously to (4-6) and (4-7), for a function ϕ on $\mathfrak{X}$, we define ϕ^+ and ϕ^- by

$$\phi^+(x) = \phi(x) + (k-2)\phi * \kappa_1(x), \quad \phi^-(x) = \phi(x) - \phi * \kappa_1(x).$$

Let $\eta \in L^1(\Omega)$ be such that $D_M\eta = \eta$. Then taking into account (3-12), we see from Proposition 3.1 that

$$(P^s\eta)^+(x) = q^{-(|x|-M)/2}\psi(|x|-M+1, s)Q_M(s)\eta^+(\omega) = (P^s\eta^+)(x) \tag{4-10}$$

and

$$\begin{aligned}(P^s\eta)^-(x) &= q^{-(|x|-M)/2}\psi(|x|-M+1, s)Q_M(s)\eta^-(\omega)\\ &\qquad\qquad -(k-1)q^{-(|x|-M)(1/2+is)}R_M(s)\eta^-(\omega)\\ &= (P^s\eta^-)(x).\end{aligned} \tag{4-11}$$

We here set

$$\psi^-(\ell+1, s) = \sum_{j=0}^{\ell} q^{i(\ell-2j)s} + q^{-1/2}(k-1)\sum_{j=1}^{\ell} q^{i(\ell-2j+1)s}.$$

Then the expression (4-11) can be written

$$(P^s\eta)^-(x) = q^{-(|x|-M)/2}\psi^-(|x|-M+1, s)q^{1/2+is}R_M(s)\eta^-(\omega). \tag{4-12}$$

We summarize these in the following corollary.

Corollary 4.1. *Let $\eta \in L^1(\Omega)$ be such that $D_M\eta = \eta$. Then*

$$(P^s\eta)^+(x) = (P^s\eta^+)(x), \qquad (P^s\eta)^-(x) = (P^s\eta^-)(x).$$

In addition, for any $x \in \mathfrak{X}$ satisfying $|x| > M$, we have

(1) $Q_M(s)^{-1}(P^s\eta)^+(x)$ *is an even entire holomorphic function on $\mathbb{C}$ with respect to the variable s,*

(2) $(q^{1/2+is}R_M(s))^{-1}(P^s\eta)^-(x)$ *is an even entire holomorphic function on $\mathbb{C}$ with respect to the variable s.*

5. The spectral projection on symmetric graphs

We first review the Helgason–Fourier transform for symmetric graphs and its inversion formula, which were introduced by Eddine [2013; 2015].

Let $(\pi_s, L^2(\Omega))$ be a spherical representation and let I_s be the intertwining operator defined in the previous section. The Helgason–Fourier transform $\tilde{f}(s,\omega)$ of $f \in C_c(\mathfrak{X})$ is defined by

$$\tilde{f}(s,\omega) = (\pi_s(f)1)(\omega) = \sum_{x\in\mathfrak{X}} f(x)p(x,\omega)^{1/2+is}. \tag{5-1}$$

Here 1 denotes the function identically one on Ω. In [Jamal Eddine 2015, Lemma 3.10], Eddine proved the following inversion formula:

$$f(x) = \frac{(k-r)_+}{k}\int_\Omega \tilde{f}(s_0,\omega)p(x,\omega)^{1/2-is_0}\,d\nu(\omega) \\ + c_G\int_\Omega\int_{\mathbb{T}} \tilde{f}(s,\omega)p(x,\omega)^{1/2-is}|c(s)|^{-2}\,ds\,d\nu(\omega),$$

where $c_G = q/\{2\tau r(k-1)\}$ and

$$c(s) = \frac{\sqrt{q}}{q+1}\cdot\frac{q^{1/2+is}-(k-1)q^{-1/2-is}+k-2}{q^{is}-q^{-is}}$$

is a c-function. Here $(k-r)_+$ stands for $k-r$ when $k>r$ and to 0 when $k \leqq r$. As described in [Cowling and Setti 1999, p. 240], we see that $I_s\tilde{f}(s,\omega) = \tilde{f}(-s,\omega)$ for almost all $s \in \mathbb{T}$ and thus we obtain the following symmetry condition:

$$\int_\Omega \tilde{f}(s,\omega)p(x,\omega)^{1/2-is}\,d\nu(\omega) = \int_\Omega \tilde{f}(-s,\omega)p(x,\omega)^{1/2+is}\,d\nu(\omega). \tag{5-2}$$

Following Bray [1996], we define the spectral projection $P_s f$ of $f \in C_c(\mathfrak{X})$ by

$$P_s f(x) = (f * \phi_s)(x) = \int_G f(g_1)\phi_s(g_1^{-1}g)\,dg_1, \tag{5-3}$$

where $x = g \cdot o$. Applying the functional equation of the spherical function [Jamal Eddine 2015, Lemma 3.9] and using Fubini's theorem, we obtain

$$\begin{aligned} P_s f(x) &= \int_G f(g_1 \cdot o) \int_\Omega p(g_1 \cdot o, \omega)^{1/2+is} p(g \cdot o, \omega)^{1/2-is} \, d\nu(\omega) \, dg_1 \\ &= \int_\Omega \tilde{f}(s, \omega) p(x, \omega)^{1/2-is} \, d\nu(\omega). \end{aligned}$$

Thus the spectral projection $P_s f(x)$ is Weyl-invariant with respect to the variable s and has the following inversion formula:

$$f(x) = \frac{(k-r)_+}{k} P_{s_0} f(x) + c_G \int_{\mathbb{T}} P_s f(x) |c(s)|^{-2} \, ds. \tag{5-4}$$

Let $a \in \mathfrak{X}$ and define the function ξ_a on Ω by $\xi_o(\omega) = 1$ and for $a \neq o$

$$\xi_a(\omega) = \nu(E(a))^{-1} \chi_{E(a)}(\omega) - \nu(E(a'))^{-1} \chi_{E(a')}(\omega).$$

Then it is easy to see that $D_{|a|} \xi_a = \xi_a$ and $B_{|a|} \xi_a = (\xi_* * \kappa_1)(a)$. For $a \in \mathfrak{X}$ and $s \in \mathbb{C}$, we define the generalized spherical function $\Phi_{a,s}$ on $\mathfrak{X}$ by

$$\Phi_{a,s}(x) = P^s \xi_a(x) = \int_\Omega p(x, \omega)^{1/2+is} \xi_a(\omega) \, d\nu(\omega).$$

By Proposition 3 in [Koizumi 2013] combined with Corollary 4.1, we have that

$$(\Delta_n P_s f)^{\pm}(x) = \int_\Omega \Phi^{\pm}_{\omega_n, -s}(x) \tilde{f}(s, \omega) \, d\nu(\omega).$$

The explicit expressions of $\Phi^{\pm}_{\omega_n,-s}$ are given by (4-10) and (4-12). We see from these that

$$\begin{aligned} (\Delta_n P_s f)^+(x) &= 0 \quad \text{when} \quad \pm s \in \Upsilon, \\ (\Delta_n P_s f)^-(x) &= 0 \quad \text{when} \quad \pm s \in \tfrac{1}{2} i + \tau \mathbb{Z}. \end{aligned}$$

Furthermore, we have the following proposition.

Proposition 5.1. *Let $f \in C_N(\mathfrak{X})$. Then $(\Delta_n P_{s_0} f)^-(x) = 0$ when $n > N$.*

Proof. By the definition of the Poisson transform, we have

$$\begin{aligned} (\Delta_n P_{s_0} f)^-(x) &= \sum_{y \in \mathfrak{X}} f(y) \biggl\{ \int_\Omega p(x, \omega)^{1/2 - is_0} \Phi^-_{\omega_n, s_0}(y) \, d\nu(\omega) \biggr\} \\ &= \sum_{y \in \mathfrak{X}} f(y) \biggl\{ \int_\Omega (1-r)^{\zeta(x,\omega)} \Bigl(\frac{1}{1-k} \Bigr)^{|y|} \frac{k(k-1)^{n-1}}{r(r-1)^{n-1}} \xi^-_{\omega_n}(\omega') \, d\nu(\omega) \biggr\}, \end{aligned}$$

where we choose $\omega' \in E(y)$. Taking into account

$$\xi^-_{\omega_n}(\omega') = \begin{cases} r(k-1)q^{n-1}, & \omega \in E(y^{(n)}), \\ -r(k-1)q^{n-1}/(k-2), & \omega \in B(y^{(n)}), \\ 0, & \text{otherwise}, \end{cases}$$

we have

$$(\Delta_n P_{s_0} f)^-(x) = k(k-1)^{2n-1} \sum_{\substack{y \in \mathfrak{X} \\ |y| \geq n}} f(y) \Big(\frac{1}{1-k}\Big)^{|y|} \int_{E(y^{(n)})} (1-r)^{\zeta(x,\omega)}\, d\nu(\omega)$$
$$- \frac{k(k-1)^{2n-1}}{k-2} \sum_{\substack{y \in \mathfrak{X} \\ |y| \geq n}} f(y) \Big(\frac{1}{1-k}\Big)^{|y|} \int_{B(y^{(n)})} (1-r)^{\zeta(x,\omega)}\, d\nu(\omega).$$

Since $f(y) = 0$ for $|y| \geq n > N$, we have $(\Delta_n P_{s_0} f)^-(x) = 0$. □

6. Paley–Wiener theorem for spectral projection

In this section, we shall characterize the image of $C_c(\mathfrak{X})$ under the spectral projection on $\mathfrak{X}$. As an application, we shall prove the Paley–Wiener theorem of the Helgason–Fourier transform for symmetric graphs.

Throughout this section, for a function ϕ on $\mathfrak{X}$, we denote $\Delta_n \phi(x)$ by $\phi_n(x)$. Let $N \in \mathbb{Z}_{\geq 0}$. We denote by $\mathscr{T}_N(\mathbb{T} \times \mathfrak{X})$ the set comprised of all functions F on $\mathbb{T} \times \mathfrak{X}$ satisfying the following conditions:

(N1) $F(s, x)$ is a Weyl-invariant smooth function on $\mathbb{R}$ with respect to the variable s.

(N2) For each $n \in \mathbb{Z}_{\geq 0}$ and $s \in \mathbb{R}$, $F_n(s, x) \in \mathscr{L}^n_{\gamma(s)}(\mathfrak{X})$,

(N3) For each $x \in \mathfrak{X}$, $F(s, x)$ extends to a Weyl-invariant holomorphic function on $\mathbb{C}$.

(N4) For each $n \in \mathbb{Z}_{\geq 0}$, $Q_n(-s)^{-1} F_n^+(s, x)$ is holomorphic on $\mathbb{C}$ and there exists a constant $C_N > 0$ which does not depend on the choice of n such that

$$|Q_n(-s)^{-1} F_n^+(s, x)| \leq C_N q^{(|x|-n+N)|\Im s|}.$$

(N5) For each $n \in \mathbb{Z}_{>0}$, $(q^{1/2-is} R_n(-s))^{-1} F_n^-(s, x)$ is holomorphic on $\mathbb{C}$ and there exists a constant $C_N > 0$ which does not depend on the choice of n such that

$$|(q^{1/2-is} R_n(-s))^{-1} F_n^-(s, x)| \leq C_N q^{(|x|-n+N)|\Im s|}.$$

(N6) $F_n^-(s_0, x) = 0$ when $n > N$.

We set

$$\mathscr{T}(\mathbb{T} \times \mathfrak{X}) = \bigcup_{N=0}^{\infty} \mathscr{T}_N(\mathbb{T} \times \mathfrak{X}).$$

The following proposition is obtained by the same arguments as in [Koizumi 2013, Proposition 4].

Proposition 6.1. *Let $f \in C_N(\mathfrak{X})$. Then $F(s,x) = P_s f(x)$ belongs to $\mathscr{T}_N(\mathbb{T} \times \mathfrak{X})$.*

To prove the sufficient condition in the Paley–Wiener theorem, we need the following lemma.

Lemma 6.2 (cf., [Koizumi 2013, Lemma 1]). *Let $N \in \mathbb{Z}_{>0}$, $F \in \mathscr{T}_N(\mathbb{T} \times \mathfrak{X})$ and $a \in \mathfrak{S}_n$. If $n > N$ then $F_n^+(s,a) = 0$ and $F_n^-(s,a) = 0$ for all $s \in \mathbb{T}$.*

Proof. We shall first show that $F_n^-(s,a) = 0$. Let us set

$$\phi^-(s) = (q^{1/2-is} R_n(-s))^{-1} F_n^-(s,a).$$

Then we see

$$\phi^-(-s) = (q^{1/2+is} R_n(s))^{-1} F_n^-(s,a) = \frac{q^{1/2-is} R_n(-s)}{q^{1/2+is} R_n(s)} \phi^-(s). \tag{6-1}$$

We put

$$c^-(n,s) = \frac{q^{1/2-is} R_n(-s)}{q^{1/2+is} R_n(s)}.$$

Obviously we have

$$c^-(n,s) = -q^{-1/2-is(2n-1)} + (1-q^{-1}) \sum_{\ell=0}^{\infty} q^{-\ell/2-is(2n+\ell)}. \tag{6-2}$$

The condition (N5) yields that $\phi^-(s)$ is an entire function of exponential type N. We use the Paley–Wiener theorem on $\mathbb{Z}$ to write

$$\phi^-(s) = \sum_{m \in \mathbb{Z}} \phi^-(m) q^{ims},$$

where $\phi^-(m) = 0$ unless $-N \le m \le N$. Substituting (6-2) to (6-1), we have

$$\sum_{m \in \mathbb{Z}} \phi^-(m) q^{-ims} = \sum_{m \in \mathbb{Z}} \left[-q^{-\frac{1}{2}-is(2n-1)} + (1-q^{-1}) \sum_{\ell=0}^{\infty} q^{-\frac{\ell}{2}-is(2n+\ell)} \right] \times \phi^-(m) q^{ims},$$

and thus we have the following recursion formula:

$$\phi^-(m) = -q^{-1/2} \phi^-(2n-1-m) + (1-q^{-1}) \sum_{\ell=0}^{\infty} q^{-\ell/2} \phi^-(2n+\ell-m). \tag{6-3}$$

From (6-3), when $n > N+1$, it is easily verified that $\phi^-(m) = 0$ for all $m \in \mathbb{Z}$. When $n = N+1$, (6-3) implies

$$\phi^-(m) = -q^{-1/2} \phi^-(2N+1-m),$$

and so $\phi^-(N) = 0$. We consequently have that $\phi^-(m) = 0$ for all $m \in \mathbb{Z}$.

We shall next show that $F_n^+(s,a)=0$. We set $\phi^+(s)=Q_n(-s)^{-1}F_n^+(s,a)$. We use the Paley–Wiener theorem on $\mathbb{Z}$ to write

$$\phi^+(s)=\sum_{m\in\mathbb{Z}}\phi^+(m)q^{ims},$$

where $\phi^+(m)=0$ unless $-N\le m\le N$. Putting

$$c^+(n,s)=\frac{Q_n(-s)}{Q_n(s)},$$

we have

$$\phi^+(-s)=c^+(n,s)\phi^+(s).$$

Because

$$c^+(n,s)=\frac{1+(k-1)q^{-1/2+is}}{1+(k-1)q^{-1/2-is}}c^-(n,s),$$

we have

$$\begin{aligned}
\text{(6-4)}\quad \phi^+&(m)\\
=&-(k-1)q^{-1}\phi^+(2n-2-m)\\
&+(k-1)(1-q^{-1})\sum_{\ell=0}^{\infty}q^{-(\ell+1)/2}\phi^+(2n+\ell-m-1)\\
&-(1-(1-k)^2q^{-1})\sum_{\ell=0}^{\infty}(1-k)^{\ell}q^{-(\ell+1)/2}\phi^+(2n+\ell-1-m)\\
&+\frac{(1-q^{-1})(1-(1-k)^2q^{-1})}{k}\sum_{\ell=0}^{\infty}(1-(1-k)^{\ell+1})q^{-\ell/2}\phi^+(2n+\ell-m).
\end{aligned}$$

From (6-4), when $n>N+1$, it is easily verified that $\phi^+(m)=0$ for all $m\in\mathbb{Z}$. In the case $n=N+1$, (6-4) implies

$$\phi^+(m)=-(k-1)q^{-1}\phi^+(2N-m),$$

and so $\phi^+(N)=0$. Therefore, in this case, $\phi^+(m)=0$ for all $m\in\mathbb{Z}$. □

The proof of the sufficient condition in the Paley–Wiener theorem is like the proof for the case of semisimple Lie groups given by Campoli [1980] and Johnson [1979]. We remark that $\Im s_0\ge 0$ when $k\le r$ and $\Im s_0<0$ when $k>r$. We also remark that the residue of $1/c(s)$ at $s=s_0$ is equal to $r(k-r)/\{ik(r-1)\log q\}$ and $1/c(-s_0)=1$. We first show the following proposition.

Proposition 6.3. *Let* $N\in\mathbb{Z}_{>0},\ F\in\mathscr{T}_N(\mathbb{T}\times\mathfrak{X})$ *and set*

$$f_0(x)=c_G\int_{\mathbb{T}}F_0(s,x)|c(s)|^{-2}\,ds.$$

Then there exists a function $J_0 \in C_N(\mathfrak{X})^{\#}$ *such that*

$$f_0(x) - J_0(x) = \frac{(k-r)_+}{k} F_0(s_0, x).$$

Proof. We put $F(s) = F_0(s, o)$. Then it follows from Lemma 3.2 and the condition (N4) that $F(s)$ is an even entire function of exponential order N and

$$F_0(s, x) = \phi_s(x) F(s).$$

We thus have

$$\begin{aligned} f(x) &= c_G \int_{\mathbb{T}} \phi_s(x) F(s) |c(s)|^{-2}\, ds \\ &= c_G \int_{\mathbb{T}} F(s) \frac{1}{c(-s)} q^{(is-1/2)|x|}\, ds + c_G \int_{\mathbb{T}} F(s) \frac{1}{c(s)} q^{(-is-1/2)|x|}\, ds. \end{aligned} \tag{6-5}$$

We write $f_1(x)$ and $f_2(x)$ for the first term and the second term of the last expression of (6-5), respectively. For a sufficiently large $\eta > 0$, let $f_{1,\eta}$ denote the formula shifting the path of integral of f_1 from $\mathbb{T}$ to $\mathbb{T} + i\eta$ and let $f_{2,-\eta}$ denote the formula shifting the path of integral of f_2 from $\mathbb{T}$ to $\mathbb{T} - i\eta$.

Suppose that $k \le r$. Because f_1 is analytic inside the rectangle with corners $\pm\tau/2$ and $\pm\tau/2 + i\eta$, we have by Cauchy's theorem that $f_1 = f_{1,\eta}$. Similarly we can also obtain that $f_2 = f_{2,-\eta}$. In case $k > r$, we have

$$\begin{aligned} f_1(x) - f_{1,\eta}(x) &= 2\pi i c_G \operatorname{Res}_{s=-s_0} \Big\{ F(s) \frac{1}{c(-s)} q^{(is-1/2)|x|} \Big\} \\ &= \frac{k-r}{2k} F(-s_0) \Big(\frac{1}{1-k} \Big)^{|x|}, \\ f_2(x) - f_{2,-\eta}(x) &= 2\pi i c_G \operatorname{Res}_{s=s_0} \Big\{ F(s) \frac{1}{c(s)} q^{(-is-1/2)|x|} \Big\} \\ &= \frac{k-r}{2k} F(s_0) \Big(\frac{1}{1-k} \Big)^{|x|}. \end{aligned}$$

Then the assertion follows immediately. □

In the nonradial case, we need slightly complicated calculations.

Proposition 6.4. *Let* $N \in \mathbb{Z}_{>0}$, $F \in \mathscr{T}_N(\mathbb{T} \times \mathfrak{X})$ *and set*

$$f_n(x) = c_G \int_{\mathbb{T}} F_n(s, x) |c(s)|^{-2}\, ds$$

for $n > 0$. *Then there exists a function* $J_n \in C_N(\mathfrak{X})$ *such that*

$$f_n(x) - J_n(x) = \frac{(k-r)_+}{k} F_n(s_0, x).$$

Proof. We put $a = x^{(n)}$ and choose $\omega \in E(x)$. Because $F_n^{\pm}$ satisfies the condition (N2), we have from (4-10) and (4-12) that

$$F_n^{+}(s,x) = q^{-(|x|-|a|)/2}\psi(|x|-|a|+1,s)Q_n(s)F_n^{+}(s,a), \tag{6-6}$$

$$F_n^{-}(s,x) = q^{-(|x|-|a|)/2}\psi^{-}(|x|-|a|+1,s)q^{1/2+is}R_n(s)F_n^{-}(s,a). \tag{6-7}$$

In the case when $n > N$, Lemma 6.2 yields that $F_n^{+}(s,a) = 0$ and $F_n^{-}(s,a) = 0$. On the other hand, since $F_n^{-}(s_0, x) = 0$, it suffices to set $J_n(x) = 0$.

In the following, we suppose that $n \le N$. Substituting (6-6) and (6-7), we obtain

$$f_n^{\pm}(x) = \xi_a^{\pm}(\omega)^{-1} c_G \int_{\mathbb{T}} h_a^{\pm}(s)\Phi_{a,s}^{\pm}(x)|c(s)|^{-2}\, ds,$$

where $h_a^{+}(s) = Q_{|a|}(s)^{-1}F_n^{+}(s,a)$, $h_a^{-}(s) = (q^{1/2+is}R_{|a|}(s))^{-1}F_n^{-}(s,a)$ and $\omega \in E(x)$. We know that $\Phi_{a,s}^{+}(x)$ has the following expansion:

$$\Phi_{a,s}^{+}(x) = \{c(s)q^{(is-1/2)|x|} - q^{i2(|a|-1)s}c(s)q^{(-is-1/2)|x|}\}\xi_a^{+}(\omega). \tag{6-8}$$

Hence we can show $f_n^{+} \in C_N(\mathfrak{X})$ for all k, r by the same arguments as in the proof of Proposition 6.3.

Hereafter we shall compute $f_n^{-}(x)$. It follows from (6-8) that

$$\begin{aligned}\Phi_{a,s}^{-}(x) = \{c(s)q^{(is-1/2)|x|} - q^{i2(|a|-1)s}c(s)q^{(-is-1/2)|x|}\\ - (k-1)R_{|a|}(s)q^{-(|x|-|a|)(1/2+is)}\}\xi_a^{-}(\omega).\end{aligned}$$

Let us set

$$\begin{aligned}
f_{n,1}^{-}(x) &= c_G\int_{\mathbb{T}} h_a^{-}(s)\frac{1}{c(-s)}q^{(is-1/2)|x|}\, ds,\\
f_{n,2}^{-}(x) &= c_G\int_{\mathbb{T}} h_a^{-}(s)\frac{1}{c(-s)}q^{i2(|a|-1)s}q^{(-is-1/2)|x|}\, ds,\\
f_{n,3}^{-}(x) &= c_G\int_{\mathbb{T}} h_a^{-}(s)R_{|a|}(s)q^{-(|x|-|a|)(1/2+is)|x|}|c(s)|^{-2}\, ds.
\end{aligned}$$

Suppose first that $k \le r$. Then as shown in Proposition 6.3, we see that $f_{n,1}^{-} \in C_N(\mathfrak{X})$. Keeping the notation in the proof of Proposition 6.3, we have

$$\begin{aligned}
f_{n,2}^{-}(x) - f_{n,2,-\eta}^{-}(x) &= 2\pi i c_G \operatorname{Res}_{s=-s_0}\Big\{h_a^{-}(s)\frac{1}{c(-s)}q^{i2(|a|-1)s}q^{(-is-1/2)|x|}\Big\}\\
&= \frac{r(k-r)}{2k}h_a^{-}(-s_0)(1-k)^{-|a|+1}(1-r)^{-|x|+|a|-1},\\
f_{n,3}^{-}(x) - f_{n,3,-\eta}^{-}(x) &= 2\pi i c_G \operatorname{Res}_{s=-s_0}\{h_a^{-}(s)R_{|a|}(s)q^{-(is+1/2)(|x|-|a|)}|c(s)|^{-2}\}\\
&= \frac{r(k-r)}{2k}h_a^{-}(-s_0)(1-k)^{-|a|}(1-r)^{-|x|+|a|-1}.
\end{aligned}$$

Setting
$$f_{n,\eta}^-(x) = f_{n,1}^-(x) - f_{n,2,-\eta}^-(x) - (k-1)f_{n,3,-\eta}^-(x),$$
we obtain that
$$f_n^-(x) - f_{n,\eta}^-(x) = 0.$$
Thus we have by contour integration arguments that $f_n^- \in C_N(\mathfrak{X})$.

Suppose next that $k > r$. In this case, by the same discussion as in the proof of Proposition 6.3, we see that $f_{n,2}^- \in C_N(\mathfrak{X})$. Moreover we have

$$\begin{aligned} f_{n,1}^-(x) - f_{n,1,\eta}^-(x) &= 2\pi i c_G \operatorname{Res}_{s=-s_0}\Big\{h_a^-(s)\frac{1}{c(-s)}q^{(is-1/2)|x|}\Big\} \\ &= \frac{k-r}{2k}\Big(\frac{1}{1-k}\Big)^{|x|} h_a^-(-s_0), \\ f_{n,3}^-(x) - f_{n,3,-\eta}^-(x) &= 2\pi i c_G \operatorname{Res}_{s=s_0}\{h_a^-(s)R_{|a|}(s)q^{-(|x|-|a|)(1/2+is)}|c(s)|^{-2}\} \\ &= \frac{k-r}{2r}h_a^-(s_0)(1-k)^{-|x|+|a|-2}(1-r)^{-|a|+1}. \end{aligned}$$
Let us set
$$f_{n,\eta}^-(x) = f_{n,1,\eta}^-(x) - f_{n,2}^-(x) - (k-1)f_{n,3,-\eta}^-(x).$$
Then we have
$$f_n^-(x) - f_{n,\eta}^-(x) = \Big(\frac{1}{1-k}\Big)^{|x|}\Big\{\frac{k-r}{2k}h_a(-s_0) + \frac{k-r}{2r}h_a^-(s_0)(1-k)^{|a|-1}(1-r)^{-|a|+1}\Big\}.$$
We see from the definition of $h_a^-(s)$ that
$$h_a^-(-s_0) = (1-k)^{|a|}F_n^-(s_0,a), \quad h_a^-(s_0) = \frac{r(1-k)(1-r)^{|a|-1}}{k}F_n^-(s_0,a).$$
We therefore obtain
$$f_n^-(x) - f_{n,\eta}^-(x) = \Big(\frac{1}{1-k}\Big)^{|x|}\frac{k-r}{k}(1-k)^{|a|}F_n^-(s_0,a) = \frac{k-r}{k}F_n^-(s_0,x).$$
By contour integration arguments, we see that there exists $J_n^- \in C_N(\mathfrak{X})$ such that
$$f_n^-(x) - J_n^-(x) = \frac{(k-r)_+}{k}F_n^-(s_0,x),$$
concluding the proof. □

Summarizing the arguments in this section, we arrive at the following theorem.

Theorem 6.5. *Let $N \in \mathbb{Z}_{>0}$ and $F \in \mathscr{T}_N(\mathbb{T}\times\mathfrak{X})$. We set*
$$f(x) = c_G\int_{\mathbb{T}} F(s,x)|c(s)|^{-2}\,ds.$$
Then there exists a function $J \in C_N(\mathfrak{X})$ such that
$$f(x) - J(x) = \frac{(k-r)_+}{k}F(s_0,x).$$

Proof. Let $x \in \mathfrak{X}$ be such that $|x| > N$. We choose a positive integer M so that $|x| \leq M$. Then $f(x)$ can be written as the following finite sum:

$$f(x) = \varepsilon_M f(x) = f_0(x) + f_1(x) + \cdots + f_M(x).$$

Applying Propositions 6.3 and 6.4 to each F_n, we have

$$\sum_{n=0}^{M} f_n(x) - \sum_{n=0}^{\min(M,N)} J_n(x) = \frac{(k-r)_+}{k} \sum_{n=0}^{\min(M,N)} F_n(s_0, x).$$

We thus have the required result. □

In the remainder of this section, as a corollary of Theorem 6.5, we shall prove the Paley–Wiener theorem of the Helgason–Fourier transform.

Let $N \in \mathbb{Z}_{\geq 0}$. Denote by $\mathscr{Z}_N(\mathbb{T} \times \Omega)$ the set of all functions F on $\mathbb{T} \times \Omega$ satisfying the following conditions:

(H1) $F(s, \omega)$ is a smooth function on $\mathbb{T}$ with respect to the variable s,

(H2) $F(s + \tau, \omega) = F(s, \omega)$,

(H3) $F(s, \omega)$ extends to a τ-periodic entire function of exponential type N,

(H4) F satisfies the symmetry condition (5-2),

(H5) $(D_n F)^-(s_0, \omega) = 0$ when $n > N$.

With the notation above, we show the following theorem.

Theorem 6.6. *Let $N \in \mathbb{Z}_{\geq 0}$, $F \in \mathscr{Z}_N(\mathbb{T} \times \Omega)$ and set*

$$f(x) = c_G \int_{\mathbb{T}} \int_{\Omega} F(s, \omega) p(x, \omega)^{1/2 - is} \, ds \, d\nu(\omega).$$

Then there exists a function $J \in C_N(\mathfrak{X})$ such that

$$f(x) - J(x) = \frac{(k-r)_+}{k} \int_{\Omega} F(s_0, \omega) p(x, \omega)^{1/2 - is_0} \, d\nu(\omega).$$

Proof. Let $F \in \mathscr{Z}_N(\mathbb{T} \times \Omega)$. It suffices to show $P^{-s} F \in \mathscr{T}_N(\mathbb{T} \times \mathfrak{X})$. The conditions (H1), (H2) and (H6) are immediate from the conditions (H1), (H2) and (H5). Noting

$$|q^{-n/2} \psi(n+1, s)| \leq \frac{q+1}{q-1} q^{n|\Im s|}, \quad |q^{-n/2} \psi^-(n+1, s)| \leq \frac{q+1}{q-1} q^{n|\Im s|},$$

we have from (4-10) that

$$\begin{aligned} |Q_n(-s)^{-1} (P^{-s} F)_n^+(s, x)| &= |q^{-(|x|-n)/2} \psi(|x| - n + 1, s)(D_n F)^+(s, \omega)| \\ &\leq C'_N q^{(|x|-n+N)|\Im s|} \end{aligned}$$

for some constant C'_N which does not depend on the choice of n. The condition (N5) is obtained in the same fashion as above. This concludes the proof. □

References

[Betori and Pagliacci 1984] W. Betori and M. Pagliacci, "Harmonic analysis for groups acting on trees", *Boll. Un. Mat. Ital. B* (6) **3**:2 (1984), 333–349. MR Zbl

[Bray 1996] W. O. Bray, "Generalized spectral projections on symmetric spaces of noncompact type: Paley–Wiener theorems", *J. Funct. Anal.* **135**:1 (1996), 206–232. MR Zbl

[Campoli 1980] O. A. Campoli, "Paley–Wiener type theorems for rank-1 semisimple Lie groups", *Rev. Un. Mat. Argentina* **29**:3 (1980), 197–221. MR Zbl

[Cowling and Setti 1999] M. Cowling and A. G. Setti, "The range of the Helgason–Fourier transformation on homogeneous trees", *Bull. Austral. Math. Soc.* **59**:2 (1999), 237–246. MR Zbl

[Cowling et al. 1998] M. Cowling, S. Meda, and A. G. Setti, "An overview of harmonic analysis on the group of isometries of a homogeneous tree", *Exposition. Math.* **16**:5 (1998), 385–423. MR Zbl

[Faraut and Picardello 1984] J. Faraut and M. A. Picardello, "The Plancherel measure for symmetric graphs", *Ann. Mat. Pura Appl.* (4) **138** (1984), 151–155. MR Zbl

[Iozzi and Picardello 1983a] A. Iozzi and M. A. Picardello, "Graphs and convolution operators", pp. 187–208 in *Topics in modern harmonic analysis, I* (Turin/Milan, 1982), edited by F. Ricci, Istituto Nazionale di Alta Matematica, Rome, 1983. MR

[Iozzi and Picardello 1983b] A. Iozzi and M. A. Picardello, "Spherical functions on symmetric graphs", pp. 344–386 in *Harmonic analysis* (Cortona, 1982), edited by F. Ricci and G. Weiss, Lecture Notes in Math. **992**, Springer, Berlin, 1983. MR

[Jamal Eddine 2013] A. Jamal Eddine, *Equations d'évolution sur certains groupes hyperboliques*, Thèse de doctorat, Université d'Orléans, 2013, http://tinyurl.com/ajeddine.

[Jamal Eddine 2015] A. Jamal Eddine, "An overview of harmonic analysis and the shifted wave equation on symmetric graphs", preprint, 2015. arXiv

[Johnson 1979] K. D. Johnson, "Paley–Wiener theorems on groups of split rank one", *J. Funct. Anal.* **34**:1 (1979), 54–71. MR Zbl

[Koizumi 2013] S. Koizumi, "An analogue of the spectral projection for homogeneous trees", *Hiroshima Math. J.* **43**:2 (2013), 207–221. MR Zbl

[Kuhn and Soardi 1983] G. Kuhn and P. M. Soardi, "The Plancherel measure for polygonal graphs", *Ann. Mat. Pura Appl.* (4) **134** (1983), 393–401. MR Zbl

[Mantero and Zappa 1983] A. M. Mantero and A. Zappa, "The Poisson transform and representations of a free group", *J. Funct. Anal.* **51**:3 (1983), 372–399. MR Zbl

Received August 28, 2016. Revised January 29, 2017.

SHIN KOIZUMI
FACULTY OF ECONOMY, MANAGEMENT AND INFORMATION SCIENCE
ONOMICHI CITY UNIVERSITY
1600 HISAYAMADA-CHO
ONOMICHI-SHI 722-8506
JAPAN
koizumi_qup_shin@live.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.139

FUNDAMENTAL DOMAINS OF ARITHMETIC QUOTIENTS OF REDUCTIVE GROUPS OVER NUMBER FIELDS

LEE TIM WENG
APPENDIX BY TAKAO WATANABE

For a connected reductive algebraic group G over a number field $\Bbbk$, we investigate the Ryshkov domain R_Q associated to a maximal $\Bbbk$-parabolic subgroup Q of G. By considering the arithmetic quotients $G(\Bbbk)\backslash G(\mathbb{A})^1/K$ and $\Gamma_i\backslash G(\Bbbk)/K_\infty$, with K a maximal compact subgroup of the adele group $G(\mathbb{A})$ and the Γ_i arithmetic subgroups of $G(\Bbbk)$, we present a method of constructing fundamental domains for $Q(\Bbbk)\backslash R_Q$ and $\Gamma_i\backslash G(\Bbbk_\infty)^1$. We also study the particular case when $G = \mathrm{GL}_n$, and subsequently construct fundamental domains for P_n, the cone of positive definite Humbert forms over $\Bbbk$, with respect to the subgroups Γ_i.

1. Introduction

Let $\Bbbk$ be an arbitrary algebraic number field with ring of integers $\mathcal{O}$. This paper mainly focuses on the determination and construction of fundamental domains associated to certain arithmetic quotients of reductive algebraic groups over $\Bbbk$.

For the first part of the paper we consider a general connected reductive isotropic algebraic group G over $\Bbbk$ and investigate fundamental domains associated to the arithmetic quotients $G(\Bbbk)\backslash G(\mathbb{A})^1/K$ and $\Gamma_i\backslash G(\Bbbk_\infty)^1/K_\infty$, with K a maximal compact subgroup of $G(\mathbb{A})$ and subgroups Γ_i of $G(\Bbbk)$ to be described below.

The discussion and results here in the preliminary sections are an extension of Watanabe's results [2014]. A maximal $\Bbbk$-parabolic subgroup Q of G is taken and we consider its associated height function H_Q and Hermite function $\mathsf{m}_Q(g) = \min_{x\in Q(\Bbbk)\backslash G(\Bbbk)} H_Q(xg)$ on $G(\mathbb{A})^1$. Watanabe [2014] introduced the Ryshkov domain of m_Q, $R_Q = \{g \in G(\mathbb{A})^1 : \mathsf{m}_Q(g) = H_Q(g)\}$, for the purpose of constructing a fundamental domain for $G(\Bbbk)\backslash G(\mathbb{A})^1$ well matched with m_Q. Watanabe also considered the case when G is of class number 1, that is, when $|G(\Bbbk)\backslash G(\mathbb{A})^1/G^1_{\mathbb{A},\infty}| = 1$, and obtained a fundamental domain for $G(\Bbbk_\infty)$ with respect to $G_{\mathcal{O}} = G(\Bbbk)\cap G_{\mathbb{A},\infty}$.

MSC2010: primary 11H55; secondary 11F06.
Keywords: fundamental domain, arithmetic quotient, Ryshkov domain, Humbert form, Voronoi reduction.

Here however, we consider algebraic groups of any general class number n_G. Particularly for class numbers higher than 1, for each $i = 1, \dots, n_G$ we are required to consider different arithmetic subgroups Γ_i of $G(\Bbbk)$ in place of just $G_{\mathcal{O}}$.

Let R_Q^* denote the closure in $G(\mathbb{A})^1$ of the interior of R_Q. It was established in [Watanabe 2014] that by starting from a fundamental domain Ω of R_Q^* with respect to $Q(\Bbbk)$, a fundamental domain of $G(\mathbb{A})^1$ with respect to $G(\Bbbk)$ can be obtained by taking the interior of Ω in $G(\mathbb{A})^1$. In order to explicitly construct such an Ω, we define groups

$$G_{\mathbb{A},\infty} = G(\Bbbk_\infty) \times K_f \quad \text{and} \quad \Gamma_i = \eta_i G^1_{\mathbb{A},\infty} \eta_i^{-1} \cap G(\Bbbk),$$

where the $\eta_1, \dots, \eta_{n_G}$ are representatives of $G(\Bbbk)\backslash G(\mathbb{A})^1/G^1_{\mathbb{A},\infty}$. Also for each i take a complete set of representatives $\{\xi_{ij}\}_{j=1}^{h_i}$ for $Q(\Bbbk)\backslash G(\Bbbk)/\Gamma_i$, define sets

$$R_{i,j,\infty} = \{g \in G(\Bbbk_\infty)^1 : \mathsf{m}_Q(g\xi_{ij}\eta_i) = H_Q(g\xi_{ij}\eta_i)\}$$

and let $Q_{i,j} = Q(\Bbbk) \cap \xi_{ij}\Gamma_i\xi_{ij}^{-1}$. By considering the action of $Q_{i,j}$ on $R_{i,j,\infty}$, we find that starting with arbitrary open fundamental domains $\Omega_{i,j,\infty}$ for $Q_{i,j}\backslash R_{i,j,\infty}$ we can construct the required Ω. From this we obtain the following results.

Theorem. $\Omega = \bigsqcup_{i=1}^{n_G} \bigsqcup_{j=1}^{h_i} \Omega_{i,j,\infty}\xi_{ij}\eta_i K_f$ *is an open fundamental domain of* R_Q^* *with respect to* $Q(\Bbbk)$.

Theorem. *For each* $i = 1, \dots, n_G$, *the set* $\bigcup_{j=1}^{h_i} \xi_{ij}^{-1}\Omega_{i,j,\infty}\xi_{ij}$ *is an open fundamental domain of* $G(\Bbbk_\infty)^1$ *with respect to* Γ_i.

In particular we can take η_1 to be the identity element of G, in which case Γ_1 coincides with the group $G_{\mathcal{O}} = G(\Bbbk) \cap G_{\mathbb{A},\infty}$ used in [Watanabe 2014] when $n_G = 1$.

The second topic of interest in this paper is the special case when G is the general linear group GL_n defined over $\Bbbk$. This time we consider the maximal $\Bbbk$-parabolic subgroup

$$Q = Q^{n,m} = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} : a \in \mathrm{GL}_m(\Bbbk),\ b \in M_{m,n-m}(\Bbbk),\ d \in \mathrm{GL}_{n-m}(\Bbbk) \right\}$$

for a fixed $1 \le m < n$. The class number of G in this case is equal to h, the class number of $\Bbbk$. Using $\{\mathfrak{a}_1, \dots, \mathfrak{a}_h\}$, a complete set of representatives for the ideal class group of $\Bbbk$, we can produce a corresponding set of matrices $\{\eta_1, \dots, \eta_h\}$ representing $\mathrm{GL}_n(\Bbbk)\backslash \mathrm{GL}_n(\mathbb{A})^1/G^1_{\mathbb{A},\infty}$. The Γ_i in this case are the subgroups of $\mathrm{GL}_n(\Bbbk)$ stabilizing the respective $\mathcal{O}$-lattices $\sum_{k=1}^{n-1} \mathcal{O}\boldsymbol{e}_k + \mathfrak{a}_i\boldsymbol{e}_n$. The main result established in this part is:

Theorem. $\quad |Q(\Bbbk)\backslash \mathrm{GL}_n(\Bbbk)/\Gamma_i| = h \quad$ *for every* $i = 1, \dots, h$.

This can be proved by identifying $Q(\Bbbk)\backslash \mathrm{GL}_n(\Bbbk)$ with the set of all m-dimensional subspaces of $\Bbbk^n$ and establishing a bijection between this set modulo Γ_i and the ideal class group of $\Bbbk$. This bijection also allows us to obtain suitable matrix representatives $\{\xi_{ij}\}_{j=1}^{h}$ for $Q(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk)/\Gamma_i$. Relations between the field class number and the number of double cosets in quotients of similar type involving other algebraic groups, e.g., SL_n, Sp_{2n} and Chevalley groups, modulo a minimal parabolic subgroup instead are noted by Borel [1962, Section 4.7].

In the final sections we consider P_n, the space of positive definite Humbert forms over $\Bbbk$, with the usual identification $P_n = \prod_\sigma P_n(\Bbbk_\sigma)$, where $P_n(\Bbbk_\sigma)$ denotes the set of $n \times n$ positive definite real symmetric/complex Hermitian matrices, depending on whether σ is real or imaginary, and the product is taken over all infinite places σ of $\Bbbk$.

If $\Bbbk = \mathbb{Q}$, then P_n is just the cone of positive definite real symmetric matrices, and fundamental domains for $P_n/\mathrm{GL}_n(\mathbb{Z})$ in this case have been historically constructed by Korkin and Zolotarev [1873], Minkowski [1905] and later on Grenier [1988]. For P_n over a general number field, Humbert [1939] previously provided a fundamental domain constructed with respect to the particular group $\mathrm{GL}_n(\mathcal{O})$. As $\mathrm{GL}_n(\mathcal{O})$ coincides with one of the Γ_i we study in this paper, the question can be raised about fundamental domains for P_n with respect to each of the groups Γ_i when $n_G > 1$.

As such, we proceed in the final sections to provide a general way of constructing fundamental domains for P_n/Γ_i given any number field. The method of construction given here follows and generalizes the example given by Watanabe [2014] for the specific case $\Bbbk = \mathbb{Q}$. As already noted in that paper, when $\Bbbk = \mathbb{Q}$ the fundamental domain for $P_n/\mathrm{GL}_n(\mathbb{Z})$ resulting from this method coincides with Grenier's [1988]. It was observed by Dutour Sikirić and Schürmann that Grenier's fundamental domain is in fact equivalent to the one previously developed by Korkin and Zolotarev. Regarding $P_n/\mathrm{GL}_n(\mathcal{O})$ for general number fields however, we note that the fundamental domain produced by the method here differs from Humbert's construction, which utilizes the matrix trace, whereas the domain here is defined using the adele norm of matrix determinants.

Using the matrix representatives $\{\eta_i\}_{i=1}^{h}$ and $\{\xi_{ij}\}_{j=1}^{h}$, we associate to each pair (η_i, ξ_{ij}) a maximal compact subgroup $K_{i,j,\infty}$ of $\mathrm{GL}_n(\Bbbk_\infty)$ and a map π_{ij} inducing an isomorphism between $\mathrm{GL}_n(\Bbbk_\infty)/K_{i,j,\infty}$ and P_n. Then the results of our discussions on GL_n can be transferred to P_n via the maps π_{ij}, which finally lead up to an iterative method of constructing fundamental domains for P_n with respect to the groups Γ_i for any general dimension n. Watanabe has also graciously provided an appendix to this paper on Voronoi reduction over general number fields that are not necessarily totally real, which settles the base case of dimension 1.

We also demonstrate that this fundamental domain construction for P_n/Γ_i is well matched with certain automorphisms of $\mathrm{GL}(\Bbbk_\infty)$. Namely we see that the fundamental domain for P_n/Γ_i constructed using a set of ideals $\{\mathfrak{a}_1, \ldots, \mathfrak{a}_h\}$ representing the ideal class group and the maximal $\Bbbk$-parabolic subgroup $Q^{n,m}$ can be directly mapped by an automorphism to the one constructed with the representative set $\{\mathfrak{a}_1^{-1}, \ldots, \mathfrak{a}_h^{-1}\}$ and $Q^{n,n-m}$.

Notation

In this paper we use $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ for the fields of rational, real, and complex numbers respectively, and $\mathbb{Z}$ for the ring of integers. $\mathbb{R}_{>0}$ will denote the set of positive reals.

For positive integers r and s, we denote by $M_{r,s}(S)$ the set of all $r \times s$ matrices with entries in the set S, and we write $M_r(S)$ for $M_{r,r}(S)$. The identity matrix of size r will be denoted by I_r. The transpose of a matrix A will be written by ${}^tA$. If $A \in M_{r,s}(\mathbb{C})$, we write $\bar{A}$ for the matrix whose entries are the complex conjugates of the original entries of A.

We will fix and consider $\mathbb{k}$, an algebraic number field of finite degree over $\mathbb{Q}$, and denote its ring of integers by $\mathcal{O}$. We denote by $\boldsymbol{p}_\infty$ and $\boldsymbol{p}_f$ the sets of infinite and finite places of $\mathbb{k}$ respectively and we let $\boldsymbol{p} = \boldsymbol{p}_\infty \cup \boldsymbol{p}_f$. For $\sigma \in \boldsymbol{p}$, we write $\mathbb{k}_\sigma$ for the completion of $\mathbb{k}$ at σ, while for any subring $\mathbb{B}$ of $\mathbb{k}$, the closure of $\mathbb{B}$ in $\mathbb{k}_\sigma$ will be denoted by $\mathbb{B}_\sigma$. We denote by $\mathbb{k}_\infty$ the étale $\mathbb{R}$-algebra $\mathbb{k} \otimes_{\mathbb{Q}} \mathbb{R}$ which we identify with $\prod_{\sigma \in \boldsymbol{p}_\infty} \mathbb{k}_\sigma$. The ideal class group of $\mathbb{k}$ will be denoted by $\mathrm{Cl}(\mathbb{k})$.

The adele ring and idele group of $\mathbb{k}$ are denoted by $\mathbb{A}$ and $\mathbb{A}^\times$ respectively. For an adele $a \in \mathbb{A}$ we write a_∞ and a_f for its infinite and finite components respectively. Similarly for any matrix $A = [a_{ij}]_{i,j}$ with elements in $\mathbb{A}$ we write A_∞ to denote the matrix $[(a_{ij})_\infty]_{i,j}$.

For each place σ, we write $|\ |_\sigma$ for the absolute value on $\mathbb{k}_\sigma$ taken as follows: at each infinite place we use the standard complex absolute value on $\mathbb{k}_\sigma$, while for $\sigma \in \boldsymbol{p}_f$ we use the normalized absolute value satisfying $|x|_\sigma = |\mathcal{O}_\sigma/\mathfrak{p}_\sigma|^{-1}$ for any arbitrary $x \in \mathfrak{p}_\sigma \setminus \mathfrak{p}_\sigma^2$, where $\mathfrak{p}_\sigma$ is the prime ideal of $\mathcal{O}_\sigma$. For an $a = (a_\sigma) \in \mathbb{A}^\times$ we write $|a|_{\mathbb{A}}$ to denote the idele norm of a, and $|a|_\infty$ for the idele norm of a restricted to $\mathbb{k}_\infty^\times$, $\prod_{\sigma \in \boldsymbol{p}_\infty} |a_\sigma|_\sigma^{[\mathbb{k}_\sigma:\mathbb{R}]}$.

Given a finite-dimensional $\mathbb{k}$-vector space V and $\sigma \in \boldsymbol{p}$, we will write V_σ for the $\mathbb{k}_\sigma$-vector space $V \otimes_{\mathbb{k}} \mathbb{k}_\sigma$. Also we will use the term *$\mathcal{O}$-lattice* in V to mean a finitely generated $\mathcal{O}$-submodule of V containing a $\mathbb{k}$-basis of V. If L is such an $\mathcal{O}$-lattice in V, we write L_σ to denote the $\mathcal{O}_\sigma$-linear span of L in V_σ when $\sigma \in \boldsymbol{p}_f$.

For an affine algebraic group G defined over $\mathbb{k}$ and any $\mathbb{k}$-algebra $\mathbb{B}$, we write $G(\mathbb{B})$ for the set of all $\mathbb{B}$-rational points of G. Also, the set of all $\mathbb{k}$-rational characters of G will be written as $\boldsymbol{X}^*(G)_{\mathbb{k}}$. We define $G(\mathbb{A})^1$ to be the set $\{g \in G(\mathbb{A}) : |\chi(g)|_{\mathbb{A}} = 1 \text{ for all } \chi \in \boldsymbol{X}^*(G)_{\mathbb{k}}\}$.

Lastly given a topological space X and a subset $Y \subset X$, we denote by Y_X° and Y_X^- (or just Y° and Y^- if the underlying space X is clear) the interior and closure of Y in X respectively.

2. The Ryshkov domain of G associated to Q

Let G denote a connected reductive isotropic affine algebraic group over $\mathbb{k}$, S a fixed maximal $\mathbb{k}$-split torus of G, and P_0 a minimal $\mathbb{k}$-parabolic subgroup of G

containing S. Let M_0 be the centralizer of S in G and U_0 the unipotent radical of P_0 so that P_0 has the Levi decomposition $P_0 = M_0U_0$. We consider a relative root system of G with respect to S and denote the set of simple roots with respect to P_0 in this system by $\Delta_{\Bbbk}$.

A $\Bbbk$-parabolic subgroup of G containing P_0 is called a standard $\Bbbk$-parabolic subgroup. A standard $\Bbbk$-parabolic subgroup R has a unique Levi subgroup M_R containing M_0, which gives the Levi decomposition $R = M_RU_R$, where U_R denotes the unipotent radical of R. We write Z_R for the largest central $\Bbbk$-split torus of M_R.

We fix a maximal compact subgroup $K = \prod_{\sigma\in \boldsymbol{p}} K_\sigma$ of $G(\mathbb{A})$, where each K_σ is a maximal compact subgroup of $G(\Bbbk_\sigma)$, satisfying the property that for every standard $\Bbbk$-parabolic subgroup R of G,

- $K \cap M_R(\mathbb{A})$ is a maximal compact subgroup in $M_R(\mathbb{A})$,
- $M_R(\mathbb{A}) = (M_R(\mathbb{A}) \cap U_0(\mathbb{A}))\, M_0(\mathbb{A})\, (K \cap M_R(\mathbb{A}))$ (Iwasawa decomposition) holds.

Consider a standard proper maximal $\Bbbk$-parabolic subgroup Q of G, which we now fix. There exists a unique simple root in $\Delta_{\Bbbk}$ that restricts nontrivially on Z_Q, which we denote by χ_0. Let m_Q be the positive integer such that $m_Q^{-1}\chi_0|_{Z_Q}$ is a $\mathbb{Z}$-basis of the $\boldsymbol{X}^*(Z_Q/Z_G)_{\Bbbk}$. We write χ_Q for the character

$$[\boldsymbol{X}^*(Z_Q/Z_G)_{\Bbbk} : \boldsymbol{X}^*(M_Q/Z_G)_{\Bbbk}]\, m_Q^{-1}(\chi_0|_{Z_Q}),$$

which is a $\mathbb{Z}$-basis for $\boldsymbol{X}^*(M_Q/Z_G)_{\Bbbk}$.

Next we define the map

$$z_Q : G(\mathbb{A}) \ni umh \longmapsto Z_G(\mathbb{A})M_Q(\mathbb{A})^1 m \in Z_G(\mathbb{A})M_Q(\mathbb{A})^1\backslash M_Q(\mathbb{A}),$$

where $u \in U_Q(\mathbb{A}),\ m \in M_Q(\mathbb{A}),\ h \in K$. This is a well-defined left $Q(\mathbb{A})^1$-invariant map, which gives rise to the following map, which we also denote by z_Q:

$$Q(\mathbb{A})^1\backslash G(\mathbb{A})^1 \ni Q(\mathbb{A})^1 g \longmapsto z_Q(g) \in M_Q(\mathbb{A})^1\backslash (M_Q(\mathbb{A}) \cap G(\mathbb{A})^1).$$

Here we have used $Z_G(\mathbb{A})^1 = Z_G(\mathbb{A}) \cap G(\mathbb{A})^1 \subset M_Q(\mathbb{A})^1$.

We can now define the *height function* $H_Q : G(\mathbb{A}) \to \mathbb{R}_{>0}$ by

$$H_Q(g) = |\chi_Q(z_Q(g))|_{\mathbb{A}}^{-1}, \quad g \in G(\mathbb{A}),$$

as well as the *Hermite function* $\mathsf{m}_Q : G(\mathbb{A})^1 \to \mathbb{R}_{>0}$ by

$$\mathsf{m}_Q(g) = \min_{x \in Q(\Bbbk)\backslash G(\Bbbk)} H_Q(xg), \quad g \in G(\mathbb{A})^1.$$

Definition [Watanabe 2014, §4]. The set R_Q defined by

$$\{g \in G(\mathbb{A})^1 : \mathsf{m}_Q(g) = H_Q(g)\}$$

is called the *Ryshkov domain of* m_Q.

3. Fundamental domains of $G(\Bbbk)\backslash G(\mathbb{A})^1$ and $\Gamma_i\backslash G(\Bbbk_\infty)^1$

Definition. Let T be a locally compact Hausdorff space and Γ a discrete group with a properly discontinuous action on T. An open subset Ω of T satisfying

(i) $T = \Gamma\Omega^-$,

(ii) $\Omega \cap \gamma\Omega^- = \varnothing$ for all $\gamma \in \Gamma \setminus \{e\}$

is called an *open fundamental domain of T with respect to* Γ. (Here we have assumed that Γ acts on T from the left. In the case of a right action the same definition holds with the group action written on the right instead.)

We call a subset F of T a *fundamental domain of T with respect to* Γ, or simply a *fundamental domain of* $\Gamma\backslash T$ (T/Γ in the case of a right action) if there exists an open fundamental domain Ω of T with respect to Γ such that $\Omega \subset F \subset \Omega^-$.

Further Notation. Hereafter we will use the following notation:

- $K_\infty = \prod_{\sigma\in \boldsymbol{p}_\infty} K_\sigma$, $K_f = \prod_{\sigma\in \boldsymbol{p}_f} K_\sigma$,
- $G_{\mathbb{A},\infty} = G(\Bbbk_\infty) \times K_f$, $G^1_{\mathbb{A},\infty} = G_{\mathbb{A},\infty} \cap G(\mathbb{A})^1$,
- $G(\Bbbk_\infty)^1 = G(\Bbbk_\infty) \cap G(\mathbb{A})^1$, where we identify $G(\Bbbk_\infty)$ with the subgroup $\{g \in G(\mathbb{A}) : g_f = e\}$ of $G(\mathbb{A})$.

We will denote the *class number of* G, i.e., the finite number $|G(\Bbbk)\backslash G(\mathbb{A})/G_{\mathbb{A},\infty}|$, by n_G. We note here that $|G(\Bbbk)\backslash G(\mathbb{A})^1/G^1_{\mathbb{A},\infty}|$ is also equal to n_G.

The case when G is of class number 1 is discussed in [Watanabe 2014], where a fundamental domain for $G(\Bbbk_\infty)^1$ with respect to the group $G(\Bbbk) \cap G_{\mathbb{A},\infty}$ is determined. In the following we discuss and obtain a similar fundamental domain in the general case.

We take a complete set of representatives $\{\eta_1, \dots, \eta_{n_G}\}$ for $G(\Bbbk)\backslash G(\mathbb{A})^1/G^1_{\mathbb{A},\infty}$. Then, for $i = 1, \dots, n_G$, define the groups

$$G_i = \eta_i G^1_{\mathbb{A},\infty} \eta_i^{-1} \quad \text{and} \quad \Gamma_i = G_i \cap G(\Bbbk).$$

We note that since $(\eta_i)_\infty G(\Bbbk_\infty)^1 (\eta_i)_\infty^{-1} = G(\Bbbk_\infty)^1$, we can also write G_i as $G(\Bbbk_\infty)^1 \times (\eta_i)_f K_f (\eta_i)_f^{-1}$ or $G(\Bbbk_\infty)^1 \eta_i K_f \eta_i^{-1}$.

From $G(\mathbb{A})^1 = \bigsqcup_{i=1}^{n_G} G(\Bbbk)\eta_i G^1_{\mathbb{A},\infty} = \bigsqcup_{i=1}^{n_G} G(\Bbbk) G_i \eta_i$ we have

$$G(\Bbbk)\backslash G(\mathbb{A})^1 = \bigsqcup_{i=1}^{n_G} \Gamma_i\backslash G_i\eta_i = \bigsqcup_{i=1}^{n_G} \Gamma_i\backslash (G(\Bbbk_\infty)^1 \eta_i K_f),$$

which gives us the isomorphism

$$G(\Bbbk)\backslash G(\mathbb{A})^1/K \simeq \bigsqcup_{i=1}^{n_G} \Gamma_i\backslash G(\Bbbk_\infty)^1/K_\infty.$$

Also for each $i = 1, \dots, n_G$ we take a complete set of representatives $\{\xi_{ij}\}_{j=1}^{h_i}$ for $Q(\Bbbk)\backslash G(\Bbbk)/\Gamma_i$ (where the number of double cosets h_i is finite; see [Borel 1963, §7]) and define groups

$$Q_{i,j} = Q \cap \xi_{ij}\Gamma_i\xi_{ij}^{-1} = Q(\Bbbk) \cap \xi_{ij}G_i\xi_{ij}^{-1}$$

and the sets

$$R_{i,j,\infty} = \{g \in G(\Bbbk_\infty)^1 : \mathsf{m}_Q(g\xi_{ij}\eta_i) = H_Q(g\xi_{ij}\eta_i)\}$$

for $j = 1, \dots, h_i$. Also since $G_i = G(\Bbbk_\infty)^1\eta_i K_f\eta_i^{-1}$ as previously noted,

$$\xi_{ij}G_i\xi_{ij}^{-1} = \xi_{ij}G(\Bbbk_\infty)^1\eta_i K_f\eta_i^{-1}\xi_{ij}^{-1} = G(\Bbbk_\infty)^1\xi_{ij}\eta_i K_f\eta_i^{-1}\xi_{ij}^{-1}.$$

Lemma 1. $$G(\mathbb{A})^1 = \bigsqcup_{i=1}^{n_G}\bigsqcup_{j=1}^{h_i} Q(\Bbbk)G(\Bbbk_\infty)^1\xi_{ij}\eta_i K_f.$$

Proof. We first show that for a fixed i the union $\bigcup_{j=1}^{h_i} Q(\Bbbk)\xi_{ij}G_i\eta_i$ is disjoint. Suppose for some $1 \le j, j' \le h_i$ that $Q(\Bbbk)\xi_{ij}G_i\eta_i \cap Q(\Bbbk)\xi_{ij'}G_i\eta_i$ is nonempty. Then there exist $q, q' \in Q(\Bbbk)$ and $g, g' \in G_i$ such that $q\xi_{ij}g = q'\xi_{ij'}g'$. Rearranging gives us $gg'^{-1} = \xi_{ij}^{-1}q^{-1}q'\xi_{ij'} \in G_i \cap G(\Bbbk) = \Gamma_i$. This shows that $Q(\Bbbk)\xi_{ij'}\Gamma_i = Q(\Bbbk)\xi_{ij}\Gamma_i$, implying $j = j'$. The result then follows from

$$\begin{aligned} G(\mathbb{A})^1 &= \bigsqcup_i G(\Bbbk)\eta_i G^1_{\mathbb{A},\infty} = \bigsqcup_i G(\Bbbk)G_i\eta_i \\ &= \bigsqcup_i\Big(\bigsqcup_j Q(\Bbbk)\xi_{ij}\Gamma_i\Big)G_i\eta_i \subset \bigsqcup_i\bigsqcup_j Q(\Bbbk)\xi_{ij}G_i\eta_i \end{aligned}$$

and $\xi_{ij}G_i\eta_i = G(\Bbbk_\infty)^1\xi_{ij}\eta_i K_f$. □

The lemma also gives us the disjointedness of the union in the following result.

Proposition 2. $$R_Q = \bigsqcup_{i=1}^{n_G}\bigsqcup_{j=1}^{h_i} Q(\Bbbk)R_{i,j,\infty}\xi_{ij}\eta_i K_f.$$

Proof. From the previous lemma, we see that any $g \in G(\mathbb{A})^1$ can be written as $qg'\xi_{ij}\eta_i h$ for some i, j and $q \in Q(\Bbbk)$, $g' \in G(\Bbbk_\infty)^1$, $h \in K_f$. Since both H_Q and m_Q are left $Q(\Bbbk)$-invariant and right K-invariant, we see that

$$H_Q(g) = H_Q(g'\xi_{ij}\eta_i), \quad \mathsf{m}_Q(g) = \mathsf{m}_Q(g'\xi_{ij}\eta_i).$$

Hence $g \in R_Q$ if and only if $g' \in R_{i,j,\infty}$. □

The following two lemmas hold for any fixed $1 \le i \le n_G$ and $1 \le j \le h_i$.

Lemma 3. *Let $q \in Q(\Bbbk)$. If the sets $q(G(\Bbbk_\infty)^1\xi_{ij}\eta_i K_f)$ and $G(\Bbbk_\infty)^1\xi_{ij}\eta_i K_f$ intersect, then $q \in Q_{i,j}$.*

Proof. Suppose that $g \in q(G(\Bbbk_\infty)^1 \xi_{ij}\eta_i K_f) \cap (G(\Bbbk_\infty)^1 \xi_{ij}\eta_i K_f)$. By rewriting $G(\Bbbk_\infty)^1 \xi_{ij}\eta_i K_f$ as $\xi_{ij} G_i \eta_i$, we have $q^{-1}g,\ g \in \xi_{ij} G_i \eta_i$, from which we get $q^{-1} \in \xi_{ij} G_i \xi_{ij}^{-1}$. Hence $q \in Q(\Bbbk) \cap \xi_{ij} G_i \xi_{ij}^{-1} = Q_{i,j}$. □

Lemma 4. $$Q_{i,j}(R_{i,j,\infty}\xi_{ij}\eta_i K_f) = R_{i,j,\infty}\xi_{ij}\eta_i K_f.$$

Proof. Consider $q \in Q_{i,j}$ and $g \in R_{i,j,\infty}$. Since $q \in G(\Bbbk_\infty)^1 \xi_{ij}\eta_i K_f \eta_i^{-1}\xi_{ij}^{-1}$, we have $q_f \in (\xi_{ij}\eta_i) K_f (\xi_{ij}\eta_i)^{-1}$. Let $q_f = (\xi_{ij}\eta_i) h (\xi_{ij}\eta_i)^{-1}$, with $h \in K_f$. Then

$$H_Q((q_\infty g)\xi_{ij}\eta_i) = H_Q(q_\infty g(\xi_{ij}\eta_i)h) = H_Q(q_\infty g q_f(\xi_{ij}\eta_i)) = H_Q(qg\xi_{ij}\eta_i),$$

which is equal to $H_Q(g\xi_{ij}\eta_i)$. Similarly

$$\mathsf{m}_Q((q_\infty g)\xi_{ij}\eta_i) = \mathsf{m}_Q(q_\infty g q_f \xi_{ij}\eta_i) = \mathsf{m}_Q(qg\xi_{ij}\eta_i) = \mathsf{m}_Q(g\xi_{ij}\eta_i);$$

thus $q_\infty g \in R_{i,j,\infty}$. Finally $q_f \xi_{ij}\eta_i K_f \subset \xi_{ij}\eta_i K_f$. Hence we get $q(g\xi_{ij}\eta_i K_f) \subset R_{i,j,\infty}\xi_{ij}\eta_i K_f$, as required. □

By taking a complete set of representatives $\{\theta_{ijk}\}_k$ for $Q(\Bbbk)/Q_{i,j}$ and using both Proposition 2 and Lemma 4, we obtain

$$\begin{aligned} R_Q &= \bigsqcup_{i=1}^{n_G}\bigsqcup_{j=1}^{h_i} Q(\Bbbk) R_{i,j,\infty}\xi_{ij}\eta_i K_f = \bigsqcup_{i=1}^{n_G}\bigsqcup_{j=1}^{h_i}\Big(\bigsqcup_k \theta_{ijk} Q_{i,j}\Big) R_{i,j,\infty}\xi_{ij}\eta_i K_f \\ &= \bigsqcup_{i=1}^{n_G}\bigsqcup_{j=1}^{h_i}\bigsqcup_k \theta_{ijk} R_{i,j,\infty}\xi_{ij}\eta_i K_f, \end{aligned} \tag{1}$$

where the final unions are disjoint as a result of Lemma 3.

Denote $(R_{i,j,\infty}^\circ)^-$ by $R_{i,j,\infty}^*$, where the interior and closure is taken in $G(\Bbbk_\infty)^1$. Similarly write R_Q^* for $(R_Q^\circ)^-$ in $G(\mathbb{A})^1$. From (1) we have

$$R_Q^* = \bigsqcup_{i=1}^{n_G}\bigsqcup_{j=1}^{h_i}\bigsqcup_k \theta_{ijk} R_{i,j,\infty}^*\xi_{ij}\eta_i K_f. \tag{2}$$

Taking open fundamental domains $\Omega_{i,j,\infty}$ of $R_{i,j,\infty}^*$ with respect to $Q_{i,j}$ for each $i = 1, \ldots, n_G$ and $j = 1, \ldots, h_i$, we consider the set

$$\Omega = \bigsqcup_{i=1}^{n_G}\bigsqcup_{j=1}^{h_i} \Omega_{i,j,\infty}\xi_{ij}\eta_i K_f.$$

Theorem 5. *Ω is an open fundamental domain of R_Q^* with respect to $Q(\Bbbk)$.*

Corollary 6. *Ω° $(= \Omega^\circ_{G(\mathbb{A})^1})$ is an open fundamental domain of $G(\mathbb{A})^1$ with respect to $G(\Bbbk)$.*

Proof. From (2) we have

$$R_Q^* = \bigsqcup_{i=1}^{n_G} \bigsqcup_{j=1}^{h_i} \bigsqcup_{k} \theta_{ijk} R_{i,j,\infty}^* \xi_{ij} \eta_i K_f = \bigsqcup_{i=1}^{n_G} \bigsqcup_{j=1}^{h_i} \bigsqcup_{k} \theta_{ijk} (Q_{i,j} \Omega_{i,j,\infty}^-) \xi_{ij} \eta_i K_f$$
$$= \bigsqcup_{i=1}^{n_G} \bigsqcup_{j=1}^{h_i} Q(\Bbbk) \Omega_{i,j,\infty}^- \eta_i K_f = Q(\Bbbk)\Omega^-.$$

Now suppose $\Omega \cap q\Omega^- \neq \varnothing$ for $q \in Q(\Bbbk)$. So for some i, i', j, j' we must have $q(\Omega_{i,j,\infty}\xi_{ij}\eta_i K_f) \cap (\Omega_{i',j',\infty}^- \xi_{i'j'}\eta_{i'} K_f) \neq \varnothing$. Writing $q = \theta_{ijk} q'$ with $q' \in Q_{i,j}$ and some k, we have

$$\theta_{ijk}(q')_\infty \Omega_{i,j,\infty} \xi_{ij} \eta_i K_f \cap \Omega_{i',j',\infty}^- \xi_{i'j'} \eta_{i'} K_f \neq \varnothing$$

since $(q')_f \xi_{ij}\eta_i K_f \subset \xi_{ij}\eta_i K_f$. Then (2) implies $i = i'$, $j = j'$, and $\theta_{ijk} = e$. Thus $\Omega_{i,j,\infty} \cap (q')_\infty \Omega_{i,j,\infty}^- = \Omega_{i,j,\infty} \cap q'\Omega_{i,j,\infty}^-$ must be nonempty, which means $q' = e$ and hence $q = e$. This proves the theorem, and the corollary follows from [Watanabe 2014, Theorem 15]. □

Finally, for any fixed $1 \le i \le n_G$, we have the following theorem.

Theorem 7. *The set* $\Omega_{i,\infty} = \bigcup_{j=1}^{h_i} \xi_{ij}^{-1} \Omega_{i,j,\infty} \xi_{ij}$ *is a fundamental domain of* $G(\Bbbk_\infty)^1$ *with respect to* Γ_i.

Proof. The following proof was suggested by Professor Watanabe. To show that $G(\Bbbk_\infty)^1 = \Gamma_i \Omega_{i,\infty}^-$, consider an arbitrary $g \in G(\Bbbk_\infty)^1$. From Corollary 6,

$$G(\mathbb{A})^1 = G(\Bbbk)\Omega^- = G(\Bbbk) \bigsqcup_{i=1}^{n_G} \bigsqcup_{j=1}^{h_i} \Omega_{i,j,\infty}^- \xi_{ij} \eta_i K_f$$
$$= G(\Bbbk) \bigsqcup_{i=1}^{n_G} \bigsqcup_{j=1}^{h_i} \xi_{ij} (\xi_{ij}^{-1} \Omega_{i,j,\infty}^- \xi_{ij}) \eta_i K_f \subset G(\Bbbk) \bigcup_{i=1}^{n_G} \Omega_{i,\infty}^- \eta_i K_f,$$

so we may write $g\eta_i = g'\omega\eta_i h$ with $g' \in G(\Bbbk)$, $\omega \in \Omega_{i,\infty}^-$ and $h \in K_f$. Rearranging we get $g' = (g\omega^{-1})(\eta_i h^{-1} \eta_i^{-1})$, which belongs to $G(\Bbbk_\infty)^1 \eta_i K_f \eta_i^{-1} = G_i$. Hence $g' \in \Gamma_i$. Since $g = (g'\omega)(\eta_i h \eta_i^{-1})$ and $g \in G(\Bbbk_\infty)^1$, we know $\eta_i h \eta_i^{-1}$ must necessarily be trivial. Thus $g \in \Gamma_i \Omega_{i,\infty}^-$.

Now suppose that $\Omega_{i,\infty}^\circ \cap g\Omega_{i,\infty}^-$ is nonempty for a $g \in \Gamma_i$. Then we must have $\xi_{ij}^{-1}\Omega_{i,j,\infty}^\circ \xi_{ij} \cap g\xi_{ij'}^{-1}\Omega_{i,,j',\infty}^- \xi_{ij'} \neq \varnothing$ for some j, j'. Since $g_f \eta_i K_f = \eta_i K_f$,

$$\xi_{ij}^{-1}\Omega_{i,j,\infty}^\circ \xi_{ij} \cap g\xi_{ij'}^{-1}\Omega_{i,,j',\infty}^- \xi_{ij'} \neq \varnothing$$
$$\Rightarrow \quad (\Omega_{i,j,\infty}\xi_{ij}\eta_i K_f)^\circ \cap \xi_{ij} g \xi_{ij'}^{-1} (\Omega_{i,j',\infty}\xi_{ij'}\eta_i K_f)^- \neq \varnothing$$
$$\Rightarrow \quad \Omega^\circ \cap (\xi_{ij} g \xi_{ij'}^{-1})\Omega^- \neq \varnothing,$$

and thus $\xi_{ij} g \xi_{ij'}^{-1} = e$ by Corollary 6. Hence $Q(\Bbbk)\xi_{ij}\Gamma_i = Q(\Bbbk)\xi_{ij'}\Gamma_i$, which implies $j = j'$ whereby $g = \xi_{ij}^{-1}\xi_{ij'} = e$. $\square$

4. The case $G = \mathrm{GL}_n$

We will now consider the case where G is a general linear group GL_n defined over $\Bbbk$. We use the group of diagonal matrices as the maximal $\Bbbk$-split torus S, and the group of upper triangular matrices in G as the minimal $\Bbbk$-parabolic subgroup P_0. Also fixing an integer $1 \le m < n$, we will consider the maximal standard $\Bbbk$-parabolic subgroup Q defined by

$$Q(\Bbbk) = \left\{ \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} : a \in \mathrm{GL}_m(\Bbbk), b \in M_{m,n-m}(\Bbbk), d \in \mathrm{GL}_{n-m}(\Bbbk) \right\}$$

and the Levi subgroup M_Q is given by

$$M_Q(\Bbbk) = \left\{ \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} : a \in \mathrm{GL}_m(\Bbbk), d \in \mathrm{GL}_{n-m}(\Bbbk) \right\}.$$

For the maximal compact subgroup K of $G(\mathbb{A})$ let $K = K_\infty \times K_f$, where

$$K_\infty = \{ g \in \mathrm{GL}_n(\Bbbk_\infty) : {}^t\bar{g} g = I_n \}, \quad K_f = \prod_{\sigma \in \boldsymbol{p}_f} \mathrm{GL}_n(\mathcal{O}_\sigma).$$

Here we identify $\mathrm{GL}_n(\Bbbk_\infty)$ with $\prod_{\sigma \in \boldsymbol{p}_\infty} \mathrm{GL}_n(\Bbbk_\sigma)$, and for $g = (g_\sigma)_{\sigma \in \boldsymbol{p}_\infty} \in \mathrm{GL}_n(\Bbbk_\infty)$ we write ${}^t\bar{g}$ for the element $({}^t\bar{g}_\sigma)_{\sigma \in \boldsymbol{p}_\infty}$ of $\mathrm{GL}_n(\Bbbk_\infty)$.

The character χ_Q described in the first section is then given by

$$\chi_Q\left(\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \right) = (\det a)^{(n-m)/l} (\det d)^{-m/l}$$

and the height function H_Q by

$$H_Q\left(\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \right) = |\det a|_{\mathbb{A}}^{-(n-m)/l} |\det d|_{\mathbb{A}}^{m/l},$$

where l is the greatest common divisor of $n - m$ and m.

We shall see that in this case the number of double cosets of $Q(\Bbbk)\backslash \mathrm{GL}_n(\Bbbk)/\Gamma_i$ for each i is invariant and equal to $|\mathrm{GL}_n(\Bbbk)\backslash \mathrm{GL}_n(\mathbb{A})^1 / G^1_{\mathbb{A},\infty}|$, the class number of GL_n.

Denote the set of all $\mathcal{O}$-lattices in $\Bbbk^r$ ($r \ge 1$) by $\mathfrak{L}_r$, and the standard unit vectors of $\Bbbk^r$ by $\boldsymbol{e}_1^{(r)}, \ldots, \boldsymbol{e}_r^{(r)}$. For this section we simply write $\mathfrak{L}$ for $\mathfrak{L}_n$ and $\boldsymbol{e}_k$ for $\boldsymbol{e}_k^{(n)}$ $(1 \le k \le n)$.

For $L \in \mathfrak{L}_r$ and $g = (g_\sigma)_{\sigma \in \boldsymbol{p}} \in \mathrm{GL}_r(\mathbb{A})$ put

$$gL = \left((\Bbbk_\infty)^r \times \prod_{\sigma \in \boldsymbol{p}_f} g_\sigma L_\sigma \right) \cap \Bbbk^r \in \mathfrak{L}_r. \tag{3}$$

This defines a transitive left action of $\mathrm{GL}_r(\mathbb{A})^1$ on $\mathfrak{L}_r$. Note that if $g \in \mathrm{GL}_r(\Bbbk)$ then gL as defined above coincides with the usual image of L under the linear transformation $v \mapsto gv$ of $\Bbbk^r$. The subset of $\mathfrak{L}$ consisting of all $\mathcal{O}$-lattices of the form gL with $g \in \mathrm{GL}_n(\Bbbk)$ will be referred to as the *$\mathcal{O}$-lattice class of L* or just the lattice class of L in $\mathfrak{L}$.

There is known to be a one-to-one correspondence between the $\mathcal{O}$-lattice classes in $\mathfrak{L}$ and the double cosets in $\mathrm{GL}_n(\Bbbk)\backslash\mathrm{GL}_n(\mathbb{A})^1/G^1_{\mathbb{A},\infty}$, which we give explicitly later on in this section. For now we note that this means the number of distinct lattice classes in $\mathfrak{L}$ and the class number $|\mathrm{GL}_n(\Bbbk)\backslash\mathrm{GL}_n(\mathbb{A})^1/G^1_{\mathbb{A},\infty}|$ are equal.

Lemma 8. *Let L be an $\mathcal{O}$-lattice in a $\Bbbk$-vector space V of dimension $s \geq 1$. Then there exists a $\Bbbk$-basis $\{x_j\}_{j=1}^s$ of V and s fractional ideals $A_1, \dots, A_s$ such that $L = A_1x_1 + \cdots + A_sx_s$. Moreover*:

(i) *If W is a $\Bbbk$-subspace of V of dimension $r \leq s$, the x_j can be chosen such that $x_1, \dots, x_r \in W$.*

(ii) *The ideal class of $A_1 \cdots A_s$ is uniquely determined by the isomorphism class of L as an $\mathcal{O}$-module. In particular, $L \simeq \left(\bigoplus_{j=1}^{s-1} \mathcal{O}\right) \oplus (A_1 \cdots A_s)$.*

(iii) *In the case $V \subseteq \Bbbk^n$ $(s \leq n)$, we can find $g \in \mathrm{GL}_n(\Bbbk)$ such that*

$$gL = \left(\sum_{j=1}^{s-1} \mathcal{O}\boldsymbol{e}_j\right) + (A_1 \cdots A_s)\boldsymbol{e}_s.$$

Proof. See [Shimura 2010, Theorem 10.19]. We prove (iii) here. Consider the case $s = 2$, where $L = A_1x_1 + A_2x_2$. We can find $k_1, k_2 \in \Bbbk^\times$ such that $A_1' = k_1A_1$ and $A_2' = k_2A_2$ are integral ideals and $A_1' + A_2' = \mathcal{O}$ [Shimura 2010, Lemma 10.15(i)]. Let g' be the matrix formed by substituting the first two columns of the $n \times n$ unit matrix with $k_1^{-1}x_1$ and $k_2^{-1}x_2$. Then $g'^{-1}L = A_1'e_1 + A_2'e_2$. Next let

$$g'' = \begin{bmatrix} 1 & 1 & \\ -a_2 & a_1 & \\ & & I_{n-2} \end{bmatrix},$$

where $a_1 \in A_1'$ and $a_2 \in A_2'$ are taken such that $a_1 + a_2 = 1$. It is easily verified that $g''(A_1'e_1 + A_2'e_2) = Oe_1 + A_1'A_2'e_2$. Hence $g = \mathrm{diag}(1, k_1^{-1}k_2^{-1}, 1, \dots, 1)g''g'^{-1}$ maps L to $\mathcal{O}e_1 + A_1A_2e_2$. The general case when $s > 2$ follows inductively from this result. □

The ideal class associated to the $\mathcal{O}$-lattice L mentioned above in (ii) is known as the *Steinitz class of L*, denoted by $\lambda(L)$. We may also speak of the Steinitz class of an entire lattice class in $\mathfrak{L}$ since every $\mathcal{O}$-lattice in a lattice class has the same Steinitz class.

It follows directly that mapping each lattice class to its Steinitz class gives a bijection between the set of lattice classes in $\mathfrak{L}$ and $\mathrm{Cl}(\Bbbk)$. As a result the class number of GL_n, which we have noted to be equivalent to the number of distinct lattice classes in $\mathfrak{L}$, is equal to the class number of $\Bbbk$, which we write as h.

We now proceed to prove that $h_i = |Q(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk)/\Gamma_i|$ is also equal to h for every $i = 1, \ldots, h$. As we did in the previous section, let $\{\eta_1, \ldots, \eta_h\}$ be a complete set of representatives for $\mathrm{GL}_n(\Bbbk)\backslash\mathrm{GL}_n(\mathbb{A})^1/G^1_{\mathbb{A},\infty}$. Then for each $i = 1, \ldots, h$ put $L_i = \eta_i(\mathcal{O}\boldsymbol{e}_1 + \cdots + \mathcal{O}\boldsymbol{e}_n) \in \mathfrak{L}$.

Next we identify $Q(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk)$ with the set of all m-dimensional linear subspaces of $\Bbbk^n$ denoted by Gr_m (the Grassmannian) via the bijection

$$Q(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk) \ni Q(\Bbbk)g \longmapsto g^{-1}\left(\sum_{k=1}^{m} \Bbbk\boldsymbol{e}_k\right) \in \mathrm{Gr}_m. \tag{4}$$

From here up to the end of Theorem 11 we fix $i \in \{1, \ldots, h\}$. Considering the left action of $\Gamma_i \subset \mathrm{GL}_n(\Bbbk)$ on Gr_m, the map (4) gives rise to the bijection

$$Q(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk)/\Gamma_i \ni Q(\Bbbk)g\Gamma_i \longmapsto \Gamma_i g^{-1}\left(\sum_{k=1}^{m} \Bbbk\boldsymbol{e}_k\right) \in \Gamma_i\backslash\mathrm{Gr}_m, \tag{5}$$

which lets us identify $Q(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk)/\Gamma_i$ with $\Gamma_i\backslash\mathrm{Gr}_m$.

Lemma 9. *Γ_i is the stabilizer of L_i in $\mathrm{GL}_n(\Bbbk)$, under the action of $\mathrm{GL}_n(\mathbb{A})^1$ on $\mathfrak{L}$, i.e.,*

$$\Gamma_i = \{g \in \mathrm{GL}_n(\Bbbk) : gL_i = L_i\}.$$

Proof. Since $\Gamma_i = \big(\mathrm{GL}_n(\Bbbk_\infty) \times \eta_i \prod_{\sigma\in\boldsymbol{p}_f} \mathrm{GL}_n(\mathcal{O}_\sigma)\eta_i^{-1}\big) \cap \mathrm{GL}_n(\Bbbk)$, this is obvious from our choice of L_i. □

Proposition 10. *Let $V_1, V_2 \in \mathrm{Gr}_m$ and put $\tilde{L}_1 = L_i \cap V_1$, $\tilde{L}_2 = L_i \cap V_2$, which are $\mathcal{O}$-lattices in V_1 and V_2 respectively. Then $\lambda(\tilde{L}_1) = \lambda(\tilde{L}_2)$ if and only if there exists $g \in \Gamma_i$ such that $V_1 = gV_2$.*

Proof. Suppose that $V_1 = gV_2$ for some $g \in \Gamma_i$. From Lemma 8 we can find a $\Bbbk$-basis $\{y_j\}_{j=1}^n$ for $\Bbbk^n$ contained in L_i with $y_1, \ldots, y_m \in V_2$. Put $x_j = gy_j$ for $j = 1, \ldots, m$. Then $\{x_j\}_{j=1}^m$ and $\{y_j\}_{j=1}^m$ span V_1 and V_2 respectively and since g stabilizes L_i, they are also contained in $\tilde{L}_1$ and $\tilde{L}_2$ respectively.

For $v \in V_1$ and $w \in V_2$, we write $(\alpha_v)_j$ and $(\beta_w)_j$ for the $\Bbbk$-coefficients of x_j and y_j in v and w respectively $\big($so $v = \sum_{j=1}^m (\alpha_v)_j x_j$ and $w = \sum_{j=1}^m (\beta_w)_j y_j\big)$. Let J_1 be the fractional ideal generated by $\{\det[(\alpha_{v_j})_l]_{j,l=1}^m \mid v_1, \ldots, v_m \in \tilde{L}_1\}$. We can show that the ideal class of J_1 in $\mathrm{Cl}(\Bbbk)$ is $\lambda(\tilde{L}_1)$ as follows: From the lemma above we have $\tilde{L}_1 = A_1x'_1 + \cdots + A_m x'_m$, with fractional ideals $A_1, \ldots, A_m$ and $\{x'_j\}_{j=i}^m$ a basis of V_1. Comparing $\bigwedge_{j=1}^m \tilde{L}_1 = A_1 \cdots A_m(x'_1 \wedge \cdots \wedge x'_m)$ with

$$\textstyle\bigwedge_{j=1}^m \tilde{L}_1 = \Bbbk\text{-span of } \{v_1 \wedge \cdots \wedge v_m \mid v_1, \ldots, v_m \in \tilde{L}_1\} = J_1(x_1 \wedge \cdots \wedge x_m),$$

we see that $A_1 \cdots A_m$ is a $\Bbbk^\times$-multiple of J_1; hence their ideal classes are equivalent.

Similarly $\lambda(\tilde{L}_2)$ is the ideal class of the fractional ideal J_2 generated by the $\det[(\beta_{w_j})_l]_{j,l=1}^m$ for all $w_1, \ldots, w_m \in \tilde{L}_2$. However, since any arbitrary $v \in \tilde{L}_1$ can be written as gw with some $w \in \tilde{L}_2$ and

$$\begin{aligned} v = gw \iff & \sum_{j=1}^m (\alpha_v)_j x_j = g\Bigl(\sum_{j=1}^m (\beta_w)_j y_j\Bigr) = \sum_{j=1}^m (\beta_w)_j g y_j = \sum_{j=1}^m (\beta_w)_j x_j \\ \iff & (\alpha_v)_j = (\beta_w)_j, \quad j = 1, \ldots, m, \end{aligned}$$

this shows that $J_1 = J_2$ and thus $\lambda(\tilde{L}_1) = \lambda(\tilde{L}_2)$.

Now suppose conversely that $\lambda(\tilde{L}_1) = \lambda(\tilde{L}_2)$. Using Lemma 8, we obtain $\Bbbk$-bases $\{x_j\}_{j=1}^n$, $\{y_j\}_{j=1}^n$ for $\Bbbk^n$ and fractional ideals $A_1, \ldots, A_n, B_1, \ldots, B_n$ such that $L_i = A_1x_1 + \cdots + A_nx_n = B_1y_1 + \cdots + B_ny_n$ and $x_1, \ldots, x_m \in V_1$, $y_1, \ldots, y_m \in V_2$. Since $\tilde{L}_1 = A_1x_1 + \cdots + A_mx_m$ and $\tilde{L}_2 = B_1y_1 + \cdots + B_my_m$, the ideal classes of $A_1 \cdots A_m$ and $B_1 \cdots B_m$ are equivalent, and hence so are those of $A_{m+1} \cdots A_n$ and $B_{m+1} \cdots B_n$. By substituting the basis vectors and fractional ideals with suitable $\Bbbk^\times$-multiples, we may assume that $A_1 \cdots A_m = B_1 \cdots B_m$ and $A_{m+1} \cdots A_n = B_{m+1} \cdots B_n$.

Finally using Lemma 8(iii) we can find $g_1, g_2 \in \mathrm{GL}_n(\Bbbk)$ satisfying

$$g_1L_i = \sum_{j=1}^{m-1} \mathcal{O}\boldsymbol{e}_j + (A_1 \cdots A_m)\boldsymbol{e}_m + \sum_{j=m+1}^{n-1} \mathcal{O}\boldsymbol{e}_j + (A_{m+1} \cdots A_n)\boldsymbol{e}_n,$$

$$g_2L_i = \sum_{j=1}^{m-1} \mathcal{O}\boldsymbol{e}_j + (B_1 \cdots B_m)\boldsymbol{e}_m + \sum_{j=m+1}^{n-1} \mathcal{O}\boldsymbol{e}_j + (B_{m+1} \cdots B_n)\boldsymbol{e}_n,$$

chosen such that

$$g_1\tilde{L}_1 = \sum_{j=1}^{m-1} \mathcal{O}\boldsymbol{e}_j + (A_1 \cdots A_m)\boldsymbol{e}_m, \quad g_2\tilde{L}_2 = \sum_{j=1}^{m-1} \mathcal{O}\boldsymbol{e}_j + (B_1 \cdots B_m)\boldsymbol{e}_m.$$

Put $g = g_1^{-1}g_2$. Since $g_1L_i = g_2L_i$, the previous lemma gives us $g \in \Gamma_i$, while $gV_2 = V_1$ follows from $gy_j \in g\tilde{L}_2 = \tilde{L}_1 \subset V_1$ $(j = 1, \ldots, m)$. ☐

Finally we consider the map

$$\lambda_i : \Gamma_i \backslash \mathrm{Gr}_m \to \mathrm{Cl}(\Bbbk), \qquad \lambda_i(\Gamma_i V) = \lambda(L_i \cap V) \quad (V \in \mathrm{Gr}_m), \tag{6}$$

which is well-defined and injective as a result of the previous proposition.

Theorem 11. $h_i = h.$

Proof. Since $h_i = |Q(\Bbbk)\backslash \mathrm{GL}_n(\Bbbk)/\Gamma_i| = |\Gamma_i \backslash \mathrm{Gr}_m|$ we only need to prove that λ_i is surjective.

Take any ideal class in $\mathrm{Cl}(\Bbbk)$ and let A be a fractional ideal representing this class. Also let B be a fractional ideal representing $\lambda(L_i)$. Lemma 8(iii) allows us to find $g \in \mathrm{GL}_n(\Bbbk)$ such that

$$gL_i = \sum_{1 \le k < n-1} \mathcal{O}\boldsymbol{e}_k + A\boldsymbol{e}_{n-1} + A^{-1}B\boldsymbol{e}_n.$$

Let V be the subspace of $\Bbbk^n$ spanned by $\boldsymbol{e}, \ldots, \boldsymbol{e}_{m-1}, \boldsymbol{e}_{n-1}$ and put $V' = g^{-1}V \in \mathrm{Gr}_m$. Then $L_i \cap V' \simeq \left(\bigoplus_{j=1}^{m-1} \mathcal{O}\right) \oplus A$ so $\lambda_i(\Gamma_i V') = \lambda(L_i \cap V')$ is the class of A in $\mathrm{Cl}(\Bbbk)$, as required. □

The one-to-one correspondence between $\mathrm{GL}_n(\Bbbk)\backslash\mathrm{GL}_n(\mathbb{A})^1/G^1_{\mathbb{A},\infty}$ and the set of $\mathcal{O}$-lattices classes in $\mathfrak{L}$ mentioned earlier in the section is given by mapping each η_i to the lattice class of L_i. That this is a bijection follows from $G^1_{\mathbb{A},\infty}$ being the stabilizer group of the $\mathcal{O}$-lattice $\mathcal{O}\boldsymbol{e}_1 + \cdots + \mathcal{O}\boldsymbol{e}_n$ under the action of $\mathrm{GL}_n(\mathbb{A})^1$ on $\mathfrak{L}$. Continuing this map to the Steinitz class of the lattice gives us the bijection

$$\mathrm{GL}_n(\Bbbk)\backslash\mathrm{GL}_n(\mathbb{A})^1/G^1_{\mathbb{A},\infty} \ni \eta_i \mapsto \lambda(L_i) \in \mathrm{Cl}(\Bbbk).$$

This gives us an explicit way to find candidates for $\{\eta_1, \ldots, \eta_h\}$ as follows. Let $\{\mathfrak{a}_1, \ldots, \mathfrak{a}_h\}$ be a complete set of fractional ideals representing the ideal class of $\Bbbk$. For each $i = 1, \ldots, h$, we shall require an element $\eta_i \in \mathrm{GL}_n(\mathbb{A})^1$ such that the Steinitz class of the resulting lattice $L_i = \eta_i\left(\sum_{k=1}^n \mathcal{O}\boldsymbol{e}_k\right)$ is the ideal class represented by $\mathfrak{a}_i$.

Let $D_n(x)$ ($x \in \mathbb{A}$) denote the unit matrix of size n with bottom-most diagonal entry replaced by x. For each $1 \le i \le h$ we can choose $\alpha_i \in \mathbb{A}^\times$ such that $\alpha_{i\sigma}$ generates the principal ideal $\mathfrak{a}_i\mathcal{O}_\sigma$ for every finite σ and $|\alpha_i|_\infty = N(\mathfrak{a}_i)$, the ideal norm of $\mathfrak{a}_i$. Then $D_n(\alpha_i) \in \mathrm{GL}_n(\mathbb{A})^1$ since $|\det D_n(\alpha_i)|_{\mathbb{A}} = |\alpha_i|_{\mathbb{A}} = 1$, and

$$D_n(\alpha_i)\left(\sum_{k=1}^n \mathcal{O}\boldsymbol{e}_k\right) = \sum_{1 \le k < n} \mathcal{O}\boldsymbol{e}_k + \mathfrak{a}_i\boldsymbol{e}_n.$$

Hence putting $\eta_i = D_n(\alpha_i)$ ($1 \le i \le h$) gives us our required set of representatives. The corresponding $\mathcal{O}$-lattice L_i and its stabilizer group Γ_i will be denoted by $L_n(\mathfrak{a}_i)$ and $\Gamma_n(\mathfrak{a}_i)$ respectively whenever we want to call to attention the fractional ideal $\mathfrak{a}_i$ or the dimension n.

We can also proceed similarly to find, for a fixed i, a suitable set of representatives for $Q(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk)/\Gamma_i$. We do this using the bijection

$$Q(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk)/\Gamma_i \ni Q(\Bbbk)g\Gamma_i \longmapsto \lambda(L_i \cap g^{-1}V_m) \in \mathrm{Cl}(\Bbbk)$$

formed by composing λ_i with the bijection (5), where $V_m = \sum_{k=1}^m \Bbbk\boldsymbol{e}_k$.

For each $j \in \{1, \ldots, h\}$ the ideal $\mathfrak{a}_i\mathfrak{a}_j^{-1}$ shares the same ideal class as a unique $\mathfrak{a}_{j'}$ ($j' \in \{1, \ldots, h\}$); that is $[\mathfrak{a}_j][\mathfrak{a}_{j'}] = [\mathfrak{a}_i]$. Putting $\tau_i(j) := j'$ defines a permutation τ_i on $\{1, \ldots, h\}$.

Call a set of matrices $\{\xi_1, \dots, \xi_h\} \subset \mathrm{GL}_n(\Bbbk)$ an (n, m)*-splitting set for* $L_n(\mathfrak{a}_i)$ if for each $j = 1, \dots, h$

$$\begin{aligned}(7)\qquad \xi_j L_n(\mathfrak{a}_i) &= \Bigl(\sum_{1\le k<m} \mathcal{O}\boldsymbol{e}_k + \mathfrak{a}_j \boldsymbol{e}_m\Bigr) + \Bigl(\sum_{m<k<n} \mathcal{O}\boldsymbol{e}_k + \mathfrak{a}_{\tau_i(j)}\boldsymbol{e}_n\Bigr)\\ &\simeq L_m(\mathfrak{a}_j) \oplus L_{n-m}(\mathfrak{a}_{\tau_i(j)}).\end{aligned}$$

Since $\lambda(L_i \cap \xi_j^{-1} V_m) = \lambda(\xi_j L_i \cap V_m) = [\mathfrak{a}_j]$ $(i \le j \le h)$, such a set of matrices completely represents $Q(\Bbbk)\backslash \mathrm{GL}_n(\Bbbk)/\Gamma_i$.

One such set is given as follows. For each $j = 1, \dots, h$, first take $\kappa_{ij} \in \Bbbk$ such that $\mathfrak{a}_j \mathfrak{a}_{\tau_i(j)} = \kappa_{ij}\mathfrak{a}_i$. Then choose elements $\alpha_{ij} \in \mathfrak{a}_j$, $\alpha'_{ij} \in \mathfrak{a}_{\tau_i(j)}$, $\beta_{ij} \in \mathfrak{a}_j^{-1}$ and $\beta'_{ij} \in \mathfrak{a}_{\tau_i(j)}^{-1}$ satisfying

$$\alpha_{ij}\beta_{ij} - \alpha'_{ij}\beta'_{ij} = 1$$

(see [Cohen 2000, §1, Proposition 1.3.12 or Algorithm 1.3.16]) and define the matrix

$$\xi_{ij} := \begin{bmatrix} I_{m-1} & & & \\ & \alpha_{ij} & & \kappa_{ij}\beta'_{ij} \\ & & I_{n-m+1} & \\ & \alpha'_{ij} & & \kappa_{ij}\beta_{ij} \end{bmatrix} \in \mathrm{GL}_n(\Bbbk).$$

By direct calculation it is easily verified that $\{\xi_{ij}\}_{j=1}^h$ is indeed an (n, m)-splitting set for $L_n(\mathfrak{a}_i)$ and thus fully represents $Q^{n,m}(\Bbbk)\backslash \mathrm{GL}_n(\Bbbk)/\Gamma_n(\mathfrak{a}_i)$.

5. Fundamental domains of $\mathrm{GL}_n(\Bbbk)\backslash \mathrm{GL}_n(\mathbb{A})^1$ and P_n/Γ_i

We use the results of Section 3 to determine suitable fundamental domains in our continued discussion of the general linear group.

5.1. *Local height functions.*

Definition. For each $\sigma \in \boldsymbol{p}$ define $H_\sigma : \bigwedge^m \Bbbk_\sigma^n \to \mathbb{R}_{>0}$ by

$$H_\sigma\Bigl(\sum_I a_I(\boldsymbol{e}_{i_1} \wedge \cdots \wedge \boldsymbol{e}_{i_m})\Bigr) = \begin{cases} \bigl(\sum_I |a_I|_\sigma^2\bigr)^{[\Bbbk_\sigma:\mathbb{R}]/2}, & \sigma \in \boldsymbol{p}_\infty,\\ \sup_I |a_I|_\sigma, & \sigma \in \boldsymbol{p}_f,\end{cases}$$

where the sum and the supremum are taken over all $I = \{i_1 < \cdots < i_m\} \subset \{1, \dots, n\}$. We call this the *local height function* at σ.

In the following we extend each H_σ to a function of $\mathrm{GL}_n(\Bbbk_\sigma)$ by putting

$$H_\sigma(\gamma) = H_\sigma(\gamma\boldsymbol{e}_1 \wedge \cdots \wedge \gamma\boldsymbol{e}_m), \quad \gamma \in \mathrm{GL}_n(\Bbbk_\sigma).$$

The following lemma allows us to express the height function H_Q (restricted to $G(\mathbb{A})^1$) in terms of these local heights.

Lemma 12. *For* $g=(g_\sigma)_{\sigma\in\boldsymbol{p}}\in \mathrm{GL}_n(\mathbb{A})^1$,

$$H_Q(g)=\prod_{\sigma\in\boldsymbol{p}} H_\sigma(g_\sigma^{-1})^{n/l}.$$

Proof. By noting that every local height H_σ as a function of $\mathrm{GL}_n(\Bbbk_\sigma)$ is left K_σ-invariant and writing

$$g=\begin{bmatrix} a & * \\ 0 & d\end{bmatrix} h \quad (a\in \mathrm{GL}_m(\mathbb{A}),\ d\in \mathrm{GL}_{n-m}(\mathbb{A}),\ h\in K),$$

we see that $H_\sigma(g_\sigma^{-1})=|\det(a_\sigma^{-1})|_\sigma^{r_\sigma}$ at every σ, where $r_\sigma=2$ when σ is an imaginary infinite place and 1 otherwise. Hence the right-hand side of our equation becomes $|\det a|_{\mathbb{A}}^{-n/l}$, while $H_Q(g)=|\det a|_{\mathbb{A}}^{-(n-m)/l}|\det d|_{\mathbb{A}}^{m/l}$ by definition. Then since $g\in \mathrm{GL}_n(\mathbb{A})^1$, we have $1=|\det g|_{\mathbb{A}}=|\det a|_{\mathbb{A}}|\det d|_{\mathbb{A}}$, which gives us our equality. □

We proceed to describe the sets $R_{i,j,\infty}$ using the matrices η_i and ξ_{ij} chosen at the end of the previous section. For the rest of this paper, for a square matrix A with entries in $\mathbb{A}$ or $\Bbbk_\infty$, we will write $|A|_{\mathbb{A}}$ and $|A|_\infty$ to denote $|\det A|_{\mathbb{A}}$ and $|\det A|_\infty$ respectively. When the size of A is at least m, we write $A^{[m]}$ for the top-left $m\times m$ submatrix of A, and use $|A|_\infty^{[m]}$ to denote $|A^{[m]}|_\infty$.

Lemma 13. *Let X_{ij} be the $n\times m$ matrix formed by the first m columns of ξ_{ij}^{-1}. Then*

$$H_Q(\xi_{ij}\gamma g\eta_i)=N(\mathfrak{a}_j)^{n/l}\left|{}^t\bar{X}_{ij}\,{}^t\bar{\gamma}^{-1}\,{}^t\bar{g}^{-1}(\eta_i)_\infty^{-2}g^{-1}\gamma^{-1}X_{ij}\right|_\infty^{n/2l} \tag{8}$$

for any $1\le i,\ j\le h,\ \gamma\in\Gamma_i$ *and* $g\in \mathrm{GL}_n(\Bbbk_\infty)^1$.

Proof. Let $x=\eta_i^{-1}g^{-1}\gamma^{-1}X_{ij}$ so that $H_\sigma((\xi_{ij}\gamma g\eta_i)_\sigma^{-1})=H_\sigma(x_\sigma\boldsymbol{e}_1\wedge\cdots\wedge x_\sigma\boldsymbol{e}_m)$. For $\sigma\in\boldsymbol{p}_\infty$, this computes to

$$\Bigg(\sum_{\substack{I\subset\{1,\dots,n\}\\ |I|=m}}|\det[x_\sigma]_I|_\sigma^2\Bigg)^{\frac12[\Bbbk_\sigma:\mathbb{R}]}=\Bigg(\sum_I \det{}^t\overline{[x_\sigma]}_I\det[x_\sigma]_I\Bigg)^{\frac12[\Bbbk_\sigma:\mathbb{R}]}=\det({}^t\bar{x}_\sigma x_\sigma)^{\frac12[\Bbbk_\sigma:\mathbb{R}]},$$

where for each $I=\{i_1<\cdots<i_m\}$ that the sums run through $[x_\sigma]_I$ denotes the $m\times n$ matrix formed by the i_1-th, …, i_m-th rows of x_σ arranged from top to bottom in that order. The final equality is due to the Cauchy–Binet formula; see [Bombieri and Gubler 2006, Proposition 2.8.8].

For $\sigma\in\boldsymbol{p}_f$, since g_σ is trivial and $\gamma_\sigma\in\eta_{i\sigma}\mathrm{GL}_n(\mathcal{O}_\sigma)\eta_{i\sigma}^{-1}$, we have $(\xi_{ij}\gamma g\eta_i)_\sigma=\xi_{ij\sigma}\eta_{i\sigma}h_\sigma$ for some $h_\sigma\in\mathrm{GL}_n(\mathcal{O}_\sigma)$. Hence $H_\sigma((\xi_{ij}\gamma g\eta_i)_\sigma^{-1})$ simplifies to

$$H_\sigma(\eta_{i\sigma}^{-1}\xi_{ij\sigma}^{-1})=H_\sigma\big(\beta_{ij}(\boldsymbol{e}_1\wedge\cdots\wedge\boldsymbol{e}_m)+\alpha_{i\sigma}^{-1}\kappa_{ij}\alpha'_{ij}(\boldsymbol{e}_1\wedge\cdots\wedge\boldsymbol{e}_{m-1}\wedge\boldsymbol{e}_n)\big)$$

or

$$\max\{|\beta_{ij}|_\sigma,|\alpha_{i\sigma}^{-1}\kappa_{ij}^{-1}\alpha'_{ij}|_\sigma\}=|\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma^{-1}\max\{|\beta_{ij}\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma,|\alpha'_{ij}\beta'_{ij}|_\sigma\}.$$

By the previous lemma, $H_Q(\xi_{ij}\gamma g\eta_i)$ is obtained by taking the n/l-th power of the product of all the $H_\sigma((\xi_{ij}\gamma g\eta_i)_\sigma^{-1})$. Thus it remains to verify that

$$\prod_{\sigma\in\boldsymbol{p}_f} |\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma^{-1} \max\{|\beta_{ij}\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma, |\alpha'_{ij}\beta'_{ij}|_\sigma\} = N(\mathfrak{a}_j).$$

First we see that $\prod_{\sigma\in\boldsymbol{p}_f} |\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma^{-1} = N(\beta'_{ij}\kappa_{ij}\mathfrak{a}_i) = N(\beta'_{ij}\mathfrak{a}_j\mathfrak{a}_{\tau_i(j)})$. It is then sufficient to show that the product of the remaining factors is $N(\beta'_{ij}\mathfrak{a}_{\tau_i(j)})^{-1}$.

Let $\mathfrak{p}_\sigma$ denote the prime ideal associated to a finite place $\sigma \in \boldsymbol{p}_f$. Write the prime ideal decompositions of $\beta_{ij}\mathfrak{a}_j$ and $\beta'_{ij}\mathfrak{a}_{\tau_i(j)}$ as $\prod_{\sigma\in\boldsymbol{p}_f}(\mathfrak{p}_\sigma\cap\mathcal{O})^{d_\sigma}$ and $\prod_{\sigma\in\boldsymbol{p}_f}(\mathfrak{p}_\sigma\cap\mathcal{O})^{e_\sigma}$ respectively, the exponents d_σ and e_σ being nonnegative.

Then $\beta_{ij}\beta'_{ij}\kappa_{ij}\mathfrak{a}_i = (\beta_{ij}\mathfrak{a}_j)(\beta'_{ij}\mathfrak{a}_{\tau_i(j)}) = \prod_{\sigma\in\boldsymbol{p}_f}(\mathfrak{p}_\sigma\cap\mathcal{O})^{d_\sigma+e_\sigma}$ and since each $\mathfrak{a}_{i\sigma}$ is generated by $\alpha_{i\sigma}$, this yields

$$|\beta_{ij}\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma = |\mathcal{O}_\sigma/\mathfrak{p}_\sigma|^{-d_\sigma-e_\sigma}, \quad \sigma\in\boldsymbol{p}_f.$$

Now $\alpha'_{ij}\beta'_{ij} \in \beta'_{ij}\mathfrak{a}_{\tau_i(j)}$ and hence $|\alpha'_{ij}\beta'_{ij}|_\sigma \le |\mathcal{O}_\sigma/\mathfrak{p}_\sigma|^{-e_\sigma}$. We have two cases.

Case 1: $d_\sigma = 0$. Then $|\alpha'_{ij}\beta'_{ij}|_\sigma \le |\beta_{ij}\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma = |O_\sigma/\mathfrak{p}_\sigma|^{-e_\sigma}$.

Case 2: $d_\mathfrak{p} > 0$. In this case

$$\alpha'_{ij}\beta'_{ij} = -1+\alpha_{ij}\beta_{ij} \in -1+\beta_{ij}\mathfrak{a}_j \subset -1+(\mathfrak{p}_\sigma\cap\mathcal{O})^{d_\mathfrak{p}}$$

shows us that $\alpha'_{ij}\beta'_{ij}\in\mathcal{O}_\sigma^\times$ and so $|\alpha'_{ij}\beta'_{ij}|_\sigma = 1 \ge |\beta_{ij}\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma$. We also note that since β_{ij} and β'_{ij} were chosen in such a way that $\beta_{ij}\mathfrak{a}_{ij}+\beta'_{ij}\mathfrak{a}_{ij} = \mathcal{O}$, the ideal $\beta_{ij}\mathfrak{a}_j$ is prime to $\beta'_{ij}\mathfrak{a}_{\tau_i(j)}$, which means $e_\sigma = 0$.

So in either case,

$$\max\{|\beta_{ij}\beta'_{ij}\kappa_{ij}\alpha_{i\sigma}|_\sigma, |\alpha'_{ij}\beta'_{ij}|_\sigma\} = |\mathcal{O}_\sigma/\mathfrak{p}_\sigma|^{-e_\sigma}$$

and thus the product over all finite places is $N(\beta'_{ij}\mathfrak{a}_{\tau_i(j)})^{-1}$, as required. □

Now fix $1\le i, j\le h$ and first consider the set $\xi_{ij}^{-1}R_{i,j,\infty}\xi_{ij}$. It is easy to directly verify that

$$\xi_{ij}^{-1}R_{i,j,\infty}\xi_{ij} = \{g\in G(\mathbb{k}_\infty)^1 : H_Q(\xi_{ij}g\eta_i) = \mathsf{m}_Q(g\eta_i)\}.$$

Hence for $g\in\xi_{ij}^{-1}R_{i,j,\infty}\xi_{ij}$ we have

$$H_Q(\xi_{ij}g\eta_i) = \mathsf{m}_Q(g\eta_i) = \min_{x\in Q(\mathbb{k})\backslash \mathrm{GL}_n(\mathbb{k})} H_Q(xg\eta_i) = \min_{\substack{1\le k\le h\\ \gamma\in\Gamma_i}} H_Q(\xi_{ik}\gamma g\eta_i),$$

which in this case can be written using (8) as

$$|{}^t\bar{X}_{ij}\,{}^t\bar{g}^{-1}(\eta_i)_\infty^{-2}g^{-1}X_{ij}|_\infty \le \left(\frac{N(\mathfrak{a}_k)}{N(\mathfrak{a}_j)}\right)^2 |{}^t\bar{X}_{ik}\,{}^t\bar{\gamma}\,{}^t\bar{g}^{-1}(\eta_i)_\infty^{-2}g^{-1}\gamma X_{ik}|_\infty$$

for all $k = 1,\dots,h$ and $\gamma\in\Gamma_i$.

Now ${}^t\overline{X}_{ik}\,{}^t\bar{\gamma}\,{}^t\bar{g}^{-1}(\eta_i)_\infty^{-2}g^{-1}\gamma X_{ik} = ({}^t\bar{\xi}_{ik}^{-1}\,{}^t\bar{\gamma}\,{}^t\bar{g}^{-1}(\eta_i)_\infty^{-2}g^{-1}\gamma\xi_{ik}^{-1})^{[m]}$, which by letting $g_{[ij]} = \xi_{ij}g\xi_{ij}^{-1}$ can be rewritten as

$$\big({}^t\overline{(\xi_{ij}\gamma\xi_{ik}^{-1})}\,{}^t\bar{g}_{[ij]}^{-1}({}^t\bar{\xi}_{ij}^{-1}(\eta_i)_\infty^{-2}\xi_{ij}^{-1})g_{[ij]}^{-1}(\xi_{ij}\gamma\xi_{ik}^{-1})\big)^{[m]}.$$

This lets us express the set $R_{i,j,\infty}$ as follows. For $g \in \mathrm{GL}_n(\Bbbk_\infty)$ let $\pi_{ij}(g)$ denote ${}^t\bar{g}^{-1}({}^t\bar{\xi}_{ij}^{-1}(\eta_i)_\infty^{-2}\xi_{ij}^{-1})g^{-1}$. Then $g \in R_{i,j,\infty}$ if and only if

$$|\pi_{ij}(g)|_\infty^{[m]} \le \left(\frac{N(\mathfrak{a}_k)}{N(\mathfrak{a}_j)}\right)^2 \left|{}^t\overline{(\xi_{ij}\gamma\xi_{ik}^{-1})}\pi_{ij}(g)(\xi_{ij}\gamma\xi_{ik}^{-1})\right|_\infty^{[m]} \tag{9}$$

for all $k = 1, \ldots, h$ and $\gamma \in \Gamma_i$.

5.2. ***Fundamental domains of P_n/Γ_i.*** For each infinite place σ of $\Bbbk$ let $P_n(\Bbbk_\sigma)$ denote the subset of $\mathrm{GL}_n(\Bbbk_\sigma)$ consisting of all positive definite real symmetric matrices when σ is real and positive definite Hermitian matrices when σ is imaginary. We consider the subset of $\mathrm{GL}_n(\Bbbk_\infty)$ defined by $P_n = \prod_{\sigma\in\boldsymbol{p}_\infty} P_n(\Bbbk_\sigma)$. This is the space of positive definite Humbert forms in $\mathrm{GL}_n(\Bbbk)$.

We have the following right action of $\mathrm{GL}_n(\Bbbk_\infty)$ on P_n:

$$A \cdot g = {}^t\bar{g}Ag \quad (g \in \mathrm{GL}_n(\Bbbk_\infty),\ A \in P_n). \tag{10}$$

To determine fundamental domains in P_n with respect to subgroups of $\mathrm{GL}_n(\Bbbk)$, we consider instead the induced action $A \cdot gZ = {}^t\bar{g}Ag$ of $\mathrm{GL}_n(\Bbbk)/Z$ on P_n, where $Z = \{z \in \Bbbk : \bar{z}z = 1\}$, the set of roots of unity in $\Bbbk$. Here $\{zI_n : z \in Z\}$ is naturally seen to be the intersection of K_∞ and the center of $\mathrm{GL}_n(\Bbbk)$.

Hence given a discrete subgroup Γ of $\mathrm{GL}_n(\Bbbk)$ acting on a subset T of P_n, a fundamental domain Ω of a T/Γ is an open subset of T satisfying

(i) $T = \Omega^- \cdot \Gamma$,

(ii) for $\gamma \in \Gamma$, if $\Omega^\circ \cap (\Omega^- \cdot \gamma) \neq \varnothing$ then $\gamma \in Z$.

Now for each $1 \le i, j \le h$, put

$$K_{i,j,\infty} = (\xi_{ij}\eta_i)_\infty K_\infty(\xi_{ij}\eta_i)_\infty^{-1}, \quad P_n^{ij} = \{A \in P_n : |A|_\infty = N(\kappa_{ij}\mathfrak{a}_i)^{-2}\},$$

and define the map $\pi_{ij} : G(\Bbbk_\infty) \ni g \mapsto {}^t\bar{g}^{-1}({}^t\bar{\xi}_{ij}^{-1}(\eta_i)_\infty^{-2}\xi_{ij}^{-1})g^{-1} \in P_n$. Note that $K_{i,j,\infty}$ is the stabilizer of ${}^t\bar{\xi}_{ij}^{-1}(\eta_i)_\infty^{-2}\xi_{ij}^{-1} \in P_n$ under the action of $\mathrm{GL}_n(\Bbbk_\infty)$ on P_n and that π_{ij} preserves this action. Thus the surjective map π_{ij} gives us the isomorphisms

$$\mathrm{GL}_n(\Bbbk_\infty)/K_{i,j,\infty} \simeq P_n \quad \text{and} \quad \mathrm{GL}_n(\Bbbk_\infty)^1/K_{i,\infty} \simeq \pi_{ij}(\mathrm{GL}_n(\Bbbk_\infty)^1) = P_n^{ij}$$

since $|{}^t\bar{\xi}_{ij}^{-1}(\eta_i)_\infty^{-2}\xi_{ij}^{-1}|_\infty = N(\kappa_{ij}\mathfrak{a}_i)^{-2}$.

Lastly let $F_{i,j}^{n,m}$ denote the following closed subset of P_n:

$$\left\{A \in P_n : |A|_\infty^{[m]} \leq \left(\frac{N(\mathfrak{a}_k)}{N(\mathfrak{a}_j)}\right)^2 \left|{}^t\overline{(\xi_{ij}\gamma\xi_{ik}^{-1})}A(\xi_{ij}\gamma\xi_{ik}^{-1})\right|_\infty^{[m]},\ 1 \leq k \leq h,\ \gamma \in \Gamma_i\right\}.$$

From (9), π_{ij} maps $R_{i,j,\infty}$ onto $F_{i,j}^{n,m} \cap P_n^{ij}$. We also note that following statement holds true, the proof of which will be given later in the section.

Proposition 14. *$F_{i,j}^{n,m}$ is right $Q_{i,j}$-invariant under the action* (10).

Thus the subgroup $Q_{i,j}$ of $\mathrm{GL}_n(\Bbbk_\infty)$ acts on $R_{i,j,\infty}$ from the left and on $F_{i,j}^{n,m}$ from the right, and π_{ij} preserves this. Hence by constructing a fundamental domain for $F_{i,j}^{n,m}/Q_{i,j}$, we can find one for $Q_{i,j}\backslash R_{i,j,\infty}$ by taking the inverse image under π_{ij}.

We start by observing that $\xi_{ij}\Gamma_i\xi_{ij}^{-1}$ is the stabilizer in $\mathrm{GL}_n(\Bbbk)$ of the $\mathcal{O}$-lattice $\xi_{ij}L_i$ described in (7). This gives us an expression for $Q_{i,j} = Q(\Bbbk) \cap \xi_{ij}\Gamma_i\xi_{ij}^{-1}$:

$$\left\{\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} : a \in \Gamma_m(\mathfrak{a}_j),\ d \in \Gamma_{n-m}(\mathfrak{a}_{\tau_i(j)}),\ bL_{n-m}(\mathfrak{a}_{\tau_i(j)}) \subset L_m(\mathfrak{a}_j)\right\}.$$

Any $A \in P_n$ can be written uniquely in the form

$$A = \begin{bmatrix} I_m & 0 \\ {}^t\overline{u_{A,m}} & I_{n-m} \end{bmatrix} \begin{bmatrix} A^{[m]} & 0 \\ 0 & A_{[n-m]} \end{bmatrix} \begin{bmatrix} I_m & u_{A,m} \\ 0 & I_{n-m} \end{bmatrix} \tag{11}$$

with $A^{[m]} \in P_m$, $A_{[n-m]} \in P_{n-m}$ and $u_{A.m} \in M_{m,n-m}(\Bbbk_\infty)$. (The symbol $A^{[m]}$ here coincides with its prior use to denote the top left $m \times m$ submatrix of A). It is easy to verify that the action of $q = \left[\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right] \in Q_{i,j}$ on A results in

$$({}^t\bar{q}Aq)^{[m]} = {}^t\bar{a}A^{[m]}a, \quad ({}^t\bar{q}Aq)_{[n-m]} = {}^t\bar{d}A_{[n-m]}d,$$
$$u_{{}^t\bar{q}Aq,m} = a^{-1}(u_{A,m}d + b).$$

These equations will determine the necessary form of our fundamental domain, as well as allow us to prove our previous proposition. Given $A \in F_{i,j}^{n,m}$ and q as above, we first see that

$$|{}^t\bar{q}Aq|_\infty^{[m]} = |{}^t\bar{a}|_\infty |A|_\infty^{[m]} |a|_\infty = |A|_\infty^{[m]}.$$

Next put $q = \xi_{ij}\gamma_q\xi_{ij}^{-1}$, $\gamma_q \in \Gamma_i$, to get

$${}^t\overline{(\xi_{ij}\gamma\xi_{ik}^{-1})}\,{}^t\bar{q}Aq(\xi_{ij}\gamma\xi_{ik}^{-1}) = {}^t\overline{(\xi_{ij}\gamma_q\gamma\xi_{ik}^{-1})}A(\xi_{ij}\gamma_q\gamma\xi_{ik}^{-1})$$

for all $\gamma \in \Gamma_i$ and every k. Together, this shows that ${}^t\bar{q}Aq \in F_{i,j}^{n,m}$ as proposed.

Now for each $k = 1, \ldots, h$ choose sets $\mathfrak{d}_k$, $\mathfrak{d}'_k$ and $\mathfrak{d}_{ik}$ that are fundamental domains for $\Bbbk_\infty$ with respect to addition by $\mathfrak{a}_k$, $\mathfrak{a}_k^{-1}$ and $\mathfrak{a}_k\mathfrak{a}_{\tau_i(k)}^{-1}$ respectively. We require each of these sets to be closed under multiplication by Z. Then choose also a subset $\tilde{\mathfrak{d}}_{ik}$ of $\mathfrak{d}_{ik}$ that is a fundamental domain for $\mathfrak{d}_{ik}$ with respect to multiplication by Z. Also if necessary (which will be the case when $m > 1$ and $n-m > 1$) take a fundamental domain $\mathfrak{d}_{\mathcal{O}}$ of $\Bbbk_\infty$ with respect to addition by $\mathcal{O}$.

Using these, we define for $1 < i,\ j < h$ the sets

$$\mathfrak{D}^{n,m}_{i,j} = \left\{ \begin{bmatrix} d_{11} & \cdots & d_{1,n-m} \\ \vdots & \ddots & \vdots \\ d_{m1} & \cdots & d_{m,n-m} \end{bmatrix} : d_{m,n-m} \in \tilde{\mathfrak{d}}_{ij},\ d_{rs} \in \begin{cases} \mathfrak{d}_{\mathcal{O}}, & r < m,\ s < n-m, \\ \mathfrak{d}'_{\tau_i(j)}, & r < m,\ s = n-m, \\ \mathfrak{d}_j, & r = m,\ s < n-m \end{cases} \right\}$$

and

$$F^{n,m}_{i,j}(S, S') = \{A \in F^{n,m}_{i,j} : A^{[m]} \in S,\ A_{[n-m]} \in S',\ u_{A,m} \in \mathfrak{D}^{n,m}_{i,j}\}$$

with arbitrary subsets $S \subset P_m$ and $S' \subset P_{n-m}$.

In particular we will want to consider $F^{n,m}_{i,j}(\mathfrak{B}_j, \mathfrak{C}_{\sigma_i(j)})$ when $\mathfrak{B}_j$ and $\mathfrak{C}_{\tau_i(j)}$ are fundamental domains for $P_m/\Gamma_m(\mathfrak{a}_j)$ and $P_{n-m}/\Gamma_{n-m}(\mathfrak{a}_{\tau_i(j)})$ respectively. In this case, based on our observations on the action of $Q_{i,j}$ on $F^{n,m}_{i,j}$, we establish the following result.

Lemma 15. *$F^{n,m}_{i,j}(\mathfrak{B}_j, \mathfrak{C}_{\tau_i(j)})$ is a fundamental domain of $F^{n,m}_{i,j}/Q_{i,j}$.*

Proof. We write $F = F^{n,m}_{i,j}(\mathfrak{B}_j, \mathfrak{C}_{\tau_i(j)})$ for short. First consider an $A \in F^{n,m}_{i,j}$. We can find $b \in \mathfrak{B}^-_j$, $c \in \mathfrak{C}^-_{\tau_i(j)}$ and $a \in \Gamma_m(\mathfrak{a}_j)$, $d \in \Gamma_{n-m}(\mathfrak{a}_{\tau_i(j)})$ such that $A^{[m]} = {}^t\bar{a}ba$ and $A_{[n-m]} = {}^t\bar{d}cd$. Also, by substituting a with a suitable Z-multiple if necessary, we can find $f \in (\mathfrak{D}^{n,m}_{i,j})^-$ and a $g \in M_{m,n-m}(\Bbbk)$ mapping $L_{n-m}(\mathfrak{a}_{\tau_i(j)})$ to $L_m(\mathfrak{a}_j)$ such that $au_{A,m}d^{-1} = f + g$. Let

$$q = \begin{bmatrix} a & gd \\ 0 & d \end{bmatrix}, \quad A' = \begin{bmatrix} I_m & 0 \\ {}^t\bar{f} & I_{n-m} \end{bmatrix} \begin{bmatrix} b & 0 \\ 0 & c \end{bmatrix} \begin{bmatrix} I_m & f \\ 0 & I_{n-m} \end{bmatrix}.$$

Then $q \in Q_{i,j}$ and $A = {}^t\bar{q}A'q$. We have from the $Q_{i,j}$-invariance of $F^{n,m}_{i,j}$ that $A' \in F^{n,m}_{i,j}$ and so $A' \in F^-$. This shows that $F^{n,m}_{i,j} = F^- \cdot Q_{i,j}$.

Next suppose $F^\circ \cap (F^- \cdot q)$ is nonempty for a $q = \left[\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right] \in Q_{i,j}$, so there exist $A \in F^\circ$ and $A' \in F^-$ such that $A = {}^t\bar{q}A'q$. We must show that $q \in Z$. From $A^{[m]} = {}^t\bar{a}A'^{[m]}a \in \mathfrak{B}_{ij}$ and $A_{[n-m]} = {}^t\bar{d}A'_{[n-m]}d \in \mathfrak{C}_{ij}$, we must have $a = a_1 I_m$ and $d = d_1 I_{n-m}$ with some $a_1, d_1 \in Z$. Since the entries of $u_{A,m}$ and $u_{A',m}$ are respectively in the interior and closure of either $\mathfrak{d}_{\mathcal{O}}$, $\mathfrak{d}_j$, $\mathfrak{d}'_{\tau_i(j)}$ or $\mathfrak{d}_{ij}$, which are all invariant under Z, we see that $b = au_{A,m} - u_{A',m}d$ must necessarily be 0. From this we get $a_1 u_{A,m} = d_1 u_{A',m}$, whose $(m, n-m)$-th entry belongs to $\tilde{\mathfrak{d}}_{ij}$, implying that $a_1 d_1^{-1} \in Z$. Hence $q \in Z$. □

As a result, the inverse image of $F^{n,m}_{i,j}(\mathfrak{B}_{ij}, \mathfrak{C}_{ij}) \cap P^{ij}_n$ under π_{ij} is a fundamental domain of $Q_{i,j} \backslash R_{i,j,\infty}$.

If we have fundamental domains $\mathfrak{B}_1, \ldots, \mathfrak{B}_h$ for P_m with respect to the groups $\Gamma_m(\mathfrak{a}_1), \ldots, \Gamma_m(\mathfrak{a}_h)$, as well as fundamental domains $\mathfrak{C}_1, \ldots, \mathfrak{C}_h$ of P_{n-m} with respect to $\Gamma_{n-m}(\mathfrak{a}_1), \ldots, \Gamma_{n-m}(\mathfrak{a}_h)$, we are able to construct the sets $F^{n,m}_{i,j}(\mathfrak{B}_j, \mathfrak{C}_{\sigma_i(j)})$

for each i and j. Then by Corollary 6 a fundamental domain for $\mathrm{GL}_n(\Bbbk)\backslash\mathrm{GL}_n(\mathbb{A})^1$ is given by the set

$$\bigsqcup_{1\le i,j\le h} \pi_{ij}^{-1}(F_{i,j}^{n,m}(\mathfrak{B}_j,\mathfrak{C}_{\tau_i(j)})\cap P_n^{ij})\xi_{ij}\eta_i K_f.$$

Also Theorem 7 shows us that $\bigcup_{j=1}^h \xi_{ij}^{-1}\pi_{ij}^{-1}(F_{i,j}^{n,m}(\mathfrak{B}_j,\mathfrak{C}_{\tau_i(j)})\cap P_n^{ij})\xi_{ij}$ is a fundamental domain for $\mathrm{GL}_n(\Bbbk_\infty)^1$ with respect to Γ_i. Now let

$$\Omega_i^{n,m}(\mathfrak{B}_1,\dots,\mathfrak{B}_h,\mathfrak{C}_1,\dots,\mathfrak{C}_h)=\bigcup_{j=1}^h {}^t\bar{\xi}_{ij}F_{i,j}^{n,m}(\mathfrak{B}_j,\mathfrak{C}_{\tau_i(j)})\xi_{ij}.$$

We have the following result.

Theorem 16. *$\Omega_i^{n,m}(\mathfrak{B}_1,\dots,\mathfrak{B}_h,\mathfrak{C}_1,\dots,\mathfrak{C}_h)\cap P_n^{ij}$ is a fundamental domain of P_n^{ij} with respect to Γ_i. In addition, by viewing $\mathbb{R}_{>0}$ as a subset of $\Bbbk_\infty$ via the usual diagonal embedding, if we assume for $k=1,\dots,h$ that*

$$\mathbb{R}_{>0}\mathfrak{B}_k=\mathfrak{B}_k,\quad \mathbb{R}_{>0}\mathfrak{C}_k=\mathfrak{C}_k,$$

then $\Omega_i^{n,m}(\mathfrak{B}_1,\dots,\mathfrak{B}_h,\mathfrak{C}_1,\dots,\mathfrak{C}_h)$ is a fundamental domain of P_n/Γ_i.

Proof. We write Ω for $\Omega_i^{n,m}(\mathfrak{B}_1,\dots,\mathfrak{B}_h,\mathfrak{C}_1,\dots,\mathfrak{C}_h)$ and Γ for Γ_i for short. If we define the map $G(\Bbbk_\infty)\ni g\mapsto {}^t\bar{g}^{-1}(\eta_i)_\infty^{-2}g^{-1}\in P_n$ we can directly verify that the image of $\bigcup_{j=1}^h \xi_{ij}^{-1}\pi_{ij}^{-1}(F_{i,j}^{n,m}(\mathfrak{B}_j,\mathfrak{C}_{\tau_i(j)})\cap P_n^{ij})\xi_{ij}$ under this map is Ω, which gives us the first result. For the second part, note that $\mathbb{R}_{>0}F_{i,j}^{n,m}=F_{i,j}^{n,m}$ and

$$(xA)^{[m]}=x(A^{[m]}),\quad (xA)_{[n-m]}=x(A_{[n-m]}),\quad u_{xA,m}=u_{A,m}$$

for any $x\in\mathbb{R}_{>0}$ and $A\in P_n$. Thus the conditions on the $\mathfrak{B}_k$ and $\mathfrak{C}_k$ imply that $\mathbb{R}_{>0}F_{i,j}^{n,m}(\mathfrak{B}_j,\mathfrak{C}_{\tau_i(j)})=F_{i,j}^{n,m}(\mathfrak{B}_j,\mathfrak{C}_{\tau_i(j)})$ for each j; hence $\mathbb{R}_{>0}\Omega=\Omega$. Since $P_n=\mathbb{R}_{>0}P_n^{ij}$, we see from $P_n^{ij}=(\Omega\cap P_n^{ij})^-\cdot\Gamma$ that $P_n=\Omega^-\cdot\Gamma$. Finally suppose that $\Omega^\circ\cap({}^t\bar{\gamma}\Omega^-\gamma)$ $(\gamma\in\Gamma)$ contains an element $g={}^t\bar{\gamma}g'\gamma$ $(g'\in\Omega^-)$. Put $x=(N(\kappa_{ij}\mathfrak{a}_i)^2|g|_\infty)^{-1/n[\Bbbk_\infty:\mathbb{R}]}$. Then $|xg|_\infty=|xg'|_\infty=N(\kappa_{ij}\mathfrak{a}_i)^{-2}$ and hence $xg={}^t\bar{\gamma}xg'\gamma\in(\Omega^\circ\cap P_n^{ij})\cap{}^t\bar{\gamma}(\Omega^-\cap P_n^{ij})\gamma$, which gives us $\gamma=I_n$, as required. □

Using the theorem, we can construct fundamental domains for P_n with respect to Γ_i for each i and $n\ge 1$. Since $\Gamma_i=\mathcal{O}^\times$ for any i when $n=1$, we can start by choosing a fixed fundamental domain, Ω^1, for P_1 with respect to $\mathcal{O}^\times/Z$ that is closed under multiplication by $\mathbb{R}_{>0}$. (The existence of such a set can be shown using Voronoi reduction; see the Appendix.) Then for each $i=1,\dots,h$, let $\Omega_i^1=\Omega^1$ and define

$$\Omega_i^n=\Omega_i^{n,n-1}(\Omega_1^{n-1}\dots,\Omega_h^{n-1},\Omega^1,\cdots,\Omega^1)$$

inductively for $n\ge 2$. By construction, $\mathbb{R}_{>0}\Omega_i^n=\Omega_i^n$ so for each $1\le i\le h$ and $n\ge 1$, Ω_i^n gives us a fundamental domain for P_n/Γ_i.

An example implementation of this construction for P_2 over the imaginary quadratic field $\mathbb{Q}(\sqrt{-5})$ of class number 2 is given in the following subsection. Similar work on fundamental domains in spaces over real quadratic fields of class number 1 can be found in [Cohn 1965].

5.3. ***An example*** **($\Bbbk = \mathbb{Q}(\sqrt{-5})$).** When $\Bbbk$ is an imaginary quadratic field, we have $\Bbbk_\infty = \mathbb{C}$. For $n = 1$ we have $P_1 = \mathbb{R}_{>0}(\subset \mathbb{C})$ and $\Gamma_i = \mathcal{O}^\times = Z$ acts trivially on P_1; hence P_1 itself is a fundamental domain for $P_1/\,\Gamma_1(\mathfrak{a}_i)$.

Consider in particular $\Bbbk = \mathbb{Q}(\sqrt{-5})$ of class number $h = 2$. We can choose representatives $\mathfrak{a}_1, \mathfrak{a}_2$ for $\mathrm{Cl}(\Bbbk)$ by putting $\mathfrak{a}_1 = \mathcal{O}$ and $\mathfrak{a}_2 = \langle 2, 1+\sqrt{-5}\rangle$. Then following the procedure at the end of Section 4, we see that

$$\mathfrak{a}_1^2 = \mathfrak{a}_1, \qquad \mathfrak{a}_2^2 = 2\mathfrak{a}_1 \qquad \left(\tau_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix},\ \kappa_{11} = 1,\ \kappa_{12} = 2\right),$$

$$\mathfrak{a}_1\mathfrak{a}_2 = \mathfrak{a}_2, \quad \mathfrak{a}_2\mathfrak{a}_1 = \mathfrak{a}_2 \qquad \left(\tau_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},\ \kappa_{21} = \kappa_{22} = 1\right),$$

and (2, 1)-splitting sets for $L_2(\mathfrak{a}_i)$ are given by

$$\left\{\xi_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ \xi_{12} = \begin{bmatrix} 2 & 2+\sqrt{-5} \\ 2 & 3+\sqrt{-5} \end{bmatrix}\right\} \quad (i = 1),$$

$$\left\{\xi_{21} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ \xi_{22} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\right\} \qquad (i = 2).$$

For $1 \le i, j, k \le 2$ denote by $\Xi_{i,j,k}$ the set of the first columns of the matrices $\xi_{ij}\gamma\xi_{ik}^{-1}$ as γ ranges over $\Gamma(\mathfrak{a}_i)$. Then for $A \in P_2$

$$\begin{aligned} \min_{\gamma\in\Gamma_i} \left|{}^t\!\left(\overline{\xi_{ij}\gamma\xi_{ik}^{-1}}\right) A \left(\xi_{ij}\gamma\xi_{ik}^{-1}\right)\right|_\infty^{[1]} &= \min_{\boldsymbol{x}\in\Xi_{i,j,k}} |{}^t\bar{\boldsymbol{x}}A\boldsymbol{x}| \\ &= \min_{\left[\begin{smallmatrix} e \\ f \end{smallmatrix}\right]\in\Xi_{i,j,k}} A^{[1]}|e + u_{A,1}f|^2 + A_{[1]}|f|^2, \end{aligned}$$

and so $F_{i,j}^{2,1}$ can be expressed as

$$F_{i,j}^{2,1} = \left\{ \begin{bmatrix} 1 & 0 \\ \bar{d} & 1 \end{bmatrix} \begin{bmatrix} b & 0 \\ 0 & c \end{bmatrix} \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} : \begin{array}{l} b, c \in \mathbb{R}_{>0},\ d \in \mathbb{C}, \\ |e + df|^2 + \dfrac{c}{b}|f|^2 \ge 1, \\ \begin{bmatrix} e \\ f \end{bmatrix} \in \dfrac{1}{N(\mathfrak{a}_j)}\Xi_{i,j,1} \cup \dfrac{2}{N(\mathfrak{a}_j)}\Xi_{i,j,2} \end{array} \right\}.$$

Now for $\alpha, \beta \in \Bbbk$ let

$$\mathfrak{d}(\alpha, \beta) = \left\{x\alpha + y\beta : -\tfrac{1}{2} < x, y \le \tfrac{1}{2}\right\}.$$

When α and β generate a fractional ideal $\mathfrak{a}$, we have $\mathfrak{d}(\alpha, \beta)$ is a fundamental domain for $\mathbb{C}$ with respect to addition by $\mathfrak{a}$. Also if we let $\tilde{\mathfrak{d}}(\alpha, \beta)$ denote the subset of $\mathfrak{d}(\alpha, \beta)$ where the range of y is restricted to $0 \le y \le \frac{1}{2}$, this gives us a fundamental domain for $\mathfrak{d}(\alpha, \beta)$ with respect to multiplication by $Z = \{\pm 1\}$.

In particular $\mathfrak{d}(1, \sqrt{-5})$, $\mathfrak{d}(2, 1+\sqrt{-5})$, $\mathfrak{d}\big(1, \frac{1}{2}(1-\sqrt{-5})\big)$ are fundamental domains for $\mathbb{C}$ with respect to addition by $\mathcal{O}$, $\mathfrak{a}_2$ and $\mathfrak{a}_2^{-1}$ respectively, and we can put $\tilde{\mathfrak{d}}_{11} = \tilde{\mathfrak{d}}_{12} = \tilde{\mathfrak{d}}(1, \sqrt{-5})$, $\tilde{\mathfrak{d}}_{21} = \tilde{\mathfrak{d}}\big(1, \frac{1}{2}(1-\sqrt{-5})\big)$ and $\tilde{\mathfrak{d}}_{22} = \tilde{\mathfrak{d}}(2, \sqrt{-5})$. Then

$$F_{i,j}^{2,1}(P_1, P_1) = \left\{ \begin{bmatrix} 1 & 0 \\ \bar{d} & 1 \end{bmatrix} \begin{bmatrix} b & 0 \\ 0 & c \end{bmatrix} \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} : \begin{array}{l} b, c \in \mathbb{R}_{>0},\ d \in \tilde{\mathfrak{d}}_{ij}, \\ |e+df|^2 + \dfrac{c}{b}|f|^2 \ge 1, \\ \begin{bmatrix} e \\ f \end{bmatrix} \in \dfrac{1}{N(\mathfrak{a}_j)} \Xi_{i,j,1} \cup \dfrac{2}{N(\mathfrak{a}_j)} \Xi_{i,j,2} \end{array} \right\}.$$

Writing $F_{i,j}^{2,1}(P_1, P_1)$ as $F_{i,j}$, we obtain the fundamental domains $\Omega_1^2 = F_{1,1} \cup {}^t\bar{\xi}_{12} F_{1,2} \xi_{12}$ for $P_2/\Gamma_2(\mathfrak{a}_1)$ and $\Omega_2^2 = F_{1,1} \cup {}^t\bar{\xi}_{22} F_{2,2} \xi_{22}$ for $P_2/\Gamma_2(\mathfrak{a}_2)$.

5.4. ***Relations between the fundamental domains.*** So far we have used a representative set $\{\mathfrak{a}_1, \dots, \mathfrak{a}_h\}$ for $\mathrm{Cl}(\Bbbk)$ and a standard parabolic subgroup $Q^{n,m}$ of GL_n in constructing our fundamental domains. This construction is of course possible with m varied and using any other representative set of fractional ideals. We will demonstrate in this section that the fundamental domain for $P_n/\Gamma_n(\mathfrak{a}_i)$ constructed using $\{\mathfrak{a}_1, \dots, \mathfrak{a}_h\}$ and $Q^{n,m}$ can be mapped by an automorphism to a fundamental domain for $P_n/\Gamma_n(\mathfrak{a}_i^{-1})$ constructed with the representative set $\{\mathfrak{a}_1^{-1}, \dots, \mathfrak{a}_h^{-1}\}$ and $Q^{n,n-m}$.

For integers n and m where $1 \le m < n$, define the outer automorphism $\phi_{n,m}$ of $\mathrm{GL}_n(\Bbbk_\infty)$ by

$$\phi_{n,m}(g) := {}^tJ_{n,m}({}^tg^{-1})J_{n,m}, \quad g \in \mathrm{GL}_n(\Bbbk_\infty), \tag{12}$$

where

$$J_{n,m} = \begin{bmatrix} 0 & I_m \\ I_{n-m} & 0 \end{bmatrix}.$$

Note that ${}^tJ_{n,m} = (J_{n,m})^{-1} = J_{n,n-m}$ so that in particular we have $\phi_{n,m}^{-1} = \phi_{n,n-m}$. Also $\phi_{n,m}$ gives a one-to-one map between these two standard parabolic subgroups of GL_n since $\phi_{n,m}(Q^{n,m}(\Bbbk)) = Q^{n,n-m}(\Bbbk)$.

Let the ideals $\mathfrak{a}_1, \dots, \mathfrak{a}_h$, the corresponding adeles $\alpha_1, \dots, \alpha_h$, and the matrices ξ_{ij} $(1 \le i, j \le h)$ be as they were chosen in the last section. Clearly $\{\mathfrak{a}_1^{-1}, \dots, \mathfrak{a}_h^{-1}\}$ is also a set of representative ideals for ideal class group. A corresponding set of matrices representing $\mathrm{GL}_n(\Bbbk)\backslash\mathrm{GL}_n(\mathbb{A})^1/(\mathrm{GL}_n)_{\mathbb{A},\infty}^1$ is given by

$$\{D_n(\alpha_1^{-1}), \dots, D_n(\alpha_h^{-1})\} = \{\eta_i^{-1}, \dots, \eta_h^{-1}\},$$

which gives us the subgroups

$$D_n(\alpha_i^{-1})(\mathrm{GL}_n(\Bbbk_\infty)^1 \times K_f)D_n(\alpha_i^{-1})^{-1} \cap \mathrm{GL}_n(\Bbbk) = \Gamma_n(\mathfrak{a}_i^{-1}),$$

which are the respective stabilizer subgroups in $\mathrm{GL}_n(\Bbbk)$ of the lattices $L_n(\mathfrak{a}_i^{-1})$ $(i=1,\dots,h)$.

Next for each $i,\ j=1,\dots,h$ set

$$\tilde{\xi}_{ij} := {}^tJ_{n,m}\,{}^t\xi^{-1}_{i\tau_i(j)} = \begin{bmatrix} & & I_{n-m+1} & \\ & -\beta'_{i\tau_i(j)} & & \kappa_{ij}^{-1}\alpha_{i\tau_i(j)} \\ I_{m-1} & & & \\ & \beta_{i\tau_i(j)} & & -\kappa_{ij}^{-1}\alpha'_{i\tau_i(j)} \end{bmatrix},$$

which is easily verified to satisfy

$$\begin{aligned}(13)\quad \tilde{\xi}_{ij}L_n(\mathfrak{a}_i^{-1}) &= \Big(\sum_{1\le k<n-m}\mathcal{O}\boldsymbol{e}_k^{(n)}+\mathfrak{a}_j^{-1}\boldsymbol{e}_m^{(n)}\Big)+\Big(\sum_{n-m<k<n}\mathcal{O}\boldsymbol{e}_k^{(n)}+\mathfrak{a}_{\tau_i(j)}^{-1}\boldsymbol{e}_n^{(n)}\Big)\\ &\simeq L_{n-m}(\mathfrak{a}_j^{-1})\oplus L_m(\mathfrak{a}_{\tau_i(j)}^{-1}).\end{aligned}$$

Thus $\{\tilde{\xi}_{ij}\}_{j=1}^h$ is an $(n,n-m)$-splitting set for $L_n(\mathfrak{a}_i^{-1})$, and hence a complete set of representatives for $Q^{n,n-m}(\Bbbk)\backslash\mathrm{GL}_n(\Bbbk)/\Gamma_n(\mathfrak{a}_i^{-1})$.

We can also define

$$\tilde{Q}^{n,n-m}_{i.j} := Q^{n,n-m}(\Bbbk)\cap\tilde{\xi}^{n,n-m}_{ij}\Gamma_n(\mathfrak{a}_i^{-1})(\tilde{\xi}^{n,n-m}_{ij})^{-1},$$

$$\tilde{F}^{n,n-m}_{i,j} = \Big\{A\in P_n : |A|_\infty^{[n-m]}\le\Big(\frac{N(\mathfrak{a}_k^{-1})}{N(\mathfrak{a}_j^{-1})}\Big)^2\big|{}^t(\overline{\tilde{\xi}_{ij}\gamma\tilde{\xi}_{ik}^{-1}})A(\tilde{\xi}_{ij}\gamma\tilde{\xi}_{ik}^{-1})\big|_\infty^{[n-m]},\ 1\le k\le h,\ \gamma\in\Gamma_n(\mathfrak{a}^{-1})\Big\},$$

$$\tilde{\mathfrak{D}}^{n,n-m}_{i,j} = \left\{\begin{bmatrix} d_{11} & \cdots & d_{1,m}\\ \vdots & \ddots & \vdots\\ d_{n-m,1} & \cdots & d_{n-m,m}\end{bmatrix} : d_{n-m,m}\in\tilde{\mathfrak{d}}_{ij},\ d_{rs}\in\begin{cases}\mathfrak{d}_{\mathcal{O}}, & r<n-m,\ s<m,\\ \mathfrak{d}_{\tau_i(j)}, & r<n-m,\ s=m,\\ \mathfrak{d}'_j, & r=n-m,\ s<m\end{cases}\right\},$$

where the fundamental domains $\mathfrak{d}_k$, $\mathfrak{d}'_k$, $\tilde{\mathfrak{d}}_{ik}$, $\mathfrak{d}_{\mathcal{O}}$ are taken as in the previous section, and

$$\tilde{F}^{n,n-m}_{i,j}(S,S') = \big\{A\in\tilde{F}^{n,n-m}_{i,j} : A^{[n-m]}\in S,\ A_{[m]}\in S',\ u_{A,n-m}\in\tilde{\mathfrak{D}}^{n,n-m}_{i,j}\big\}$$

for arbitrary subsets $S\subset P_{n-m}$, $S'\subset P_m$. These are precisely the groups $Q^{n,m}_{i,j}$ and sets $F^{n,m}_{i,j}$, $\mathfrak{D}^{n,m}_{i,j}$ and $F^{n,m}_{i,j}(S,S')$ from the previous section with $\mathfrak{a}_i^{-1}$ and $\tilde{\xi}_{ik}$ in place of the $\mathfrak{a}_i$ and ξ_{ik} respectively, when $m=n-m$. It is easily verified that $\phi_{n,m}(Q^{n,m}_{i,j})=\tilde{Q}^{n,n-m}_{i,\tau_i(j)}$.

Lemma 17. *For $A\in P_n$,*

$$\phi_{n,m}(A)^{[n-m]} = {}^tA^{-1}_{[n-m]},\quad \phi_{n,m}(A)_{[m]} = {}^t(A^{[m]})^{-1},$$

$$u_{\phi_{n,m}(A),n-m} = -{}^tu_{A,m}.$$

Proof. Apply the automorphism $\phi_{n,m}$ to both sides of (11). □

Given a set S consisting of invertible matrices, denote the set $\{{}^ts^{-1} : s \in S\}$ by ${}^tS^{-1}$.

Lemma 18. *For $S \subset P_m$ and $S' \subset P_{n-m}$,*

$$\phi_{n,m}(F_{i,j}^{n,m}(S, S')) = \tilde{F}_{i,\tau_i(j)}^{n,n-m}({}^tS'^{-1}, {}^tS^{-1}).$$

Proof. We first show that $\phi_{n,m}(F_{i,j}^{n,m}) = \tilde{F}_{i,\tau_i(j)}^{n,n-m}$. First consider $A \in F_{i,j}^{n,m}$. Put

$$A(k, \gamma) = {}^t\overline{(\xi_{ij}\gamma\xi_{ik}^{-1})}A(\xi_{ij}\gamma\xi_{ik}^{-1})$$

for $1 \le k \le h$ and $\gamma \in \Gamma_i$. We have

$$|A(k, \gamma)| = \left(\frac{\kappa_{ij}}{\kappa_{ik}}\right)^2 |A| = \left(\frac{\kappa_{ij}}{\kappa_{ik}}\right)^2 |A^{[m]}|\,|A_{[n-m]}|.$$

Substitute this and $|A(k, \gamma)^{[m]}| = |A(k, \gamma)|\,|A(k, \gamma)_{[n-m]}|^{-1}$ into the inequality

$$|A^{[m]}|_\infty \le \left(\frac{N(\mathfrak{a}_k)}{N(\mathfrak{a}_j)}\right)^2 |A(k, \gamma)^{[m]}|_\infty.$$

Rearranging, we get

$$|A_{[n-m]}|_\infty^{-1} \le \left(\frac{|\kappa_{ik}^{-1}|_\infty N(\mathfrak{a}_k)}{|\kappa_{ij}^{-1}|_\infty N(\mathfrak{a}_j)}\right)^2 |A(k, \gamma)_{[n-m]}|_\infty^{-1},$$

which, using the previous lemma, becomes

$$|\phi_{n,m}(A)|_\infty^{[n-m]} \le \left(\frac{N(\mathfrak{a}_{\tau_i(k)}^{-1})}{N(\mathfrak{a}_{\tau_i(j)}^{-1})}\right)^2 |\phi_{n,m}(A(k, \gamma))|_\infty^{[n-m]},$$

and since

$$\phi_{n,m}(A(k, \gamma)) = {}^t\overline{(\tilde{\xi}_{i\tau_i(j)}{}^t\gamma^{-1}\tilde{\xi}_{i\tau_i(k)}^{-1})}\phi_{n,m}(A)(\tilde{\xi}_{i\tau_i(j)}{}^t\gamma^{-1}\tilde{\xi}_{i\tau_i(k)}^{-1}),$$

this shows that $\phi_{n,m}(A) \in \tilde{F}_{i,\tau_i(j)}^{n,n-m}$. Thus $\phi_{n,m}(F_{i,j}^{n,m}) \subset \tilde{F}_{i,\tau_i(j)}^{n,n-m}$ and similarly $\phi_{n,n-m}(\tilde{F}_{i,\tau_i(j)}^{n,n-m}) \subset F_{i,j}^{n,m}$. The rest of our result follows from the previous lemma. □

Lemma 19. *Let Γ be a subgroup of $\mathrm{GL}_n(\Bbbk_\infty)$ acting on a subset X of P_n, the action being the one defined in* (10). *If F is a given fundamental domain for X/Γ and ϕ a group automorphism of $\mathrm{GL}_n(\Bbbk_\infty)$ that is also a topological isomorphism, then $\phi(F)$ is a fundamental domain for $\phi(X)/\phi(\Gamma)$.*

Proof. Since ϕ is both a group homomorphism and a topological isomorphism, $X = F^- \cdot \Gamma$ implies $\phi(X) = \phi(F)^- \cdot \phi(\Gamma)$. Also, for $g \in \Gamma$, if the intersection of $\phi(F)^\circ$ and $\phi(F)^- \cdot \phi(g)$ is nonempty, then so is $F^\circ \cap F^- \cdot g$, implying $g \in Z$. Since Z consists of all roots of unity in $\Bbbk$, we have $\phi(g) \in Z$. □

In particular, if for $k = 1, \dots, h$ we let $\mathfrak{B}_k$ and $\mathfrak{C}_k$ be fundamental domains for $P_m/\Gamma_m(\mathfrak{a}_k)$ and $P_{n-m}/\Gamma_{n-m}(\mathfrak{a}_k)$ respectively as in the end of the previous section, then ${}^t\mathfrak{B}_k^{-1}$ and ${}^t\mathfrak{C}_k^{-1}$ are respectively fundamental domains for $P_{n-m}/\Gamma_{n-m}(\mathfrak{a}_k^{-1})$ and $P_m/\Gamma_m(\mathfrak{a}_k^{-1})$. Also we have:

Corollary 20. *The set*

$$\phi_{n,m}(F^{n,m}_{i,j}(\mathfrak{B}_j, \mathfrak{C}_{\tau_i(j)}))$$

is a fundamental domain for $\tilde{F}^{n,n-m}_{i,\tau_i(j)}/\tilde{Q}^{n,n-m}_{i,\tau_i(j)}$.

Corollary 21. *The set*

$${}^t\big(\Omega^{n,m}_i(\mathfrak{B}_1, \dots \mathfrak{B}_h, \mathfrak{C}_1, \dots, \mathfrak{C}_h)\big)^{-1}$$

is a fundamental domain for $P_n/\,\Gamma_n(\mathfrak{a}_i^{-1})$.

Since

$$\tilde{F}^{n,n-m}_{i.j}({}^t\mathfrak{C}_j^{-1}, {}^t\mathfrak{B}^{-1}_{\tau_i(j)}) = \phi_{n,m}(F^{n,m}_{i,\tau_i(j)}(\mathfrak{B}_{\tau_i(j)}, \mathfrak{C}_j)),$$

the first corollary is consistent with Lemma 15 in the previous section.

Similarly if we put

$$\tilde{\Omega}^{n,n-m}_i(\mathfrak{C}_1, \dots \mathfrak{C}_h, \mathfrak{B}_1, \dots, \mathfrak{B}_h) = \bigcup_{j=1}^{h} {}^t\bar{\tilde{\xi}}_{ij}\, \tilde{F}^{n,m}_{i,j}({}^t\mathfrak{C}_j^{-1}, {}^t\mathfrak{B}^{-1}_{\tau_i(j)})\tilde{\xi}_{ij}$$

then $\tilde{\Omega}^{n,n-m}_i(\mathfrak{C}_1, \dots \mathfrak{C}_h, \mathfrak{B}_1, \dots, \mathfrak{B}_h) = {}^t\big(\Omega^{n,m}_i(\mathfrak{B}_1, \dots \mathfrak{B}_h, \mathfrak{C}_1, \dots, \mathfrak{C}_h)\big)^{-1}$ and according to Theorem 16, this set is indeed a fundamental domain for $P_n/\,\Gamma_n(\mathfrak{a}_i^{-1})$.

Appendix: Voronoi reduction

by Takao Watanabe

We present here generalizations of results from [Watanabe et al. 2013, §4], without the assumption that the underlying number field is totally real.

Let $\Bbbk$, $\mathcal{O}$ and P_n be as previously defined in this paper. We consider the space of self-adjoint matrices in $M_n(\Bbbk_\infty)$ (with respect to the inner product $\langle\ ,\ \rangle$ as defined in [Watanabe et al. 2013, §1]), which we denote here by H_n. Identifying H_n with $\prod_{\sigma\in \boldsymbol{p}_\infty} H_n(\Bbbk_\sigma)$, where $H_n(\Bbbk_\sigma)$ denotes the set of $n\times n$ real symmetric (complex Hermitian) matrices when σ is real (imaginary respectively), we see that P_n is the set of positive definite matrices in H_n.

Also as per [Watanabe et al. 2013, §1], we use the inner product $(\ ,\ )$ on H_n defined by

$$(A, B) = \sum_{\sigma\in \boldsymbol{p}_\infty} \mathrm{Tr}_{\Bbbk_\sigma/\mathbb{R}}(\mathrm{Tr}(A_\sigma B_\sigma))$$

for $A = (A_\sigma)_{\sigma\in\boldsymbol{p}_\infty}$, $B = (B_\sigma)_{\sigma\in\boldsymbol{p}_\infty} \in H_n$.

Following [Watanabe et al. 2013, §2], we fix a projective $\mathcal{O}$-module $\Lambda\subset \Bbbk^n$ of rank n and consider the arithmetical minimum function

$$\mathsf{m}_\Lambda(A) = \inf_{x\in\Lambda\setminus\{0\}} \langle Ax, x\rangle$$

on P_n^-. The set

$$K_1(\mathsf{m}_\Lambda) = \{A \in P_n^- : \mathsf{m}_\Lambda(A) \geq 1\},$$

known as the Ryshkov polyhedron of m_Λ, is a locally finite polyhedron contained in P_n [Watanabe et al. 2013, Lemma 2.1 and Proposition 2.2]. The set of 0-dimensional faces of $K_1(\mathsf{m}_\Lambda)$, denoted by $\partial^0 K_1(\mathsf{m}_\Lambda)$, is characterized in [Watanabe et al. 2013, Theorem 2.5].

Now for a given $A \in P_n$ and a positive constant θ, define the sets

$$H_{A,\theta} = \{B \in H_n : (A, B) \leq \theta\},$$
$$[A]_\theta = \partial^0 K_1(\mathsf{m}_{\Lambda_0}) \cap H_{A,\theta}.$$

Lemma A1. *$[A]_\theta$ is a finite set.*

Proof. Since $H_{A,\theta} \cap P_n^-$ is compact [Faraut and Korányi 1994, Corollary I.1.6] and $K_1(\mathsf{m}_\Lambda)$ is a locally finite polyhedron, it follows that their intersection $K_1(\mathsf{m}_\Lambda) \cap H_{A,\theta}$ is a polytope. Hence $[A]_\theta$ must be finite. □

Lemma A2. *For an $A \in P_n$, there exists $B_0 \in \partial^0 K_1(\mathsf{m}_\Lambda)$ such that*

$$\inf_{B \in K_1(\mathsf{m}_\Lambda)} (A, B) = (A, B_0)$$

and hence A is in D_{B_0}, the perfect domain of B_0 [Watanabe et al. 2013, §3]. *Here*

$$D_{B_0} = \left\{ \sum_{x \in S_\Lambda(B_0)} \lambda_x x\, {}^t\bar{x} : \lambda_x \geq 0 \right\},$$

where

$$S_\Lambda(B_0) = \{x \in \Lambda : \mathsf{m}_\Lambda(B_0) = \langle B_0 x, x\rangle\}.$$

Proof. Take a sufficiently large $\theta > 0$ whereby $[A]_\theta$ is nonempty. Since $K_1(\mathsf{m}_\Lambda)$ is the convex hull of $\partial K_1(\mathsf{m}_\Lambda)$ [Watanabe et al. 2013, Theorem 2.6], we have

$$\inf_{B \in K_1(\mathsf{m}_\Lambda)} (A, B) = \inf_{B \in \partial K_1(\mathsf{m}_\Lambda)} (A, B) = \inf_{B \in [A]_\theta} (A, B),$$

which together with the previous lemma proves the existence of B_0. The proof that $A \in D_{B_0}$ is the same as in [Watanabe et al. 2013, Lemma 4.8]. □

Next consider the set

$$\Bbbk_\infty^+ = \{(\alpha_\sigma)_{\sigma \in \boldsymbol{p}_\infty} : \alpha_\sigma > 0 \text{ for all } \sigma \in \boldsymbol{p}_\infty\}.$$

Lemma A3. *The subset $\{\beta\bar{\beta} : \beta \in \Bbbk^\times\}$ of $\Bbbk_\infty$ is dense in $\Bbbk_\infty^+$.*

Proof. Define the norm $\|\cdot\|$ on $\Bbbk_\infty$ by

$$\|\alpha\| = \max_{\sigma \in \boldsymbol{p}_\infty} \sqrt{\alpha_\sigma \bar{\alpha}_\sigma}, \quad \alpha = (\alpha_\sigma) \in \Bbbk_\infty.$$

Now given a $\alpha \in \Bbbk_\infty^+$ there is an element $\sqrt{\alpha} \in \Bbbk_\infty^+$ such that $(\sqrt{\alpha})^2 = \alpha$. Since $\Bbbk$ is dense in $\Bbbk_\infty$, for a sufficiently small $\epsilon > 0$ we can find $\beta \in \Bbbk^\times$ such that

$$\|\sqrt{\alpha} - \beta\| < \frac{\epsilon}{2\|\sqrt{\alpha}\| + 1} < 1.$$

From $\|\beta\| < \|\sqrt{\alpha}\| + 1$, we have $\|\sqrt{\alpha} + \beta\| < 2\|\sqrt{\alpha}\| + 1$, and thus

$$\begin{aligned}\|\alpha - \beta\bar{\beta}\| &= \tfrac{1}{2}\left\|(\sqrt{\alpha} - \beta)(\sqrt{\alpha} + \bar{\beta}) + (\sqrt{\alpha} + \beta)(\sqrt{\alpha} - \bar{\beta})\right\| \\ &\le \tfrac{1}{2}\big(\|\sqrt{\alpha} - \beta\|\,\|\sqrt{\alpha} + \beta\| + \|\sqrt{\alpha} + \beta\|\,\|\sqrt{\alpha} - \beta\|\big) < \epsilon.\end{aligned}$$ □

Lemma A4. $$\Bbbk_\infty^+ \cup \{0\} = \left\{\sum_{k=1}^{l} \lambda_k \beta_k {}^t\bar{\beta}_k : 1 \le l \in \mathbb{Z},\ \lambda_k \in \mathbb{R}_{\ge 0},\ \beta_k \in \Bbbk^\times\right\}.$$

Proof. See the proof of [Watanabe et al. 2013, Lemma 4.2]. □

As a result of the previous lemma, if we define the subsets

$$\Omega_1 = \left\{\sum_{k=1}^{l} \alpha_k x_k {}^t\bar{x}_k : 1 \le l \in \mathbb{Z},\ \alpha_k \in \Bbbk_\infty^+ \cup \{0\},\ x_i \in \Bbbk^n\right\},$$

$$\Omega_2 = \left\{\sum_{k=1}^{l} \lambda_k x_k {}^t\bar{x}_k : 1 \le l \in \mathbb{Z},\ \lambda_k \in \mathbb{R}_{\ge 0},\ x_i \in \Bbbk^n\right\}$$

of P_n^-, we have $\Omega_1 = \Omega_2$. Also by Lemma A2, $P_n \subset \Omega_2 = \Omega_1$.

Lemma A5. $$\Omega_2 = \bigcup_{B \in \partial^0 K_1(\mathsf{m}_\Lambda)} D_B.$$

Proof. For any $A \in \Omega_2 \backslash \{0\}$, following the same arguments as in the proofs of [Watanabe et al. 2013, Lemmas 4.7 and 4.8], we can find an element $B_0 \in \partial^0 K_1(\mathsf{m}_\Lambda)$ such that $\inf_{B \in K_1(\mathsf{m}_\Lambda)}(A, B) = (A, B_0)$ and hence $A \in D_{B_0}$. □

Finally take a complete set of representatives $B_1, \dots, B_t$ for $\partial^0 K_1(\mathsf{m}_\Lambda)/\mathrm{GL}(\Lambda)$, where the right action is the same one as (10), and for each $k = 1, \dots, t$ define the subgroups $\Gamma_{B_k} = \{\gamma \in \mathrm{GL}(\Lambda) : B_k \cdot {}^t\bar{\gamma} = B_k\}$. Since for any $A \in \partial^0 K_1(\mathsf{m})$ and $\gamma \in \mathrm{GL}(\Lambda)$ we have $S_\Lambda(A \cdot \gamma) = \gamma^{-1} S_\Lambda(A)$ and hence $D_{A \cdot {}^t\bar{\gamma}} = (D_A) \cdot \gamma^{-1}$, we see that Γ_{B_k} stabilizes D_{B_k} for each k. Thus we conclude from the previous lemma the following result.

Theorem A6. $$\Omega_2/\mathrm{GL}(\Lambda) = \bigcup_{k=1}^{t} D_{B_k}/\Gamma_{B_k}.$$

This is analogous to [Watanabe et al. 2013, Theorem 4.9]. In particular when $n = 1$, if we take $\Lambda = O$, we have $\mathrm{GL}(\Lambda) = O^\times$ and $P_1 = \Bbbk_\infty^+ = \Omega_1 \backslash \{0\} = \Omega_2 \backslash \{0\}$.

Since the action of $\mathcal{O}^\times$ on $\Bbbk_\infty^+$ is simply $x \cdot \epsilon = \bar{\epsilon}\epsilon x$ ($x \in \Bbbk_\infty^+$, $\epsilon \in \mathcal{O}^\times$), we have $\Gamma_{B_k} = Z$ acts trivially on D_{B_k} for each k. Thus we obtain the decomposition

$$P_1/\mathcal{O}^\times = \Bbbk_\infty^+/\mathcal{O}^\times = \bigcup_{k=1}^{t} (D_{B_k} \backslash \{0\}).$$

By definition each $D_{B_k} \backslash \{0\}$ is invariant under multiplication by $\mathbb{R}_{>0}$, so this establishes the existence of the fundamental domain Ω^1 in the conclusion of Section 5.2.

Acknowledgements

I would like to express my utmost gratitude to Professor Takao Watanabe for his patience and invaluable guidance in my research, as well as his work in [Watanabe 2014], which has served as a basis of the results in this paper. I also would like to thank Professor Achill Schürmann and Dr. Mathieu Dutour Sikirić for confirming via private communication their observation on the fundamental domains of Korkin and Zolotarev and Grenier.

References

[Bombieri and Gubler 2006] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs **4**, Cambridge Univ. Press, 2006. MR Zbl

[Borel 1962] A. Borel, "Ensembles fondamentaux pour les groupes arithmétiques", pp. 23–40 in *Colloque sur la théorie des groupes algébriques* (Brussels, 1962), Librairie Universitaire, Louvain, 1962. MR Zbl

[Borel 1963] A. Borel, "Some finiteness properties of adele groups over number fields", *Inst. Hautes Études Sci. Publ. Math.* **16**:1 (1963), 5–30. MR Zbl

[Cohen 2000] H. Cohen, *Advanced topics in computational number theory*, Graduate Texts in Mathematics **193**, Springer, New York, 2000. MR Zbl

[Cohn 1965] H. Cohn, "A numerical survey of the floors of various Hilbert fundamental domains", *Math. Comp.* **19** (1965), 594–605. MR Zbl

[Faraut and Korányi 1994] J. Faraut and A. Korányi, *Analysis on symmetric cones*, Oxford University Press, New York, 1994. MR Zbl

[Grenier 1988] D. Grenier, "Fundamental domains for the general linear group", *Pacific J. Math.* **132**:2 (1988), 293–317. MR Zbl

[Humbert 1939] P. Humbert, "Théorie de la réduction des formes quadratiques définies positives dans un corps algébrique K fini", *Comment. Math. Helv.* **12** (1939), 263–306. MR Zbl

[Korkin and Zolotarev 1873] A. Korkine and G. Zolotareff, "Sur les formes quadratiques", *Math. Ann.* **6**:3 (1873), 366–389. MR JFM

[Minkowski 1905] H. Minkowski, "Diskontinuitätsbereich für arithmetische Äquivalenz", *J. Reine Angew. Math.* **129** (1905), 220–274. MR JFM

[Shimura 2010] G. Shimura, *Arithmetic of quadratic forms*, Springer, New York, 2010. MR Zbl

[Watanabe 2014] T. Watanabe, "Ryshkov domains of reductive algebraic groups", *Pacific J. Math.* **270**:1 (2014), 237–255. MR Zbl

[Watanabe et al. 2013] T. Watanabe, S. Yano, and T. Hayashi, “Voronoï’s reduction theory of GL_n over a totally real number field”, pp. 213–232 in *Diophantine methods, lattices, and arithmetic theory of quadratic forms*, edited by W. K. Chan et al., Contemporary Mathematics **587**, Amer. Math. Soc., Providence, RI, 2013. MR Zbl

Received May 31, 2015. Revised January 23, 2017.

LEE TIM WENG
GRADUATE SCHOOL OF SCIENCE
OSAKA UNIVERSITY
TOYONAKA
JAPAN
lt-weng@cr.math.sci.osaka-u.ac.jp
now_then_here_there@hotmail.com

TAKAO WATANABE
GRADUATE SCHOOL OF SCIENCE
OSAKA UNIVERSITY
TOYONAKA
JAPAN
twatanabe@math.sci.osaka-u.ac.jp

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.169

GROWTH AND DISTORTION THEOREMS FOR SLICE MONOGENIC FUNCTIONS

GUANGBIN REN AND XIEPING WANG

We establish the sharp growth and distortion theorems for slice monogenic extensions of univalent functions on the unit disc $\mathbb{D} \subset \mathbb{C}$ in the setting of Clifford algebras, based on a new convex combination identity. The analogous results are also valid in the quaternionic setting for slice regular functions and we can even prove a Koebe type one-quarter theorem in this case. Our growth and distortion theorems for slice regular (slice monogenic) extensions to higher dimensions of univalent holomorphic functions hold without extra geometric assumptions, in contrast to the setting of several complex variables in which the growth and distortion theorems fail in general and hold only for some subclasses with the starlike or convex assumption.

1. Introduction

In geometric function theory of holomorphic functions of one complex variable, the following well-known growth and distortion theorems (see, e.g., [Duren 1983; Graham and Kohr 2003]) mark the beginning of the systematic study of univalent functions.

Theorem 1.1 (growth and distortion theorems). *Let F be a univalent function on the open unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ such that $F(0) = 0$ and $F'(0) = 1$. Then for each $z \in \mathbb{D}$, the following inequalities hold*:

$$\frac{|z|}{(1+|z|)^2} \leq |F(z)| \leq \frac{|z|}{(1-|z|)^2}; \tag{1-1}$$

$$\frac{1-|z|}{(1+|z|)^3} \leq |F'(z)| \leq \frac{1+|z|}{(1-|z|)^3}; \tag{1-2}$$

$$\frac{1-|z|}{1+|z|} \leq \left|\frac{zF'(z)}{F(z)}\right| \leq \frac{1+|z|}{1-|z|}. \tag{1-3}$$

This work was supported by the NNSF of China (11371337).
MSC2010: primary 30G35; secondary 30C45.
Keywords: quaternions, Clifford algebra, slice regular (slice monogenic) functions, growth and distortion theorems, Koebe one-quarter theorem.

Moreover, equality holds for one of these six inequalities at some point $z_0 \in \mathbb{D} \setminus \{0\}$ *if and only if* F *is a rotation of the Koebe function, i.e.,*

$$F(z) = \frac{z}{(1 - e^{i\theta}z)^2}, \qquad \forall z \in \mathbb{D},$$

for some $\theta \in \mathbb{R}$.

The extension of geometric function theory to higher dimensions was suggested by H. Cartan [1933], but the first meaningful result was only made in 1991 by Barnard, Fitzgerald and Gong [Barnard et al. 1991]. Since then, the geometric function theory in several complex variables has been extensively studied, see, for example, [Gong 1998; Graham and Kohr 2003]. In particular, the growth theorem holds for starlike mappings on starlike circular domains [Liu and Ren 1998a], and for convex mappings on convex circular domains [Liu and Ren 1998b].

However, as far as we know, nearly nothing has been done about the corresponding theory for other classes of functions, such as the classical regular (monogenic) functions in the sense of Fueter and the recently introduced slice regular (slice monogenic) functions, mainly because both regularity (monogenicity) and slice regularity (slice monogenicity) of functions are seldom preserved under multiplication and composition, because of the noncommutativity of the underlying algebras on which these functions are defined.

In this paper, we shall focus on slice regular and slice monogenic functions and aim to generalize Theorem 1.1 to the noncommutative setting for slice regular and slice monogenic extensions of univalent functions on the unit disc $\mathbb{D} \subset \mathbb{C}$. The theory of slice regular functions of one quaternionic variable was initiated recently by Gentili and Struppa [2006; 2007], and was also extended by the same authors to octonions [2010] for octonionic slice regular functions. The related theory of slice monogenic functions on domains in the paravector space $\mathbb{R}^{n+1}$ with values in the Clifford algebra $\mathbb{R}_n$ was introduced in [Colombo et al. 2009; 2010]. For a more complete insight and further references, we refer the reader to the monographs [Gentili et al. 2013; Colombo et al. 2011a]. These function theories were also unified and generalized in [Ghiloni and Perotti 2011a] using the concept of slice functions on the so-called quadratic cone of a real alternative *-algebra, based on a slight modification of a well-known construction due to Fueter. The theory of slice regular functions on real alternative *-algebras is now well-developed through a series of papers mainly due to Ghiloni and Perotti after their seminal work [Ghiloni and Perotti 2011a]. It is also well worth mentioning that this recently introduced theory of slice regular (slice monogenic) functions is significantly different from the more classical theory of regular (monogenic) functions in the sense of Fueter (cf. [Brackx et al. 1982; Colombo et al. 2004; Gürlebeck et al. 2008]), and has elegant applications to the functional calculus for noncommutative operators [Colombo et al. 2011a], to Schur analysis [Alpay et al. 2016], and to the construction and classification of orthogonal complex structures on dense open subsets of $\mathbb{R}^4 \simeq \mathbb{H}$ [Gentili et al. 2014].

We are now in a position to state one of our main results in the case of the Clifford algebra $\mathbb{R}_n$ for slice monogenic extensions to the open unit ball

$$\mathbb{B} := \{x \in \mathbb{R}^{n+1} : |x| < 1\}$$

of univalent functions on the unit disc $\mathbb{D} \subset \mathbb{C}$.

Theorem 1.2. *Let $F : \mathbb{D} \to \mathbb{C}$ be a univalent function such that $F(0) = 0$ and $F'(0) = 1$, and let $f : \mathbb{B} \to \mathbb{R}_n$ be the slice monogenic extension of F. Then for each $x \in \mathbb{B}$, the following inequalities hold*:

$$\frac{|x|}{(1+|x|)^2} \leq |f(x)| \leq \frac{|x|}{(1-|x|)^2}; \tag{1-4}$$

$$\frac{1-|x|}{(1+|x|)^3} \leq |f'(x)| \leq \frac{1+|x|}{(1-|x|)^3}; \tag{1-5}$$

$$\frac{1-|x|}{1+|x|} \leq \left|xf'(x) * f^{-*}(x)\right| \leq \frac{1+|x|}{1-|x|}. \tag{1-6}$$

Moreover, equality holds for one of these six inequalities at some point $x_0 \in \mathbb{B} \setminus \{0\}$ if and only if

$$f(x) = x(1 - xe^{i\theta})^{-*2}, \qquad \forall x \in \mathbb{B},$$

for some $\theta \in \mathbb{R}$.

Although Theorem 1.2 coincides in form with Theorem 1.1, the classical approach to Theorem 1.1 cannot be directly applied in this new case of the Clifford algebra $\mathbb{R}_n$, since there lacks a fruitful theory of compositions for slice monogenic functions. We shall reduce Theorem 1.2 to Theorem 1.1 via a new convex combination identity; see (3-11). We remark that in contrast to the setting of several complex variables in which the growth and distortion theorems fail to hold in general [Cartan 1933] and can only be restricted to the starlike or convex subclasses, our result for slice monogenic extensions of univalent functions holds without extra geometric assumptions. This new phenomenon is in a certain sense related to the rigidity of the functions under consideration. There is a significant difference between slice monogenic functions and holomorphic functions of several complex variables, although they are both the generalizations in higher dimensions of holomorphic functions of one complex variable. The former are closer to holomorphic functions of one complex variable, and each of them can be completely determined by its values on a set that lies in a complex slice and has an accumulation point in its domain of definition. However, this is not the case for the latter, each of which is not always determined by its values on a complex submanifold of positive codimensions in its domain of definition. From this perspective, we realize that holomorphic functions of several complex variables are less rigid than slice monogenic functions so that certain extra geometric assumptions such as starlikeness and convexity are naturally present in the geometric function theory in several complex variables.

A result analogous to Theorem 1.2 also holds in the setting of quaternions (see Theorem 4.7). As an application, we can prove a covering theorem, i.e., the so-called Koebe type one-quarter theorem (see Theorem 4.10, a generalization of [Gal et al. 2015, Theorem 3.11 (1)]), with the help of the open mapping theorem, which is now known to hold only for slice regular functions defined on symmetric slice domains in $\mathbb{H}$ with values in $\mathbb{H}$ rather than slice monogenic functions defined on symmetric slice domains in paravector space $\mathbb{R}^{n+1}$ with values in the Clifford algebra $\mathbb{R}_n$.

We now describe in more detail the structure of the paper. In Section 2, we set up basic notation and give some preliminary results. In Section 3, we first prove in Proposition 3.1 a general formula to express the squared norm of a slice monogenic function defined on a symmetric slice domain in the paravector space $\mathbb{R}^{n+1}$, in terms of the values of the function at two conjugate points on some fixed slice of the domain. For slice monogenic functions that preserve one slice, we provide in Lemma 3.2 the aforementioned convex combination identity, which is the key ingredient in proving Theorem 1.2. Section 4 is devoted to the detailed proofs of the analogous results and the Koebe type one-quarter theorem (Theorem 4.10) for slice regular functions in the quaternionic setting. Thanks to the specialty of quaternions, we can also provide in Corollary 4.4 a sufficient and necessary condition under which the aforementioned convex combination identity holds identically. Finally, Section 5 provides a concluding remark and an open question.

2. Preliminaries

We recall in this section some necessary definitions and preliminary results on real Clifford algebras and slice monogenic functions. For a more complete insight, we refer the reader to the monograph [Colombo et al. 2011a].

The real Clifford algebra $\mathbb{R}_n = \mathrm{Cl}_{0,n}$ is an associative algebra over $\mathbb{R}$ generated by n basis elements $e_1, e_2, \dots, e_n$, subject to the relation

$$e_i e_j + e_j e_i = -2\delta_{ij}, \quad i, j = 1, 2, \dots, n.$$

As a real vector space, $\mathbb{R}_n$ has dimension 2^n. Each element b in $\mathbb{R}_n$ can be represented *uniquely* as

$$b = \sum_A b_A e_A,$$

where $b_A \in \mathbb{R}$, $e_0 = 1$, $e_A := e_{h_1} e_{h_2} \cdots e_{h_r}$, and $A = h_1 \cdots h_r$ is a multi-index such that $1 \le h_1 < \cdots < h_r \le n$. The real number b_0 is called the *scalar* part of

b and is denoted by $\mathrm{Sc}(b)$ as usual. The *Clifford conjugate* of each generator e_i, $i = 1, 2, \dots, n$, is defined to be $\bar{e}_i = -e_i$, and thus extends to each e_A by setting

$$\bar{e}_A := \bar{e}_{h_r}\bar{e}_{h_{r-1}}\cdots\bar{e}_{h_1} = (-1)^r e_{h_r}e_{h_{r-1}}\cdots e_{h_1} = (-1)^{r(r+1)/2}e_A,$$

and further extends by linearity to each element $b = \sum_A b_A e_A \in \mathbb{R}_n$ so that

$$\bar{b} = \sum_A b_A \bar{e}_A.$$

Therefore, the Clifford conjugate is an antiautomorphism of $\mathbb{R}_n$, i.e., $\overline{ab} = \bar{b}\bar{a}$ for any $a, b \in \mathbb{R}_n$. Moreover, the Euclidean inner product on $\mathbb{R}_n \simeq \mathbb{R}^{2^n}$ is given by

$$\langle a, b\rangle := \mathrm{Sc}(a\bar{b}) = \sum_A a_A b_A \tag{2-1}$$

for any $a = \sum_A a_A e_A,\ b = \sum_A b_A e_A \in \mathbb{R}_n$, so it follows from the simple identity

$$\langle a, b\rangle = \tfrac{1}{2}\big(|a+b|^2 - |a|^2 - |b|^2\big)$$

that

$$\langle a, b\rangle = \langle b, a\rangle = \langle \bar{a}, \bar{b}\rangle = \langle \bar{b}, \bar{a}\rangle. \tag{2-2}$$

It is worth remarking here that for $\mathbb{R}_n (n \geq 3)$ the multiplicative property of the Euclidean norm fails in general, and holds only for some special cases; see [Colombo et al. 2011a, Proposition 2.1.17] or [Gürlebeck et al. 2008, Theorem 3.14 (ii)]. In particular, it holds that

$$|ab| = |ba| = |a||b| \tag{2-3}$$

whenever one of a and b is a paravector (see below for this definition). This simple fact will be useful for our argument in Section 3.

For convenience, some specific elements in $\mathbb{R}_n$ can be identified with vectors in the Euclidean space $\mathbb{R}^{n+1}$: an element $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ will be identified with a so-called 1-*vector* in the Clifford algebra $\mathbb{R}_n$ through the map

$$(x_1, x_2, \dots, x_n) \mapsto \underline{x} = x_1e_1 + e_2x_2 + \cdots + x_ne_n;$$

and an element $(x_0, x_1, \dots, x_n) \in \mathbb{R}^{n+1}$ will be identified with

$$x = x_0 + \underline{x} = x_0 + x_1e_1 + \cdots + x_ne_n,$$

which is called a *paravector*. Now for any two 1-vectors $x, y \in \mathbb{R}^n$, the Euclidean inner product becomes

$$\langle x, y\rangle = \mathrm{Sc}(x\bar{y}) = -\tfrac{1}{2}(xy + yx),$$

and consequently,

$$xy = -\langle x, y\rangle + x \wedge y,$$

where

$$x \wedge y := \tfrac{1}{2}(xy - yx)$$

is called the *outer product* (see [Brackx et al. 1982, p.4; Gürlebeck et al. 2008, p.58]) or *wedge product* (see [Colombo et al. 2004, p.218, Definition 4.1.9; Colombo et al. 2011a, p.21]) of x and y. It is noteworthy here that *in general the operator* $\wedge$ *is a mapping from* $\mathbb{R}^n \times \mathbb{R}^n$ *to* $\mathbb{R}_n$, *not to* $\mathbb{R}^n$. Furthermore, under the identifications above, a vector x in $\mathbb{R}^{n+1}$ can be taken as a Clifford number

$$x = x_0 + \sum_{i=1}^{n} x_i e_i$$

so that it has inverse

$$x^{-1} = \frac{\bar{x}}{|x|^2},$$

where $\bar{x}$ is the *conjugate* of x given by $\bar{x} = x_0 - \sum_{i=1}^{n} x_i e_i$, and the norm of x is induced by the inner product given above, that is, $|x| = \langle x, x\rangle^{\frac{1}{2}}$. Every $x = x_0 + x_1e_1 + \cdots + x_ne_n \in \mathbb{R}^{n+1}$ is composed by the *scalar* part $\mathrm{Sc}(x) = x_0 \in \mathbb{R}$ and the *vector* part $\underline{x} = x_1e_1 + \cdots + x_ne_n \in \mathbb{R}^n$, and it can be expressed alternatively as $x = u + Iv$, where $u, v \in \mathbb{R}$ and

$$I = \frac{\underline{x}}{|\underline{x}|}$$

if $\underline{x} \neq 0$, otherwise we take I arbitrarily in $\mathbb{R}^n$ such that $I^2 = -1$. Then I is an element of the unit $(n-1)$-sphere of 1-vectors in $\mathbb{R}^n$,

$$\mathbb{S} = \big\{\underline{x} = x_1e_1 + \cdots + x_ne_n \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\big\}.$$

For every $I \in \mathbb{S}$ we will denote by $\mathbb{C}_I$ the plane $\mathbb{R} \oplus I\mathbb{R}$, isomorphic to $\mathbb{C}$, and, if $U \subseteq \mathbb{R}^{n+1}$, by U_I the intersection $U \cap \mathbb{C}_I$. Also, for $R > 0$, we will denote the open ball of $\mathbb{R}^{n+1}$ centered at the origin with radius R by

$$B(0, R) = \{x \in \mathbb{R}^{n+1} : |x| < R\}.$$

We can now recall the definition of slice monogenicity.

Definition 2.1. Let U be a domain in $\mathbb{R}^{n+1}$. A function $f : U \to \mathbb{R}_n$ is called *slice monogenic* if, for all $I \in \mathbb{S}$, its restriction f_I to U_I is *holomorphic*, i.e., it has continuous partial derivatives and satisfies

$$\bar{\partial}_I f(u + vI) := \tfrac{1}{2}\Big(\frac{\partial}{\partial u} + I\frac{\partial}{\partial v}\Big) f_I(u + vI) = 0,$$

for all $u + vI \in U_I$.

For slice monogenic functions, the natural domains of definition are symmetric slice domains.

Definition 2.2. Let U be a domain in $\mathbb{R}^{n+1}$.

(i) U is called a *slice domain* if it intersects the real axis and if for each $I \in \mathbb{S}$, U_I is a domain in $\mathbb{C}_I$.

(ii) U is called an *axially symmetric domain* if for every point $u + vI \in U$, with $u, v \in \mathbb{R}$ and $I \in \mathbb{S}$, the entire sphere $u + v\mathbb{S}$ is contained in U.

A domain in $\mathbb{R}^{n+1}$ is called a *symmetric slice domain* if it is not only a slice domain, but also an axially symmetric domain. By the very definition, an open ball $B(0, R)$ is a typical symmetric slice domain. From now on, we will focus mainly on slice monogenic functions on $B(0, R)$. In most cases, the following results hold, with appropriate changes, for symmetric slice domains more general than open balls of the type $B(0, R)$. For slice monogenic functions a natural definition of derivative is given by the following.

Definition 2.3. Let $f : B(0, R) \to \mathbb{R}_n$ be a slice monogenic function. The *slice derivative* of f is defined to be

$$\partial_I f(u + vI) := \tfrac{1}{2}\Big(\frac{\partial}{\partial u} - I\frac{\partial}{\partial v}\Big) f_I(u + vI).$$

Notice that the operators ∂_I and $\bar{\partial}_I$ commute, and

$$\partial_I f(u + vI) = \frac{\partial}{\partial u} f(u + vI)$$

holds for slice monogenic functions. Therefore, the slice derivative of a slice monogenic function is still slice monogenic so we can iterate the differentiation to obtain the k-th slice derivative,

$$\partial_I^k f(u + vI) = \Big(\frac{\partial}{\partial u}\Big)^k f(u + vI), \quad \forall k \in \mathbb{N}.$$

In what follows, for the sake of simplicity, we will directly denote the k-th slice derivative $\partial_I^k f$ by $f^{(k)}$ for every $k \in \mathbb{N}$.

As shown in [Colombo et al. 2009], a paravector power series $\sum_{k=0}^{\infty} x^k a_k$ with $\{a_k\}_{k\in\mathbb{N}} \subset \mathbb{R}_n$ defines a slice monogenic function in its domain of convergence, which proves to be an open ball $B(0, R)$ with R equal to the radius of convergence of the power series. The converse result is also true.

Theorem 2.4. *A function f is slice monogenic on $B = B(0, R)$ if and only if f has a power series expansion*

$$f(x) = \sum_{k=0}^{\infty} x^k a_k \quad \textit{with} \quad a_k = \frac{f^{(k)}(0)}{k!}.$$

A fundamental result in the theory of slice monogenic functions is described by the splitting lemma, which relates the notion of slice monogenicity to the classical notion of holomorphicity; see [Colombo et al. 2009].

Lemma 2.5. *Let f be a slice monogenic function on $B = B(0, R)$. For each $I_1 = I \in \mathbb{S}$, let $I_2, \dots, I_n$ be a completion to a basis of $\mathbb{R}^n$ satisfying the defining relations $I_i I_j + I_j I_i = -2\delta_{ij}$. Then there exist 2^{n-1} holomorphic functions $F_A : B_I \to \mathbb{C}_I$ such that for every $z = u + vI \in B_I$,*

$$f_I(z) = \sum_{|A|=0}^{n-1} F_A(z) I_A,$$

where $I_0 = 1$ when $r = 0$, or $I_A = I_{i_1} I_{i_2} \cdots I_{i_r}$, with $A = i_1 i_2 \cdots i_r$ a multi-index such that $2 \le i_1 < \cdots < i_r \le n$ when $r > 0$.

The following version of the identity principle is one of the direct consequences of the preceding lemma; see [Colombo et al. 2009].

Theorem 2.6. *Let f be a slice monogenic function on $B = B(0, R)$. Denote by $\mathcal{Z}_f$ the zero set of f,*

$$\mathcal{Z}_f = \{x \in B : f(x) = 0\}.$$

If there exists an $I \in \mathbb{S}$ such that $B_I \cap \mathcal{Z}_f$ has an accumulation point in B_I, then f vanishes identically on B.

Another useful result is Theorem 2.7; see [Colombo and Sabadini 2009].

Theorem 2.7. *Let f be a slice monogenic function on a symmetric slice domain $U \subseteq \mathbb{R}^{n+1}$ and let $I \in \mathbb{S}$. Then for all $u + vJ \in U$ with $J \in \mathbb{S}$,*

$$f(u + vJ) = \tfrac{1}{2}\big(f(u + vI) + f(u - vI)\big) + \tfrac{1}{2} JI\big(f(u - vI) - f(u + vI)\big).$$

In particular, for each sphere of the form $u + v\mathbb{S}$ contained in U, there exist $b, c \in \mathbb{R}_n$ such that $f(u + vI) = b + Ic$ for all $I \in \mathbb{S}$.

Thanks to this result, it is possible to recover the values of a slice monogenic function on symmetric slice domains, which are more general than open balls centered at the origin, from its values on a single slice. This yields an extension theorem that, in the special case of functions that are slice monogenic on $B(0, R)$, can be obtained by means of their power series expansions.

Remark 2.8. Fix an element $I \in \mathbb{S}$ and denote by B_I the intersection $B(0, R) \cap \mathbb{C}_I$ of the open ball $B(0, R)$ with the complex plane $\mathbb{C}_I$. Given a holomorphic function $f_I : B_I \to \mathbb{C}_I$ with the power series expansion taking the form

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

where $\{a_k\}_{k \in \mathbb{N}} \subset \mathbb{C}_I$, the unique slice monogenic extension of f_I to the whole ball $B(0, R)$ is the function given by

$$f(x) := \mathrm{ext}(f_I)(x) = \sum_{k=0}^{\infty} x^k a_k,$$

which takes values in $\mathbb{R}_n$. The uniqueness is guaranteed by the identity principle; that is, Theorem 2.6. In Section 3, we will establish the growth and distortion theorems for such a class of slice monogenic functions that are injective on B_I.

Since slice monogenicity is not preserved under the usual pointwise product of two slice monogenic functions, a new multiplication operation, called the slice monogenic product (or $*$-product), appears via a suitable modification of the usual operation, subject to the noncommutative setting, and plays a key role in the theory of slice monogenic functions. On open balls centered at the origin, the slice monogenic product of two slice monogenic functions is defined by means of their power series expansions; see [Colombo et al. 2010; 2011a].

Definition 2.9. Let $f, g : B = B(0, R) \to \mathbb{R}_n$ be two slice monogenic functions and let

$$f(x) = \sum_{k=0}^{\infty} x^k a_k, \qquad g(x) = \sum_{k=0}^{\infty} x^k b_k$$

be their power series expansions. The *slice monogenic product* (*$*$-product*) of f and g is the function defined by

$$f * g(x) = \sum_{k=0}^{\infty} x^k \Bigl(\sum_{j=0}^{k} a_j b_{k-j}\Bigr),$$

which is slice monogenic on B.

We now recall more definitions (see, e.g., [Colombo et al. 2010; 2011a; Ghiloni and Perotti 2011a; 2011b]).

Definition 2.10. Letting $f(x) = \sum_{k=0}^{\infty} x^k a_k$ be a slice monogenic function on $B = B(0, R)$, we define the *slice monogenic conjugate* of f as

$$f^c(x) = \sum_{k=0}^{\infty} x^k \bar{a}_k,$$

and the *symmetrization* of f as

$$f^s(x) := \sum_{k=0}^{\infty} x^k \mathrm{Sc}\Bigl(\sum_{j=0}^{k} a_j \bar{a}_{k-j}\Bigr). \tag{2-4}$$

Moreover, we define the *normal function* of f as

$$N(f)(x) := f * f^c(x) = \sum_{k=0}^{\infty} x^k \Bigl(\sum_{j=0}^{k} a_j \bar{a}_{k-j}\Bigr). \tag{2-5}$$

These three functions are slice monogenic on B.

Remark 2.11. Here are several useful remarks concerning Definitions 2.9 and 2.10:

(i) The slice monogenic product ($*$-product), the slice monogenic conjugate, and symmetrization can also be defined for slice monogenic functions f on symmetric slice domains U in $\mathbb{R}^{n+1}$ (we refer the interested reader to [Colombo et al. 2010] or [Colombo et al. 2011a, Section 2.6] for details). Moreover, for any two slice monogenic functions $f, g : U \to \mathbb{R}_n$ and each point $x_0 \in \mathbb{R}$, we can define two slice monogenic functions f_{x_0} and g_{x_0} on the symmetric slice domain $U_{x_0} := U - x_0$ by setting

$$f_{x_0}(x) = f(x + x_0), \qquad g_{x_0}(x) = g(x + x_0)$$

for each $x \in U_{x_0}$. Then we have the following identity

$$(f * g)_{x_0} = f_{x_0} * g_{x_0}.$$

This follows from the identity principle together with the fact that when restricted to the real axis, the slice monogenic product is just the usual pointwise one.

(ii) For slice monogenic functions on open balls of type $B := B(0, R)$, the notion of slice monogenic conjugate coincides with the one introduced in Definition 5.4 of [Colombo et al. 2010] (see also Proposition 5.5 therein). Further, the notion of symmetrization given here is equivalent to the one introduced in Definition 5.6. of that paper. To see this, we proceed as follows: For a slice monogenic function $f : B \to \mathbb{R}_n$, we denote by f^{s} the symmetrization of f according to [Colombo et al. 2010, Definition 5.6]. By considering the power series expansion of f^{s}, we may assume that

$$f^{\mathrm{s}}(x) = \sum_{k=0}^{\infty} x^k \alpha_k. \tag{2-6}$$

We also fix an element $I \in \mathbb{S}$. Then according to [Colombo et al. 2010, p.386] or [Colombo et al. 2011a, p.50], for each $x \in B_I$, we have

$$f^{\mathrm{s}}(x) = \mathrm{Sc}\big(f * f^c(x)\big) + \big\langle f * f^c(x), I \big\rangle I.$$

Now substituting (2-5) and (2-6) into the preceding equality, we see that for each $x \in B \cap \mathbb{R}$,

$$\sum_{k=0}^{\infty} x^k \alpha_k = \sum_{k=0}^{\infty} x^k \mathrm{Sc}\Big(\sum_{j=0}^{k} a_j \bar{a}_{k-j}\Big) + \sum_{k=0}^{\infty} x^k \Big\langle \sum_{j=0}^{k} a_j \bar{a}_{k-j}, I \Big\rangle I.$$

For each $k \in \mathbb{N}$, since $\sum_{j=0}^{k} a_j \bar{a}_{k-j}$ is invariant under the Clifford conjugate (see (2-9) below), the second summation on the right-hand side of the preceding equality must vanish identically. Indeed, in view of (2-2),

$$\Big\langle \sum_{j=0}^{k} a_j \bar{a}_{k-j}, I \Big\rangle = \Big\langle \sum_{j=0}^{k} \overline{a_j \bar{a}_{k-j}}, \bar{I} \Big\rangle = -\Big\langle \sum_{j=0}^{k} a_j \bar{a}_{k-j}, I \Big\rangle,$$

which must be zero. Consequently, we deduce that the equality

$$\sum_{k=0}^{\infty} x^k \alpha_k = \sum_{k=0}^{\infty} x^k \mathrm{Sc}\Big(\sum_{j=0}^{k} a_j \bar{a}_{k-j}\Big)$$

holds for all $x \in B \cap \mathbb{R}$. By uniqueness,

$$\alpha_k = \mathrm{Sc}\Big(\sum_{j=0}^{k} a_j \bar{a}_{k-j}\Big), \qquad \forall k \in \mathbb{N}.$$

This shows that f^{s} is the same as f^s defined in (2-4).

(iii) In view of (i), the definition $N(f) := f * f^c$ is also valid for slice monogenic functions f on symmetric slice domains in $\mathbb{R}^{n+1}$.

(iv) The notation $N(f)$ in the definition of normal functions is chosen in accordance with [Ghiloni and Perotti 2011a, Definition 11], which treated the case of slice functions on symmetric open subsets of the so-called *quadratic cone* of a finite-dimensional real alternative *-algebra.

(v) For each slice monogenic function f on a symmetric slice domain $U \subseteq \mathbb{R}^{n+1}$ and each element $I \in \mathbb{S}$, the restriction $N(f)_I$ of $N(f)$ to $U_I := U \cap \mathbb{C}_I$ coincides with the function $f_I * f_I^c : U_I \to \mathbb{R}_n$ considered in [Colombo et al. 2010] or [Colombo et al. 2011a, Section 2.6].

With parts (i) and (iii) of Remark 2.11 in mind, the inverse element of a non-identically vanishing slice monogenic functions with respect to the $*$-product can be defined under a suitable condition.

Definition 2.12. Let f be a slice monogenic function on a symmetric slice domain $U \subseteq \mathbb{R}^{n+1}$ such that

$$N(f)(U_I) \subseteq \mathbb{C}_I$$

for *some* $I \in \mathbb{S}$. If f does not vanish identically, its *slice monogenic inverse* is the function defined by

$$f^{-*}(x) := f^s(x)^{-1} f^c(x),$$

which is slice monogenic on $U \setminus \mathcal{Z}_{f^s}$. Here $\mathcal{Z}_{f^s}$ denotes the zero set of the symmetrization f^s of f.

Remark 2.13. Two useful remarks concerning Definition 2.12 are in order:

(i) For each function f as described in Definition 2.12, the requirement that
$$N(f)(U_I) \subseteq \mathbb{C}_I$$
for some $I \in \mathbb{S}$ guarantees that f_I^s coincides with $N(f)_I = f_I * f_I^c$, see [Colombo et al. 2011a, Definition 2.6.10], although this fact is not explicitly proven in that paper.

(ii) Also we will see, in the proof of Proposition 2.14, that for each function f as described in Definition 2.12, the coefficients which appeared in (2-5) are *real numbers*. This implies that for each such function f, its normal function $N(f)$ is the same as its symmetrization f^s, which is a slice preserving function so that its slice monogenic inverse
$$f^{-*}(x) = f^s(x)^{-1} f^c(x) = \big(N(f)(x)\big)^{-1} f^c(x)$$
is indeed slice monogenic on $U \setminus \mathcal{Z}_{f^s}$. Furthermore, it is well worth noting that in view of [Colombo et al. 2011a, Remark 2.6.8 and Lemma 2.5.12], the zero set $\mathcal{Z}_{f^s}$ of f^s is precisely the union of isolated spheres of the form $u + v\mathbb{S}$ with $u, v \in \mathbb{R}$. This implies that $U \setminus \mathcal{Z}_{f^s}$ is a symmetric slice domain in $\mathbb{R}^{n+1}$.

The function f^{-*} defined in Definition 2.12 deserves the name of slice monogenic inverse of f due to the following:

Proposition 2.14. *Let f be as described in Definition 2.12. Then we have*

$$f|_{U\setminus\mathcal{Z}_{f^s}} * f^{-*} = f^{-*} * f|_{U\setminus\mathcal{Z}_{f^s}} = 1, \tag{2-7}$$

and

$$(f^{-*})^{-*} = f|_{U\setminus\mathcal{Z}_{f^s}}. \tag{2-8}$$

This proposition is quite important in the theory of slice monogenic functions. The equalities in (2-7) first appeared in [Colombo et al. 2010, Proposition 5.9], but the proofs given there and in [Colombo et al. 2011a, Proposition 2.6.11] seem incomplete — the equality $f_I * f_I^c = f_I^c * f_I$ (which is equivalent to $N(f) = N(f^c)$, in view of Remark 2.11 (v) and the identity principle) is used without being proven. A different approach has been used in [Colombo et al. 2011b, Proposition 3.2]. A complete treatment has been given in [Ghiloni et al. 2016, Section 2] in the case of slice functions, which subsumes the case of slice monogenic functions. To keep our presentation self-contained, we provide here a detailed proof of Proposition 2.14.

Proof. We first prove (2-7). To this end, we need the following well known facts:

Fact 1: *For any $a, b \in \mathbb{R}_n$, $ab = 1$ if and only if $ba = 1$.*

Fact 2: *For each $a \in \mathbb{R}_n$, $a\bar{a} = 0$ if and only if $a = 0$.*

Indeed, Fact 1 holds for all finite-dimensional associative algebras (see, e.g., [Drozd and Kirichenko 1994, Theorem 1.2.1]), and Fact 2, which immediately follows from (2-1), is called *nonsingularity* of $\mathbb{R}_n$.

Note that f does not vanish identically on U, and neither does the restriction $f|_{U\cap\mathbb{R}}$ of f to $U\cap\mathbb{R}$, in view of the identity principle. Thus we can find one point $x_0 \in U\cap\mathbb{R}$ and a positive number $R > 0$ such that the open ball $B(x_0, R)$ is contained in U and f is nowhere vanishing on $B(x_0, R)$. Thanks to Remark 2.11(i), we may further assume that $x_0 = 0$ without loss of generality. Now we expand f on $B := B(0, R)$ as

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

Since there exists an element $I \in \mathbb{S}$ such that $N(f) = f * f^c$ maps U_I into $\mathbb{C}_I$ (and also maps B_I into $\mathbb{C}_I$), and

$$\sum_{j=0}^{k} \overline{a_j \bar{a}_{k-j}} = \sum_{j=0}^{k} a_{k-j}\, \bar{a}_j \overset{j\to k-j}{=} \sum_{j=0}^{k} a_j \bar{a}_{k-j}, \tag{2-9}$$

we see that $\sum_{j=0}^{k} a_j \bar{a}_{k-j}$ must be a real number for each $k \in \mathbb{N}$. Therefore, $f * f^c$ is slice preserving and maps $B\cap\mathbb{R}$ into $\mathbb{R}$. We next show that

$$f^c * f = f * f^c. \tag{2-10}$$

We proceed as follows. In view of Definition 2.10,

$$f * f^c|_{B\cap\mathbb{R}} = (f\bar{f})|_{B\cap\mathbb{R}}.$$

Since $f * f^c(B\cap\mathbb{R}) \subseteq \mathbb{R}$, we deduce that the restriction $(f\bar{f})|_{B\cap\mathbb{R}}$ takes values in $\mathbb{R}$ as well. This together with Facts 1 and 2 implies that

$$(f\bar{f})|_{B\cap\mathbb{R}} = (\bar{f}f)|_{B\cap\mathbb{R}}.$$

The right-hand side is no other than the restriction $f^c * f|_{\mathbb{B}\cap\mathbb{R}}$, according to Definitions 2.9 and 2.10. Now we obtain that $f * f^c$ coincides with $f^c * f$ on $B\cap\mathbb{R} \subset U$, and hence on U by the identity principle. Now by using [Colombo et al. 2011a, Proposition 2.6.9], Remark 2.13 (ii) and equality (2-10), we can conclude the proof of equality (2-7) as follows:

$$f^{-*} * f = \frac{1}{f^s}(f^c * f) = \frac{1}{f^s}(f * f^c) = \frac{1}{f^s}N(f) = 1,$$

and

$$f * f^{-*} = f * (\frac{1}{f^s} f^c) = \frac{1}{f^s}(f * f^c) = 1.$$

Now it remains to prove (2-8). In view of the very definition, we first need to show that f^{-*} satisfies the condition given in Definition 2.12. To see this, let I

be an element of $\mathbb{S}$ such that f satisfies the assumption therein. From the above argument, we know that $f^s = N(f)$ is slice preserving. This, together with (2-10) and [Colombo et al. 2011a, Proposition 2.6.9], implies that

$$f^{-*} * (f^{-*})^c = \frac{1}{N(f)}$$

so that f^{-*} satisfies the assumption in Definition 2.12 and hence $(f^{-*})^{-*}$ is well defined on $U \setminus \mathcal{Z}_{f^s}$. Now (2-8) follows from (2-7) and uniqueness of $(f^{-*})^{-*}$. □

3. Growth and distortion theorems for slice monogenic functions

In this section, in the setting of the Clifford algebra $\mathbb{R}_n$, we establish the growth and distortion theorems for slice monogenic extensions to the open unit ball $\mathbb{B} := \{x \in \mathbb{R}^{n+1} : |x| < 1\}$ of univalent functions on the unit disc $\mathbb{D} \subset \mathbb{C}$. We begin with a technical proposition. To present it more generally, we will digress for a moment to slice monogenic functions on general symmetric slice domains.

Proposition 3.1. *Let $U \subseteq \mathbb{R}^{n+1}$ be a symmetric slice domain and $f : U \to \mathbb{R}_n$ a slice monogenic function. Then for every $x = u + vJ \in U$ and every $I \in \mathbb{S}$, there holds the identity*

$$|f(x)|^2 = \frac{1+\langle I, J\rangle}{2}|f(y)|^2 + \frac{1-\langle I, J\rangle}{2}|f(\bar{y})|^2 - \big\langle f(y)\overline{f(\bar{y})}, I \wedge J\big\rangle, \tag{3-1}$$

where $y = u + vI$ and $\bar{y} = u - vI$.

Proof. Fix an arbitrary point $x = u+vJ \in U$ and an element $I \in \mathbb{S}$. Set $y := u+vI$ and $\bar{y} := u - vI$. It follows from Theorem 2.7 that

$$f(x) = \tfrac{1}{2}\big(f(y) + f(\bar{y})\big) - \tfrac{1}{2}JI\big(f(y) - f(\bar{y})\big). \tag{3-2}$$

Notice that, in vector notation,

$$\langle I, J\rangle = \mathrm{Sc}(I\bar{J}) = -\tfrac{1}{2}(IJ + JI), \tag{3-3}$$

and

$$I \wedge J = \tfrac{1}{2}(IJ - JI). \tag{3-4}$$

We shall use the simple identity that

$$|a+b|^2 = |a|^2 + |b|^2 + 2\langle a, b\rangle \tag{3-5}$$

for any $a, b \in \mathbb{R}_n \simeq \mathbb{R}^{2^n}$.

Observe that I and J are 1-vectors and hence are paravectors. In view of (2-3), it holds that

$$|JI(f(y) - f(\bar{y}))| = |f(y) - f(\bar{y})|.$$

Take the modulus on both sides of (3-2) and apply (3-5) to obtain

$$
\begin{aligned}
|f(x)|^2 &= \tfrac{1}{4}\big(|f(y)+f(\bar{y})|^2+|f(y)-f(\bar{y})|^2\big)\\
&\qquad -\tfrac{1}{2}\big\langle f(y)+f(\bar{y}), JI(f(y)-f(\bar{y}))\big\rangle\\
&=: A-\tfrac{1}{2}B.
\end{aligned}
\tag{3-6}
$$

Again applying (3-5), it is evident that

$$
A=\tfrac{1}{2}(|f(y)|^2+|f(\bar{y})|^2).
\tag{3-7}
$$

To calculate the term B, it first follows from the very definition of inner product (see (2-1)) that

$$
B=\big\langle\big(f(y)+f(\bar{y})\big)\big(\,\overline{f(y)}-\overline{f(\bar{y})}\,\big), JI\big\rangle =: B_1+B_2,
\tag{3-8}
$$

where $B_1=\big\langle f(y)\overline{f(y)}-f(\bar{y})\overline{f(\bar{y})}, JI\,\big\rangle$, and $B_2=\big\langle f(\bar{y})\overline{f(y)}-f(y)\overline{f(\bar{y})}, JI\,\big\rangle$.

We next claim that

$$
B_1=-\langle I,J\rangle\big(|f(y)|^2-|f(\bar{y})|^2\big),
\tag{3-9}
$$

and

$$
B_2=2\big\langle f(y)\overline{f(\bar{y})}, I\wedge J\,\big\rangle.
\tag{3-10}
$$

Indeed, applying the fact that $\langle a,b\rangle=\langle\bar{a},\bar{b}\rangle$ from (2-2) to B_1 yields that

$$
B_1=\big\langle f(y)\overline{f(y)}-f(\bar{y})\overline{f(\bar{y})}, IJ\,\big\rangle.
$$

Combining this, (3-3) and the initial notion of B_1, we thus obtain

$$
\begin{aligned}
B_1 &= \tfrac{1}{2}\big\langle f(y)\overline{f(y)}-f(\bar{y})\overline{f(\bar{y})}, IJ+JI\,\big\rangle\\
&= -\big\langle f(y)\overline{f(y)}-f(\bar{y})\overline{f(\bar{y})}, \langle I,J\,\rangle\big\rangle\\
&= -\langle I,J\rangle\big\langle f(y)\overline{f(y)}-f(\bar{y})\overline{f(\bar{y})}, 1\big\rangle\\
&= -\langle I,J\rangle\big(|f(y)|^2-|f(\bar{y})|^2\big).
\end{aligned}
$$

Similarly,

$$
B_2=\Big\langle\,\overline{f(\bar{y})\overline{f(y)}}, \overline{JI}\,\Big\rangle-\big\langle f(y)\overline{f(\bar{y})}, JI\,\big\rangle=2\big\langle\, f(y)\overline{f(\bar{y})}, I\wedge J\,\big\rangle
$$

as desired. In the second equality we have used (3-4). Now substituting (3-7)–(3-10) into (3-6) yields that

$$
|f(x)|^2=\frac{1+\langle I,J\rangle}{2}|f(y)|^2+\frac{1-\langle I,J\rangle}{2}|f(\bar{y})|^2-\big\langle f(y)\overline{f(\bar{y})}, I\wedge J\big\rangle,
$$

which completes the proof. □

Proposition 3.1 shows that when f preserves at least one slice, the squared norm of f can thus be expressed as a convex combination of those in the preserved slice.

Lemma 3.2. *Letting f be a slice monogenic function on a symmetric slice domain $U \subseteq \mathbb{R}^{n+1}$ such that $f(U_I) \subseteq \mathbb{C}_I$ for some $I \in \mathbb{S}$, the convex combination identity*

$$|f(u+vJ)|^2 = \frac{1+\langle I,J\rangle}{2}|f(u+vI)|^2 + \frac{1-\langle I,J\rangle}{2}|f(u-vI)|^2 \tag{3-11}$$

holds for every $u+vJ \in U$.

Proof. As mentioned before, this lemma is a direct consequence of the preceding proposition. But here, we would like to provide an alternative easier approach to it, making no use of Proposition 3.1.

First, we have the following simple fact, which can be easily verified:

Fact: *For any $I, J \in \mathbb{S}$, the set*

$$\{1, I, I \wedge J, I(I \wedge J)\}$$

is an orthogonal set of $\mathbb{R}_n \simeq \mathbb{R}^{2^n}$.

As in the preceding proposition, it follows from Theorem 2.7 that

$$f(x) = \tfrac{1}{2}\big(f(y) + f(\bar{y})\big) - \tfrac{1}{2}JI\big(f(y) - f(\bar{y})\big) \tag{3-12}$$

for every $x = u+vJ \in U$ with $y = u+vI$ and $\bar{y} = u-vI$. We can rewrite (3-12), in terms of the relation that

$$JI = -\langle I,J\rangle + J \wedge I,$$

as

$$\begin{aligned} f(x) = {} & \tfrac{1}{2}\big((1+\langle I,J\rangle)f(y) + (1-\langle I,J\rangle)f(\bar{y})\big) + \tfrac{1}{2}(J\wedge I)(f(\bar{y}) - f(y)) \\ =: {} & \tfrac{1}{2}A + \tfrac{1}{2}(J \wedge I)B. \end{aligned}$$

By assumption $f(U_I) \subseteq \mathbb{C}_I$, we thus have

$$A \in \mathbb{C}_I, \qquad B \in \mathbb{C}_I.$$

From the fact above and equality (2-3), taking the modulus on both sides yields

$$|f(x)|^2 = \tfrac{1}{4}|A|^2 + \tfrac{1}{4}|J \wedge I|^2|B|^2. \tag{3-13}$$

A simple calculation shows that

$$\begin{aligned} |A|^2 = {} & (1+\langle I,J\rangle)^2|f(y)|^2 + (1-\langle I,J\rangle)^2|f(\bar{y})|^2 \\ & + 2\big(1-\langle I,J\rangle^2\big)\langle f(y), f(\bar{y})\rangle \end{aligned} \tag{3-14}$$

and

$$|B|^2 = |f(y)|^2 + |f(\bar{y})|^2 - 2\langle f(y), f(\bar{y})\rangle. \tag{3-15}$$

Notice that

$$|J \wedge I|^2 = 1 - \langle I,J\rangle^2. \tag{3-16}$$

Now inserting (3-14), (3-15) and (3-16) into (3-13) yields

$$|f(x)|^2 = \frac{1+\langle I, J\rangle}{2}|f(y)|^2 + \frac{1-\langle I, J\rangle}{2}|f(\bar{y})|^2,$$

which completes the proof. □

Remark 3.3. The counterpart of the convex combination identity (3-11) from Lemma 3.2 also holds for slice regular functions defined on octonions or more general real alternative algebras under the extra assumption that f preserves at least one slice. This can be verified much as in the proof of Proposition 3.1; see [Wang 2015; Ren et al. 2016] for details.

As a direct consequence of Lemma 3.2, we conclude that the maximum and minimum moduli of f are actually attained on the preserved slice.

Corollary 3.4. *Let f be a slice monogenic function on a symmetric slice domain $U \subseteq \mathbb{R}^{n+1}$ such that $f(U_I) \subseteq \mathbb{C}_I$ for some $I \in \mathbb{S}$. Then for each sphere $u+v\mathbb{S} \subset U$, we have the equalities*:

$$\max_{J\in\mathbb{S}}|f(u+vJ)| = \max\big(|f(u+vI)|, |f(u-vI)|\big),$$
$$\min_{J\in\mathbb{S}}|f(u+vJ)| = \min\big(|f(u+vI)|, |f(u-vI)|\big).$$

We can now state the growth and distortion theorems for slice monogenic functions.

Theorem 3.5 (growth and distortion theorems for paravectors). *Let f be a slice monogenic function on $\mathbb{B}$ such that its restriction f_I to $\mathbb{B}_I$ is injective and such that $f(\mathbb{B}_I) \subseteq \mathbb{C}_I$ for some $I \in \mathbb{S}$. If $f(0) = 0$ and $f'(0) = 1$, then for all $x \in \mathbb{B}$, the following inequalities hold*:

$$\frac{|x|}{(1+|x|)^2} \leq |f(x)| \leq \frac{|x|}{(1-|x|)^2}; \tag{3-17}$$

$$\frac{1-|x|}{(1+|x|)^3} \leq |f'(x)| \leq \frac{1+|x|}{(1-|x|)^3}; \tag{3-18}$$

$$\frac{1-|x|}{1+|x|} \leq \big|xf'(x) * f^{-*}(x)\big| \leq \frac{1+|x|}{1-|x|}. \tag{3-19}$$

Moreover, equality holds for one of these six inequalities at some point $x_0 \in \mathbb{B}\setminus\{0\}$ if and only if f is of the form

$$f(x) = x\big(1-xe^{I\theta}\big)^{-*2}, \quad \forall x \in \mathbb{B},$$

for some $\theta \in \mathbb{R}$.

Proof. Notice that $f_I : \mathbb{B}_I \to \mathbb{C}_I$ is a univalent function by our assumption. Theorem 1.1 with F replaced by f_I implies that the inequalities

$$\frac{|z|}{(1+|z|)^2} \le |f(z)| \le \frac{|z|}{(1-|z|)^2}, \tag{3-20}$$

$$\frac{1-|z|}{(1+|z|)^3} \le |f'(z)| \le \frac{1+|z|}{(1-|z|)^3}, \tag{3-21}$$

$$\frac{1-|z|}{1+|z|} \le \left|\frac{zf'(z)}{f(z)}\right| \le \frac{1+|z|}{1-|z|} \tag{3-22}$$

hold for every $z=u+vI\in\mathbb{B}_I$. On the other hand, it follows from Lemma 3.2 that

$$|f(x)|^2=\frac{1+\langle I,J\rangle}{2}|f(z)|^2+\frac{1-\langle I,J\rangle}{2}|f(\bar{z})|^2$$

holds for every $x=u+vJ\in\mathbb{B}$. Since (3-20) holds for all $z=u+vI$, $\bar{z}=u-vI\in\mathbb{B}_I$, it immediately follows that the inequalities in (3-17) hold for all $x=u+vJ\in\mathbb{B}$, by virtue of the convex combination identity above. Since the condition $f'(\mathbb{B}_I)\subseteq\mathbb{C}_I$ holds trivially, Lemma 3.2 can also be used so that the inequalities in (3-18) can be proved in the same manner.

Now it remains to prove the inequalities in (3-19). To this end, we first need to show that the slice monogenic function $xf'(x)*f^{-*}(x)$ is well-defined on the whole ball $\mathbb{B}$. We proceed as follows. First of all, since $f(0)=0$, by considering the Taylor expansion of f at the origin 0 (see Theorem 2.4) and using the Cauchy–Hadamard formula for the radius of convergence of power series (which is valid in the situation here by following the classical proof and making use of (2-3)), or by Remark 2.8, we can write

$$f(x)=xg(x), \tag{3-23}$$

where g is a slice monogenic function on $\mathbb{B}$. This together with the injectivity of f_I and $f'(0)=1$ implies that g has no zeros on $\mathbb{B}_I$. Moreover, g maps $\mathbb{B}_I$ into $\mathbb{C}_I$, since f does by our assumption. Secondly, again from the assumption that $f(\mathbb{B}_I)\subseteq\mathbb{C}_I$, i.e., all the coefficients of the Taylor expansion of f at the origin belong to the complex plane $\mathbb{C}_I$, it follows that

$$f_I^c(z)=\overline{f_I(\bar{z})},$$

and hence

$$N(f)_I(z)=f_I(z)\overline{f_I(\bar{z})}=z^2g_I(z)\overline{g_I(\bar{z})}=z^2N(g)_I(z). \tag{3-24}$$

This implies that

$$N(f)(\mathbb{B}_I)\subseteq\mathbb{C}_I.$$

Furthermore, since g maps $\mathbb{B}_I$ into $\mathbb{C}_I$ and has no zeros on $\mathbb{B}_I$, we obtain that g_I^s is exactly $g_I\overline{g_I(\bar{\cdot})}$ and is zero free on $\mathbb{B}_I$. Thus it follows from Remark 2.13 (ii) and [Colombo et al. 2011a, Remark 2.6.8 and Lemma 2.5.12] that g^s is zero free on $\mathbb{B}$ as well. This, together with the fact that

$$f^s(x) = x^2 g^s(x), \qquad \forall x \in \mathbb{B}, \tag{3-25}$$

(as obtained easily from (3-23)), implies that 0 is the only zero of f^s. Therefore, according to Definition 2.12, f^{-*} and g^{-*} can be defined on $\mathbb{B} \setminus \{0\}$ and $\mathbb{B}$, respectively. Finally, in view of (3-23),

$$f^c(x) = x g^c(x), \qquad \forall x \in \mathbb{B}, \tag{3-26}$$

from which and (3-25) it follows that the relation

$$x f'(x) * f^{-*}(x) = (f' * g^{-*})(x)$$

holds for all $x \in \mathbb{B} \setminus \{0\}$. Since the right-hand side is well-defined on the whole ball $\mathbb{B}$, the left-hand side can extend regularly to the whole ball $\mathbb{B}$, as desired.

Notice also that $x f'(x) * f^{-*}(x)$ is just the slice monogenic extension to $\mathbb{B}$ of the holomorphic function $z f_I'(z)/f_I(z)$, which also maps the unit disc $\mathbb{B}_I$ into $\mathbb{C}_I$. Now inequalities in (3-19) immediately follow from (3-22) and

$$\left| x f' * f^{-*}(x) \right|^2 = \frac{1+\langle I, J\rangle}{2} \left| \frac{z f'(z)}{f(z)} \right|^2 + \frac{1-\langle I, J\rangle}{2} \left| \frac{\bar{z} f'(\bar{z})}{f(\bar{z})} \right|^2,$$

in view of Lemma 3.2.

Furthermore, if equality holds for one of six inequalities in (3-17), (3-18) and (3-19) at some point $x_0 = u_0 + v_0 J \neq 0$ with $J \in \mathbb{S}$, then the corresponding equality also holds at $z_0 = u_0 + v_0 I$ or $\bar{z}_0 = u_0 - v_0 I$. Then from Theorem 1.1, we obtain

$$f_I(z) = \frac{z}{\left(1 - e^{I\theta} z\right)^2}, \qquad \forall z \in \mathbb{B}_I,$$

for some $\theta \in \mathbb{R}$, which implies

$$f(x) = x\left(1 - x e^{I\theta}\right)^{-*2}, \qquad \forall x \in \mathbb{B}.$$

The converse part is obvious. Now the proof is complete. □

Remark 3.6. The right-hand inequalities in (3-17) and (3-18) can follow alternatively from the well-known but highly nontrivial Bieberbach–de Branges theorem for univalent functions on the open unit disc $\mathbb{D} \subset \mathbb{C}$.

Let $F : \mathbb{D} \to \mathbb{C}$ be a univalent function on the unit disc $\mathbb{D}$ of the complex plane with Taylor expansion

$$F(z) = z + \sum_{m=2}^{\infty} z^m a_m, \qquad a_m \in \mathbb{C}.$$

We consider the canonical imbedding $\mathbb{C} \subset \mathbb{R}^{n+1}$ by expanding the basis $\{1, i\}$ of $\mathbb{C}$ to the basis $\{1, e_1, \dots, e_n\}$ of $\mathbb{R}^{n+1}$ with $e_1 = i$. Therefore we can construct a natural extension of F to $\mathbb{B}$ by setting

$$f(x) = x + \sum_{m=2}^{\infty} x^m a_m, \qquad x \in \mathbb{B}.$$

It is evident that f is a slice monogenic function on the open unit ball $\mathbb{B} = B(0, 1)$ such that its restriction $f|_{\mathbb{D}} = F$ is injective and satisfies that $F(\mathbb{D}) \subseteq \mathbb{C}$. Clearly, $f(0) = 0$ and $f'(0) = 1$. Thus f satisfies all the assumptions of Theorem 3.5 and thus Theorem 1.2 immediately follows.

Remark 3.7. The slice monogenic extension of holomorphic functions on the unit disc $\mathbb{D}$ of the complex plane can result in the theory of slice monogenic elementary functions. We refer to [Colombo et al. 2011a] for the corresponding functional calculus and applications.

The following proposition is of independent interest.

Proposition 3.8. *Let f be a slice monogenic function on a symmetric slice domain $U \subseteq \mathbb{R}^{n+1}$ such that its restriction f_I to U_I is injective and $f(U_I) \subseteq \mathbb{C}_I$ for some $I \in \mathbb{S}$. Then the restriction $f_J : U_J \to \mathbb{R}_n$ is also injective for every $J \in \mathbb{S}$.*

Proof. Suppose that there are two points $x = \alpha + \beta J$ and $y = \gamma + \delta J$ such that $f(x) = f(y)$, then it suffices to prove that $x = y$. If $J = \pm I$, the result follows from the assumption. Otherwise, from Theorem 2.7 one can deduce that

$$f(x) = \tfrac{1}{2}(f(z) + f(\bar{z})) - \tfrac{1}{2} JI(f(z) - f(\bar{z}))$$

and

$$f(y) = \tfrac{1}{2}(f(w) + f(\bar{w})) - \tfrac{1}{2} JI(f(w) - f(\bar{w})).$$

Here $z = \alpha + \beta I$ and $w = \gamma + \delta I$ for the given $I \in \mathbb{S}$. Therefore,

$$\big((f(z) + f(\bar{z})) - (f(w) + f(\bar{w}))\big) - JI\big((f(z) - f(\bar{z})) - (f(w) - f(\bar{w}))\big) = 0.$$

Since $f(U_I) \subseteq \mathbb{C}_I$, 1 and J are linearly independent on $\mathbb{C}_I$ we obtain that

$$f(z) + f(\bar{z}) = f(w) + f(\bar{w})$$

and

$$f(z) - f(\bar{z}) = f(w) - f(\bar{w}),$$

which imply that $f(z) = f(w)$. Thus it follows from the injectivity of f_I that $z = w$ and consequently, $x = y$. □

Remark 3.9. Let f be as described in Theorem 3.5. Then $f_J : \mathbb{B}_J \to \mathbb{R}_n$ is injective for any $J \in \mathbb{S}$ by the preceding proposition. Unfortunately, the authors do not know whether $f : U \to \mathbb{R}_n$ is injective.

4. Growth, distortion and covering theorems for slice regular functions

Let $\mathbb{H}$ denote the noncommutative, associative, real algebra of quaternions with standard basis $\{1, i, j, k\}$, subject to the multiplication rules

$$i^2 = j^2 = k^2 = ijk = -1.$$

Let $\langle\ ,\ \rangle$ denote the standard inner product on $\mathbb{H} \cong \mathbb{R}^4$, i.e.,

$$\langle p, q\rangle = \mathrm{Re}(p\bar{q}) = \sum_{n=0}^{3} x_n y_n$$

for any

$$p = x_0 + x_1 i + x_2 j + x_3 k, \qquad q = y_0 + y_1 i + y_2 j + y_3 k \in \mathbb{H}.$$

In this section, we shall consider slice regular functions defined on domains in quaternions $\mathbb{H}$ with values also in $\mathbb{H}$. These functions are not slice monogenic functions obtained by setting $n = 2$ in the Clifford algebra $\mathbb{R}_n$. Such a class of functions enjoys many nice properties similar to those of classical holomorphic functions of one complex variable. For example, the open mapping theorem holds for slice regular functions on symmetric slice domains in $\mathbb{H}$, but fails for slice monogenic functions even in the quaternionic setting. A simple counterexample is the imbedding map $\imath : \mathbb{R}^3 \hookrightarrow \mathbb{R}_2 \simeq \mathbb{H}$. The open mapping theorem allows us to prove a Koebe type one-quarter theorem (see Theorem 4.10 below). Furthermore, in the quaternionic setting only, we have an explicit formula to express the regular product and regular quotient in terms of the usual pointwise product and quotient. It is exactly this explicit formula which plays a crucial role in many arguments; see the monograph [Gentili et al. 2013] and the recent papers [Ren and Wang 2017; Wang 2015] for more details. In higher dimensions, the formulas to express slice products and slice quotients in terms of the usual pointwise products hold true only under some special cases; see [Ghiloni et al. 2016, Corollary 3.5 and Theorem 3.7] for details. In a certain sense, this phenomenon distinguishes quaternions from other real alternative algebras.

To introduce the theory of slice regular functions, we will denote by $\mathbb{S}$ the unit 2-sphere of purely imaginary quaternions, i.e.,

$$\mathbb{S} = \big\{q \in \mathbb{H} : q^2 = -1\big\}.$$

For every $I \in \mathbb{S}$ we will denote by $\mathbb{C}_I$ the plane $\mathbb{R} \oplus I\mathbb{R}$, isomorphic to $\mathbb{C}$, and, if $\Omega \subseteq \mathbb{H}$, by Ω_I the intersection $\Omega \cap \mathbb{C}_I$. Also, we will denote by $\mathbb{B}$ the open unit ball centered at the origin in $\mathbb{H}$, i.e.,

$$\mathbb{B} = \big\{q \in \mathbb{H} : |q| < 1\big\}.$$

We can now recall the definition of slice regularity.

Definition 4.1. Let Ω be a domain in $\mathbb{H}$. A function $f : \Omega \to \mathbb{H}$ is called *slice regular* if, for all $I \in \mathbb{S}$, its restriction f_I to Ω_I is *holomorphic*, i.e., it has continuous partial derivatives and satisfies

$$\bar{\partial}_I f(x+yI) := \tfrac{1}{2}\Big(\frac{\partial}{\partial x} + I\frac{\partial}{\partial y}\Big) f_I(x+yI) = 0,$$

for all $x + yI \in \Omega_I$.

The notions of slice domain, of symmetric slice domain and of slice derivative are similar to those already given in Section 2. Moreover, the corresponding results still hold for the slice regular functions in the setting of quaternions, such as the splitting lemma, the representation formula, the power series expansion and so on.

Now we can establish the following result by some obvious modifications of the proof of Proposition 3.1.

Proposition 4.2. *Let f be a slice regular function on a symmetric slice domain $\Omega \subseteq \mathbb{H}$. Then for every $q = x + yJ \in \Omega$ and every $I \in \mathbb{S}$, there holds the identity*

$$|f(q)|^2 = \frac{1+\langle I,J\rangle}{2}|f(z)|^2 + \frac{1-\langle I,J\rangle}{2}|f(\bar z)|^2 - \big\langle \mathrm{Im}\big(f(z)\overline{f(\bar z)}\big), I \wedge J\big\rangle, \tag{4-1}$$

where $z = x + yI$ and $\bar z = x - yI$.

Before presenting the key ingredient in establishing the growth and distortion theorems, we first make an equivalent characterization of the vanishing of the third term on the right-hand side of (4-1), thanks to the specialty of quaternions.

Theorem 4.3. *Let f be a slice regular function on a symmetric slice domain $\Omega \subseteq \mathbb{H}$ and let $I \in \mathbb{S}$. Then,*

$$\big\langle \mathrm{Im}\big(f(z)\overline{f(\bar z)}\big), I \wedge J\big\rangle = 0,$$

for all $J \in \mathbb{S}$ and all $z \in \Omega_I$ if and only if there exist $u \in \partial\mathbb{B}$ and a slice regular function g on Ω that preserves the slice Ω_I such that

$$f(q) = g(q)u$$

on Ω.

Proof. We only prove the necessity, since the sufficiency is obvious. Let

$$f_I = F + GK$$

be the splitting of f_I, where $K \in \mathbb{S}$ is perpendicular to I, and $F, G : \Omega_I \to \mathbb{C}_I$ are holomorphic functions. Take $L \in \mathbb{S}$ such that $\{1, I, K, L\}$ is an orthonormal basis of quaternions $\mathbb{H}$ and let V denote the real vector space generated by the set $\{I \wedge J : J \in \mathbb{S}\}$. Then it is clear that

$$V = K\mathbb{R} \oplus L\mathbb{R}. \tag{4-2}$$

Moreover, a simple calculation gives

$$f(z)\overline{f(\bar z)} = \big(F(z)\overline{F(\bar z)} + G(z)\overline{G(\bar z)}\big) + \big(F(\bar z)G(z) - F(z)G(\bar z)\big)K,$$

and from this combined with (4-2) it follows that

$$\left\langle \mathrm{Im}\big(f(z)\overline{f(\bar{z})}\big), I \wedge J \right\rangle = 0, \qquad \forall J \in \mathbb{S},$$

if and only if

$$F(z)G(\bar{z}) = F(\bar{z})G(z), \qquad \forall z \in \Omega_I. \tag{4-3}$$

If $G \equiv 0$ on Ω_I, there is nothing to prove and the desired result follows. Otherwise, $G \not\equiv 0$. Then by the identity principle, its zero set $\mathcal{Z}_G$ has no accumulation points in Ω_I, and neither does

$$\overline{\mathcal{Z}}_G := \{\bar{z} \in \Omega_I : z \in \mathcal{Z}_G\}.$$

Thus, by (4-3),

$$\frac{F(z)}{G(z)} = \frac{F(\bar{z})}{G(\bar{z})}$$

is both holomorphic and antiholomorphic on $\Omega_I \setminus \big(\mathcal{Z}_G \cup \overline{\mathcal{Z}}_G\big)$, which is still a domain of $\mathbb{C}_I$, therefore there exists a constant $\lambda \in \mathbb{C}_I$ such that

$$\frac{F(z)}{G(z)} = \frac{F(\bar{z})}{G(\bar{z})} = \lambda,$$

which implies that $F = \lambda G$ on $\Omega_I \setminus \big(\mathcal{Z}_G \cup \overline{\mathcal{Z}}_G\big)$ and hence on Ω_I by the identity principle.

Now let

$$g := \big(1 + |\lambda|^2\big)^{\frac{1}{2}} \mathrm{ext}(G),$$

and set

$$u := \big(1 + |\lambda|^2\big)^{-\frac{1}{2}} (\lambda + K) \in \partial\mathbb{B}.$$

Then g is a slice regular function on Ω such that $g(\Omega_I) \subseteq \mathbb{C}_I$ and $f = gu$, which completes the proof. □

As a direct consequence, we obtain Corollary 4.4.

Corollary 4.4. *Let I be an element of $\mathbb{S}$ and f a slice regular function on a symmetric slice domain $\Omega \subseteq \mathbb{H}$. Then the convex combination identity*

$$|f(x+yJ)|^2 = \frac{1+\langle I, J\rangle}{2}|f(x+yI)|^2 + \frac{1-\langle I, J\rangle}{2}|f(x-yI)|^2 \tag{4-4}$$

holds for every $x + yJ \in \Omega$ if and only if there exists some $u \in \partial\mathbb{B}$ such that $f(\Omega_I) \subseteq \mathbb{C}_I u$.

In particular, each element f from the slice regular automorphism group of the open unit ball $\mathbb{B}$ of $\mathbb{H}$

$$\mathrm{Aut}(\mathbb{B}) = \big\{f(q) = (1-q\bar{a})^{-*} * (q-a)u : a \in \mathbb{B}, u \in \partial\mathbb{B}\big\}$$

satisfies the condition that there exists some $u \in \partial\mathbb{B}$ such that $f(\Omega_I) \subseteq \mathbb{C}_I u$ so that equality (4-4) holds for such an f.

From Corollary 4.4, we also conclude that the maximum and minimum moduli of every slice regular function on a symmetric slice domain in $\mathbb{H}$ that preserves one slice are actually attained on its preserved slice.

Corollary 4.5. *Let f be a slice regular function on a symmetric slice domain $\Omega \subseteq \mathbb{H}$ such that $f(\Omega_I) \subseteq \mathbb{C}_I$ for some $I \in \mathbb{S}$. Then for each sphere $x + y\mathbb{S} \subset \Omega$, the following equalities hold*:

$$\max_{J\in\mathbb{S}} |f(x+yJ)| = \max\big(|f(x+yI)|, |f(x-yI)|\big), \tag{4-5}$$

$$\min_{J\in\mathbb{S}} |f(x+yJ)| = \min\big(|f(x+yI)|, |f(x-yI)|\big). \tag{4-6}$$

Consequently,

$$\sup_{q\in\Omega} |f(q)| = \sup_{z\in\Omega_I} |f(z)| \tag{4-7}$$

and

$$\inf_{q\in\Omega} |f(q)| = \inf_{z\in\Omega_I} |f(z)|. \tag{4-8}$$

Remark 4.6. Equalities (4-5) and (4-6) were first proved in [Sarfatti 2013, Proposition 1.13] and [de Fabritiis et al. 2015, Proposition 2.6]. Together with the classical growth and distortion theorems, Corollary 4.5 is sufficient to prove Theorem 4.7 even without Corollary 4.4. Despite this trivial fact, Corollary 4.4 is of independent interest and has its own intrinsic value. It presents, additionally, a new convex combination identity (4-4) and provides a sufficient and necessary condition under which (4-4) holds identically. This convex combination identity is also quite useful for other purposes. For instance, it provides an effective approach to a quaternionic version of a well-known Forelli–Rudin estimate, which will play a fundamental role in the theory of various spaces of slice regular functions [Ren and Xu 2016].

Now we state the growth and distortion theorems for slice regular functions.

Theorem 4.7 (growth and distortion theorems for quaternions). *Let f be a slice regular function on $\mathbb{B}$ such that its restriction f_I to $\mathbb{B}_I$ is injective and $f(\mathbb{B}_I) \subseteq \mathbb{C}_I$ for some $I \in \mathbb{S}$. If $f(0) = 0$ and $f'(0) = 1$, then for all $q \in \mathbb{B}$, the following inequalities hold*:

$$\frac{|q|}{(1+|q|)^2} \le |f(q)| \le \frac{|q|}{(1-|q|)^2}; \tag{4-9}$$

$$\frac{1-|q|}{(1+|q|)^3} \le |f'(q)| \le \frac{1+|q|}{(1-|q|)^3}; \tag{4-10}$$

$$\frac{1-|q|}{1+|q|} \le \big|qf'(q) * f^{-*}(q)\big| \le \frac{1+|q|}{1-|q|}. \tag{4-11}$$

Moreover, equality holds for one of these six inequalities at some point $q_0 \in \mathbb{B} \setminus \{0\}$ *if and only if* f *is of the form*

$$f(q) = q\big(1 - qe^{I\theta}\big)^{-*2}, \quad \forall q \in \mathbb{B},$$

for some $\theta \in \mathbb{R}$.

Let $F : \mathbb{D} \to \mathbb{C}$ be a univalent function on the unit disc $\mathbb{D}$ of the complex plane with Taylor expansion

$$F(z) = z + \sum_{n=2}^{\infty} z^n a_n, \qquad a_n \in \mathbb{C}.$$

As in Section 3, with a canonical imbedding $\mathbb{C} \subset \mathbb{H}$, we can construct a natural slice regular extension of F to $\mathbb{B}$ via

$$f(q) = q + \sum_{n=2}^{\infty} q^n a_n, \qquad q \in \mathbb{B}.$$

It is evident that f is a slice regular function on the open unit ball $\mathbb{B} = B(0, 1)$ such that its restriction $f|_{\mathbb{D}} = F$ is injective and satisfies $F(\mathbb{D}) \subseteq \mathbb{C}$. Clearly, $f(0) = 0$ and $f'(0) = 1$. Thus f satisfies all the assumptions of Theorem 4.7 and this results in Theorem 4.8.

Theorem 4.8. *Let* $F : \mathbb{D} \to \mathbb{C}$ *be a univalent function on* $\mathbb{D}$ *such that* $F(0) = 0$ *and* $F'(0) = 1$, *and let* $f : \mathbb{B} \to \mathbb{H}$ *be the slice regular extension of F. Then for all* $q \in \mathbb{B}$, *the following inequalities hold*:

$$\frac{|q|}{(1+|q|)^2} \le |f(q)| \le \frac{|q|}{(1-|q|)^2}; \tag{4-12}$$

$$\frac{1-|q|}{(1+|q|)^3} \le |f'(q)| \le \frac{1+|q|}{(1-|q|)^3}; \tag{4-13}$$

$$\frac{1-|q|}{1+|q|} \le \big|qf'(q) * f^{-*}(q)\big| \le \frac{1+|q|}{1-|q|}. \tag{4-14}$$

Moreover, equality holds for one of these six inequalities at some point $q_0 \in \mathbb{B} \setminus \{0\}$ *if and only if*

$$f(q) = q\big(1 - qe^{i\theta}\big)^{-*2}, \qquad \forall q \in \mathbb{B}.$$

Next we digress to the Koebe one-quarter theorem for slice regular functions on the open unit ball $\mathbb{B} \subset \mathbb{H}$. We recall the following definition (see [Gentili et al. 2013, Definition 7.5]):

Definition 4.9. Let f be a slice regular function on a symmetric slice domain $\Omega \subset \mathbb{H}$. The *degenerate set* of f is defined to be the union D_f of the 2-dimensional spheres $S = x + y\mathbb{S}$ (with $y \neq 0$) such that $f|_S$ is constant.

Now as a direct consequence of the open mapping theorem and the first inequality in (4-9), we have the following result, which is a generalization of [Gal et al. 2015, Theorem 3.11 (1)].

Theorem 4.10 (Koebe one-quarter theorem). *Let f be a slice regular function on $\mathbb{B}$ such that its restriction f_I to $\mathbb{B}_I$ is injective and $f(\mathbb{B}_I) \subseteq \mathbb{C}_I$ for some $I \in \mathbb{S}$. If $f(0) = 0$ and $f'(0) = 1$, then $B\big(0, \frac{1}{4}\big) \subset f(\mathbb{B})$.*

Proof. By assumption, the degenerate set D_f of f is empty. Then f is open by the open mapping theorem (see [Gentili et al. 2013, Theorem 7.7]). This together with the first inequality in (4-9) shows that the image set $f(\mathbb{B})$, containing the origin 0, is an open subset of $\mathbb{H}$, whose boundary $\partial f(\mathbb{B})$ lies outside of the ball $B(0, 1/4)$. Indeed, for each point $w \in \partial f(\mathbb{B})$, there exists a sequence $\{q_n\}_{n=1}^{\infty}$ in $\mathbb{B}$ such that $\lim_{n\to\infty} f(q_n) = w$. By passing to a subsequence, we may assume that the sequence $\{q_n\}_{n=1}^{\infty}$ itself converges to one point, say $q_\infty \in \overline{\mathbb{B}}$. By the openness of f, q_∞ must lie on the boundary $\partial\mathbb{B}$. Thus in view of the first inequality in (4-9),

$$|w| = \lim_{n\to\infty} |f(q_n)| \geq \lim_{n\to\infty} \frac{|q_n|}{(1 + |q_n|)^2} = \frac{1}{4}.$$

Consequently, $f(\mathbb{B})$ must contain the ball $B(0, 1/4)$. This completes the proof. □

Let $\mathcal{SR}(\mathbb{B})$ denote the set of slice regular functions on the open unit ball $\mathbb{B} \subset \mathbb{H}$. We define

$$\mathcal{S} := \big\{ f \in \mathcal{SR}(\mathbb{B}) : \exists\, I \in \mathbb{S} \text{ such that } f_I \text{ is injective and } f_I(\mathbb{B}_I) \subseteq \mathbb{C}_I \big\}$$

and

$$\mathcal{S}_0 := \big\{ f \in \mathcal{S} : f(0) = 0,\ f'(0) = 1 \big\}.$$

For each $f \in \mathcal{S}_0$, we use $r_0(f)$ to denote the radius of the smallest ball $B(0, r)$ contained in $f(\mathbb{B})$. Also for every $\theta \in \mathbb{R}$ and every $I \in \mathbb{S}$, denote by $k_{I,\theta}$ the slice regular function given by

$$k_{I,\theta}(q) = q\big(1 - qe^{I\theta}\big)^{-*2}, \quad \forall q \in \mathbb{B}, \tag{4-15}$$

which obviously belongs to the class $\mathcal{S}_0$. The image set of the unit disc $\mathbb{B}_I$ under $k_{I,\theta}$ is exactly the complex plane except for a radial slit from ∞ to $-e^{I\theta}/4$. This fact together with Theorem 4.10 gives the following result:

Theorem 4.11. *Let the notation be as above.*

(i) *For each $f \in \mathcal{S}_0$,*

$$r_0(f) \geq \tfrac{1}{4},$$

with equality if and only if $f = k_{I,\theta}$ for some $I \in \mathbb{S}$ and some $\theta \in \mathbb{R}$.

(ii) $$\bigcap_{f \in \mathcal{S}_0} f(\mathbb{B}) = B\big(0, \tfrac{1}{4}\big).$$

Proof. We only prove (i). It suffices to consider the extremal case, since the remainder is clear. If $r_0(f) = 1/4$, from the proof of Theorem 4.10 and inequality (4-8), we conclude that there exists some $I_0 \in \mathbb{S}$ such that $1/4$ is exactly the radius of the smallest disc $\mathbb{B}_{I_0}(0, r)$ contained in the image set $f_{I_0}(\mathbb{B}_{I_0})$ of the unit disc $\mathbb{B}_{I_0}$ under the classical univalent function $f_{I_0} : \mathbb{B}_{I_0} \to \mathbb{C}_{I_0}$. This is possible only if $f = k_{I_0,\theta}$ for some $\theta \in \mathbb{R}$ (see the proof of [Graham and Kohr 2003, Theorem 1.1.5] or [Duren 1983, Theorem 2.3]). Now the proof is complete. □

Remark 4.12. Two remarks are in order:

(i) It is noteworthy here that Gal et al. [2015] dealt with the growth, distortion and covering theorems for *slice preserving* and *injective* slice regular functions on the open unit ball $\mathbb{B} \subset \mathbb{H}$ with certain normalized conditions. More precisely, they focused on injective slice functions f on $\mathbb{B}$ of the form

$$f(q) = q + \sum_{n=2}^{\infty} q^n a_n,$$

with $\{a_n\}_{n\geq 2}$ being a sequence of real numbers; see [Gal et al. 2015, Theorem 3.11] for details, while, in the present paper we consider slice regular functions $f(q) = q + \sum_{n=2}^{\infty} q^n a_n$ on $\mathbb{B}$ for which there exists some $I \in \mathbb{S}$ such that the restriction f_I is injective and $\{a_n\}_{n\geq 2}$ is a sequence of numbers in the complex plane $\mathbb{C}_I$ determined by I. Thus our result properly includes the former case. Moreover, our approach to the Koebe type one-quarter theorem (Theorem 4.10), which can be specialized to the complex case, depends only on the open mapping theorem and the first inequality in (4-9), and does not involve compositions of functions. We refer the interested reader to [Graham and Kohr 2003, p.14; Duren 1983, p.31] for a standard proof of the classical Koebe one-quarter theorem for univalent functions.

(ii) Functions $k_{I,\theta}$ of the form in (4-15) are specific examples in $\mathcal{S}_0$. In view of Theorem 4.10, the image of $\mathbb{B}$ under the function $k_{I,\pi/2}$ contains the open ball $B(0, 1/4)$. However, it does not seem so easy to directly deduce this fact from the classical complex result, without using the open mapping theorem and the first inequality in (4-9).

The following proposition is the quaternionic version of Proposition 3.8 for slice regular functions.

Proposition 4.13. *Let f be a slice regular function on a symmetric slice domain $\Omega \subseteq \mathbb{H}$ such that its restriction f_I to Ω_I is injective and $f(\Omega_I) \subseteq \mathbb{C}_I$ for some $I \in \mathbb{S}$. Then its restriction $f_J : \Omega_J \to \mathbb{H}$ is also injective for every $J \in \mathbb{S}$.*

Remark 4.14. Let f be as described in Theorem 4.7. Then according to the preceding proposition, $f_J : \mathbb{B}_J \to \mathbb{H}$ is injective for every $J \in \mathbb{S}$. It is well worth knowing whether $f : \mathbb{B} \to \mathbb{H}$ is injective. If it is indeed the case, together with the first inequality in (4-9) and invariance of domain theorem, it would provide an alternative approach to Theorem 4.10.

5. Concluding remarks

As pointed out in Remark 3.3, the counterpart of the convex combination identity (3-11) in Lemma 3.2 also holds for slice regular functions defined on octonions or more general real alternative algebras under the extra assumption that f preserves at least one slice. Therefore some of the results given in the preceding sections can be easily generalized by slight modification to these new settings. Finally, we conclude with an open question connected with the subject of this paper.

Recall that $\mathcal{SR}(\mathbb{B})$ is the set of slice regular functions on the open unit ball $\mathbb{B} \subset \mathbb{H}$. We denote

$$\mathcal{SR}_0(\mathbb{B}) := \big\{ f \in \mathcal{SR}(\mathbb{B}) : f(0) = 0, f'(0) = 1 \big\}$$

and

$$\mathcal{S}_0 := \big\{ f \in \mathcal{SR}_0(\mathbb{B}) : \exists I \in \mathbb{S} \text{ such that } f_I \text{ is injective and } f_I(\mathbb{B}_I) \subseteq \mathbb{C}_I \big\}.$$

Open question:[1] Is the class $\mathcal{S}_0$ the largest subclass of $\mathcal{SR}_0(\mathbb{B})$ in which the corresponding growth, distortion and covering theorems hold?

Acknowledgment

The authors sincerely thank the referee for a careful reading of the paper and several valuable suggestions and comments, which have significantly improved the presentation of this paper.

References

[Alpay et al. 2016] D. Alpay, F. Colombo, and I. Sabadini, *Slice hyperholomorphic Schur analysis*, Operator Theory, Advances and Applications **256**, Birkhäuser, Basel, 2016. Zbl

[Barnard et al. 1991] R. W. Barnard, C. H. FitzGerald, and S. Gong, "The growth and $1/4$-theorems for starlike mappings in $\mathbb{C}^n$", *Pacific J. Math.* **150**:1 (1991), 13–22. MR Zbl

[Brackx et al. 1982] F. Brackx, R. Delanghe, and F. Sommen, *Clifford analysis*, Research Notes in Mathematics **76**, Pitman, Boston, 1982. MR Zbl

[Cartan 1933] H. Cartan, "Sur la possibilité d'étendre aux fonctions de plusieurs variables complexes la théorie des fonctions univalents", pp. 129–155 in *Leçons sur les fonctions univalents ou mutivalents*, edited by P. Montel, Gauthier-Villars, Paris, 1933. JFM

[Colombo and Sabadini 2009] F. Colombo and I. Sabadini, "A structure formula for slice monogenic functions and some of its consequences", pp. 101–114 in *Hypercomplex analysis*, edited by I. Sabadini et al., Birkhäuser, Basel, 2009. MR Zbl

[1] Very recently, Xu has found a negative answer to this question; see [Xu 2016, Example 3.1, Theorems 5.1 and 5.6] for details.

[Colombo et al. 2004] F. Colombo, I. Sabadini, F. Sommen, and D. C. Struppa, *Analysis of Dirac systems and computational algebra*, Progress in Mathematical Physics **39**, Birkhäuser, Boston, 2004. MR Zbl

[Colombo et al. 2009] F. Colombo, I. Sabadini, and D. C. Struppa, "Slice monogenic functions", *Israel J. Math.* **171** (2009), 385–403. MR Zbl

[Colombo et al. 2010] F. Colombo, I. Sabadini, and D. C. Struppa, "An extension theorem for slice monogenic functions and some of its consequences", *Israel J. Math.* **177** (2010), 369–389. MR Zbl

[Colombo et al. 2011a] F. Colombo, I. Sabadini, and D. C. Struppa, *Noncommutative functional calculus: theory and applications of slice hyperholomorphic functions*, Progress in Mathematics **289**, Birkhäuser, Basel, 2011. MR Zbl

[Colombo et al. 2011b] F. Colombo, I. Sabadini, and D. C. Struppa, "The Runge theorem for slice hyperholomorphic functions", *Proc. Amer. Math. Soc.* **139**:5 (2011), 1787–1803. MR Zbl

[Drozd and Kirichenko 1994] Y. A. Drozd and V. V. Kirichenko, *Finite-dimensional algebras*, Springer, Berlin, 1994. MR Zbl

[Duren 1983] P. L. Duren, *Univalent functions*, Grundlehren Math. Wissenschaften **259**, Springer, New York, 1983. MR Zbl

[de Fabritiis et al. 2015] C. de Fabritiis, G. Gentili, and G. Sarfatti, "Quaternionic Hardy spaces", preprint, 2015. arXiv

[Gal et al. 2015] S. G. Gal, J. O. González-Cervantes, and I. Sabadini, "Univalence results for slice regular functions of a quaternion variable", *Complex Var. Elliptic Equ.* **60**:10 (2015), 1346–1365. MR Zbl

[Gentili and Struppa 2006] G. Gentili and D. C. Struppa, "A new approach to Cullen-regular functions of a quaternionic variable", *C. R. Math. Acad. Sci. Paris* **342**:10 (2006), 741–744. MR Zbl

[Gentili and Struppa 2007] G. Gentili and D. C. Struppa, "A new theory of regular functions of a quaternionic variable", *Adv. Math.* **216**:1 (2007), 279–301. MR Zbl

[Gentili and Struppa 2010] G. Gentili and D. C. Struppa, "Regular functions on the space of Cayley numbers", *Rocky Mountain J. Math.* **40**:1 (2010), 225–241. MR Zbl

[Gentili et al. 2013] G. Gentili, C. Stoppato, and D. C. Struppa, *Regular functions of a quaternionic variable*, Springer, Heidelberg, 2013. MR Zbl

[Gentili et al. 2014] G. Gentili, S. Salamon, and C. Stoppato, "Twistor transforms of quaternionic functions and orthogonal complex structures", *J. Eur. Math. Soc.* **16**:11 (2014), 2323–2353. MR Zbl

[Ghiloni and Perotti 2011a] R. Ghiloni and A. Perotti, "Slice regular functions on real alternative algebras", *Adv. Math.* **226**:2 (2011), 1662–1691. MR Zbl

[Ghiloni and Perotti 2011b] R. Ghiloni and A. Perotti, "Zeros of regular functions of quaternionic and octonionic variable: a division lemma and the camshaft effect", *Ann. Mat. Pura Appl.* (4) **190**:3 (2011), 539–551. MR Zbl

[Ghiloni et al. 2016] R. Ghiloni, A. Perotti, and C. Stoppato, "The algebra of slice functions", 2016. To appear in *Trans. Amer. Math. Soc.*

[Gong 1998] S. Gong, *Convex and starlike mappings in several complex variables*, Mathematics and its Applications (China) **435**, Kluwer, Dordrecht, 1998. MR Zbl

[Graham and Kohr 2003] I. Graham and G. Kohr, *Geometric function theory in one and higher dimensions*, Monographs and Textbooks in Pure and Applied Mathematics **255**, Marcel Dekker, New York, 2003. MR Zbl

[Gürlebeck et al. 2008] K. Gürlebeck, K. Habetha, and W. Sprößig, *Holomorphic functions in the plane and n-dimensional space*, Birkhäuser, Basel, 2008. MR Zbl

[Liu and Ren 1998a] T. Liu and G. Ren, "The growth theorem for starlike mappings on bounded starlike circular domains", *Chinese Ann. Math. Ser. B* **19**:4 (1998), 401–408. MR Zbl

[Liu and Ren 1998b] T. Liu and G. Ren, "Growth theorem of convex mappings on bounded convex circular domains", *Sci. China Ser. A* **41**:2 (1998), 123–130. MR Zbl

[Ren and Wang 2017] G. Ren and X. Wang, "Julia theory for slice regular functions", *Trans. Amer. Math. Soc.* **369**:2 (2017), 861–885. MR Zbl

[Ren and Xu 2016] G. Ren and Z. Xu, "Slice Lebesgue measure of quaternions", *Adv. Appl. Clifford Algebr.* **26**:1 (2016), 399–416. MR Zbl

[Ren et al. 2016] G. Ren, X. Wang, and Z. Xu, "Slice regular functions on regular quadratic cones of real alternative algebras", pp. 227–245 in *Modern trends in hypercomplex analysis*, edited by S. Bernstein et al., Birkhäuser, Basel, 2016.

[Sarfatti 2013] G. Sarfatti, *Elements of function theory in the unit ball of quaternions*, Ph.D. thesis, University of Florence, 2013, https://webusers.imj-prg.fr/~giulia.sarfatti/PhDGiulia.pdf.

[Wang 2015] X. Wang, "On geometric aspects of quaternionic and octonionic slice regular functions", 2015. To appear in *J. Geom. Anal.* arXiv

[Xu 2016] Z. Xu, "Geometric function theory for certain slice regular functions", preprint, 2016. arXiv

Received July 21, 2014. Revised January 13, 2017.

GUANGBIN REN
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
HEFEI, 230026
CHINA
rengb@ustc.edu.cn

XIEPING WANG
SCHOOL OF MATHEMATICAL SCIENCES
UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA
HEFEI, 230026
CHINA
pwx@mail.ustc.edu.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.199

REMARKS ON METAPLECTIC TENSOR PRODUCTS FOR COVERS OF GL_r

SHUICHIRO TAKEDA

We had previously constructed a metaplectic tensor product of automorphic representations of covers of GL_r. To be precise, let $M = \mathrm{GL}_{r_1} \times \cdots \times \mathrm{GL}_{r_k} \subseteq \mathrm{GL}_r$ be a Levi subgroup of GL_r, where $r = r_1 + \cdots + r_k$, and $\widetilde{M}$ its metaplectic preimage in the n-fold metaplectic cover $\widetilde{\mathrm{GL}}_r$ of GL_r. For automorphic representations $\pi_1, \dots, \pi_k$ of $\widetilde{\mathrm{GL}}_{r_1}(\mathbb{A}), \dots, \widetilde{\mathrm{GL}}_{r_k}(\mathbb{A})$, we had constructed (under certain technical assumptions, which are always satisfied when $n = 2$) an automorphic representation π of $\widetilde{M}$ that can be considered as the "tensor product" of the representations $\pi_1, \dots, \pi_k$.

Here we significantly simplify and generalize our previous construction without the technical assumptions mentioned above.

1. Introduction

Let F be a number field and $\mathbb{A}$ be the ring of adeles. For a partition $r = r_1 + \cdots + r_k$ of r, one has the Levi subgroup

$$M(\mathbb{A}) := \mathrm{GL}_{r_1}(\mathbb{A}) \times \cdots \times \mathrm{GL}_{r_k}(\mathbb{A}) \subseteq \mathrm{GL}_r(\mathbb{A})$$

of the $(r_1, \dots, r_k)$-parabolic. Let $\pi_1, \dots, \pi_k$ be automorphic representations of $\mathrm{GL}_{r_1}(\mathbb{A}), \dots, \mathrm{GL}_{r_k}(\mathbb{A})$, respectively. It is a trivial construction to obtain the automorphic representation $\pi_1 \otimes \cdots \otimes \pi_k$ of the Levi $M(\mathbb{A})$ simply by taking the usual tensor product. Though highly trivial, this construction is of great importance in the theory of automorphic forms, especially when one would like to formulate Eisenstein series.

Now if one considers the metaplectic n-fold cover $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ constructed by Kazhdan and Patterson [1984], the analogous construction turns out to be far from trivial. Namely for the metaplectic preimage $\widetilde{M}(\mathbb{A})$ of $M(\mathbb{A})$ in $\mathrm{GL}_r(\mathbb{A})$ and automorphic representations $\pi_1, \dots, \pi_k$ of the metaplectic n-fold covers $\widetilde{\mathrm{GL}}_{r_1}(\mathbb{A}), \dots, \widetilde{\mathrm{GL}}_{r_k}(\mathbb{A})$, one cannot construct a representation of $\widetilde{M}(\mathbb{A})$ simply by taking the tensor product

---

MSC2010: 11F70.
Keywords: automorphic representations, metaplectic groups.

$\pi_1 \otimes \cdots \otimes \pi_k$, because $\widetilde{M}(\mathbb{A})$ is not the direct product of $\widetilde{\mathrm{GL}}_{r_1}(\mathbb{A}), \ldots, \widetilde{\mathrm{GL}}_{r_k}(\mathbb{A})$, namely

$$\widetilde{M}(\mathbb{A}) \ncong \widetilde{\mathrm{GL}}_{r_1}(\mathbb{A}) \times \cdots \times \widetilde{\mathrm{GL}}_{r_k}(\mathbb{A}),$$

and even worse there is no natural map between them.

For the local case, P. Mezo [2004], whose work, we believe, is based on the work by Kable [2001], carried out a construction of an irreducible admissible representation of the Levi $\widetilde{M}$ starting with representations $\pi_1, \ldots, \pi_k$ of $\widetilde{\mathrm{GL}}_{r_1}, \ldots, \widetilde{\mathrm{GL}}_{r_k}$, which can be called the "metaplectic tensor product" of $\pi_1, \ldots, \pi_k$, and characterized it uniquely up to certain character twists.

In [Takeda 2016], we carried out an analogous construction for the global case and defined the global metaplectic tensor product. Further, we showed that the global metaplectic tensor product satisfies various expected properties. We, however, needed to impose certain technical assumptions for the group $\widetilde{M}$, most notably Hypothesis $(*)$ in [Takeda 2016, p. 202]. In this paper, we will modify the construction of that work so that the metaplectic tensor product can be defined without those technical assumptions and show that the new version also satisfies all the expected properties. Indeed, it seems our previous construction was unnecessarily complicated, and here we will give a simpler construction. To be more precise:

Main Theorem. *Let $M = \mathrm{GL}_{r_1} \times \cdots \times \mathrm{GL}_{r_k}$ be a Levi subgroup of GL_r, and let $\pi_1, \ldots, \pi_k$ be automorphic subrepresentations of $\widetilde{\mathrm{GL}}_{r_1}(\mathbb{A}), \ldots, \widetilde{\mathrm{GL}}_{r_k}(\mathbb{A})$. Then there exists an automorphic subrepresentation π of $\widetilde{M}(\mathbb{A})$ such that*

$$\pi \cong \widetilde{\bigotimes_v}' \pi_v,$$

where each π_v is the local metaplectic tensor product of Mezo. Moreover, if $\pi_1, \ldots, \pi_k$ are cuspidal (alternatively, square-integrable modulo center), then so is π. Further the metaplectic tensor product satisfies various expected properties.

In the above theorem, $\widetilde{\bigotimes}'_v$ indicates the metaplectic restricted tensor product, the meaning of which will be explained later in the paper. Also we require π_i be an automorphic subrepresentation, so that it is realized in a subspace of automorphic forms and hence each element in π_i is indeed an automorphic form. (Note that in general an automorphic representation is a subquotient.)

As we will see, strictly speaking the metaplectic tensor product of $\pi_1, \ldots, \pi_k$ might not be unique even up to equivalence but is dependent on a character ω on the center $Z_{\widetilde{\mathrm{GL}}_r}$ of $\widetilde{\mathrm{GL}}_r$. Hence we write

$$\pi_\omega := (\pi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_k)_\omega$$

for the metaplectic tensor product to emphasize the dependence on ω.

Notation. Throughout the paper, F is a number field and $\mathbb{A}$ is the ring of adeles of F. For each place v, F_v is the corresponding local field and $\mathcal{O}_{F_v}$ is the ring of integers of F_v. For each algebraic group G over a global F, and $g \in G(\mathbb{A})$, by g_v we mean the v-th component of g, and so $g_v \in G(F_v)$. For any group G, we denote its center by Z_G.

For a positive integer r, we denote by I_r the $r \times r$ identity matrix. Throughout we fix an integer $n \geq 2$, and we let μ_n be the group of n-th roots of unity in the algebraic closure of the prime field. We always assume that $\mu_n \subseteq F$.

We fix an ordered partition $r_1 + \cdots + r_k = r$ of r, and we let

$$M = \mathrm{GL}_{r_1} \times \cdots \times \mathrm{GL}_{r_k} \subseteq \mathrm{GL}_r$$

and assume it is embedded diagonally as usual.

If π is a representation of a group G, we denote the space of π by V_π, though we often conflate π with V_π when there is no danger of confusion. We say π is unitary if V_π is equipped with a Hermitian structure invariant under the action of G, but we do not necessarily assume that the space V_π is complete. Now assume that the space V_π is a space of functions or maps on the group G and π is the representation of G on V_π defined by right-translation. (This is the case, for example, if π is an automorphic subrepresentation.) Let $H \subseteq G$ be a subgroup. We define $\pi\|_H$ to be the representation of H realized in the space

$$V_{\pi\|_H} := \{f|_H : f \in V_\pi\}$$

of restrictions of $f \in V_\pi$ to H, on which H acts by right translation. Namely $\pi\|_H$ is the representation obtained by restricting the functions in V_π. Occasionally, we confuse $\pi\|_H$ with its space when there is no danger of confusion. Note that there is an H-intertwining surjection $\pi|_H \to \pi\|_H$, where $\pi|_H$ is the (usual) restriction of π to H. Also for any subset $X \subseteq G$ and any $f \in V_\pi$, we denote by $\pi(X)f$ the vector space generated by $\pi(x)f$ for all $x \in X$. If X is a subgroup, this gives rise to a representation of X, which is a subrepresentation of $\pi|_X$.

2. The metaplectic cover $\widetilde{\mathrm{GL}}_r$ of GL_r

The groups. In this subsection, we set up our notations for the metaplectic n-fold cover $\widetilde{\mathrm{GL}}_r$ of GL_r for both local and global cases. Most of the time, we work both locally and globally at the same time. Hence we let

$$R = \begin{cases} F_v & \text{in the local case,} \\ \mathbb{A} & \text{in the global case.} \end{cases}$$

By the metaplectic n-fold cover $\widetilde{\mathrm{GL}}_r(R)$ of $\mathrm{GL}_r(R)$ with a fixed parameter $c \in \{0, \ldots, n-1\}$, we mean the central extension of $\mathrm{GL}_r(R)$ by μ_n as constructed by Kazhdan and Patterson in [1984]. More concretely, as a set,

$$\widetilde{\mathrm{GL}}_r(R) = \mathrm{GL}_r(R) \times \mu_n = \{(g, \xi) : g \in \mathrm{GL}_r(R), \xi \in \mu_n\},$$

whereas the multiplication is defined by

$$(g, \xi) \cdot (g', \xi') = (gg', \tau_r(g, g')\xi\xi'),$$

where τ_r is a certain 2-cocycle. (See [Takeda 2016, Sections 2 and 3] more about various issues on cocycles.)

If P is a parabolic subgroup of GL_r whose Levi part is $M = \mathrm{GL}_{r_1} \times \cdots \times \mathrm{GL}_{r_k}$, we often write

$$\widetilde{M}(R) = \widetilde{\mathrm{GL}}_{r_1}(R) \widetilde{\times} \cdots \widetilde{\times} \widetilde{\mathrm{GL}}_{r_k}(R)$$

for the metaplectic preimage of $M(R)$. Next let

$$\mathrm{GL}_r^{(n)}(R) = \{g \in \mathrm{GL}_r(R) : \det g \in R^{\times n}\},$$

and $\widetilde{\mathrm{GL}}_r^{(n)}(R)$ be its metaplectic preimage. Also we define

$$M^{(n)}(R) = \{(g_1, \ldots, g_k) \in M(R) : \det g_i \in R^{\times n}\}$$

and often denote its preimage by

$$\widetilde{M}^{(n)}(R) = \widetilde{\mathrm{GL}}_{r_1}^{(n)}(R) \widetilde{\times} \cdots \widetilde{\times} \widetilde{\mathrm{GL}}_{r_k}^{(n)}(R).$$

The groups $M^{(n)}(R)$ and $\widetilde{M}^{(n)}(R)$ are normal subgroups of $M(R)$ and $\widetilde{M}(R)$, respectively. Indeed, if we define

$$\mathrm{Det}_M : M(R) = \mathrm{GL}_{r_1}(R) \times \cdots \times \mathrm{GL}_{r_k}(R) \to \underbrace{R^\times \times \cdots \times R^\times}_{k \text{ times}} \tag{2.1}$$

to be the map given by determinant on each factor GL_{r_i}, then $M^{(n)}(R)$ is the kernel of the composition of Det_M with projection to $R^{\times n} \setminus R^\times \times \cdots \times R^{\times n} \setminus R^\times$. Hence for the local case ($R = F_v$), the groups $M^{(n)}(R)$ and $\widetilde{M}^{(n)}(R)$ are of finite index.

An important observation is that the metaplectic preimage of the center $Z_{\mathrm{GL}_r}(R)$ of $\mathrm{GL}_r(R)$ does not in general coincide with the center of $\widetilde{\mathrm{GL}}_r(R)$. (It might not be even commutative for $n > 2$.) The center, which we denote by $Z_{\widetilde{\mathrm{GL}}_r(R)}$, is

$$\begin{aligned} Z_{\widetilde{\mathrm{GL}}_r(R)} &= \{(aI_r, \xi) : a^{r-1+2rc} \in R^{\times n}, \xi \in \mu_n\} \\ &= \{(aI_r, \xi) : a \in R^{\times n/d}, \xi \in \mu_n\}, \end{aligned} \tag{2.2}$$

where $d = \gcd(r - 1 + 2rc, n)$. The second equality is proven in [Chinta and Offen 2013, Lemma 1].

Also the center $Z_{\widetilde{M}(R)}$ of $\widetilde{M}(R)$ is described as

$$Z_{\widetilde{M}(R)} = \left\{ \begin{pmatrix} a_1 I_{r_1} & & \\ & \ddots & \\ & & a_k I_{r_k} \end{pmatrix} : a_i^{r-1+2cr} \in R^{\times n} \text{ and } a_1 \equiv \cdots \equiv a_r \bmod R^{\times n} \right\}.$$

See Proposition 3.10 of [Takeda 2016]. Let us mention that the above descriptions of $Z_{\widetilde{\mathrm{GL}}_r(R)}$ and $Z_{\widetilde{M}(R)}$ give

$$Z_{\widetilde{\mathrm{GL}}_r(R)} \widetilde{M}^{(n)}(R) = Z_{\widetilde{M}(R)} \widetilde{M}^{(n)}(R). \tag{2.3}$$

Let π be a representation of a subgroup $H \subseteq \widetilde{\mathrm{GL}}_r(R)$ containing μ_n. We say π is "genuine" if each element $(1, \xi) \in H$ acts as multiplication by ξ, where we view ξ as an element of $\mathbb{C}$ in the natural way.

We will revisit the question, considered in [Takeda 2016], of how the metaplectic tensor product behaves under restriction to a smaller Levi. Some relevant notation: if

$$I = \{i_1, \ldots, i_l\} \subseteq \{1, \ldots, k\}$$

is a nonempty subset with $i_1 < \cdots < i_l$, we set

$$M_I(R) = \mathrm{GL}_{r_{i_1}}(R) \times \cdots \times \mathrm{GL}_{r_{i_l}}(R) \tag{2.4}$$

which is embedded into $M(R)$ in the obvious way and hence viewed as a subgroup of $M(R)$. Let $\widetilde{M}_I(R)$ be the metaplectic preimage of $M_I(R)$, so we have

$$\widetilde{M}_I(R) \subseteq \widetilde{M}(R).$$

Also set

$$\widetilde{M}_I^{(n)}(R) := \widetilde{M}_I(R) \cap \widetilde{M}^{(n)}(R).$$

The global metaplectic cover $\widetilde{\mathbf{GL}}_r(\mathbb{A})$. In this subsection we only consider the global case, i.e., $R = \mathbb{A}$.

First let us mention that both the F-rational points $\mathrm{GL}_r(F)$ and the unipotent radical $N_B(\mathbb{A})$ of the Borel subgroup B split in $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ via a certain partial map $\mathbf{s} : \mathrm{GL}_r(\mathbb{A}) \to \widetilde{\mathrm{GL}}_r(\mathbb{A})$. Via this splitting we identify $\mathrm{GL}_r(F)$ with a subgroup of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$. Let us mention, however, that this partial map is not given by the map $g \mapsto (g, 1)$ for our choice of cocycle τ_r. But rather the map $g \mapsto (g, 1)$ splits some compact subgroup. For our purpose here, we have only to mention the following. Let S be a finite set of places containing all Archimedean places and those v with $v \mid n$. Then we have a group homomorphism

$$\prod_{v \notin S} \mathrm{GL}_r(\mathcal{O}_{F_v}) \to \widetilde{\mathrm{GL}}_r(\mathbb{A}) \tag{2.5}$$

under the map $g \mapsto (g, 1)$.

We can also describe $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ as a quotient of a restricted direct product of the groups $\widetilde{\mathrm{GL}}_r(F_v)$ as follows. Consider the restricted direct product $\prod'_v \widetilde{\mathrm{GL}}_r(F_v)$ with respect to the groups K_v for all v with $v \nmid n$ and $v \nmid \infty$. If we denote each element in this restricted direct product by $\Pi'_v(g_v, \xi_v)$ so that $g_v \in K_v$ and $\xi_v = 1$ for almost all v, we have the surjection

$$\rho : \prod_v{}' \widetilde{\mathrm{GL}}_r(F_v) \to \widetilde{\mathrm{GL}}_r(\mathbb{A}), \quad \Pi'_v(g_v, \xi_v) \mapsto (\Pi'_v g_v, \Pi_v \xi_v), \tag{2.6}$$

where the product $\Pi_v \xi_v$ is literally the product inside μ_n. This is indeed a group homomorphism and

$$\prod_v{}' \widetilde{\mathrm{GL}}_r(F_v)/\ker\rho \cong \widetilde{\mathrm{GL}}_r(\mathbb{A}),$$

where $\ker\rho$ consists of the elements of the form $(1, \xi)$ with $\xi \in \prod'_v \mu_n$ and $\Pi_v \xi_v = 1$. We set

$$\widetilde{\prod_v}{}' \widetilde{\mathrm{GL}}_r(F_v) := \prod_v{}' \widetilde{\mathrm{GL}}_r(F_v)/\ker\rho$$

and call it the metaplectic restricted direct product. Let us note that each $\widetilde{\mathrm{GL}}_r(F_v)$ has a natural embedding into $\prod'_v \widetilde{\mathrm{GL}}_r(F_v)$. By composing it with ρ, we have the natural inclusion

$$\widetilde{\mathrm{GL}}_r(F_v) \hookrightarrow \widetilde{\mathrm{GL}}_r(\mathbb{A}), \tag{2.7}$$

which allows us to view $\widetilde{\mathrm{GL}}_r(F_v)$ as a subgroup of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$.

Let us mention that all the discussions above on $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ can be generalized to $\widetilde{M}(\mathbb{A})$, though there is a subtle issue on cocycles for $\widetilde{M}(\mathbb{A})$, which is discussed in detail in [Takeda 2016, Sect. 3]. This issue will not play any role in this paper.

We have the notion of automorphic representations as well as automorphic forms on $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ or $\widetilde{M}(\mathbb{A})$. In this paper, by an automorphic form, we mean a smooth automorphic form instead of a K-finite one, namely an automorphic form is K_f-finite, $\mathcal{Z}$-finite and of uniformly moderate growth; see [Cogdell 2004, p. 17]. Hence if π is an automorphic representation of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ (or $\widetilde{M}(\mathbb{A})$), the full group $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ (or $\widetilde{M}(\mathbb{A})$) acts on π. An automorphic form f on $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ (or $\widetilde{M}(\mathbb{A})$) is said to be genuine if $f(g, \xi) = \xi f(g, 1)$ for all $(g, \xi) \in \widetilde{\mathrm{GL}}_r(\mathbb{A})$ (or $\widetilde{M}(\mathbb{A})$). In particular every automorphic form in the space of a genuine automorphic representation is genuine. We denote the space of genuine automorphic forms on $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ (resp. $\widetilde{M}(\mathbb{A})$) by $\mathcal{A}(\widetilde{\mathrm{GL}}_r)$ (resp. $\mathcal{A}(\widetilde{M})$).

Suppose we are given a collection of irreducible admissible representations π_v of $\widetilde{\mathrm{GL}}_r(F_v)$ such that π_v is K_v-spherical for almost all v. Then we can form an irreducible admissible representation of $\prod'_v \widetilde{\mathrm{GL}}_r(F_v)$ by taking a restricted tensor product $\bigotimes'_v \pi_v$ as usual. Suppose further that $\ker\rho$ acts trivially on $\bigotimes'_v \pi_v$, which is always the case if each π_v is genuine. Then it descends to an irreducible admissible representation of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$, which we denote by $\widetilde{\bigotimes}'_v\pi_v$, and call it the "metaplectic restricted tensor product". Let us emphasize that the space for $\widetilde{\bigotimes}'_v\pi_v$ is the same as that for $\bigotimes'_v \pi_v$. Conversely, if π is an irreducible genuine admissible representation of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$, it is written as $\widetilde{\bigotimes}'_v\pi_v$ where π_v is an irreducible genuine admissible representation of $\widetilde{\mathrm{GL}}_r(F_v)$, and for almost all v, π_v is K_v-spherical. (To see this, view π as a representation of the restricted direct product $\prod'_v \widetilde{\mathrm{GL}}_r(F_v)$ by pulling it back by ρ in (2.6) and apply the usual tensor product theorem for the restricted direct product. This gives the restricted tensor product $\bigotimes'_v \pi_v$, where each π_v is genuine, and hence it descends to $\widetilde{\bigotimes}'_v\pi_v$.)

We now list some important properties of various groups we consider.

Lemma 2.8 [Takeda 2016, Lemma 14]. *Let S be a finite set of places containing all the Archimedean ones, and set*

$$\mathcal{O}_S^\times := \prod_{v\notin S} \mathcal{O}_{F_v}^\times .$$

Then the set $F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times/\mathcal{O}_S^\times$ is finite.

Lemma 2.9. *The group $F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times$ is compact.*

Proof. Let S be any finite set of places containing all the Archimedean ones. By the above lemma, we know $F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times$ is a finite union of sets of the form $F^\times\mathbb{A}^{\times n}a\mathcal{O}_S^\times$ for $a\in\mathbb{A}^\times$. But this set, which is the image of the compact set $a\mathcal{O}_S^\times$ under the quotient map $\mathbb{A}^\times\to F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times$, is compact in the topology of $F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times$. Hence the lemma follows. ☐

This in turn implies:

Lemma 2.10. *The group $M(F)\widetilde{M}^{(n)}(\mathbb{A})$ is a closed normal subgroup of $\widetilde{M}^{(n)}(\mathbb{A})$ whose quotient $M(F)\widetilde{M}^{(n)}(\mathbb{A})\setminus\widetilde{M}(\mathbb{A})$ is a compact abelian group. Indeed, we have an isomorphism*

$$M(F)\widetilde{M}^{(n)}(\mathbb{A})\setminus\widetilde{M}(\mathbb{A})\cong\underbrace{F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times\times\cdots\times F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times}_{k\ \textit{times}}$$

of topological groups.

Proof. That it is closed is [Takeda 2016, Proposition A.4]. To show it is normal, one can check that the group $M(F)\widetilde{M}^{(n)}(\mathbb{A})$ is indeed the kernel of the composite

$$\widetilde{M}(\mathbb{A})\to\mathbb{A}^\times\times\cdots\times\mathbb{A}^\times\to F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times\times\cdots\times F^\times\mathbb{A}^{\times n}\setminus\mathbb{A}^\times$$

where the first map is the determinant map Det_M as in (2.1). By the previous lemma, the last group on the right-hand side is compact. ☐

Lemma 2.11. *Let S be a finite set of places containing all the Archimedean ones and those v with $v \mid n$. Define*

$$K^S := \prod_{v \notin S} M(\mathcal{O}_{F_v}), \tag{2.12}$$

which can be viewed as a subgroup of $\widetilde{M}(\mathbb{A})$ as in (2.5). *Then the set*

$$M(F)\widetilde{M}^{(n)}(\mathbb{A}) \setminus \widetilde{M}(\mathbb{A})/K^S$$

is finite.

Proof. This is immediate from Lemma 2.8, because $M(F)\widetilde{M}^{(n)}(\mathbb{A}) \setminus \widetilde{M}(\mathbb{A})/K^S$ is a product of k copies of $F^\times \mathbb{A}^{\times n} \setminus \mathbb{A}^\times / \mathcal{O}_S^\times$ □

We next state a lemma from general topology and an important consequence of it.

Lemma 2.13. *Let A be a Hausdorff compact abelian group, and $m_1, \dots, m_k$ be positive integers. Define*

$$H := \{(a^{m_1}, \dots, a^{m_k}) : a \in A\} = A^{m_1} \times \cdots \times A^{m_k} \subseteq \underbrace{A \times \cdots \times A}_{k \text{ times}}.$$

Then H is a closed subgroup of $A \times \cdots \times A$.

Proof. Note that for each $i \in \{1, \dots, k\}$, the m_i-th power map $A \to A^{m_i} \subseteq A$ is continuous, and hence the image A^{m_i} of the compact A is compact. Recall that in a Hausdorff topological group, every compact subgroup is closed by, say, [Deitmar and Echterhoff 2009, Lemma 1.1.4]. So each A^{m_i} is closed. Hence H is closed. □

Proposition 2.14. *We have*

$$M(F)Z_{\widetilde{M}(\mathbb{A})}\widetilde{M}^{(n)}(\mathbb{A}) = M(F)Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\widetilde{M}^{(n)}(\mathbb{A})$$

and this group is a closed (*hence locally compact*) *subgroup of $\widetilde{M}(\mathbb{A})$.*

Proof. The equality is immediate from (2.3).

To prove this group is closed, it suffices to show that the image of $Z_{\widetilde{M}(\mathbb{A})}$ in the quotient $M(F)\widetilde{M}^{(n)}(\mathbb{A}) \setminus \widetilde{M}(\mathbb{A})$ is closed. But one can see that the image of $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$ under the isomorphism

$$M(F)\widetilde{M}^{(n)}(\mathbb{A}) \setminus \widetilde{M}(\mathbb{A}) = F^\times \mathbb{A}^{\times n} \setminus \mathbb{A}^\times \times \cdots \times F^\times \mathbb{A}^{\times n} \setminus \mathbb{A}^\times$$

is the subgroup of the form

$$\{(a^{nr_1/d}, \dots, a^{nr_k/d}) : a \in F^\times \mathbb{A}^{\times n} \setminus \mathbb{A}^\times\},$$

where $d = \gcd(r - 1 + 2rc, n)$. By Lemma 2.9, we know that $F^\times \mathbb{A}^{\times n} \setminus \mathbb{A}^\times$ is compact, and hence by the previous lemma, this is closed. □

3. The metaplectic tensor product

In this section, after reviewing the local metaplectic tensor product of Mezo [2004] with the modification made by the author in [Takeda 2016], we will construct the global metaplectic tensor product.

Mezo's local metaplectic tensor product. In this subsection, all the groups are over the local field F_v, and accordingly we simply write $\widetilde{\mathrm{GL}}_r$, $\widetilde{M}$, etc, instead of $\widetilde{\mathrm{GL}}_r(F_v)$, $\widetilde{M}(F_v)$, etc.

Let $\pi_1, \dots, \pi_k$ be irreducible admissible genuine representations of $\widetilde{\mathrm{GL}}_{r_1}, \dots, \widetilde{\mathrm{GL}}_{r_k}$, respectively. For each $i = 1, \dots, k$, let

$$\sigma_i := \pi_i|_{\widetilde{\mathrm{GL}}_{r_i}^{(n)}}.$$

Note that σ_i, as a representation of $\widetilde{\mathrm{GL}}_{r_i}^{(n)}$, is completely reducible, and the multiplicities of all the irreducible constituents are all equal. Namely, we have

$$\sigma_i = m_i \bigoplus_j \tau_{i,j}, \tag{3.1}$$

where $\tau_{i,j}$ is an irreducible representation of $\widetilde{\mathrm{GL}}_{r_i}^{(n)}$ such that $\tau_{i,j} \not\cong \tau_{i,k}$ for $j \neq k$, and m_i is a positive multiplicity which is independent of $\tau_{i,j}$. For the non-Archimedean case, this is precisely [Gelbart and Knapp 1982, Lemma 2.1], and the Archimedean case can be proven in the same way as the non-Archimedean case because the index of $\widetilde{\mathrm{GL}}_{r_i}^{(n)}$ in $\widetilde{\mathrm{GL}}_{r_i}$ is at most 2. Mezo [2004] first picks up an irreducible constituent τ_i of σ_i and considers the (usual) tensor product

$$V_{\tau_1} \otimes \cdots \otimes V_{\tau_k},$$

which, of course, gives a representation of the direct product $\widetilde{\mathrm{GL}}_{r_1}^{(n)} \times \cdots \times \widetilde{\mathrm{GL}}_{r_k}^{(n)}$. The genuineness of the representations $\tau_1, \dots, \tau_k$ implies that this tensor product representation descends to a representation of the group $\widetilde{M}^{(n)} := \widetilde{\mathrm{GL}}_{r_1}^{(n)} \widetilde{\times} \cdots \widetilde{\times} \widetilde{\mathrm{GL}}_{r_k}^{(n)}$, i.e., the representation factors through the natural surjection

$$\widetilde{\mathrm{GL}}_{r_1}^{(n)} \times \cdots \times \widetilde{\mathrm{GL}}_{r_k}^{(n)} \twoheadrightarrow \widetilde{\mathrm{GL}}_{r_1}^{(n)} \widetilde{\times} \cdots \widetilde{\times} \widetilde{\mathrm{GL}}_{r_k}^{(n)}.$$

We denote this representation of $\widetilde{M}^{(n)}$ by

$$\tau := \tau_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \tau_k.$$

Let us emphasize that the space V_τ of τ is the usual tensor product $V_{\tau_1} \otimes \cdots \otimes V_{\tau_k}$.

In this paper, however, we will take a different approach. Instead of picking up a τ_i, we will consider all of σ_i at the same time and define the representation

$$\sigma := \sigma_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \sigma_k$$

of $\widetilde{M}^{(n)}$ in the same way as τ. Let us again emphasize that the space V_σ of σ is the usual tensor product $V_{\sigma_1} \otimes \cdots \otimes V_{\sigma_k}$. Further note that because of (3.1), we have

$$\sigma = m \bigoplus_{\tau} \tau, \tag{3.2}$$

where $m = m_1 \cdots m_k$ and the sum is over all possible equivalence classes of representations of the form $\tau = \tau_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \tau_k$.

Then we define

$$\Pi = \Pi(\pi_1, \ldots, \pi_k) := \operatorname{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}} \sigma.$$

By (3.2), we have

$$\Pi = m \bigoplus_{\tau} \operatorname{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}} \tau,$$

where the sum is over all the equivalence classes of irreducible subrepresentations τ of σ. Note that since σ is completely reducible and the index of $\widetilde{M}^{(n)}$ in $\widetilde{M}$ is finite, one can see that the representation Π is completely reducible. Certainly, it is highly unlikely that Π is irreducible. Rather it contains all the metaplectic tensor products constructed by Mezo. To see it, we need to take the action of the center $Z_{\widetilde{\mathrm{GL}}_r}$ into account. For this purpose, let us first define

$$\Omega = \Omega(\pi_1, \ldots, \pi_k) := \{\omega : \omega \text{ is a character on } Z_{\widetilde{\mathrm{GL}}_r} \text{ which appears in } \sigma\}, \tag{3.3}$$

where we say "ω appears in σ" if there is an irreducible constituent $\tau \subseteq \sigma$ that agrees with ω on the overlap, namely

$$\omega|_{Z_{\widetilde{\mathrm{GL}}_r} \cap \widetilde{M}^{(n)}} = \tau|_{Z_{\widetilde{\mathrm{GL}}_r} \cap \widetilde{M}^{(n)}}.$$

Now Mezo's construction can be summarized as follows. Let $\tau \subseteq \sigma$ be an irreducible representation, so $\tau = \tau_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \tau_k$ for some τ_i, and let $\omega \in \Omega$ be such that it agrees with τ on the overlap. Then we can extend τ to the representation

$$\tau_\omega := \omega\tau$$

of $Z_{\widetilde{\mathrm{GL}}_r} \widetilde{M}^{(n)}$ by letting $Z_{\widetilde{\mathrm{GL}}_r}$ act by ω. Then if we induce it to the group $\widetilde{M}$, it is isotypic (though possibly reducible), and we denote this isomorphism class by

$$(\pi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_k)_\omega$$

and call it the metaplectic tensor product of $\pi_1, \ldots, \pi_k$ with respect to ω. With this notation, we have

$$\operatorname{Ind}_{Z_{\widetilde{\mathrm{GL}}_r} \widetilde{M}^{(n)}}^{\widetilde{M}} \tau_\omega = m'(\pi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_k)_\omega$$

for some finite multiplicity m', which will be seen to be independent of τ and ω but only dependent on the representations $\pi_1, \ldots, \pi_k$.

Clearly we have the inclusions

$$\text{(3.4)}\quad \mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}^{\widetilde{M}}\tau_\omega \hookrightarrow \mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}}\tau \hookrightarrow \mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}}\sigma = \Pi(\pi_1,\dots,\pi_k) = m\bigoplus_{\tau}\mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}}\tau,$$

because $\tau \subseteq \sigma$. Further we have:

Proposition 3.5. *For each fixed τ, let*

$$\Omega(\tau) := \{\omega \in \Omega : \omega|_{Z_{\widetilde{\mathrm{GL}}_r}\cap\widetilde{M}^{(n)}} = \tau|_{Z_{\widetilde{\mathrm{GL}}_r}\cap\widetilde{M}^{(n)}}\}.$$

Then

$$\mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}}\tau = \bigoplus_{\omega\in\Omega(\tau)}\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}^{\widetilde{M}}\tau_\omega = m'\bigoplus_{\omega\in\Omega(\tau)}(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega,$$

where m' is the positive multiplicity of $(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega$ in $\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}^{\widetilde{M}}\tau_\omega$, which is independent of τ and ω but is only dependent on $\pi_1,\dots,\pi_k$.

Proof. The proof is an elementary exercise in representation theory. But we will give a brief explanation for each equality. First, by inducing in stages, we have

$$\mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}}\tau = \mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}^{\widetilde{M}}\mathrm{Ind}_{\widetilde{M}^{(n)}}^{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}\tau.$$

Then similarly to [Mezo 2004, Lemma 4.1], one can see

$$\mathrm{Ind}_{\widetilde{M}^{(n)}}^{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}\tau = \bigoplus_{\omega\in\Omega(\tau)}\tau_\omega,$$

because the quotient $\widetilde{M}^{(n)}\setminus Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)} = Z_{\widetilde{\mathrm{GL}}_r}\cap\widetilde{M}^{(n)}\setminus Z_{\widetilde{\mathrm{GL}}_r}$ has the same size as $\Omega(\tau)$. To be more precise, for a fixed $\omega\in\Omega(\tau)$ we can write

$$\Omega(\tau) = \{\omega\chi : \chi \text{ is in the dual of } Z_{\widetilde{\mathrm{GL}}_r}\cap\widetilde{M}^{(n)}\setminus Z_{\widetilde{\mathrm{GL}}_r}\}.$$

To show the next equality, the only nontrivial part is to show that the multiplicity m' is independent of τ and ω. The independence from ω follows from the fact that the restrictions $(\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}^{\widetilde{M}}\tau_\omega)|_{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}$ and $(\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}^{\widetilde{M}}\tau_\omega)|_{\widetilde{M}^{(n)}}$ have the same number of constituents and the latter is independent of ω, and further the restrictions $(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega|_{Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}}$ and $(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega|_{\widetilde{M}^{(n)}}$ have the same number of constituents and again the latter is independent of ω. To show it is independent of τ, let us note that m' is indeed equal to $[\widetilde{H} : Z_{\widetilde{\mathrm{GL}}_r}\widetilde{M}^{(n)}]$, where $\widetilde{H}$ is a maximal subgroup of $\widetilde{M}$ such that τ_ω can be extended to $\widetilde{H}$ so that Mackey's irreducible criterion is satisfied as constructed in [Mezo 2004, pp. 89–90]. This can be proven in the same way as in [Takeda 2016, Proposition 4.7]. From the construction of $\widetilde{H}$, one can see that $\widetilde{H}$ is independent of the choice of $\tau_1,\dots,\tau_k$ but only dependent on $\pi_1,\dots,\pi_k$. Also see [Cai 2016, Section 3.4] for this issue. □

The main theorem for local metaplectic tensor product follows easily:

Theorem 3.6. *Keeping the above notation, we have*

$$\Pi = \Pi(\pi_1, \dots, \pi_k) = \bigoplus_{\omega\in\Omega} m(\omega)(\pi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_k)_\omega,$$

where Ω is as in (3.3) *and $m(\omega)$ is the positive multiplicity of $(\pi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_k)_\omega$.*

Proof. By the previous proposition, we have

$$\Pi = mm' \bigoplus_{\tau} \bigoplus_{\omega\in\Omega(\tau)} (\pi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_k)_\omega,$$

which implies the theorem. □

Remark 3.7. In the above theorem, one may wonder if $m(\omega) = mm'$. This is certainly the case if the $\Omega(\tau)$ are all distinct for distinct τ. But it may be the case that $\Omega(\tau) \cap \Omega(\tau') \neq \varnothing$ even when $\tau \neq \tau'$. Also it should be mentioned that if $\Omega(\tau) \cap \Omega(\tau') \neq \varnothing$, then necessarily $\Omega(\tau) = \Omega(\tau')$

With this theorem, one can tell that the presentation Π contains all the metaplectic tensor products, and one can call each irreducible constituent of Π a metaplectic tensor product.

Next we consider the behavior of metaplectic tensor products upon restriction to a smaller Levi. Let I, $\widetilde{M}_I$, etc., be as on page 203.

Proposition 3.8. *Let $\pi \subseteq \Pi(\pi_1, \dots, \pi_k)$ be a metaplectic tensor product. Then the restriction $\pi|_{\widetilde{M}_I}$ is completely reducible* (*with most likely infinite multiplicity*). *Further each constituent of $\tau|_{\widetilde{M}_I}$ is of the form $(\pi_{i_1} \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_{i_l})_{\omega'}$ for some $\omega' \in \Omega(\pi_{i_1}, \dots, \pi_{i_l})$.*

Proof. Note that $\pi \hookrightarrow \mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}} \tau$ for some irreducible representation τ of $\widetilde{M}^{(n)}$. Hence it suffices to show the restriction $\big(\mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}} \tau\big)|_{\widetilde{M}_I}$ is completely irreducible. But since the group $\widetilde{M}^{(n)} \setminus \widetilde{M}$ is finite, one has the following Mackey type theorem:

$$\big(\mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}} \tau\big)|_{\widetilde{M}_I} = \bigoplus_{g\in \widetilde{M}^{(n)}\setminus \widetilde{M}/\widetilde{M}_I} \mathrm{Ind}_{\widetilde{M}_I \cap g\widetilde{M}^{(n)}g^{-1}}^{\widetilde{M}_I} (\tau^g),$$

where, as usual, τ^g is the representation of τ twisted by g viewed as a representation of $\widetilde{M}_I \cap g\widetilde{M}^{(n)}g^{-1}$ by restriction. But note that $\widetilde{M}^{(n)} \setminus \widetilde{M}/\widetilde{M}_I = M^{(n)} \setminus M/M_I$ and each element in this double coset is represented by an element in M which has the identity on all the components for the GL_{r_i} factors for $i \in I$. Hence $\widetilde{M}_I \cap g\widetilde{M}^{(n)}g^{-1} = \widetilde{M}_I^{(n)}$, and $\tau^g = \tau$ as a representation of $\widetilde{M}_I^{(n)}$. Hence we have

$$\big(\mathrm{Ind}_{\widetilde{M}^{(n)}}^{\widetilde{M}} \tau\big)|_{\widetilde{M}_I} = \bigoplus_{g\in \widetilde{M}^{(n)}\setminus \widetilde{M}/\widetilde{M}_I} \mathrm{Ind}_{\widetilde{M}_I^{(n)}}^{\widetilde{M}_I} \big(\tau|_{\widetilde{M}_I^{(n)}}\big).$$

But note that the space V_τ of τ is of the form $V_{\tau_1} \otimes \cdots \otimes V_{\tau_k}$ for some irreducible representations $\tau_1, \dots, \tau_k$ of $\widetilde{\mathrm{GL}}_{r_1}^{(n)}, \dots, \widetilde{\mathrm{GL}}_{r_k}^{(n)}$, which are irreducible constituents of the restrictions $\pi_1|_{\widetilde{\mathrm{GL}}_{r_1}^{(n)}}, \dots, \pi_k|_{\widetilde{\mathrm{GL}}_{r_k}^{(n)}}$. Hence when it is restricted to $\widetilde{M}_I^{(n)}$, it is completely reducible. (Yet note that the multiplicity is infinite unless all the σ_i for $i \notin I$ are one dimensional.) Indeed $\tau|_{\widetilde{M}_I^{(n)}}$ is isotypic with all the irreducible constituents equivalent to $\tau_{i_1}\widetilde{\otimes}\cdots\widetilde{\otimes}\tau_{i_l}$. Hence first of all, $\pi|_{\widetilde{M}_I}$ is completely reducible. Second of all, each irreducible constituent in $\pi|_{\widetilde{M}_I}$ is contained in the induced representation

$$\mathrm{Ind}_{\widetilde{M}_I^{(n)}}^{\widetilde{M}}\ \tau_{i_1}\widetilde{\otimes}\cdots\widetilde{\otimes}\tau_{i_l}.$$

This, together with Theorem 3.6 applied to the group $\widetilde{M}_I$, implies that each constituent of $\pi|_{\widetilde{M}_I}$ is of the form $(\pi_{i_1}\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_{i_l})_{\omega'}$. □

The global metaplectic tensor product. By essentially following the local metaplectic tensor product, the global metaplectic tensor product was constructed in [Takeda 2016] with some technical assumptions, most notably Hypothesis $(*)$ on page 202 of that work. Here we will simplify our previous construction and remove the technical assumptions imposed there. Throughout this subsection, let $\pi_1, \dots, \pi_k$ be automorphic subrepresentations of the groups $\widetilde{\mathrm{GL}}_{r_1}(\mathbb{A}), \dots, \widetilde{\mathrm{GL}}_{r_k}(\mathbb{A})$ realized in the spaces of automorphic forms. Namely we assume

$$V_{\pi_i} \subseteq \mathcal{A}(\widetilde{\mathrm{GL}}_{r_i}).$$

Also let

$$H_i := \mathrm{GL}_{r_i}(F)\widetilde{\mathrm{GL}}_{r_i}^{(n)}(\mathbb{A}).$$

Note that by Lemma 2.10, H_i is a closed normal subgroup of $\mathrm{GL}_{r_i}(\mathbb{A})$ whose quotient is a compact abelian group.

First let

$$\sigma_i := \pi_i\|_{H_i},$$

where we recall the notation $\|$ from the notation section. Each element φ in the space of V_{σ_i} is a restriction to H_i of an automorphic form on $\widetilde{\mathrm{GL}}_{r_i}(\mathbb{A})$, and hence we may view it as a function on H_i with the property that $\varphi(\gamma g) = \varphi(g)$ for all $\gamma \in \mathrm{GL}_{r_i}(F)$ and $g \in \widetilde{\mathrm{GL}}_{r_i}^{(n)}(\mathbb{A})$. Namely the representation σ_i is a representation of the group H_i realized in a space of "automorphic forms on H_i".

We should mention

Proposition 3.9. *Let π be an irreducible smooth representation of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$. Then the restriction $\pi|_{\mathrm{GL}_r(F)\widetilde{\mathrm{GL}}_r^{(n)}(\mathbb{A})}$ is completely reducible, and hence $\pi\|_{\mathrm{GL}_r(F)\widetilde{\mathrm{GL}}_r^{(n)}(\mathbb{A})}$ is a subrepresentation of $\pi|_{\mathrm{GL}_r(F)\widetilde{\mathrm{GL}}_r^{(n)}(\mathbb{A})}$.*

Proof. In this proof, let us write $H = \mathrm{GL}_r(F)\widetilde{\mathrm{GL}}_r^{(n)}(\mathbb{A})$. We will prove the proposition by modifying the proof of [Gelbart and Knapp 1982, Lemma 2.1].

First we will show that the restriction $\pi|_H$ has an irreducible subrepresentation. For this, consider the contragredient $\hat{\pi}$ of π. Since π is irreducible, so is $\hat{\pi}$. Let

$\varphi \in \hat{\pi}$ be nonzero. Then $\hat{\pi}$ is generated by φ as a representation of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$. One of the key points is that the restriction $\hat{\pi}|_H$ is also finitely generated as a representation of H. To see it, let S be a sufficiently large finite set of places such that φ is fixed by $K^S := \prod_{v \notin S} \mathrm{GL}_r(\mathcal{O}_{F_v})$. We know that the set $H \setminus \widetilde{\mathrm{GL}}(\mathbb{A})/K^S$ is finite by Lemma 2.11. Let $\{g_1, \ldots, g_l\}$ be a complete set of representatives of this finite set. Then one can see that the vectors $\hat{\pi}(g_i)\varphi$ generate $\hat{\pi}|_H$, i.e., $\hat{\pi}|_H$ is finitely generated. Hence $\hat{\pi}|_H$ has an irreducible quotient. (It is an elementary exercise of Zorn's lemma to show that every finitely generated representation of any group has an irreducible quotient.) Let W be the kernel of the surjection from $V_{\hat{\pi}}$ to this irreducible quotient. Let

$$\mathrm{Ann}(W) := \{f \in V_\pi : \langle f, \varphi \rangle = 0 \text{ for all } \varphi \in W\}$$

be the annihilator of W, which gives a representation of H. Then one can see $\mathrm{Ann}(W)$ is an irreducible subrepresentation of $\pi|_H$ as follows. Let $X \subseteq \mathrm{Ann}(W)$ be any nonzero subrepresentation of $\mathrm{Ann}(W)$. Consider the annihilator $\mathrm{Ann}(X)$ of X, which is a subrepresentation of $V_{\hat{\pi}}$. Note that $W \subseteq \mathrm{Ann}(X) \subseteq V_{\hat{\pi}}$. But since the pairing $V_\pi \times V_{\hat{\pi}} \to \mathbb{C}$ is nondegenerate, we have $\mathrm{Ann}(X) \neq V_{\hat{\pi}}$. Hence $W = \mathrm{Ann}(X)$ by the irreducibility of $V_{\hat{\pi}}/W$. Hence the pairing $V_\pi \times V_{\hat{\pi}} \to \mathbb{C}$ gives rise to a nondegenerate pairing

$$X \times V_{\hat{\pi}}/W \to \mathbb{C},$$

which is H invariant. This implies that X is canonically isomorphic to the representation realized in the space

$$\{\langle f, - \rangle : f \in X\},$$

where $\langle f, - \rangle : V_{\hat{\pi}}/W \to \mathbb{C}$ is the functional given by $\varphi \mapsto \langle f, \varphi \rangle$. But this space is independent of X. Hence $X = \mathrm{Ann}(W)$, which shows $\mathrm{Ann}(W)$ is irreducible.

Now let V be an irreducible subrepresentation of $\pi|_H$ and let $f \in V$ be a fixed nonzero vector. As above there exists a sufficiently large S, possibly (most likely) different from the above one, such that the group K^S fixes f. Again let $\{g_1, \ldots, g_l\}$ be a complete set of representatives of $H \setminus \widetilde{\mathrm{GL}}(\mathbb{A})/K^S$, which is most likely different from the one above. Then one can see

$$V_\pi = \sum_{i=1}^{l} \pi(Hg_iK^S)f = \sum_{i=1}^{l} \pi(Hg_i)f.$$

Note that each space $\pi(Hg_i)f$ gives rise to a representation of H, which is equivalent to the g_i twist of $\pi(H)f$. But $\pi(H)f = V$ because V is a space of an irreducible representation of H, and hence each $\pi(Hg_i)f$ is irreducible.

Let $\{g_{i_1}, \dots, g_{i_N}\}$ be the smallest subset of $\{g_1, \dots, g_l\}$ such that

$$V_\pi = \sum_{j=1}^{N} \pi(Hg_{i_j})f.$$

This is actually a direct sum because for each $k \in \{1, \dots, N\}$, if the intersection

$$\pi(Hg_{i_k})f \cap \sum_{j \neq k} \pi(Hg_{i_j})f,$$

which is a representation of H, is nonzero, then it is actually equal to $\pi(Hg_{i_k})f$ by irreducibility, which contradicts to the minimality of the set $\{g_{i_1}, \dots, g_{i_N}\}$. This competes the proof. $\square$

Next note that each element in H_i is of the form (h_i, ξ_i) for $h_i \in \mathrm{GL}_{r_i}(F)\,\mathrm{GL}_{r_i}(\mathbb{A})$ and $\xi_i \in \mu_n$. As in [Takeda 2016, p. 215], we have the natural surjection

$$H_1 \times \cdots \times H_k \to M(F)\widetilde{M}^{(n)}(\mathbb{A}) \tag{3.10}$$

given by the map $((h_1, \xi_1), \dots, (h_k, \xi_k)) \mapsto (h_1 \cdots h_k, \xi_1 \cdots \xi_k)$. Then consider the space

$$V_{\sigma_1} \otimes \cdots \otimes V_{\sigma_k}$$

of functions on the direct product $H_1 \times \cdots \times H_k$, which gives rise to a representation of the direct product $H_1 \times \cdots \times H_k$. But each element in V_σ, which is a function on this direct product, descends to a function on $M(F)\widetilde{M}^{(n)}(\mathbb{A})$, which is "automorphic" in the sense that it is left-invariant on $M(F)$. (It should be mentioned that this is not as immediate as it looks, especially due to some issues on cocycles. See [Takeda 2016, Proposition 5.2] for details.) If $\varphi_i \in V_{\sigma_i}$ for $i = 1, \dots, k$, we denote this function by

$$\varphi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \varphi_k,$$

and denote the space generated by those functions by V_σ. We call each function in V_σ an "automorphic form on $M(F)\widetilde{M}^{(n)}(\mathbb{A})$". The group $M(F)\widetilde{M}^{(n)}(\mathbb{A})$ acts on V_σ by right-translation, and denote this representation by σ. We define

$$\sigma_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \sigma_k := \sigma.$$

Proposition 3.11. *With the above notation, σ is completely reducible. Further, if all of $\pi_1, \dots, \pi_k$ are unitary, so is σ.*

Proof. Each σ_i is completely reducible by Proposition 3.9. Hence one can see σ is completely reducible. If $\pi_1, \dots, \pi_k$ are unitary and each σ_i is a subrepresentation of $\pi_i|_{H_i}$, the unitary structure on π_i descends to σ_i. Hence one can define a unitary structure on $\sigma_1 \otimes \cdots \otimes \sigma_k$, which descends to σ. $\square$

Now just as we did for the local case, consider the smooth induced representation

$$\Pi = \Pi(\pi_1, \dots, \pi_k) := \mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})} \sigma. \tag{3.12}$$

Then we have the obvious map

$$\mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})} \sigma \to \mathcal{A}(\widetilde{M}), \quad f \mapsto \tilde{f} \tag{3.13}$$

where $\tilde{f}$ is defined by

$$\tilde{f}(m) = f(m)(1) \quad \text{for } m \in \widetilde{M}(\mathbb{A})$$

and $\mathcal{A}(\widetilde{M})$ is the space of automorphic forms on $\widetilde{M}(\mathbb{A})$. Further this map is one-to-one, and one can identify Π as a subspace of $\mathcal{A}(\widetilde{M})$, namely we have

$$\Pi \subseteq \mathcal{A}(\widetilde{M}).$$

Proposition 3.14. *If all of $\pi_1, \dots, \pi_k$ are cuspidal, then Π, viewed as a subspace of $\mathcal{A}(\widetilde{M})$, is in the space of cusp forms. If all of $\pi_1, \dots, \pi_k$ are realized in the spaces of square integrable automorphic forms, then Π is also in the space of square integrable automorphic forms. Also if all of $\pi_1, \dots, \pi_k$ are unitary, so is Π.*

Proof. The proofs for the first two parts (cuspidality and square-integrability) are simple modifications of the proofs of Theorems 5.12 and 5.13 of [Takeda 2016]. But let us repeat the key points. For this purpose, note that for $f \in \Pi$ and $m \in \widetilde{M}(\mathbb{A})$, we have $f(m) \in V_\sigma$, and hence $f(m)$ is a sum of functions of the form $\varphi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \varphi_k$, where each φ_i is a restriction of a function in V_{π_i} to H_i.

Assume that $\pi_1, \dots, \pi_k$ are cuspidal, and $N := N_1 \times \cdots \times N_k$ is a unipotent radical of a parabolic of M, where each N_i is a unipotent radical of a parabolic of GL_{r_i}. We view $N(\mathbb{A})$ as a subgroup of $\widetilde{M}(\mathbb{A})$ via the splitting $N(\mathbb{A}) \to \widetilde{M}(\mathbb{A})$. Noting that $N(\mathbb{A}) \subseteq M(F)\widetilde{M}^{(n)}(\mathbb{A})$, we have

$$\begin{aligned}
\int_{N(F)\backslash N(\mathbb{A})} \tilde{f}(nm)\, dn &= \int_{N(F)\backslash N(\mathbb{A})} f(nm)(1)\, dn \\
&= \int_{N(F)\backslash N(\mathbb{A})} f(m)(n)\, dn \\
&= \sum \int_{N(F)\backslash N(\mathbb{A})} (\varphi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \varphi_k)(n)\, dn \\
&= \sum \int_{N(F)\backslash N(\mathbb{A})} \varphi_1(n) \cdots \varphi_k(n)\, dn \\
&= \sum \int_{N_1(F)\backslash N_1(\mathbb{A})} \varphi_1(n_1)\, dn_1 \cdots \int_{N_k(F)\backslash N_k(\mathbb{A})} \varphi_k(n_k)\, dn_k \\
&= 0,
\end{aligned}$$

where the last equality follows from the cuspidality of the φ_i.

Next let us show the square integrability. By [Takeda 2016, Lemma 5.17], it suffices to show that

$$\int_{Z^{(n)}_{M(\mathbb{A})}M(F)\backslash M(\mathbb{A})} |\tilde{f}(m)|^2\,dm<\infty$$

for each $\tilde{f}\in\Pi$, where

$$Z^{(n)}_{M(\mathbb{A})}=\left\{\begin{pmatrix} a_1^n I_{r_1} & & \\ & \ddots & \\ & & a_k^n I_{r_k}\end{pmatrix} : a_i\in\mathbb{A}^\times\right\}.$$

But

$$\begin{aligned}\int_{Z^{(n)}_{M(\mathbb{A})}M(F)\backslash M(\mathbb{A})} &|\tilde{f}(m)|^2\,dm\\ &=\int_{Z^{(n)}_{M(\mathbb{A})}M(F)M^{(n)}(\mathbb{A})\backslash M(\mathbb{A})}\int_{Z^{(n)}_{M(\mathbb{A})}M^{(n)}(F)\backslash M^{(n)}(\mathbb{A})} |\tilde{f}(m'm)|^2\,dm'\,dm\\ &=\int_{Z^{(n)}_{M(\mathbb{A})}M(F)M^{(n)}(\mathbb{A})\backslash M(\mathbb{A})}\int_{Z^{(n)}_{M(\mathbb{A})}M^{(n)}(F)\backslash M^{(n)}(\mathbb{A})} |f(m'm)(1)|^2\,dm'\,dm\\ &=\int_{Z^{(n)}_{M(\mathbb{A})}M(F)M^{(n)}(\mathbb{A})\backslash M(\mathbb{A})}\int_{Z^{(n)}_{M(\mathbb{A})}M^{(n)}(F)\backslash M^{(n)}(\mathbb{A})} |f(m)(m')|^2\,dm'\,dm,\end{aligned}$$

Note that the outer integral is over a compact set, and hence we only need to show the convergence of the inner integral. But this follows from the square integrability of the function $f(m)\in V_\sigma$ as an "automorphic form on $M(F)\widetilde{M}^{(n)}(\mathbb{A})$".

Finally, assume $\pi_1,\dots,\pi_k$ are unitary. By Proposition 3.11, we know σ is unitary. But by Lemma 2.11, we know the induction defining Π is a compact induction, which makes Π unitary. □

Remark 3.15. In the above proof for square integrability, we implicitly used the fact that the group $M(F)Z^{(n)}_{M(\mathbb{A})}M^{(n)}(\mathbb{A})$ is closed, which can be shown by the same argument as Proposition 2.14. This justifies the existence of the quotient measure for $Z^{(n)}_{M(\mathbb{A})}M(F)M^{(n)}(\mathbb{A})\setminus M(\mathbb{A})$. The author has to admit that this subtle point was not addressed in the proof of [Takeda 2016, Theorem 5.13]. Also there we, for some reason, did not realize that the group $Z^{(n)}_{M(\mathbb{A})}M(F)M^{(n)}(\mathbb{A})\setminus M(\mathbb{A})$ is compact when writing our previous paper, which made the proof there unnecessarily long.

We would like to have that the representation $\Pi=\Pi(\pi_1,\dots,\pi_k)$ is completely reducible, as in the local case. And this is immediate if $\pi_1,\dots,\pi_k$ are cuspidal because then Π is in the space of cusp forms. We do not know if this is true in general, but the following formulation is enough for our purposes:

Proposition 3.16. *Let $\tau \subseteq \sigma$ be an irreducible subspace. Then the space*

$$\mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})} \tau$$

has an irreducible subrepresentation. Hence Π has an irreducible subrepresentation.

Proof. In this proof, let us write $H = M(F)\widetilde{M}^{(n)}(\mathbb{A})$. First note that since σ is completely reducible by Proposition 3.9, an irreducible $\tau \subseteq \sigma$ always exists. Let $\varphi \in V_\tau$ be nonzero. Since $\pi_1, \ldots, \pi_k$ are smooth, there exists a finite set of places such that φ is fixed by the group $H \cap K^S$, where $K^S = \prod_{v \notin S} M(\mathcal{O}_{F_v})$. Let $g_1, \ldots, g_l$ be a complete set of representatives of the double cosets $H \setminus \widetilde{M}(\mathbb{A})/K^S$, which, we know, is finite by Lemma 2.11, where we assume $g_1 = 1$. Hence each vector in $\mathrm{Ind}_H^{\widetilde{M}(\mathbb{A})} \tau$ fixed by K^S is completely determined by its values at $g_1, \ldots, g_l$. With this said, let us define an element $f : \widetilde{M}(\mathbb{A}) \to V_\tau$ in $\mathrm{Ind}_H^{\widetilde{M}(\mathbb{A})} \tau$ by setting

$$f(hg_ik) = \begin{cases} \tau(h)\varphi & \text{if } i = 1, \\ 0 & \text{otherwise,} \end{cases}$$

where $h \in H$ and $k \in K^S$. This is well defined because φ is fixed by $H \cap K^S$, and has the property that $f(hm) = \tau(h)f(m)$ for all $h \in H$ and $m \in \widetilde{M}(\mathbb{A})$. To show f is indeed in $\mathrm{Ind}_H^{\widetilde{M}(\mathbb{A})} \tau$, we need to show that f is smooth. This can be checked at each v by viewing $\widetilde{M}(F_v)$ as a subgroup of $\widetilde{M}(\mathbb{A})$ as in (2.7) (or its $\widetilde{M}(\mathbb{A})$ analogue). Namely for each $v \notin S$, clearly f is fixed by $M(\mathcal{O}_{F_v})$ and hence f is smooth at v. If v is Archimedean, then since the Lie algebra of $\widetilde{M}(F_v)$ is the same as that of $\widetilde{M}^{(n)}(F_v)$, the smoothness follows from that of φ. Finally let $v \in S$ be a non-Archimedean place in S. Then by the smoothness of φ, there is an open compact subgroup U of $\widetilde{M}^{(n)}(F_v)$. Since $\widetilde{M}^{(n)}(F_v)$ is an open subgroup of $\widetilde{M}(F_v)$, U is also an open compact subgroup of $\widetilde{M}(F_v)$. Then one can see that the intersection of all $g_i^{-1}Ug_i$, which is also an open compact subgroup of $\widetilde{M}(F_v)$, fixes f. Hence f is smooth and indeed in $\mathrm{Ind}_H^{\widetilde{M}(\mathbb{A})} \tau$.

Now consider the space $\Pi(\widetilde{M}(\mathbb{A}))f$ generated by f. Then we can write

$$\Pi(\widetilde{M}(\mathbb{A}))f = \sum_{i=1}^{l} \Pi(Hg_i)f,$$

where each space $\Pi(Hg_i)f$ is H invariant and hence a subrepresentation of $\Pi(\widetilde{M}(\mathbb{A}))f|_H$. Now to prove the proposition, it suffices to show that $\Pi(Hg_i)f$ is irreducible, because, then, $\Pi(\widetilde{M}(\mathbb{A}))f|_H$ has only finite length, and hence *a fortiori* $\Pi(\widetilde{M}(\mathbb{A}))f$ is of finite length, which implies that $\Pi(\widetilde{M}(\mathbb{A}))f$ has an irreducible subrepresentation. Moreover, one can see that, as abstract representations, each $\Pi(Hg_i)f$ is equivalent to the g_i twist of $\Pi(H)f$. Hence it suffices to show that $\Pi(H)f$ is irreducible.

To show each $\Pi(H)f$ is irreducible, consider the evaluation map at 1, namely

$$\Pi(H)f \to \tau, \quad f' \mapsto f'(1)$$

for $f' \in \Pi(H)f$, which is H-intertwining. Since $f(1) = \varphi \neq 0$, this map is nonzero. But note that each nonzero f' is supported on HK^S, which implies $f'(1) \neq 0$ for all nonzero $f' \in \Pi(H)f$. Therefore $\Pi(H)f \cong \tau$, which shows $\Pi(H)f$ is irreducible. □

Now as in the local case, $\Pi(\pi_1, \ldots, \pi_k)$ essentially contains all the possible metaplectic tensor products. To see it, we need to carry out a construction analogous to the local metaplectic tensor product of Mezo by taking the central character into account. Namely, we now need to consider an irreducible subrepresentation of σ and extend it to a representation of $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F)\widetilde{M}^{(n)}(\mathbb{A})$ by letting the center $Z_{\widetilde{\mathrm{GL}}_r}(\mathbb{A})$ act as a character.

First note that since σ is completely reducible by Proposition 3.11, it has an irreducible subrepresentation

$$\tau \subseteq \sigma.$$

Fix such τ from now on. We need

Lemma 3.17. *For each irreducible $\tau \subseteq \sigma$, the abelian group*

$$Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \cap M(F)\widetilde{M}^{(n)}(\mathbb{A})$$

acts as a character, which we denote by ω_τ.

Proof. By Proposition 3.16, there exists an irreducible subrepresentation π of $\mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})} \tau$. Let ω be the central character of π. Now by Frobenius reciprocity we have an $M(F)\widetilde{M}^{(n)}(\mathbb{A})$-intertwining map $\pi \to \tau$, which shows that the group $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \cap M(F)\widetilde{M}^{(n)}(\mathbb{A})$ acts via the character ω on τ. □

Remark 3.18. Of course, if τ is unitary, which is the case if $\pi_1, \cdots, \pi_k$ are, then τ actually has a central character because $M(F)\widetilde{M}^{(n)}(\mathbb{A})$ is locally compact. But the author does not know if τ admits a central character in general.

By the "automorphy" of each element in V_τ, we can see that the character ω_τ in the above lemma is "automorphic" in the sense that

$$\omega_\tau(\gamma z) = \omega_\tau(z)$$

for all $z \in Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \cap M(F)\widetilde{M}^{(n)}(\mathbb{A})$ and $\gamma \in M(F) \cap (Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \cap M(F)\widetilde{M}^{(n)}(\mathbb{A}))$. Then we can find a "Hecke character" ω on $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$ by extending ω_τ; namely ω is a character on $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$ such that

$$\omega(z) = \omega_\tau(z) \quad \text{for all } z \in Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \cap M(F)\widetilde{M}^{(n)}(\mathbb{A}).$$

Such ω always exists because both $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$ and $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \cap M(F)\widetilde{M}^{(n)}(\mathbb{A})$ are locally compact abelian groups. Also note that any such ω is indeed a "Hecke character" in the sense that

$$\omega(\gamma z) = \omega(z)$$

for all $z \in Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$ and $\gamma \in \mathrm{GL}_r(F) \cap Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$, simply because

$$\mathrm{GL}_r(F) \cap Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \subseteq Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \cap M(F)\widetilde{M}^{(n)}(\mathbb{A}).$$

For each $f \in V_\tau$, which is a function on $M(F)\widetilde{M}^{(n)}(\mathbb{A})$, we can extend it to a function

$$f_\omega : Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F)\widetilde{M}^{(n)}(\mathbb{A}) \to \mathbb{C}$$

by

$$f_\omega(zm) = \omega(z) f(m) \quad \text{for all } z \in Z_{\widetilde{\mathrm{GL}}_r}(\mathbb{A}) \text{ and } m \in M(F)\widetilde{M}^{(n)}(\mathbb{A}).$$

This is well defined because of our choice of ω, and is considered as an "automorphic form on the group $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F)\widetilde{M}^{(n)}(\mathbb{A})$". We define

$$V_{\tau_\omega} := \{f_\omega : f \in V_\tau\}.$$

The group $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F)\widetilde{M}^{(n)}(\mathbb{A})$ irreducibly acts on this space, giving rise to an "automorphic representation" τ_ω of $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F)\widetilde{M}^{(n)}(\mathbb{A})$. As abstract representations, we have

$$\tau_\omega \cong \omega\tau \tag{3.19}$$

where by $\omega\tau$ we mean the representation of the group $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F)\widetilde{M}^{(n)}(\mathbb{A})$ extended from τ by letting $Z_{\widetilde{\mathrm{GL}}(\mathbb{A})}$ act via the character ω.

As we did before, let us consider the smooth induced representation

$$\Pi(\tau_\omega) := \mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})} \tau_\omega.$$

Note that we have the obvious inclusion

$$\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})} \tau_\omega \hookrightarrow \mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})} \sigma,$$

which allows us to view $\Pi(\tau_\omega)$ as a subrepresentation of Π realized in the space of automorphic forms on $\widetilde{M}(\mathbb{A})$, namely

$$\Pi(\tau_\omega) \subseteq \Pi \subseteq \mathcal{A}(\widetilde{M}).$$

Then we have

Proposition 3.20. *The representation $\Pi(\tau_\omega)$ has an irreducible subrepresentation.*

Proof. This can be proven identically to Proposition 3.16. $\square$

Finally, we can define our metaplectic tensor product as follows.

Definition 3.21. Keeping the above notations, let $\pi_\omega \subseteq \Pi(\tau_\omega)$ be an irreducible subrepresentation. Then we write

$$\pi_\omega = (\pi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_k)_\omega$$

and call it a metaplectic tensor product of $\pi_1, \ldots, \pi_k$ with respect to the character ω.

The definition of metaplectic tensor product along with Proposition 3.14 immediately implies the following.

Proposition 3.22. *If all of $\pi_1, \cdots, \pi_k$ are cuspidal (alternatively, unitary, or square integrable modulo center), then so is $(\pi_1 \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_k)_\omega$.*

This π_ω is precisely the metaplectic tensor product we want:

Theorem 3.23. *The representation π_ω constructed above has the desired local-global compatibility. Namely if we write $\pi_\omega = \widetilde{\otimes}'_v \pi_{\omega,v}$, then for each v we have*

$$\pi_{\omega,v} = (\pi_{1,v} \widetilde{\otimes} \cdots \widetilde{\otimes} \pi_{k,v})_{\omega_v}.$$

Thus π_ω is unique up to equivalence, and depends only on $\pi_1, \ldots, \pi_k$ and ω.

Proof. Note that the uniqueness assertion follows from the corresponding local statement that the local metaplectic tensor product only depends on $\pi_{1,v}, \ldots, \pi_{k,v}$ and ω_v. Hence we have only to show the local-global compatibility.

First, note that, since $\pi_\omega \subseteq \mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F) \widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})} \tau_\omega$, we have the natural surjection

$$\pi_\omega|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} M(F) \widetilde{M}^{(n)}(\mathbb{A})} \to \tau_\omega.$$

Recall that as abstract representations, we have $\tau_\omega \cong \omega\tau$, where τ is an irreducible representation of $M(F)\widetilde{M}^{(n)}(\mathbb{A})$. So by restricting further down to $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \widetilde{M}^{(n)}(\mathbb{A})$, we have

$$\pi_\omega|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \widetilde{M}^{(n)}(\mathbb{A})} \to \omega\tau|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \widetilde{M}^{(n)}(\mathbb{A})}.$$

Now by Lemma A.1 in the Appendix, $\pi_\omega|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \widetilde{M}^{(n)}(\mathbb{A})}$ is completely reducible. Hence $\omega\tau|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \widetilde{M}^{(n)}(\mathbb{A})}$ is completely reducible. Let

$$\omega\pi^{(n)} \subseteq \tau_\omega|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \widetilde{M}^{(n)}(\mathbb{A})}$$

be an irreducible subrepresentation, where $\pi^{(n)}$ is an irreducible representation of $\widetilde{M}^{(n)}(\mathbb{A})$. By complete reducibility, this is also a quotient, and hence we have a surjection

$$\pi_\omega|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})} \widetilde{M}^{(n)}(\mathbb{A})} \to \omega\pi^{(n)}.$$

Recall that τ is realized as a space of "automorphic forms on $M(F)\widetilde{M}^{(n)}(\mathbb{A})$" and is written as

$$V_\tau = V_{\tau_1}\widetilde{\otimes}\cdots\widetilde{\otimes}V_{\tau_k},$$

where each V_{τ_i} is a space of restrictions of automorphic forms in the space V_{π_i}. By the automorphy, one can see that

$$\tau_\omega|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\widetilde{M}^{(n)}(\mathbb{A})} = \tau_\omega\|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\widetilde{M}^{(n)}(\mathbb{A})}.$$

Hence we have

$$V_{\pi^{(n)}} = V_{\pi_1^{(n)}}\widetilde{\otimes}\cdots\widetilde{\otimes}V_{\pi_k^{(n)}},$$

where $V_{\pi_i^{(n)}} \subseteq V_{\tau_i}$ for each i, and indeed we have

$$\pi_i^{(n)} \subseteq \tau_i|_{\widetilde{\mathrm{GL}}_{r_i}^{(n)}(\mathbb{A})} = \tau_i\|_{\widetilde{\mathrm{GL}}_{r_i}^{(n)}(\mathbb{A})} \subseteq \pi_i\|_{\widetilde{\mathrm{GL}}_{r_i}^{(n)}(\mathbb{A})}.$$

Therefore we can write

$$\omega\pi^{(n)} = \widetilde{\otimes}'_v\omega_v(\pi_{1,v}^{(n)}\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_{k,v}^{(n)}),$$

where $\omega_v(\pi_{1,v}^{(n)}\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_{k,v}^{(n)})$ is the irreducible representation of $Z_{\widetilde{\mathrm{GL}}_r(F_v)}\widetilde{M}^{(n)}(F_v)$ constructed from $\pi_{1,v}^{(n)},\dots,\pi_{k,v}^{(n)}$ as is done for the local metaplectic tensor product.

Then if we write

$$\pi_\omega = \widetilde{\bigotimes_v}' \pi_{\omega,v},$$

where $\pi_{\omega,v}$ is an irreducible representation of $\widetilde{M}(F_v)$, we have the surjection

$$\Big(\widetilde{\bigotimes_v}'\pi_{\omega,v}\Big)|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\widetilde{M}^{(n)}(\mathbb{A})} \to \widetilde{\bigotimes_v}'\omega_v\big(\pi_{1,v}^{(n)}\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_{k,v}^{(n)}\big).$$

Hence by Lemma 5.5 of [Takeda 2016], we conclude that at each place v, the representation $\omega_v\big(\pi_{1,v}^{(n)}\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_{k,v}^{(n)}\big)$ is a quotient of $\pi_{\omega,v}|_{Z_{\widetilde{\mathrm{GL}}_r(F_v)}\widetilde{M}^{(n)}(F_v)}$. By Frobenius reciprocity, we have

$$\pi_{\omega,v} \subseteq \mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(F_v)}\widetilde{M}^{(n)}(F_v)}^{\widetilde{M}(F_v)}\,\omega_v\big(\pi_{1,v}^{(n)}\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_{k,v}^{(n)}\big).$$

Thus by the definition of local metaplectic tensor product, we have

$$\pi_{\omega,v} = (\pi_{1,v}\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_{k,v})_{\omega_v}.$$

Hence we have the desired local-global compatibility. □

Remark 3.24. With the theorem, we can say that the notation $(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega$ is unambiguous in the sense that it only depends on $\pi_1,\dots,\pi_k$ and ω as long as we consider the metaplectic tensor product as an equivalence class of representations, which we usually do.

This theorem immediately implies the following.

Corollary 3.25. *For fixed ω, all the irreducible subrepresentations of*

$$\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})}\tau_\omega$$

are equivalent.

Next we will show that $\Pi(\pi_1,\ldots,\pi_k)$ contains all the possible metaplectic tensor products of $\pi_1,\ldots,\pi_k$. For this purpose, let us define

$\Omega=\Omega(\pi_1,\ldots,\pi_k):=\{\omega:\omega$ is a Hecke character on $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$ which appears in $\sigma\}$,

where we say "ω appears in σ" if there exists a nonzero function $\varphi\in\sigma$ such that

$$\varphi(zm)=\omega(z)\varphi(m) \tag{3.26}$$

for all $z\in Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\cap M(F)\widetilde{M}^{(n)}(\mathbb{A})$ and $m\in M(F)\widetilde{M}^{(n)}(\mathbb{A})$.

We need

Proposition 3.27. *Let $\omega\in\Omega$ be as above, i.e., ω appears in σ. Then there exists a metaplectic tensor product $\pi_\omega=(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega$ such that $\pi_\omega\subseteq\Pi$.*

Proof. Since ω appears in σ, there exists $\varphi\in V_\sigma$ with the property (3.26). Consider the space $\sigma(M(F)\widetilde{M}^{(n)}(\mathbb{A}))\varphi$ generated by φ inside V_σ. Because each $z\in Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\cap M(F)\widetilde{M}^{(n)}(\mathbb{A})$ is in the center of $M(F)\widetilde{M}^{(n)}(\mathbb{A})$, one can see that $\sigma(z)\varphi'=\omega(z)\varphi'$ for all $\varphi'\in\sigma(M(F)\widetilde{M}^{(n)}(\mathbb{A}))\varphi$. Hence if we pick up an irreducible $\tau\subseteq\sigma(M(F)\widetilde{M}^{(n)}(\mathbb{A}))\varphi$, we can extend it to τ_ω, and an irreducible subrepresentation of $\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})}\tau_\omega$ is the desired metaplectic tensor product. $\square$

With this proposition, we can state the global analogue of Proposition 3.5 as follows:

Proposition 3.28. *First we have the decomposition*

$$\Pi(\pi_1,\ldots,\pi_k)=\bigoplus_\tau m(\tau)\,\mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})}\tau,$$

where the sum is over all the equivalence classes $\tau\subseteq\sigma$ and $m(\tau)$ is the positive multiplicity of τ in σ. Further for each fixed τ, let

$$\Omega(\tau):=\left\{\omega\in\Omega:\omega|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\cap M(F)\widetilde{M}^{(n)}(\mathbb{A})}=\tau|_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\cap M(F)\widetilde{M}^{(n)}(\mathbb{A})}\right\}.$$

Then we have

$$\mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})}\tau\supseteq\bigoplus_{\omega\in\Omega(\tau)}\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})}\tau_\omega\supseteq\bigoplus_{\omega\in\Omega(\tau)}m(\tau,\omega)(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega,$$

where $m(\tau,\omega)$ is the positive multiplicity.

Proof. The proposition can be proven in the same way as the local case. Yet, we should mention that unlike the local case, we do not seem to know the precise information on the multiplicities. □

From the proposition, the following is immediate.

Theorem 3.29. *We have*

$$\Pi(\pi_1, \dots, \pi_k) \supseteq \bigoplus_{\omega\in\Omega} m(\omega)(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega,$$

where $m(\omega)$ is the multiplicity of $(\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega$. Also if all of $\pi_1, \dots, \pi_k$ are cuspidal, then the inclusion is actually an equality.

Proof. The first part is obvious from the above proposition. The second part follows from Proposition 3.14 because if Π is in the cuspidal spectrum, it is completely reducible. □

Let us note that if we know the multiplicity-one property for the group $\widetilde{M}(\mathbb{A})$, we could set $m(\omega) = 1$. Yet, the author does not know if the multiplicity-one property holds even for the cuspidal spectrum.

Restriction to a smaller Levi. As we did for the local case, we will discuss the restriction of our metaplectic tensor products to a smaller Levi $\widetilde{M}_I$, where M_I is as in (2.4). Recall

$$\sigma = \sigma_1\widetilde{\otimes}\cdots\widetilde{\otimes}\sigma_k,$$

whose space V_σ is essentially identified with $V_{\sigma_1}\otimes\cdots\otimes V_{\sigma_k}$, which is a space of functions on the direct product $H_1\times\cdots\times H_k$. Hence if we let

$$\sigma_I := \sigma_{i_1}\widetilde{\otimes}\cdots\widetilde{\otimes}\sigma_{i_l},$$

one can see that if $\varphi\in V_\sigma$, we have $\varphi|_{M_I(F)\widetilde{M}_I^{(n)}(\mathbb{A})} \in V_{\sigma_I}$. Indeed, we have an $M_I(F)\widetilde{M}_I^{(n)}(\mathbb{A})$-intertwining surjection

$$\sigma\to\sigma_I, \quad \varphi\mapsto\varphi|_{M_I(F)\widetilde{M}_I^{(n)}(\mathbb{A})}.$$

In other words, we have $\sigma_I = \sigma\|_{M_I(F)\widetilde{M}_I^{(n)}(\mathbb{A})}$.

Now we can prove

Theorem 3.30. *For each $\pi_\omega = (\pi_1\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_k)_\omega \subseteq \Pi(\pi, \dots, \pi_k)$, we have*

$$\pi_\omega\|_{\widetilde{M}_I(\mathbb{A})} \subseteq \bigoplus_{\omega'\in\Omega_I} m(\omega')(\pi_{i_1}\widetilde{\otimes}\cdots\widetilde{\otimes}\pi_{i_l})_{\omega'},$$

where $\Omega_I = \Omega(\pi_{i_1}, \cdots, \pi_{i_k})$ is defined analogously to Ω.

Proof. Let us view π_ω as a subspace of the induced space of (3.12). Then we have the commutative diagram

$$\begin{array}{ccccc} \pi_\omega & \subseteq & \mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})}\,\sigma & \hookrightarrow & \mathcal{A}(\widetilde{M}) \\ \downarrow & & \downarrow & & \downarrow \\ \pi_\omega\|_{\widetilde{M}_I(\mathbb{A})} & \subseteq & \mathrm{Ind}_{M_I(F)\widetilde{M}_I^{(n)}(\mathbb{A})}^{\widetilde{M}_I(\mathbb{A})}\,\sigma_I & \hookrightarrow & \mathcal{A}(\widetilde{M}_I) \end{array}$$

where all the vertical arrows are restriction of functions and the hooked arrow on the top is the "automorphic realization map" as in (3.13), and the one on the bottom is its analogue for $\widetilde{M}_I$. By Lemma A.1, we know that $\pi_\omega|_{\widetilde{M}_I(\mathbb{A})}$ is completely irreducible, and hence so is $\pi_\omega\|_{\widetilde{M}_I(\mathbb{A})}$. Note that every irreducible subrepresentation of $\mathrm{Ind}_{M_I(F)\widetilde{M}_I^{(n)}(\mathbb{A})}^{\widetilde{M}_I(\mathbb{A})}\,\sigma_I$ is a metaplectic tensor product of $\pi_{i_1},\dots,\pi_{i_l}$ with respect to some ω'. Hence the theorem follows by identifying $\mathrm{Ind}_{M_I(F)\widetilde{M}_I^{(n)}(\mathbb{A})}^{\widetilde{M}_I(\mathbb{A})}\,\sigma_I$ with a subspace of $\mathcal{A}(\widetilde{M}_I)$. □

Other properties. In [Takeda 2016], a couple of other properties of metaplectic tensor product are discussed. To be precise, they have the expected behavior under the Weyl group action and the compatibility with parabolic induction, which are, respectively, Theorems 5.19 and 5.22 in that work. But both of them follow from the corresponding local statements, and hence they also hold in our new construction.

It should also be mentioned that recently it has been shown by W. T. Gan [2016] that the metaplectic tensor product can be interpreted as an instance of Langlands functoriality by using the L-group formalism of covering groups developed by Weissman. (See [Weissman 2016; Gan and Gao 2014] for this formalism.) This shows that the construction of the metaplectic tensor product is indeed a natural one.

Some remarks on past literature. The notion of metaplectic tensor product has been implicitly used in many of the past works on automorphic forms on $\widetilde{\mathrm{GL}}_r(\mathbb{A})$, especially when one would like to construct Eisenstein series on $\widetilde{\mathrm{GL}}_r(\mathbb{A})$. But there are various discrepancies in the past literature in this subject, which, we believe, was due to the lack of a foundation on the metaplectic tensor product. In this final subsection, let us briefly discuss some of the previous works and how they can be reconciled with the theory developed in this paper.

The first work that considered Eisenstein series on $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ is, of course, the important work of Kazhdan and Patterson [1984]. There they only considered those Eisenstein series which are induced from the Borel subgroup B. Namely they only considered the case $M = \mathrm{GL}_1 \times \cdots \times \mathrm{GL}_1$. In this case, one can show that the group $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})$ is a maximal abelian subgroup of $\widetilde{M}$, and accordingly, from the outset they considered a character on $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})$

instead of starting with characters on $\widetilde{\mathrm{GL}}_1(\mathbb{A})$ (see their Section II.1). Yet, one can see that this is the same as constructing some τ_ω in our notation. Then Kazhdan and Patterson considered the induced representation $\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})N_B}^{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\tau_\omega$, to construct Eisenstein series. By inducing in stages,

$$\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})N_B}^{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\tau_\omega=\mathrm{Ind}_{\widetilde{B}(\mathbb{A})}^{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})N_B}^{\widetilde{M}(\mathbb{A})N_B}\tau_\omega,$$

and hence the metaplectic tensor product that is implicitly used in [Kazhdan and Patterson 1984] is our $\Pi(\tau_\omega)=\mathrm{Ind}_{Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})}\tau_\omega$. Further the fact that $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})$ is maximal abelian implies that $\Pi(\tau_\omega)$ is irreducible (see Section 0.3 in the same reference).

The next important set of works on this subject is probably the one by Bump and Ginzburg [1992] on the symmetric square L-function, and the work by Banks [1997] on the twisted case for GL_3, both of which dealt with only the case $n=2$. There are two main parabolic subgroups considered there: the Borel and the $(r-1,1)$-parabolic. For the Borel, they use the same formulation as [Kazhdan and Patterson 1984]. For the $(r-1,1)$-parabolic, they first start with a representation of $\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})$ viewed as a subgroup of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ and they extend it to a representation of $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})$ by letting $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$ act by an appropriate character. Now if r is odd (and $n=2$), this gives a representation of $\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})\widetilde{\times}\widetilde{\mathrm{GL}}_1(\mathbb{A})$ because $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})=\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})\widetilde{\times}\widetilde{\mathrm{GL}}_1(\mathbb{A})$. But if r is even, we only have $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})=\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})\widetilde{\times}\widetilde{\mathrm{GL}}_1^{(2)}(\mathbb{A})$. Then, in [Bump and Ginzburg 1992], they induced the representation of $\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})\widetilde{\times}\widetilde{\mathrm{GL}}_1^{(2)}(\mathbb{A})$ to $\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})\widetilde{\times}\widetilde{\mathrm{GL}}_1(\mathbb{A})$ (see the middle of page 159 in the same reference.) However, it seems to the author that one cannot show the automorphy of this induced representation if it is simply induced from $\widetilde{\mathrm{GL}}_{r-1}(\mathbb{A})\widetilde{\times}\widetilde{\mathrm{GL}}_1^{(2)}(\mathbb{A})$, and probably this is another technical issue to be addressed in that work. At any rate, one can see that at least if r is odd this construction is also obtained as our metaplectic tensor product, say, by first restricting to $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}M(F)\widetilde{M}^{(n)}(\mathbb{A})$ and then inducing one of the constituents to $\widetilde{M}(\mathbb{A})$. It should be also mentioned that in [Bump and Ginzburg 1992; Banks 1997] various properties of metaplectic tensor product, such as the behavior of metaplectic tensor product upon restriction to a smaller Levi, are implicitly used.

From those two works, the author generalized in [Takeda 2014], in which the parabolic subgroups considered are mainly $(2,\dots,2)$ and $(r-1,1)$ parabolic. For the $(2,\dots,2)$-parabolic, the inducing representation for each $\widetilde{\mathrm{GL}}_2$ factor is only the Weil representation, and hence by using the Schrödinger model for $\widetilde{\mathrm{GL}}_2^{(2)}$, we explicitly constructed what we called the "Weil representation of $\widetilde{M}_P$". One can see that this is also an instance of our metaplectic tensor product, because for the case at hand we have $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\subseteq\widetilde{M}^{(n)}(\mathbb{A})$, which means that the central character does not play any role in the formation of metaplectic tensor product and hence the metaplectic tensor product only depends locally on restrictions to

$\widetilde{M}^{(n)}(F_v)$. For the $(r-1, 1)$-parabolic case in [Takeda 2014], however, depending on the parity of r, we took a different approach. For r odd, we did just as in [Bump and Ginzburg 1992; Banks 1997]. For r even, we directly constructed a representation of the Levi $\widetilde{\mathrm{GL}}_{r-1}\widetilde{\times}\widetilde{\mathrm{GL}}_1$ as residues of Eisenstein series induced from the Borel, instead of starting with representations of $\widetilde{\mathrm{GL}}_{r-1}$ and $\widetilde{\mathrm{GL}}_1$ separately. One can show that this construction is the same as our metaplectic tensor product by using the compatibility of our metaplectic tensor product with parabolic induction as discussed in the previous subsection.

Besides those applications to symmetric square L-functions, the works of Suzuki [1997; 1998] should be mentioned. In the first of these works, he considered the (r_1, r_2)-parabolic for $r_1+r_2=r$. To construct an automorphic form on the Levi part, he uses what he calls "partial Eisenstein series" (see his Sec. 5.4). This construction is essentially the same as the r = even case of [Takeda 2014] mentioned above, and again our metaplectic tensor product encompasses this construction of Suzuki. Also in [Suzuki 1998], Eisenstein series induced from the $(\ell, \dots, \ell)$-parabolic are considered. There it seems that what he considers is our Π, namely the whole induced representation $\mathrm{Ind}_{M(F)\widetilde{M}^{(n)}(\mathbb{A})}^{\widetilde{M}(\mathbb{A})}\sigma$. Yet, it should be mentioned that first of all he assumes that each automorphic representation of $\widetilde{\mathrm{GL}}_\ell(\mathbb{A})$ is already induced from $\mathrm{GL}_\ell(F)\widetilde{\mathrm{GL}}_\ell^{(n)}(\mathbb{A})$ (page 750 of that work), and second of all it is claimed, without proof, that the representation on the Levi thus constructed is irreducible with local-global compatibility. (See the beginning of [Suzuki 1998, p. 752].) At any rate, since no proofs or no detailed explanations are given for his assertions, it is not completely clear to the author that what kind of construction is carried out there and even that his construction is legitimate.

Finally, more recently Brubaker and Friedberg [2015] considered metaplectic Eisenstein series not just on the group $\widetilde{\mathrm{GL}}_r$ but on other covering groups in general. Although they use the language of "S-integers", what they use to construct representations of the Levi amounts to our Π, the whole induced representation. Also the same convention is used in the even more recent [Friedberg and Ginzburg 2016].

Probably which convention to use might be a matter of taste or the nature of the problem one works on. But it seems to the author that for the purpose of constructing Eisenstein series, using the whole induced space Π, which contains all the metaplectic tensor products, is an easy choice, especially because then, the inducing data is essentially the same as the usual tensor product. One should, however, be careful that usually the representation $\Pi=\Pi(\pi_1, \dots, \pi_k)$ is reducible. Hence for example we do not know if we can express it as a restricted tensor product as $\Pi=\widetilde{\bigotimes}'_v\Pi_v$, which is often crucial when one would like to find out analytic properties of intertwining operators. Therefore, it might be more convenient to pick an irreducible subrepresentation $\pi_\omega\subseteq\Pi$, although this requires one to take care of the dependence of π_ω on ω. Nonetheless, probably many of the important properties (especially analytic ones) of Eisenstein series constructed from different π_ω might usually be independent of ω, because, after all, the characters ω differ by characters on $Z_{\widetilde{\mathrm{GL}}(\mathbb{A})} \cap M(F)\widetilde{M}^{(n)}(\mathbb{A}) \backslash Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}$, which is compact, and hence it seems unlikely that a difference in a character on a compact group affects analytic properties of Eisenstein series. Indeed, for example, in [Takeda 2015] the author studied some analytic properties of some Eisenstein series using the formalism of metaplectic tensor product of [Takeda 2016], and all the results there hold independently of the choice of ω.

Appendix: A lemma on complete reducibility

We prove the result used on pages 219 and 223.

Lemma A.1. *Let $\widetilde{\pi} = \widetilde{\bigotimes}'_v \pi_v$ be an irreducible admissible representation of $\widetilde{M}(\mathbb{A})$. Let $\widetilde{H}$ be a group of the form $\widetilde{H} = \widetilde{\prod}'_v H_v$, where $H_v \subseteq \widetilde{M}(F_v)$ and the restricted direct product is with respect to the group $H_v \cap M(\mathcal{O}_{F_v})$. (The groups $\widetilde{M}^{(n)}(\mathbb{A})$, $Z_{\widetilde{\mathrm{GL}}_r(\mathbb{A})}\widetilde{M}^{(n)}(\mathbb{A})$ and $\widetilde{M}_I(\mathbb{A})$ are such examples of $\widetilde{H}$.) Further assume that for each v, the restriction $\pi_v|_{H_v}$ is completely reducible. Then the restriction $\widetilde{\pi}|_{\widetilde{H}}$ is completely reducible.*

Proof. We argue "semilocally" using the definition of the restricted metaplectic tensor product $\widetilde{\pi} = \widetilde{\bigotimes}'_v \pi_v$. First note that the space of $\widetilde{\bigotimes}'_v \pi_v$ is actually $\bigotimes'_v V_{\pi_v}$ (usual restricted tensor product) on which not only the group $\widetilde{\prod}'_v \widetilde{M}(F_v)$, but also $\prod'_v \widetilde{M}(F_v)$ acts. Accordingly we set

$$\pi := \bigotimes_v{}' \pi_v \quad \text{(usual restricted tensor product)},$$

$$H := \prod_v{}' H_v \quad \text{(usual restricted direct product)},$$

and it suffices to show that the restriction $\pi|_H$ is completely reducible.

Now let us recall the definition of $\bigotimes'_v \pi_v$. For almost all v, we choose a spherical vector $\xi_v^\circ \in \pi_v$. Let S be a sufficiently large finite set of places so that each π_v is spherical for $v \notin S$. Let

$$\pi_S = \bigotimes_{v \in S} \pi_v,$$

which gives a representation of $\prod_{v\in S} \widetilde{M}(F_v)$. For each $S' \supseteq S$ we have the inclusion $\pi_S \to \pi_{S'}$ by tensoring the chosen spherical vectors ξ_v° for $v \in S' \setminus S$. Then the system $\{\pi_S\}_S$ is a directed system and by definition $\bigotimes'_v \pi_v = \varinjlim_S \pi_S$. For each S, let us define $H_S := \prod_{v\in S} H_v$. Then one can see that

$$\pi|_H = \varinjlim_S \pi_S|_{H_S}.$$

For each v, the restriction $\pi_v|_{H_v}$ is completely reducible by our assumption.

Hence let us fix the decomposition

$$\pi_v|_{H_v} = \bigoplus_{i_v \in I_v} \pi_{i_v}$$

for some finite indexing set I_v, where each π_{i_v} is irreducible. (We do not assume that the restriction $\pi_v|_{H_v}$ is multiplicity free, and hence this decomposition might not be unique even up to ordering. So we "fix" the decomposition for each $\pi_v|_{H_v}$ once and for all.) We let

$$\mathrm{pr}_{i_v} : \pi_v \to \pi_{i_v}$$

be the projection map. Further we let

$$I_v^\circ := \{i_v \in I_v : \mathrm{pr}_{i_v}(\xi_v^\circ) \neq 0\}.$$

Note that if $i_v \in I_v^\circ$ then π_{i_v} is spherical in the sense that it contains a vector fixed by $H_v \cap M(\mathcal{O}_{F_v})$. (The author does not know if I_v° has only one element, and probably it does have more than one in general. This makes the following argument a bit delicate.) Let us define

$$I_S = \prod_{v \in S} I_v \quad \text{and} \quad I = \prod_v{}' I_v = \Big\{ i \in \prod_v I_v : i_v \in I_v^\circ \text{ for almost all } v \Big\},$$

where for each $i \in I$, we denote its v-th component by i_v. Namely I is the restricted direct product of I_v with respect to I_v°. For each $i \in I$, we write

$$i_S := (i_v)_{v \in S} \in I_S.$$

With this notation, we can write

$$\pi_S|_{H_S} = \bigoplus_{i \in I_S} \pi_{i_S},$$

where $\pi_{i_S} = \otimes_{v \in S} \pi_{i_v}$.

Now for each $i \in I$, let us define

$$\pi_i := \varinjlim_S \pi_{i_S}$$

by using $\mathrm{pr}_{i_v}(\xi_v^\circ)$ for our spherical vector for $i_v \in I_v^\circ$. Note that each π_i is an irreducible representation of H. To prove the lemma, it suffices to show we have an isomorphism

$$\varinjlim_S \pi_S|_{H_S} \cong \bigoplus_{i \in I} \varinjlim_S \pi_{i_S}, \tag{A.2}$$

namely $\pi|_H \cong \bigoplus_{i \in I} \pi_i$, which will show that $\pi|_H$ is completely reducible. To show there is such an isomorphism, first note that for each $S \subseteq S'$, the following diagram commutes:

$$\text{(A.3)} \qquad \begin{array}{ccc} \pi_S|_{H_S} & \longrightarrow & \pi_{S'}|_{H_{S'}} \\ \downarrow & & \downarrow \\ \bigoplus_{i_S \in I_S} \pi_{i_S} & \longrightarrow & \bigoplus_{i_{S'} \in I_{S'}} \pi_{i_{S'}} \end{array}$$

where the vertical arrows are actually an equality and the top horizontal arrow is given by tensoring with $\otimes_{v \in S' \setminus S} \xi_v^\circ$ and the bottom horizontal arrow is given as follows: for each $i_S \in I_S$, define a map

$$\pi_{i_S} \to \pi_{i_S} \otimes \bigoplus_{i_{S'\setminus S} \in I_{S' \setminus S}} \pi_{i_{S'\setminus S}} = \pi_{i_S} \otimes \bigotimes_{v \in S' \setminus S} \bigoplus_{i_v \in I_v} \pi_{i_v}$$

by

$$v_{i_S} \mapsto v_{i_S} \otimes \bigotimes_{v \in S' \setminus S} \bigoplus_{i_v \in I_v} \mathrm{pr}_{i_v}(\xi_v^\circ),$$

where recall that pr_{i_v} is the projection from π_v to π_{i_v}. Then the bottom horizontal arrow is given by combining all those maps for all the i_S.

Next, one can see that for each S there is an obvious injection

$$\text{(A.4)} \qquad \bigoplus_{i \in I_S} \pi_{i_S} \hookrightarrow \bigoplus_{i \in I} \varinjlim_S \pi_{i_S},$$

which makes the diagram

$$\begin{array}{ccc} \bigoplus_{i_S \in I_S} \pi_{i_S} & \longrightarrow & \bigoplus_{i_{S'} \in I_{S'}} \pi_{i_{S'}} \\ & \searrow \quad \swarrow & \\ & \bigoplus_{i \in I} \varinjlim_S \pi_{i_S} & \end{array}$$

commute. This diagram and the diagram (A.3) together with the universal property of $\varinjlim_S \pi_S|_{H_S}$ give a unique map

$$T : \pi|_H = \varinjlim_S \pi_S|_{H_S} \longrightarrow \bigoplus_{i \in I} \varinjlim_S \pi_{i_S} = \bigoplus_{i \in I} \pi_i,$$

which "commutes with the directed system". This map is injective because each $\varphi \in \pi$ is in $\varphi \in \pi_S$ for some S, which maps to $\bigoplus_{i \in I} \pi_i$ via the map in (A.4), and hence there is no kernel for T. Also one can see that T is surjective because if $\varphi_i \in \pi_i$, one can find S such that $\varphi_i \in \pi_{i_S}$, which comes from some vector π_S under the vertical map in (A.3). This completes the proof. $\square$

Remark A.5. As our last remark, let us mention that in [Takeda 2016, Lemma 5.1], it is erroneously claimed that the complete reducibility of a unitary automorphic representation of $\widetilde{\mathrm{GL}}_r(\mathbb{A})$ to $\widetilde{\mathrm{GL}}_r^{(n)}(\mathbb{A})$ follows from the admissibility and unitarity, but actually the restriction to $\widetilde{\mathrm{GL}}_r^{(n)}(\mathbb{A})$ is most likely not admissible, and hence the argument in the proof there does not work. But the proof of the above lemma, we hope, fixes the mistake.

Acknowledgements

This paper grew out of a question asked by Solomon Friedberg on the construction of the metaplectic tensor product of [Takeda 2016], when both he and the author were attending the conference in honor of Professor Roger Howe on the occasion of his 70th birthday. The author would like to thank him for asking the question and the organizers of the conference (James Cogdell, Ju-Lee Kim, Jian-Shu Li, David Manderscheid, Gregory Margulis, Cheng-Bo Zhu and Gregg Zuckerman) for creating a good opportunity for the interaction. Also a part of the paper was completed when the author was visiting the National University of Singapore in the spring of 2016, and he greatly appreciates their hospitality. Finally, he would like to thank the anonymous referee for his/her careful reading and letting him know about the independence of the multiplicities in Proposition 3.5.

The author is partially supported by Simons Foundations Collaboration Grants #35952 and Young Investigator Grant H98230-16-1-0312.

References

[Banks 1997] W. D. Banks, "Twisted symmetric-square L-functions and the nonexistence of Siegel zeros on GL(3)", *Duke Math. J.* **87**:2 (1997), 343–353. MR Zbl

[Brubaker and Friedberg 2015] B. Brubaker and S. Friedberg, "Whittaker coefficients of metaplectic Eisenstein series", *Geom. Funct. Anal.* **25**:4 (2015), 1180–1239. MR Zbl

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

[Cai 2016] Y. Cai, "Fourier coefficients for theta representations on covers of general linear groups", preprint, 2016. arXiv

[Chinta and Offen 2013] G. Chinta and O. Offen, "A metaplectic Casselman–Shalika formula for GL_r", *Amer. J. Math.* **135**:2 (2013), 403–441. MR Zbl

[Cogdell 2004] J. W. Cogdell, "Lectures on L-functions, converse theorems, and functoriality for GL_n", pp. 1–96 in *Lectures on automorphic L-functions*, edited by J. W. Cogdell et al., Fields Inst. Monogr. **20**, Amer. Math. Soc., Providence, RI, 2004. MR

[Deitmar and Echterhoff 2009] A. Deitmar and S. Echterhoff, *Principles of harmonic analysis*, Springer, New York, 2009. MR Zbl

[Friedberg and Ginzburg 2016] S. Friedberg and D. Ginzburg, "Criteria for the existence of cuspidal theta representations", *Res. Number Theory* **2** (2016), art. id. 16, 16 pp. MR Zbl

[Gan 2016] W. T. Gan, "The metaplectic tensor product as an instance of Langlands functoriality", preprint, 2016, available at http://www.math.nus.edu.sg/~matgwt/metaplectic-tensor.pdf.

[Gan and Gao 2014] W. T. Gan and F. Gao, "The Langlands–Weissman program for Brylinski–Deligne extensions", 2014. To appear in *Astérisque*. arXiv

[Gelbart and Knapp 1982] S. S. Gelbart and A. W. Knapp, "L-indistinguishability and R groups for the special linear group", *Adv. in Math.* **43**:2 (1982), 101–121. MR Zbl

[Kable 2001] A. C. Kable, "The tensor product of exceptional representations on the general linear group", *Ann. Sci. École Norm. Sup.* (4) **34**:5 (2001), 741–769. MR Zbl

[Kazhdan and Patterson 1984] D. A. Kazhdan and S. J. Patterson, "Metaplectic forms", *Inst. Hautes Études Sci. Publ. Math.* 59 (1984), 35–142. MR Zbl

[Mezo 2004] P. Mezo, "Metaplectic tensor products for irreducible representations", *Pacific J. Math.* **215**:1 (2004), 85–96. MR Zbl

[Suzuki 1997] T. Suzuki, "Metaplectic Eisenstein series and the Bump–Hoffstein conjecture", *Duke Math. J.* **90**:3 (1997), 577–630. MR Zbl

[Suzuki 1998] T. Suzuki, "Distinguished representations of metaplectic groups", *Amer. J. Math.* **120**:4 (1998), 723–755. MR Zbl

[Takeda 2014] S. Takeda, "The twisted symmetric square L-function of $\mathrm{GL}(r)$", *Duke Math. J.* **163**:1 (2014), 175–266. MR Zbl

[Takeda 2015] S. Takeda, "On a certain metaplectic Eisenstein series and the twisted symmetric square L-function", *Math. Z.* **281**:1-2 (2015), 103–157. MR Zbl

[Takeda 2016] S. Takeda, "Metaplectic tensor products for automorphic representation of $\widetilde{\mathrm{GL}}(r)$", *Canad. J. Math.* **68**:1 (2016), 179–240. MR Zbl

[Weissman 2016] M. H. Weissman, "L-groups and parameters for covering groups", 2016. To appear in *Astérisque*. arXiv

Received October 17, 2016. Revised February 10, 2017.

SHUICHIRO TAKEDA
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MISSOURI
202 MATH SCIENCE BLDG
COLUMBIA, MO 65211
UNITED STATES

takedas@missouri.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.231

ON RELATIVE RATIONAL CHAIN CONNECTEDNESS OF THREEFOLDS WITH ANTI-BIG CANONICAL DIVISORS IN POSITIVE CHARACTERISTICS

YUAN WANG

Let X be a projective klt threefold over an algebraically closed field of positive characteristic, and $f : X \to Y$ a morphism from X to a projective variety Y of dimension 1 or 2. We study how bigness and relative bigness of $-K_X$ influences the rational chain connectedness of X and fibers of f, respectively. We construct a canonical bundle formula and use it as well as the minimal model program to prove two results in this context.

1. Introduction

It is widely recognized that the geometry of a higher-dimensional variety is closely related to the geometry of rational curves on it. A classical result by Campana [1992] and Kollár, Miyaoka and Mori [Kollár et al. 1992] says that smooth Fano varieties are rationally connected in characteristic zero and are rationally chain connected in positive characteristics. This was generalized in characteristic zero in [Zhang 2006; Hacon and McKernan 2007]. More recently, using the minimal model program of [Hacon and Xu 2015; Birkar 2016], Gongyo, Li, Patakfalvi, Schwede, Tanaka and Zong [Gongyo et al. 2015a] proved that projective globally F-regular threefolds in characteristic ≥ 11 are rationally chain connected and this was later generalized to threefolds of log Fano type by Gongyo, Nakamura and Tanaka [Gongyo et al. 2015b].

The main result of Hacon and McKernan is as follows:

Theorem 1.1 [Hacon and McKernan 2007, Theorem 1.2]. *Let (X, Δ) be a log pair, and let $f : X \to S$ be a proper morphism such that $-K_X$ is relatively big and $-(K_X + \Delta)$ is relatively semiample. Let $g : Y \to X$ be any birational morphism. Then the connected components of every fiber of $f \circ g$ are rationally chain connected modulo the inverse image of the locus of log canonical singularities of (X, Δ).*

The author was supported in part by the FRG grant DMS-#1265261.
MSC2010: primary 14M22; secondary 14E30.
Keywords: rational chain connectedness, positive characteristic, minimal model program, weak positivity, canonical bundle formula.

In this paper we prove a theorem similar to Theorem 1.1 for morphisms from a klt threefold to a variety of dimension ≥ 1. More precisely, we have

Theorem 3.1. *Let X be a normal $\mathbb{Q}$-factorial threefold over an algebraically closed field k of characteristic ≥ 7 and (X, D) a klt pair. Let $f : X \to Z$ be a proper morphism such that $f_*\mathcal{O}_X = \mathcal{O}_Z$, $\dim(Z)$ is 1 or 2, Z is klt, $-K_X$ is relatively big, $-(K_X + D)$ is relatively semiample, and (X_z, D_z) is klt for general $z \in Z$. Let $g : Y \to X$ be any birational morphism. Then the connected components of every fiber of $f \circ g$ are rationally chain connected.*

Motivated by Theorem 3.1, we construct a global version of rational chain connectedness for threefolds.

Theorem 5.1. *Let X be a projective threefold over an algebraically closed field k of characteristic $p > 0$, $f : X \to Y$ a projective surjective morphism from X to a projective variety Y such that $f_*\mathcal{O}_X = \mathcal{O}_Y$. Let D be an effective $\mathbb{Q}$-divisor, and $X_{\bar{\eta}}$ the geometric generic fiber of f. Assume that the following conditions hold*:

(1) *(X, D) is klt, $-K_X$ is big, and f-ample, $K_X + D \sim_{\mathbb{Q}} 0$, and the general fibers of f are smooth.*

(2) *$p > 2/\delta$, where δ is the minimum nonzero coefficient of D.*

(3) *$D = E + f^*L$ where E is an effective $\mathbb{Q}$-Cartier divisor such that $p \nmid \operatorname{ind}(E)$, $(X_{\bar{\eta}}, E|_{X_{\bar{\eta}}})$ is globally F-split, and L is a big $\mathbb{Q}$-divisor on Y.*

(4) *$\dim(Y)$ is 1 or 2.*

Then X is rationally chain connected.

Here $\operatorname{ind}(E)$ means the Cartier index of E.

The main ingredients of the proofs of Theorems 3.1 and 5.1 are the minimal model program constructed in [Hacon and Xu 2015; Birkar 2016; Gongyo et al. 2015a]; some facts, especially [Gongyo et al. 2015a, Theorem 2.1]; some positivity results [Patakfalvi 2014; Ejiri 2015]; a canonical bundle formula constructed in Section 4 in the spirit of [Prokhorov and Shokurov 2009]. Note that condition (3) in Theorem 5.1 is used in order to apply the result [Ejiri 2015, Theorem 1.1] to deduce that $-K_Y$ is big, and to apply Theorem 4.3 when $\dim Y = 2$. This creates enough rational curves on Y. Note that by [Ejiri 2015, Example 3.4], $(X_{\bar{\eta}}, E|_{X_{\bar{\eta}}})$ being globally F-split is equivalent to $S^0(X_{\bar{\eta}}, E|_{X_{\bar{\eta}}}, \mathcal{O}_{X_{\bar{\eta}}}) = H^0(X_{\bar{\eta}}, \mathcal{O}_{X_{\bar{\eta}}})$.

We note that although its proof is independent, Theorem 3.1 is implied by [Gongyo et al. 2015b, Theorem 4.1], which was put on arXiv before this paper. The proof of that result relies on the minimal model program in dimension 3 in positive characteristic, which is only established in characteristic ≥ 7 so far. On the other hand, Theorem 5.1 covers some cases in characteristic < 7. It does not rely on the minimal model program and is not implied by [Gongyo et al. 2015b].

2. Preliminaries

We work over an algebraically closed field k of characteristic $p > 0$.

Preliminaries on rational connected varieties and the minimal model program.

Definition 2.1. For a variety X and a $\mathbb{Q}$-Weil divisor on X such that $K_X + \Delta$ is $\mathbb{Q}$-Cartier. Let $f : Y \to X$ be a log resolution of (X, Δ) and write

$$K_Y = f^*(K_X + \Delta) + \sum_i a_i E_i$$

where E_i is a prime divisor. We say that (X, Δ) is

- *sub-Kawamata log terminal* (*sub-klt* for short) if $a_i > -1$ for any i;
- *Kawamata log terminal* (*klt* for short) if $a_i > -1$ for any i and $\Delta \geq 0$;
- *log canonical* if $a_i \geq -1$ for any i and $\Delta \geq 0$.

Definition 2.2. [Kollár 1996, IV.3.2] Suppose that X is a variety over k.

(1) We say that X is *rationally chain connected* (*RCC*) if there is a family of proper and connected algebraic curves $g : U \to Y$ whose geometric fibers have only rational components and there is a cycle morphism $u : U \to X$ such that $u^{(2)} : U \times_Y U \to X \times_k X$ is dominant.

(2) We say that X is *rationally connected (RC)* if (1) holds and moreover the geometric fibers of g in (1) are irreducible.

Proposition 2.3. *Let X be a klt $\mathbb{Q}$-factorial threefold over an algebraically closed field k and* $\mathrm{char}(k) \geq 7$. *Let $g : W \to X$ be a log resolution and assume that $K_W + E = g^*K_X + B$, where E and B are exceptional divisors and the coefficients in E are all* 1. *Then relative minimal model for* (W, E) *over X exists. Denote this process by*

$$W = W_0 \overset{f_0}{\dashrightarrow} W_1 \overset{f_1}{\dashrightarrow} \cdots \overset{f_{N-1}}{\dashrightarrow} W_N = W'.$$

Then we actually have $W' = X$. Moreover if we have a morphism $h : X \to Y$ such that every fiber of h is RCC, then every fiber of $h \circ g$ is RCC.

Proof. The existence of this minimal model program is by [Gongyo et al. 2015a, Theorem 3.2]. So we have a morphism $g' : W' \to X$ and we want to show that g' is the identity. Denote the strict transform of E by E', then $K_{W'} + E' = g'^*K_X + B'$ for some exceptional $\mathbb{Q}$-divisor B'. By construction of the minimal model program we know that $g'^*K_X + B'$ is nef over X which means that B' is g'-nef and since X is klt the support of B' is the whole exceptional locus of g'. So we can get that $B' = 0$ by the negativity lemma, and since X is $\mathbb{Q}$-factorial we will get $W' = X$.

The proof of the last statement follows the proof of Proposition 3.6 in the same reference. Without loss of generality we can do a base change and assume that the base field k is uncountable. Define F in the following way: if f_i is a divisorial contraction, then let $E_0 = E$, $E_{i+1} = f_{i,*}E_i$, and F be an arbitrary component of E_i; if f_i is a flip and C is any flipping curve then let F be a component of E_i that contains C. Let $K_F + \Delta_F := \bigl(K_{W_i} + E_i - \frac{1}{n}(E_i - F)\bigr)\big|_F$, where $n \gg 0$. By assumption $K_{W_i} + E_i - \frac{1}{n}(E_i - F)$ is plt, then by adjunction $K_F + \Delta_F$ is klt, hence by [Tanaka 2014, Theorem 14.4] F is $\mathbb{Q}$-factorial. We also know that $-(K_{W_i} + E_i)$ is f_i-ample by assumption, then $-(K_F + \Delta_F)$ is ample. Moreover by [Prokhorov 2001, Corollary 2.2.8] the coefficients of Δ_F are in the standard set $\bigl\{1 - \frac{1}{n} \bigm| n \in \mathbb{N}\bigr\}$. Let $\tilde{F}$ be the normalization of F. Then by [Hacon and Xu 2015, Theorem 3.1] we know that $(\tilde{F}, \Delta_{\tilde{F}})$ is strongly F-regular and by Theorem 4.1 from that reference F is a normal surface.

Next we consider three cases.

Case 1: If f_i is a divisorial contraction and the exceptional divisor is contracted to a point, then since $-(K_F + \Delta_F)$ is ample, by [Kawamata 1994, Lemma 2.2] F is a rational surface, in particular it is rationally connected.

Case 2: If f_i is a divisorial contraction and the exceptional divisor is contracted to a curve, then let $p : F \to B$ be the Stein factorization of $f_i|_F$. By assumption $-(K_F + \Delta_F)$ is f_i-ample, so it is p-ample. Then for a general fiber D of p,

$$(K_F + D) \cdot D = (K_F + \Delta_F + D - \Delta_F) \cdot D = (K_F + \Delta_F) \cdot D - \Delta_F \cdot D < 0.$$

Here D is reduced and irreducible by [Bădescu 2001, Theorem 7.1], hence by [Tanaka 2014, Theorem 5.3] $D \cong \mathbb{P}^1$. Therefore every component of every fiber of f_i is a rational curve.

Case 3: If f_i is a flip, then let C be an arbitrary flipping curve. By assumption we have $(K_F + \Delta_F) \cdot C < 0$, $C^2 < 0$, and $0 \le \operatorname{coeff}_C \Delta_F < 1$, so $(K_F + C) \cdot C < 0$. Again by [op. cit., Theorem 5.3] $C \cong \mathbb{P}^1$.

We denote a fiber of h over $y \in Y$ by $F_{X,y}$. There is a morphism from W_i to Y for every i, and we denote the fiber of this morphism over y as $F_{W_i,y}$. Then there is a rational map $F_{W_i,y} \dashrightarrow F_{W_{i+1},y}$. From the above Cases 1–3 we see that compared to $F_{W_i,y}$, there are only rational curves or a rational surface generated in $F_{W_{i+1},y}$. So the RCC-ness of $F_{W_{i+1},y}$ implies the RCC-ness of $F_{W_i,y}$. By assumption $F_{X,y}$ is RCC, so $F_{W,y}$ is RCC. □

Proposition 2.4. *Let X be a klt $\mathbb{Q}$-factorial threefold over an algebraically closed field k and* $\operatorname{char}(k) \ge 7$. *Let $f : X \to Y$ be a morphism from X to a normal surface Y. Suppose we run a K_X-minimal model program and it terminates at $g : X' \to Y$. If every fiber of g is RCC then every fiber of f is RCC.*

Proof. This can be easily deduced from Proposition 2.3 by taking a common resolution of X and X'. The proof of [Gongyo et al. 2015a, Proposition 3.6] works as well. □

Preliminaries on F-singularities. In this article, for a proper variety X, a $\mathbb{Q}$-divisor Δ, and the line bundle M, we will use the concepts of *strongly F-regular*, the *non-F-pure ideal* $\sigma(X, \Delta)$ and $S^0(X, \sigma(X, \Delta) \otimes M)$. The definitions of these can be found in many papers related to F-singularities, e.g., [Hacon and Xu 2015]. For a pair (X, Δ) where Δ is a $\mathbb{Q}$-Cartier divisor we also follow the definition of *globally F-split* in [Ejiri 2015].

Lemma 2.5. *Let X be a surface, D an effective $\mathbb{Q}$-divisor on X, $f : X \to C$ a morphism from X to a smooth curve C, and (X_c, D_c) is a strongly F-regular pair for general $c \in C$. Assume that $-K_X$ is big, $K_X + D \sim_{\mathbb{Q}} 0$, then $C \cong \mathbb{P}^1$.*

Proof. By Kodaira's lemma we can write $D \sim_{\mathbb{Q}} \epsilon f^*H + E$ where H is an ample $\mathbb{Q}$-divisor on C, $0 < \epsilon \in \mathbb{Q}$, E is an effective $\mathbb{Q}$-divisor on X and (X_c, E_c) is also strongly F-regular for general $c \in C$ (since X_c is a curve). Suppose that C is not isomorphic to $\mathbb{P}^1$. We know that $K_{X/C} + E \sim_{\mathbb{Q}} f^*(-K_C - \epsilon H)$ is f-nef and $K_{X_c} + E_c$ is semiample for general $c \in C$, so by [Patakfalvi 2014, Theorem 3.16], $K_{X/C} + E = K_X - f^*K_C + E$ is nef. Since we have assumed that $g(C) > 0$ we have that $K_X + E$ is nef. However this is impossible since $K_X + E \sim_{\mathbb{Q}} -\epsilon f^*H$ where H is ample and $\epsilon > 0$. □

Weak positivity. Let Y be a nonsingular projective variety, $\mathcal{F}$ a torsion-free coherent sheaf on Y. We take $i : \hat{Y} \to Y$ to be the biggest open subvariety such that $\mathcal{F}|_{\hat{Y}}$ is locally free. Let $\hat{S}^k(\mathcal{F}) := i_* S^k(i^*\mathcal{F})$.

Definition 2.6 [Viehweg 1983, Definition 1.2]. We call $\mathcal{F}$ *weakly positive*, if there is an open subset $U \subseteq Y$ such that for every ample line bundle $\mathcal{H}$ on Y and every positive number α there exists some positive number β such that $\hat{S}^{\alpha\cdot\beta}(\mathcal{F}) \otimes \mathcal{H}^\beta$ is generated by global sections over U.

Lemma 2.7. *Weakly positive line bundles are nef.*

Proof. This easily follows from Definition 2.6. □

3. Relative rational chain connectedness

In this section we prove the following

Theorem 3.1. *Let X be a normal $\mathbb{Q}$-factorial threefold over an algebraically closed field k of characteristic ≥ 7 and (X, D) a klt pair. Let $f : X \to Z$ be a proper morphism such that $f_*\mathcal{O}_X = \mathcal{O}_Z$, $\dim(Z)$ is 1 or 2, Z is klt, $-K_X$ is relatively big, $-(K_X + D)$ is relatively semiample, and (X_z, D_z) is klt for general $z \in Z$. Let $g : Y \to X$ be any birational morphism. Then the connected components of every fiber of $f \circ g$ are rationally chain connected.*

Remark 3.2. In Theorem 3.1, if $\dim Z = 2$, by adjunction and a theorem of Tate (see [Liedtke 2013, Theorem 5.1]) we have that the generic fiber of f is smooth. So in this case the condition that (X_z, D_z) is klt for general $z \in Z$ is not necessary.

Proof. First we observe that (X_z, D_z) being klt implies that X_z is normal (in particular reduced) and irreducible.

Next we prove that if every fiber of f is RCC, then every fiber of $f \circ g$ is RCC. We take a log resolution of Y and denote it by $p : Y' \to Y$ and let $q = g \circ p$. If $K_{Y'} = q^* K_X + \tilde{B}$ then $K_{Y'} - \tilde{B} = q^* K_X$ and the coefficients of $-\tilde{B}$ are < 1. Then we can add another effective divisor to make all the coefficients 1, and we denote this divisor by $\tilde{E}$. Now we run a relative $(K_{Y'} + \tilde{E})$-minimal model program of Y' over X. By Proposition 2.3 we see that if every fiber of f is RCC then every fiber of $f \circ g \circ p$ is RCC, hence every fiber of $f \circ g$ is RCC.

Therefore it suffices to show that every fiber of f is RCC. We consider the cases of $\dim(Z) = 2$ and $\dim(Z) = 1$, respectively.

Case 1: $\dim(Z) = 2$. If $\dim(Z) = 2$ then a general fiber of f being normal and $-K_X$ being relatively big implies that a general fiber of f is a smooth rational curve. Next we run a relative minimal model program over Z and denote this process as

$$X = X_0 \overset{f_0}{\dashrightarrow} X_1 \overset{f_1}{\dashrightarrow} \cdots \overset{f_{N-1}}{\dashrightarrow} X_n = X'.$$

Since $-K_X$ is relatively big we end up with a Mori fiber space $X' \xrightarrow{h} Z' \xrightarrow{p} Z$ where Z' is also a surface. Then the general fibers of h are rational curves. Moreover since $p_* \mathcal{O}_{Z'} = \mathcal{O}_Z$ we know that p is birational.

Now we prove that h is equidimensional. Suppose that this is not the case, then there is a fiber $\tilde{F}$ of h over a point $\tilde{z} \in Z'$ which contains a 2-dimensional irreducible component. If $\tilde{F}$ is reducible then let $\tilde{F}_1$ be a 2-dimensional component of $\tilde{F}$ and $\tilde{F}_2$ another component which intersects $\tilde{F}_1$. We can choose a curve $\tilde{C}_2 \subseteq \tilde{F}_2$ such that $\tilde{F}_1 \cdot \tilde{C}_2 > 0$. On the other hand if we take a general point $z' \in Z'$ then $h^{-1}(z')$ is an irreducible curve and $h^{-1}(z') \cdot \tilde{F}_2 = 0$. This contradicts the fact that $\rho(X'/Z') = 1$. If $\tilde{F}$ is irreducible, by Bertini's theorem we have a very ample divisor $H \subset X'$ such that $H \cap \tilde{F}$ is an irreducible curve which we denote by $\tilde{C}$. We do the Stein factorization of $h|_H$ and denote the process as

$$H \xrightarrow{h_1} Z'' \xrightarrow{h_2} Z',$$

then h_1 is birational and $\tilde{C}$ is an exceptional curve of h_1. After possibly replacing Z'' by its normalization we can assume that Z'' is normal. Now $\tilde{F} \cdot \tilde{C}$ is equal to $\tilde{C}^2$, viewed as the self-intersection of $\tilde{C}$ in H, so by the negativity lemma it is negative. On the other hand we can still take a general point $z' \in Z'$ as above such that $h^{-1}(z') \cdot \tilde{F} = 0$. This also contradicts the fact that $\rho(X'/Z') = 1$.

Since h is equidimensional, by [Debarre 2001, Lemma 3.7] the components of every fiber of h are rational curves. Then by Proposition 2.4 every fiber of f is RCC.

Case 2: $\dim(Z) = 1$. Without loss of generality we can do a base change and assume that the base field k is uncountable. By passing to the normalization of Z we can assume that Z is smooth. Then since every closed point of Z is a Cartier divisor, every fiber of f is also Cartier, hence f is equidimensional.

We first show that the general fibers of f are rationally chain connected. Let F be a general fiber of f. Since we assume that $(F, D|_F)$ is klt, by adjunction

$$K_X|_F \equiv_{\text{num}} (K_X + F)|_F = K_F + \mathrm{Diff}_F(0),$$

where $\mathrm{Diff}_F(0) \geq 0$; see [Kollár 1992, Proposition-Definition 16.5]. Therefore, $-(K_F + \mathrm{Diff}_F(0))$ is big, hence $-K_F$ as well. As a result, $\kappa(F) = -\infty$ and F is birationally ruled by classification of surfaces. To prove that the general fibers of f are RCC it suffices to prove that F is rational. By assumption $-(K_F + D|_F) = -(K_X + D)|_F$ is semiample, so there exists an effective $\mathbb{Q}$-divisor H such that $H \sim_{\mathbb{Q}} -(K_F + D|_F)$ and $(F, D|_F + H)$ is klt. We define $\Delta := D|_F + H$. Let $\pi : F' \to F$ be a minimal resolution of $(F, \mathrm{Diff}_F(0))$, then F' maps to a ruled surface F'' over a smooth curve B via a sequence of blowdowns and we denote the morphism by ψ. The situation is as follows:

$$\begin{tikzcd} & F' \arrow[dl, "\pi"'] \arrow[dr, "\psi"] & \\ F & & F'' \arrow[d, "q"] \\ & & B \end{tikzcd}$$

Since (F, Δ) is klt, by [Kollár and Mori 1998, Theorem 4.7] π and ψ only contract copies of $\mathbb{P}^1$. So F is RCC if and only if F'' is RCC. Define Δ'' on F'' via

$$K_{F''} + \Delta'' = \psi_* \pi^*(K_F + \Delta).$$

Then (F, Δ) being klt implies that (F'', Δ'') is klt.

We denote a general fiber of q by R. By construction $R \cong \mathbb{P}^1$, so we know that $(R, \Delta''|_R)$ is klt and hence strongly F-regular. Then by applying Lemma 2.5 on F'' we know that $B = \mathbb{P}^1$. So F is rational. Therefore we have proven that the general fibers of f are RCC.

Since we have assumed that the base field k is uncountable, by [Kollár 1996, Chapter IV, Corollary 3.5.2] we know that every fiber of f is RCC. $\square$

4. A canonical bundle formula for threefolds in positive characteristics

In this section following the idea of the proof of [Prokhorov and Shokurov 2009] we construct a canonical bundle formula in characteristic p for a morphism from a threefold to a surface, whose general fibers are $\mathbb{P}^1$. There are similar constructions in [Cascini et al. 2015, 6.7; Das and Hacon 2016, Theorem 4.8].

Let $\overline{\mathcal{M}}_{0,n}$ be the moduli space of n-pointed stable curves of genus 0, let $f_{0,n} : \overline{\mathcal{U}}_{0,n} \to \overline{\mathcal{M}}_{0,n}$ be the universal family, and let $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_n$ be the sections of $f_{0,n}$ which correspond to the marked points. Let $d_j\,(j = 1, 2, \dots, n)$ be the rational numbers such that $0 < d_j \le 1$ for all j, $\sum_j d_j = 2$, and $\mathcal{D} = \sum_j d_j \mathcal{P}_j$.

Lemma 4.1 [Das and Hacon 2016, Lemma 4.6; Kawamata 1997, Theorem 2].

(1) *There exists a smooth projective variety $\mathcal{U}^*_{0,n}$, a $\mathbb{P}^1$-bundle $g_{0,n} : \mathcal{U}^*_{0,n} \to \overline{\mathcal{M}}_{0,n}$, and a sequence of blowups with smooth centers*

$$\overline{\mathcal{U}}_{0,n} = \mathcal{U}^{(1)} \xrightarrow{\sigma_2} \mathcal{U}^{(2)} \xrightarrow{\sigma_3} \cdots \xrightarrow{\sigma_{n-2}} \mathcal{U}^{(n-2)} = \mathcal{U}^*_{0,n}$$

(2) *Let $\sigma : \overline{\mathcal{U}_{0,n}} \to \mathcal{U}^*_{0,n}$ be the induced morphism, and let $\mathcal{D}^* = \sigma_*\mathcal{D}$. Then $K_{\overline{\mathcal{U}}_{0,n}} + \mathcal{D} - \sigma^*(K_{\mathcal{U}^*_{0,n}} + \mathcal{D}^*)$ is effective.*

(3) *There exists a semiample $\mathbb{Q}$-divisor $\mathcal{L}$ on $\overline{\mathcal{M}}_{0,n}$ such that*

$$K_{\mathcal{U}^*_{0,n}} + \mathcal{D}^* \sim_{\mathbb{Q}} g^*_{0,n}(K_{\overline{\mathcal{M}}_{0,n}} + \mathcal{L}).$$

Definition 4.2. Let $f : X \to Y$ be a surjective proper morphism between two normal varieties and $K_X + D \sim_{\mathbb{Q}} f^*L$, where D is a boundary divisor on X and L is a $\mathbb{Q}$-Cartier $\mathbb{Q}$-divisor on Y. Let (X, D) be log canonical near the generic fiber of f, i.e., $(f^{-1}U, D|_{f^{-1}U})$ is log canonical for some Zariski dense open subset $U \subseteq Y$. We define

$$D_{\text{div}} := \sum (1 - c_Q) Q,$$

where $Q \subset Z$ are prime Weil divisors on Z and

$c_Q = \sup\{c \in \mathbb{R} : (X, D + cf^*Q)\text{is log canonical over the generic point } \eta_Q \text{ of } Q\}.$

Next we define

$$D_{\text{mod}} := L - K_Y - D_{\text{div}},$$

so $K_X + D = f^*(K_Y + D_{\text{div}} + D_{\text{mod}})$.

Theorem 4.3. *Let $f : X \to Y$ be a proper surjective morphism, where X is a normal threefold and Y is a normal surface over an algebraically closed field k of characteristic $p > 0$. Assume that $Q = \sum_i Q_i$ is a divisor on Y such that f is smooth over $(Y - \operatorname{Supp}(Q))$ with fibers isomorphic to $\mathbb{P}^1$. Let $D = \sum_i d_i D_i$ be a $\mathbb{Q}$-divisor on X where $d_i = 0$ is allowed, which satisfies the following conditions*:

(1) $(X, D \geq 0)$ *is klt on a general fiber of* f.

(2) *Suppose* $D = D^h + D^v$ *where* D^h *is the horizontal part and* D^v *is the vertical part of* D. *Then* $p = \mathrm{char}(k) > 2/\delta$, *where* δ *is the minimum nonzero coefficient of* D^h.

(3) $K_X + D \sim_{\mathbb{Q}} f^*(K_Y + M)$ *for some* $\mathbb{Q}$*-Cartier divisor* M *on* Y.

Then we have that D_{mod} *is* $\mathbb{Q}$*-linearly equivalent to an effective* $\mathbb{Q}$*-divisor. Here* D_{mod} *is defined as in Definition 4.2. Moreover if* (X, D) *is klt then there exists an effective* $\mathbb{Q}$*-divisor* $\bar{\mathcal{D}}_{\mathrm{mod}}$ *on* Y *such that* $\bar{\mathcal{D}}_{\mathrm{mod}} \sim_{\mathbb{Q}} D_{\mathrm{mod}}$ *and* $(Y, D_{\mathrm{div}} + \bar{\mathcal{D}}_{\mathrm{mod}})$ *is klt.*

Proof. First we reduce the problem to the case where all components of D^h are sections. Let D_{i_0} be a horizontal component of D and $D_{i_0} \to D_{i_0}^\flat \to Y$ be the Stein factorization of $f|_{D_{i_0}}$. Let $Y' \to D_{i_0}^\flat$ be the normalization of $D_{i_0}^\flat$, then $Y' \to Y$ is a finite surjective morphism of normal surfaces. Let X' be the normalization of the component of $X \times_Y Y'$ dominating Y.

$$\begin{tikzcd} X & X' \arrow[l, "\nu'"'] \arrow[d, "f'"] \\ Y \arrow[from=u, "f"'] & Y' \arrow[l, "\nu"] \end{tikzcd}$$

Let $m = \deg(\mu : Y' \to Y)$ and l be a general fiber of f. Then

$$m = D_i \cdot l \leq \frac{1}{d_i}(D \cdot l) = \frac{1}{d_i}(-K_X \cdot l) = \frac{2}{d_i} \leq \frac{2}{\delta} < \mathrm{char}(k). \tag{4-1}$$

Therefore ν is a separable and tamely ramified morphism.

Let D' be the log pullback of D under ν', i.e.,

$$K_{X'} + D' = \nu'^*(K_X + D).$$

More precisely by [Kollár 1992, 20.2],

$$D' = \sum_{i,j} d'_{ij} D'_{ij}, \qquad \nu'(D'_{ij}) = D_i, \qquad d'_{ij} = 1 - (1 - d_i)e_{ij},$$

where e_{ij} is the ramification indices along D'_{ij}.

By construction X dominates Y. Also, since ν is étale over a dense open subset of Y, say $\nu^{-1}U \to U$, and étale morphisms are stable under base change, the map $(f' \circ \nu)^{-1}U \to f^{-1}U$ is étale. Thus the ramification locus Λ of ν' does not contain any horizontal divisor f', i.e., $f'(\Lambda) \neq Y'$. Therefore D' is a boundary near the generic fiber of f', i.e., D'^h is effective. We observe that the coefficients of D'^h can be computed by intersecting with a general fiber of $f' : X' \to Y'$, hence they are equal to the coefficient of $D^h \subseteq X$. Thus the condition $p > 2/\delta$ remains true for D' on X'.

After finitely many such base changes we get a family $f'' : X'' \to Y''$, such that all of the horizontal components of D'' are rational sections of f''. Here D'' is the log pullback of D via the induced finite morphism $\alpha : X'' \to X$, i.e., $K_{X''} + D'' = \alpha^*(K_X + D)$.

By construction of $\overline{\mathcal{M}}_{0,n}$ there is a generically finite rational map $Y'' \dashrightarrow \overline{\mathcal{M}}_{0,n}$. Let $\beta_0 : \tilde{Y} \to Y''$ be a morphism that resolves the indeterminacies of $Y'' \to \overline{\mathcal{M}}_{0,n}$ and $\tilde{X}$ the normalization of $X'' \times_{Y''} \tilde{Y}$. We have a morphism $\tilde{Y} \to \overline{\mathcal{M}}_{0,n}$ and let $\hat{X} = \tilde{Y} \times_{\overline{\mathcal{M}}_{0,n}} \overline{\mathcal{U}}_{0,n}$. Let $X^\sharp$ be a common resolution of $\tilde{X}$ and $\hat{X}$. We have the following diagram:

$$\begin{tikzcd}
 & & & X^\sharp \arrow[dlll, bend right, "\pi"'] \arrow[dl, "\lambda"'] \arrow[dr, "\mu"] \arrow[dd, "f^\sharp"] & & & \\
X \arrow[d, "f"'] & X'' \arrow[l, "\alpha"] \arrow[d, "f''"'] & \tilde{X} \arrow[l, "\beta"] \arrow[ll, bend right, "\psi"'] \arrow[rr, dashed] \arrow[dr, "\tilde{f}"'] & & \hat{X} \arrow[r] \arrow[dl, "\hat{f}"] \arrow[rr, bend left, "\hat{\phi}"] & \overline{\mathcal{U}}_{0,n} \arrow[r, "\sigma"] \arrow[d, "f_{0,n}"'] & \mathcal{U}^*_{0,n} \arrow[dl, "g_{0,n}"] \\
Y & Y'' \arrow[l, "\alpha_0"] & & \tilde{Y} \arrow[ll, "\beta_0"] \arrow[rr, "\phi_0"] \arrow[lll, bend left, "\psi_0"] & & \overline{\mathcal{M}}_{0,n} &
\end{tikzcd}$$

Let $D^\sharp$ and $\hat{D}$ be $\mathbb{Q}$-divisors on $X^\sharp$ and $\hat{X}$ respectively, defined by

$$K_{X^\sharp} + D^\sharp = \pi^*(K_X + D) \quad \text{and} \quad K_{\hat{X}} + \hat{D} = \mu_*(K_{X^\sharp} + D^\sharp).$$

We also define D''_{mod} and D''_{div} on Y'' for (X'', D'') as in Definition 4.2, such that

$$K_{X''} + D'' = f''^*(K_{Y''} + D''_{\mathrm{mod}} + D''_{\mathrm{div}}),$$

and we define $\tilde{D}_{\mathrm{mod}}$ and $\tilde{D}_{\mathrm{div}}$ on $\tilde{Y}$ in a similar way. Since $K_{X^\sharp} + D^\sharp$ is the pullback of some $\mathbb{Q}$-divisor from the base $\tilde{Y}$ we get

$$K_{X^\sharp} + D^\sharp = \mu^*(K_{\hat{X}} + \hat{D}).$$

Since D_{div} does not depend on the birational modification of the family [Prokhorov and Shokurov 2009, Remark 7.3], we will define it with respect to $\hat{f} : \hat{X} \to \tilde{Y}$.

Since $\hat{\phi}$ is generically finite and $\mathcal{D}^*$ is horizontal it follows that $\hat{\phi}^*\mathcal{D}^*$ is horizontal too. Since $\hat{D}^h$ is also horizontal,

$$\hat{D}^h = \hat{\phi}^*\mathcal{D}^*. \tag{4-2}$$

From the construction of the map $\sigma : \overline{\mathcal{U}}_{0,n} \to \mathcal{U}^*_{0,n}$ we see that $(F, \mathcal{D}^*|_F)$ is log canonical for any fiber F of $g_{0,n} : \mathcal{U}^*_{0,n} \to \overline{\mathcal{M}}_{0,n}$. Since the fibers of $\hat{f} : \hat{X} \to \tilde{Y}$ are isomorphic to the fiber of $g_{0,n}$, we see that $(\hat{F}, \hat{D}^h|_{\hat{F}})$ is also log canonical, where $\hat{F}$ is any fiber of $\hat{f}$. Let $\hat{D}^v_i$ be a component of $\hat{D}^v$ and η the generic point of $\hat{f}(\hat{D}^v_i)$.

Then by inversion of adjunction we know that $(\hat{X}_\eta, (\hat{D}_i^v + \hat{D}^h)|_\eta)$ is log canonical. Since the fibers of $\hat{f}$ are reduced, the log canonical threshold of $(\hat{X}, \hat{D}; \hat{D}_i^v)$ over the generic point of $\hat{D}_i^v$ is $(1 - \operatorname{coeff}_{\hat{D}_i^v} \hat{D})$. Hence we get $\hat{D}^v = \hat{f}^* \tilde{D}_{\text{div}}$. Note that the coefficients of $\hat{D}^v$ can be > 1. By definition of $\tilde{D}_{\text{mod}}$ we have

$$K_{\hat{X}} + \hat{D}^h \sim_{\mathbb{Q}} \hat{f}^*(K_{\tilde{Y}} + \tilde{D}_{\text{mod}}). \tag{4-3}$$

Then

$$K_{\hat{X}} + \hat{D}^h - f^*(K_{\tilde{Y}} + \phi_0^*\mathcal{L}) = K_{\hat{X}/\tilde{Y}} + \hat{D}^h - \hat{\phi}^* K_{\mathcal{U}_{0,n}^*/\overline{\mathcal{M}}_{0,n}} - \hat{\phi}^*\mathcal{D}^* \sim_{\mathbb{Q}} 0, \tag{4-4}$$

where the first equality follows from (4-3) and Lemma 4.1(3), and the second relation from (4-2) and [Liu 2002, Chapter 6, Theorem 4.9(b) and Example 3.18].

Since $\hat{f}$ has connected fibers, by (4-3) and (4-4) and projection formula we get

$$\tilde{D}_{\text{mod}} \sim_{\mathbb{Q}} \phi_0^*\mathcal{L}, \tag{4-5}$$

i.e., $\tilde{D}_{\text{mod}}$ is semiample.

Now since $\alpha_0 : Y'' \to Y$ is a composition of finite morphisms of degree strictly less than $\operatorname{char}(k)$ and β_0 is a birational morphism, by [Ambro 1999, Theorem 3.2 and Example 3.1],

$$K_{Y''} + D''_{\text{div}} \sim_{\mathbb{Q}} \alpha_0^*(K_Y + D_{\text{div}})$$

and

$$K_{\tilde{Y}} + \tilde{D}_{\text{div}} \sim_{\mathbb{Q}} \beta_0^*(K_{Y''} + D''_{\text{div}}).$$

So $\alpha_0^* D_{\text{mod}} \sim_{\mathbb{Q}} D''_{\text{mod}}$, and $\beta_0^* D''_{\text{mod}} \sim_{\mathbb{Q}} \tilde{D}_{\text{mod}}$. By the projection formula we have

$$D''_{\text{mod}} \sim_{\mathbb{Q}} \beta_{0,*}\tilde{D}_{\text{mod}}.$$

Then since α_0 is finite,

$$\psi_{0,*}\tilde{D}_{\text{mod}} \sim_{\mathbb{Q}} \alpha_{0,*}\beta_{0,*}\tilde{D}_{\text{mod}} \sim_{\mathbb{Q}} \alpha_{0,*}D''_{\text{mod}} \sim_{\mathbb{Q}} \alpha_{0,*}\alpha_0^* D_{\text{mod}} \sim_{\mathbb{Q}} D_{\text{mod}}.$$

Here we view the pushforward through α_0 as pushforward of cycles. Therefore D_{mod} is $\mathbb{Q}$-linearly equivalent to an effective divisor.

Next we prove the second statement. Since α is finite, by [Kollár 2013, Corollary 2.42] we know that (X'', D'') is klt, and as β, λ, and μ are birational we know that $(\hat{X}, \hat{D})$ is sub-klt, in particular $\hat{D}^v$ has coefficients < 1. Since $\hat{f}$ is a $\mathbb{P}^1$ fibration and $(\tilde{Y}, \tilde{D}_{\text{div}})$ is log smooth we have that $(\tilde{Y}, \tilde{D}_{\text{div}})$ is sub-klt. By construction $\tilde{D}_{\text{mod}}$ is semiample, so by [Tanaka 2015, Theorem 1] we know that $(\tilde{Y}, \tilde{D}_{\text{div}} + \tilde{D}_{\text{mod}})$ is sub-klt up to $\mathbb{Q}$-linear equivalence. Then $K_{Y''} + D''_{\text{mod}} + D''_{\text{div}} \sim_{\mathbb{Q}} \beta_{0,*}(K_{\tilde{Y}} + \tilde{D}_{\text{div}} + \tilde{D}_{\text{mod}})$ is also sub-klt. Finally using [Kollár 2013, Corollary 2.42] again and the fact that $D_{\text{mod}} + D_{\text{div}} \geq 0$ we get that $(Y, D_{\text{mod}} + D_{\text{div}})$ is klt. □

5. Global rational chain connectedness

In this section we prove the following theorem.

Theorem 5.1. *Let X be a projective threefold over an algebraically closed field k of characteristic $p > 0$, and $f : X \to Y$ a projective surjective morphism from X to a projective variety Y such that $f_*\mathcal{O}_X = \mathcal{O}_Y$. Let D be an effective $\mathbb{Q}$-divisor, and $X_{\bar{\eta}}$ the geometric generic fiber of f. Assume that the following conditions hold*:

(1) *(X, D) is klt, $-K_X$ is big and f-ample, $K_X + D \sim_{\mathbb{Q}} 0$ and the general fibers of f are smooth.*

(2) *$p > 2/\delta$, where δ is the minimum nonzero coefficient of D.*

(3) *$D = E + f^*L$ where E is an effective $\mathbb{Q}$-Cartier divisor such that $p \nmid \operatorname{ind}(E)$, $(X_{\bar{\eta}}, E|_{X_{\bar{\eta}}})$ is globally F-split, and L is a big $\mathbb{Q}$-divisor on Y.*

(4) *$\dim(Y)$ is 1 or 2.*

Then X is rationally chain connected.

Remark 5.2. Under the assumptions of Theorem 5.1, the smoothness of the general fibers of f holds in characteristic $p \geq 11$ when $\dim Y = 1$ by [Hirokado 2004, Theorem 5.1(2)], and in characteristic $p \geq 5$ when $\dim Y = 2$, as is explained in Remark 3.2.

Proposition 5.3. *Let $f : X \to Y$ be a projective surjective morphism between normal varieties with $f_*\mathcal{O}_X = \mathcal{O}_Y$. Assume that the following conditions hold*:

(1) *The general fibers of f are isomorphic to $\mathbb{P}^1$.*

(2) *Y is rationally chain connected.*

Then X is rationally chain connected.

Proof. The proof is essentially the same as [Gongyo et al. 2015a, Lemma 3.12 and Proposition 3.13]. We take two general points $x_1, x_2 \in X$ and let $y_1 = f(x_1)$, $y_2 = f(x_2)$, so by construction $f^{-1}(y_1) \cong f^{-1}(y_2) \cong \mathbb{P}^1$. By assumption y_1 and y_2 can be connected by a chain of rational curves, say $C_1, C_2, \ldots, C_n$. Let $\overline{C_i} \to C_i$ be the normalization for each C_i, $S_i := f^{-1}(C_i)$, $\overline{S_i} := S_i \times_{\overline{C_i}} C_i$, and $g_i : \overline{S_i} \to S_i$ the induced morphisms. Now the morphism $\overline{S_i} \to \overline{C_i}$ is a flat projective morphism whose general fibers are $\mathbb{P}^1$, by [de Jong and Starr 2003, Theorem] it has a section which we denote by $\tilde{C}_i$. Then x_1 and x_2 is connected by $f^{-1}(y_1)$, $f^{-1}(y_2)$, $g_i(\tilde{C}_i)$ and the fibers of f over the intersection points of $\{C_i\}$, which is a union of rational curves by [Debarre 2001, Lemma 3.7]. □

Proof of Theorem 5.1. We first prove the following lemma.

Lemma 5.4. *Under the condition of Theorem 5.1, $-K_Y$ is big.*

Proof. By assumption $m(K_{X_{\bar\eta}} + E|_{X_{\bar\eta}}) \sim_{\mathbb{Q}} 0$ for sufficiently large and divisible m; in particular, the $k(\bar\eta)$-algebra

$$\bigoplus_{m\geq 0} H^0\big(m(K_{X_{\bar\eta}} + E|_{X_{\bar\eta}})\big)$$

is finitely generated. On the other hand since $(X_{\bar\eta}, E|_{X_{\bar\eta}})$ is globally F-split we have that

$$S^0\big(X_{\bar\eta}, \sigma(X_{\bar\eta}, E|_{X_{\bar\eta}}) \otimes \mathcal{O}_{X_{\bar\eta}}(m(K_{X_{\bar\eta}} + E|_{X_{\bar\eta}}))\big) = H^0\big(X_{\bar\eta}, \mathcal{O}_{X_{\bar\eta}}(m(K_{X_{\bar\eta}} + E|_{X_{\bar\eta}}))\big).$$

Here we would like to mention that for a line bundle M and a $\mathbb{Q}$-Cartier divisor Δ, the notation $S^0(X, \Delta, M)$ is the same as the standard notation $S^0(X, \sigma(X, \Delta)\otimes M)$; see [Hacon and Xu 2015, between Lemma 2.2 and Proposition 2.3]. Therefore by [Ejiri 2015, Theorem 1.1],

$$f_*\mathcal{O}_X(m(K_{X/Y} + E)) \cong f_*\mathcal{O}_X(f^*(-m(K_Y + L))) = \mathcal{O}_Y(-m(K_Y - L))$$

is weakly positive for m sufficiently large and divisible. By Lemma 2.7, $-K_Y - L$ is nef, so $-K_Y$ is big. □

Next we consider the following two cases.

<u>Case 1: Y is 1-dimensional.</u> After possibly taking the normalization of Y we can assume that Y is smooth. Then Lemma 5.4 implies that $g(Y) = 0$, i.e., $Y \cong \mathbb{P}^1$. Let F be a general fiber of f. By assumption F is smooth and K_F is anti-ample, hence F is separably rationally connected. By [de Jong and Starr 2003, Theorem] we know that f has a section which we denote by s. Then $s(Y)$ is a rational curve in X which dominates Y. Therefore we get that X is rationally chain connected.

<u>Case 2: Y is 2-dimensional.</u> By assumption, a general fiber of f is isomorphic to $\mathbb{P}^1$. Now by Lemma 5.4 we know that $-K_Y$ is big. On the other hand since (X, D) is klt, by Theorem 4.3 there is a nonzero effective $\mathbb{Q}$-Cartier divisor M on Y such that $K_Y + M \sim_{\mathbb{Q}} 0$ and (Y, M) is klt. Then by the proof of Case 2 of Theorem 3.1 we know that Y is rational. Finally by Proposition 5.3 we get that X is rationally chain connected. □

Acknowledgements

The author would like to express his gratitude to Christopher Hacon for suggesting this direction of research and a lot of valuable suggestions, comments, support and encouragement. He would like to thank Karl Schwede for answering many questions about F-singularities. He also thanks Omprokash Das, Honglu Fan and Zsolt Patakfalvi for helpful discussions. Finally he would like to thank the referees for many useful suggestions.

References

[Ambro 1999] F. Ambro, *The adjunction conjecture and its applications*, Ph.D. thesis, Johns Hopkins University, 1999, available at http://search.proquest.com/docview/304508793.

[Birkar 2016] C. Birkar, "Existence of flips and minimal models for 3-folds in char p", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:1 (2016), 169–212. MR Zbl

[Bădescu 2001] L. Bădescu, *Algebraic surfaces*, Springer, New York, 2001. MR Zbl

[Campana 1992] F. Campana, "Connexité rationnelle des variétés de Fano", *Ann. Sci. École Norm. Sup.* (4) **25**:5 (1992), 539–545. MR Zbl

[Cascini et al. 2015] P. Cascini, H. Tanaka, and C. Xu, "On base point freeness in positive characteristic", *Ann. Sci. Éc. Norm. Supér.* (4) **48**:5 (2015), 1239–1272. MR Zbl

[Das and Hacon 2016] O. Das and C. D. Hacon, "On the adjunction formula for 3-folds in characteristic $p > 5$", *Math. Z.* **284**:1-2 (2016), 255–269. MR Zbl

[Debarre 2001] O. Debarre, *Higher-dimensional algebraic geometry*, Springer, New York, 2001. MR Zbl

[Ejiri 2015] S. Ejiri, "Weak positivity theorem and Frobenius stable canonical rings of geometric generic fibers", preprint, 2015. arXiv

[Gongyo et al. 2015a] Y. Gongyo, Z. Li, Z. Patakfalvi, K. Schwede, H. Tanaka, and R. Zong, "On rational connectedness of globally F-regular threefolds", *Adv. Math.* **280** (2015), 47–78. MR Zbl

[Gongyo et al. 2015b] Y. Gongyo, Y. Nakamura, and H. Tanaka, "Rational points on log Fano threefolds over a finite field", 2015. To appear in *J. Eur. Math. Soc.* arXiv

[Hacon and McKernan 2007] C. D. Hacon and J. McKernan, "On Shokurov's rational connectedness conjecture", *Duke Math. J.* **138**:1 (2007), 119–136. MR Zbl

[Hacon and Xu 2015] C. D. Hacon and C. Xu, "On the three dimensional minimal model program in positive characteristic", *J. Amer. Math. Soc.* **28**:3 (2015), 711–744. MR Zbl

[Hirokado 2004] M. Hirokado, "Deformations of rational double points and simple elliptic singularities in characteristic p", *Osaka J. Math.* **41**:3 (2004), 605–616. MR Zbl

[de Jong and Starr 2003] A. J. de Jong and J. Starr, "Every rationally connected variety over the function field of a curve has a rational point", *Amer. J. Math.* **125**:3 (2003), 567–580. MR Zbl

[Kawamata 1994] Y. Kawamata, "Semistable minimal models of threefolds in positive or mixed characteristic", *J. Algebraic Geom.* **3**:3 (1994), 463–491. MR Zbl

[Kawamata 1997] Y. Kawamata, "Subadjunction of log canonical divisors for a subvariety of codimension 2", pp. 79–88 in *Birational algebraic geometry* (Baltimore, 1996), edited by Y. Kawamata and V. V. Shokurov, Contemp. Math. **207**, Amer. Math. Soc., Providence, RI, 1997. MR Zbl

[Kollár 1992] J. Kollár (editor), *Flips and abundance for algebraic threefolds* (Salt Lake City, 1991), Astérisque **211**, Société Mathématique de France, Paris, 1992. MR Zbl

[Kollár 1996] J. Kollár, *Rational curves on algebraic varieties*, Ergebnisse der Mathematik (3) **32**, Springer, Berlin, 1996. MR Zbl

[Kollár 2013] J. Kollár, *Singularities of the minimal model program*, Cambridge Tracts in Mathematics **200**, Cambridge Univ. Press, 2013. MR Zbl

[Kollár and Mori 1998] J. Kollár and S. Mori, *Birational geometry of algebraic varieties*, Cambridge Tracts in Mathematics **134**, Cambridge Univ. Press, 1998. MR Zbl

[Kollár et al. 1992] J. Kollár, Y. Miyaoka, and S. Mori, "Rational connectedness and boundedness of Fano manifolds", *J. Differential Geom.* **36**:3 (1992), 765–779. MR Zbl

[Liedtke 2013] C. Liedtke, "Algebraic surfaces in positive characteristic", pp. 229–292 in *Birational geometry, rational curves, and arithmetic*, edited by F. Bogomolov et al., Springer, New York, 2013. MR Zbl

[Liu 2002] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Graduate Texts in Mathematics **6**, Oxford Univ. Press, 2002. MR Zbl

[Patakfalvi 2014] Z. Patakfalvi, "Semi-positivity in positive characteristics", *Ann. Sci. Éc. Norm. Supér.* (4) **47**:5 (2014), 991–1025. MR Zbl

[Prokhorov 2001] Y. G. Prokhorov, *Lectures on complements on log surfaces*, MSJ Memoirs **10**, Mathematical Society of Japan, Tokyo, 2001. MR Zbl

[Prokhorov and Shokurov 2009] Y. G. Prokhorov and V. V. Shokurov, "Towards the second main theorem on complements", *J. Algebraic Geom.* **18**:1 (2009), 151–199. MR Zbl

[Tanaka 2014] H. Tanaka, "Minimal models and abundance for positive characteristic log surfaces", *Nagoya Math. J.* **216** (2014), 1–70. MR Zbl

[Tanaka 2015] H. Tanaka, "Semiample perturbations for log canonical varieties over an F-finite field containing an infinite perfect field", preprint, 2015. arXiv

[Viehweg 1983] E. Viehweg, "Weak positivity and the additivity of the Kodaira dimension for certain fibre spaces", pp. 329–353 in *Algebraic varieties and analytic varieties* (Tokyo, 1981), edited by S. Iitaka, Adv. Stud. Pure Math. **1**, North-Holland, Amsterdam, 1983. MR Zbl

[Zhang 2006] Q. Zhang, "Rational connectedness of log **Q**-Fano varieties", *J. Reine Angew. Math.* **590** (2006), 131–142. MR Zbl

Received January 25, 2016. Revised November 23, 2016.

YUAN WANG
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
155 SOUTH 1400 EAST, ROOM 233
SALT LAKE CITY, UT 84112-0090
UNITED STATES
ywang@math.utah.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 290, No. 1, 2017
https://doi.org/10.2140/pjm.2017.290.247

AN ORTHOGONALITY RELATION FOR SPHERICAL CHARACTERS OF SUPERCUSPIDAL REPRESENTATIONS

CHONG ZHANG

We show that, in the setting of Galois pairs, the spherical characters of unitary supercuspidal representations satisfy an orthogonality relation.

1. Main result

Let F be a finite extension field of $\mathbb{Q}_p$ for an odd prime p, and E a quadratic field extension of F. Let G be a connected reductive group over F, and $\boldsymbol{G} = \mathrm{R}_{E/F} G$ the Weil restriction of G with respect to E/F. The nontrivial automorphism in $\mathrm{Gal}(E/F)$ induces an involution θ, defined over F, on $\boldsymbol{G}$. The pair $(\boldsymbol{G}, G)$ is called a Galois pair, which is a kind of symmetric pair.

Let π be an irreducible admissible unitary representation of $G(E) = \boldsymbol{G}(F)$. We say that π is G-distinguished if the space $\mathrm{Hom}_{G(F)}(\pi, \mathbb{C})$ is nonzero. We fix a Haar measure $\mathrm{d}g$ on $G(E)$. Given an element ℓ in $\mathrm{Hom}_{G(F)}(\pi, \mathbb{C})$, the *spherical character* $\Phi_{\pi,\ell}$ associated to ℓ is the distribution on $G(E)$ defined by

$$\Phi_{\pi,\ell}(f) := \sum_{v \in \mathrm{ob}(\pi)} \ell(\pi(f)v)\overline{\ell(v)}, \quad f \in C_c^\infty(G(E)),$$

where $\mathrm{ob}(\pi)$ is an orthonormal basis of the representation space V_π of π. In this note, our main goal is to show that spherical characters satisfy an orthogonality relation when π is unitary supercuspidal.

Before stating our result, we introduce some notation. Recall that an element $g \in G(E)$ is called θ-regular if $s(g) := g^{-1}\theta(g)$ is regular semisimple in $G(E)$ in the usual sense; a θ-regular element g is called θ-elliptic if the identity component of the centralizer of $s(g)$ in G is an elliptic F-torus. We denote by $G(E)_{\mathrm{reg}}$ (resp. $G(E)_{\mathrm{ell}}$) the subset of θ-regular (resp. θ-elliptic) elements of $G(E)$.

Theorem 1.1 [Hakim 1994, Theorem 1]. *The spherical character $\Phi_{\pi,\ell}$ is locally integrable on $G(E)$ and locally constant on the θ-regular locus $G(E)_{\mathrm{reg}}$.*

MSC2010: 11F70.
Keywords: spherical character, supercuspidal representation.

We denote by $\phi_{\pi,\ell}$ the locally integrable function on $G(E)$ representing the distribution $\Phi_{\pi,\ell}$, that is,

$$\Phi_{\pi,\ell}(f)=\int_{G(E)}\phi_{\pi,\ell}(g)f(g)\,\mathrm{d}g,\qquad f\in C_c^\infty(G(E)).$$

We will also call $\phi_{\pi,\ell}$ a *spherical character*. Note that $\phi_{\pi,\ell}$ is bi-$G(F)$-invariant and independent of the choice of Haar measures $\mathrm{d}g$. Theorem 1.1 is analogous to the classical result of Harish-Chandra [1999, Theorem 16.3] on admissible invariant distributions on connected reductive p-adic groups.

When π is unitary supercuspidal, we will show that the spherical characters $\phi_{\pi,\ell}$ satisfy an orthogonality relation (see Theorem 1.2). Before stating this relation, we need to review the Weyl integration formula in the setting of symmetric pairs. We refer the reader to [Rader and Rallis 1996, §3] or [Hakim 2003, §6] for the notation below and more details on this integration formula.

Let $\mathscr{T}$ be a set of representatives for the equivalence classes of Cartan subsets of $G(E)$ with respect to the involution θ. For $T\in\mathscr{T}$, denote $T_{\mathrm{reg}}=T\cap G(E)_{\mathrm{reg}}$. For $T\in\mathscr{T}$, the map

$$\mu: G(F)\times T_{\mathrm{reg}}\times G(F)\to G(E)_{\mathrm{reg}},\qquad (h_1,t,h_2)\mapsto h_1th_2,$$

is submersive and

$$G(E)_{\mathrm{reg}}=\coprod_{T\in\mathscr{T}}G(F)T_{\mathrm{reg}}G(F).$$

Let A be the split component of the center of G. For each θ-regular element g, we choose a Haar measure on $G_\gamma(F)$ where $\gamma=s(g)$ and G_γ is the split component of the centralizer of γ in G. Fix Haar measures on $A(F)$ and $G(F)$. For $\phi\in C_c^\infty(G(E)/A(F))$ and $g\in G(E)_{\mathrm{reg}}$, the orbital integral $O(g,\phi)$ of ϕ at g is defined to be

$$O(g,\phi)=\int_{A(F)\backslash G(F)}\int_{G_\gamma(F)\backslash G(F)}\phi(h_1gh_2)\,\mathrm{d}h_1\,\mathrm{d}h_2,$$

where $\gamma=s(g)$ and the measures inside the integral are quotient measures. From the definition we see that orbital integrals are bi-$G(F)$-invariant functions on $G(E)_{\mathrm{reg}}$. For $T\in\mathscr{T}$, the group $G_{s(t)}$ is the same for each $t\in T_{\mathrm{reg}}$. Let $D_{G(E)}$ be the usual Weyl discriminant function on $G(E)$. Then the Weyl integration formula reads as follows: with suitably normalized measures, for each $\phi\in C_c^\infty(G(E)/A(F))$, we have

$$\int_{G(E)/A(F)}\phi(g)\,\mathrm{d}g=\sum_{T\in\mathscr{T}}\frac{1}{w_T}\int_T|D_{G(E)}(s(t))|_E\cdot O(t,\phi)\,\mathrm{d}t,\tag{1}$$

where w_T are some positive constants only depending on T (see [Rader and Rallis 1996, Theorem 3.4] and [Hakim 2003, Lemma 5]). Let $\mathscr{T}_{\mathrm{ell}}$ be the subset of $\mathscr{T}$ consisting of elliptic Cartan subsets, that is, for $T \in \mathscr{T}$, T belongs to $\mathscr{T}_{\mathrm{ell}}$ if and only if $T_{\mathrm{reg}} \subset G(E)_{\mathrm{ell}}$.

Theorem 1.2. (1) *Suppose that π is unitary supercuspidal and G-distinguished. Let ℓ be a nonzero element of* $\mathrm{Hom}_{G(F)}(\pi, \mathbb{C})$. *Then*

$$\sum_{T \in \mathscr{T}_{\mathrm{ell}}} \frac{1}{w_T} \int_T |D_{G(E)}(s(t))|_E \cdot |\phi_{\pi,\ell}(t)|^2 \,\mathrm{d}t$$

is nonzero.

(2) *Suppose that π and π' are two unitary supercuspidal representations of $G(E)$ and $\pi \ncong \pi'$. Then for any $\ell \in$* $\mathrm{Hom}_{G(F)}(\pi, \mathbb{C})$ *and $\ell' \in$* $\mathrm{Hom}_{G(F)}(\pi', \mathbb{C})$, *we have*

$$\sum_{T \in \mathscr{T}_{\mathrm{ell}}} \frac{1}{w_T} \int_T |D_{G(E)}(s(t))|_E \cdot \phi_{\pi,\ell}(t) \cdot \overline{\phi_{\pi',\ell'}(t)} \,\mathrm{d}t = 0.$$

Theorem 1.2 is an analog of the classical orthogonality relation for characters of discrete series (see [Clozel 1991] or [Kazhdan 1986] for this classical result). The following corollary, which has potential application in simple relative trace formula, is a direct consequence of Theorem 1.2.

Corollary 1.3. *Suppose that π is unitary supercuspidal and G-distinguished. Let ℓ be a nonzero element of* $\mathrm{Hom}_{G(F)}(\pi, \mathbb{C})$. *Then the spherical character $\Phi_{\pi,\ell}$ does not vanish identically on $G(E)_{\mathrm{ell}}$.*

2. Proof of Theorem 1.2

Lemma 2.1. *Suppose that $\gamma = s(g)$ with $g \in G(E)$ lies in an F-Levi subgroup M of G. Then there exists $m \in M(E)$ such that $\gamma = s(m)$.*

Proof. First we recall some basic facts about symmetric spaces. Denote $\boldsymbol{G} = \mathrm{R}_{E/F} G$ and $\boldsymbol{M} = \mathrm{R}_{E/F} M$. Let $\boldsymbol{X} = \boldsymbol{G}/G$ and $\boldsymbol{X}_M = \boldsymbol{M}/M$ be the quotient varieties. As F-varieties, $\boldsymbol{X}$ and $\boldsymbol{X}_M$ are isomorphic to the identity components of the varieties defined by the equations

$$\tilde{\boldsymbol{X}} = \{x \in \boldsymbol{G} : x\theta(x) = 1\} \quad \text{and} \quad \tilde{\boldsymbol{X}}_M = \{x \in \boldsymbol{M} : x\theta(x) = 1\}$$

respectively [Richardson 1982, 2.1–2.4]. The exact sequences

$$1 \to G \to \boldsymbol{G} \to \boldsymbol{X} \to 1 \quad \text{and} \quad 1 \to M \to \boldsymbol{M} \to \boldsymbol{X}_M \to 1$$

induce the following exact cohomology sequences:

$$1 \to s(\boldsymbol{G}(F)) \to \boldsymbol{X}(F) \to H^1(F, G)$$

and

$$1 \to s(\boldsymbol{M}(F)) \to \boldsymbol{X}_M(F) \to H^1(F, M),$$

where we use the standard notation $H^1(F, \bullet)$ to denote the Galois cohomology of algebraic groups [Serre 1997, Chapter III. §2]. However, the above exact sequences have little to do with our assertion. What we need are the following exact sequences [Carmeli 2015, Lemma 4.1.1]:

$$\begin{array}{ccccccccc} 1 & \to & s(\boldsymbol{M}(F)) & \to & \tilde{\boldsymbol{X}}_M(F) & \to & H^1(\theta, \boldsymbol{M}(F)) & \to & 1 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 1 & \to & s(\boldsymbol{G}(F)) & \to & \tilde{\boldsymbol{X}}(F) & \to & H^1(\theta, \boldsymbol{G}(F)) & \to & 1, \end{array}$$

where

$$H^1(\theta, \boldsymbol{M}(F)) := H^1(\mathrm{Gal}(E/F), M(E))$$

and

$$H^1(\theta, \boldsymbol{G}(F)) := H^1(\mathrm{Gal}(E/F), G(E)).$$

Note that $\gamma \in \tilde{\boldsymbol{X}}_M(F)$, and Lemma 2.1 asserts that $\gamma \in s(\boldsymbol{M}(F))$. Thus it suffices to show that the image $[\gamma]_M$ of γ in $H^1(\theta, \boldsymbol{M}(F))$ is trivial. On the other hand, we know that the image $[\gamma]_G$ of γ in $H^1(\theta, \boldsymbol{G}(F))$ is trivial, and $[\gamma]_G$ is also the image of $[\gamma]_M$ under the natural map

$$\iota : H^1(\theta, \boldsymbol{M}(F)) \to H^1(\theta, \boldsymbol{G}(F)).$$

We claim that ι is injective, which implies that $[\gamma]_M$ is trivial. Consider the exact sequences [Serre 1997, Chapter I. §5.8(a)]:

$$\begin{array}{ccccccc} 1 & \to & H^1(\theta, \boldsymbol{M}(F)) & \to & H^1(F, M) & \to & H^1(E, M)^{\mathrm{Gal}(E/F)} \\ & & \downarrow & & \downarrow & & \downarrow \\ 1 & \to & H^1(\theta, \boldsymbol{G}(F)) & \to & H^1(F, G) & \to & H^1(E, G)^{\mathrm{Gal}(E/F)}. \end{array}$$

Let P an F-parabolic subgroup of G such that $P = M \ltimes U$ where U is the unipotent radical of P. We have natural isomorphisms (see [Gille 2007, Lemma 16.2])

$$H^1(F, P) \xrightarrow{\simeq} H^1(F, M), \quad \text{and} \quad H^1(E, P) \xrightarrow{\simeq} H^1(E, M),$$

and natural injections [Serre 1997, Chapter III. §2.1]

$$H^1(F, P) \hookrightarrow H^1(F, G) \quad \text{and} \quad H^1(E, P) \hookrightarrow H^1(E, G).$$

In summary we have the following commutative diagram of exact sequences:

$$\begin{array}{ccccccc} 1 & \to & H^1(\theta, \boldsymbol{G}(F)) & \to & H^1(F, G) & \to & H^1(E, G) \\ & & \uparrow & & \uparrow & & \uparrow \\ 1 & \to & H^1(\theta, \boldsymbol{P}(F)) & \to & H^1(F, P) & \to & H^1(E, P) \\ & & \downarrow & & \downarrow & & \downarrow \\ 1 & \to & H^1(\theta, \boldsymbol{M}(F)) & \to & H^1(F, M) & \to & H^1(E, M), \end{array}$$

which implies that ι is injective. □

Lemma 2.2. *Suppose that ϕ is a matrix coefficient of a unitary supercuspidal G-distinguished representation. Then, for any $g \in G(E)_{\mathrm{reg}}$, the orbital integral $O(g, \phi)$ vanishes unless g is θ-elliptic.*

Proof. Since ϕ is a matrix coefficient of a unitary supercuspidal G-distinguished representation, it belongs to $C_c^\infty(G(E)/A(F))$ and is a supercusp form [Harish-Chandra 1970, Part I. §3]. In particular, for any unipotent radical N of a proper parabolic subgroup P of G, we have

$$\int_{N(E)} \phi(gn)\, \mathrm{d}n = 0$$

for any $g \in G(E)$. Write $\gamma = s(g)$. Suppose that g is not θ-elliptic, which means that γ is not elliptic by definition. Therefore there exists a Levi subgroup M of a proper parabolic subgroup P of G such that $G_\gamma \subset M$. According to Lemma 2.1 there exists $m \in M(E)$ such that $\gamma = s(m)$. Since

$$O(g, \phi) = O(m, \phi),$$

we assume that g is in $M(E)$ from now on. Let N be the unipotent radical of P, and K a maximal open compact subgroup of $G(F)$ such that $G(F) = M(F)N(F)K$. Fix Haar measures $\mathrm{d}m$, $\mathrm{d}n$ and $\mathrm{d}k$ on $M(F)/A(F)$, $N(F)$ and $K/K \cap A(F)$ respectively so that $\mathrm{d}h = \mathrm{d}k\, \mathrm{d}n\, \mathrm{d}m$ on $G(F)/A(F)$. Denote $\bar{K} = K/K \cap A(F)$. Then the orbital integral $O(g, \phi)$ can be written as follows:

$$\begin{aligned} O(g, \phi) &= \int_{A(F)\backslash G(F)} \int_{G_\gamma(F)\backslash G(F)} \phi(h_1^{-1} g h_2)\, \mathrm{d}h_2\, \mathrm{d}h_1 \\ &= \int_{(A(F)\backslash M(F)) \times N(F) \times \bar{K}} \int_{(G_\gamma(F)\backslash M(F)) \times N(F) \times \bar{K}} \phi(k_1^{-1} n_1^{-1} m_1^{-1} g m_2 n_2 k_2) \\ &\qquad\qquad \cdot \mathrm{d}k_1\, \mathrm{d}k_2\, \mathrm{d}n_1\, \mathrm{d}n_2\, \mathrm{d}m_1\, \mathrm{d}m_2 \\ &= \int_{(A(F)\backslash M(F)) \times N(F)} \int_{(G_\gamma(F)\backslash M(F)) \times N(F)} \phi'(n_1^{-1} m_1^{-1} g m_2 n_2) \\ &\qquad\qquad \cdot \mathrm{d}n_1\, \mathrm{d}n_2\, \mathrm{d}m_1\, \mathrm{d}m_2, \end{aligned}$$

where

$$\phi'(x) := \int_{\bar{K}\times\bar{K}} \phi(k_1 x k_2)\, \mathrm{d}k_1\, \mathrm{d}k_2, \quad x \in G(E)/A(F).$$

Note that ϕ' is still a supercusp form on $G(E)$. From now on, for convenience, we write ϕ instead of ϕ' and g instead of $m_1^{-1} g m_2$. Let $\gamma = s(g)$ for this "new" g. We claim that:

$$\int_{N(F)\times N(F)} \phi(n_1^{-1} g n_2)\, \mathrm{d}n_1\, \mathrm{d}n_2 = 0. \tag{2}$$

It is clear that this claim implies the lemma directly.

Now we begin to prove claim (2). Note that

$$\int_{N(F)\times N(F)} \phi(n_1^{-1} g n_2)\, \mathrm{d}n_1\, \mathrm{d}n_2 = \int_{N(F)\times N(F)} \phi(g \cdot g^{-1} n_1 g n_2)\, \mathrm{d}n_1\, \mathrm{d}n_2.$$

Denote $\boldsymbol{N} = \mathrm{R}_{E/F} N$. Consider the morphism of the algebraic varieties:

$$\eta_g : N \times N \to \boldsymbol{N}, \quad (n_1, n_2) \mapsto g^{-1} n_1 g n_2.$$

We will show that η_g is an isomorphism. If $g^{-1} n_1 g n_2 = g^{-1} n_1' g n_2'$, we have the relation

$$n_2^{-1} \gamma n_2 = s(n_1 g n_2) = s(n_1' g n_2') = n_2'^{-1} \gamma n_2'. \tag{3}$$

Since γ is regular, according to [Harish-Chandra 1970, Lemma 22], the equality (3) implies $n_2 = n_2'$, and thus $n_1 = n_1'$. Hence η_g is injective. To show η_g is surjective, consider the Lie algebras $\mathfrak{n}' = \mathrm{Lie}(N')$, $\mathfrak{n}'' = \mathrm{Lie}(N)$ and $\mathfrak{n} = \mathrm{Lie}(\boldsymbol{N})$, where N' is the unipotent subgroup $g^{-1} N g$. Since

$$2 \dim_F \mathfrak{n}' = 2 \dim_F \mathfrak{n}'' = \dim_F \mathfrak{n}$$

and $\mathfrak{n}' \cap \mathfrak{n}'' = \{0\}$ by the injectivity of η_g, we have $\mathfrak{n} = \mathfrak{n}' \oplus \mathfrak{n}''$. Therefore η_g is submersive and thus $N' \cdot N$ is open in $\boldsymbol{N}$. On the other hand, since N' and N are unipotent groups, the orbit $N' \cdot N$ of 1 under the left and right translations of N' and N is closed in $\boldsymbol{N}$. Hence $\boldsymbol{N} = N' \cdot N$, that is, η_g is surjective. It turns out that

$$\int_{N(E)} \phi(gn)\, \mathrm{d}n = \int_{N(F)\times N(F)} j_g(n_1, n_2) \cdot \phi(g \cdot g^{-1} n_1 g n_2)\, \mathrm{d}n_1\, \mathrm{d}n_2,$$

where $j_g(n_1, n_2)$ is the Jacobian of η_g at (n_1, n_2). Note that

$$j_g(n_1, n_2) = |\mathrm{ad}(g)|_{\mathfrak{n}(F)}|_E,$$

which is independent of (n_1, n_2). At last, the claim (2) follows from the condition that ϕ is a supercusp form. $\square$

Proof of Theorem 1.2. Let π be a unitary supercuspidal representation of $G(E)$ and $\ell \in \mathrm{Hom}_{G(F)}(\pi, \mathbb{C})$. By [Zhang 2016, Theorem 1.5], there exists a vector u_0 in the space V_π such that $\ell = \mathscr{L}_{u_0}$, where the $G(F)$-invariant linear form $\mathscr{L}_{u_0}$ is defined by

$$\mathscr{L}_{u_0}(v) := \int_{G(F)/A(F)} \langle \pi(h)v, u_0 \rangle \,\mathrm{d}h, \quad v \in V_\pi.$$

Set

$$\phi(g) = \langle \pi(g)u_0, u_0 \rangle,$$

which is a matrix coefficient of π. Then, according to [Zhang 2016, Corollary 1.11], the spherical character $\Phi_{\pi,\ell}$ has the following expression:

$$\Phi_{\pi,\ell}(f) = \int_{G(F)/A(F)} \int_{G(F)/A(F)} \int_{G(E)} \phi(h_1 g h_2) f(g) \,\mathrm{d}g \,\mathrm{d}h_1 \,\mathrm{d}h_2. \tag{4}$$

Note that $G_{s(g)} = A$ for $g \in G(E)_{\mathrm{ell}}$. Therefore, when $f \in C_c^\infty(G(E)_{\mathrm{ell}})$, we get

$$\Phi_{\pi,\ell}(f) = \int_{G(E)} O(g, \phi) f(g) \,\mathrm{d}g.$$

On the other hand, by Theorem 1.1, we have

$$\Phi_{\pi,\ell}(f) = \int_{G(E)} \phi_{\pi,\ell}(g) f(g) \,\mathrm{d}g.$$

Therefore, for $g \in G(E)_{\mathrm{ell}}$, we obtain

$$\phi_{\pi,\ell}(g) = O(g, \phi). \tag{5}$$

Now let π' be another unitary supercuspidal representation of $G(E)$ and $\ell' \in \mathrm{Hom}_{G(F)}(\pi', \mathbb{C})$. Let ϕ' be a matrix coefficient of π' such that the distribution $\Phi_{\pi',\ell'}$ can be expressed as

$$\Phi_{\pi',\ell'}(f) = \int_{G(F)/A(F)} \int_{G(F)/A(F)} \int_{G(E)} \phi'(h_1 g h_2) f(g) \,\mathrm{d}g \,\mathrm{d}h_1 \,\mathrm{d}h_2$$

for any $f \in C_c^\infty(G(E))$. Thus

$$\phi_{\pi',\ell'}(g) = O(g, \phi') \tag{6}$$

for any $g \in G(E)_{\mathrm{ell}}$. We choose $f_1 \in C_c^\infty(G(E))$ so that $\phi_{f_1} = \bar{\phi}'$, where

$$\phi_{f_1}(g) := \int_{A(F)} f_1(ag) \,\mathrm{d}a.$$

Then, by the Weyl integration formula (1), we have

$$\Phi_{\pi,\ell}(f_1) = \int_{G(E)/A(F)} \phi_{\pi,\ell}(g) \cdot \phi_{f_1}(g)\,\mathrm{d}g$$
$$= \sum_{T \in \mathscr{T}} \frac{1}{w_T} \int_T |D_{G(E)}(s(t))|_E \cdot \phi_{\pi,\ell}(t) \cdot O(t, \bar{\phi}')\,\mathrm{d}t.$$

Combining Lemma 2.2 and (6), we get

$$\Phi_{\pi,\ell}(f_1) = \sum_{T \in \mathscr{T}_{\mathrm{ell}}} \frac{1}{w_T} \int_T |D_{G(E)}(s(t))|_E \cdot \phi_{\pi,\ell}(t) \cdot \overline{\phi_{\pi',\ell'}(t)}\,\mathrm{d}t. \tag{7}$$

For the first assertion, we take $\pi' = \pi$, $\ell' = \ell$ and $\phi' = \phi$. Then (7) implies

$$\Phi_{\pi,\ell}(f_1) = \sum_{T \in \mathscr{T}_{\mathrm{ell}}} \frac{1}{w_T} \int_T |D_{G(E)}(s(t))|_E \cdot |\phi_{\pi,\ell}(t)|^2\,\mathrm{d}t.$$

On the other hand, we set

$$v_0 = \frac{1}{\sqrt{\langle u_0, u_0 \rangle}} u_0$$

and choose $\{v_i\}_{i \in \mathbb{N}}$ such that $\{v_i\}_{i \geq 0}$ is an orthonormal basis of V_π. Then

$$\pi(\bar{\phi}) v_0 = \lambda v_0 \quad \text{for some nonzero } \lambda, \quad \text{and} \quad \pi(\bar{\phi}) v_i = 0 \quad \text{for } i \geq 1.$$

Therefore

$$\Phi_{\pi,\ell}(f_1) = \lambda |\ell(v_0)|^2.$$

From the proof of [Zhang 2016, Theorem 1.4] (page 1542), we see that

$$\overline{\ell(u_0)} = c \int_{A(E)G(F)\backslash G(E)} |\ell(\pi(g)u_0)|^2\,\mathrm{d}g = c' \langle u_0, u_0 \rangle,$$

where c and c' are some nonzero numbers. Hence $\Phi_{\pi,\ell}(f_1)$ is nonzero. This completes the proof of the first assertion.

As for the second assertion, note that

$$\Phi_{\pi,\ell}(f_1) = \int_{G(F)/A(F)} \int_{G(F)/A(F)} \int_{G(E)/A(F)} \phi(h_1 g h_2) \bar{\phi}'(g)\,\mathrm{d}g\,\mathrm{d}h_1\,\mathrm{d}h_2 = 0,$$

since the inner integral over $G(E)/A(F)$ vanishes by the Schur orthogonality relation. Hence the assertion is deduced from (7) directly. □

Acknowledgements

This work was partially supported by NSFC 11501033 and the Fundamental Research Funds for the Central Universities. The author thanks Wen-Wei Li for helpful discussions, especially for his help on some arguments in the proof of Lemma 2.2. He also thanks the anonymous referees for their helpful suggestions to improve this article.

References

[Carmeli 2015] S. Carmeli, "On the stability and Gelfand property of symmetric pairs", preprint, 2015. arXiv

[Clozel 1991] L. Clozel, "Invariant harmonic analysis on the Schwartz space of a reductive p-adic group", pp. 101–121 in *Harmonic analysis on reductive groups* (Brunswick, ME, 1989), edited by W. Barker and P. Sally, Progr. Math. **101**, Birkhäuser, Boston, 1991. MR Zbl

[Gille 2007] P. Gille, "Rationality properties of linear algebraic groups and Galois cohomology", course notes, 2007, available at http://math.univ-lyon1.fr/homes-www/gille/prenotes/mcm.pdf.

[Hakim 1994] J. Hakim, "Admissible distributions on p-adic symmetric spaces", *J. Reine Angew. Math.* **455** (1994), 1–19. MR Zbl

[Hakim 2003] J. Hakim, "Supercuspidal Gelfand pairs", *J. Number Theory* **100**:2 (2003), 251–269. MR Zbl

[Harish-Chandra 1970] Harish-Chandra, *Harmonic analysis on reductive p-adic groups*, Lecture Notes in Mathematics **162**, Springer, Berlin, 1970. MR Zbl

[Harish-Chandra 1999] Harish-Chandra, *Admissible invariant distributions on reductive p-adic groups*, University Lecture Series **16**, American Mathematical Society, Providence, RI, 1999. MR Zbl

[Kazhdan 1986] D. Kazhdan, "Cuspidal geometry of p-adic groups", *J. Analyse Math.* **47** (1986), 1–36. MR Zbl

[Rader and Rallis 1996] C. Rader and S. Rallis, "Spherical characters on p-adic symmetric spaces", *Amer. J. Math.* **118**:1 (1996), 91–178. MR Zbl

[Richardson 1982] R. W. Richardson, "Orbits, invariants, and representations associated to involutions of reductive groups", *Invent. Math.* **66**:2 (1982), 287–312. MR Zbl

[Serre 1997] J.-P. Serre, *Galois cohomology*, Springer, Berlin, 1997. MR Zbl

[Zhang 2016] C. Zhang, "Local periods for discrete series representations", *J. Funct. Anal.* **271**:6 (2016), 1525–1543. MR Zbl

Received January 11, 2017. Revised February 9, 2017.

CHONG ZHANG
SCHOOL OF MATHEMATICAL SCIENCES
BEIJING NORMAL UNIVERSITY
BEIJING 100875
CHINA

zhangchong@bnu.edu.cn

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 290 No. 1 September 2017
