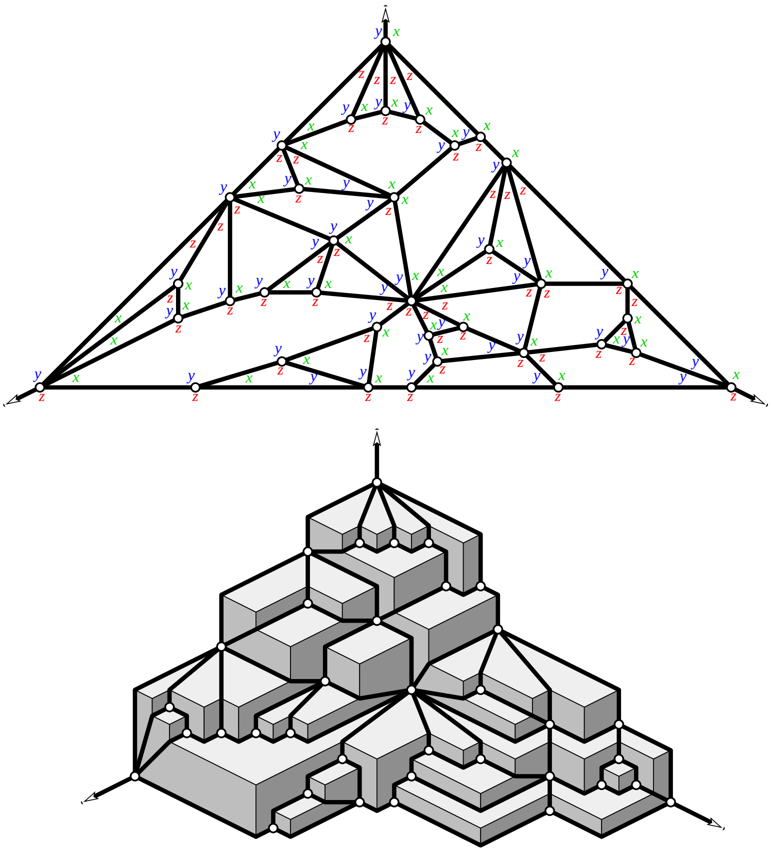
Documenta Mathematica

Gegründet 1996 durch die

Deutsche Mathematiker-Vereinigung



Orthogonal coloring and rigid embedding
of an extended map, see page 48

Band 7 · 2002

Documenta Mathematica veröffentlicht Forschungsarbeiten aus allen mathematischen Gebieten und wird in traditioneller Weise referiert.

Documenta Mathematica erscheint am World Wide Web unter:
`http://www.mathematik.uni-bielefeld.de/documenta`

Artikel können als TEX-Dateien per E-Mail bei einem der Herausgeber eingereicht werden. Hinweise für die Vorbereitung der Artikel können unter der obigen WWW-Adresse gefunden werden.

Documenta Mathematica publishes research manuscripts out of all mathematical fields and is refereed in the traditional manner.

Documenta Mathematica is published on the World Wide Web under:
`http://www.mathematik.uni-bielefeld.de/documenta`

Manuscripts should be submitted as TEX -files by e-mail to one of the editors. Hints for manuscript preparation can be found under the above WWW-address.

ISSN 1431-0635 (Print), ISSN 1431-0643 (Internet)

SPARC
Leading Edge

Documenta Mathematica is a Leading Edge Partner of SPARC, the Scholarly Publishing and Academic Resource Coalition of the Association of Research Libraries (ARL), Washington DC, USA.

Address of Technical Managing Editor: Ulf Rehmann, Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Typesetting in TEX, Printing: Schury Druck & Verlag, 83064 Raubling, Germany.

Documenta Mathematica

Band 7, 2002

C^*-Algebras Associated with Presentations of Subshifts

Kengo Matsumoto

Received: May 28, 2001
Revised: March 4, 2002

Communicated by Joachim Cuntz

Abstract. A λ-graph system is a labeled Bratteli diagram with an upward shift except the top vertices. We construct a continuous graph in the sense of V. Deaconu from a λ-graph system. It yields a Renault's groupoid C^*-algebra by following Deaconu's construction. The class of these C^*-algebras generalize the class of C^*-algebras associated with subshifts and hence the class of Cuntz-Krieger algebras. They are unital, nuclear, unique C^*-algebras subject to operator relations encoded in the structure of the λ-graph systems among generating partial isometries and projections. If the λ-graph systems are irreducible (resp. aperiodic), they are simple (resp. simple and purely infinite). K-theory formulae of these C^*-algebras are presented so that we know an example of a simple and purely infinite C^*-algebra in the class of these C^*-algebras that is not stably isomorphic to any Cuntz-Krieger algebra.

2000 Mathematics Subject Classification: Primary 46L35, Secondary 37B10.
Keywords and Phrases: C^*-algebras, subshifts, groupoids, Cuntz-Krieger algebras

1. Introduction

In [CK], J. Cuntz-W. Krieger have presented a class of C^*-algebras associated to finite square matrices with entries in $\{0, 1\}$. The C^*-algebras are simple if the matrices satisfy condition (I) and irreducible. They are also purely infinite if the matrices are aperiodic. There are many directions to generalize the Cuntz-Krieger algebras (cf. [An],[De],[De2],[EL],[KPRR],[KPW],[Pi], [Pu],[T], etc.). The Cuntz-Krieger algebras have close relationships to topological Markov shifts by Cuntz-Krieger's observation in [CK]. Let Σ be a finite set, and let σ be the shift on the infinite product space $\Sigma^{\mathbb{Z}}$ defined by

$\sigma((x_n)_{n\in\mathbb{Z}}) = (x_{n+1})_{n\in\mathbb{Z}}, (x_n)_{n\in\mathbb{Z}} \in \Sigma^{\mathbb{Z}}$. For a closed σ-invariant subset Λ of $\Sigma^{\mathbb{Z}}$, the topological dynamical system Λ with σ is called a subshift. The topological Markov shifts form a class of subshifts. In [Ma], the author has generalized the class of Cuntz-Krieger algebras to a class of C^*-algebras associated with subshifts. He has formulated several topological conjugacy invariants for subshifts by using the K-theory for these C^*-algebras ([Ma5]). He has also introduced presentations of subshifts, that are named symbolic matrix system and λ-graph system ([Ma5]). They are generalized notions of symbolic matrix and λ-graph (= labeled graph) for sofic subshifts respectively.
We henceforth denote by $\mathbb{Z}_+$ and $\mathbb{N}$ the set of all nonnegative integers and the set of all positive integers respectively. A symbolic matrix system $(\mathcal{M}, I)$ over a finite set Σ consists of two sequences of rectangular matrices $(\mathcal{M}_{l,l+1}, I_{l,l+1}), l \in \mathbb{Z}_+$. The matrices $\mathcal{M}_{l,l+1}$ have their entries in the formal sums of Σ and the matrices $I_{l,l+1}$ have their entries in $\{0,1\}$. They satisfy the following relations

$$I_{l,l+1}\mathcal{M}_{l+1,l+2} = \mathcal{M}_{l,l+1}I_{l+1,l+2}, \qquad l \in \mathbb{Z}_+. \tag{1.1}$$

It is assumed for $I_{l,l+1}$ that for i there exists j such that the (i,j)-component $I_{l,l+1}(i,j) = 1$ and that for j there uniquely exists i such that $I_{l,l+1}(i,j) = 1$. A λ-graph system $\mathfrak{L} = (V, E, \lambda, \iota)$ consists of a vertex set $V = V_0 \cup V_1 \cup V_2 \cup \cdots$, an edge set $E = E_{0,1} \cup E_{1,2} \cup E_{2,3} \cup \cdots$, a labeling map $\lambda : E \to \Sigma$ and a surjective map $\iota_{l,l+1} : V_{l+1} \to V_l$ for each $l \in \mathbb{Z}_+$. It naturally arises from a symbolic matrix system. For a symbolic matrix system $(\mathcal{M}, I)$, a labeled edge from a vertex $\mathfrak{v}_i^l \in V_l$ to a vertex $\mathfrak{v}_j^{l+1} \in V_{l+1}$ is given by a symbol appearing in the (i,j)-component $\mathcal{M}_{l,l+1}(i,j)$ of the matrix $\mathcal{M}_{l,l+1}$. The matrix $I_{l,l+1}$ defines a surjection $\iota_{l,l+1}$ from V_{l+1} to V_l for each $l \in \mathbb{Z}_+$. The symbolic matrix systems and the λ-graph systems are the same objects. They give rise to subshifts by looking the set of all label sequences appearing in the labeled Bratteli diagram (V, E, λ). A canonical method to construct a symbolic matrix system and a λ-graph system from an arbitrary subshift has been introduced in [Ma5]. The obtained symbolic matrix system and the λ-graph system are said to be canonical for the subshift. For a symbolic matrix system $(\mathcal{M}, I)$, let $A_{l,l+1}$ be the nonnegative rectangular matrix obtained from $\mathcal{M}_{l,l+1}$ by setting all the symbols equal to 1 for each $l \in \mathbb{Z}_+$. The resulting pair (A, I) satisfies the following relations from (1.1)

$$I_{l,l+1}A_{l+1,l+2} = A_{l,l+1}I_{l+1,l+2}, \qquad l \in \mathbb{Z}_+. \tag{1.2}$$

We call (A, I) the nonnegative matrix system for $(\mathcal{M}, I)$.
In the present paper, we introduce C^*-algebras from λ-graph systems. If a λ-graph system is the canonical λ-graph system for a subshift Λ, the C^*-algebra coincides with the C^*-algebra $\mathcal{O}_\Lambda$ associated with the subshift. Hence the class of the C^*-algebras in this paper generalize the class of Cuntz-Krieger algebras. Let $\mathfrak{L} = (V, E, \lambda, \iota)$ be a λ-graph system over alphabet Σ. We first construct a continuous graph from $\mathfrak{L}$ in the sense of V. Deaconu ([D2],[De3],[De4]). We

then define the C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ associated with $\mathfrak{L}$ as the Renault's C^*-algebra of a groupoid constructed from the continuous graph. For an edge $e \in E_{l,l+1}$, we denote by $s(e) \in V_l$ and $t(e) \in V_{l+1}$ its source vertex and its terminal vertex respectively. Let Λ^l be the set of all words of length l of symbols appearing in the labeled Bratteli diagram of $\mathfrak{L}$. We put $\Lambda^* = \cup_{l=0}^{\infty} \Lambda^l$ where Λ^0 denotes the empty word. Let $\{\mathfrak{v}_1^l, \dots, \mathfrak{v}_{m(l)}^l\}$ be the vertex set V_l. We denote by $\Gamma_l^-(\mathfrak{v}_i^l)$ the set of all words in Λ^l presented by paths starting at a vertex of V_0 and terminating at the vertex $\mathfrak{v}_i^l$. $\mathfrak{L}$ is said to be left-resolving if there are no distinct edges with the same label and the same terminal vertex. $\mathfrak{L}$ is said to be predecessor-separated if $\Gamma_l^-(\mathfrak{v}_i^l) \neq \Gamma_l^-(\mathfrak{v}_j^l)$ for distinct i, j and for all $l \in \mathbb{N}$. Assume that $\mathfrak{L}$ is left-resolving and satisfies condition (I), a mild condition generalizing Cuntz-Krieger's condition (I). We then prove:

Theorem A (Theorem 3.6 and Theorem 4.3). *Suppose that a λ-graph system $\mathfrak{L}$ satisfies condition (I). Then the C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ is the universal concrete unique C^*-algebra generated by partial isometries $S_\alpha, \alpha \in \Sigma$ and projections $E_i^l, i = 1, 2, \dots, m(l),\ l \in \mathbb{Z}_+$ satisfying the following operator relations:*

$$\sum_{\alpha \in \Sigma} S_\alpha S_\alpha^* = 1, \tag{1.3}$$

$$\sum_{i=1}^{m(l)} E_i^l = 1, \qquad E_i^l = \sum_{j=1}^{m(l+1)} I_{l,l+1}(i,j) E_j^{l+1}, \tag{1.4}$$

$$S_\alpha S_\alpha^* E_i^l = E_i^l S_\alpha S_\alpha^*, \tag{1.5}$$

$$S_\alpha^* E_i^l S_\alpha = \sum_{j=1}^{m(l+1)} A_{l,l+1}(i,\alpha,j) E_j^{l+1}, \tag{1.6}$$

for $i = 1, 2, \dots, m(l), l \in \mathbb{Z}_+, \alpha \in \Sigma$, where

$$A_{l,l+1}(i,\alpha,j) = \begin{cases} 1 & \text{if } s(e) = \mathfrak{v}_i^l, \lambda(e) = \alpha, t(e) = \mathfrak{v}_j^{l+1} \text{ for some } e \in E_{l,l+1}, \\ 0 & \text{otherwise,} \end{cases}$$
$$I_{l,l+1}(i,j) = \begin{cases} 1 & \text{if } \iota_{l,l+1}(\mathfrak{v}_j^{l+1}) = \mathfrak{v}_i^l, \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1, 2, \dots, m(l),\ j = 1, 2, \dots, m(l+1),\ \alpha \in \Sigma$.

If $\mathfrak{L}$ is predecessor-separated, the following relations:

$$E_i^l = \prod_{\substack{\mu,\nu \in \Lambda^l \\ \mu \in \Gamma_l^-(\mathfrak{v}_i^l), \nu \notin \Gamma_l^-(\mathfrak{v}_i^l)}} S_\mu^* S_\mu (1 - S_\nu^* S_\nu), \qquad l \in \mathbb{N}, \tag{1.7}$$
$$E_i^0 = \sum_{\alpha \in \Sigma} \sum_{j=1}^{m(1)} A_{0,1}(i,\alpha,j) S_\alpha E_j^1 S_\alpha^*$$

hold for $i = 1, 2, \dots, m(l)$, where $S_\mu = S_{\mu_1} \cdots S_{\mu_k}$ for $\mu = (\mu_1, \dots, \mu_k), \mu_1, \dots, \mu_k \in \Sigma$. In this case, $\mathcal{O}_{\mathfrak{L}}$ is generated by only the partial isometries $S_\alpha, \alpha \in \Sigma$.
If $\mathfrak{L}$ comes from a finite directed graph G, the algebra $\mathcal{O}_{\mathfrak{L}}$ becomes the Cuntz-Krieger algebra $\mathcal{O}_{A_G}$ associated to its adjacency matrix A_G with entries in $\{0,1\}$.
We generalize irreducibility and aperiodicity for finite directed graphs to λ-graph systems. Then simplicity arguments of the Cuntz algebras in [C], the Cuntz-Krieger algebras in [CK] and the C^*-algebras associated with subshifts in [Ma] are generalized to our C^*-algebras $\mathcal{O}_{\mathfrak{L}}$ so that we have

Theorem B (Theorem 4.7 and Proposition 4.9). *If $\mathfrak{L}$ satisfies condition (I) and is irreducible, the C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ is simple. In particular if $\mathfrak{L}$ is aperiodic, $\mathcal{O}_{\mathfrak{L}}$ is simple and purely infinite.*

There exists an action $\alpha_{\mathfrak{L}}$ of the torus group $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ on the algebra $\mathcal{O}_{\mathfrak{L}}$ that is called the gauge action. It satisfies $\alpha_{\mathfrak{L}z}(S_\alpha) = zS_\alpha, \alpha \in \Sigma$ for $z \in \mathbb{T}$. The fixed point subalgebra ${\mathcal{O}_{\mathfrak{L}}}^{\alpha_{\mathfrak{L}}}$ of $\mathcal{O}_{\mathfrak{L}}$ under $\alpha_{\mathfrak{L}}$ is an AF-algebra $\mathcal{F}_{\mathfrak{L}}$, that is stably isomorphic to the crossed product $\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T}$. Let $(A, I) = (A_{l,l+1}, I_{l,l+1})_{l \in \mathbb{Z}_+}$ be the nonnegative matrix system for the symbolic matrix system corresponding to the λ-graph system $\mathfrak{L}$. In [Ma5], its dimension group $(\Delta_{(A,I)}, \Delta^+_{(A,I)}, \delta_{(A,I)})$, its Bowen-Franks groups $BF^i(A, I), i = 0, 1$ and its K-groups $K_i(A, I), i = 0, 1$ have been formulated. They are related to topological conjugacy invariants of subshifts. The following K-theory formulae are generalizations of the K-theory formulae for the Cuntz-Krieger algebras and the C^*-algebras associated with subshifts ([Ma2],[Ma4],[Ma5],[Ma6], cf.[C2],[C3],[CK]).

Theorem C (Proposition 5.3, Theorem 5.5 and Theorem 5.9).

$$\begin{aligned}(K_0(\mathcal{F}_{\mathfrak{L}}), K_0(\mathcal{F}_{\mathfrak{L}})_+, \widehat{\alpha_{\mathfrak{L}}}_*) &\cong (\Delta_{(A,I)}, \Delta^+_{(A,I)}, \delta_{(A,I)}), \\ K_i(\mathcal{O}_{\mathfrak{L}}) &\cong K_i(A, I), \qquad i = 0, 1, \\ \mathrm{Ext}^{i+1}(\mathcal{O}_{\mathfrak{L}}) &\cong BF^i(A, I), \qquad i = 0, 1\end{aligned}$$

where $\widehat{\alpha_{\mathfrak{L}}}$ denotes the dual action of the gauge action $\alpha_{\mathfrak{L}}$ on $\mathcal{O}_{\mathfrak{L}}$.

We know that the C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ is nuclear and satisfies the Universal Coefficient Theorem (UCT) in the sense of Rosenberg and Schochet (Proposition 5.7)([RS], cf. [Bro2]). Hence, if $\mathfrak{L}$ is aperiodic, $\mathcal{O}_{\mathfrak{L}}$ is a unital, separable, nuclear, purely infinite, simple C^*-algebra satisfying the UCT, that lives in a classifiable class by K-theory of E. Kirchberg [Kir] and N. C. Phillips [Ph]. By Rørdam's result [Rø;Proposition 6.7], one sees that $\mathcal{O}_{\mathfrak{L}}$ is isomorphic to the C^*-algebra of an inductive limit of a sequence $B_1 \to B_2 \to B_3 \to \cdots$ of simple Cuntz-Krieger algebras (Corollary 5.8).
We finally present an example of a λ-graph system for which the associated C^*-algebra is not stably isomorphic to any Cuntz-Krieger algebra $\mathcal{O}_A$ and any Cuntz-algebra $\mathcal{O}_n$ for $n = 2, 3, \dots, \infty$. The example is a λ-graph system $\mathfrak{L}(\mathcal{S})$ constructed from a certain Shannon graph $\mathcal{S}$ (cf.[KM]). We obtain

Theorem D (Theorem 7.7). *The C^*-algebra $\mathcal{O}_{\mathfrak{L}(\mathcal{S})}$ is unital, simple, purely infinite, nuclear and generated by five partial isometries with mutually orthogonal ranges. Its K-groups are*

$$K_0(\mathcal{O}_{\mathfrak{L}(\mathcal{S})}) = 0, \qquad K_1(\mathcal{O}_{\mathfrak{L}(\mathcal{S})}) = \mathbb{Z}.$$

In [Ma7], among other things, relationships between ideals of $\mathcal{O}_{\mathfrak{L}}$ and sub λ-graph systems of $\mathfrak{L}$ are studied so that the class of C^*-algebras associated with λ-graph systems is closed under quotients by its ideals.

Acknowledgments: The author would like to thank Yasuo Watatani for his suggestions on groupoid C^*-algebras and C^*-algebras of Hilbert C^*-bimodules. The author also would like to thank the referee for his valuable suggestions and comments for the presentation of this paper.

2. Continuous graphs constructed from λ-graph systems

We will construct Deaconu's continuous graphs from λ-graph systems. They yield Renault's r-discrete groupoid C^*-algebras by Deaconu ([De],[De2],[De3]). Following V. Deaconu in [De3], by a continuous graph we mean a closed subset $\mathcal{E}$ of $\mathcal{V} \times \Sigma \times \mathcal{V}$ where $\mathcal{V}$ is a compact metric space and Σ is a finite set. If in particular $\mathcal{V}$ is zero-dimensional, that is, the set of all clopen sets form a basis of the open sets, we say $\mathcal{E}$ to be zero-dimensional or Stonean.
Let $\mathfrak{L} = (V, E, \lambda, \iota)$ be a λ-graph system over Σ with vertex set $V = \cup_{l\in\mathbb{Z}_+} V_l$ and edge set $E = \cup_{l\in\mathbb{Z}_+} E_{l,l+1}$ that is labeled with symbols in Σ by $\lambda : E \to \Sigma$, and that is supplied with surjective maps $\iota(= \iota_{l,l+1}) : V_{l+1} \to V_l$ for $l \in \mathbb{Z}_+$. Here the vertex sets $V_l, l \in \mathbb{Z}_+$ are finite disjoint sets. Also $E_{l,l+1}, l \in \mathbb{Z}_+$ are finite disjoint sets. An edge e in $E_{l,l+1}$ has its source vertex $s(e)$ in V_l and its terminal vertex $t(e)$ in V_{l+1} respectively. Every vertex in V has a successor and every vertex in V_l for $l \in \mathbb{N}$ has a predecessor. It is then required that there exists an edge in $E_{l,l+1}$ with label α and its terminal is $v \in V_{l+1}$ if and only if there exists an edge in $E_{l-1,l}$ with label α and its terminal is $\iota(v) \in V_l$. For $u \in V_{l-1}$ and $v \in V_{l+1}$, we put

$$\begin{aligned} E^{\iota}(u,v) &= \{e \in E_{l,l+1} \mid t(e) = v, \iota(s(e)) = u\}, \\ E_{\iota}(u,v) &= \{e \in E_{l-1,l} \mid s(e) = u, t(e) = \iota(v)\}. \end{aligned}$$

Then there exists a bijective correspondence between $E^{\iota}(u,v)$ and $E_{\iota}(u,v)$ that preserves labels for each pair of vertices u, v. We call this property the local property of $\mathfrak{L}$. Let $\Omega_{\mathfrak{L}}$ be the projective limit of the system $\iota_{l,l+1} : V_{l+1} \to V_l, l \in \mathbb{Z}_+$, that is defined by

$$\Omega_{\mathfrak{L}} = \{(v^l)_{l\in\mathbb{Z}_+} \in \prod_{l\in\mathbb{Z}_+} V_l \mid \iota_{l,l+1}(v^{l+1}) = v^l, l \in \mathbb{Z}_+\}.$$

We endow $\Omega_{\mathfrak{L}}$ with the projective limit topology so that it is a compact Hausdorff space. An element v in $\Omega_{\mathfrak{L}}$ is called an ι-orbit or also a vertex. Let $E_{\mathfrak{L}}$ be the set of all triplets $(u,\alpha,v) \in \Omega_{\mathfrak{L}} \times \Sigma \times \Omega_{\mathfrak{L}}$ such that for each $l \in \mathbb{Z}_+$, there exists $e_{l,l+1} \in E_{l,l+1}$ satisfying

$$u^l = s(e_{l,l+1}), \quad v^{l+1} = t(e_{l,l+1}) \quad \text{and} \quad \alpha = \lambda(e_{l,l+1})$$

where $u = (u^l)_{l\in\mathbb{Z}_+}, v = (v^l)_{l\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}}$.

Proposition 2.1. *The set $E_{\mathfrak{L}} \subset \Omega_{\mathfrak{L}} \times \Sigma \times \Omega_{\mathfrak{L}}$ is a zero-dimensional continuous graph.*

Proof. It suffices to show that $E_{\mathfrak{L}}$ is closed. For $(u,\beta,v) \in \Omega_{\mathfrak{L}} \times \Sigma \times \Omega_{\mathfrak{L}}$ with $(u,\beta,v) \notin E_{\mathfrak{L}}$, one finds $l \in \mathbb{N}$ such that there does not exist any edge e in $E_{l,l+1}$ with $s(e) = u^l, t(e) = v^{l+1}$ and $\lambda(e) = \beta$. Put

$$U_{u^l} = \{(w^i)_{i\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}} \mid w^l = u^l\}, \quad U_{v^{l+1}} = \{(w^i)_{i\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}} \mid w^{l+1} = v^{l+1}\}.$$

They are open sets in $\Omega_{\mathfrak{L}}$. Hence $U_{u^l} \times \{\beta\} \times U_{v^{l+1}}$ is an open neighborhood of (u,β,v) that does not intersect with $E_{\mathfrak{L}}$ so that $E_{\mathfrak{L}}$ is closed. $\square$

We denote by $\{\mathfrak{v}_1^l, \dots, \mathfrak{v}_{m(l)}^l\}$ the vertex set V_l. Put for $\alpha \in \Sigma, i = 1, \dots, m(1)$

$$U_i^1(\alpha) = \{(u,\alpha,v) \in E_{\mathfrak{L}} \mid v^1 = \mathfrak{v}_i^1 \text{ where } v = (v^l)_{l\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}}\}.$$

Then $U_i^1(\alpha)$ is a clopen set in $E_{\mathfrak{L}}$ such that

$$\cup_{\alpha\in\Sigma} \cup_{i=1}^{m(1)} U_i^1(\alpha) = E_{\mathfrak{L}}, \qquad U_i^1(\alpha) \cap U_j^1(\beta) = \emptyset \quad \text{if} \quad (i,\alpha) \neq (j,\beta).$$

Put $t(u,\alpha,v) = v$ for $(u,\alpha,v) \in E_{\mathfrak{L}}$. Suppose that $\mathfrak{L}$ is left-resolving. It is easy to see that if $U_i^1(\alpha) \neq \emptyset$, the restriction of t to $U_i^1(\alpha)$ is a homeomorphism onto $U_{\mathfrak{v}_i^1} = \{(v^l)_{l\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}} \mid v^1 = \mathfrak{v}_i^1\}$. Hence $t : E_{\mathfrak{L}} \to \Omega_{\mathfrak{L}}$ is a local homeomorphism.

Following Deaconu [De3], we consider the set $X_{\mathfrak{L}}$ of all one-sided paths of $E_{\mathfrak{L}}$:

$$X_{\mathfrak{L}} = \{(\alpha_i, u_i)_{i=1}^\infty \in \prod_{i=1}^\infty (\Sigma \times \Omega_{\mathfrak{L}}) \mid (u_i, \alpha_{i+1}, u_{i+1}) \in E_{\mathfrak{L}} \text{ for all } i \in \mathbb{N} \\ \text{and } (u_0, \alpha_1, u_1) \in E_{\mathfrak{L}} \text{ for some } u_0 \in \Omega_{\mathfrak{L}}\}.$$

The set $X_{\mathfrak{L}}$ has the relative topology from the infinite product topology of $\Sigma \times \Omega_{\mathfrak{L}}$. It is a zero-dimensional compact Hausdorff space. The shift map $\sigma : (\alpha_i, u_i)_{i=1}^\infty \in X_{\mathfrak{L}} \to (\alpha_{i+1}, u_{i+1})_{i=1}^\infty \in X_{\mathfrak{L}}$ is continuous. For $v = (v^l)_{l\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}}$ and $\alpha \in \Sigma$, the local property of $\mathfrak{L}$ ensures that if there exists $e_{0,1} \in E_{0,1}$ satisfying $v^1 = t(e_{0,1}), \alpha = \lambda(e_{0,1})$, there exist $e_{l,l+1} \in E_{l,l+1}$ and $u = (u^l)_{l\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}}$ satisfying $u^l = s(e_{l,l+1}), v^{l+1} = t(e_{l,l+1}), \alpha = \lambda(e_{l,l+1})$ for each $l \in \mathbb{Z}_+$. Hence if $\mathfrak{L}$ is left-resolving, for any $x = (\alpha_i, v_i)_{i=1}^\infty \in X_{\mathfrak{L}}$, there uniquely exists $v_0 \in \Omega_{\mathfrak{L}}$ such that $(v_0, \alpha_1, v_1) \in E_{\mathfrak{L}}$. Denote by $v(x)_0$ the unique vertex v_0 for $x \in X_{\mathfrak{L}}$.

Lemma 2.2. *For a λ-graph system $\mathfrak{L}$, consider the following conditions*

(i) *$\mathfrak{L}$ is left-resolving.*

(ii) *$E_{\mathfrak{L}}$ is left-resolving, that is, for $(u,\alpha,v),(u',\alpha,v') \in E_{\mathfrak{L}}$, the condition $v = v'$ implies $u = u'$.*

(iii) *σ is a local homeomorphism on $X_{\mathfrak{L}}$.*

Then we have

$$(i) \Leftrightarrow (ii) \Rightarrow (iii).$$

Proof. The implications $(i) \Leftrightarrow (ii)$ are direct. We will see that $(ii) \Rightarrow (iii)$. Suppose that $\mathfrak{L}$ is left-resolving. Let $\{\gamma_1,\dots,\gamma_m\} = \Sigma$ be the list of the alphabet. Put

$$X_{\mathfrak{L}}(k) = \{(\alpha_i, v_i)_{i=1}^{\infty} \in X_{\mathfrak{L}} \mid \alpha_1 = \gamma_k\}$$

that is a clopen set of $X_{\mathfrak{L}}$. Since the family $X_{\mathfrak{L}}(k), k = 1,\dots,m$ is a disjoint covering of $X_{\mathfrak{L}}$ and the restriction of σ to each of them $\sigma|_{X_{\mathfrak{L}}(k)} : X_{\mathfrak{L}}(k) \to X_{\mathfrak{L}}$ is a homeomorphism, the continuous surjection σ is a local homeomorphism on $X_{\mathfrak{L}}$. $\square$

Remark. We will remark that a continuous graph coming from a left-resolving, predecessor-separated λ-graph system is characterized as in the following way. Let $\mathcal{E} \subset \mathcal{V} \times \Sigma \times \mathcal{V}$ be a continuous graph. Following [KM], we define the l-past context of $v \in \mathcal{V}$ as follows:

$$\begin{aligned}\Gamma_l^-(v) = \{(\alpha_1,\dots,\alpha_l) \in \Sigma^l \mid \exists v_0, v_1, \dots, v_{l-1} \in \mathcal{V};\\ (v_{i-1},\alpha_i,v_i) \in \mathcal{E}, i = 1,2,\dots,l-1, (v_{l-1},\alpha_l,v) \in \mathcal{E}\}.\end{aligned}$$

We say $\mathcal{E}$ to be predecessor-separated if for two vertices $u, v \in \mathcal{V}$, there exists $l \in \mathbb{N}$ such that $\Gamma_l^-(u) \neq \Gamma_l^-(v)$. The following proposition can be directly proved by using an idea of [KM]. Its result will not be used in our further discussions so that we omit its proof.

Proposition 2.3. *Let $\mathcal{E} \subset \mathcal{V} \times \Sigma \times \mathcal{V}$ be a zero-dimensional continuous graph such that $\mathcal{E}$ is left-resolving, predecessor-separated. If the map $t : \mathcal{E} \to \mathcal{V}$ defined by $t(u,\alpha,v) = v$ is a surjective open map, there exists a λ-graph system $\mathcal{L}^{\mathcal{E}}$ over Σ and a homeomorphism Φ from $\mathcal{V}$ onto $\Omega_{\mathfrak{L}^{\mathcal{E}}}$ such that the map $\Phi \times \mathrm{id} \times \Phi : \mathcal{V} \times \Sigma \times \mathcal{V} \to \Omega_{\mathfrak{L}^{\mathcal{E}}} \times \Sigma \times \Omega_{\mathfrak{L}^{\mathcal{E}}}$ satisfies $(\Phi \times \mathrm{id} \times \Phi)(\mathcal{E}) = E_{\mathfrak{L}^{\mathcal{E}}}$.*

3. The C^*-algebra $\mathcal{O}_{\mathfrak{L}}$.

In what follows we assume $\mathfrak{L}$ to be left-resolving. Following V. Deaconu [De2],[De3],[De4], one may construct a locally compact r-discrete groupoid from a local homeomorphism σ on $X_{\mathfrak{L}}$ as in the following way (cf. [An],[Re]). Set

$$G_{\mathfrak{L}} = \{(x,n,y) \in X_{\mathfrak{L}} \times \mathbb{Z} \times X_{\mathfrak{L}} \mid \exists k,l \geq 0;\ \sigma^k(x) = \sigma^l(y), n = k - l\}.$$

The range map and the domain map are defined by

$$r(x,n,y) = x, \qquad d(x,n,y) = y.$$

The multiplication and the inverse operation are defined by

$$(x,n,y)(y,m,z) = (x,n+m,z), \qquad (x,n,y)^{-1} = (y,-n,x).$$

The unit space $G^0_{\mathfrak{L}}$ is defined to be the space $X_{\mathfrak{L}} = \{(x,0,x) \in G_{\mathfrak{L}} \mid x \in X_{\mathfrak{L}}\}$. A basis of the open sets for $G_{\mathfrak{L}}$ is given by

$$Z(U,V,k,l) = \{(x, k-l, (\sigma^l|_V)^{-1} \circ (\sigma^k(x)) \in G_{\mathfrak{L}} \mid x \in U\}$$

where U,V are open sets of $X_{\mathfrak{L}}$, and $k,l \in \mathbb{N}$ are such that $\sigma^k|_U$ and $\sigma^l|_V$ are homeomorphisms with the same open range. Hence we see

$$Z(U,V,k,l) = \{(x,k-l,y) \in G_{\mathfrak{L}} \mid x \in U, y \in V, \sigma^k(x) = \sigma^l(y)\}.$$

The groupoid C^*-algebra $C^*(G_{\mathfrak{L}})$ for the groupoid $G_{\mathfrak{L}}$ is defined as in the following way ([Re], cf. [An],[De2],[De3],[De4]). Let $C_c(G_{\mathfrak{L}})$ be the set of all continuous functions on $G_{\mathfrak{L}}$ with compact support that has a natural product structure of $*$-algebra given by

$$(f*g)(s) = \sum_{\substack{t \in G_{\mathfrak{L}}, \\ r(t)=r(s)}} f(t)g(t^{-1}s) = \sum_{\substack{t_1,t_2 \in G_{\mathfrak{L}}, \\ s=t_1t_2}} f(t_1)g(t_2),$$

$$f^*(s) = \overline{f(s^{-1})}, \qquad f,g \in C_c(G_{\mathfrak{L}}), \quad s \in G_{\mathfrak{L}}.$$

Let $C_0(G^0_{\mathfrak{L}})$ be the C^*-algebra of all continuous functions on $G^0_{\mathfrak{L}}$ that vanish at infinity. The algebra $C_c(G_{\mathfrak{L}})$ is a $C_0(G^0_{\mathfrak{L}})$-module, endowed with a $C_0(G^0_{\mathfrak{L}})$-valued inner product by

$$\begin{aligned}(\xi f)(x,n,y) &= \xi(x,n,y)f(y), \qquad \xi \in C_c(G_{\mathfrak{L}}), \quad f \in C_0(G^0_{\mathfrak{L}}), \quad (x,n,y) \in G_{\mathfrak{L}}, \\ <\xi,\eta>(y) &= \sum_{\substack{x,n \\ (x,n,y) \in G_{\mathfrak{L}}}} \overline{\xi(x,n,y)}\eta(x,n,y), \qquad \xi,\eta \in C_c(G_{\mathfrak{L}}), \quad y \in X_{\mathfrak{L}}.\end{aligned}$$

Let us denote by $l^2(G_{\mathfrak{L}})$ the completion of the inner product $C_0(G^0_{\mathfrak{L}})$-module $C_c(G_{\mathfrak{L}})$. It is a Hilbert C^*-right module over the commutative C^*-algebra $C_0(G^0_{\mathfrak{L}})$. We denote by $B(l^2(G_{\mathfrak{L}}))$ the C^*-algebra of all bounded adjointable $C_0(G^0_{\mathfrak{L}})$-module maps on $l^2(G_{\mathfrak{L}})$. Let π be the $*$-homomorphism of $C_c(G_{\mathfrak{L}})$ into $B(l^2(G_{\mathfrak{L}}))$ defined by $\pi(f)\xi = f*\xi$ for $f,\xi \in C_c(G_{\mathfrak{L}})$. Then the closure of $\pi(C_c(G_{\mathfrak{L}}))$ in $B(l^2(G_{\mathfrak{L}}))$ is called the (reduced) C^*-algebra of the groupoid $G_{\mathfrak{L}}$, that we denote by $C^*(G_{\mathfrak{L}})$.

Definition. *The C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ associated with λ-graph system $\mathfrak{L}$ is defined to be the C^*-algebra $C^*(G_{\mathfrak{L}})$ of the groupoid $G_{\mathfrak{L}}$.*

We will study the algebraic structure of the C^*-algebra $\mathcal{O}_{\mathfrak{L}}$. Recall that Λ^k denotes the set of all words of Σ^k that appear in $\mathfrak{L}$. For $x = (\alpha_n, u_n)_{n=1}^\infty \in X_{\mathfrak{L}}$, we put $\lambda(x)_n = \alpha_n \in \Sigma$, $v(x)_n = u_n \in \Omega_{\mathfrak{L}}$ respectively. The ι-orbit $v(x)_n$

is written as $v(x)_n = (v(x)^l_n)_{l\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}} = \varprojlim V_l$. Now $\mathfrak{L}$ is left-resolving so that there uniquely exists $v(x)_0 \in \Omega_{\mathfrak{L}}$ satisfying $(v(x)_0, \alpha_1, u_1) \in E_{\mathfrak{L}}$. Set for $\mu = (\mu_1, \dots, \mu_k) \in \Lambda^k$,

$$U(\mu) = \{(x,k,y) \in G_{\mathfrak{L}} \mid \sigma^k(x) = y, \lambda(x)_1 = \mu_1, \dots, \lambda(x)_k = \mu_k\}$$

and for $\mathfrak{v}^l_i \in V_l$,

$$U(\mathfrak{v}^l_i) = \{(x,0,x) \in G_{\mathfrak{L}} \mid v(x)^l_0 = \mathfrak{v}^l_i\}$$

where $v(x)_0 = (v(x)^l_0)_{l\in\mathbb{Z}_+} \in \Omega_{\mathfrak{L}}$. They are clopen sets of $G_{\mathfrak{L}}$. We set

$$S_\mu = \pi(\chi_{U(\mu)}), \qquad E^l_i = \pi(\chi_{U(\mathfrak{v}^l_i)}) \qquad \text{in} \qquad \pi(C_c(G_{\mathfrak{L}}))$$

where $\chi_F \in C_c(G_{\mathfrak{L}})$ denotes the characteristic function of a clopen set F on the space $G_{\mathfrak{L}}$. Then it is straightforward to see the following lemmas.

Lemma 3.1.

(i) *S_μ is a partial isometry satisfying $S_\mu = S_{\mu_1}\cdots S_{\mu_k}$, where $\mu = (\mu_1, \dots, \mu_k) \in \Lambda^k$.*

(ii) *$\sum_{\mu\in\Lambda^k} S_\mu S^*_\mu = 1$ for $k \in \mathbb{N}$. We in particular have*

$$\sum_{\alpha\in\Sigma} S_\alpha S^*_\alpha = 1. \tag{3.1}$$

(iii) *E^l_i is a projection such that*

$$\sum_{i=1}^{m(l)} E^l_i = 1, \qquad E^l_i = \sum_{j=1}^{m(l+1)} I_{l,l+1}(i,j) E^{l+1}_j, \tag{3.2}$$

where $I_{l,l+1}$ is the matrix defined in Theorem A in Section 1, corresponding to the map $\iota_{l,l+1} : V_{l+1} \to V_l$.

Take $\mu = (\mu_1, \dots, \mu_k) \in \Lambda^k$, $\nu = (\nu_1, \dots, \nu_{k'}) \in \Lambda^{k'}$ and $\mathfrak{v}^l_i \in V_l$ with $k, k' \le l$ such that there exist paths ξ, η in $\mathfrak{L}$ satisfying $\lambda(\xi) = \mu, \lambda(\eta) = \nu$ and $t(\xi) = t(\eta) = \mathfrak{v}^l_i$. We set

$$\begin{aligned} U(\mu, \mathfrak{v}^l_i, \nu) =& \{(x, k-k', y) \in G_{\mathfrak{L}} \mid \sigma^k(x) = \sigma^{k'}(y),\ v(x)^l_k = v(y)^l_{k'} = \mathfrak{v}^l_i, \\ & \lambda(x)_1 = \mu_1, \dots, \lambda(x)_k = \mu_k,\ \lambda(y)_1 = \nu_1, \dots, \lambda(y)_{k'} = \nu_{k'}\ \}. \end{aligned}$$

The sets $U(\mu, \mathfrak{v}^l_i, \nu), \mu \in \Lambda^k, \nu \in \Lambda^{k'}, i = 1, \dots, m(l)$ are clopen sets and generate the topology of $G_{\mathfrak{L}}$.

Lemma 3.2.

$$S_\mu E^l_i S^*_\nu = \pi(\chi_{U(\mu,\mathfrak{v}^l_i,\nu)}) \in \pi(C_c(G_{\mathfrak{L}})).$$

Hence the C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ is generated by $S_\alpha, \alpha \in \Sigma$ and $E^l_i, i = 1, \dots, m(l), l \in \mathbb{Z}_+$.

The generators S_α, E^l_i satisfy the following operator relations, that are straightforwardly checked.

LEMMA 3.3.

$$S_\alpha S_\alpha^* E_i^l = E_i^l S_\alpha S_\alpha^*, \tag{3.3}$$

$$S_\alpha^* E_i^l S_\alpha = \sum_{j=1}^{m(l+1)} A_{l,l+1}(i,\alpha,j) E_j^{l+1}, \tag{3.4}$$

for $\alpha \in \Sigma$, $i = 1,2,\dots,m(l), l \in \mathbb{Z}_+$, *where* $A_{l,l+1}(i,\alpha,j)$ *is defined in Theorem A in Section 1.*

The four operator relations (3.1),(3.2),(3.3),(3.4) are called the relations $(\mathfrak{L})$. Let $\mathcal{A}_l, l \in \mathbb{Z}_+$ be the C^*-subalgebra of $\mathcal{O}_{\mathfrak{L}}$ generated by the projections $E_i^l, i = 1,\dots,m(l)$, that is,

$$\mathcal{A}_l = \mathbb{C}E_1^l \oplus \cdots \oplus \mathbb{C}E_{m(l)}^l.$$

The projections $S_\alpha^* S_\alpha, \alpha \in \Sigma$ and $S_\mu^* S_\mu, \mu \in \Lambda^k, k \le l$ belong to $\mathcal{A}_l, l \in \mathbb{N}$ by (3.4) and the first relation of (3.2). Let $\mathcal{A}_{\mathfrak{L}}$ be the C^*-subalgebra of $\mathcal{O}_{\mathfrak{L}}$ generated by all the projections $E_i^l, i = 1,\dots,m(l), l \in \mathbb{Z}_+$. By the second relation of (3.2), the algebra $\mathcal{A}_l$ is naturally embedded in $\mathcal{A}_{l+1}$ so that $\mathcal{A}_{\mathfrak{L}}$ is a commutative AF-algebra. We note that there exists an isomorphism between $\mathcal{A}_l$ and $C(V_l)$ for each $l \in \mathbb{Z}_+$ that is compatible with the embeddings $\mathcal{A}_l \hookrightarrow \mathcal{A}_{l+1}$ and $I_{l,l+1}^t (= \iota_{l,l+1}^*) : C(V_l) \hookrightarrow C(V_{l+1})$. Hence there exists an isomorphism between $\mathcal{A}_{\mathfrak{L}}$ and $C(\Omega_{\mathfrak{L}})$. Let k,l be natural numbers with $k \le l$. We set

$\mathcal{D}_{\mathfrak{L}}$ =The C^*-subalgebra of $\mathcal{O}_{\mathfrak{L}}$ generated by $S_\mu a S_\mu^*, \mu \in \Lambda^*, a \in \mathcal{A}_{\mathfrak{L}}$.

$\mathcal{F}_k^l$ =The C^*-subalgebra of $\mathcal{O}_{\mathfrak{L}}$ generated by $S_\mu a S_\nu^*, \mu,\nu \in \Lambda^k, a \in \mathcal{A}_l$.

$\mathcal{F}_k^\infty$ =The C^*-subalgebra of $\mathcal{O}_{\mathfrak{L}}$ generated by $S_\mu a S_\nu^*, \mu,\nu \in \Lambda^k, a \in \mathcal{A}_{\mathfrak{L}}$.

$\mathcal{F}_{\mathfrak{L}}$ =The C^*-subalgebra of $\mathcal{O}_{\mathfrak{L}}$ generated by $S_\mu a S_\nu^*, \mu,\nu \in \Lambda^*$, $|\mu| = |\nu|, a \in \mathcal{A}_{\mathfrak{L}}$.

The algebra $\mathcal{D}_{\mathfrak{L}}$ is isomorphic to $C(X_{\mathfrak{L}})$. It is obvious that the algebra $\mathcal{F}_k^l$ is finite dimensional and there exists an embedding $\iota_{l,l+1} : \mathcal{F}_k^l \hookrightarrow \mathcal{F}_k^{l+1}$ through the preceding embedding $\mathcal{A}_l \hookrightarrow \mathcal{A}_{l+1}$. Define a homomorphism $c : (x,n,y) \in G_{\mathfrak{L}} \to n \in \mathbb{Z}$. We denote by $F_{\mathfrak{L}}$ the subgroupoid $c^{-1}(0)$ of $G_{\mathfrak{L}}$. Let $C^*(F_{\mathfrak{L}})$ be its groupoid C^*-algebra. It is also immediate that the algebra $\mathcal{F}_{\mathfrak{L}}$ is isomorphic to $C^*(F_{\mathfrak{L}})$. By (3.1),(3.3),(3.4), the relations:

$$E_i^l = \sum_{\alpha \in \Sigma} \sum_{j=1}^{m(l+1)} A_{l,l+1}(i,\alpha,j) S_\alpha E_j^{l+1} S_\alpha^*, \qquad i = 1,2,\dots,m(l)$$

hold. They yield

$$S_\mu E_i^l S_\nu^* = \sum_{\alpha \in \Sigma} \sum_{j=1}^{m(l+1)} A_{l,l+1}(i,\alpha,j) S_{\mu\alpha} E_j^{l+1} S_{\nu\alpha}^* \qquad \text{for } \mu,\nu \in \Lambda^k,$$

that give rise to an embedding $\mathcal{F}_k^l \hookrightarrow \mathcal{F}_{k+1}^{l+1}$. It induces an embedding of $\mathcal{F}_k^\infty$ into $\mathcal{F}_{k+1}^\infty$ that we denote by $\lambda_{k,k+1}$.

Proposition 3.4.

(i) $\mathcal{F}_k^\infty$ *is an AF-algebra defined by the inductive limit of the embeddings* $\iota_{l,l+1} : \mathcal{F}_k^l \hookrightarrow \mathcal{F}_k^{l+1}, l \in \mathbb{N}$.

(ii) $\mathcal{F}_{\mathfrak{L}}$ *is an AF-algebra defined by the inductive limit of the embeddings* $\lambda_{k,k+1} : \mathcal{F}_k^\infty \hookrightarrow \mathcal{F}_{k+1}^\infty, k \in \mathbb{Z}_+$.

Let $U_z, z \in \mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ be an action of $\mathbb{T}$ to the unitary group of $B(l^2(G_{\mathfrak{L}}))$ defined by

$$(U_z\xi)(x,n,y) = z^n\xi(x,n,y) \quad \text{for} \quad \xi \in l^2(G_{\mathfrak{L}}), (x,n,y) \in G_{\mathfrak{L}}.$$

The action $Ad(U_z)$ on $B(l^2(G_{\mathfrak{L}}))$ leaves $\mathcal{O}_{\mathfrak{L}}$ globally invariant. It gives rise to an action on $\mathcal{O}_{\mathfrak{L}}$. We denote it by $\alpha_{\mathfrak{L}}$ and call it the gauge action. Let $E_{\mathfrak{L}}$ be the expectation from $\mathcal{O}_{\mathfrak{L}}$ onto the fixed point subalgebra ${\mathcal{O}_{\mathfrak{L}}}^{\alpha_{\mathfrak{L}}}$ under $\alpha_{\mathfrak{L}}$ defined by

$$E_{\mathfrak{L}}(X) = \int_{z\in\mathbb{T}} {\alpha_{\mathfrak{L}}}_z(X)dz, \qquad X \in \mathcal{O}_{\mathfrak{L}}. \tag{3.5}$$

Let $\mathcal{P}_{\mathfrak{L}}$ be the $*$-algebra generated algebraically by $S_\alpha, \alpha \in \Sigma$ and $E_i^l, i = 1,\dots,m(l)$, $l \in \mathbb{Z}_+$. For $\mu = (\mu_1,\dots,\mu_k) \in \Lambda^k$, it follows that by (3.3), $E_i^l S_{\mu_1}\cdots S_{\mu_k} = S_{\mu_1}S_{\mu_1}^* E_i^l S_{\mu_1}\cdots S_{\mu_k}$. As $S_{\mu_1}^* E_i^l S_{\mu_1}$ is a linear combination of $E_j^{l+1}, j = 1,\dots,m(l+1)$ by (3.4), one sees $S_{\mu_1}^* E_i^l S_{\mu_1} S_{\mu_2} = S_{\mu_2}S_{\mu_2}^* S_{\mu_1}^* E_i^l S_{\mu_1} S_{\mu_2}$ and inductively

$$E_i^l S_\mu = S_\mu S_\mu^* E_i^l S_\mu, \qquad E_i^l S_\mu S_\mu^* = S_\mu S_\mu^* E_i^l. \tag{3.6}$$

By the relations (3.6), each element $X \in \mathcal{P}_{\mathfrak{L}}$ is expressed as a finite sum

$$X = \sum_{|\nu|\ge 1} X_{-\nu}S_\nu^* + X_0 + \sum_{|\mu|\ge 1} S_\mu X_\mu \qquad \text{for some} \quad X_{-\nu}, X_0, X_\mu \in \mathcal{F}_{\mathfrak{L}}.$$

Then the following lemma is routine.

Lemma 3.5. *The fixed point subalgebra* ${\mathcal{O}_{\mathfrak{L}}}^{\alpha_{\mathfrak{L}}}$ *of* $\mathcal{O}_{\mathfrak{L}}$ *under* $\alpha_{\mathfrak{L}}$ *is the AF-algebra* $\mathcal{F}_{\mathfrak{L}}$.

We can now prove a universal property of $\mathcal{O}_{\mathfrak{L}}$.

Theorem 3.6. *The C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ is the universal C^*-algebra subject to the relations* $(\mathfrak{L})$.

Proof. Let $\mathcal{O}_{[\mathfrak{L}]}$ be the universal C^*-algebra generated by partial isometries $s_\alpha, \alpha \in \Sigma$ and projections $e_i^l, i = 1,\dots,m(l), l \in \mathbb{Z}_+$ subject to the operator relations $(\mathfrak{L})$. This means that $\mathcal{O}_{[\mathfrak{L}]}$ is generated by $s_\alpha, \alpha \in \Sigma$ and $e_i^l, i = 1,\dots,m(l), l \in \mathbb{Z}_+$, that have only operator relations $(\mathfrak{L})$. The C^*-norm of $\mathcal{O}_{[\mathfrak{L}]}$ is given by the universal C^*-norm. Let us denote by $\mathcal{F}_{[k]}^{[l]}, \mathcal{F}_{[\mathfrak{L}]}$

the similarly defined subalgebras of $\mathcal{O}_{[\mathfrak{L}]}$ to $\mathcal{F}_k^l, \mathcal{F}_{\mathfrak{L}}$ respectively. The algebra $\mathcal{F}_{[k]}^{[l]}$ as well as $\mathcal{F}_k^l$ is a finite dimensional algebra. Since $s_\mu e_i^l s_\nu^* \neq 0$ if and only if $S_\mu E_i^l S_\nu^* \neq 0$, the correspondence $s_\mu e_i^l s_\nu^* \to S_\mu E_i^l S_\nu^*, |\mu| = |\nu| = k \le l$ yields an isomorphism from $\mathcal{F}_{[k]}^{[l]}$ to $\mathcal{F}_k^l$. It induces an isomorphism from $\mathcal{F}_{[\mathfrak{L}]}$ to $\mathcal{F}_{\mathfrak{L}}$. By the universality, for $z \in \mathbb{C}, |z| = 1$ the correspondence $s_\alpha \to z s_\alpha, \alpha \in \Sigma$, $e_i^l \to e_i^l, i = 1, \dots, m(l), l \in \mathbb{Z}_+$ gives rise to an action of the torus group $\mathbb{T}$ on $\mathcal{O}_{[\mathfrak{L}]}$, which we denote by $\alpha_{[\mathfrak{L}]}$. Let $E_{[\mathfrak{L}]}$ be the expectation from $\mathcal{O}_{[\mathfrak{L}]}$ onto the fixed point subalgebra ${\mathcal{O}_{[\mathfrak{L}]}}^{\alpha_{[\mathfrak{L}]}}$ under $\alpha_{[\mathfrak{L}]}$ similarly defined to (3.5). The algebra ${\mathcal{O}_{[\mathfrak{L}]}}^{\alpha_{[\mathfrak{L}]}}$ is nothing but the algebra $\mathcal{F}_{[\mathfrak{L}]}$. By the universality of $\mathcal{O}_{[\mathfrak{L}]}$, the correspondence $s_\alpha \to S_\alpha, \alpha \in \Sigma$, $e_i^l \to E_i^l, i = 1, \dots, m(l), l \in \mathbb{Z}_+$ extends to a surjective homomorphism from $\mathcal{O}_{[\mathfrak{L}]}$ to $\mathcal{O}_{\mathfrak{L}}$, which we denote by $\pi_{\mathfrak{L}}$. The restriction of $\pi_{\mathfrak{L}}$ to $\mathcal{F}_{[\mathfrak{L}]}$ is the preceding isomorphism. As we see that $E_{\mathfrak{L}} \circ \pi_{\mathfrak{L}} = \pi_{\mathfrak{L}} \circ E_{[\mathfrak{L}]}$ and $E_{[\mathfrak{L}]}$ is faithful, we conclude that $\pi_{\mathfrak{L}}$ is isomorphic by a similar argument to [CK; 2.12. Proposition]. $\square$

4. Uniqueness and simplicity

We will prove that $\mathcal{O}_{\mathfrak{L}}$ is the unique C^*-algebra subject to the operator relations $(\mathfrak{L})$ under a mild condition on $\mathfrak{L}$, called (I). The condition (I) is a generalization of condition (I) for a finite square matrix with entries in $\{0,1\}$ defined by Cuntz-Krieger in [CK] and condition (I) for a subshift defined in [Ma4]. A related condition for a Hilbert C^*-bimodule has been introduced by Kajiwara-Pinzari-Watatani in [KPW]. For an infinite directed graph, such a condition is defined by Kumjian-Pask-Raeburn-Renault in [KPRR]. For a vertex $\mathfrak{v}_i^l \in V_l$, let $\Gamma^+(\mathfrak{v}_i^l)$ be the set of all label sequences in $\mathfrak{L}$ starting at $\mathfrak{v}_i^l$. That is,

$$\Gamma^+(\mathfrak{v}_i^l) = \{(\alpha_1, \alpha_2, \dots,) \in \Sigma^{\mathbb{N}} \mid \exists e_{n,n+1} \in E_{n,n+1} \text{ for } n = l, l+1, \dots; \\ \mathfrak{v}_i^l = s(e_{l,l+1}),\ t(e_{n,n+1}) = s(e_{n+1,n+2}),\ \lambda(e_{n,n+1}) = \alpha_{n-l+1}\}.$$

DEFINITION. *A λ-graph system $\mathfrak{L}$ satisfies condition (I) if for each $\mathfrak{v}_i^l \in V$, the set $\Gamma^+(\mathfrak{v}_i^l)$ contains at least two distinct sequences.*

For $\mathfrak{v}_i^l \in V_l$ set $F_i^l = \{x \in X_{\mathfrak{L}} \mid v(x)_0^l = \mathfrak{v}_i^l\}$ where $v(x)_0 = (v(x)_0^l)_{l \in \mathbb{Z}_+} \in \Omega_{\mathfrak{L}} = \varprojlim V_l$ is the unique ι-orbit for $x \in X_{\mathfrak{L}}$ such that $(v(x)_0, \lambda(x)_1, v(x)_1) \in E_{\mathfrak{L}}$ as in the preceding section. By a similar discussion to [Ma4; Section 5] (cf.[CK; 2.6.Lemma]), we know that if $\mathfrak{L}$ satisfies (I), for $l, k \in \mathbb{N}$ with $l \ge k$, there exists $y_i^l \in F_i^l$ for each $i = 1, 2, \dots, m(l)$ such that

$$\sigma^m(y_i^l) \neq y_j^l \qquad \text{for all} \quad 1 \le i, j \le m(l), \quad 1 \le m \le k.$$

By the same manner as the proof of [Ma4;Lemma 5.3], we obtain

LEMMA 4.1. *Suppose that $\mathfrak{L}$ satisfies condition (I). Then for $l, k \in \mathbb{N}$ with $l \ge k$, there exists a projection $q_k^l \in \mathcal{D}_{\mathfrak{L}}$ such that*

(i) $q_k^l a \neq 0$ *for all nonzero* $a \in \mathcal{A}_l$,
(ii) $q_k^l \phi_{\mathfrak{L}}^m(q_k^l) = 0$ *for all* $m = 1, 2, \dots, k$, *where* $\phi_{\mathfrak{L}}^m(X) = \sum_{\mu \in \Lambda^m} S_\mu X S_\mu^*$.

Now we put $Q_k^l = \phi_{\mathfrak{L}}^k(q_k^l)$ a projection in $\mathcal{D}_{\mathfrak{L}}$. Note that each element of $\mathcal{D}_{\mathfrak{L}}$ commutes with elements of $\mathcal{A}_{\mathfrak{L}}$. As we see $S_\mu \phi_{\mathfrak{L}}^j(X) = \phi_{\mathfrak{L}}^{j+|\mu|}(X) S_\mu$ for $X \in \mathcal{D}_{\mathfrak{L}}, j \in \mathbb{Z}_+, \mu \in \Lambda^*$, a similar argument to [CK;2.9.Proposition] leads to the following lemma.

Lemma 4.2.

(i) *The correspondence:* $X \in \mathcal{F}_k^l \longrightarrow Q_k^l X Q_k^l \in Q_k^l \mathcal{F}_k^l Q_k^l$ *extends to an isomorphism from* $\mathcal{F}_k^l$ *onto* $Q_k^l \mathcal{F}_k^l Q_k^l$.

(ii) $Q_k^l X - X Q_k^l \to 0$, $\|Q_k^l X\| \to \|X\|$ *as* $k, l \to \infty$ *for* $X \in \mathcal{F}_{\mathfrak{L}}$.

(iii) $Q_k^l S_\mu Q_k^l, Q_k^l S_\mu^* Q_k^l \to 0$ *as* $k, l \to \infty$ *for* $\mu \in \Lambda^*$.

We then prove the uniqueness of the algebra $\mathcal{O}_{\mathfrak{L}}$ subject to the relations $(\mathfrak{L})$.

Theorem 4.3. *Suppose that* $\mathfrak{L}$ *satisfies condition (I). Let* $\widehat{S}_\alpha, \alpha \in \Sigma$ *and* $\widehat{E}_i^l, i = 1, 2, \dots, m(l), l \in \mathbb{Z}_+$ *be another family of nonzero partial isometries and nonzero projections satisfying the relations* $(\mathfrak{L})$. *Then the map* $S_\alpha \to \widehat{S}_\alpha, \alpha \in \Sigma$, $E_i^l \to \widehat{E}_i^l, i = 1, \dots, m(l), l \in \mathbb{Z}_+$ *extends to an isomorphism from* $\mathcal{O}_{\mathfrak{L}}$ *onto the* C^**-algebra* $\widehat{\mathcal{O}}_{\mathfrak{L}}$ *generated by* $\widehat{S}_\alpha, \alpha \in \Sigma$ *and* $\widehat{E}_i^l, i = 1, \dots, m(l), l \in \mathbb{Z}_+$.

Proof. We may define C^*-subalgebras $\widehat{\mathcal{D}}_{\mathfrak{L}}, \widehat{\mathcal{F}}_k^l, \widehat{\mathcal{F}}_{\mathfrak{L}}$ of $\widehat{\mathcal{O}}_{\mathfrak{L}}$ by using the elements $\widehat{S}_\mu \widehat{E}_i^l \widehat{S}_\nu^*$ by the same manners as the constructions of the C^*-subalgebras $\mathcal{D}_{\mathfrak{L}}, \mathcal{F}_k^l, \mathcal{F}_{\mathfrak{L}}$ of $\mathcal{O}_{\mathfrak{L}}$ respectively. As in the proof of Theorem 3.6, the map $S_\mu E_i^l S_\nu^* \in \mathcal{F}_k^l \to \widehat{S}_\mu \widehat{E}_i^l \widehat{S}_\nu^* \in \widehat{\mathcal{F}}_k^l, |\mu| = |\nu| = k \le l$ extends to an isomorphism from the AF-algebra $\mathcal{F}_{\mathfrak{L}}$ onto the AF-algebra $\widehat{\mathcal{F}}_{\mathfrak{L}}$. By Theorem 3.6, the algebra $\mathcal{O}_{\mathfrak{L}}$ has a universal property subject to the relations $(\mathfrak{L})$ so that there exists a surjective homomorphism $\hat{\pi}$ from $\mathcal{O}_{\mathfrak{L}}$ onto $\widehat{\mathcal{O}}_{\mathfrak{L}}$ satisfying $\hat{\pi}(S_\alpha) = \widehat{S}_\alpha$ and $\hat{\pi}(E_i^l) = \widehat{E}_i^l$. The restriction of $\hat{\pi}$ to $\mathcal{F}_{\mathfrak{L}}$ is the preceding isomorphism onto $\widehat{\mathcal{F}}_{\mathfrak{L}}$. Now $\mathfrak{L}$ satisfies (I). Let Q_k^l be the sequence of projections as in Lemma 4.2. We put $\widehat{Q}_k^l = \hat{\pi}(Q_k^l) \in \widehat{\mathcal{D}}_{\mathfrak{L}}$ that has the corresponding properties to Lemma 4.2 for the algebra $\widehat{\mathcal{F}}_{\mathfrak{L}}$. Let $\widehat{\mathcal{P}}_{\mathfrak{L}}$ be the $*$-algebra generated algebraically by $\widehat{S}_\alpha, \alpha \in \Sigma$ and $\widehat{E}_i^l, i = 1, \dots, m(l), l \in \mathbb{Z}_+$. By the relations $(\mathfrak{L})$, each element $X \in \widehat{\mathcal{P}}_{\mathfrak{L}}$ is expressed as a finite sum

$$X = \sum_{|\nu| \ge 1} X_{-\nu} \widehat{S}_\nu^* + X_0 + \sum_{|\mu| \ge 1} \widehat{S}_\mu X_\mu \qquad \text{for some} \qquad X_{-\nu}, X_0, X_\mu \in \widehat{\mathcal{F}}_{\mathfrak{L}}.$$

By a similar argument to [CK;2.9.Proposition], it follows that the map $X \in \widehat{\mathcal{P}}_{\mathfrak{L}} \to X_0 \in \widehat{\mathcal{F}}_{\mathfrak{L}}$ extends to an expectation $\widehat{E}_{\mathfrak{L}}$ from $\widehat{\mathcal{O}}_{\mathfrak{L}}$ onto $\widehat{\mathcal{F}}_{\mathfrak{L}}$, that satisfies $\widehat{E}_{\mathfrak{L}} \circ \hat{\pi} = \hat{\pi} \circ E_{\mathfrak{L}}$. As $E_{\mathfrak{L}}$ is faithful, we conclude that $\hat{\pi}$ is isomorphic. □

Remark. Let e_x be a vector assigned to $x \in X_{\mathfrak{L}}$. Let $\mathfrak{H}_{\mathfrak{L}}$ be the Hilbert space spanned by the vectors $e_x, x \in X_{\mathfrak{L}}$ such that the vectors $e_x, x \in X_{\mathfrak{L}}$ form its complete orthonormal basis. For $x = (\alpha_i, v_i)_{i=1}^\infty \in X_{\mathfrak{L}}$, take $v_0 = v(x)_0 \in \Omega_{\mathfrak{L}}$. For a symbol $\beta \in \Sigma$, if there exists a vertex $v_{-1} \in \Omega_{\mathfrak{L}}$ such that $(v_{-1}, \beta, v_0) \in E_{\mathfrak{L}}$, we define $\beta x \in X_{\mathfrak{L}}$, by putting $\alpha_0 = \beta$, as

$$\beta x = (\alpha_{i-1}, v_{i-1})_{i=1}^\infty \in X_{\mathfrak{L}}.$$

Put

$$\Gamma_1^-(x) = \{\gamma \in \Sigma \mid (v_{-1}, \gamma, v(x)_0) \in E_{\mathfrak{L}} \text{ for some } v_{-1} \in \Omega_{\mathfrak{L}}\}.$$

We define the creation operators $\widetilde{S}_\beta, \beta \in \Sigma$ on $\mathfrak{H}_{\mathfrak{L}}$ by

$$\widetilde{S}_\beta e_x = \begin{cases} e_{\beta x} & \text{if } \beta \in \Gamma_1^-(x), \\ 0 & \text{if } \beta \notin \Gamma_1^-(x). \end{cases}$$

PROPOSITION 4.4. *Suppose that $\mathfrak{L}$ satisfies condition (I). If $\mathfrak{L}$ is predecessor-separated, $\mathcal{O}_{\mathfrak{L}}$ is isomorphic to the C^*-algebra $C^*(\widetilde{S}_\beta, \beta \in \Sigma)$ generated by the partial isometries $\widetilde{S}_\beta, \beta \in \Sigma$ on the Hilbert space $\mathfrak{H}_{\mathfrak{L}}$.*

Proof. Suppose that $\mathfrak{L}$ is predecessor-separated. Define a sequence of projections $\widetilde{E}_i^l, i = 1, \dots, m(l), l \in \mathbb{N}$ and $\widetilde{E}_i^0, i = 1, \dots, m(0)$ by using the formulae (1.7) from the partial isometries $\widetilde{S}_\beta, \beta \in \Sigma$. It is straightforward to see that $\widetilde{E}_i^l, i = 1, \dots, m(l), l \in \mathbb{Z}_+$ are nonzero. The partial isometries $\widetilde{S}_\beta$ and the projections $\widetilde{E}_i^l$ satisfy the relations $(\mathfrak{L})$. □

Let Λ be a subshift and $\mathfrak{L}^\Lambda$ its canonical λ-graph system, that is left-resolving and predecessor-separated. It is easy to see that Λ satisfies condition (I) in the sense of [Ma4] if and only if $\mathfrak{L}^\Lambda$ satisfies condition (I).

COROLLARY 4.5(CF.[MA],[CAM]). *The C^*-algebra $\mathcal{O}_{\mathfrak{L}^\Lambda}$ associated with λ-graph system $\mathfrak{L}^\Lambda$ is canonically isomorphic to the C^*-algebra $\mathcal{O}_\Lambda$ associated with subshift Λ.*

We next refer simplicity and purely infiniteness of the algebra $\mathcal{O}_{\mathfrak{L}}$. We introduce the notions of irreducibility and aperiodicity for λ-graph system

DEFINITION.

(i) *A λ-graph system $\mathfrak{L}$ is said to be irreducible if for a vertex $v \in V_l$ and $x = (x_1, x_2, \dots) \in \Omega_{\mathfrak{L}} = \varprojlim V_l$, there exists a path in $\mathfrak{L}$ starting at v and terminating at x_{l+N} for some $N \in \mathbb{N}$.*

(ii) *A λ-graph system $\mathfrak{L}$ is said to be aperiodic if for a vertex $v \in V_l$ there exists an $N \in \mathbb{N}$ such that there exist paths in $\mathfrak{L}$ starting at v and terminating at all the vertices of V_{l+N}.*

Aperiodicity automatically implies irreducibility. Define a positive operator $\lambda_{\mathfrak{L}}$ on $\mathcal{A}_{\mathfrak{L}}$ by

$$\lambda_{\mathfrak{L}}(X) = \sum_{\alpha \in \Sigma} S_\alpha^* X S_\alpha \qquad \text{for} \qquad X \in \mathcal{A}_{\mathfrak{L}}.$$

We say that $\lambda_{\mathfrak{L}}$ is *irreducible* if there exists no non-trivial ideal of $\mathcal{A}_{\mathfrak{L}}$ invariant under $\lambda_{\mathfrak{L}}$, and $\lambda_{\mathfrak{L}}$ is *aperiodic* if for a projection $E_i^l \in \mathcal{A}_l$ there exists $N \in \mathbb{N}$ such that $\lambda_{\mathfrak{L}}^N(E_i^l) \geq 1$. The following lemma is easy to prove (cf.[Ma4]).

Lemma 4.6.

(i) *A λ-graph system $\mathfrak{L}$ is irreducible if and only if $\lambda_{\mathfrak{L}}$ is irreducible.*
(ii) *A λ-graph system $\mathfrak{L}$ is aperiodic if and only if $\lambda_{\mathfrak{L}}$ is aperiodic.*

We thus obtain

Theorem 4.7. *Suppose that a λ-graph system $\mathfrak{L}$ satisfies condition (I). If $\mathfrak{L}$ is irreducible, $\mathcal{O}_{\mathfrak{L}}$ is simple.*

Proof. Suppose that there exists a nonzero ideal $\mathcal{I}$ of $\mathcal{O}_{\mathfrak{L}}$. As $\mathfrak{L}$ satisfies condition (I), by uniqueness of the algebra $\mathcal{O}_{\mathfrak{L}}$, $\mathcal{I}$ must contain a projection E_i^l for some l, i. Hence $\mathcal{I} \cap \mathcal{A}_{\mathfrak{L}}$ is a nonzero ideal of $\mathcal{A}_{\mathfrak{L}}$ that is invariant under $\lambda_{\mathfrak{L}}$. This leads to $\mathcal{I} = \mathcal{A}_{\mathfrak{L}}$ so that $\mathcal{O}_{\mathfrak{L}}$ is simple. $\square$

The above theorem is a generalization of [CK; 2.14.Theorem] and [Ma;Theorem 6.3]. We next see that $\mathcal{O}_{\mathfrak{L}}$ is purely infinite (and simple) if $\mathfrak{L}$ is aperiodic. Assume that the subshift presented by λ-graph system $\mathfrak{L}$ is not a single point. Note that if $\mathfrak{L}$ is aperiodic, it satisfies condition (I). By [Bra;Corollary 3.5], the following lemma is straightforward.

Lemma 4.8. *A λ-graph system $\mathfrak{L}$ is aperiodic if and only if the AF-algebra $\mathcal{F}_{\mathfrak{L}}$ is simple.*

As in the proof of [C3;1.6 Proposition], we conclude

Proposition 4.9 (cf.[C;1.13 Theorem]). *If a λ-graph system $\mathfrak{L}$ is aperiodic, $\mathcal{O}_{\mathfrak{L}}$ is simple and purely infinite.*

5. K-Theory

The K-groups for the C^*-algebras associated with subshifts have been computed in [Ma2] by using an analogous idea to the Cuntz's paper [C3]. The discussion given in [Ma2] well works for our algebras $\mathcal{O}_{\mathfrak{L}}$ associated with λ-graph systems. Let (A, I) be the nonnegative matrix system of the symbolic matrix system for $\mathfrak{L}$. We first study the K_0-group for the AF-algebra $\mathcal{F}_{\mathfrak{L}}$. We denote by $\Lambda^k(\mathfrak{v}_i^l)$ the set of words of length k that terminate at the vertex $\mathfrak{v}_i^l$. Let $\mathcal{F}_k^{l,i}$ be the C^*-subalgebra of $\mathcal{F}_k^l$ generated by the elements $S_\mu E_i^l S_\nu^*, \mu, \nu \in \Lambda^k$. It is isomorphic to the full matrix algebra $M_{n_i^l(k)}(\mathbb{C})$ of size $n_i^l(k)$ where $n_i^l(k)$ denotes the number of the set $\Lambda^k(\mathfrak{v}_i^l)$, so that one sees

$$\mathcal{F}_k^l \cong M_{n_1^l(k)}(\mathbb{C}) \oplus \cdots \oplus M_{n_{m(l)}^l(k)}(\mathbb{C}).$$

The map $\Phi_k^l : [S_\mu E_i^l S_\mu^*] \in K_0(\mathcal{F}_k^l) \to [E_i^l] \in K_0(\mathcal{A}_l)$ for $i = 1, 2, \dots, m(l), \mu \in \Lambda^k(\mathfrak{v}_i^l)$ yields an isomorphism between $K_0(\mathcal{F}_k^l)$ and $K_0(\mathcal{A}_l) = \mathbb{Z}^{m(l)} = \sum_{i=1}^{m(l)} \mathbb{Z}[E_i^l]$. The isomorphisms $\Phi_k^l, l \in \mathbb{N}$ induce an isomorphism $\Phi_k = \varinjlim_l \Phi_k^l$ from $K_0(\mathcal{F}_k^\infty) = \varinjlim_{\iota_{l,l+1*}} K_0(\mathcal{F}_k^l)$ onto $K_0(\mathcal{A}_{\mathfrak{L}}) = \varinjlim_{\iota_{l,l+1*}} K_0(\mathcal{A}_l)$ in a natural way.

The latter group is denoted by $\mathbb{Z}_{I^t}$, that is isomorphic to the abelian group $\varinjlim_l \{\mathbb{Z}^{m(l)}, I^t_{l,l+1}\}$ of the inductive limit of the homomorphisms $I^t_{l,l+1} : \mathbb{Z}^{m(l)} \longrightarrow \mathbb{Z}^{m(l+1)}, l \in \mathbb{N}$. The embedding $\lambda_{k,k+1}$ of $\mathcal{F}_k^\infty$ into $\mathcal{F}_{k+1}^\infty$ given in Proposition 3.4 (ii) induces a homomorphism $\lambda_{k,k+1_*}$ from $K_0(\mathcal{F}_k^\infty)$ to $K_0(\mathcal{F}_{k+1}^\infty)$ that satisfies

$$\lambda_{k,k+1_*}([S_\mu E_i^l S_\mu^*]) = \sum_{\alpha\in\Sigma}[S_{\mu\alpha} S_\alpha^* E_i^l S_\alpha S_{\mu\alpha}^*], \quad \mu \in \Lambda^k(\mathfrak{v}_i^l), \quad i = 1, 2, \dots, m(l).$$

Define a homomorphism λ_l from $K_0(\mathcal{A}_l)$ to $K_0(\mathcal{A}_{l+1})$ by

$$\lambda_l([E_i^l]) = \sum_{j=1}^{m(l+1)} A_{l,l+1}(i,j)[E_j^{l+1}].$$

As $A_{l,l+1}(i,j) = \sum_{\alpha\in\Sigma} A_{l,l+1}(i,\alpha,j)$, we see $\lambda_l([E_i^l]) = \sum_{\alpha\in\Sigma}[S_\alpha^* E_i^l S_\alpha]$ by (3.4). The homomorphisms $\lambda_l : K_0(\mathcal{A}_l) \to K_0(\mathcal{A}_{l+1}), l \in \mathbb{N}$ act as the transposes $A^t_{l,l+1}$ of the matrices $A_{l,l+1} = [A_{l,l+1}(i,j)]_{i,j}$, that are compatible with the embeddings $\iota_{l,l+1_*} (= I^t_{l,l+1}) : K_0(\mathcal{A}_l) \to K_0(\mathcal{A}_{l+1})$ by (1.2). They define an endomorphism on $\mathbb{Z}_{I^t} (\cong K_0(\mathcal{A}_\mathfrak{L})$. We denote it by $\lambda_{(A,I)}$. Since the diagram

$$\begin{array}{ccc}
K_0(\mathcal{F}_k^\infty) & \xrightarrow{\lambda_{k,k+1_*}} & K_0(\mathcal{F}_{k+1}^\infty) \\
\Phi_k \big\downarrow & & \big\downarrow \Phi_{k+1} \\
K_0(\mathcal{A}_\mathfrak{L}) & \xrightarrow[\lambda_{(A,I)}]{} & K_0(\mathcal{A}_\mathfrak{L})
\end{array}$$

is commutative, one obtains

Proposition 5.1. $K_0(\mathcal{F}_\mathfrak{L}) = \varinjlim\{\mathbb{Z}_{I^t}, \lambda_{(A,I)}\}$.

The group $\varinjlim\{\mathbb{Z}_{I^t}, \lambda_{(A,I)}\}$ is the dimension group $\Delta_{(A,I)}$ for the nonnegative matrix system (A, I) defined in [Ma5]. The dimension group for a nonnegative square finite matrix has been introduced by W. Krieger in [Kr] and [Kr2]. It is realized as the K_0-group for the canonical AF-algebra inside the Cuntz-Krieger algebra associated with the matrix ([C2],[C3]). If a λ-graph system $\mathfrak{L}$ is arising from the finite directed graph associated with the matrix, the C^*-algebras $\mathcal{O}_\mathfrak{L}$ and $\mathcal{F}_\mathfrak{L}$ coincide with the Cuntz-Krieger algebra and the canonical AF-algebra respectively (cf. Section 7). Hence in this case, $K_0(\mathcal{F}_\mathfrak{L})$ coincides with the Krieger's dimension group for the matrix.

Let $p_0 : \mathbb{T} \to \mathcal{O}_\mathfrak{L}$ be the constant function whose value everywhere is the unit 1 of $\mathcal{O}_\mathfrak{L}$. It belongs to the algebra $L^1(\mathbb{T}, \mathcal{O}_\mathfrak{L})$ and hence to the crossed product $\mathcal{O}_\Lambda \rtimes_{\alpha_\mathfrak{L}} \mathbb{T}$. By [Ro], the fixed point subalgebra ${\mathcal{O}_\mathfrak{L}}^{\alpha_\mathfrak{L}}$ is isomorphic to the algebra $p_0(\mathcal{O}_\mathfrak{L} \rtimes_{\alpha_\mathfrak{L}} \mathbb{T})p_0$ through the correspondence : $x \in {\mathcal{O}_\mathfrak{L}}^{\alpha_\mathfrak{L}} \to \hat{x} \in L^1(\mathbb{T}, \mathcal{O}_\mathfrak{L}) \subset \mathcal{O}_\mathfrak{L} \rtimes_{\alpha_\mathfrak{L}} \mathbb{T}$ where the function $\hat{x}$ is defined by $\hat{x}(t) = x, t \in \mathbb{T}$. Then as in [Ma2;Section 4], the projection p_0 is full in $\mathcal{O}_\mathfrak{L} \rtimes_{\alpha_\mathfrak{L}} \mathbb{T}$. Since the AF-algebra $\mathcal{F}_\mathfrak{L}$ is realized as ${\mathcal{O}_\mathfrak{L}}^{\alpha_\mathfrak{L}}$, one sees, by [Bro;Corollary 2.6]

Lemma 5.2. *$\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T}$ is stably isomorphic to $\mathcal{F}_{\mathfrak{L}}$.*

The natural inclusion $\iota : p_0(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T})p_0 \to \mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T}$ induces an isomorphism $\iota_* : K_0(p_0(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T})p_0) \to K_0(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T})$ on K-theory (cf.[Ri;Proposition 2.4]). Denote by $\widehat{\alpha_{\mathfrak{L}}}$ the dual action of $\alpha_{\mathfrak{L}}$ on $\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T}$. Under the identification between $\mathcal{F}_{\mathfrak{L}}$ and $p_0(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T})p_0$, we define an automorphism β on $K_0(\mathcal{F}_{\mathfrak{L}})$ by $\beta = {\iota_*}^{-1} \circ \widehat{\alpha_{\mathfrak{L}}}_* \circ \iota_*$. By a similar argument to [Ma2;Lemma 4.5], [Ma2; Lemma 4.6] and [Ma2;Corollary 4.7], the automorphism $\beta^{-1} : K_0(\mathcal{F}_{\mathfrak{L}}) \to K_0(\mathcal{F}_{\mathfrak{L}})$ corresponds to the shift σ on $\varinjlim\{\mathbb{Z}_{I^t}, \lambda_{(A,I)}\}$. That is, if $x = (x_1, x_2, \dots)$ is a sequence representing an element of $\varinjlim\{\mathbb{Z}_{I^t}, \lambda_{(A,I)}\}$, then $\beta^{-1}x$ is represented by $\sigma(x) = (x_2, x_3, \dots)$. As the dimension automorphism $\delta_{(A,I)}$ of $\Delta_{(A,I)}$ is defined to be the shift of the inductive limit $\varinjlim\{\mathbb{Z}_{I^t}, \lambda_{(A,I)}\}$ ([Ma5]), we obtain

Proposition 5.3. $(K_0(\mathcal{F}_{\mathfrak{L}}), K_0(\mathcal{F}_{\mathfrak{L}})_+, \widehat{\alpha_{\mathfrak{L}}}_*) \cong (\Delta_{(A,I)}, \Delta^+_{(A,I)}, \delta_{(A,I)})$.

We will next present the K-theory formulae for $\mathcal{O}_{\mathfrak{L}}$. As $K_1(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T}) = 0$, the Pimsner-Voiculescu's six term exact sequence of the K-theory for the crossed product $(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T}) \rtimes_{\widehat{\alpha_{\mathfrak{L}}}} \mathbb{Z}$ [PV] says the following lemma:

Lemma 5.4.

(i) $K_0(\mathcal{O}_{\mathfrak{L}}) \cong K_0(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T})/(\mathrm{id} - \widehat{\alpha_{\mathfrak{L}}}_*^{-1})K_0(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T})$.
(ii) $K_1(\mathcal{O}_{\mathfrak{L}}) \cong \mathrm{Ker}(\mathrm{id} - \widehat{\alpha_{\mathfrak{L}}}_*^{-1})$ *on* $K_0(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T})$.

Therefore we have the K-theory formulae for $\mathcal{O}_{\mathfrak{L}}$ by a similar argument to [C3;3.1.Proposition].

Theorem 5.5.

(i)

$$\begin{aligned} K_0(\mathcal{O}_{\mathfrak{L}}) &\cong \mathbb{Z}_{I^t}/(\mathrm{id} - \lambda_{(A,I)})\mathbb{Z}_{I^t} \\ &\cong \varinjlim\{\mathbb{Z}^{m(l+1)}/(I^t_{l,l+1} - A^t_{l,l+1})\mathbb{Z}^{m(l)}; \bar{I}^t_{l,l+1}\}, \end{aligned}$$

(ii)

$$\begin{aligned} K_1(\mathcal{O}_{\mathfrak{L}}) &\cong \mathrm{Ker}(\mathrm{id} - \lambda_{(A,I)}) \ \text{in } \mathbb{Z}_{I^t} \\ &\cong \varinjlim\{\mathrm{Ker}(I^t_{l,l+1} - A^t_{l,l+1}) \ \text{in } \mathbb{Z}^{m(l)}; I^t_{l,l+1}\} \end{aligned}$$

where $\bar{I}^t_{l,l+1}$ is the homomorphism from $\mathbb{Z}^{m(l)}/(I^t_{l-1,l} - A^t_{l-1,l})\mathbb{Z}^{m(l-1)}$ to $\mathbb{Z}^{m(l+1)}/(I^t_{l,l+1} - A^t_{l,l+1})\mathbb{Z}^{m(l)}$ induced by $I^t_{l,l+1}$. More precisely, for the minimal projections $E^l_1, \dots, E^l_{m(l)}$ of $\mathcal{A}_l$ with $\sum_{i=1}^{m(l)} E^l_i = 1$ and the canonical basis $e^l_1, \dots, e^l_{m(l)}$ of $\mathbb{Z}^{m(l)}$, the map $[E^l_i] \to e^l_i$ extends to an isomorphism from $K_0(\mathcal{O}_{\mathfrak{L}})$ onto $\varinjlim\{\mathbb{Z}^{m(l+1)}/(I^t_{l,l+1} - A^t_{l,l+1})\mathbb{Z}^{m(l)}; \bar{I}^t_{l,l+1}\}$. Hence we have

$$K_i(\mathcal{O}_{\mathfrak{L}}) \cong K_i(A, I) \qquad i = 0, 1.$$

Since the double crossed product $(\mathcal{O}_{\mathfrak{L}} \rtimes_{\alpha_{\mathfrak{L}}} \mathbb{T}) \times_{\widehat{\alpha_{\mathfrak{L}}}} \mathbb{Z}$ is stably isomorphic to $\mathcal{O}_{\mathfrak{L}}$, the following proposition is immediate from Lemma 5.2 (cf.[RS],[Bl;p.287]).

Proposition 5.6. *The C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ is nuclear and satisfies the Universal Coefficient Theorem in the sense of Rosenberg and Schochet [RS] (also [Bro2]).*

Hence, for an aperiodic λ-graph system $\mathfrak{L}$, $\mathcal{O}_{\mathfrak{L}}$ is a unital, separable, nuclear, purely infinite, simple C^*-algebra satisfying the UCT so that it lives in a classifiable class of nuclear C^*-algebras by Kirchberg [Kir] and Phillips [Ph]. As the K-groups $K_0(\mathcal{O}_{\mathfrak{L}}), K_1(\mathcal{O}_{\mathfrak{L}})$ are countable abelian groups with $K_1(\mathcal{O}_{\mathfrak{L}})$ torsion free by Theorem 5.5, Rørdam's result [Rø;Proposition 6.7] says the following:

Corollary 5.7. *For an aperiodic λ-graph system $\mathfrak{L}$, the C^*-algebra $\mathcal{O}_{\mathfrak{L}}$ is isomorphic to the C^*-algebra of an inductive limit of a sequence $B_1 \to B_2 \to B_3 \to \cdots$ of simple Cuntz-Krieger algebras.*

Set the Ext-groups

$$\mathrm{Ext}^1(\mathcal{O}_{\mathfrak{L}}) = \mathrm{Ext}(\mathcal{O}_{\mathfrak{L}}), \qquad \mathrm{Ext}^0(\mathcal{O}_{\mathfrak{L}}) = \mathrm{Ext}(\mathcal{O}_{\mathfrak{L}} \otimes C_0(\mathbb{R})).$$

As the UCT holds for our algebras as in the lemma below, it is now easy to compute the Ext-groups by using Theorem 5.5.

Lemma 5.8([RS],[Bro2]). *There exist short exact sequences*

$$0 \longrightarrow \mathrm{Ext}^1_{\mathbb{Z}}(K_0(\mathcal{O}_{\mathfrak{L}}),\mathbb{Z}) \longrightarrow \mathrm{Ext}^1(\mathcal{O}_{\mathfrak{L}}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(K_1(\mathcal{O}_{\mathfrak{L}}),\mathbb{Z}) \longrightarrow 0,$$
$$0 \longrightarrow \mathrm{Ext}^1_{\mathbb{Z}}(K_1(\mathcal{O}_{\mathfrak{L}}),\mathbb{Z}) \longrightarrow \mathrm{Ext}^0(\mathcal{O}_{\mathfrak{L}}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(K_0(\mathcal{O}_{\mathfrak{L}}),\mathbb{Z}) \longrightarrow 0$$

that split unnaturally.

We denote by $\mathbb{Z}_I$ the abelian group defined by the projective limit $\varprojlim_l \{I_{l,l+1} : \mathbb{Z}^{m(l+1)} \to \mathbb{Z}^{m(l)}\}$. The sequence $A_{l,l+1}, l \in \mathbb{Z}_+$ naturally acts on $\mathbb{Z}_I$ as an endomorphism that we denote by A. The identity on $\mathbb{Z}_I$ is denoted by I. Then the cokernel and the kernel of the endomorphism $I - A$ on $\mathbb{Z}_I$ are the Bowen-Franks groups $BF^0(A,I)$ and $BF^1(A,I)$ for (A,I) respectively ([Ma5]). By [Ma5;Theorem 9.6], there exists a short exact sequence

$$0 \longrightarrow \mathrm{Ext}^1_{\mathbb{Z}}(K_0(A,I),\mathbb{Z}) \longrightarrow BF^0(A,I) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(K_1(A,I),\mathbb{Z}) \longrightarrow 0$$

that splits unnaturally. And also

$$BF^1(A,I) \cong \mathrm{Hom}_{\mathbb{Z}}(K_0(A,I),\mathbb{Z}).$$

As in the proof of [Ma5;Lemma 9.7], we see that $\mathrm{Ext}^1_{\mathbb{Z}}(\mathbb{Z}_{I^t},\mathbb{Z}) = 0$ so that $\mathrm{Ext}^1_{\mathbb{Z}}(\mathrm{Ker}(\mathrm{id} - \lambda_{(A,I)})$ in $\mathbb{Z}_{I^t},\mathbb{Z}) = 0$. This means that $\mathrm{Ext}^1_{\mathbb{Z}}(K_1(A,I)),\mathbb{Z}) = 0$. Theorem 5.5 says that $K_i(A,I) \cong K_i(\mathcal{O}_{\mathfrak{L}})$ so that we conclude by Lemma 5.8,

Theorem 5.9.

(i) $\mathrm{Ext}(\mathcal{O}_{\mathfrak{L}}) = \mathrm{Ext}^1(\mathcal{O}_{\mathfrak{L}}) \cong BF^0(A,I) = \mathbb{Z}_I/(I-A)\mathbb{Z}_I$,
(ii) $\mathrm{Ext}^0(\mathcal{O}_{\mathfrak{L}}) \cong BF^1(A,I) = \mathrm{Ker}(I-A)$ *in* $\mathbb{Z}_I$.

Theorem 5.9 is a generalization of [CK;5.3 Theorem] and [Ma6].

6. Realizations as endomorphism crossed products and Hilbert C^*-bimodule algebras

Following Deaconu's discussions in [De2],[De3],[De4], we will realize the algebra $\mathcal{O}_{\mathfrak{L}}$ as an endomorphism crossed product $\mathcal{F}_{\mathfrak{L}} \times_{\beta_{\mathfrak{L}}} \mathbb{N}$. Recall that the algebra $\mathcal{F}_{\mathfrak{L}}$ is isomorphic to the C^*-algebra $C^*(F_{\mathfrak{L}})$ of the groupoid $F_{\mathfrak{L}}$. The groupoid $F_{\mathfrak{L}}$ is written as

$$\{(x,y) \in X_{\mathfrak{L}} \times X_{\mathfrak{L}} \mid \sigma^k(x) = \sigma^k(y) \text{ for some } k \in \mathbb{Z}_+\}.$$

Put

$$\beta_{\mathfrak{L}}(f)(x,y) = \frac{1}{\sqrt{p(\sigma(x))p(\sigma(y))}} f(\sigma(x),\sigma(y)), \quad f \in C_c(F_{\mathfrak{L}}), x, y \in X_{\mathfrak{L}}$$

where $p(x)$ is the number of the paths z such that $\sigma(z) = x$, and for $(x,n,y) \in G_{\mathfrak{L}}$

$$v(x,n,y) = \begin{cases} \frac{1}{\sqrt{p(\sigma(x))}}, & \text{if } n = 1 \text{ and } y = \sigma(x), \\ 0 & \text{otherwise.} \end{cases}$$

Regarding $C^*(F_{\mathfrak{L}})$ as a subalgebra of $C^*(G_{\mathfrak{L}})$, one sees that v is a nonunitary isometry satisfying $\beta_{\mathfrak{L}}(f) = vfv^*$ ([De2],[De3],[De4]). Then $\beta_{\mathfrak{L}}$ is a proper corner endomorphism of $C^*(F_{\mathfrak{L}})$ such that $C^*(G_{\mathfrak{L}})$ is isomorphic to the crossed product $C^*(F_{\mathfrak{L}}) \times_{\beta_{\mathfrak{L}}} \mathbb{N}$ (cf.[Rø2]). We will write the isometry v in terms of the generators S_α, E_i^l. For $l \in \mathbb{N}, i = 1, \dots, m(l)$, we denote by n_i^l the number of the edges e in $E_{l-1,l}$ such that $t(e) = \mathfrak{v}_i^l$. As $\mathfrak{L}$ is left-resolving, it is the number of the symbols $\alpha \in \Sigma$ such that $S_\alpha^* S_\alpha E_i^l \neq 0$. It follows that $n_i^l E_i^l = \sum_{\alpha \in \Sigma} S_\alpha^* S_\alpha E_i^l$. Note that if $I_{l,l+1}(i,j) = 1$, then $n_i^l = n_j^{l+1}$. Then one obtains

$$v = \sum_{i=1}^{m(l)} \frac{1}{\sqrt{n_i^l}} \sum_{\alpha \in \Sigma} S_\alpha E_i^l \tag{6.1}$$

where the right-hand side does not depend on the choice of $l \in \mathbb{N}$. We can immediately see that $\mathcal{O}_{\mathfrak{L}}$ is generated by the C^*-algebra $\mathcal{F}_{\mathfrak{L}}$ and the above isometry v, that satisfies

$$v^*v = 1, \quad v\mathcal{F}_{\mathfrak{L}}v^* \subset \mathcal{F}_{\mathfrak{L}}, \quad v^*\mathcal{F}_{\mathfrak{L}}v \subset \mathcal{F}_{\mathfrak{L}}. \tag{6.2}$$

The universality (Theorem 3.6) of the algebra $\mathcal{O}_{\mathfrak{L}}$ corresponds to the universality of the crossed product $C^*(F_{\mathfrak{L}}) \times_{\beta_{\mathfrak{L}}} \mathbb{N}$. It needs however a slightly complicated argument to directly determine the operator relations ($\mathfrak{L}$) by using (6.1) and (6.2), as it is possible. There are some merits to realize $\mathcal{O}_{\mathfrak{L}}$ as $\mathcal{F}_{\mathfrak{L}} \times_{\beta_{\mathfrak{L}}} \mathbb{N}$. One is the fact that its purely infiniteness is immediately deduced from Rørdam's result [Rø2;Theorem 3.1] under the condition that $\mathcal{F}_{\mathfrak{L}}$ is simple. The other one is K-theory formulae. Rørdam also in [Rø2; Corollary 2.2] showed that

(i) $K_0(\mathcal{F}_{\mathfrak{L}} \times_{\beta_{\mathfrak{L}}} \mathbb{N}) \cong K_0(\mathcal{F}_{\mathfrak{L}})/(\mathrm{id} - \beta_{\mathfrak{L}*})K_0(\mathcal{F}_{\mathfrak{L}})$.
(ii) $K_1(\mathcal{F}_{\mathfrak{L}} \times_{\beta_{\mathfrak{L}}} \mathbb{N}) \cong \mathrm{Ker}(\mathrm{id} - \beta_{\mathfrak{L}*})$ on $K_0(\mathcal{F}_{\mathfrak{L}})$

(cf.Paschke [Pa], Deaconu [De3]). These are precisely the formulae of Lemma 5.4 with Lemma 5.2.

In [De3; Section 3], Deaconu showed that the groupoid C^*-algebras of continuous graphs are realized as C^*-algebras constructed from Hilbert C^*-bimodules defined in [Pi] (see also [Kat]). A special case of continuous graphs was studied in Kajiwara-Watatani [KW]. We identify the algebra $C(\Omega_{\mathfrak{L}})$ of all continuous functions on $\Omega_{\mathfrak{L}}$ with the commutative C^*-algebra $\mathcal{A}_{\mathfrak{L}}$. Let $X_{\mathcal{A}_{\mathfrak{L}}}$ be the set $C(E_{\mathfrak{L}})$ of all continuous functions on $E_{\mathfrak{L}}$, that is identified with $\sum_{\alpha\in\Sigma}^{\oplus}\mathbb{C}S_\alpha\mathcal{A}_{\mathfrak{L}}$, because $\mathfrak{L}$ is left-resolving. We endow $X_{\mathcal{A}_{\mathfrak{L}}}$ with a Hilbert C^*-bimodule structure over $\mathcal{A}_{\mathfrak{L}}$ defined by

$$(S_\alpha a)\cdot b = S_\alpha \cdot ab, \qquad < S_\alpha a, S_\beta b >_{\mathcal{A}_{\mathfrak{L}}} = a^* S_\alpha^* S_\beta b,$$
$$\phi_{\mathfrak{L}}(b) S_\alpha a = b S_\alpha a = S_\alpha \cdot S_\alpha^* b S_\alpha a$$

for $a, b \in \mathcal{A}_{\mathfrak{L}}, \alpha, \beta \in \Sigma$. A special case of this construction of the Hilbert C^*-bimodules is seen in the proof of [PWY;Theorem 4.2] for the C^*-algebras associated with subshifts. The above mentioned Deaconu's result says the following proposition:

PROPOSITION 6.1. *The C^*-algebra constructed from the Hilbert C^*-bimodule $(\phi_{\mathfrak{L}}, X_{\mathcal{A}_{\mathfrak{L}}})$ over $\mathcal{A}_{\mathfrak{L}}$ is isomorphic to the C^*-algebra $\mathcal{O}_{\mathfrak{L}}$.*

7. EXAMPLES

In this section, we give two kinds of examples of λ-graph systems and study their associated C^*-algebras. The first ones appear as presentations of sofic shifts. The second one is defined by a Shannon graph with countable infinite vertices.

Presentations of sofic shifts come from labeled graphs with finite vertices that are called λ-graphs (cf. [Fi],[Kr4],[Kr5],[LM],[We], ...). Let $G = (V, E)$ be a finite directed graph with finite vertex set V and finite edge set E. Let $\mathcal{G} = (G, \lambda)$ be a labeled graph over Σ defined by G and a labeling map $\lambda : E \to \Sigma$. Suppose that it is left-resolving and predecessor-separated. Let A_G be the adjacency matrix of G, that is defined by

$$A_G(e, f) = \begin{cases} 1 & \text{if } t(e) = s(e), \\ 0 & \text{otherwise} \end{cases}$$

for $e, f \in E$. The matrix A_G defines a shift of finite type by regarding its edges as its alphabet. Since the matrix A_G is of entries in $\{0, 1\}$, we have the Cuntz-Krieger algebra $\mathcal{O}_{A_G}$ defined by A_G ([CK] cf.[KPPR],[Rø]). By putting $V_l^{\mathcal{G}} = V$, $E_{l,l+1}^{\mathcal{G}} = E$ for $l \in \mathbb{Z}_+$, and $\lambda^{\mathcal{G}} = \lambda, \iota^{\mathcal{G}} = \mathrm{id}$, we have a λ-graph system $\mathfrak{L}_{\mathcal{G}} = (V^{\mathcal{G}}, E^{\mathcal{G}}, \lambda^{\mathcal{G}}, \iota^{\mathcal{G}})$. Then we have

Proposition 7.1. *The C^*-algebra $\mathcal{O}_{\mathfrak{L}_{\mathcal{G}}}$ is isomorphic to the Cuntz-Krieger algebra $\mathcal{O}_{A_G}$.*

Proof. Let $V = \{v_1, \dots, v_m\}$ be the vertex set of G. Let $S_\alpha, \alpha \in \Sigma$ be the canonical generating partial isometries of $\mathcal{O}_{\mathfrak{L}_{\mathcal{G}}}$. We denote by $E_1, E_2, \dots, E_m$ the set of all minimal projections of $\mathcal{A}_l = \mathcal{A}_{\mathfrak{L}_{\mathcal{G}}}, l \in \mathbb{N}$ corresponding to the vertices $v_1, \dots, v_m$. As the labeled graph $\mathcal{G}$ is predecessor-separated, they are written in terms of $S_\alpha, \alpha \in \Sigma$ as in (1.7). Note that in the algebra $\mathcal{O}_{\mathfrak{L}_{\mathcal{G}}}$, $S_\alpha E_i \neq 0$ if and only if there exists an edge $e \in E$ satisfying $\lambda(e) = \alpha$ and $t(e) = v_i$. As $\mathcal{G}$ is left-resolving, the correspondence

$$e \in E \longleftrightarrow (\lambda(e), t(e)) \in \{(\alpha, v_i) \in \Sigma \times V \mid S_\alpha E_i \neq 0\}$$

is bijective. For $e \in E$, put $s_e = S_{\lambda(e)} E_{t(e)} \in \mathcal{O}_{\mathfrak{L}_{\mathcal{G}}}$, where E_{v_i} denotes E_i. As $E_{t(e)} = s_e^* s_e$ for $e \in E$ and $S_\alpha = \sum_{\substack{e \in E, \\ \lambda(e)=\alpha}} s_e$ for $\alpha \in \Sigma$, the algebra $\mathcal{O}_{\mathfrak{L}_{\mathcal{G}}}$ is generated by the partial isometries $s_e, e \in E$. It is immediate to see that the following relations hold:

$$\sum_{e \in E} s_e s_e^* = 1, \qquad s_e^* s_e = \sum_{f \in E} A_G(e, f) s_f s_f^*.$$

This means that the C^*-algebra generated by $s_e, e \in E$ is the Cuntz-Krieger algebra $\mathcal{O}_{A_G}$ defined by the matrix A_G. $\square$

If, in particular, a labeled graph $\mathcal{G} = (G, \lambda)$ has different labels for different edges, it defines a shift of finite type. In this case, one may identify the edge set E with the alphabet Σ. Let $\mathfrak{L}_G$ be the λ-graph system $\mathfrak{L}_{\mathcal{G}}$ as in the above one. Let $S_\alpha, \alpha \in \Sigma$ be the generating partial isometries of $\mathcal{O}_{\mathfrak{L}_G}$. It is obvious that the relations (1.4),(1.5) and (1.6) give rise to the following relations:

$$S_\alpha^* S_\alpha = \sum_{\beta \in E} A_G(\alpha, \beta) S_\beta S_\beta^*, \qquad \alpha \in \Sigma.$$

Remark. While completing this paper, Toke M. Carlsen let the author know his preprint [Ca], where he shows that the C^*-algebra associated with sofic shifts are isomorphic to the Cuntz-Krieger algebras of their left Krieger cover graphs. His result is a special case of the above proposition.

We will next present a λ-graph system for which the associated C^*-algebra is not stably isomorphic to any Cuntz-Krieger algebra and any Cuntz algebra. There is a method introduced in [KM] to construct λ-graph systems from Shannon graphs. By a Shannon graph we mean here a left-resolving labeled directed graph with countable vertices and finite labels.

Let us consider a Shannon graph defined as follows: Let $V = \{v_1, v_2, \dots, \}$ be its countable infinite vertex set. Its alphabet Σ consist of the five symbols $\{\alpha, \beta, \gamma, \delta, \epsilon\}$. The edges labeled α are from v_{n+1} to v_n for $n = 1, 2, \dots$. The

edges labeled β are from v_1 to v_2 and from v_{2n} to v_{2n+2} for $n = 1,2,\dots$. The edges labeled γ are self-loops at v_n for $n = 2,3,\dots$. The edge labeled δ is a self-loop at v_1. The edges labeled ϵ are from v_1 to v_n for $n = 1,2,\dots$. The resulting labeled graph is left-resolving and hence it is a Shannon graph. We denote it by $\mathcal{S}$. We will construct a λ-graph system $\mathfrak{L}(\mathcal{S})$ from the Shannon graph $\mathcal{S}$ by a method introduced in [KM] as in the following way. For a vertex $v \in V$ and for $l \in \mathbb{N}$, let $\Gamma_l^-(v)$ be the set of all label sequences of length l terminating at v. Define an equivalence relation $v \approx_{(l)} v'$ for vertices $v, v' \in V$ by $\Gamma_l^-(v) = \Gamma_l^-(v')$. For $l = 0$, define $v \approx_{(l)} v'$ for all $v, v' \in V$. The vertex set V_l is then defined by the set of $\approx_{(l)}$-equivalence classes of V. We denote by $V_l = \{\mathfrak{v}_1^l, \dots, \mathfrak{v}_{m(l)}^l\}$. The vertices $\mathfrak{v}_i^l, i = 1,\dots,m(l)$ of V_l may be identified with $\{\Gamma_l^-(v) : v \in V\}$. We define a map $\iota_{l,l+1} : V_{l+1} \to V_l$ by $\iota_{l,l+1}(\mathfrak{v}_j^{l+1}) = \mathfrak{v}_i^l$ if $\mathfrak{v}_j^{l+1} \subset \mathfrak{v}_i^l$. We define an edge labeled $\omega \in \Sigma$ from $\mathfrak{v}_i^l$ to $\mathfrak{v}_j^{l+1}$ if there exists an edge labeled ω in $\mathcal{S}$ from a vertex in $\mathfrak{v}_i^l$ to a vertex in $\mathfrak{v}_j^{l+1}$. Then the resulting labeled graph with vertex sets V_l, labeled edges from V_l to V_{l+1} and surjective maps $\iota_{l,l+1} : V_{l+1} \to V_l$ for $l \in \mathbb{Z}_+$ defines a λ-graph system over Σ ([KM]). We denote it by $\mathfrak{L}(\mathcal{S})$.
The vertex sets $V_l, l \in \mathbb{Z}_+$ are written as in the following way:

$$
\begin{aligned}
&V_0 : \mathfrak{v}_1^0 = \{v_n \mid n = 1,2,\dots\}.\\
&V_1 : \mathfrak{v}_1^1 = \{v_1\}, \mathfrak{v}_2^1 = \{v_{2n} \mid n = 1,2,\dots\}, \mathfrak{v}_3^1 = \{v_{2n+1} \mid n = 1,2,\dots\}.\\
&V_2 : \mathfrak{v}_1^2 = \{v_1\}, \mathfrak{v}_2^2 = \{v_2\}, \mathfrak{v}_3^2 = \{v_{2n} \mid n = 2,3,\dots\},\\
&\quad \mathfrak{v}_4^2 = \{v_{2n+1} \mid n = 1,2,\dots\}.\\
&V_3 : \mathfrak{v}_1^3 = \{v_1\}, \mathfrak{v}_2^3 = \{v_2\}, \mathfrak{v}_3^3 = \{v_4\}, \mathfrak{v}_4^3 = \{v_{2n} \mid n = 3,4,\dots\},\\
&\quad \mathfrak{v}_5^3 = \{v_{2n+1} \mid n = 1,2,\dots\}.\\
&V_4 : \mathfrak{v}_1^4 = \{v_1\}, \mathfrak{v}_2^4 = \{v_2\}, \mathfrak{v}_3^4 = \{v_4\}, \mathfrak{v}_4^4 = \{v_6\}, \mathfrak{v}_5^4 = \{v_{2n} \mid n = 4,5,\dots\},\\
&\quad \mathfrak{v}_6^4 = \{v_3\}, \mathfrak{v}_7^4 = \{v_{2n+1} \mid n = 2,3,\dots\}.\\
&V_5 : \mathfrak{v}_1^5 = \{v_1\}, \mathfrak{v}_2^5 = \{v_2\}, \mathfrak{v}_3^5 = \{v_4\}, \mathfrak{v}_4^5 = \{v_6\}, \mathfrak{v}_5^5 = \{v_8\},\\
&\quad \mathfrak{v}_6^5 = \{v_{2n} \mid n = 5,6,\dots\}, \mathfrak{v}_7^5 = \{v_3\}, \mathfrak{v}_8^5 = \{v_5\},\\
&\quad \mathfrak{v}_9^5 = \{v_{2n+1} \mid n = 3,4,\dots\}.\\
&V_6 : \mathfrak{v}_1^6 = \{v_1\}, \mathfrak{v}_2^6 = \{v_2\}, \mathfrak{v}_3^6 = \{v_4\}, \mathfrak{v}_4^6 = \{v_6\}, \mathfrak{v}_5^6 = \{v_8\}, \mathfrak{v}_6^6 = \{v_{10}\},\\
&\quad \mathfrak{v}_7^6 = \{v_{2n} \mid n = 6,7,\dots\}, \mathfrak{v}_8^6 = \{v_3\}, \mathfrak{v}_9^6 = \{v_5\}, \mathfrak{v}_{10}^6 = \{v_7\},\\
&\quad \mathfrak{v}_{11}^6 = \{v_{2n+1} \mid n = 4,5,\dots\}.\\
&V_7 : \mathfrak{v}_1^7 = \{v_1\}, \mathfrak{v}_2^7 = \{v_2\}, \mathfrak{v}_3^7 = \{v_4\}, \mathfrak{v}_4^7 = \{v_6\}, \mathfrak{v}_5^7 = \{v_8\}, \mathfrak{v}_6^7 = \{v_{10}\},\\
&\quad \mathfrak{v}_7^7 = \{v_{12}\}, \mathfrak{v}_8^7 = \{v_{2n} \mid n = 7,8,\dots\}, \mathfrak{v}_9^7 = \{v_3\}, \mathfrak{v}_{10}^7 = \{v_5\}, \mathfrak{v}_{11}^7 = \{v_7\},\\
&\quad \mathfrak{v}_{12}^7 = \{v_9\}, \mathfrak{v}_{13}^7 = \{v_{2n+1} \mid n = 5,6,\dots\}.\\
&\quad \cdots\cdots\cdots.
\end{aligned}
$$

Lemma 7.2. *The λ-graph system $\mathfrak{L}(\mathcal{S})$ is aperiodic.*

Proof. Take and fix an arbitrary vertex $\mathfrak{v}_i^l \in V_l$, let k be the minimum number such that $v_k \in \mathfrak{v}_i^l$. There exists a path of length k in $\mathfrak{L}(\mathcal{S})$ that starts at $\mathfrak{v}_i^l$ and ends at $\mathfrak{v}_1^{l+k} = \{v_1\} \in V_{l+k}$ whose label is $(\alpha, \alpha, \dots, \alpha)$. There exist edges from $\mathfrak{v}_1^{l+k}$ to $\mathfrak{v}_j^{l+k+1}$ for all $j = 1, 2, \dots, m(l+1)$ whose labels are ϵ. This means that $\mathfrak{L}(\mathcal{S})$ is aperiodic. $\square$

Therefore the associated C^*-algebra $\mathcal{O}_{\mathfrak{L}(\mathcal{S})}$ is simple and purely infinite.

Let us compute its K-groups. We first present the matrices $A^t_{l,l+1}$ and $I^t_{l,l+1}$ for all $l \geq 5$. As in the preceding table, we see that $m(l) = 2l - 1$ for $l \geq 5$. We write the matrices along the ordered basis $\mathfrak{v}_1^l, \mathfrak{v}_2^l, \dots, \mathfrak{v}_{m(l)}^l$, where blanks denote zeros. For $l = 5$, we have

$$A^t_{5,6} = \begin{bmatrix}
2 & 1 & & & & & & & \\
2 & 1 & 1 & & & & & & \\
1 & & 1 & 1 & & & & & \\
1 & 1 & & 1 & 1 & & & & \\
1 & & & & 1 & 1 & & & \\
1 & & & 1 & & 1 & & 1 & \\
1 & & & & & & 1 & 1 & \\
1 & & & & & 1 & 1 & 1 & \\
1 & & & & & & 1 & 1 & 1 \\
1 & & & & & & & 1 & 1 \\
1 & & & & & & & 1 & 2
\end{bmatrix}, \quad I^t_{5,6} = \begin{bmatrix}
1 & & & & & & & & \\
 & 1 & & & & & & & \\
 & & 1 & & & & & & \\
 & & & 1 & & & & & \\
 & & & & 1 & & & & \\
 & & & & & 1 & & & \\
 & & & & & & & 1 & \\
 & & & & & & 1 & & \\
 & & & & & & & & 1 \\
 & & & & & & & 1 & \\
 & & & & & & & & 1
\end{bmatrix},$$

so that

$$A^t_{5,6} - I^t_{5,6} = \begin{bmatrix}
1 & 1 & & & & & & & \\
2 & & 1 & & & & & & \\
1 & & & 1 & & & & & \\
1 & 1 & & & 1 & & & & \\
1 & & & & & 1 & & & \\
1 & & & 1 & & & & 1 & \\
1 & & & & & & 1 & & \\
1 & & & & & 1 & & 1 & \\
1 & & & & & & 1 & 1 & \\
1 & & & & & & & & 1 \\
1 & & & & & & & 1 & 1
\end{bmatrix}.$$

For general $l \geq 5$, we see

$$
A^t_{l,l+1} = \begin{bmatrix}
2 & 1 & & & & & & & & & & & & & \\
2 & 1 & 1 & & & & & & & & & & & & \\
1 & & 1 & 1 & & & & & & & & & & & \\
1 & 1 & & 1 & 1 & & & & & & & & & & \\
1 & & & & 1 & 1 & & & & & & & & & \\
1 & & & 1 & & 1 & 1 & & & & & & & & \\
1 & & & & & & 1 & 1 & & & & & & & \\
1 & & & & & 1 & & 1 & 1 & & & & & & \\
\vdots & & & & & & & & \ddots & \ddots & & & & & \\
\vdots & & & & & & & \ddots & & \ddots & \ddots & & & & \\
1 & & & & & & & & & & 1 & 1 & & & \\
1 & & & & & & & & & 1 & & 1 & & 1 & \\
1 & & & & & & & & & & & & 1 & 1 & \\
1 & & & & & & & & & & & 1 & 1 & 1 & \\
1 & & & & & & & & & & & & 1 & 1 & 1 \\
1 & & & & & & & & & & & & & 1 & 1 \\
1 & & & & & & & & & & & & & 1 & 2
\end{bmatrix},
$$

$$
I^t_{l,l+1} = \begin{bmatrix}
1 & & & & & & & & & & & & & & \\
& 1 & & & & & & & & & & & & & \\
& & 1 & & & & & & & & & & & & \\
& & & 1 & & & & & & & & & & & \\
& & & & 1 & & & & & & & & & & \\
& & & & & 1 & & & & & & & & & \\
& & & & & & 1 & & & & & & & & \\
& & & & & & & 1 & & & & & & & \\
& & & & & & & & \ddots & & & & & & \\
& & & & & & & & & \ddots & & & & & \\
& & & & & & & & & & 1 & & & & \\
& & & & & & & & & & & 1 & & & \\
& & & & & & & & & & & & & 1 & \\
& & & & & & & & & & & & 1 & & \\
& & & & & & & & & & & & & & 1 \\
& & & & & & & & & & & & & 1 & \\
& & & & & & & & & & & & & & 1
\end{bmatrix},
$$

and hence

$$A^t_{l,l+1} - I^t_{l,l+1} = \begin{bmatrix}
1 & 1 & & & & & & & & & & & & & \\
2 & & 1 & & & & & & & & & & & & \\
1 & & & 1 & & & & & & & & & & & \\
1 & 1 & & & 1 & & & & & & & & & & \\
1 & & & & & 1 & & & & & & & & & \\
1 & & & 1 & & & 1 & & & & & & & & \\
1 & & & & & & & 1 & & & & & & & \\
1 & & & & & 1 & & & 1 & & & & & & \\
\vdots & & & & & & & & & \ddots & & & & & \\
\vdots & & & & & & & \ddots & & & \ddots & & & & \\
1 & & & & & & & & & & & 1 & & & \\
1 & & & & & & & & & 1 & & & & 1 & \\
1 & & & & & & & & & & & & 1 & & \\
1 & & & & & & & & & & & 1 & & 1 & \\
1 & & & & & & & & & & & & 1 & 1 & \\
1 & & & & & & & & & & & & & & 1 \\
1 & & & & & & & & & & & & & 1 & 1
\end{bmatrix}.$$

Lemma 7.3. *For $l \geq 4$ and $z = \begin{bmatrix} z_1 \\ \vdots \\ z_{2l+1} \end{bmatrix} \in \mathbb{Z}^{2l+1}$, put*

$$\begin{aligned}
\varphi_l([z_i]_{i=1}^{2l+1}) &= z_{2l-7} - z_{2l-5} - z_{2l-4} + z_{2l-2}, \\
\psi_l([z_i]_{i=1}^{2l+1}) &= z_{2l-7} - z_{2l-4} - z_{2l-3} + z_{2l-1}, \\
\xi_l([z_i]_{i=1}^{2l+1}) &= z_{2l-7} - z_{2l-4} - z_{2l} + z_{2l+1}.
\end{aligned}$$

Then there exists $x = \begin{bmatrix} x_1 \\ \vdots \\ x_{2l-1} \end{bmatrix} \in \mathbb{Z}^{2l-1}$ such that

$$\begin{bmatrix} z_1 \\ \vdots \\ z_{2l-1} \\ z_{2l} \\ z_{2l+1} \end{bmatrix} = (A^t_{l,l+1} - I^t_{l,l+1}) \begin{bmatrix} x_1 \\ \vdots \\ x_{2l-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \varphi_l(z) \\ \psi_l(z) \\ 0 \\ \xi_l(z) \end{bmatrix}.$$

Lemma 7.4. *The map $z = [z_i]_{i=1}^{2l+1} \in \mathbb{Z}^{2l+1} \to (\varphi_l(z), \psi_l(z), \xi_l(z)) \in \mathbb{Z}^3$ induces an isomorphism from $\mathbb{Z}^{2l+1}/(A^t_{l,l+1} - I^t_{l,l+1})\mathbb{Z}^{2l-1}$ onto $\mathbb{Z}^3$.*

Proof. It suffices to show the surjectivity of the induced map

$$z = [z_i]_{i=1}^{2l+1} \in \mathbb{Z}^{2l+1}/(A^t_{l,l+1} - I^t_{l,l+1})\mathbb{Z}^{2l-1} \longrightarrow (\varphi_l(z), \psi_l(z), \xi_l(z)) \in \mathbb{Z}^3.$$

For $(m, n, k) \in \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$, put $z = [z_i]_{i=1}^{2l+1}$ where

$$z_i = \begin{cases} 0 & \text{for } i = 1, 2, \dots, 2l-3, 2l, \\ m & \text{for } i = 2l-2, \\ n & \text{for } i = 2l-1, \\ k & \text{for } i = 2l+1. \end{cases}$$

Then we see that $\varphi_l(z) = m, \psi_l(z) = n, \xi_l(z) = k$. $\square$

We denote by ρ_{l+1} the above isomorphism from $\mathbb{Z}^{2l+1}/(A^t_{l,l+1} - I^t_{l,l+1})\mathbb{Z}^{2l-1}$ onto $\mathbb{Z}^3$. Let L be the matrix $\begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$. Since the following diagram is commutative:

$$\begin{array}{ccc} \mathbb{Z}^{2l-1}/(A^t_{l-1,l} - I^t_{l-1,l})\mathbb{Z}^{2l-3} & \xrightarrow{I^t_{l,l+1}} & \mathbb{Z}^{2l+1}/(A^t_{l,l+1} - I^t_{l,l+1})\mathbb{Z}^{2l-1} \\ {\scriptstyle \rho_l}\downarrow & & {\scriptstyle \rho_{l+1}}\downarrow \\ \mathbb{Z}^3 & \xrightarrow{L} & \mathbb{Z}^3, \end{array}$$

we obtain

PROPOSITION 7.5. $K_0(\mathcal{O}_{\mathfrak{L}(\mathcal{S})}) \cong 0$.

Proof. As $L^3 = 0$, by Theorem 5.5, it follows that

$$K_0(\mathcal{O}_{\mathfrak{L}(\mathcal{S})}) = \varinjlim\{\mathbb{Z}^{2l+1}/(A^t_{l,l+1} - I^t_{l,l+1})\mathbb{Z}^{2l-1}, \bar{I}^t_{l+1,l+2}\} = \varinjlim\{\mathbb{Z}^3, L\} \cong 0.$$

$\square$

Concerning the group $K_1(\mathcal{O}_{\mathfrak{L}(\mathcal{S})})$, one sees

PROPOSITION 7.6. $K_1(\mathcal{O}_{\mathfrak{L}(\mathcal{S})}) \cong \mathbb{Z}$.

Proof. For $l \geq 5$, put $x(l) = [x_i]_{i=1}^{2l-1} \in \mathbb{Z}^{2l-1}$ where

$$\begin{aligned} &x_1 = 1, \quad x_2 = -1, \quad x_3 = -2, \\ &x_i = -1 \quad \text{for} \quad i = 4, 6, 8, \dots, 2l-4, 2l-3, 2l-1, \\ &x_i = 0 \quad \text{for} \quad i = 5, 7, 9, \dots, 2l-5, 2l-2. \end{aligned}$$

It is easy to see that

$$\operatorname{Ker}(A^t_{l,l+1} - I^t_{l,l+1}) = \mathbb{Z}x(l), \qquad I^t_{l,l+1}x(l) = x(l+1).$$

Hence we obtain $K_1(\mathcal{O}_{\mathfrak{L}(\mathcal{S})}) \cong \mathbb{Z}$ by Theorem 5.5. $\square$

Therefore we conclude

Theorem 7.7. *The C^*-algebra $\mathcal{O}_{\mathfrak{L}(\mathcal{S})}$ is unital, simple, purely infinite, nuclear and generated by five partial isometries with mutually orthogonal ranges. Its K-groups are*

$$K_0(\mathcal{O}_{\mathfrak{L}(\mathcal{S})}) \cong 0, \qquad K_1(\mathcal{O}_{\mathfrak{L}(\mathcal{S})}) \cong \mathbb{Z}.$$

As the K_1-group of a Cuntz-Krieger algebra is the torsion-free part of its K_0-group, the algebra $\mathcal{O}_{\mathfrak{L}(\mathcal{S})}$ lives outside the Cuntz-Krieger algebras (cf.[Ma3]).

Remark. M. Tomforde in [T] considered C^*-algebras associated to labeled graphs as a generalization of Cuntz-Krieger algebras (cf.[T2]). He deals with labeled directed graphs with (generally) infinite vertices. If the labeled graphs have finite vertices, the resulting graphs are ones in the first examples of this section. In this case, his C^*-algebras coincide with our C^*-algebras. The referee informed to the author that his algebras in general are not ours of λ-graph systems.

References

[An] C. Anantharaman-Delaroche, *Purely infinite C^*-algebras arising from dynamical systems*, Bull. Soc. Math. France **125(2)** (1997), 199–225.

[Bl] B. Blackadar, *K-Theory for operator algebras*, Springer-Verlag, Berlin, Heidelberg and New York, 1986.

[Bra] O. Bratteli, *Inductive limits of finite-dimensional C^*-algebras*, Trans. Amer. Math. Soc. **171** (1972), 195–234.

[Bro] L. G. Brown, *Stable isomorphism of hereditary subalgebras of C^*-algebras*, Pacific. J. Math. **71** (1977), 335–348.

[Bro2] L. G. Brown, *The universal coefficient theorem for Ext and quasidiagonality*, Operator Algebras and Group Representation, Pitman Press **17** (1983), 60–64.

[Ca] T. M. Carlsen, *On C^*-algebras associated with sofic shifts*, preprint 2000.

[CaM] T. M. Carlsen and K. Matsumoto, *Some remarks on the C^*-algebras associated with subshifts*, preprint 2001.

[C] J. Cuntz, *Simple C^*-algebras generated by isometries*, Commun. Math. Phys. **57** (1977), 173–185.

[C2] J. Cuntz, *K-Theory for certain C^*-algebras*, Ann. Math. **113** (1981), 181–197.

[C3] J. Cuntz, *A class of C^*-algebras and topological Markov chains II: reducible chains and the Ext-functor for C^*-algebras*, Invent. Math. **63** (1980), 25–40.

[CK] J. Cuntz and W. Krieger, *A class of C^*-algebras and topological Markov chains*, Invent. Math. **56** (1980), 251–268.

[De] V. Deaconu, *Groupoids associated with endomorphisms*, Trans. Amer. Math. Soc. **347** (1995), 1779–1786.

[De2] V. Deaconu, *Generalized Cuntz-Krieger algebras*, Proc. Amer. Math. Soc. **124** (1996), 3427–3435.

[De3] V. Deaconu, *Generalized solenoids and C^*-algebras*, Pacific J. Math. **190** (1999), 247–260.

[De4] V. Deaconu, *Continuous graphs and C^*-algebras*, Operator Theoretical Methods (Timiç soara, 1998) Theta Found., Bucharest (2000), 137–149.

[EL] R. Excel and M. Laca, *Cuntz-Krieger algebras for infinite matrices*, J. reine. angew. Math. **512** (1999), 119–172.

[Fi] R. Fischer, *Sofic systems and graphs*, Monats. für Math. **80** (1975), 179–186.

[H] D. Huang, *Flow equivalence of reducible shifts of finite type and Cuntz-Krieger algebras*, J. reine angew. Math. **462** (1995), 185–217.

[KPW] T. Kajiwara, C. Pinzari and Y. Watatani, *Ideal structure and simplicity of the C^*-algebras generated by Hilbert modules*, J. Funct. Anal. **159** (1998), 295–322.

[KPW2] T. Kajiwara, C. Pinzari and Y. Watatani, *Hilbert C^*-bimodules and countably generated Cuntz-Krieger algebras*, J. Operator Theory **45** (2001), 3–18.

[KW] T. Kajiwara and Y. Watatani, *Hilbert C^*-bimodules and continuous Cuntz-Krieger algebras considered by Deaconu*, J. Math. Soc. Japan **54** (2002), 35–60.

[Kat] Y. Katayama, *Generalized Cuntz algebras $\mathcal{O}_N^M$*, RIMS kokyuroku **858** (1994), 87–90.

[Kir] E. Kirchberg, *The classification of purely infinite C^*-algebras using Kasparov's theory*, preprint,1994.

[Kit] B. P. Kitchens, *Symbolic dynamics*, Springer-Verlag, Berlin, Heidelberg and New York, 1998.

[Kr] W. Krieger, *On dimension functions and topological Markov chains*, Invent. Math. **56** (1980), 239–250.

[Kr2] W. Krieger, *On dimension for a class of homeomorphism groups*, Math. Ann **252** (1980), 87–95.

[Kr3] W. Krieger, *On sofic systems I*, Israel J. Math. **48** (1984), 305–330.

[Kr4] W. Krieger, *On sofic systems II*, Israel J. Math **60** (1987), 167–176.

[KM] W. Krieger and K. Matsumoto, *Shannon graphs, subshifts and lambda-graph systems*, to appear in J. Math. Soc. Japan.

[KPR] A. Kumjian, D. Pask and I. Raeburn, *Cuntz-Krieger algebras of directed graphs*, Pacific J. Math. **184** (1998), 161–174.

[KPRR] A. Kumjian, D. Pask, I. Raeburn and J. Renault, *Graphs, groupoids and Cuntz-Krieger algebras*, J. Funct. Anal. **144** (1997), 505–541.

[LM] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press., 1995.

[Ma] K. Matsumoto, *On C^*-algebras associated with subshifts*, Internat. J. Math. **8** (1997), 357–374.

[Ma2] K. Matsumoto, *K-theory for C^*-algebras associated with subshifts*, Math. Scand. **82** (1998), 237–255.

[Ma3] K. Matsumoto, *A simple C^*-algebra arising from a certain subshift*, J. Operator Theory **42** (1999), 351-370.

[Ma4] K. Matsumoto, *Dimension groups for subshifts and simplicity of the associated C^*-algebras*, J. Math. Soc. Japan **51** (1999), 679-698.

[Ma5] K. Matsumoto, *Presentations of subshifts and their topological conjugacy invariants*, Doc. Math. **4** (1999), 285-340.

[Ma6] K. Matsumoto, *Bowen-Franks groups for subshifts and Ext-groups for C^*-algebras*, K-Theory **23** (2001), 67–104.

[Ma7] K. Matsumoto, *C^*-algebras associated with presentations of subshifts II. Ideal structure and lambda-graph subsystems*, preprint, 2001.

[Pas] W. Paschke, *The crossed product of a C^*-algebra by an endomorphism*, Proc. Amer. Math. Soc. **80** (1980), 113–118.

[Ph] N. C. Phillips, *A classification theorem for nuclear purely infinite simple C^*-algebras*, Doc. Math. **5** (2000), 49–114.

[Pi] M. V. Pimsner, *A class of C^*-algebras, generalizing both Cuntz-Krieger algebras and crossed product by $\mathbb{Z}$*, in Free Probability Theory, Fields Institute Communications **12** (1996), 189–212.

[PV] M. Pimsner and D. Voiculescu, *Exact sequences for K-groups and Ext-groups of certain cross-products C^*-algebras*, J. Operator Theory **4** (1980), 93–118.

[PWY] C. Pinzari, Y. Watatani and K. Yonetani, *KMS states, entropy and the variational principle in full C^*-dynamical systems*, Commun. Math. Phys. **213** (2000), 331–381.

[Pu] I. Putnam, *C^*-algebras from Smale spaces*, Canad. J. Math. **48** (1996), 175–195.

[Re] J. N. Renault, *A groupoid approach to C^*-algebras*, Lecture Notes in Math. Springer **793** (1980).

[Ri] M. Rieffel, *C^*-algebras associated with irrational rotations*, Pacific J. Math. **93** (1981), 415–429.

[Rø] M. Rørdam, *Classification of Cuntz-Krieger algebras,*, K-Theory **9** (1995), 31–58.

[Rø2] M. Rørdam, *Classification of certain infinite simple C^*-algebras,*, J. Funct. Anal. **131** (1995), 415–458.

[Ro] J. Rosenberg, *Appendix to O. Bratteli's paper "Crossed products of UHF algebras "*, Duke Math. J. **46** (1979), 25–26.

[RS] J. Rosenberg and C. Schochet, *The Künneth theorem and the universal coefficient theorem for Kasparov's generalized K-functor*, Duke Math. J. **55** (1987), 431–474.

[T] M. Tomforde, *C^*-algebras of labeled graphs*, preprint.

[T2] M. Tomforde, *A unified approach to Exel-Laca algebras and C^*-algebras associated to directed graphs*, preprint.

[We] B. Weiss, *Subshifts of finite type and sofic systems*, Monats. Math. **77** (1973), 462–474.

Kengo Matsumoto
Department of Mathematical Science
Yokohama City University
Sato 22-2, Kanazawa-ku, Yokohama
236-0027, Japan
kengo@yokohama-cu.ac.jp

Effective Freeness of Adjoint Line Bundles

Gordon Heier

Received: March 19, 2002

Communicated by Thomas Peternell

Abstract. This note shows how two existing approaches to providing effective (quadratic) bounds for the freeness of adjoint line bundles can be linked to establish a new effective bound which approximately differs from the linear bound conjectured by Fujita only by a factor of the cube root of the dimension of the underlying manifold. As an application, a new effective statement for pluricanonical embeddings is derived.

2000 Mathematics Subject Classification: 14C20, 14F17, 14B05.
Keywords and Phrases: adjoint linear system – vanishing theorem – Fujita conjecture.

1 Introduction and statement of the Main Theorem

Let L be an ample line bundle over a compact complex projective manifold X of complex dimension n. Let K_X be the canonical line bundle of X. The following conjecture is due to Fujita [Fuj87].

Conjecture 1.1 (Fujita). The adjoint line bundle $K_X + mL$ is base point free (i.e. spanned by global holomorphic sections) for $m \geq n+1$. It is very ample for $m \geq n+2$.

The standard example of the hyperplane line bundle on $X = \mathbb{P}^n$ shows that the conjectured numerical bounds are optimal. In the case of X being a compact Riemann surface, the conjecture is easily verified by means of the Riemann-Roch theorem. Moreover, Reider [Rei88] was able to validate the conjecture also in the case $n = 2$. In higher dimensions, the very ampleness part of the conjecture has proved to be quite intractable so far. In fact, no further results seem to be known here. On the other hand, several further results have been established towards the freeness conjecture. The case $n = 3$ was solved by

Ein and Lazarsfeld [EL93] (see also [Fuj93]), and $n = 4$ is due to Kawamata [Kaw97]. In arbitrary dimension n, the state-of-the-art is that $K_X + mL$ is base point free for any integer m that is no less than a number roughly of order n^2 (see below for exact statements).
To the author's knowledge, [Dem00] constitutes the most recent survey on the subject under discussion. It contains an extensive list of references (see also the references at the end of this article) and, furthermore, introduces the reader to various other effective results in algebraic geometry.
The above-mentioned bound in the case of arbitrary dimension n can be derived from each of the following two theorems, due to Angehrn and Siu [AS95] (see also [Siu96]) and Helmke [Hel97], [Hel99], respectively. Although the proofs of these two theorems adhere to the same inductive approach, the key ideas at their cores are of a different nature. In Proposition 3.5, we will show how to use the techniques in question seamlessly in a back-to-back manner. This insight, together with the numerical considerations in Section 2, will lead to the improved bound asserted in the Main Theorem and proved in Section 3.
First, let us state the bound given by [AS95].

THEOREM 1.2 ([AS95]). *The line bundle $K_X + mL$ is base point free for $m \geq \frac{1}{2}n(n+1) + 1$.*

Secondly, we state Helmke's result. Due to the nature of his technique, the assumptions of his theorem are formulated in a slightly different way. We quote the result in the way it is presented in [Hel97], because the slight improvement achieved in [Hel99] is not relevant for our purposes.

THEOREM 1.3 ([HEL97]). *Assume that L has the additional properties that*

$$L^n > n^n$$

and for all $x \in X$:

$$L^d.Z \geq m_x(Z) \cdot n^d$$

for all subvarieties $Z \subset X$ with $x \in Z$, $d = \dim Z \leq n-1$ and multiplicity $m_x(Z) \leq \binom{n-1}{d-1}$ at x. Then $K_X + L$ is base point free.

If $n \geq 3$, it is clear that we need to set

$$m_0 := \max\{n \cdot \sqrt[d]{\binom{n-1}{d-1}} : d \in \mathbb{N} \text{ and } 1 \leq d \leq n\}$$

to determine the minimal bound m_0 deducible from Theorem 1.3 such that $K_X + mL$ is base point free for any integer $m \geq m_0$. Since

$$\sqrt[d]{\binom{n-1}{d-1}} \geq \frac{1}{\sqrt[d]{3}} \frac{n^{1-\frac{1}{d}}}{d}$$

according to our Lemma 2.3, we find that the m_0 which one can derive from Theorem 1.3 is essentially also of the order n^2.
We conclude this section with the statement of our Main Theorem, which asserts that the bound m_0 can be chosen to be a number of the order $n^{\frac{4}{3}}$.

Theorem 1.4 (Main Theorem). *The line bundle $K_X + mL$ is base point free for any integer m with*

$$m \geq (e + \frac{1}{2})n^{\frac{4}{3}} + \frac{1}{2}n^{\frac{2}{3}} + 1,$$

where $e \approx 2.718$ is Euler's number.

2 Estimates for binomial coefficients

In order to understand precisely the nature of the numerical conditions in the assumptions of Theorem 1.3, we prove some auxiliary estimates in this section. We begin with the following lemma.

Lemma 2.1. *For all $x \in]0,1[$: $1 < \left(\frac{1}{1-x}\right)^{\frac{1-x}{x}} < e$.*

Proof. It is obvious that 1 is a strict lower bound of the given expression, so it remains to show that

$$\left(\frac{1}{1-x}\right)^{\frac{1-x}{x}} < e.$$

Taking log on both sides of the inequality, we see that we are done if we can show that

$$g(x) := \frac{x-1}{x}\log(1-x) < 1$$

on the open unit interval. However, for this it suffices to prove that $\lim_{x\to 0^+} g(x) = 1$ and $g'(x) < 0$. The former is easily verified using L'Hôpital's rule, while the latter follows readily from a simple computation. □

In the proof of the subsequent Lemma 2.3, we will employ Lemma 2.1 in the form of the following corollary.

Corollary 2.2. *Let n be an integer ≥ 2. Let d be an integer with $1 \leq d \leq n-1$. Then*

$$1 < \left(\frac{n}{n-d}\right)^{\frac{n-d}{d}} < e.$$

Proof. We have

$$\left(\frac{n}{n-d}\right)^{\frac{n-d}{d}} = \left(\frac{1}{1-\frac{d}{n}}\right)^{\frac{1-\frac{d}{n}}{\frac{d}{n}}}.$$

Thus the corollary follows immediately from Lemma 2.1. □

The preceding considerations allow us to estimate the binomial coefficients from Theorem 1.3 in the form of the following lemma.

Lemma 2.3. *Let* $1 \leq d \leq n-1$. *Then*

$$\frac{1}{\sqrt[d]{3}} \frac{n^{1-\frac{1}{d}}}{d} \leq \sqrt[d]{\binom{n-1}{d-1}} \leq e\frac{n}{d}.$$

Proof. In [Ahl78], page 206, Stirling's formula is stated as

$$\Gamma(x) = \sqrt{2\pi}\, x^{x-\frac{1}{2}} e^{-x} e^{\frac{\theta(x)}{12x}}$$

for $x > 0$ with $0 < \theta(x) < 1$. In particular,

$$\sqrt{2\pi}\, x^{x-\frac{1}{2}} e^{-x} \leq \Gamma(x) \leq \sqrt{2\pi}\, x^{x-\frac{1}{2}} e^{-x} e^{\frac{1}{12}}$$

for any $x \geq 1$. Thus, for the proof of the desired estimate from above, Stirling's formula enables us to proceed as follows.

$$\begin{aligned}
\binom{n-1}{d-1} &= \frac{(n-1)!}{(d-1)!(n-d)!} = \frac{1}{n-d}\frac{\Gamma(n)}{\Gamma(d)\Gamma(n-d)} \\
&\leq \frac{1}{n-d}\frac{\sqrt{2\pi}\, n^{n-\frac{1}{2}} e^{-n} e^{\frac{1}{12}}}{\sqrt{2\pi}\, d^{d-\frac{1}{2}} e^{-d} \sqrt{2\pi}\,(n-d)^{n-d-\frac{1}{2}} e^{-(n-d)}} \\
&= \frac{e^{\frac{1}{12}}}{\sqrt{2\pi}} \sqrt{\frac{d}{(n-d)n}} \left(\frac{n}{d}\right)^d \left(\frac{n}{n-d}\right)^{n-d} \\
&\leq \left(\frac{n}{d}\right)^d \left(\frac{n}{n-d}\right)^{n-d}.
\end{aligned}$$

Resorting to Corollary 2.2, we eventually conclude:

$$\binom{n-1}{d-1}^{\frac{1}{d}} \leq \frac{n}{d}\left(\frac{n}{n-d}\right)^{\frac{n-d}{d}} \leq e\frac{n}{d}.$$

The desired estimate from below is proved analogously:

$$\begin{aligned}
\binom{n-1}{d-1} &= \frac{(n-1)!}{(d-1)!(n-d)!} = \frac{1}{n-d}\frac{\Gamma(n)}{\Gamma(d)\Gamma(n-d)} \\
&\geq \frac{1}{n-d}\frac{\sqrt{2\pi}\, n^{n-\frac{1}{2}} e^{-n}}{\sqrt{2\pi}\, d^{d-\frac{1}{2}} e^{-d} e^{\frac{1}{12}} \sqrt{2\pi}\,(n-d)^{n-d-\frac{1}{2}} e^{-(n-d)} e^{\frac{1}{12}}} \\
&= \frac{1}{e^{\frac{1}{6}}\sqrt{2\pi}} \sqrt{\frac{d}{(n-d)n}} \left(\frac{n}{d}\right)^d \left(\frac{n}{n-d}\right)^{n-d} \\
&\geq \frac{1}{3}\frac{1}{n} \left(\frac{n}{d}\right)^d \left(\frac{n}{n-d}\right)^{n-d}.
\end{aligned}$$

Using Corollary 2.2 again, we obtain:

$$\binom{n-1}{d-1}^{\frac{1}{d}} \geq \sqrt[d]{\frac{1}{3n}}\frac{n}{d} = \frac{1}{\sqrt[d]{3}}\frac{n^{1-\frac{1}{d}}}{d}.$$

□

3 Proof of the Main Theorem

The following theorem states the improved effective freeness bound which we shall prove at the end of this section.

Theorem 3.1 (Main Theorem). *The line bundle $K_X + mL$ is base point free for any integer m with*

$$m \geq (e + \frac{1}{2})n^{\frac{4}{3}} + \frac{1}{2}n^{\frac{2}{3}} + 1,$$

where $e \approx 2.718$ is Euler's number.

First of all, let us recall how a result of this type can be proved by means of multiplier ideal sheaves.
Let $x \in X$ be an arbitrary but fixed point. The key idea of both [AS95] and [Hel97] is to find an integer m_0 (as small as possible) and a singular metric h of the line bundle m_0L with the following two properties:

1. Let h be given locally by $e^{-\varphi}$. Then the curvature current $i\partial\bar{\partial}\varphi$ dominates a positive definite smooth $(1,1)$-form on X in the sense of currents.
2. Let the *multiplier ideal sheaf* of h be defined stalk-wise by ($\chi \in X$):

$$(\mathcal{I}_h)_\chi := \{f \in \mathcal{O}_{X,\chi} : |f|^2 e^{-\varphi} \text{ is locally integrable at } \chi\}.$$

Then, in a neighborhood of x, the zero set of $\mathcal{I}_h$, which we denote by $V(\mathcal{I}_h)$, is just the point x. (This is the key property we are looking for. Note that the support of $V(\mathcal{I}_h)$ is just the set of points where h is not locally integrable.)

The first property implies that

$$H^q(X, \mathcal{I}_h(K_X + m_0L)) = (0) \quad (q \geq 1),$$

due to the vanishing theorem of Nadel [Nad89], [Nad90]. (In the special case when the singular metric is algebraic geometrically defined, Nadel's vanishing theorem is the same as the theorem of Kawamata and Viehweg [Kaw82], [Vie82].) With this information and the second property, it is easy to obtain an element of $\Gamma(X, K_X + m_0L)$ which does not vanish at x. Namely, consider the short exact sequence

$$0 \to \mathcal{I}_h(K_X + m_0L) \to K_X + m_0L \to (\mathcal{O}_X/\mathcal{I}_h)(K_X + m_0L) \to 0.$$

The relevant part of the pertaining long exact sequence reads:

$$\Gamma(X, K_X + m_0L) \to \Gamma(V(\mathcal{I}_h), (\mathcal{O}_X/\mathcal{I}_h)(K_X + m_0L)) \to 0,$$

which implies by virtue of the second property that

$$\Gamma(X, K_X + m_0L) \overset{\text{restr.}}{\longrightarrow} \Gamma(\{x\}, \mathcal{O}_{\{x\}}(K_X + m_0L)) \to 0,$$

meaning that the restriction map to $\{x\}$ is surjective, which is what we intended to prove. Note that, since L is ample, there trivially exists a metric with the two aforementioned properties for every line bundle mL with $m \geq m_0$ (just multiply the metric for m_0L by the $(m - m_0)$-th power of a smooth positive metric of L).
In both [AS95] and [Hel97], the sought-after metric h is produced by an inductive method. First, here is the key statement proved in sections 7–9 of [AS95]. The cornerstone of its proof is a clever application of the theorem of Ohsawa and Takegoshi on the extension of L^2 holomorphic functions [OT87]. Note that, in contrast to [AS95], we are only concerned with freeness, and not point separation, so we can do without the complicated formulations found there.

Proposition 3.2 ([AS95]). *Let d be an integer with $1 \leq d \leq n-1$. Let k_d be a positive rational number, and let h_d be a singular metric of the line bundle k_dL. Assume that $x \in V(\mathcal{I}_{h_d})$ and $x \notin V(\mathcal{I}_{(h_d)^\gamma})$ for $\gamma < 1$. Moreover, assume that the dimensions of those components of $V(\mathcal{I}_{h_d})$ which contain x do not exceed d. Then there exist integers $d', k_{d'}$ with $0 \leq d' < d$ and $k_d < k_{d'} < k_d + d + \varepsilon$ (ε denotes a positive rational number which can be chosen to be arbitrarily small) and a singular metric $h_{d'}$ of $k_{d'}L$ such that $h_{d'}$ possesses the same properties as h_d, but with d and k_d replaced by d' and $k_{d'}$.*

Second, the key statement of [Hel97] is the following proposition. It is stated in such a way that it unites [Hel97], Proposition 3.2 (the inductive statement), and [Hel97], Corollary 4.6 (the multiplicity bound), into one ready-to-use statement. In its proof, the use of the aforementioned L^2 extension theorem is avoided by an explicit bound on the multiplicity of the minimal centers occurring in the inductive procedure.

Proposition 3.3 ([Hel97]). *Let d be an integer with $1 \leq d \leq n - 1$. Let*

$$L^n > n^n$$

and

$$L^{\tilde{d}}.Z \geq m_x(Z) \cdot n^{\tilde{d}}$$

for all subvarieties $Z \subset X$ such that $x \in Z$, $d \leq \tilde{d} = \dim Z \leq n - 1$ and multiplicity $m_x(Z) \leq \binom{n-1}{\tilde{d}-1}$ at x. Then there exists an integer $0 \leq d' < d$, a rational number $0 < c < 1$ and an effective $\mathbb{Q}$-divisor D such that D is $\mathbb{Q}$-linearly equivalent to cL, the pair (X, D) is log canonical at x and the minimal center of (X, D) at x is of dimension d'.

Let us briefly recall the definitions of some of the terms occurring in Proposition 3.3. First of all, for a pair (X, D) of a variety X and a $\mathbb{Q}$-divisor D, an *embedded resolution* is a proper birational morphism $\pi : Y \to X$ from a smooth variety Y such that the union of the support of the strict transform of D and the exceptional divisor of π is a normal crossing divisor. Next, we define the following notions, which are fundamental in the study of the birational geometry of pairs (X, D).

Definition 3.4. Let X be a normal variety and $D = \sum_i d_i D_i$ an effective $\mathbb{Q}$-divisor such that $K_X + D$ is $\mathbb{Q}$-Cartier. If $\pi : Y \to X$ is a birational morphism (in particular, an embedded resolution of the pair (X, D)), we define the *discrepancy divisor* of (X, D) under π to be

$$\sum_j b_j F_j := K_Y - \pi^*(K_X + D).$$

The pair (X, D) is called *log canonical* (resp. *Kawamata log terminal*) at x, if there exists an embedded resolution π such that $b_j \geq -1$ (resp. $b_j > -1$) for all j with $x \in \pi(F_j)$. Moreover, a subvariety Z of X containing x is said to be a *center of a log canonical singularity* at x, if there exists a birational morphism $\pi : Y \to X$ and a component F_j with $\pi(F_j) = Z$ and $b_j \leq -1$.

It follows from Shokurov's connectedness lemma in [Sho86] that the intersection of two centers of a log canonical singularity is again a center of a log canonical singularity (for a proof, see [Kaw97]). Thus there exists a unique minimal center of a log canonical singularity at x with respect to the inclusion of subvarieties on X.

In order to connect Proposition 3.2 and Proposition 3.3 for our purposes, we derive from the conclusion of Proposition 3.3 a statement about the existence of a certain singular metric:

Proposition 3.5. *Let (X, D) be a pair of a smooth projective variety X and an effective $\mathbb{Q}$-divisor D. Let $x \in X$ be an arbitrary but fixed point. Assume that the pair (X, D) is a log canonical at x with its minimal center at x being non-empty. Let $0 < c < 1$ be a rational number such that D is $\mathbb{Q}$-linearly equivalent to cL. Then there exists a singular metric h_D and a rational number c' (which can be chosen to be arbitrarily close to c) such that h_D is a metric of $c'L$, $x \in V(\mathcal{I}_{h_D})$ and $V(\mathcal{I}_{h_D})$ is contained in the minimal center of (X, D) at x in a neighborhood of x. Moreover, $x \notin V(\mathcal{I}_{(h_D)^\gamma})$ for $\gamma < 1$.*

Proof. Let s be a multivalued holomorphic section of cL whose $\mathbb{Q}$-divisor is D. This means that for some positive integer p with cp being an integer, the p-th power of s is the canonical holomorphic section of pcL with divisor pD. Let Z denote the minimal center of (X, D) and $\pi : Y \to X$ a log resolution of (X, D) with discrepancy divisor $\sum_j b_j F_j$. We choose π such that there exists at least one index j_0 with $b_{j_0} = -1$ and $\pi(F_{j_0}) = Z$. Furthermore, we set $\sum_j \delta_j F_j := \pi^*(D)$.

Since L is ample, we can choose a finite number of multivalued holomorphic sections $s_1, \ldots, s_q$ of L whose common zero set is exactly Z. Let $\delta_{i,j}$ denote the vanishing order of $\pi^* s_i$ along F_j at a generic point of F_j. If we set $\delta := \min\{\delta_{i,j_0} : i = 1, \ldots, q\}$, then $\delta > 0$ holds because all s_i vanish on Z.
For small positive rational numbers $\varepsilon, \varepsilon'$, we define the following singular metric of, say, $\tilde{c}L$:

$$\tilde{h}_D := \frac{1}{|s|^{2(1-\varepsilon)}} \frac{1}{(\sum_{i=1}^q |s_i|^2)^{\varepsilon'}}.$$

Whatever the choice of $\varepsilon, \varepsilon'$ may be, $\tilde{h}_D$ is locally integrable outside of Z in a small neighborhood of x. Here is how to choose $\varepsilon, \varepsilon'$ in order to make $\tilde{h}_D$ not integrable at x. In a small neighborhood U of x, the integrability of $\tilde{h}_D$ is equivalent to the integrability of

$$\pi^* \tilde{h}_D |\mathrm{Jac}(\pi)|^2 = \frac{1}{|s \circ \pi|^{2(1-\varepsilon)}} \frac{1}{(\sum_{i=1}^q |s_i \circ \pi|^2)^{\varepsilon'}} |\mathrm{Jac}(\pi)|^2$$

over every small open subset W of $\pi^{-1}(U)$. Note that $|\mathrm{Jac}(\pi)|^2$ can only be defined locally, and over W we take it to be the quotient

$$\frac{\pi^*(\omega_U \wedge \bar{\omega}_U)}{\omega_W \wedge \bar{\omega}_W},$$

where ω_U, ω_W are arbitrary but fixed nowhere vanishing local holomorphic n-forms on U and W, respectively. As we continue, we observe that there exists a small open subset W of $\pi^{-1}(U)$ such that $W \cap F_{j_0} \neq \emptyset$ and $\pi^* \tilde{h}_D |\mathrm{Jac}(\pi)|^2$ has a pole along $W \cap F_{j_0}$, with its order at a generic point of $W \cap F_{j_0}$ being

$$b_{j_0} + \varepsilon \delta_{j_0} - \varepsilon' \delta.$$

This number equals -1 if we choose ε arbitrarily and set $\varepsilon' := \frac{1}{\delta} \varepsilon \delta_{j_0}$. We conclude that, with these choices for ε and ε', $\tilde{h}_D$ is not integrable at x.
Finally, we set $h_D := (\tilde{h}_D)^r$ with

$$r := \min\{\rho : 0 < \rho \leq 1,\ (\tilde{h}_D)^\rho \text{ is not integrable at } x\}$$

to obtain the desired singular metric for $c'L$. Notice that if we let $\varepsilon \to 0$, then $r \to 1$, $\tilde{c} \to c$ and $c' \to c$. □

Now we are in a position to prove our Main Theorem.

Proof. Fix $x \in X$. Our goal is to prove that, if m_0 is the smallest integer no less than

$$(e + \frac{1}{2}) n^{\frac{4}{3}} + \frac{1}{2} n^{\frac{2}{3}} + 1,$$

there exists a singular metric h of the line bundle $m_0 L$ such that the two properties listed at the beginning of this section are satisfied. As was explained before, this is all that is necessary to prove the Main Theorem.

Let a be the smallest integer which is no less than $e\,n^{\frac{4}{3}}$. Let d_0 be the integral part of $n^{\frac{2}{3}}$. According to Lemma 2.3, we have

$$\binom{n-1}{\tilde{d}-1}^{\frac{1}{\tilde{d}}} n \leq e\,\frac{n^2}{\tilde{d}}$$

for all integers $\tilde{d}$ with $1 \leq \tilde{d} \leq n-1$. Furthermore,

$$e\,\frac{n^2}{\tilde{d}} \leq e\,n^{\frac{4}{3}} \leq a$$

for $\tilde{d} \geq n^{\frac{2}{3}}$. Thus we can use Proposition 3.3 to produce an effective $\mathbb{Q}$-divisor D such that D is $\mathbb{Q}$-linearly equivalent to caL for some $0 < c < 1$, the pair (X, D) is log canonical at x and its minimal center at x is of dimension d' for some integer d' with $0 \leq d' \leq d_0$. By Proposition 3.5, this translates into the existence of a singular metric h_1 of $c'aL$ $(0 < c' < 1)$ such that $x \in V(\mathcal{I}_{h_1})$, $x \notin V(\mathcal{I}_{(h_1)^\gamma})$ for $\gamma < 1$ and the dimensions of those components of $V(\mathcal{I}_{h_1})$ that contain x do not exceed d'.

From this point onwards, we can use the method of [AS95] in the form of Proposition 3.2 to produce inductively a singular metric h_2 such that $V(\mathcal{I}_{h_2})$ is isolated at x. If the constructed metric h_2 is a metric for, say, kL, then

$$k \leq c'a + 1 + 2 + \ldots + d_0 + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_{d_0}.$$

Since the ε_i can be chosen to be arbitrarily small positive rational numbers and since $c' < 1$, we can assume that

$$c'a + 1 + 2 + \ldots + d_0 + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_{d_0} < a + 1 + 2 + \ldots + d_0.$$

To obtain a metric of m_0L with the additional property that its curvature current dominates a positive definite smooth $(1,1)$-form on X in the sense of currents, we can simply multiply h_2 by the $(m_0 - k)$-th power of a smooth positive metric of L to obtain the desired metric h of m_0L. Note that $m_0 - k$ is a positive number because

$$\begin{aligned}
m_0 - k &> m_0 - (a + 1 + 2 + \ldots + d_0) \\
&\geq m_0 - (e\,n^{\frac{4}{3}} + 1 + 1 + 2 + \ldots + d_0) \\
&= m_0 - (e\,n^{\frac{4}{3}} + 1 + \frac{1}{2}d_0(d_0+1)) \\
&\geq m_0 - (e\,n^{\frac{4}{3}} + 1 + \frac{1}{2}n^{\frac{2}{3}}(n^{\frac{2}{3}}+1)) \\
&= m_0 - ((e + \frac{1}{2})n^{\frac{4}{3}} + \frac{1}{2}n^{\frac{2}{3}} + 1) \geq 0.
\end{aligned}$$

The proof of the Main Theorem is now complete. $\square$

4 Applications

As was indicated before, not much is known about the very ampleness part of the Fujita conjecture. The theorems and techniques mentioned in the previous sections do not seem to be directly applicable to it. However, Angehrn and Siu [AS95] were able to prove the following weaker analog to the very ampleness part of Fujita's conjecture, in which they assume that L, in addition to being ample, is also base point free. Their result improves on previous results of Ein, Küchle and Lazarsfeld [EKL95] and Kollar [Kol93].

Theorem 4.1 ([AS95]). *Let L be an ample line bundle over a compact complex manifold X of complex dimension n such that L is free. Let A be an ample line bundle. Then $(n+1)L + A + K_X$ is very ample.*

In conjunction with our Main Theorem, Theorem 4.1 can readily be applied to the case of an ample canonical line bundle in order to give the following effective statement on pluricanonical embeddings. As far as the author knows, this is the best effective statement on pluricanonical embeddings currently on hand. Note that Fujita's conjecture indicates that the statement of Corollary 4.2 should hold true for any integer $m \geq n+3$.

Corollary 4.2. *If X is a compact complex manifold of complex dimension n whose canonical bundle K_X is ample, then mK_X is very ample for any integer $m \geq (e+\frac{1}{2})n^{\frac{7}{3}} + \frac{1}{2}n^{\frac{5}{3}} + (e+\frac{1}{2})n^{\frac{4}{3}} + 3n + \frac{1}{2}n^{\frac{2}{3}} + 5$.*

Proof. Let m_0 be the smallest integer no less than $(e+\frac{1}{2})n^{\frac{4}{3}} + \frac{1}{2}n^{\frac{2}{3}} + 1$. According to our Main Theorem, $m_0K_X + K_X = (m_0+1)K_X$ is base point free (and, of course, ample). Thus we can apply Theorem 4.1 with $L = (m_0+1)K_X$ and $A = K_X$ to obtain that mK_X is very ample for any integer $m \geq (n+1)(m_0+1)+2$. A simple estimate yields the following upper bound for $(n+1)(m_0+1)+2$:

$$\begin{aligned}
& (n+1)(m_0+1)+2 \\
= \;& m_0(n+1)+n+3 \\
\leq \;& ((e+\frac{1}{2})n^{\frac{4}{3}} + \frac{1}{2}n^{\frac{2}{3}} + 1 + 1)(n+1) + n + 3 \\
= \;& (e+\frac{1}{2})n^{\frac{7}{3}} + \frac{1}{2}n^{\frac{5}{3}} + (e+\frac{1}{2})n^{\frac{4}{3}} + 3n + \frac{1}{2}n^{\frac{2}{3}} + 5.
\end{aligned}$$

□

Finally, we remark that our effective statement on pluricanonical embeddings can be used to sharpen the best known bound for the number of dominant holomorphic maps from a fixed compact complex manifold with ample canonical bundle to any variable compact complex manifold with big and numerically effective canonical bundle.

Acknowledgement. The author is supported by a doctoral student fellowship of the Studienstiftung des deutschen Volkes (German National Merit Foundation). The results published in this article were obtained as part of the author's research for his Ph.D. dissertation at the Ruhr-Universität Bochum under the auspices of Professor A. T. Huckleberry. It is a great pleasure to thank Professor Y.-T. Siu for introducing the author to the field of effective algebraic geometry, and for numerous discussions on the subject, which took place during visits to the Mathematics Department of Harvard University and the Institute of Mathematical Research at the University of Hong Kong. Special thanks go to Professor N. Mok for making an extended visit to Hong Kong possible. The author is a member of the Forschungsschwerpunkt "Globale Methoden in der komplexen Geometrie" of the Deutsche Forschungsgemeinschaft.

References

[Ahl78] L. Ahlfors. *Complex analysis*. McGraw-Hill Book Co., New York, third edition, 1978. An introduction to the theory of analytic functions of one complex variable, International Series in Pure and Applied Mathematics.

[AS95] U. Angehrn and Y.-T. Siu. Effective freeness and point separation for adjoint bundles. *Invent. Math.*, 122(2):291–308, 1995.

[Dem93] J.-P. Demailly. A numerical criterion for very ample line bundles. *J. Differential Geom.*, 37(2):323–374, 1993.

[Dem00] J.-P. Demailly. Méthodes L^2 et résultats effectifs en géométrie algébrique. *Astérisque*, (266):Exp. No. 852, 3, 59–90, 2000. Séminaire Bourbaki, Vol. 1998/99.

[EKL95] L. Ein, O. Küchle, and R. Lazarsfeld. Local positivity of ample line bundles. *J. Differential Geom.*, 42(2):193–219, 1995.

[EL93] L. Ein and R. Lazarsfeld. Global generation of pluricanonical and adjoint linear series on smooth projective threefolds. *J. Amer. Math. Soc.*, 6(4):875–903, 1993.

[Fuj87] T. Fujita. On polarized manifolds whose adjoint bundles are not semipositive. In *Algebraic geometry, Sendai, 1985*, pages 167–178. North-Holland, Amsterdam, 1987.

[Fuj93] T. Fujita. Remarks on Ein-Lazarsfeld criterion of spannedness of adjoint bundles of polarized threefolds. LANL-preprint alg-geom/9311013, 1993.

[Hel97] S. Helmke. On Fujita's conjecture. *Duke Math. J.*, 88(2):201–216, 1997.

[Hel99] S. Helmke. On global generation of adjoint linear systems. *Math. Ann.*, 313(4):635–652, 1999.

[Kaw82] Y. Kawamata. A generalization of Kodaira-Ramanujam's vanishing theorem. *Math. Ann.*, 261(1):43–46, 1982.

[Kaw97] Y. Kawamata. On Fujita's freeness conjecture for 3-folds and 4-folds. *Math. Ann.*, 308(3):491–505, 1997.

[Kol93] J. Kollár. Effective base point freeness. *Math. Ann.*, 296(4):595–605, 1993.

[Kol97] J. Kollár. Singularities of pairs. In *Algebraic geometry—Santa Cruz 1995*, pages 221–287. Amer. Math. Soc., Providence, RI, 1997.

[Nad89] A. Nadel. Multiplier ideal sheaves and existence of Kähler-Einstein metrics of positive scalar curvature. *Proc. Nat. Acad. Sci. U.S.A.*, 86(19):7299–7300, 1989.

[Nad90] A. Nadel. Multiplier ideal sheaves and Kähler-Einstein metrics of positive scalar curvature. *Ann. of Math. (2)*, 132(3):549–596, 1990.

[OT87] T. Ohsawa and K. Takegoshi. On the extension of L^2 holomorphic functions. *Math. Z.*, 195(2):197–204, 1987.

[Rei88] I. Reider. Vector bundles of rank 2 and linear systems on algebraic surfaces. *Ann. of Math. (2)*, 127(2):309–316, 1988.

[Sho86] V. V. Shokurov. The non-vanishing theorem. *Math. USSR, Izv.*, 26:591–604, 1986.

[Siu96] Y.-T. Siu. The Fujita conjecture and the extension theorem of Ohsawa-Takegoshi. In *Geometric complex analysis (Hayama, 1995)*, pages 577–592. World Sci. Publishing, River Edge, NJ, 1996.

[Smi97] K. Smith. Fujita's freeness conjecture in terms of local cohomology. *J. Algebraic Geom.*, 6(3):417–429, 1997.

[Vie82] E. Viehweg. Vanishing theorems. *J. Reine Angew. Math.*, 335:1–8, 1982.

Gordon Heier
Ruhr-Universität Bochum
Fakultät für Mathematik
D-44780 Bochum
Germany
heier@cplx.ruhr-uni-bochum.de

Current address:
Harvard University
Department of Mathematics
Cambridge, MA 02138
USA
heier@math.harvard.edu

Planar Graphs as Minimal Resolutions of Trivariate Monomial Ideals

Ezra Miller

Received: May 4, 2001
Revised: March 11, 2002

Communicated by Günter M. Ziegler

Abstract. We introduce the notion of rigid embedding in a grid surface, a new kind of plane drawing for simple triconnected planar graphs. Rigid embeddings provide methods to (1) find well-structured (cellular, here) minimal free resolutions for arbitrary monomial ideals in three variables; (2) strengthen the Brightwell–Trotter bound on the order dimension of triconnected planar maps by giving a geometric reformulation; and (3) generalize Schnyder's angle coloring of planar triangulations to arbitrary triconnected planar maps via geometry. The notion of rigid embedding is stable under duality for planar maps, and has certain uniqueness properties.

2000 Mathematics Subject Classification: 05C10, 13D02, 06A07, 13F55, 68R10, 52Cxx
Keywords and Phrases: planar graph, monomial ideal, free resolution, order dimension, (rigid) geodesic embedding

Contents

INTRODUCTION

Simple triconnected planar graphs admit numerous characterizations. Two famous examples include Steinitz' theorem on the edge graphs of 3-polytopes, and the Koebe–Andreev–Thurston circle packing theorem (see [Zie95] for both). These results produce "correct" planar (or spherical) drawings of the graphs in question, from which a great deal of geometric and combinatorial information flows readily.

This paper introduces a new kind of plane drawing for simple triconnected planar graphs, from which a great deal of *algebraic* and combinatorial information flows readily. These **geodesic embeddings** inside **grid surfaces** provide methods to

- solve the problem of finding well-structured (cellular, in this case) minimal free resolutions for arbitrary monomial ideals in three variables;

- strengthen the Brightwell–Trotter bound on the order dimension of triconnected planar maps [BT93] by giving a geometric reformulation; and
- generalize Schnyder's angle coloring for planar triangulations [Tro92, Chapter 6] to arbitrary triconnected planar maps via geometry.

We note that Felsner's generalization of Schnyder's angle coloring [Fel01] coincides with the **orthogonal colorings** independently discovered here as consequences of geometric considerations. In parallel with circle-packed and polyhedral graph drawings, additional evidence for the naturality of geodesic embeddings comes from their stability under duality, and the uniqueness properties enjoyed by "correct" geodesic embeddings—called **rigid embeddings** in what follows—for a given planar map.

The plan of the paper is as follows. Immediately following this Introduction is a section containing two theorems summarizing the equivalences and constructions forming main results of the paper. After that, the paper is divided into three Parts.

Part I lays the groundwork for geodesic and rigid embeddings in grid surfaces, and is geared almost entirely toward proving Theorem 5.1: the rigid embedding theorem. Terminology for the rest of the paper is set in Section 1, which also states a standard criterion for triconnectivity under edge contraction that serves as an inductive tool in the proof of Theorem 5.1. Then Section 2 presents the definition of grid surfaces, as well as the vertex and edge axioms for geodesic and rigid embeddings. Their consequences, the region and rigid region axioms, appear in Propositions 2.3 and 2.4. The first connection with order dimension comes in Corollary 2.5.

Sections 3 and 4 consist of stepping stones to the rigid embedding theorem. The basic inductive step for abstract planar maps is Lemma 3.1, which motivates the preliminary grid surface construction of Lemma 3.2. Induction for grid surfaces occupies the three Propositions in Section 4. They have been worded so that their rather technical proofs (particularly that of Proposition 4.2) may be skipped the first time through; instead, the Figures should provide ample intuition.

Section 5 completes the induction with a few more arguments about abstract planar maps. Corollary 5.2 recovers the Brightwell–Trotter bound on order dimension from rigid embedding.

The focus shifts in Part II to the algebra of monomial ideals in three variables, specifically their minimal free resolutions. A review of the standard tools occupies Section 6, while Section 7 recaps the more recent theory of cellular resolutions, along with a triconnectivity result (Proposition 7.2) suited to the applications here. Theorem 8.4 says how geodesic embeddings become minimal free resolutions. Corollary 8.5 then characterizes triconnectivity as the condition guaranteeing that a planar map supports a minimal free resolution of some artinian monomial ideal.

Section 9 displays even more reasons why rigid embeddings are better than arbitrary geodesic embeddings: they have a strong uniqueness property (Corollary 9.1), which implies in particular that every minimal cellular resolution of the corresponding monomial ideal is planar. Surprisingly, there can exist *non-planar* cell complexes supporting minimal free resolutions of trivariate artinian monomial ideals that are sufficiently nonrigid; Example 9.2 illustrates one.

Sections 10–12 are devoted to producing minimal cellular free resolutions of arbitrary monomial ideals in three variables (Theorem 11.1). The deformations reviewed in Section 10 serve as part of the algorithmic solution pseudocoded in Algorithm 11.2. The proof of correctness for the algorithm and the theorem, which occupy Section 12, are rather technical and delicate. As with Section 4, the pictures may give a better feeling for the methods than the proofs themselves, at least upon first reading.

Part III continues where Part I left off, with more combinatorial theory for planar maps. Section 13 introduces orthogonal coloring, which generalizes Schnyder's angle coloring and abstracts the notion of geodesic embedding (Proposition 13.1). Then, Section 14 shows how orthogonal coloring encodes the abstract versions of the orthogonal flows that played crucial roles in Section 2. As a consequence, Proposition 14.2 shows that orthogonal flows are examples of—but somewhat better than—normal families of paths, connecting once again with the work of Brightwell and Trotter on order dimension. Section 15 demonstrates how Alexander duality for grid surfaces (or monomial ideals) manifests itself as duality for planar maps geodesically embedded in grid surfaces.

Finally, Section 16 presents some open problems related to the notions developed in earlier sections, including a conjecture on orthogonal colorings and some problems on classifying cell complexes supporting minimal resolutions. Further questions concern applications of the present results to broader combinatorial algebraic problems, notably how to describe the "moduli space" of all minimal free (or injective) resolutions of ideals generated by a fixed number of monomials.

After completing an earlier version of this paper, the author was informed that Stefan Felsner had independently discovered the theory in Sections 13 and 14 [Fel01, Sections 1 and 2]. In addition, Felsner proved Conjecture 16.3 in [Fel02] after reading the preliminary version of this paper. See Section 16.3 for details and consequences.

Part III is almost logically independent of Part II, the only exceptions being Lemmas 8.2 and 8.3. Thus, the reader interested primarily in the combinatorics of planar graphs (as opposed to resolutions of monomial ideals) can read Parts I and III, safely skipping everything in Part II except for these two lemmas. The reader interested primarily in resolutions of monomial ideals should skip everything in Sections 3–5 except for the statement of Theorem 5.1.

Acknowledgements

This paper grew out of conjectures developed with Bernd Sturmfels during a memorable train ride through the Alps, and subsequent discussions resulting in the expository paper [MS99], where some of the results were announced without proof. The Alfred P. Sloan Foundation and the National Science Foundation funded various stages of this project.

Summary theorems

For the sake of perspective and completeness, we collect the main ideas of the paper into a pair of precisely stated summary theorems. Their proofs are included, in the sense that the appropriate results from later on are cited. All of the notions appearing in Theorems A and B will be introduced formally in due time; until then, brief descriptions along with Figure 1 should suffice.

Let M be a connected simple planar map—that is, a graph embedded in a surface S homeomorhpic to the plane $\mathbb{R}^2$. All graphs in this paper have finitely many vertices and edges. Fix a point $\infty \in S$ far from M, and define the exterior region of M to be the connected component of $S \smallsetminus M$ containing ∞. Given three vertices $\dot{x}, \dot{y}, \dot{z} \in M$ bordering the exterior region, form the extended map $M_\infty(\dot{x}, \dot{y}, \dot{z})$ by connecting $\dot{x}, \dot{y}, \dot{z}$ to ∞. Call a graph triconnected either if it is a triangle, or if it has at least four vertices of which deleting any pair along with their incident edges leaves a connected graph. A set of paths leaving a fixed vertex $\nu \in M$ is said to be independent if their pairwise intersection is $\{\nu\}$.

Let $k[x, y, z]$ be the polynomial ring in three variables over a field k, and let $I \subset k[x, y, z]$ be an ideal generated by monomials. The grid surface $\mathcal{S}$ corresponding to I is the boundary of the staircase diagram of I, which is drawn (as usual) as the stack of cubes corresponding to monomials not in I. Rigid embedding of a planar map M in $\mathcal{S}$ involves identifying the edges of M as certain piecewise linear geodesics in $\mathcal{S}$, and constitutes an inclusion of the vertex-edge-face poset of M into $\mathbb{N}^3$. Orthogonal coloring M involves coloring the angles in M with three colors according to certain rules. Since it would take too long to do real justice to the definitions of 'rigid embedding' and 'orthogonal coloring' here, Figure 1 will have to do for now. The outer corners in the orthogonal coloring and the vectors on the axes in $\mathcal{S}$ are called axial vertices. The grid surface $\mathcal{S}$ is called axial when I is artinian.

Suppose M is a cell complex (finite CW complex) whose faces are labeled by vectors in $\mathbb{N}^3$, in such a way that the union $M_{\preceq\alpha}$ of faces whose labels precede $\alpha \in \mathbb{N}^3$ is a subcomplex of M for every α. Roughly speaking, M supports a cellular free resolution of I if the boundary complex of $M_{\preceq\alpha}$ with coefficients in k is the $\mathbb{N}^3$-degree α piece of a free resolution of I, for every $\alpha \in \mathbb{N}^3$.

Theorem A *Let M be a planar map. The following are equivalent.*

1. *M has three vertices $\dot{x}, \dot{y}, \dot{z}$ bordering its exterior region for which $M_\infty(\dot{x}, \dot{y}, \dot{z})$ is triconnected.*
2. *M has three vertices $\dot{x}, \dot{y}, \dot{z}$ bordering its exterior region to which every vertex of M has independent paths.*
3. *M has an orthogonal coloring with axial vertices $\dot{x}, \dot{y}, \dot{z}$.*
4. *M can be rigidly embedded in an axial grid surface.*

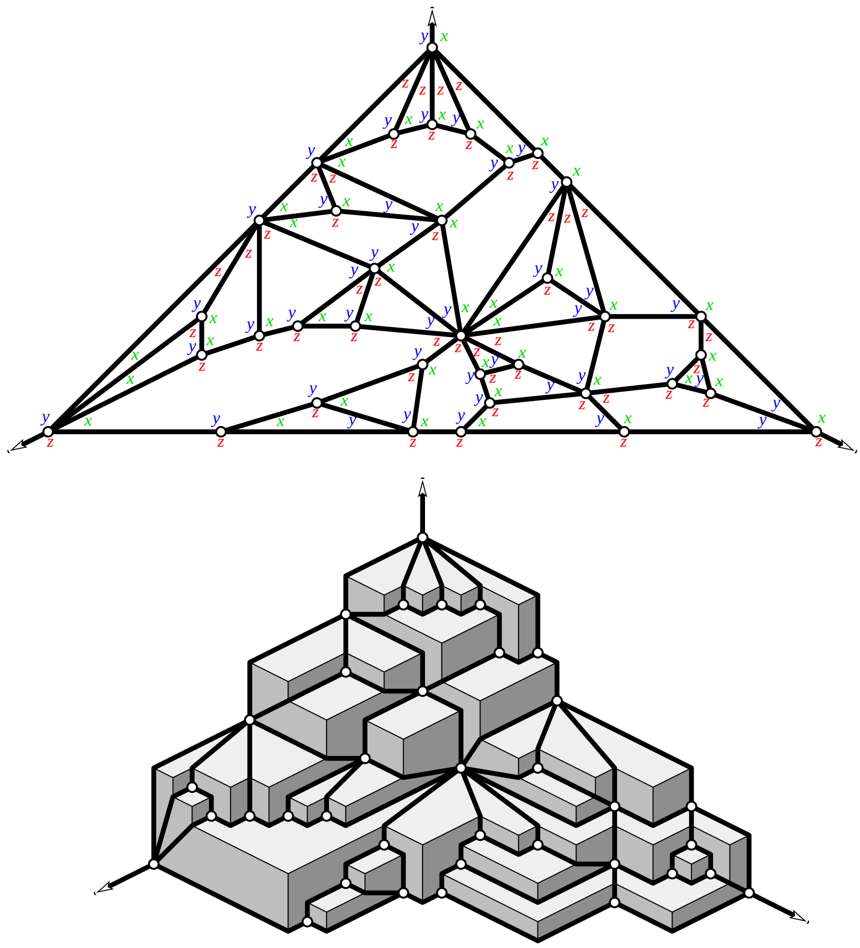


Figure 1: Orthogonal coloring and rigid embedding of an extended map

5. *M supports a cellular minimal free resolution of some artinian monomial ideal in $k[x, y, z]$.*

Every artinian monomial ideal in $k[x, y, z]$ has a minimal cellular resolution supported on a cell complex M satisfying these conditions; in fact, Algorithm 11.2 produces such an M automatically.

Proof. $4 \Rightarrow 3$ follows from Proposition 13.1.
$3 \Rightarrow 2$ follows from Proposition 14.2.
$2 \Rightarrow 1$ follows easily from the definitions.
$1 \Rightarrow 4$ is Theorem 5.1.
$4 \Rightarrow 5$ follows from Theorem 8.4.
$5 \Rightarrow 1$ is Proposition 7.2.

The final statement comes from Theorem 11.1 and Proposition 12.4. □

Similar—but weaker—statements apply to minimal cellular free resolutions of arbitrary (not necessarily artinian) monomial ideals in $k[x, y, z]$.

Theorem B *Let N be a planar map. The following two conditions are equivalent.*

1. *N can be rigidly embedded in some grid surface.*
2. *N can be obtained by deleting $\dot{x}, \dot{y}, \dot{z}$ and all edges incident to them from some planar map M satisfying the equivalent conditions in Theorem A.*

These conditions imply that

3. *N supports a minimal free resolution of some monomial ideal in $k[x, y, z]$.*

Every monomial ideal in $k[x, y, z]$ has a minimal free resolution supported on a planar map N satisfying conditions 1 and 2; such an N can be produced algorithmically.

Proof. $1 \Rightarrow 2$ follows from Theorem 8.4 and Lemma 8.2.
$2 \Rightarrow 1$ follows from Theorem 5.1 and Lemma 8.2.
$1 \Rightarrow 3$ follows from Theorem 8.4.
The first half of the final statement is Theorem 11.1 along with the first paragraph of its proof on p. 80; add in Proposition 12.4 for the algorithmic part. □

In reality, the more detailed versions later on are considerably more precise, demonstrating how some of the equivalent descriptions naturally give rise to others.

Part I
Geodesic embedding in grid surfaces

1 Planar maps

Let $\mathcal{V} = \{\nu_1, \ldots, \nu_r\}$ be a finite set. A **graph** G with vertex set $\mathcal{V}$ is uniquely determined by a collection $\mathcal{E} \subseteq \binom{\mathcal{V}}{2}$ of edges, each consisting of a pair of vertices. Except for one paragraph at the beginning of Section 15, we consider only **simple graphs**—that is, without loops or multiple edges—so G is an abstract simplicial complex of dimension 1 having vertex set $\mathcal{V}$. Thus G can be regarded as a topological space, via any geometric realization.
Let S be a surface homeomorphic to the Euclidean plane $\mathbb{R}^2$. A **plane drawing** of G in S is a continuous morphism $G \hookrightarrow S$ of topological spaces that is a homeomorphism onto its image. If G is connected, the image M is called a **planar map**. Deleting the images of the vertices and edges of G from S leaves several connected components whose closures are the **regions** of M. The unique unbounded region is called the **exterior region** of M. Two planar maps

are **isomorphic** if they result from plane drawings of the same graph G, their regions have the same boundaries in G, and the boundaries of their exterior regions correspond. We often blur the distinction between a planar map and the underlying graph, by not distinguishing a vertex (resp. edge) of G from the corresponding point (resp. arc) of M in the surface S.

A graph G is **k-connected** either if G is the complete graph on k vertices, or if G has at least $k+1$ vertices, and given any $k-1$ vertices $\nu_1, \ldots, \nu_{k-1}$ of G, the **deletion** $\mathrm{del}(G; \nu_1, \ldots, \nu_{k-1})$ is connected. Here, the deletion is obtained by removing $\nu_1, \ldots, \nu_{k-1}$ as well as all edges containing them from G. In case $k = 2$ or 3, the graph G is called **biconnected** or **triconnected**, respectively.

Suppose that e is an edge of a planar map M, and that none of the (one or two) regions containing e is a triangle. The **contraction** M/e of M along e is obtained by removing the edge e and identifying the two vertices of e. The underlying graph of M/e is the topological quotient G/e; it is still simple because e is the only edge connecting its vertices in G (so G/e has no loops) and no triangles contain e in G (so G/e has no multiple edges). Some plane drawing of M/e is obtained by literally contracting the edge e in M (technically: there is a homotopy $G \times [0,1] \to S$ such that $G \times t \to S$ is a plane drawing of G for $t < 1$, while $G \times 1 \to S$ is a composition $G \to G/e \to S$ with the second map being a plane drawing). Contraction will be a crucial inductive tool, via a well-known criterion for triconnectivity under contraction:

Proposition 1.1 *Let M be a triconnected planar map with at least four vertices, and let e be an edge. If there exist two regions F, F' of M such that*

1. *$e \cap F$ and $e \cap F'$ are the two vertices of e, and*
2. *$F \cap F'$ is nonempty,*

then either e borders a triangle or the contraction M/e fails to be triconnected. Conversely, if e borders no triangles and M/e is triconnected, then no such F, F' exist.

Thinking of the surface $S \cong \mathbb{R}^2$ as the 2-sphere minus ∞, many of the planar maps M in this paper result by embedding some graph G_∞ in the sphere with ∞ as a vertex, and then considering the induced plane drawing M of $\mathrm{del}(G_\infty; \infty)$. When this is the case, we frequently need to consider the subset $M_\infty \subset S$ obtained by omitting the point ∞ from the plane drawing of G_∞ in the sphere; thus some of the vertices in M connect to the missing point ∞ by unbounded arcs in S. More generally, define an **extended map** $M_\infty \subset S$ to be the union of a planar map M and a set of infinite nonintersecting arcs connecting some of its vertices to ∞. The closure $\overline{M}_\infty$ of M_∞ in the sphere need not be a simple graph because it can have doubled edges: some vertex in M could have two or more unbounded arcs in M_∞ containing it.

Suppose the edges contained in the exterior region of M form a simple closed curve, called the **exterior cycle**. This occurs, for instance, when M is triconnected. Three vertices $\dot{x}, \dot{y}, \dot{z} \in M$ are called **axial** if they are encountered (in order) proceeding counterclockwise around the exterior cycle. Having chosen axial vertices, define $M_\infty(\dot{x}, \dot{y}, \dot{z}) \subset S$ to be the union of M and three unbounded arcs, called the x, y, and z-**axes**, connecting $\dot{x}, \dot{y}, \dot{z}$ to ∞. We sometimes blur the distinction between $M_\infty(\dot{x}, \dot{y}, \dot{z})$ and its closure $\overline{M}_\infty(\dot{x}, \dot{y}, \dot{z})$ in the sphere. For instance, we say that $M_\infty(\dot{x}, \dot{y}, \dot{z})$ is triconnected if the graph underlying $\overline{M}_\infty(\dot{x}, \dot{y}, \dot{z})$ is.

2 Grid surfaces

Let $\mathbb{R}$ denote the real numbers. Write vectors in $\mathbb{R}^3$ as $\alpha = (\alpha_x, \alpha_y, \alpha_z)$, and partially order $\mathbb{R}^3$ by setting $\alpha \preceq \beta$ (read 'α **precedes** β') whenever $\alpha_u \leq \beta_u$ for all $u \in \{x, y, z\}$. Say that $\alpha \in \mathbb{R}^3$ **strongly precedes** $\beta \in \mathbb{R}^3$ when $\alpha_u < \beta_u$ for all $u = x, y, z$; this is stronger than saying $\alpha \prec \beta$. (Throughout this paper, the letter u denotes any one of x, y, z, in the same way that x_i denotes one of $x_1, \ldots, x_n$.) Use $\alpha \vee \beta$ and $\alpha \wedge \beta$ to denote the **join** (componentwise maximum) and **meet** (componentwise minimum) of $\alpha, \beta \in \mathbb{R}^3$.

Let $\mathcal{V} \subset \mathbb{N}^3 \subset \mathbb{R}^3$ be a set of pairwise incomparable elements, where $\mathbb{N}$ denotes the set of nonnegative integers. The order filter

$$\langle \mathcal{V} \rangle \quad = \quad \{\alpha \in \mathbb{R}^3 \mid \alpha \succeq \nu \text{ for some } \nu \in \mathcal{V}\}$$

generated by $\mathcal{V}$ is a closed subset of the topological space $\mathbb{R}^3$. Its boundary $\mathcal{S}_\mathcal{V}$ is called a **grid surface** or **staircase**. Orthogonal projection onto the plane $x + y + z = 0$ restricts to a homeomorphism $\mathcal{S}_\mathcal{V} \cong \mathbb{R}^2$. (This homeomorphism gives the correspondence between rhombic tilings of the orthogonal projection of the $|\dot{x}| \times |\dot{y}| \times |\dot{z}|$ parallelepiped and plane partitions of the $|\dot{x}| \times |\dot{y}|$ grid with parts at most $|\dot{z}|$. The grid surface in Figure 1 clearly demonstrates the homeomorphism: the diagram is, after all, drawn faithfully on the two-dimensional page.)

One of the basic properties of grid surfaces is that $\alpha \in \mathcal{S}_\mathcal{V}$ whenever $\rho, \sigma \in \mathcal{S}_\mathcal{V}$ and $\rho \preceq \alpha \preceq \sigma$. Therefore, if $\rho, \sigma \in \mathcal{S}_\mathcal{V}$ and $\rho \preceq \sigma$, then $\mathcal{S}_\mathcal{V}$ contains the line segment in $\mathbb{R}^3$ connecting ρ to σ. In particular, if $\nu, \omega \in \mathcal{V}$ satisfy $\nu \vee \omega \in \mathcal{S}_\mathcal{V}$, then $\mathcal{S}_\mathcal{V}$ contains the union $[\nu, \omega]$ of the two line segments joining ν and ω to $\nu \vee \omega$; we refer to such arcs as **elbow geodesics**[1] in $\mathcal{S}_\mathcal{V}$. When ν and ω are the *only* vectors in $\mathcal{V}$ preceding $\nu \vee \omega$, the arc $[\nu, \omega]$ is called a **rigid geodesic**.

Denote the nonnegative rays of the coordinate axes in $\mathbb{R}^3$ by X, Y, and Z, and use the letter U to refer to any of X, Y, Z. The ray $\nu + U$ intersects $\mathcal{S}_\mathcal{V}$ in an oriented line segment U_ν called the **orthogonal ray** leaving ν in the direction of U. Thus every point in $\mathcal{V}$ has precisely three orthogonal rays, one parallel to each coordinate axis and all contained in $\mathcal{S}_\mathcal{V}$, although some orthogonal rays may be unbounded while others are bounded.

If U_ν is bounded, so it has an endpoint besides ν, then the other endpoint of U_ν can always be expressed as a join $\nu \vee \omega$ for some $\omega \in \mathcal{V}$. When there is

[1] Elbow geodesics do minimize length for the metric on $\mathcal{S}_\mathcal{V}$ induced by the usual metric on $\mathbb{R}^3$, but this fact has no practical application in this paper.

exactly one such point ω, so $[\nu,\omega]$ is a rigid geodesic, we say that ν or U_ν **points toward** ω.

Observe that $\nu\vee\omega$ must share two coordinates with at least one (and perhaps both) of ν and ω, so every elbow geodesic contains at least one orthogonal ray. Making compatible choices of elbow geodesics containing all orthogonal rays yields a planar map. To be precise, a plane drawing $M \hookrightarrow \mathcal{S}_\mathcal{V}$ is a **geodesic grid surface embedding**, or simply a **geodesic embedding** in $\mathcal{S}_\mathcal{V}$, if the following two axioms are satisfied:

(Vertex axiom) The vertices of M coincide with $\mathcal{V}$.

(Elbow geodesic axiom) Every edge of M is an elbow geodesic in $\mathcal{S}_\mathcal{V}$, and every bounded orthogonal ray in $\mathcal{S}_\mathcal{V}$ is part of an edge of M.

With the following stronger edge axiom instead, $M \hookrightarrow \mathcal{S}_\mathcal{V}$ is a **rigid embedding**, which we sometimes phrase by saying that M is **rigidly embedded** in $\mathcal{S}_\mathcal{V}$:

(Rigid geodesic axiom) The elbow geodesic exiom holds, and every edge of M is a rigid geodesic in $\mathcal{S}_\mathcal{V}$.

The rigid geodesic axiom really consists of three parts, each of which puts nontrivial restrictions on $\mathcal{S}_\mathcal{V}$ or M: every bounded orthogonal ray in $\mathcal{S}_\mathcal{V}$ is part of a rigid geodesic (*a priori*, this has nothing to do with M); every rigid geodesic in $\mathcal{S}_\mathcal{V}$ is an edge of M; and every edge of M is a rigid geodesic in $\mathcal{S}_\mathcal{V}$.

Lemma 2.1 *Let $M \hookrightarrow \mathcal{S}_\mathcal{V}$ be a geodesic or rigid embedding. Suppose $\mathcal{V}$ is in order-preserving bijection with another set $\widetilde{\mathcal{V}}$ of vertices via $\nu \leftrightarrow \widetilde{\nu}$, so that $\nu_u \leq \omega_u \Leftrightarrow \widetilde{\nu}_u \leq \widetilde{\omega}_u$ for all $\nu,\omega \in \mathcal{V}$ and $u \in \{x,y,z\}$. Then the elbow or rigid geodesics in $\mathcal{S}_{\widetilde{\mathcal{V}}}$ constitute another geodesic or rigid embedding of M. In particular, linearly scaling one or more coordinate axes by integer factors preserves geodesic or rigid embeddings.*

Proof. Purely order-theoretic properties of $\mathcal{V}$ determine whether ν and ω are the endpoints of an elbow geodesic, or whether ν points toward ω. $\square$

Any geodesic embedding $M \hookrightarrow \mathcal{S}_\mathcal{V}$ determines an extended map

$$M_\infty \quad = \quad M \cup (\text{unbounded orthogonal rays}).$$

A special case occurs when $\mathcal{S}_\mathcal{V}$ is **axial**, having axial vectors

$$\dot{x} = (|\dot{x}|,0,0), \quad \dot{y} = (0,|\dot{y}|,0), \quad \text{and} \quad \dot{z} = (0,0,|\dot{z}|)$$

in $\mathcal{V}$ for nonzero $|\dot{x}|,|\dot{y}|,|\dot{z}| \in \mathbb{N}$. Thus, if M is geodesically embedded in an axial grid surface $\mathcal{S}_\mathcal{V}$, we can define the axial vertices of M to be the axial vectors in $\mathcal{V}$, and set $M_\infty(\dot{x},\dot{y},\dot{z}) = M \cup X_{\dot{x}} \cup Y_{\dot{y}} \cup Z_{\dot{z}}$. (Precisely two bounded orthogonal rays leave each axial vertex, while all three orthogonal rays leaving any other vertex are bounded). Conversely, if M comes equipped with axial

vertices $\dot{x}, \dot{y}, \dot{z}$, then we require any geodesic embedding $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ to send these axial vertices to axial vectors in $\mathcal{V}$.
Suppose M is geodesically embedded in the axial grid surface $\mathcal{S}_{\mathcal{V}}$. The edge of M leaving any vertex $\nu \neq \dot{z}$ along the vertical orthogonal ray Z_ν connects ν to another vertex ω with strictly larger z-coordinate, but weakly smaller x and y-coordinates. Continuing in this manner constructs an **orthogonal flow** $[\nu, \dot{z}]$ from ν to $\dot{z}$ that is increasing in z, but weakly decreasing in x and y. It follows that $[\nu, \dot{z}]$ and the similarly constructed paths $[\nu, \dot{x}]$ and $[\nu, \dot{y}]$ are **independent**, meaning that they intersect pairwise only at ν itself. Since $[\dot{x}, \dot{y}]$, $[\dot{y}, \dot{z}]$, and $[\dot{z}, \dot{x}]$ partition the exterior cycle of M into three arcs, the contractible sets bounded by

$$[\dot{x}, \nu, \dot{y}] \quad := \quad [\nu, \dot{x}] \cup [\nu, \dot{y}] \cup [\dot{x}, \dot{y}]$$

and its cyclically permuted analogues partition the regions of M.

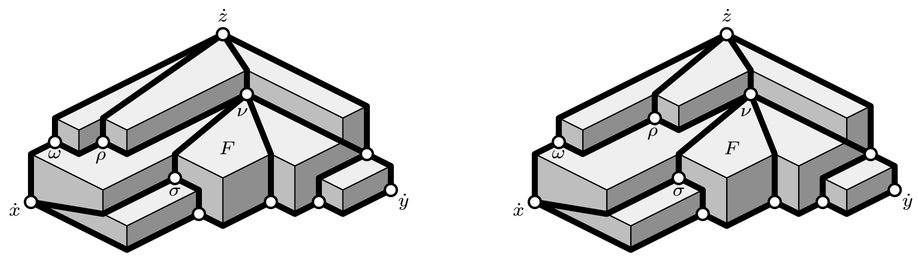
Lemma 2.2 *Suppose $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ is an axial geodesic embedding, and $\nu \in \mathcal{V}$ borders a region contained in $[\dot{x}, \omega, \dot{y}]$. Then $\nu_z \leq \omega_z$, with strict inequality if $\nu \notin [\omega, \dot{x}] \cup [\omega, \dot{y}]$. A similar statement holds for arbitrary permutations of x, y, z.*

Proof. The orthogonal flow $[\nu, \dot{z}]$ must cross $[\omega, \dot{x}]$ or $[\omega, \dot{y}]$, at $\nu' \in [\omega, \dot{x}]$, say. Concatenating the part of $[\omega, \dot{x}]$ from ω to ν' with the part of $[\nu, \dot{z}]$ from ν' to ν yields a path from ω to ν that is weakly decreasing in z. This path is strictly decreasing if $\nu \notin [\omega, \dot{x}] \cup [\omega, \dot{y}]$, for then it traverses (downwards) the vertical orthogonal ray Z_ν. □

Proposition 2.3 (Region axiom) *Let $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ be an axial geodesic embedding, and F a bounded region of M. If α_F is the join of the vertices of F, then $\alpha_F \in \mathcal{S}_{\mathcal{V}}$, and every vertex $\nu \in F$ shares precisely one coordinate with α_F.*

Proof. If $\omega \in \mathcal{V}$, then Lemma 2.2 implies there is some $u \in \{x, y, z\}$ such that $\nu_u \leq \omega_u$ for all $\nu \in F$. This shows ω cannot strongly precede α_F, so $\alpha_F \in \mathcal{S}_{\mathcal{V}}$; every vertex $\nu \in F$ therefore shares at least one coordinate with α_F. Suppose by symmetry that $\nu_z = (\alpha_F)_z$. The two edges of F containing ν cannot increase in z, so they exit ν counterclockwise of X_ν and clockwise of Y_ν. At least one of these edges strictly increases in x, and another strictly increases in y, completing the proof. □

The next proposition says that regions in rigid axial embeddings are 2-dimensional analogues of rigid geodesics. In addition to its applications throughout Sections 3 and 4, this fact will play a crucial role in Corollary 5.2, by way of Corollary 2.5, below. Refer to Figure 2 for an illustration of the rigid region axiom as well as its failure for nonrigid embeddings; ν and σ are vertices of an elbow geodesic, but $\rho \preceq \nu \vee \sigma$ in the latter case. The diagram is labeled as in the coming proof of Proposition 2.4.



rigid embedding nonrigid embedding

Figure 2: The rigid region axiom and its failure for nonrigid embeddings

Proposition 2.4 (Rigid region axiom) *Let $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ be a rigid axial embedding, and F a region of $M_\infty(\dot x, \dot y, \dot z)$. If α_F is the join of the vertices of F and $\omega \in \mathcal{S}_{\mathcal{V}}$, then $\omega \in F \Leftrightarrow \omega \preceq \alpha_F$.*

Proof. The claim is obvious if F is one of the three unbounded regions, so assume F is bounded. Since $\omega \in F \Rightarrow \omega \preceq \alpha_F$ by definition, let $\omega \notin F$, and assume F is contained in $[\dot x, \omega, \dot y]$ by symmetry. Either $\omega_z > \nu_z$ for some vertex $\nu \in F$ with maximal z-coordinate $\nu_z = (\alpha_F)_z$, in which case the proof is trivial, or all vertices of F with z-coordinate $(\alpha_F)_z$ lie on $[\omega, \dot x] \cup [\omega, \dot y]$, by Lemma 2.2. Assume there is one on $[\omega, \dot y]$ by transposing x and y if necessary, and let $\nu \in F \cap [\omega, \dot y]$ be closest to ω.
The vertex $\rho \in [\omega, \dot y]$ pointing toward ν has the same z-coordinate as ν (because $\omega_z \geq \rho_z \geq \nu_z$ and $\omega_z = \nu_z$), so the rigid geodesic $[\rho, \nu]$ consists of the orthogonal rays Y_ρ and X_ν. Of ν's two neighbors in F, let σ have smaller y-coordinate. Since $\rho \neq \sigma$ and $\sigma \notin [\omega, \dot y]$, the edge connecting ν to σ exits ν strictly counterclockwise of X_ν and strictly clockwise of Y_ν. Thus $\sigma_z < \nu_z$, whence $\sigma_x = (\alpha_F)_x$ by the region axiom (Proposition 2.3). But $\rho \not\preceq \nu \vee \sigma$ by the rigid geodesic axiom, while $\rho_z = \nu_z = (\nu \vee \sigma)_z$ and $\rho_y < \sigma_y = (\nu \vee \sigma)_y$ by construction. Therefore $\rho_x > (\nu \vee \sigma)_x = \sigma_x = (\alpha_F)_x$. The proof is complete because ω decreases in x along $[\omega, \dot y]$ to ρ. □

Corollary 2.5 *Let $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ be an axial rigid embedding. If P is a vertex, edge, or bounded region of M, let α_P denote join of the vertices in P. The map sending $P \mapsto \alpha_P$ constitutes an embedding in $\mathbb{N}^3$ of the vertex-edge-face poset of M.*

Proof. Immediate from the vertex, rigid geodesic, and rigid region axioms. □

Recall that the **order dimension** of a partially ordered set (poset) $\mathcal{P}$ is the smallest $d \in \mathbb{N}$ such that $\mathcal{P}$ includes into the poset $\mathbb{R}^d$. The previous corollary says that the order dimension of an axial rigidly embedded planar map is no greater than 3. See [Tro92] for more on order dimension.

3 Gluing geodesic embeddings

Let M be a planar map with axial vertices $\dot{x}, \dot{y}, \dot{z}$ and extended map $M_\infty = M_\infty(\dot{x}, \dot{y}, \dot{z})$. Suppose C is a simple cycle in M having three counterclockwise ordered vertices $\ddot{x}, \ddot{y}, \ddot{z}$, and furthermore that C bounds a closed disk $R \subset S$ that is a union of bounded regions in M. Following Brightwell and Trotter (cf. [Tro92, Chapter 6], although our definition differs slightly), we call C a **ring** if every edge of M not contained in R intersects R in a (possibly empty) subset of $\{\ddot{x}, \ddot{y}, \ddot{z}\}$. The double-dotted vertices play the roles of axial vertices for a smaller map $N = M \cap R$ "glued into" M by external edges emanating from N at $\ddot{x}, \ddot{y}, \ddot{z}$. Although we allow $\ddot{u}$ for $u \in \{x, y, z\}$ to equal the original axial vertex $\dot{u} \in M$, we exclude the case where C is the exterior cycle of M by referring to a **proper ring**.
Assume for each $u = x, y, z$ that the vertex $\ddot{u}$ meets at least one edge in M_∞ not contained in R (this occurs when M_∞ is triconnected). If there are at least two such edges then set $\underline{\ddot{u}} = \ddot{u}$. Otherwise, call the unique edge $e_{\ddot{u}}$ and name its other endpoint $\underline{\ddot{u}}$. Here, $\underline{\ddot{u}} = \infty$ is allowed because $\ddot{u} = \dot{u}$ is; but if C is a *proper* ring, then at most one of $\underline{\ddot{x}}, \underline{\ddot{y}}, \underline{\ddot{z}}$ can equal ∞, because there are no proper rings containing two axial vertices $\dot{u} \in \{\dot{x}, \dot{y}, \dot{z}\}$ such that R contains all of their edges in M. Indeed, it would be impossible to choose the third vertex $\ddot{u}$ from the pair of last points on the exterior cycle of M going from the two axial vertices on C toward the third.
The closure in the 2-sphere of the subset $\overline{M}_\infty \smallsetminus R$ is a planar map whose intersection with R equals $\{\ddot{x}, \ddot{y}, \ddot{z}\}$. Construct the **contraction** $\overline{M}_\infty/R$ by leaving off the edges $\{e_{\ddot{u}} \mid \underline{\ddot{u}} \neq \ddot{u}\}$ as well as their endpoints $\ddot{u}$ on C, and then connecting $\underline{\ddot{x}}, \underline{\ddot{y}}, \underline{\ddot{z}}$ to a new vertex τ inside R. View $M_\infty/R := (\overline{M}_\infty/R) \smallsetminus \infty$ as being the extension of a map $M/R = \mathrm{del}(\overline{M}_\infty/R; \infty)$. Thus $M_\infty/R = (M/R)_\infty$ has τ as a vertex, and still has axes drawn to ∞, although τ might replace one of $\dot{x}, \dot{y}, \dot{z}$ as an axial vertex. When τ replaces $\dot{u}$, however, we are free to choose $\tau = \dot{u}$, so we still write $(M/R)_\infty(\dot{x}, \dot{y}, \dot{z})$.

Lemma 3.1 *Let $M_\infty(\dot{x}, \dot{y}, \dot{z})$ be triconnected and M contain a proper ring C as above. Then both $M \cap R$ and M/R are planar maps, with axial vertices, whose extended maps $(M \cap R)_\infty(\ddot{x}, \ddot{y}, \ddot{z})$ and $(M/R)_\infty(\dot{x}, \dot{y}, \dot{z})$ are triconnected. Each of $M \cap R$ and M/R contains fewer edges and strictly fewer regions than M.*

Proof. Deleting from M_∞ any pair of vertices in M leaves every remaining vertex $\nu \in M \cap R$ connected to $\{\ddot{x}, \ddot{y}, \ddot{z}\}$, because every path connecting ν to ∞ in M_∞ passes through $\{\ddot{x}, \ddot{y}, \ddot{z}\}$. By the same argument, every vertex in M_∞ that remains after deleting any pair of vertices in M is connected to R—and hence to τ—in the deletion. (The removal of the edges $e_{\ddot{u}}$ ensures that M/R has no bivalent vertices on the way to τ.) The fact that $\dot{x}, \dot{y}, \dot{z}$ and $\ddot{x}, \ddot{y}, \ddot{z}$ can be chosen as axial vertices follows from the triconnectivity of the extended maps of M/R and $M \cap R$.
Now M has at least one region inside (resp. outside) R, because C is a simple cycle (resp. a proper ring). Thus M/R (resp. $M \cap R$) has strictly fewer regions

than M. The edge number inequality is obvious for $M \cap R$. For M/R, the number of edges is at most $E+3$, counting the edges to τ, where E is the number of edges in $M \smallsetminus R$. But the number of edges in M is at least $E+3$, because R contains the cycle C. $\square$

Let M and N be planar maps with axial vertices $(\dot{x}, \dot{y}, \dot{z})$ and $(\ddot{x}, \ddot{y}, \ddot{z})$, respectively, such that $M_\infty(\dot{x}, \dot{y}, \dot{z})$ and $N_\infty(\ddot{x}, \ddot{y}, \ddot{z})$ are both triconnected. We now show how to **glue** N **into** M at a vertex $\tau \in M$ that is trivalent in M_∞. Let the counterclockwise ordered neighbors of τ be α, β, γ (one of which might be ∞) in M_∞. (Think of M and N as M/R and $M \cap R$ from Lemma 3.1, respectively.) Start by replacing τ with a small triangle in M_∞ (a 'Y–Δ' transformation), adding three new vertices in the process. This action requires working in M_∞ rather than M if $\tau \in \{\dot{x}, \dot{y}, \dot{z}\}$. Next, replace the new triangle and its interior with N, in such a way that $\alpha, \beta, \gamma \in M$ connect to the axial vertices $\ddot{x}, \ddot{y}, \ddot{z} \in N$ via edges $e_{\ddot{x}}, e_{\ddot{y}}, e_{\ddot{z}}$, called **tethers** in M_∞. The result is an extended map for $M \cup_\tau N$, the **tethered gluing** of N into M at τ. Contracting some or all of the tethers yields a **gluing** of N into M, provided the resulting map is simple and triconnected.

The construction of the tethered gluing works at the level of grid surfaces. For instance, the hypotheses in the next lemma can easily be attained by scaling M. This is a key observation, making the induction in the proof of Theorem 5.1 possible. The left columns of Figures 5 and 4 illustrate examples of $\mathcal{S}_{\mathcal{V}_M}, \mathcal{S}_{\mathcal{V}_N}, \tau$, and $M \cup_\tau N \hookrightarrow \mathcal{S}_\mathcal{V}$.

Lemma 3.2 *Let $M \hookrightarrow \mathcal{S}_{\mathcal{V}_M}$ and $N \hookrightarrow \mathcal{S}_{\mathcal{V}_N}$ be rigid embeddings with respective axial vertices $\dot{x}, \dot{y}, \dot{z}$ and $\ddot{x}, \ddot{y}, \ddot{z}$, and suppose $\tau \in M_\infty(\dot{x}, \dot{y}, \dot{z})$ is trivalent. If U_τ has length at least $m+1$ for $U = X, Y, Z$, then τ is the unique vector in $\mathcal{V}_M$ preceding $\tau + m\mathbf{1}$, where $\mathbf{1} = (1,1,1)$. If, in addition, $|\ddot{u}| \leq m$ for $u = x, y, z$ and*

$$\mathcal{V} \quad = \quad (\mathcal{V}_M \smallsetminus \tau) \cup (\tau + \mathcal{V}_N)$$

then the rigid geodesics in $\mathcal{S}_\mathcal{V}$ provide a rigid embedding of the map $M \cup_\tau N$.

Proof. The orthogonal rays X_τ, Y_τ, Z_τ point toward α, β, γ (one of these may be ∞) in $\mathcal{S}_{\mathcal{V}_M}$ because τ is trivalent. Each vertex $\nu \in \mathcal{V}_M$ with $\nu \neq \tau$ has $\nu_x \geq \alpha_x$, $\nu_y \geq \beta_y$, or $\nu_z \geq \gamma_z$. Indeed, if ν lies in $[\dot{x}, \tau, \dot{y}]$ (say), then considering where the orthogonal flow $[\nu, \dot{z}]$ intersects $[\tau, \dot{x}] \cup [\tau, \dot{y}]$ shows that either $\nu_x \geq \alpha_x$ or $\nu_y \geq \beta_y$. Thus $\tau \preceq \tau + m\mathbf{1}$ is unique in $\mathcal{S}_{\mathcal{V}_M}$; the vertex axiom for $\mathcal{V}$ is immediate.

The part of $\mathcal{S}_\mathcal{V}$ preceding $\tau + m\mathbf{1}$ equals $\tau +$ (the part of $\mathcal{S}_{\mathcal{V}_N}$ preceding $m\mathbf{1}$), by the uniqueness in M of $\tau \preceq \tau + m\mathbf{1}$. Thus every vertex, rigid geodesic, or bounded orthogonal ray in $N \hookrightarrow \mathcal{S}_{\mathcal{V}_N}$ gets translated by τ to the corresponding feature in $\mathcal{S}_\mathcal{V}$. Similarly, the parts of $\mathcal{S}_\mathcal{V}$ and $\mathcal{S}_{\mathcal{V}_M}$ not preceded by τ agree, so any vertex, rigid geodesic, or bounded orthogonal ray in $M \hookrightarrow \mathcal{S}_{\mathcal{V}_M}$ survives in $\mathcal{S}_\mathcal{V}$, as long as it is contained in a (perhaps unbounded) region of M_∞ not containing τ, by the rigid region axiom.

The only orthogonal rays unaccounted for as yet for the rigid geodesic axiom are those leaving $\tau+\ddot{u}$ and α,β,γ. Observe that $\tau+\ddot{u}$ is the unique element of $\mathcal{V}$ on the orthogonal ray $U_\tau \subset \mathcal{S}_{\mathcal{V}_M}$. Thus an orthogonal ray leaving $\tau+\ddot{u}$ either points toward the corresponding one of α,β,γ whenever the latter is not ∞, or it points away from $\mathcal{V}_M$. An orthogonal ray leaving α,β,γ either points back toward $\ddot{u}$ along a rigid geodesic $e_{\ddot{u}}$, or it points away from $\tau+\mathcal{V}_N$. We conclude that the rigid geodesics in $\mathcal{S}_\mathcal{V}$ form a planar map isomorphic to $M \cup_\tau N$. □

4 Contracting rigid geodesics

Proposition 4.1 *Let M be a planar map with axial vertices $\dot{x},\dot{y},\dot{z}$. Suppose e is the edge in the exterior cycle of M leaving $\dot{x}$ toward $\dot{y}$, and that e borders no triangles in $M_\infty(\dot{x},\dot{y},\dot{z})$. If M/e can be rigidly embedded in some grid surface, then so can M.*

Proof. Letting $\nu \in M$ be the other endpoint of e, we have $\nu \neq \dot{y}$ because the unbounded region of M_∞ containing e is not a triangle. The edge in M leaving ν clockwise from e determines an edge f in any rigid embedding $N \hookrightarrow \mathcal{S}_\mathcal{V}$ isomorphic to M/e. Note that $f \in N$ does not contain the orthogonal ray $Y_{\dot{x}}$ because $f \neq e$; and $f \not\supset Z_{\dot{x}}$ because ν sits between $\dot{x}$ and $\dot{y}$. Therefore the orthogonal ray X_ω at the other endpoint ω of f in N points toward $\dot{x}$.
Assume all coordinates of vectors in $\mathcal{V}$ are even, by scaling. The claim is that the rigid geodesics in $\mathcal{S}_{\mathcal{V}\cup\nu}$ constitute a rigid embedding isomorphic to M, where the coordinates of ν are defined by

$$\nu \quad = \quad (\nu_x,\nu_y,\nu_z) \quad = \quad (|\dot{x}|-1,\omega_y,0).$$

The addition of ν to $\mathcal{V}$ affects at most the rigid geodesics in N containing one of the following: an orthogonal ray X_σ for some vertex $\sigma \in \mathcal{V}$ pointing toward $\dot{x}$; an orthogonal ray at ν; or $Y_{\dot{x}}$. All other rigid geodesics lie behind the plane $x = |\dot{x}|-1$.
If $\sigma_y < \omega_y$, then X_σ is unaffected by ν, while if $\sigma_y \geq \omega_y$, then $\dot{x}$ and ν are the only elements of $\mathcal{V}$ preceding $\sigma\vee\dot{x} = \sigma\vee\nu + (1,0,0)$. Thus X_σ points toward ν if $\sigma_y \geq \omega_y$, because $\dot{x} \not\preceq \sigma\vee\nu$. The three orthogonal rays X_ν, Z_ν, and Y_ν leaving ν point respectively toward $\dot{x}$, ω, and the vertex to which $Y_{\dot{x}} \subset \mathcal{S}_\mathcal{V}$ points in N. Finally, $Y_{\dot{x}} \subset \mathcal{S}_{\mathcal{V}\cup\nu}$ points toward ν. (Figure 3 illustrates the transition $M/e \rightsquigarrow M$.) □

In the situation of Lemma 3.1, gluing $M\cap R$ into M/R may involve contracting some of the tethers in $(M/R) \cup_\tau (M \cap R)$. For the rest of this section, let $M \hookrightarrow \mathcal{S}_{\mathcal{V}_M}$ and $N \hookrightarrow \mathcal{S}_{\mathcal{V}_N}$ be rigid embeddings having respective axial vertices $\dot{x},\dot{y},\dot{z}$ and $\ddot{x},\ddot{y},\ddot{z}$, with $\tau \in M_\infty(\dot{x},\dot{y},\dot{z})$ a trivalent vertex having neighbors α,β,γ. Let B be the region of $M \cup_\tau N$ containing $e_{\ddot{y}}$ and $e_{\ddot{z}}$.

Proposition 4.2 *Assume that $\tau \neq \dot{y}$, that $Y_{\dot{x}}$ points toward τ, and that no edge in M has vertices $\{\dot{x},\gamma\}$. Contracting $e_{\ddot{x}}$ along with neither, either, or*

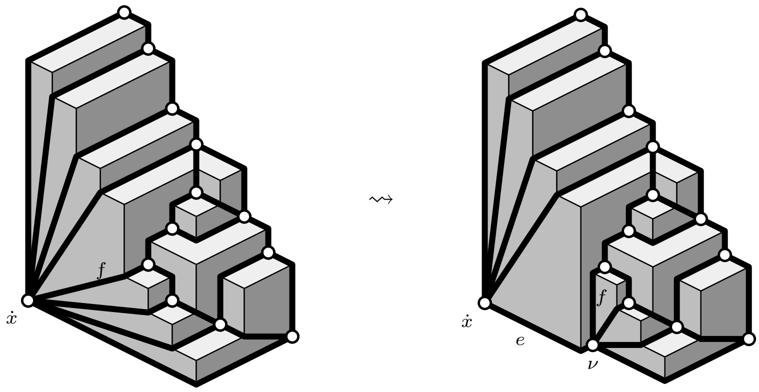


Figure 3: Uncontracting the lower-left edge

(if B has at least five vertices) both of $e_{\ddot{z}}$ and $e_{\ddot{y}}$ in $M \cup_\tau N$ yields a planar map possessing a geodesic grid surface embedding.

Proof. Paragraph headings are included below to make parts of the proof easier to follow and cross-reference.

PLAN OF PROOF

Given the conditions of the present proposition, assume all hypotheses and notation of Lemma 3.2, as well; this is possible by Lemma 2.1. If $|\mathcal{V}_N| = n$, scale M so that all coordinates of vectors in $\mathcal{V}_M$ are divisible by $n+1$. Until further notice (see the special construction for (8), below), assume in addition that the orthogonal rays X_τ, Y_τ, Z_τ all have length exactly $m+1$. Set

$$\nu' \quad := \quad \tau + \nu \quad \text{ for } \quad \nu \in \mathcal{V}_N.$$

The plan of the proof is to split into a number of cases, each of which demands slightly different treatment. In every case, Lemma 2.1 allows a judicious choice of coordinates for vectors in $\mathcal{V}_N$. Most often, omitting one or more of the vertices $\{\ddot{x}', \ddot{y}', \ddot{z}', \beta\}$ from $\mathcal{V}$ leaves a set $\underline{\mathcal{V}}$ such that the desired contraction of $M \cup_\tau N$ geodesically embeds into $\mathcal{S}_{\underline{\mathcal{V}}}$; two of the cases require additional fiddling with the surviving vertices to get the desired grid surface.

EIGHT CASES

Let A and B be the bounded regions of $M \cup_\tau N$ containing $e_{\ddot{x}}$ and $e_{\ddot{y}}$, respectively. By the region axiom, we think of $A = (A_x, A_y, A_z)$ and $B = (B_x, B_y, B_z)$ as the vectors given by the joins of their vertices. Denote by $A(u)$ the set of

vertices of A having u-coordinate A_u. For instance, $A(x) = \{\dot{x}\} = \{\alpha\}$, and $\beta \in B(y)$.
Here is the list of constructions yielding $\mathcal{S}_{\underline{\mathcal{V}}}$. In each of (1)–(7), choose $\mathcal{V}_N$ so that every coordinate of every vector in $\mathcal{V}_N$ is at most n; we treat the last case (8) separately later, since it involves somewhat different choices. Construct $\underline{\mathcal{V}}$ from $\mathcal{V} = (\mathcal{V}_M \smallsetminus \tau) \cup (\tau + \mathcal{V}_N)$ by omitting the indicated vectors, and (in (5) and (8)) making the specified alterations.

(1) To contract only $e_{\ddot{x}}$: omit $\ddot{x}'$.
(2) To contract $e_{\ddot{x}}$ and $e_{\ddot{y}}$: omit $\ddot{x}', \ddot{y}'$.
To contract $e_{\ddot{x}}$ and $e_{\ddot{z}}$...
(3) if no edge in N has endpoints $\{\ddot{x}, \ddot{z}\}$: omit $\ddot{x}', \ddot{z}'$.
if $\{\ddot{x}, \ddot{z}\}$ are the endpoints of an edge in N...
(4) and $A_z > \gamma_z$: omit $\ddot{x}', \ddot{z}'$.
(5) and $A_z = \gamma_z$: omit $\ddot{x}', \ddot{z}'$; then add 1 to ν_z for all $\nu \in A(z) \smallsetminus \gamma$.
To contract $e_{\ddot{x}}$, $e_{\ddot{y}}$, and $e_{\ddot{z}}$...
(6) if no edge in N has endpoints $\{\ddot{y}, \ddot{z}\}$: omit $\ddot{y}'$ after (3)–(5).
if $\{\ddot{y}, \ddot{z}\}$ are the endpoints of an edge in N...
(7) and $\beta_x \geq \gamma_x$: omit β after (3)–(5).
(8) and $\beta_x < \gamma_x$: make the special construction below.

In general, observe that the only vertices connected to $\ddot{x}$, $\ddot{y}$, or $\ddot{z}$ in $M \cup_\tau N$ are α, β, γ, and some vertices in N. Also, one of (6)–(8) must occur if B has at least 5 vertices. Representative instances of the cases (1)–(8) appear in Figures 4 and 5.

Omitting vertices

In general, omitting one or more elements from $\mathcal{V}$ always leaves a set of pairwise incomparable vectors. The vertex axiom will follow immediately in the applications below, in the sense that the surviving vertex vectors $\underline{\mathcal{V}}$ are in obvious bijection with the vertices of the desired map. To check the rigid geodesic axiom after omitting one vertex ν, we must verify that any orthogonal ray U pointing toward ν before the omission of ν points to some other uniquely determined surviving vertex afterwards. This will show that the surviving vertices $\underline{\mathcal{V}}$ define a rigid embedding $L \hookrightarrow \mathcal{S}_{\underline{\mathcal{V}}}$ for some map L.
Specifically, L is obtained from $M \cup_\tau N$ by first deleting ν along with all of its incident edges, and then reconnecting the neighbors of ν to other surviving vertices according to where their orthogonal rays point. It is important to remember that some edges incident to ν may fail to reappear upon reconnecting its neighbors: these non-reappearing edges are precisely those rigid geodesics that, before the omission of ν, do not contain any orthogonal ray pointing toward ν. (There are at most three such edges, because each one must contain an orthogonal ray leaving ν.)

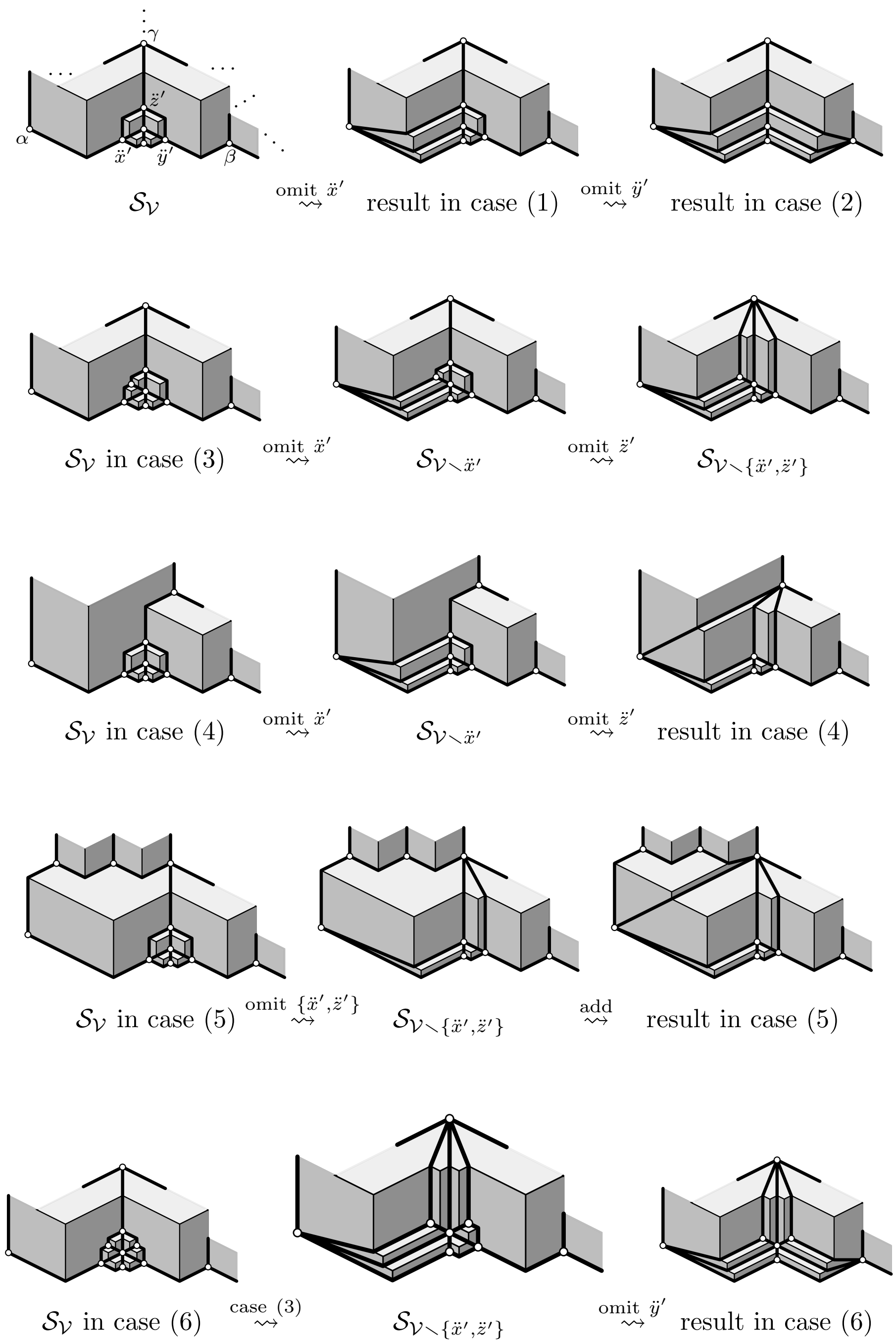


Figure 4: Gluing grid surfaces: cases (1)–(6)

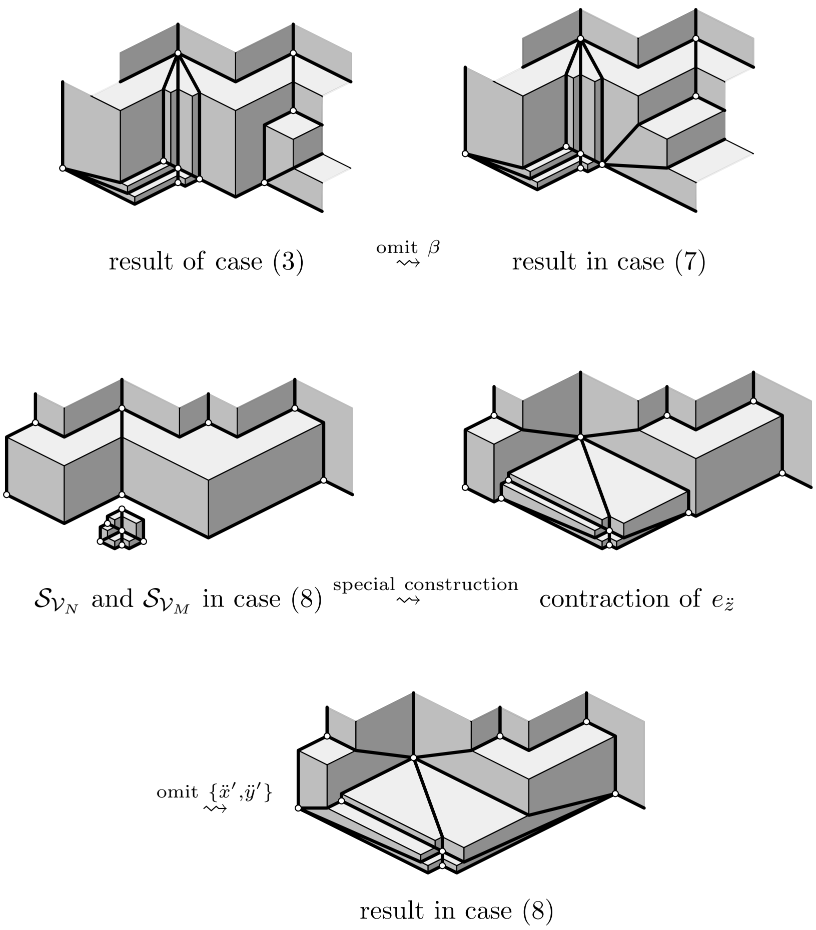


Figure 5: Gluing grid surfaces: cases (7)–(8)

Scale principle

We chose the relative sizes of $\mathcal{V}_N$ and $\mathcal{V}_M$ so that some arguments in what follows can rely on the following principle: If ω, σ are two vectors whose coordinates are all divisible by $n+1$, then $\omega \preceq \sigma$ if and only if $\omega \preceq \sigma + n\mathbf{1}$.

Rays leaving $\tau + \mathcal{V}_N$

An orthogonal ray leaving $\nu' \in \tau + \mathcal{V}_N$ and pointing toward $\ddot{z}'$ in $\mathcal{S}_{\mathcal{V}}$ must be $Z_{\nu'} \subset \mathcal{S}_{\mathcal{V}}$. We claim that after omitting $\ddot{z}'$, the ray $Z_{\nu'}$ points toward γ, regardless of whether or not one or both of $\ddot{x}', \ddot{y}'$ has already been omitted. Moreover, this statement remains valid after permuting the roles of x, y, z.
To see why, suppose that $\omega \preceq \nu' \vee \gamma$ for $\omega \in \mathcal{V}$. We must show that

$\omega \in \{\nu', \ddot{z}', \gamma\}$. If $\omega \in \tau + \mathcal{V}_N$ then $\omega_z \leq \tau_z + |\ddot{z}|$ and thus $\omega \preceq \nu' \vee \ddot{z}'$, so $\omega \in \{\nu', \ddot{z}'\}$ because $[\nu', \ddot{z}']$ is a rigid geodesic (Lemma 3.2). When $\omega \in \mathcal{V}_M \smallsetminus \tau$, apply the scale principle with $\sigma = \gamma \vee \ddot{z}'$, using the fact that $\gamma \vee \nu' \preceq \gamma \vee \ddot{z}' + n\mathbf{1}$ because $\tau \preceq \ddot{z}', \nu' \preceq \tau + n\mathbf{1}$. The argument is invariant under permutation of x, y, z.

Proof of (1)

The rigid geodesics containing orthogonal rays $X_{\nu'}$ pointing toward $\ddot{x}'$ before the omission and toward α afterwards account for all of the necessary edges in the contraction. It remains only to verify that $Y_\alpha \subset \mathcal{S}_{\mathcal{V} \smallsetminus \ddot{x}'}$ points toward the next vertex after $\ddot{x}'$ whose z-coordinate is zero—that is, the counterclockwise next vertex after $\ddot{x}'$ on the exterior cycle of $M \cup_\tau N$. This easy argument is left to the reader, completing the proof of (1). Until the proof of (8), assume $\ddot{x}'$ has been omitted.

Proof of (2)

The orthogonal rays pointing toward $\ddot{y}'$ in $\mathcal{S}_{\mathcal{V} \smallsetminus \ddot{x}'}$ are $Y_{\nu'}$ for some vertices $\nu \in \mathcal{V}_N$, the ray X_β, and possibly Y_α (if $\ddot{x}$ and $\ddot{y}$ are the vertices of an edge in N). The arguments in 'Rays in $\tau + \mathcal{V}_N$' and 'Proof of (1)' apply as well to the omission of $\ddot{y}'$, including the fact that the ray X_β points toward the next vertex after $\ddot{y}'$ whose z-coordinate is zero (which may be α). None of the edges incident to $\ddot{y}'$ vanish (see 'omitting vertices'), although X_β and the Y-orthogonal ray at height $z = 0$ pointing toward $\ddot{y}'$ point toward each other after omitting $\ddot{y}'$, effectively contracting $e_{\ddot{y}}$.

Cases (3)–(5)

Every rigid geodesic incident to $\ddot{z}'$ in $\mathcal{S}_{\mathcal{V} \smallsetminus \ddot{x}'}$ contains an orthogonal ray pointing toward $\ddot{z}'$, except the geodesic connecting α to $\ddot{z}'$, if there is one (this occurs in (4) and (5) only, where $X_{\ddot{z}'}$ points toward α in $\mathcal{S}_{\mathcal{V} \smallsetminus \ddot{x}'}$), and sometimes the geodesic connecting γ to $\ddot{z}'$.
After omitting $\ddot{z}'$, we will verify the rigid geodesic axiom at γ separately for each of (3)–(5), by checking *only* that an orthogonal ray leaving γ and pointing toward $\ddot{z}'$ (if there is one) points instead to another surviving vertex after omitting $\ddot{z}'$. Any other orthogonal ray in $\mathcal{S}_{\mathcal{V} \smallsetminus \ddot{x}'}$ either leaves some vector in $\tau + \mathcal{V}_N$ before omitting $\ddot{z}'$ (these have been dealt with in a paragraph above), or points toward another vertex in $\mathcal{V}_M \smallsetminus \tau$ both before and after omitting $\ddot{z}'$.
Verifying the rigid geodesic axiom will complete the proof unless the edges connecting α and γ to $\ddot{z}'$ (if these exist) *both* vanish upon omitting $\ddot{z}'$, for otherwise all of the edges in the contraction of $e_{\ddot{z}}$ are accounted for as in the proofs of (1) and (2). Both geodesics vanish in (5) only: in (3), there is no edge connecting α to $\ddot{z}'$; while in (4), the region axiom implies $\gamma_y = A_y$, whence X_γ points toward $\ddot{z}'$. In (5), we show that the addition procedure reconstructs a rigid geodesic connecting α to γ.

Proof of (3)

If X_γ points toward $\ddot{z}'$ in $\mathcal{S}_{\mathcal{V}\smallsetminus\ddot{x}'}$, and ν' is the vertex to which $X_{\ddot{z}'}$ points in $\mathcal{S}_{\mathcal{V}\smallsetminus\ddot{x}'}$, then X_γ points toward $\nu' \in \tau + \mathcal{V}_N$ after omitting $\ddot{z}'$ by the scale principle (the hypothesis for (3) guarantees that $\nu' \neq \ddot{x}'$). A similar statement holds by switching the roles of x and y (but $\nu' \in \tau + \mathcal{V}_N$ is always guaranteed to exist, since $\ddot{y}'$ has not been omitted). This verifies the rigid geodesic axiom at γ.

Proof of (4)

The assumption $A_z > \gamma_z$ implies $A_y = \gamma_y$, so that X_γ points toward $\ddot{z}'$ in $\mathcal{S}_{\mathcal{V}\smallsetminus\ddot{x}'}$. If $\nu \in \mathcal{V} \smallsetminus \dot{x}$ and $\nu_y < \gamma_y$, then $\nu_z > \gamma_z$, by the region axiom. The omission of $\ddot{z}'$ therefore causes the ray $X_\gamma \subset \mathcal{S}_{\mathcal{V}\smallsetminus\{\ddot{x}',\ddot{z}'\}}$ to point toward α.

Proof of (5)

The region axiom implies $\gamma_y < \tau_y$, so X_γ does not point toward $\ddot{z}'$. When Y_γ points toward $\ddot{z}'$, the rigid geodesic axiom holds for $\mathcal{S}_{\mathcal{V}\smallsetminus\{\ddot{x}',\ddot{z}'\}}$ by the argument in the proof of (3), although no elbow geodesic connects α to γ in $\mathcal{S}_{\mathcal{V}\smallsetminus\{\ddot{x}',\ddot{z}'\}}$. Now we verify that the addition procedure outputs a rigid geodesic embedding, and that an edge connecting α to γ is the only new rigid geodesic. We can safely ignore all orthogonal rays contained in rigid geodesics on the positive side of the plane $y = \gamma_y$. All of the vertices $\omega \notin A$ satisfying $\omega_y \leq \gamma_y$ must also satisfy $\omega_z > \gamma_z$ by the rigid region axiom. The adding rule therefore causes X_γ to point toward α after omitting $\ddot{z}'$. The orthogonal rays leaving vertices originally in $A(z) \smallsetminus \gamma$ still point to the same vertices, by the scale principle. If $\omega \notin A(z) \cup \{\alpha\}$, then any orthogonal ray U_ω pointing toward $\nu \in A(z)$ before the addition still points toward the same vertex ν afterwards, because $\omega_z > \nu_z$, whence the join $\nu \vee \omega$ remains unaffected. Finally, if Z_α points toward a vertex in $A(z)$ before the addition procedure, then Z_α still points to the same vertex afterwards, by the scale principle, while Y_α remains unaffected.

Proof of (6)

No new phenomena occur here; see Proof of (3).

Technical lemma

The following result will be applied in the proofs of (7) and (8). For the proof of (8), note that it holds after any rescaling of $\mathcal{V}_M$ as in Lemma 2.1.

> Let $\omega \in \mathcal{V}_M$. If $\tau_x \geq \omega_x \geq \beta_x$, then $\omega \in \{\tau, \beta\}$ or $\omega_z \geq \gamma_z$. Similarly, if $\tau_y \geq \omega_y \geq \alpha_y = 0$, then $\omega \in \{\tau, \alpha\}$ or $\omega_z \geq \gamma_z$.

Proof of technical lemma. Suppose $\tau_x \geq \omega_x \geq \beta_x$, but $\omega_z < \gamma_z$. If $\omega \neq \tau$, then $\omega_y \geq \beta_y$ by the uniqueness in Lemma 3.2 of τ among vectors preceding $\tau + m\mathbf{1}$

in $\mathcal{V}_M$. Thus $\beta \preceq \omega$, so $\beta = \omega$. Swap the roles of x and y to prove the other statement.

Proof of (7)

First claim: The only orthogonal rays pointing toward β are $Y_{\ddot{y}'}$, and X_ν for some vertices $\nu \in M$. These vertices ν all have $\nu_z < m+1$.
The final sentence of the first claim is easy, because otherwise $\gamma \preceq \beta \vee \nu$. For the rest of the claim, use the inequalities $B_x = \tau_x > \beta_x \geq \gamma_x$ and $B_y = \beta_y > \tau_y \geq \gamma_y$, which follow from the hypotheses of (7). These imply $\gamma_z = m+1 = B_z$, thanks to the region axiom. By the technical lemma, any orthogonal ray Y_ν pointing toward β in $\mathcal{S}_\mathcal{V}$ (and therefore in $\mathcal{S}_{\mathcal{V} \smallsetminus \{\ddot{x}', \ddot{z}'\}}$) must have either $\nu_x \geq \tau_x$ or $\nu_z \geq m+1$. When $\nu_x \geq \tau_x$, we get $\nu_y < \beta_y$ and hence $\ddot{y}' \preceq \nu \vee \beta$, so $\nu = \ddot{y}'$. The case $\nu_z \geq m+1$ is actually impossible, for it implies $\gamma \preceq \nu \vee \beta$, so $\nu = \gamma$ is connected to β by an edge in $M \cup_\tau N$; this cannot happen in (7) if B has at least five vertices. The fact that $\beta_z = 0$ rules out Z_ν pointing toward β, completing the proof of the first claim.
Now we verify that $X_\nu \subset \mathcal{S}_{\mathcal{V} \smallsetminus \{\ddot{x}', \beta, \ddot{z}'\}}$ points toward $\ddot{y}'$ whenever $X_\nu \subset \mathcal{S}_{\mathcal{V} \smallsetminus \{\ddot{x}', \ddot{z}'\}}$ points toward β. In other words, we need $\omega \preceq \nu \vee \ddot{y}'$ for $\omega \in \mathcal{V}$ to imply $\omega \in \{\ddot{x}', \ddot{z}', \ddot{y}', \beta, \nu\}$. If $\omega \in \tau + \mathcal{V}_N$, then $\omega_x = \ddot{y}'_x = \tau_x$ implies $\omega \in \{\ddot{y}', \ddot{z}'\}$. If $\omega \in \mathcal{V}_M \smallsetminus \tau$, then either $\omega_x \leq \beta_x$, in which case $\omega \preceq \beta \vee \nu$ implies $\omega \in \{\beta, \nu\}$, or $\omega_z < m+1$ in addition to $\tau_x \geq \omega_x \geq \beta_x$, in which case $\omega = \beta$ by the technical lemma.
It is easy to verify that $\ddot{y}'$ points toward the next vertex after β having z-coordinate zero. Note that such a next vertex must exist, since the rigid geodesic leaving β and containing Z_β strictly decreases in x. This completes the proof of (7).

Special construction for (8)

The meet (componentwise minimum) of τ and γ is $\tau \wedge \gamma = (\gamma_x, \gamma_y, 0) = \gamma - (0, 0, m+1)$ since Z_τ points toward γ in $\mathcal{S}_{\mathcal{V}_M}$. Observe that

$$\omega \in \mathcal{V}_M \text{ and } \tau \wedge \gamma \preceq \omega \;\text{ implies }\; \omega \in \{\tau, \gamma\}. \tag{9}$$

Indeed, if $\tau \not\preceq \omega$ but $\tau \wedge \gamma \preceq \omega$, then either $\tau_x \geq \omega_x \geq \gamma_x$ or $\tau_y \geq \omega_y \geq \gamma_y$. The hypothesis $\gamma_x > \beta_x$ of (8) plus the technical lemma imply $\omega_z \geq \gamma_z$, whence $\gamma \preceq \omega$.
It follows from (9) that the set

$$\underline{\mathcal{V}}_M \quad = \quad (\mathcal{V}_M \smallsetminus \{\tau, \gamma\}) \cup \tau \wedge \gamma$$

of vectors determines a grid surface $\mathcal{S}_{\underline{\mathcal{V}}_M}$. Use the freedom afforded by Lemma 2.1 to rechoose $\mathcal{V}_M$ and m so that $n+3$ divides all coordinates of vectors therein, while γ_z as well as the lengths of the orthogonal rays $X_{\tau \wedge \gamma}$ and $Y_{\tau \wedge \gamma} \subset \mathcal{S}_{\underline{\mathcal{V}}_M}$ equal $m+2$. Applying (9) again, further alter $\mathcal{V}_M$ by moving

τ so that $\tau - \tau\wedge\gamma$ equals $(1,0,0)$, $(0,1,0)$, or $(1,1,0)$, depending on whether $\tau_y = \gamma_y$, $\tau_x = \gamma_x$, or neither.
Now choose $\mathcal{V}_N$ so that all of its *nonzero* x and y-coordinates lie in the interval $[m-n, m]$, but all of its z-coordinates are no greater than n. Let

$$\underline{\mathcal{V}} \quad = \quad (\underline{\mathcal{V}}_M \smallsetminus \tau\wedge\gamma) \cup (\tau\wedge\gamma + \mathcal{V}_N),$$

and denote by $\underline{\nu}$ the vector $\tau\wedge\gamma+\nu$ for $\nu \in \mathcal{V}_N$. Our (final) goal is to show that the rigid geodesics in $\mathcal{S}_{\underline{\mathcal{V}}}$ constitute an embedding of $(M \cup_\tau N)/e_{\ddot{z}}$. After that, $e_{\ddot{x}}$ and $e_{\ddot{y}}$ can be contracted by omitting $\underline{\ddot{x}}$ and $\underline{\ddot{y}}$, using the same arguments appearing in Scale principle, Rays leaving $\tau+\mathcal{V}_N$, Proof of (1), and Proof of (2).

Proof of (8)

Begin by mimicking as closely as possible the proof of Lemma 3.2. First, $\tau\wedge\gamma$ is the unique vector in $\underline{\mathcal{V}}_M$ preceding $\tau\wedge\gamma + m\mathbf{1}$, by Lemma 3.2 applied to $\mathcal{V}_M$; this is why τ needs to be so close to $\tau\wedge\gamma$. The vertex axiom for $\mathcal{S}_{\underline{\mathcal{V}}}$ is immediate. Moreover, the part of $\mathcal{S}_{\underline{\mathcal{V}}}$ that precedes $\tau\wedge\gamma+m\mathbf{1}$ equals $\tau\wedge\gamma+$(the part of $\mathcal{S}_{\mathcal{V}_N}$ preceding $m\mathbf{1}$). Thus every vertex, rigid geodesic, or bounded orthogonal ray in $N \hookrightarrow \mathcal{S}_{\mathcal{V}_N}$ gets translated by $\tau\wedge\gamma$ to the corresponding feature in $\mathcal{S}_{\underline{\mathcal{V}}}$. Similarly, the parts of $\mathcal{S}_{\underline{\mathcal{V}}}$ and $\mathcal{S}_{\underline{\mathcal{V}}_M}$ not preceded by $\tau\wedge\gamma$ agree, so any vertex, rigid geodesic, or bounded orthogonal ray in $M \hookrightarrow \mathcal{S}_{\underline{\mathcal{V}}_M}$ survives in $\mathcal{S}_{\mathcal{V}}$ whenever it is contained in a (perhaps unbounded) region of M_∞ not containing τ or γ, by the rigid region axiom.
The only orthogonal rays unaccounted for as yet for the rigid geodesic axiom are $X_{\underline{\ddot{x}}}$, $Y_{\underline{\ddot{y}}}$, Y_α, X_β, $Z_{\underline{\ddot{z}}}$, and any orthogonal ray $U_\nu \subset \mathcal{S}_{\underline{\mathcal{V}}}$ such that $U_\nu \subset \mathcal{S}_{\mathcal{V}_M}$ points toward γ and $\nu \neq \tau$. (Neither the rigid geodesic connecting τ to γ in M nor the orthogonal rays leaving γ in $\mathcal{S}_{\mathcal{V}_M}$ play roles in this verification.) The only case requiring significant effort are the U_ν rays, which must point toward $\underline{\ddot{z}}$ in $\mathcal{S}_{\mathcal{V}}$.
Suppose $\omega \preceq \nu\vee\underline{\ddot{z}}$ for some $\omega \in \underline{\mathcal{V}}$. The technical lemma and (9) imply that $\nu_z \geq \gamma_z$ for any $\tau \neq \nu \in \mathcal{V}_M$ pointing toward γ, whence $\nu\vee\underline{\ddot{z}} = \nu\vee\gamma$ for any such ν. Therefore $\omega \neq \nu$ implies $\omega \in \tau\wedge\gamma + \mathcal{V}_N$. On the other hand, either $\nu_y < \beta_y$ or $\nu_x < \alpha_x$, because otherwise β or α precedes $\nu\vee\gamma = \nu\vee\underline{\ddot{z}}$. By the choice of scaling, $\omega_y \leq \nu_y < \beta_y$ forces $\omega_y \leq \nu_y \leq \beta_y - (n+3) < \gamma_y + m - n$, whence $\omega = \underline{\ddot{z}}$ whenever $\omega \in \tau\wedge\gamma + \mathcal{V}_N$. This argument also works with the roles of x and y switched.
The above reasoning proves that the rigid geodesics in $\mathcal{S}_{\underline{\mathcal{V}}}$ embed *some* planar map L. To conclude that $L \cong (M \cup_\tau N)/e_{\ddot{z}}$, one last item remains: show that no geodesics in M vanish. More precisely, whenever X_γ or Y_γ does not point toward τ in $\mathcal{V}_M$, we require it to point toward a vertex in $\mathcal{V}_M$ that points back toward γ in $\mathcal{S}_{\mathcal{V}_M}$. Suppose $Y_\gamma \subset \mathcal{S}_{\mathcal{V}_M}$ does not point toward τ. Then $\gamma_x < \tau_x = B_x$, whence $\gamma_z = B_z$ by the region axiom, because $\gamma_y < \beta_y = B_y$. If $B(z) = \{\gamma\}$, then $Y_\gamma \subset \mathcal{S}_{\mathcal{V}_M}$ must point toward β, because $\beta_x < \gamma_x$. This is impossible whenever the region B in $M \cup_\tau N$ has at least five vertices, by the hypothesis of (8) stipulating that $[\ddot{x}, \ddot{z}]$ is an edge in N. The analogous

argument works for X_γ, but the reason why X_γ cannot point toward α is different: it is ruled out by the statement of the Proposition. □

Recall the conventions set before the statement of Proposition 4.2.

Proposition 4.3 *Suppose $\tau = \dot{x}$ is trivalent. Contracting neither, either, or (if B has at least five vertices) both of $e_{\ddot{z}}$ and $e_{\ddot{y}}$ in $M \cup_\tau N$ yields a planar map possessing a rigid embedding.*

Proof. Pretend $e_{\ddot{x}}$ has already been contracted, and apply the constructions in the proofs of (6)–(8) in Proposition 4.2. (In reality, only (6) and (7) are required here, given the symmetry switching the roles of y and z). □

5 Triconnectivity and rigid embedding

Theorem 5.1 (Rigid embedding) *A planar map M with given axial vertices $\dot{x}, \dot{y}, \dot{z}$ can be rigidly embedded in a grid surface if and only if the extended map $M_\infty(\dot{x}, \dot{y}, \dot{z})$ is triconnected. In particular, every triconnected planar map can be rigidly embedded.*

Proof. ($\Rightarrow$) Let $M \hookrightarrow \mathcal{S}_\mathcal{V}$ be an axial rigid embedding, and delete two vertices ν, ω from M_∞. If $\{\nu, \omega\} \subset \{\dot{x}, \dot{y}, \dot{z}\}$ then each remaining vertex has an orthogonal flow to the third axial vertex, by independence of the orthogonal flows to $\dot{x}, \dot{y}, \dot{z}$. If $\{\nu, \omega\} \not\subset \{\dot{x}, \dot{y}, \dot{z}\}$ then what remains of the exterior cycle is connected, and every vertex in the deletion still has an orthogonal flow to the exterior cycle.

($\Leftarrow$) Induct on the sum of the number of regions and the number of edges in M, observing that the minimal sum of five is attained only when M is a triangle. Assume the notation of Section 3, and suppose M is not a triangle. Letting e be the edge leaving $\dot{x}$ towards $\dot{y}$ on the exterior cycle of M, we claim that at least one of the following occurs:

1. The endpoints of e are $\dot{x}$ and $\dot{y}$.
2. The edge e does not border a triangle in $M_\infty(\dot{x}, \dot{y}, \dot{z})$, and M/e is triconnected.
3. The edge e does not contain $\dot{y}$, and M contains a proper ring C for which $\underline{\ddot{x}} = \infty$.
4. The edge e does not contain $\dot{y}$, and M contains a proper ring C for which
   (a) $\underline{\ddot{x}} = \ddot{x} = \dot{x}$;
   (b) $\ddot{y}$ is the other endpoint of e, and $\ddot{y} \neq \dot{y}$ (that is, $\underline{\ddot{y}} \neq \infty$); and
   (c) $\underline{\ddot{z}}$ does not lie between $\dot{x}$ and $\dot{z}$ on the exterior cycle of M, and no single edge outside the region bounded by C has endpoints $\{\dot{x}, \underline{\ddot{z}}\}$.

If the first three cases do not occur, then Proposition 1.1 produces a ring C in which $\ddot{x}$ and $\ddot{y}$ are the two endpoints of e, while $\ddot{z}$ is a vertex in $F \cap F'$ as in

Proposition 1.1.2. Note that C is proper because $\ddot{y} \neq \dot{y}$. Let C be a maximal such ring.

The first half of 4(c) holds; if not, construct a ring satisfying option 3 as follows: replace the arc of C connecting $\dot{x}$ to $\ddot{z}$ with the arc traversing the exterior cycle from $\dot{x}$ to $\underline{\ddot{z}}$ and then $e_{\ddot{z}}$ (the latter only if $\ddot{z} \neq \underline{\ddot{z}}$). The second half of 4(c) also holds, for if an edge f outside C connects $\dot{x}$ to $\underline{\ddot{z}}$, then replace the arc of C connecting $\dot{x}$ to $\ddot{z}$ in C by f and $e_{\ddot{z}}$ (if $\underline{\ddot{z}} \neq \ddot{z}$). The resulting cycle is a larger ring satisfying the condition defining C, contradicting maximality. Finally, 4(a) holds by the failure of option 3: C does not contain the edge of M leaving $\dot{x}$ toward $\dot{z}$ on the exterior cycle of M.

Given the first option, M has a proper ring C, with $\underline{\ddot{z}} = \ddot{z} = \dot{z}$, containing every bounded region of M except the one containing e. Let $M \cap R \hookrightarrow \mathcal{S}_{\mathcal{V}}$ be a geodesic embedding, where R is the union of regions contained within C. Leaving $\dot{x}, \dot{y}$ fixed while adding 1 to the z-coordinates of every other vector in $\mathcal{V}$ yields a grid surface $\mathcal{S}_{\mathcal{V}'}$ whose rigid geodesics constitute an embedding of M; the easy proof is omitted.

Given the second option, use Proposition 4.1. Given the third or fourth option, use Lemma 3.1 with $M = M/R$ and $N = M \cap R$, along with Proposition 4.3 for option 3, or Lemma 3.2 and Proposition 4.2 for option 4. The 'five vertex' conditions in Propositions 4.2 and 4.3 are always satisfied when reconstructing M from the tethered gluing $(M/R) \cup_\tau (M \cap R)$, because M is simple and triconnected. □

The next corollary clarifies the close connection between grid surfaces and order dimension for posets. It shows that Theorem 5.1 generalizes the three-variable special case of [BPS98, Theorem 6.4], which is presented in the equivalent language of monomial ideals.

Corollary 5.2 (Brightwell–Trotter [BT93]) *The vertices, edges, and bounded regions of any triconnected planar map form a partially ordered set of order dimension* ≤ 3.

Proof. Theorem 5.1 and Corollary 2.5. □

Example 5.3 Theorem 5.1 is stronger than Corollary 5.2, even for triconneted maps. In general, every inclusion of the vertex-edge poset of M into $\mathbb{N}^3$ yields an inclusion of the vertex set $\mathcal{V} \hookrightarrow \mathcal{S}_{\mathcal{V}}$ such that each edge of M is a rigid geodesic in $\mathcal{S}_{\mathcal{V}}$. What fails is that there may be orthogonal rays in $\mathcal{S}_{\mathcal{V}}$ that are not contained in any edges of M. Faces such as the central face in Figure 6 are then forced to lie off of $\mathcal{S}_{\mathcal{V}}$. □

Remark 5.4 Rigid embeddings give a fresh perspective on a standard fact, known as Menger's theorem: If G is a triconnected planar graph and $\nu, \omega \in G$ are distinct vertices, then there are three independent paths from ν to ω in G. To explain Menger's theorem via Theorem 5.1, let M be a plane drawing of G, and suppose $e_1, \ldots, e_r$ are the edges of M containing ν, in cyclic order. Form

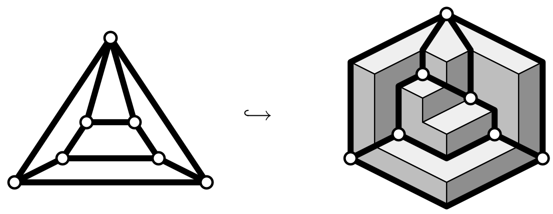


Figure 6: A vertex-edge-face poset embedding that is not a geodesic embedding

a new map M' by drawing a small circle C around ν and adding new vertices $\nu_1, \ldots, \nu_r$ where $e_1, \ldots, e_r$ intersect C. Then set $M^\nu = \mathrm{del}(M'; \nu)$, with underlying graph G^ν. Clearly G^ν is triconnected.

Choose a plane drawing of G^ν in which C is the exterior cycle, and let $G^\nu \hookrightarrow \mathcal{S}$ be an axial geodesic embedding compatible with some (any) choice of axial vertices $\dot{x}, \dot{y}, \dot{z} \in C$. The orthogonal flows from ω to $\dot{x}, \dot{y}, \dot{z}$ in G^ν first intersect C at points $\nu_{i_x}, \nu_{i_y}, \nu_{i_z}$, giving rise to truncated orthogonal flows $[\omega, \nu_{i_u}]$ for $u \in \{x, y, z\}$. Connecting $[\omega, \nu_{i_u}]$ to ν via the arc in G between ν_{i_u} and ν yields independent paths in G from ω to ν.

Part II
Monomial ideals

6 Betti numbers

Let k be a field, and consider the polynomial ring $R = k[x, y, z]$ with the $\mathbb{Z}^3$-grading in which $\deg(x) = (1, 0, 0)$, $\deg(y) = (0, 1, 0)$, and $\deg(z) = (0, 0, 1)$. Use $I_{\mathcal{V}} = \langle m^\nu \mid \nu \in \mathcal{V}\rangle$ to denote the ideal generated the monomials $m^\nu = x^{\nu_x} y^{\nu_y} z^{\nu_z}$ for $\nu \in \mathcal{V}$. The integer points in $\langle \mathcal{V} \rangle$ coincide with the exponent vectors on monomials in $I_{\mathcal{V}}$, and $\mathcal{V}$ is axial if and only if $I_{\mathcal{V}}$ is artinian, containing a power of each variable.

Any principal monomial ideal $\langle m \rangle$ is a free $\mathbb{Z}^3$-graded R-module of rank 1. If ϕ is $\mathbb{Z}^3$-graded homomorphism $\bigoplus \langle m_q \rangle \longleftarrow \bigoplus \langle m_p \rangle$ of degree zero, then we can express ϕ as a **monomial matrix**. This is a matrix whose entries $\lambda_{pq} \in k$ are scalars, and whose p^{th} row (resp. q^{th} column) is labeled by the monomial m_p (resp. m_q) that generates the corresponding p^{th} source (resp. q^{th} target) summand. Of course, $\lambda_{pq} = 0$ whenever m_q does not divide m_p, because then there are no nonzero $\mathbb{Z}^3$-graded maps $\langle m_q \rangle \leftarrow \langle m_p \rangle$. The map ϕ is called **minimal** if also $\lambda_{pq} = 0$ whenever $m_p = m_q$. See [Mil00a, Section 2] for more on monomial matrices.

We consider **free resolutions** of $I_{\mathcal{V}}$ that are exact sequences having the form

$$\mathcal{F}.: \quad 0 \leftarrow I_{\mathcal{V}} \xleftarrow{\phi_0} \mathcal{F}_0 \xleftarrow{\phi_1} \mathcal{F}_1 \xleftarrow{\phi_2} \mathcal{F}_2 \leftarrow 0, \tag{10}$$

in which $\mathcal{F}_i \cong \bigoplus_p \langle m_{ip} \rangle$ for some (finite set of) monomials $m_{ip} \in R$. We call $\mathcal{F}.$

minimal if ϕ_1 and ϕ_2 are minimal (for any such direct sum decomposition). The **Betti number** $\beta_{i,\alpha}(I_{\mathcal{V}})$ is the number of m_{ip} equal to m^α, when $\mathcal{F}.$ is minimal. This homological data reflects the local properties of $\mathcal{S}_{\mathcal{V}}$ near the vector α via the **Koszul simplicial complex** of $\mathcal{V}$ at $\alpha \in \mathbb{N}^3$,

$$K_\alpha(\mathcal{V}) \quad = \quad \{\sigma \in \{0,1\}^3 \mid \alpha - \sigma \in \langle \mathcal{V} \rangle\},$$

which is a subcomplex of the abstract triangle $(\{0,1\}^3, \preceq)$.

Proposition 6.1 ([Hoc77, Roz70]) *$\beta_{i,\alpha}(I_{\mathcal{V}}) = \dim_k \widetilde{H}_{i-1}(K_\alpha(\mathcal{V}); k)$ is the dimension of the $(i-1)^{\mathrm{st}}$ reduced simplicial homology of $K_\alpha(\mathcal{V})$ with coefficients in k.*

The small number of simplicial complexes on three vertices seriously limits the possibilities for nonzero Betti numbers.

Lemma 6.2 *If $\mathcal{V} \subset \mathbb{N}^3$ and $i \in \mathbb{N}$ then $\beta_{i,\alpha}(I_{\mathcal{V}}) \neq 0$ for at most one $\alpha \in \mathbb{Z}^3$. If $\beta_{i,\alpha}$ is nonzero then $\beta_{i,\alpha} = 1$ unless $K_\alpha(\mathcal{V})$ has 3 vertices and no edges (so $\beta_{1,\alpha} = 2$).*

Proof. Use the previous proposition, and list all simplicial complexes on 3 vertices. □

7 Cellular resolutions

Suppose M is a cell complex (precisely, a finite CW complex) of dimension 2 whose cells P have vector labels $\alpha_P \in \mathbb{N}^3$, in such a way that $\alpha_P \preceq \alpha_{P'}$ if P lies in the closure of P'. For instance, if the vertices have natural labels, then define α_P to be the join of the labels on the vertices of P. Such a **labeled cell complex** determines monomial matrices ϕ_{vertex}, ϕ_{edge}, and ϕ_{region} for a **cellular free complex** $\mathcal{F}_M$, by labeling the rows and columns of matrices for the boundary map of the ordinary chain complex of M with the monomials m^{α_P}. Using P also to denote the basis vector of a rank 1 free R-module $\langle m^{\alpha_P} \rangle$, the cellular free complex $\mathcal{F}_M$ takes the form

$$0 \leftarrow I_{\mathcal{V}} \xleftarrow{\phi_{\mathrm{vertex}}} \bigoplus_{\text{vertices } v} R \cdot v \xleftarrow{\phi_{\mathrm{edge}}} \bigoplus_{\text{edges } e} R \cdot e \xleftarrow{\phi_{\mathrm{region}}} \bigoplus_{\text{regions } F} R \cdot F \leftarrow 0. \tag{11}$$

We say M **supports** $\mathcal{F}_M$; see [BS98, Mil00a] for more on cellular monomial matrices.

Our main examples of labeled cell complexes are of course the geodesically embedded planar maps $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$, whose labels are determined by the obvious vertex labels by taking joins. We shall see in the next section that this cellular free complex is exact and minimal, so it provides a **cellular minimal free resolution** of $I_{\mathcal{V}}$. Unfortunately, there are monomial ideals whose associated grid surfaces contain no geodesically embedded map.

Example 7.1 Let $I_{\mathcal{V}} = \langle x,y,z\rangle^2 = \langle x^2,y^2,z^2,xy,xz,yz\rangle$. The orthogonal rays X_{yz}, Y_{xz}, Z_{xy} meet at a single point not in $\mathcal{V}$, so there can be no planar map geodesically embedded in $\mathcal{S}_{\mathcal{V}}$. However, $\langle x,y,z\rangle^2$ still has a minimal cellular resolution: connect the midpoints of the edges of a triangle, and delete any one of the three interior edges. Label the resulting planar map M with x^2,y^2,z^2 on the corners of the outside triangle; xy between x^2 and y^2; yz between y^2 and z^2; and xz between x^2 and z^2. Label the edges and regions of M by the joins of their vertex labels. $\square$

Although the map M in the above example fails to embed geodesically in $\mathcal{S}_{\mathcal{V}}$, the extended map $M_\infty(x^2,y^2,z^2)$ is still triconnected. This phenomenon is general. In the following proposition, we do not require planarity of M, so we use $G_\infty(\dot{x},\dot{y},\dot{z})$ to mean the abstract graph obtained from G by adding a new vertex ∞ connected to each of $\dot{x},\dot{y},\dot{z}$.

Proposition 7.2 *If the labeled cell complex M supports a minimal free resolution of an artinian ideal $I_{\mathcal{V}}$, and the 1-skeleton of M is a graph G, then the extended graph $G_\infty(\dot{x},\dot{y},\dot{z})$ is triconnected, where $\dot{x},\dot{y},\dot{z}$ are the vertices whose labels lie on the axes.*

Proof. Given $\nu \in \mathcal{V}$ with $\nu \neq \dot{z}$, the orthogonal ray Z_ν leaving ν has its head at some vector $\alpha \in \mathbb{N}^3$ for which $K_\alpha(\mathcal{V})$ is disconnected; indeed, the vertex $(0,0,1)$ is isolated in $K_\alpha(\mathcal{V})$. Choose a vertex $\omega' \neq \nu$ preceding α, so that $m^{\alpha-\nu}\nu - m^{\alpha-\omega'}\omega' \in \ker(\phi_{\text{vertex}})$. Since $\alpha - \nu = Z(\nu)$ lies on the z-axis Z, there must be an edge $e \in M$ connecting ν to a vertex ω (possibly different from ω') such that $\phi_{\text{edge}}(e) = z^d\nu - m^{\nu\vee\omega-\omega}\omega$ for some $d \in \mathbb{N}$. Clearly $\omega_x \preceq \nu_x$ and $\omega_y \preceq \nu_y$.
Repeating the procedure with ω in place of ν, and with x or y in place of z, we find that M contains paths analogous to orthogonal flows from ν to each axial vertex $\dot{x},\dot{y},\dot{z}$. As in Section 2, these paths are independent, intersecting only at ν. $\square$

8 Graphs to minimal resolutions

Lemma 8.1 *If $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ is a geodesic embedding, then $\beta_{1,\alpha}(I_{\mathcal{V}}) \neq 0$ if and only if $\alpha = \nu\vee\omega$ for some elbow geodesic $[\nu,\omega] \in M$, and $\beta_{1,\nu\vee\omega}(I_{\mathcal{V}}) = 1$ in this case.*

Proof. Assume $\beta_{1,\alpha}(I_{\mathcal{V}}) \neq 0$. Then $K_\alpha(\mathcal{V})$ is disconnected by Proposition 6.1, and this occurs if and only if $K_\alpha(\mathcal{V})$ contains an isolated vertex. An isolated vertex of $K_\alpha(\mathcal{V})$ occurs if and only if α lies on an orthogonal ray leaving some vertex $\nu \in \mathcal{V}$. Therefore, $\alpha = \nu\vee\omega$ for some $\omega \neq \nu$ by the edge axiom. Since $[\nu,\omega]$ is an elbow geodesic, $K_{\nu\vee\omega}(\mathcal{V})$ cannot have three isolated points because three orthogonal rays cannot meet at the point $\nu\vee\omega$ in the relative interior of the edge $[\nu,\omega]$ of M. $\square$

Here is a result that sometimes reduces statements about arbitrary geodesic or rigid embeddings to axial ones. Its straightforward proof is omitted.

Lemma 8.2 *Append axial vertices to* $\mathcal{V}$ *by letting* $\overline{\mathcal{V}} = \mathcal{V} \cup \{\dot{u} \mid \mathcal{S}_{\mathcal{V}} \cap U = \varnothing\}$ *for sufficiently large* $|\dot{u}|$. *A planar map* M *is geodesically embedded in* $\mathcal{S}_{\overline{\mathcal{V}}}$ *if and only if* $N = \mathrm{del}(M; \overline{\mathcal{V}} \smallsetminus \mathcal{V})$ *is geodesically embedded in* $\mathcal{S}_{\mathcal{V}}$. *Furthermore,* $M \hookrightarrow \mathcal{S}_{\overline{\mathcal{V}}}$ *is rigidly embedded if and only if* $N \hookrightarrow \mathcal{S}_{\mathcal{V}}$ *is rigidly embedded.*

Lemma 8.3 *Let* $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ *be a geodesic embedding, and* $\mathcal{S}_{\mathcal{V}}^{\max}$ *the set of points in* $\mathcal{S}_{\mathcal{V}}$ *maximal under the partial order induced by the relation* $\preceq$ *on* $\mathbb{R}^3$. *Then* $\alpha \in \mathcal{S}_{\mathcal{V}}^{\max} \Leftrightarrow \beta_{2,\alpha}(I_{\mathcal{V}}) \neq 0 \Leftrightarrow \alpha = \alpha_F$ *is the join of the vertices in a bounded region* F *of* M.

Proof. For the equivalence $\alpha \in \mathcal{S}_{\mathcal{V}}^{\max} \Leftrightarrow \beta_{2,\alpha}(I_{\mathcal{V}}) \neq 0$, use the fact that $K_\alpha(\mathcal{V})$ is the boundary of the triangle; the easy details are omitted. The equivalence $\alpha \in \mathcal{S}_{\mathcal{V}}^{\max} \Leftrightarrow \alpha = \alpha_F$ holds for all vertex sets $\mathcal{V}$ if it holds when $\mathcal{V}$ is axial. Indeed, using the notation and result of Lemma 8.2, the maximal points of $\mathcal{S}_{\mathcal{V}}$ are still maximal in $\mathcal{S}_{\overline{\mathcal{V}}}$, while the points in $\mathcal{S}_{\overline{\mathcal{V}}}^{\max} \smallsetminus \mathcal{S}_{\mathcal{V}}^{\max}$ are exactly those having u-coordinate $|\dot{u}|$ for some u such that $\mathcal{S}_{\mathcal{V}} \cap U = \varnothing$, by the region axiom applied to $\mathcal{S}_{\overline{\mathcal{V}}}$. The bounded regions of N having such joins disappear upon deleting $\dot{u}$.
Assume henceforth that $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ is axial. If $\rho \in \mathcal{S}_{\mathcal{V}}$ has some coordinate $\rho_u = 0$, then $\rho \notin \mathcal{S}_{\mathcal{V}}^{\max}$ because adding ε to any other coordinate of ρ yields another point in $\mathcal{S}_{\mathcal{V}}$. Therefore each maximal point of $\mathcal{S}_{\mathcal{V}}$ lies in a bounded region of M. When F is such a bounded region, the region axiom implies $\alpha_F \in \mathcal{S}_{\mathcal{V}}^{\max}$, because some vertex $\nu \in \mathcal{V}$ strongly precedes $\alpha_F + \varepsilon\dot{u}$ for any $\varepsilon > 0$ and $u \in \{x, y, z\}$.
Every point ρ on a given elbow geodesic $[\nu, \omega]$ precedes $\nu \vee \omega$ by definition, so $\rho \preceq \nu \vee \omega \preceq \alpha_F$ whenever $[\nu, \omega] \subseteq F$. Any point σ on the line segment in $\mathbb{R}^3$ connecting ρ to α_F therefore satisfies $\rho \preceq \sigma \preceq \alpha_F$, whence $\sigma \in \mathcal{S}_{\mathcal{V}}$. It follows that F is the union of such line segments, so every point of F precedes α_F. □

Theorem 8.4 *Given a geodesic embedding* $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$, *the cellular free complex* $\mathcal{F}_M$ *is a minimal free resolution of* $I_{\mathcal{V}}$.

Proof. Since $\beta_{i,\alpha}(I_{\mathcal{V}}) \neq 0$ if and only if $\beta_{i,\alpha}(I_{\mathcal{V}}) = 1$ by Lemmas 6.2 and 8.1, it makes sense simply to speak of the i^{th} **Betti degrees** α, for which $\beta_{i,\alpha} = 1$. The zeroth, first, and second Betti degrees are the labels on the vertices, edges, and regions of M, respectively, by Lemmas 6.2, 8.1, and 8.3. Any minimal free resolution $\mathcal{F}.$ of $I_{\mathcal{V}}$ as in (10) therefore takes the form of (11), at least as a homologically graded module; that is, $\mathcal{F}. \cong \mathcal{F}_M$ abstractly as modules. We need to show that some choice of this abstract isomorphism is a homomorphism of *complexes*.
Identifying the homological degree zero parts of $\mathcal{F}.$ and $\mathcal{F}_M$, the zeroth homology of $\mathcal{F}_M$ surjects onto $I_{\mathcal{V}}$ because the image of ϕ_{edge} is clearly contained in the kernel of ϕ_0. Since $\mathcal{F}.$ is exact and $\mathcal{F}_M$ is a complex of free modules, there

exists a homomorphism $\psi : \mathcal{F}_M \to \mathcal{F}$. lifting the surjection on zeroth homology and the isomorphism in homological degree zero.

Suppose $e \in \mathcal{F}_M$ maps to $\psi(e) = \sum m_j e'_j$, where each $m_j \in R$ is a monomial with nonzero scalar coefficient, and each e'_j denotes the generator of $\mathcal{F}_1$ corresponding to the edge $e_j \in M$. The elbow geodesic $e = [\nu, \omega]$ in M contains an orthogonal ray U_ν, so that $\pm\phi_1(\sum m_j e'_j) = m^{\nu\vee\omega-\nu}\nu - m^{\nu\vee\omega-\omega}\omega$, where the first term is $m^{\nu\vee\omega-\nu}\nu = m^{U(\nu)}\nu$. Thus $m^{U(\nu)}\nu = u^{|U_\nu|}\nu$ appears with nonzero scalar coefficient in $\phi_1(m_j e'_j)$ for some j. Since there is a *unique* first Betti degree α satisfying $\alpha - \nu \in U$, namely $\alpha_e = \nu\vee\omega$, it must be that $e'_j = e'$ and m_j is a nonzero scalar. Nakayama's lemma implies $\psi_1 : (\mathcal{F}_M)_1 \to \mathcal{F}_1$ is surjective, and hence an isomorphism by rank considerations.

No summand $R \cdot F \subset \mathcal{F}_M$ can map to zero in $\mathcal{F}_2$ because $\phi_{\text{region}}(F)$ is nonzero in $\mathcal{F}_M$, and ψ is an isomorphism in homological degree 1. On the other hand, the second Betti degrees are pairwise incomparable by Lemma 8.3. Thus $\psi(F)$ is some nonzero scalar multiple of the unique generator of $\mathcal{F}_2$ in degree α_F. The map ψ is therefore an isomorphism in homological degree 2, completing the proof. □

Corollary 8.5 *A planar map M with axial vertices $\dot{x}, \dot{y}, \dot{z}$ supports a minimal free resolution of an artinian monomial ideal if and only if $M_\infty(\dot{x}, \dot{y}, \dot{z})$ is triconnected. In particular, every triconnected planar map supports a minimal free resolution.*

Proof. 'Only if' is Proposition 7.2; apply Theorem 8.4 to Theorem 5.1 for 'if'. □

9 Uniqueness vs. nonplanarity

Continuing with the analogy at the beginning of the Introduction, circle packings and polytopes that realize planar graphs are unique up to Möbius transformation and spherical rotation, respectively (see [Zie95] for discussion and references). Rigid embeddings $M \hookrightarrow \mathcal{S}_\mathcal{V}$ for a fixed planar map are similarly not unique: at the very least, any order-preserving bijection of $\mathcal{V}$ as in Lemma 2.1 gives another rigid embedding. Of course, such bijections affect neither the combinatorics nor the algebra. In fact, rigid embeddings are uniquely determined by the algebraic properties of the grid surface in question, specifically the minimal free resolution of the corresponding monomial ideal.

Corollary 9.1 *When $M \hookrightarrow \mathcal{S}_\mathcal{V}$ is rigidly embedded, M is the* unique *cell complex supporting a minimal cellular free resolution of $I_\mathcal{V}$.*

Proof. Let N be a labeled cell complex supporting a minimal cellular free resolution of $I_\mathcal{V}$. The abstract graph underlying N (the 1-skeleton) coincides with that of M by Theorem 8.4 and rigidity. The label α_F on any region F of N is a second Betti degree of $I_\mathcal{V}$, and hence coincides with the label on a region F' of M by Theorem 8.4. The boundary of F in N is a cycle of edges

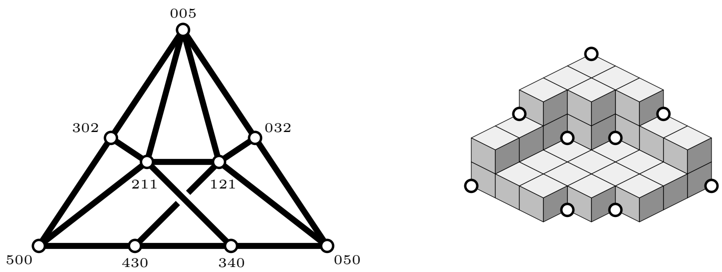


Figure 7: Nonplanar minimal free resolution

whose degrees precede α_F. The only such cycle consists of all the edges whose degrees precede α_F, by the rigid region axiom for F' in M (by Lemma 8.2 the rigid region axiom holds for nonaxial grid surfaces). $\square$

Nonrigid monomial ideals can have many distinct isomorphism classes of minimal cellular resolutions; [MS99, Figure 4] depicts an example of this phenomenon. In fact, the bad behavior gets much worse.

Example 9.2 Minimal cellular resolutions of ideals in $k[x, y, z]$ need not be supported planar cell complexes. In fact, explicit examples crop up with even the smallest violations of rigidity. For instance, let

$$\mathcal{V} \quad = \quad \{(4,3,0),\ (3,4,0),\ (3,0,2),\ (2,1,1),\ (1,2,1),\ (0,3,2)\}$$

and $\overline{\mathcal{V}} = \mathcal{V} \cup \{(5,0,0),\ (0,5,0),\ (0,0,3)\}$. The cell complex M depicted in Figure 7 consists of five triangles in addition to the three quadrilaterals with vertices

$$\{500, 430, 121, 211\},\ \{430, 121, 211, 341\},\ \{050, 121, 211, 340\}.$$

Label the edges and regions by the joins of their vertex labels. That M supports a minimal free resolution of $I_{\overline{\mathcal{V}}}$ can be checked by verifying for each $\alpha \preceq (5,5,3)$ that $M_{\preceq\alpha}$ is acyclic [BS98, Corollary 1.3]. That M cannot be planar follows by contracting the edges labeled 530 and 350 while deleting the edges labeled 312 and 132 to get the complete graph K_5 as a minor of the 1-skeleton. $\square$

10 Deformation and genericity

Given a finite subset $\mathcal{W} \subset \mathbb{R}^3$, write $\vee\mathcal{W}$ for the join of the vectors in $\mathcal{W}$, and set $m^{\mathcal{W}} = m^{\vee\mathcal{W}}$. Following [BPS98], define the **Scarf complex** of $\mathcal{V}$,

$$\Delta_{\mathcal{V}} \quad = \quad \{\mathcal{W} \subseteq \mathcal{V} \mid \text{ if } \vee\mathcal{W}' = \vee\mathcal{W} \text{ for some } \mathcal{W}' \subseteq \mathcal{V} \text{ then } \mathcal{W}' = \mathcal{W}\},$$

to consist of the subsets whose joins are uniquely attained. It is an easy (but not obvious) fact that $\Delta_{\mathcal{V}}$ is a simplicial complex. Each face $\mathcal{W} \in \Delta_{\mathcal{V}}$ comes with a natural label $\vee\mathcal{W}$, and the resulting cellular free complex $\mathcal{F}_{\Delta_\mathcal{V}}$ is called the **free Scarf complex** of $I_\mathcal{V}$. The Scarf complex is planar by virtue of its containment in the **hull complex** [BS98], so the union of its edges and vertices is a planar map. This planar map has already appeared, in Section 2: two vertices $\nu, \omega \in \mathcal{V}$ are connected by a rigid geodesic if and only if $\{\nu, \omega\} \in \Delta_\mathcal{V}$.

Under special circumstances, the Scarf complex is rigidly embedded in $\mathcal{S}_\mathcal{V}$. To be precise, call $\mathcal{V}$ **strongly generic** if no two distinct elements of $\mathcal{V}$ share a nonzero coordinate. In other words, $\nu_u = \omega_u \neq 0$ for some $u \in \{x, y, z\}$ implies $\nu = \omega$.

Corollary 10.1 (Bayer–Peeva–Sturmfels [BPS98, §3]) *The free Scarf complex $\mathcal{F}_{\Delta_\mathcal{V}}$ minimally resolves $I_\mathcal{V}$ when $\mathcal{V}$ is strongly generic.*

Proof. Strong genericity easily implies that every orthogonal ray is contained in a rigid geodesic, so the Scarf graph rigidly embeds in $\mathcal{S}_\mathcal{V}$. It is straightforward to verify that the labels on triangles (2-dimensional faces) in $\Delta_\mathcal{V}$ are maximal in $\mathcal{S}_\mathcal{V}$. Furthermore, every maximal point of $\mathcal{S}_\mathcal{V}$ has exactly three vectors in $\mathcal{V}$ preceding it by Lemma 8.3, the region axiom, and strong genericity. Therefore, all maximal points are labels on regions in $\Delta_\mathcal{V}$, and the result holds by Theorem 8.4. $\square$

Since the definition of the Scarf complex depends only on the coordinatewise order of the exponents of the generators, it also makes sense for (formal) monomials with real exponents in $\mathbb{R}^n$. This makes way for the following definition. Let $\mathbb{Q}$ denote the rational numbers. A **deformation** ϵ of $\mathcal{V}$ is a choice of vectors $\epsilon^\nu = (\epsilon^\nu_x, \epsilon^\nu_y, \epsilon^\nu_z) \in \mathbb{Q}^3$ for each $\nu \in \mathcal{V}$ satisfying

$$\nu_u < \omega_u \ \Rightarrow\ \nu_u + \epsilon^\nu_u < \omega_u + \epsilon^\omega_u, \qquad \text{and} \qquad \nu_u = 0 \ \Rightarrow\ \epsilon^\nu_u = 0$$

for $u \in \{x, y, z\}$. In practice, everything we do is invariant under scaling of $\mathcal{V}$, so we will always assume $\mathcal{V}^\epsilon = \{\nu + \epsilon^\nu \mid \nu \in \mathcal{V}\}$ consists of integer vectors. Set $\nu^\epsilon = \nu + \epsilon^\nu$.

The sole purpose of the ϵ vectors is to break ties one way or the other between equal nonzero coordinates of vectors in $\mathcal{V}$. In this manner, deformations of $\mathcal{V}$ are closer to being generic than $\mathcal{V}$ is. The verb **specialize** is used here to indicate that a deformation (generization) is being reversed; thus $\mathcal{V}$ is a **specialization** of $\mathcal{V}^\epsilon$ if the latter is a deformation of the former.

One particular deformation will play a key role in the coming sections. To define it, let $\mathcal{V}(u, a) = \{\nu \in \mathcal{V} \mid \nu_u = a\}$ for each $0 < a \in \mathbb{N}$ and $u \in \{x, y, z\}$. Up to order-preserving bijection (as in Lemma 2.1), there is a unique deformation ϵ satisfying the following condition as well as its analogues via cyclic permutation of x, y, z:

$$\text{If the elements of } \mathcal{V}(z, a) \text{ satisfy } \nu_x > \cdots > \omega_x, \text{ then } a = \nu^\epsilon_z < \cdots < \omega^\epsilon_z. \quad (12)$$

Note that $\nu_x > \cdots > \omega_x$ is equivalent to $\nu_y < \cdots < \omega_y$ for elements of $\mathcal{V}(z, a)$, by pairwise incomparability. Thus, looking down the x-axis, ϵ raises the vectors in $\mathcal{V}(z, a)$ higher as they move to the right.

11 Ideals to graphs: algorithm

Arbitrary monomial ideals in more than three variables need not have minimal cellular free resolutions [RW01], but limiting to three variables forces better behavior.

Theorem 11.1 *Any monomial ideal $I_\mathcal{V} \subset k[x,y,z]$ has a cellular minimal free resolution supported on a labeled planar map M.*

The proof (at the end of Section 12) will reduce to the artinian case, for which Algorithm 11.2 produces M. The justification of Algorithm 11.2 appears in Section 12. Heuristically, the idea is to apply the deformation ϵ of (12), and show that specializing $\mathcal{V}^\epsilon \rightsquigarrow \mathcal{V}$ step by step makes the spurious edges in the Scarf triangulation disappear one at a time.
More precisely, the algorithm specializes $\mathcal{V}^\epsilon$ back to $\mathcal{V}$ by making strict inequalities $\nu_u^\epsilon < \omega_u^\epsilon$ into equalities $\nu_u = \omega_u$, judiciously and one at a time. Before each specialization step, the (already partially specialized) ideal has a cellular minimal resolution by induction; after each specialization step, the same planar map still supports a cellular free resolution, although it may not be minimal. However, in the nonminimal case, minimality is achieved by removing exactly one edge.

Algorithm 11.2

input an artinian ideal $I_\mathcal{V} \subset k[x,y,z]$
output a planar map M supporting a cellular minimal free resolution of $I_\mathcal{V}$
initialize $\epsilon :=$ the deformation of $I_\mathcal{V}$ in (12) to a strongly generic ideal $I_{\mathcal{V}^\epsilon}$
$M :=$ Scarf complex of $\mathcal{V}^\epsilon$
while $\mathcal{V}^\epsilon \neq \mathcal{V}$ do
choose $\nu \in \mathcal{V}$ and $u \in \{x,y,z\}$ such that $\nu_u^\epsilon \neq \nu_u$ and ν_u^ϵ is minimal; for ease of notation, assume $u = x$ by applying a *cyclic* permutation of $\{x,y,z\}$ translating u to x, if necessary
$\gamma_x := \nu_x$
$\gamma_y := \nu_y^\epsilon + |Y_{\nu^\epsilon}|$
$\gamma_z := \min_{\omega\neq\nu}\{\omega_z^\epsilon \mid \omega_x^\epsilon \le \nu_x \text{ and } \omega_y^\epsilon < \gamma_y\}$
$\rho :=$ the element of $\{\omega^\epsilon \in \mathcal{V}^\epsilon \mid \omega_y^\epsilon = \gamma_y \text{ and } \omega_z^\epsilon < \gamma_z\}$ with maximal ω_z^ϵ
redefine ϵ by replacing the x-coordinate ϵ_x^ν with 0
M by changing the label on ν^ϵ accordingly
if $\rho_x = \nu_x$
then redefine M by removing edge labeled $\gamma = (\gamma_x, \gamma_y, \gamma_z)$
else leave M unchanged
end if-then-else
end while-do
output M

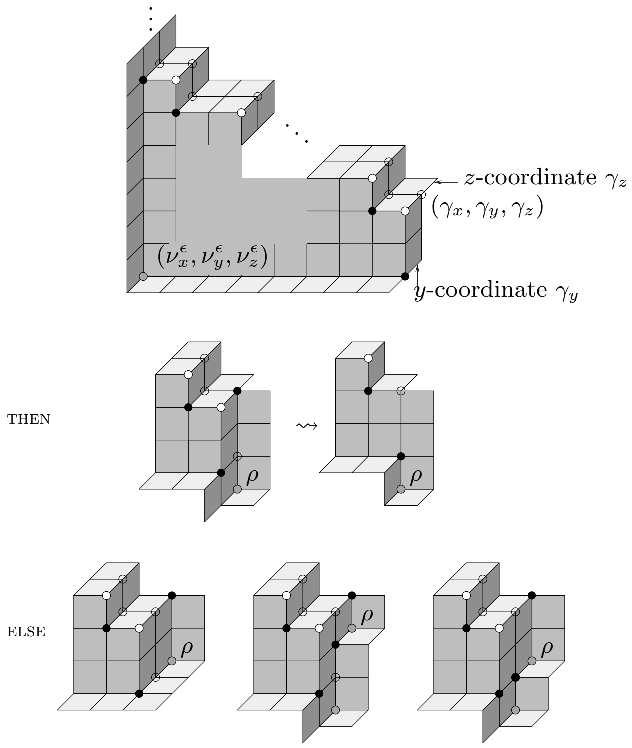


Figure 8: The geometry of Algorithm 11.2

Remark 11.3 Here are some elementary observations to aid in parsing the algorithm. See also Figure 8, which illustrates the geometry.

(i) It may be necessary to scale $\mathcal{V}$ in order to choose $\epsilon^\nu \in \mathbb{N}^3$ for all ν, because of the condition $a = \nu^\epsilon_u$.

(ii) Note that $\gamma_x = \nu_x$, not ν^ϵ_x.

(iii) The orthogonal ray Y_{ν^ϵ} used to define γ_y is bounded because $\nu^\epsilon \neq \nu$, so that ν cannot be the axial vertex $\dot{y}$.

(iv) The set used to define γ_z is nonempty because the axial vector $\dot{z}$ is in the set; indeed, $\nu^\epsilon \neq \nu$ implies $\nu \neq \dot{z}$.

(v) $\gamma_z > \nu^\epsilon_z$ because $\mathcal{V}^\epsilon$ consists of pairwise incomparable vectors.

(vi) The set defining ρ is nonempty because $\nu^\epsilon + Y_{\nu^\epsilon} = \nu^\epsilon \vee \omega^\epsilon$ for some ω^ϵ in this set. Uniqueness of ρ follows from pairwise incomparability of elements in $\mathcal{V}^\epsilon$.

(vii) Relabeling M at the REDEFINE step yields a cellular free resolution of the resulting ideal with the new ϵ by [GPW00, Theorem 3.3], though it need not be minimal.

Example 11.4 If $I_{\mathcal{V}} = \langle x^4, x^2y^2, x^2z^2, y^4, y^2z^2, z^4\rangle$ is obtained from Example 7.1 by a scale factor of 2, then the generic deformation $I_{\mathcal{V}^\epsilon} = \langle x^4, x^2y^3, x^3z^2, y^4, y^2z^3, x^4\rangle$ satisfies the condition of Algorithm 11.2. Furthermore, the Scarf complex of $I_{\mathcal{V}^\epsilon}$ is the triangle with its edge-midpoints connected, as in Example 7.1. If Algorithm 11.2 is run on this $I_{\mathcal{V}}$, then one of the three nonminimal edges is removed on the first iteration of the while loop. Precisely which of the nonminimal edges is removed depends on which $u \in \{x, y, z\}$ is chosen first; any u will work, not just $u = x$. In the remaining two iterations of the while loop, no further edges are removed. It is instructive to work out this example by hand; see Figure 8 for general pictures. □

12 Ideals to graphs: proof

The gruntwork in proving that the algorithm accomplishes its goal is contained in the following two technical lemmas, whose hypotheses are designed to be satisfied by the deformation taking place in one pass of the while loop (after a cyclic permutation of (x, y, z) and an order-preserving bijection as in Lemma 2.1, perhaps).

Lemma 12.1 *Suppose $I_{\mathcal{V}} \subset k[x, y, z]$ is artinian, and that $\nu \in \mathcal{V}$ has $\nu_x \neq 0$ and satisfies $\omega_y \geq \nu_y$ whenever $\omega_x = \nu_x$. Suppose further that $\epsilon = \{\epsilon^\omega\}_{\omega\in\mathcal{V}}$ is a deformation of $\mathcal{V}$ with $\epsilon^\omega = \mathbf{0}$ for $\omega \neq \nu$ and $\epsilon^\nu = (1, 0, 0)$. Let $\gamma_y = \nu^\epsilon_y + |Y_{\nu^\epsilon}|$. If $\alpha \in \mathbb{N}^3$ then $K_\alpha(\mathcal{V}^\epsilon) = K_\alpha(\mathcal{V})$ unless*

$$\nu_z \leq \alpha_z \quad \text{and} \quad \nu_y \leq \alpha_y \leq \gamma_y \quad \text{and} \quad \alpha_x \in \{\nu_x, 1 + \nu_x\}. \tag{13}$$

If $\alpha_y \neq \gamma_y$ and α satisfies (13) with $\alpha_x = \nu_x$, then $K_{\alpha+\epsilon^\nu}(\mathcal{V}^\epsilon) = K_\alpha(\mathcal{V})$ while both $K_\alpha(\mathcal{V}^\epsilon)$ and $K_{\alpha+\epsilon^\nu}(\mathcal{V})$ have no reduced homology.

The last sentence takes care of the case where α satisfies (13) and $\alpha_x = 1 + \nu_x$, because $\alpha + \epsilon^\nu$ has x-coordinate $1 + \nu_x$; the case $\alpha_y = \gamma_y$ will be covered in Lemma 12.2.

The idea comes from Figure 8, where the grey dots represent elements of $\mathcal{V}^\epsilon$, the white dots represent maximal points of $\mathcal{S}_{\mathcal{V}^\epsilon}$ (= second syzygies of $I_{\mathcal{V}^\epsilon}$ = irreducible components of $I_{\mathcal{V}^\epsilon}$), and the black dots represent first syzygies of $I_{\mathcal{V}^\epsilon}$. Looking from far down the x-axis, the vector ν^ϵ has a vertical plateau behind it: the big medium-grey wall, parallel to the yz-plane. Pushing ν^ϵ back to ν moves that vertical wall back a single unit. The only places where the topology of $K_\alpha(\mathcal{V})$ can possibly change are at lattice points α that sit either on the original wall in $\mathcal{S}_{\mathcal{V}^\epsilon}$ or its pushed-back image in $\mathcal{S}_{\mathcal{V}}$; these are the vectors α described in (13).

For vectors $\alpha + \epsilon^\nu$ that sit on the original wall but to the left of its right-hand edge (i.e. those with $\alpha_y \neq \gamma_y$), the Koszul simplicial complex $K_{\alpha+\epsilon^\nu}(\mathcal{V}^\epsilon)$ gets carried along for the ride to $K_\alpha(\mathcal{V})$ as the wall gets pushed back; the empty circles denote where the filled (black and white) dots get moved to. On the other hand, if α sits on the pushed-back image of the wall in $\mathcal{S}_{\mathcal{V}}$ (as the empty circles strictly to the left of γ_y do), then $\mathcal{S}_{\mathcal{V}^\epsilon}$ is translation-invariant in the x-direction near α, making $K_\alpha(\mathcal{V}^\epsilon)$ a cone; the same goes for $K_{\alpha+\epsilon^\nu}(\mathcal{V})$. Cones have no reduced homology. Having this geometry in mind, here's the official proof of the lemma.

Proof. First assume α fails to satisfy (13). Use σ to denote an element of $\{0,1\}^3$. To start off with, $\omega \preceq \alpha - \sigma$ if and only if $\omega^\epsilon \preceq \alpha - \sigma$ whenever $\omega \in \mathcal{V} \smallsetminus \nu$, so the only possible differences in the simplicial complexes $K_\alpha(\mathcal{V})$ and $K_\alpha(\mathcal{V}^\epsilon)$ come from the placement of ν and ν^ϵ relative to the vectors $\alpha-\sigma$. If $\alpha_x \notin \{\nu_x, 1+\nu_x\}$ then clearly $\nu \preceq \alpha-\sigma$ if and only if $\nu^\epsilon \preceq \alpha-\sigma$. And if $\alpha_y < \nu_y$ or $\alpha_z < \nu_z$, then neither ν nor ν^ϵ precedes $\alpha - \sigma$. The only remaining case of α not satisfying (13) has $\alpha_x \geq \nu_x$ and $\alpha_y > \gamma_y$ and $\alpha_z \geq \nu_z$. Suppose $\nu^\epsilon + Y_{\nu^\epsilon} = \nu^\epsilon \vee \omega$. Then $\nu^\epsilon \preceq \alpha \Rightarrow \omega \preceq \alpha$, and also $\nu \preceq \alpha \Rightarrow \omega \preceq \alpha$, so $K_\alpha(\mathcal{V})$ and $K_\alpha(\mathcal{V}^\epsilon)$ do not depend on ν or ν^ϵ.

Now assume α satisfies (13) and $\alpha_y < \gamma_y$. Then $\omega \preceq \alpha - \sigma$ if and only if $\omega \preceq \alpha - (\sigma \cup \epsilon^\nu)$ whenever $\omega \in \mathcal{V} \smallsetminus \nu$ by the assumption '$\omega_y \geq \nu_y$ whenever $\omega_x = \nu_x$'. If $\alpha_x = \nu_x$ then $\nu^\epsilon \not\preceq \alpha$ and thus $K_\alpha(\mathcal{V}^\epsilon)$ is a cone with vertex ϵ^ν. If $\alpha_x = 1 + \nu_x$ then $\nu \preceq \alpha - \sigma$ if and only if $\nu \preceq \alpha - (\sigma \cup \epsilon^\nu)$, so $K_\alpha(\mathcal{V})$ is another cone with vertex ϵ^ν. Finally, suppose that $\alpha_x = \nu_x$ in addition to (13) and $\alpha_y < \gamma_y$. Then $\omega \preceq \alpha + \epsilon^\nu - \sigma$ if and only if $\omega \preceq \alpha - \sigma$ for $\omega \in \mathcal{V} \smallsetminus \nu$ by the assumption '$\omega_y \geq \nu_y$ whenever $\omega_x = \nu_x$'. And clearly $\nu^\epsilon \preceq \alpha + \epsilon^\nu - \sigma$ if and only if $\nu \preceq \alpha - \sigma$, since $\nu^\epsilon = \nu + \epsilon^\nu$. Therefore $K_{\alpha+\epsilon^\nu}(\mathcal{V}^\epsilon) = K_\alpha(\mathcal{V})$ in this case. □

Lemma 12.2 *Assume the hypotheses and notation from Lemma 12.1, let $\gamma_x = \nu_x$, and let*

$$\gamma_z \quad = \quad \min_{\omega \neq \nu}\{\omega_z \mid \omega_x \leq \nu_x \text{ and } \omega_y < \gamma_y\},$$

which exists because $I_{\mathcal{V}}$ is artinian. Denote by $\mathcal{F}^\epsilon_.$ a minimal free resolution of $I_{\mathcal{V}^\epsilon}$ and by $\mathcal{F}.$ the specialized free resolution of $I_{\mathcal{V}}$ via [GPW00, Theorem 3.3]. (This amounts to the redefine *step applied to any cellular resolution supported on a complex with vertex set $\mathcal{V}^\epsilon$.) Then at most two syzygies of $\mathcal{F}^\epsilon_.$ become nonminimal in $\mathcal{F}.$: a second syzygy s_2^ϵ in degree $\gamma + \epsilon^\nu = (1+\gamma_x, \gamma_y, \gamma_z)$ and a first syzygy s_1^ϵ in degree γ. Choose the unique $\rho \in \mathcal{V}$ with $\rho_y = \gamma_y$ such that $\rho_z < \gamma_z$ is maximal. Then the specializations (s_1, s_2) of $(s_1^\epsilon, s_2^\epsilon)$ are nonminimal if and only if $\rho_x = \gamma_x$.*

Proof. The only possible nonminimal summands of $\mathcal{F}.$ occur in degrees $(\nu_x, \gamma_y, \alpha_z)$ or $(\nu_x^\epsilon, \gamma_y, \alpha_z)$ for some value of $\alpha_z \geq \nu_z$, because the other Betti numbers of $I_{\mathcal{V}}$ and $I_{\mathcal{V}^\epsilon}$ are in bijection by Lemma 12.1. Furthermore, nonminimal summands cannot come from zeroth syzygies of $I_{\mathcal{V}^\epsilon}$, since these are in bijection with those of $I_{\mathcal{V}}$ (no elements of $\mathcal{V}^\epsilon$ disappear when the ϵ is removed). Therefore, nonminimal syzygies in $\mathcal{F}.$ can only be first or second syzygies.

It is a general fact about nonminimal free resolutions that nonminimal summands come in pairs (s_1, s_2) consisting of a first and second syzygy. In the present case, such pairs arise from minimal first and second syzygies $(s_1^\epsilon, s_2^\epsilon)$ of $I_{\mathcal{V}^\epsilon}$. Since $\deg(s_1) = \deg(s_2)$ but $\deg(s_1^\epsilon) \neq \deg(s_2^\epsilon)$, and the only change is occurring in the x-direction, it must be that $\deg(s_2^\epsilon) = \epsilon^\nu + \deg(s_1^\epsilon)$. Lemma 12.1 therefore implies that any minimal syzygy s_2^ϵ becoming nonminimal in $\mathcal{F}$. must have $\deg(s_2^\epsilon)$ along the vertical ray $(\nu_x^\epsilon, \gamma_y, \alpha_z)$ for varying $\alpha_z \geq \nu_z$. Furthermore, there can be at most one value γ_z for α_z, since there can be only one second syzygy along any line parallel to an axis. This proves all but the last sentence.

The two specialized syzygies are nonminimal if and only if either one of them is. The specialization of s_2^ϵ is a second syzygy in degree γ that is minimal if and only if $K_\gamma(\mathcal{V})$ is the boundary of a triangle by Proposition 6.1. In any case, $\gamma-(1,1,1) \notin \langle\mathcal{V}\rangle$ by minimality of γ_z. Suppose $\rho_x \neq \gamma_x$. Then $(0,1,1) \in K_\gamma(\mathcal{V})$ because $\gamma_z > \nu_z$ and $\gamma_y > \nu_y$. Any vector $\omega \in \mathcal{V}$ whose z-coordinate was used to define γ_z has $\omega_x < \nu_x$ by the assumption '$\omega_y \geq \nu_y$ whenever $\omega_x = \nu_x$'; thus $(1,1,0) \in K_\gamma(\mathcal{V})$. And $(1,0,1) \in K_\gamma(\mathcal{V})$ because of ρ. Therefore, $K_\gamma(\mathcal{V})$ is the boundary of the triangle when $\rho_x \neq \gamma_x$ (in this case, there is no first syzygy of $I_{\mathcal{V}^\epsilon}$ in degree γ waiting to cancel s_2^ϵ as it specializes to s_2). Finally, if $\rho_x = \gamma_x$, then $(1,0,1) \notin K_\gamma(\mathcal{V})$, whence $K_\gamma(\mathcal{V})$ cannot be the boundary of a triangle. □

Example 12.3 Some possible combinatorial types for $K_\alpha(\mathcal{V})$, where $\alpha = \deg(s_1)$ is the degree of the specialized first syzygy of Lemma 12.2, are depicted in Figure 8. The headings 'else' and 'then' correspond to the cases in Algorithm 11.2. Observe that in the single then case, the white dot s_2^ϵ at $(1+\gamma_x, \gamma_y, \gamma_z)$ gets smashed into the vertical plane during specialization and cancels the black dot s_1^ϵ at $(\gamma_x, \gamma_y, \gamma_z)$. On the other hand, the topology remains constant in the first two else cases. In the final else case, two of the black dots merge to become a "double" black dot, since the resulting Koszul simplicial complex (3 disjoint vertices), has 2-dimensional $\widetilde{H}_0$ after the wall is pushed back. □

Proposition 12.4 *At every iteration of the line* end while-do *in Algorithm 11.2, the labeled map M provides a* minimal *cellular free resolution of $I_{\mathcal{V}^\epsilon}$.*

Proof. This has two parts, of course: then and else. Both follow from Lemma 12.2, given Remark 11.3(vii). Indeed, removing the unique nonminimal edge automatically destroys the unique nonminimal region by merging it with an adjacent region.

This argument implicitly uses Proposition 7.2, which guarantees that the deleted edge equals the entire intersection of the two regions containing it, so that matrices for the ordinary boundary complex of the deletion are obtained from those for M by removing the appropriate rows and columns. It should also be reiterated that we can choose the deformation in Lemma 12.1 to be the one occurring in each pass of the while loop; indeed it is here that the precise condition (12) on the deformation ϵ in Algorithm 11.2 is used in an essential way. □

Proof of Theorem 11.1. It remains only to reduce to the artinian case. Let $\overline{\mathcal{V}} = \mathcal{V} \cup \{\dot{u} \mid \mathcal{S}_{\mathcal{V}} \cap U = \varnothing\}$ for sufficiently large $|\dot{u}|$, as in Lemma 8.2. Given any labeled cell complex $\overline{M}$ supporting a minimal cellular resolution of $I_{\overline{\mathcal{V}}}$, taking the subcomplex $M \subseteq \overline{M}$ whose labels precede the join $\vee\mathcal{V}\}$ produces a cell complex supporting a minimal resolution of M. This is the content of [BS98, Corollary 1.3]. □

Part III
Planar maps revisited

13 Orthogonal coloring

Let M be a planar map with vertex set $\mathcal{V}$ and M_∞ an extended map. Since M_∞ is embedded in a surface S homeomorphic to the plane, it makes sense to order the angles at any of its vertices cyclically, and to say that a list of angles at a vertex is **consecutive**. The same comment applies as well to the angles in any bounded region of M_∞, to the four angles having any fixed bounded edge as a leg (two at each vertex), and to the unbounded edges (read as the hands on an analog clock). With these cyclically ordered sets in mind, let A be any finite set of objects arranged cyclically in the surface S. Given three **colors** x, y, z, the set A is **trichromatic** if

- there is an element in A colored u, for each $u = x, y, z$;
- the elements in A colored u are consecutive, for each $u = x, y, z$; and
- the block of elements colored z is immediately counterclockwise from the block of elements colored y.

Deleting 'counter' from the last item defines **clockwise trichromatic** instead.

An **orthogonal coloring** $\mathcal{O}$ of M_∞ is a labeling of the angles in M_∞ at every vertex in M by three **colors** x, y, z such that

(i) all vertices of M are trichromatic in M_∞;
(ii) all bounded edges of M are clockwise trichromatic in M_∞;
(iii) all bounded regions of M are trichromatic in M_∞;
(iv) the two angles adjacent to each unbounded edge have different colors; and
(v) attaching to each unbounded edge the color missing from its two angles makes the set of unbounded edges in M_∞ trichromatic.

Suppose, in addition, that M has axial vertices $\dot{x}, \dot{y}, \dot{z}$, so $M_\infty = M_\infty(\dot{x}, \dot{y}, \dot{z})$. The angles interior to the three unbounded regions of M_∞ are called **exterior angles** of M_∞; besides the obvious pair of angles at each of $\dot{x}, \dot{y}, \dot{z}$, they include one angle at each nonaxial vertex lying on the exterior cycle. The **interior angles** at vertices on the exterior cycle are the angles lying inside M—that is, the non-exterior angles. Call an orthogonal coloring of M **axial** if the following boundary conditions, which imply axioms (iv) and (v), are satisfied for $u = x, y, z$:

(iv)′ all interior angles at the axial vertex $\dot{u}$ are colored u;

(v)′ all exterior angles of M_∞ in the unbounded region *not* touching $\dot{u}$ are colored u.

For example, the orthogonal coloring in Figure 1 is axial. Roughly speaking, even axioms (iv)′ and (v)′ say that something is trichromatic: (iv)′ says that the exterior cycle, thought of as a triangle with vertices $\dot{x}, \dot{y}, \dot{z}$, is trichromatic; while (v)′ says that the set of exterior angles is trichromatic. This observation makes it particularly easy to remember the axioms defining an orthogonal coloring.
Felsner independently defined what he called **Schnyder colorings** in [Fel01, Section 1]. These are the *same thing* as axial orthogonal colorings. Felsner also pointed out that axiom (ii) follows from the others [Fel01, Lemma 1]. When the planar map is a triangulation of a simplex with no new vertices on the boundary, axial orthogonal coloring reduces to the angle-coloring of Schnyder [Tro92, Theorem 6.2.1], justifying Felsner's terminology. The next result and its proof justify the adjective "orthogonal".

Proposition 13.1 *Any (axial) geodesic embedding $M \hookrightarrow \mathcal{S}_\mathcal{V}$ induces an (axial) orthogonal coloring of $M_\infty = M \cup$ (unbounded orthogonal rays).*

Proof. Any angle in a geodesic embedding, whether between bounded or unbounded edges or both, locally lies in a plane u = constant for some $u \in \{x, y, z\}$. Color such an angle by u. The orthogonality axioms follow readily from the definition of axial geodesic embedding and the region axiom (which holds for nonaxial grid surfaces by Lemma 8.2) in Section 2. □

14 Orthogonal flows

In this section the planar map M has axial vertices $\dot{x}, \dot{y}, \dot{z}$, and $M_\infty(\dot{x}, \dot{y}, \dot{z})$ is orthogonally colored. We derive properties of orthogonal flows in grid surfaces (Section 2) from the axioms for axial orthogonal coloring, for comparison with [BT93]. In an earlier version of this paper, the current section was intended to serve as a possible proof method for Conjecture 16.3, below. Felsner in fact carried out this program [Fel02], having independently found the results in this section already [Fel01].
To begin, interpret an orthogonal coloring as a family of three **orthogonal vector fields** on M_∞: for each vertex of M_∞, assign precisely three arrows pointing away from it—one of each color—along the edges separating the blocks of differently colored angles. Thus, for example, the z-colored arrows point upward along an edge if and only if the angles around the edge have colors ${}^{z}_{y}|^{z}_{x}$, ${}^{y}_{y}|^{z}_{x}$, ${}^{z}_{y}|^{x}_{x}$, or ${}_{y}|_{x}$ (the z-axis). The first of these edges has no arrow pointing downward from its top vertex, while the second and third have downward arrows colored x and y, respectively.
The u-colored vector fields for $u = x, y, z$ can be "integrated" to get **orthogonal flow lines**: the u-colored flow line from $\nu \in M$ is a directed path in M_∞, beginning with ν, that is a union of edges underlying u-colored arrows. Thus

the next vertex after ν is at the other end of the edge whose u-colored arrow points away from ν.

Orthogonal flow lines can only meet in certain orientations. To make a precise statement, let L be a directed path passing through a vertex ν, and K a directed path containing an edge $\vec{e}$ pointing toward ν. Then K **approaches** L **from the left** at ν if $\vec{e}$ is distinct from L's two arrows at ν, and these three arrows are oriented as in Eq. (14), below (ignoring the labels for the moment). Similarly, given the mirror orientations, K approaches L from the right.

Lemma 14.1 *A flow line colored x never approaches a flow line colored z from the left. A flow line colored y never approaches a flow line colored z from the right. These statements remain true for* cyclic *permutations of x, y, z.*

Proof. Suppose $\vec{e}$ approaches the z-colored flow line L from the left at ν. The angle coloration around ν looks like the following diagram,

$$\begin{array}{ccc}
 & & L \\
 & & \uparrow \\
 & & {}_{y}\,|\,{}_{x} \\
{}^{c}_{b} & \xrightarrow{\ \vec{e}\ } & \nu \\
 & & {}_{a}\uparrow \\
 & & {}_{y}\,|\,{}_{x}
\end{array} \tag{14}$$

in which $a \neq x$ by edge trichromatics. (There may be other edges containing ν but not shown.) Since none of the angles going clockwise between a and y at ν can be x-colored (by vertex trichromatics), it is impossible to have $\{b, c\} \subseteq \{y, z\}$, by edge trichromatics. The other case is similar, and the symmetry is obvious. $\square$

The observation in Lemma 14.1 imposes useful conditions on flow lines. The next proposition says that orthogonal flow lines satisfy conditions slightly stronger than the five "path properties" defining a **normal family of paths**, as introduced by Brightwell and Trotter [BT93] (see [Tro92, Chapter 6] for an exposition). A set of paths whose pairwise intersections consist of a single vertex is called **independent**.

Proposition 14.2 (Path Properties) *Endow $M_\infty(\dot{x}, \dot{y}, \dot{z})$ with an axial orthogonal coloring. Suppose $\nu \in M$ is a vertex and $u \in \{x, y, z\}$ is a color.*

1. *There is a unique u-colored flow line beginning at ν; it connects ν to ∞ via $\dot{u}$.*
2. *If $[\nu, \dot{u}]$ denotes the part of the u-colored flow line starting with ν and ending with $\dot{u}$, then $[\nu, \dot{x}]$, $[\nu, \dot{y}]$, and $[\nu, \dot{z}]$ are independent paths from ν.*
3. *The six paths $[\dot{x}, \dot{y}], [\dot{y}, \dot{x}], [\dot{y}, \dot{z}], [\dot{z}, \dot{y}], [\dot{x}, \dot{z}], [\dot{z}, \dot{x}]$ are on the exterior cycle of M.*
4. *If $\omega \in M$ is a vertex such that $\omega \in [\nu, \dot{u}]$, then $[\omega, \dot{u}] \subseteq [\nu, \dot{u}]$.*

5. *If ω is in the union $[\dot{x}, \nu, \dot{y}]$ of $[\nu, \dot{x}]$, $[\nu, \dot{y}]$, $[\dot{x}, \dot{y}]$, and any regions they enclose, then $[\dot{x}, \omega, \dot{y}] \subseteq [\dot{x}, \nu, \dot{y}]$; the same holds with $\dot{z}$ in place of $\dot{x}$ or $\dot{y}$.*

Proof. Existence and uniqueness in part 1 are obvious. Since the only arrows in M_∞ pointing out of M are on the axes by orthogonality axioms (iv)′ and (v)′, part 1 is equivalent to statement that flow lines contain no cycles. Any flow cycle C containing a vertex in the interior of the region it bounds contains an entire flow cycle of some other color in its interior: one of the other colors cannot escape by Lemma 14.1. Thus we may assume C has no vertices interior to it. But then the argument in the proof of Lemma 14.1 forces the coloration of the angles in the interior of C to omit one color entirely, violating region trichromatics.

Now suppose two flow lines from ν—colored x and y, say—intersect, and consider the cycle C formed by their arcs connecting ν to their first intersection point. Assume for the moment that ω lies interior to C. Depending on the orientations of the two flow lines around C, either z-flow lines cannot escape C, or the u-flow line from ω exits through the u-colored arc of C, for $u = x, y$. The first case contradicts acyclicity, while the second case produces a smaller cycle C. Again, we may therefore assume C contains no vertices interior to it. In the first orientation, no interior angle of C is colored z, while in the second orientation, all interior angles of C are colored z.

Part 3 follows easily by applying orthogonality axiom (ii) to the edges on the exterior cycle of M, each of which has two of its four colors specified by (v)′.

Part 4 is obvious from the definition of flow line.

To prove part 5, first observe that a flow line colored x or y cannot escape $[\dot{x}, \nu, \dot{y}]$ if it originates at a vertex $\omega \in [\dot{x}, \nu, \dot{y}]$ that is on neither $[\nu, \dot{x}]$ nor $[\nu, \dot{y}]$. Indeed, if $[\omega, \dot{x}]$ intersects $[\nu, \dot{x}]$, then these two flow lines agree thereafter; and $[\omega, \dot{x}]$ cannot even approach $[\nu, \dot{y}]$, thanks to Lemma 14.1. On the other hand, if $\nu \neq \omega \in [\nu, \dot{y}]$, say, then $[\omega, \dot{y}] \subset [\nu, \dot{y}]$ by part 4. Moreover, vertex trichromatics force the first edge of $[\omega, \dot{x}]$ to exit ω clockwise from the y-colored arrow pointing away ω, but counterclockwise from the other edge leaving ω and in $[\nu, \dot{y}]$. Therefore, $[\omega, \dot{x}]$ remains inside $[\dot{x}, \nu, \dot{y}]$ either by part 4 or the first sentence of this paragraph. □

Example 14.3 Orthogonal flow lines are better behaved than arbitrary normal families of paths, since their strong local properties imply Path Property 2, and especially the crucial Path Property 5, which are global. Figure 9 depicts two triples of vector fields determined by trichromatic angle colorings whose corresponding flow lines satisfy the conclusion of Proposition 14.2, and therefore constitute normal families of paths. The vector fields in the left diagram are not orthogonal because of the edge connecting the rightmost interior vertices, the edge connecting the remaining interior vertex to $\dot{x}$, the interior region, and the bottom region bordering $[\dot{x}, \dot{z}]$. Recoloring the angles at the leftmost interior vertices yields the orthogonal vector fields at right. □

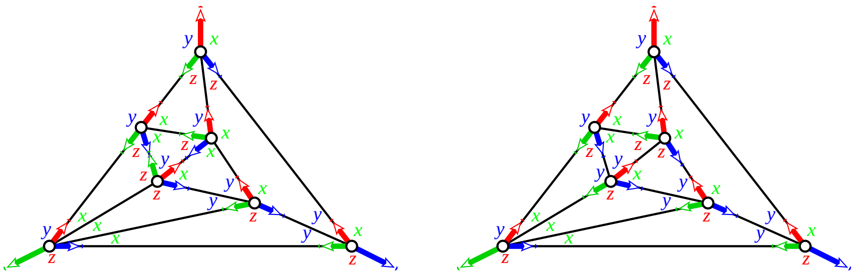


Figure 9: Nonorthogonal and orthogonal vector fields

Corollary 14.4 *Let M be a planar with axial vertices $\dot{x}, \dot{y}, \dot{z}$. The extended map $M_\infty(\dot{x}, \dot{y}, \dot{z})$ can be orthogonally colored if and only if it is triconnected.*

Proof. Proposition 14.2.2 proves the 'only if' direction, while Theorem 5.1 and Proposition 13.1 prove the 'if' direction. □

Remark 14.5 The results of this section can be massaged to work for nonaxial orthogonally colored extended maps M_∞, as well. In fact, everything reduces to the axial case: draw a large triangle containing all of M—not M_∞—in its interior, and call its three vertices $\dot{x}, \dot{y}, \dot{z}$, in counterclockwise order. Then connect each u-colored unbounded edge to $\dot{u}$. It is straightforward to verify the axioms for axial orthogonal colorings, given the ordinary axioms for M_∞. Observe the analogy with Lemma 8.2.

15 Duality for geodesic embeddings

Let G be a graph embedded in the sphere $S \cup \{\infty\}$, where $S \cong \mathbb{R}^2$, and assume that $G \cap S$ is a planar map or extended map. For this paragraph only, we allow graphs and planar maps to have multiple edges, although we assume G has no loops, and that the edges of G in each one of its regions form a simple cycle. Define the **spherical dual** $\hat{G}$ of G as usual: place a vertex $\hat{A}$ in each region A of G, and draw an edge connecting $\hat{A}$ to $\hat{B}$ through each edge contained in $A \cap B$. Then $\hat{G}$ also satisfies the no-loop and simple-cycle conditions, which are dual to each other. Assume that ∞ is a vertex of either G or $\hat{G}$. When $\infty \notin G$, so G is a planar map $M \subset S$, then $\hat{G}$ is an extended map that we denote by $\hat{M}_\infty$ (with associated planar map $\mathrm{del}(\hat{G}; \infty) = \hat{M}$) and call the **planar dual** of M. When $\infty \in G$, so $G \cap S = N_\infty$ is an extended map satisfying $N = \mathrm{del}(G; \infty)$, then $\hat{G}$ is a planar map that we denote by $\hat{N}$ and call the **planar dual** of N_∞. For an arbitrary grid surface $\mathcal{S}_\mathcal{V}$, let $\hat{\alpha}$ be any vector preceded by $\mathbf{1} +$ the join of the vectors in $\mathcal{V}$, where $\mathbf{1} = (1, 1, 1)$. Define $\overline{\mathcal{V}}$ by throwing in axial vertices missing from $\mathcal{V}$:

$$\overline{\mathcal{V}} \quad = \quad \mathcal{V} \cup \{\dot{u} \mid \mathcal{S}_\mathcal{V} \cap U = \varnothing\}, \quad \text{where} \quad \dot{u} \in U \text{ has length } |\dot{u}| = \hat{\alpha}_u.$$

This notation agrees with that in Lemma 8.2, but specifies the length of $\dot{u}$. Define the **Alexander dual** grid surface $\mathcal{S}_{\hat{\mathcal{V}}}$ by

$$\hat{\mathcal{V}} \quad = \quad \{\hat{\alpha} - \rho \mid \rho \in \mathcal{S}_{\overline{\mathcal{V}}}^{\max}\},$$

where the notation comes from Lemma 8.3. Neither $\mathcal{S}_{\hat{\mathcal{V}}}$ nor $\mathcal{S}_{\overline{\mathcal{V}}}$ depends combinatorially on $\hat{\alpha}$, in the sense of Lemma 2.1.

By definition, $\overline{\mathcal{V}} = \mathcal{V}$ and $\hat{\alpha} \succeq \dot{x} + \dot{y} + \dot{z} + \mathbf{1}$ when $\mathcal{S}_{\mathcal{V}}$ is axial. Define $\mathcal{S}_{\mathcal{V}}$ to be **radial** if $\nu_u \neq 0$ for all $\nu \in \mathcal{V}$ and $u \in \{x, y, z\}$. In particular, $\overline{\mathcal{V}} = \mathcal{V} \cup \{\dot{x}, \dot{y}, \dot{z}\}$ is a disjoint union when $\mathcal{S}_{\mathcal{V}}$ is radial. Although the following duality theorem is stated only for axial and radial grid surfaces, similar (but slightly harder to state) considerations apply to arbitrary geodesic and rigid embeddings; the definition of Alexander dual grid surface extends verbatim. Recall that $M_\infty = M \cup$ (unbounded orthogonal rays in $\mathcal{S}_{\mathcal{V}}$) for geodesic embeddings $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$.

Theorem 15.1 *Let $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ be an axial geodesic embedding, and $N \hookrightarrow \mathcal{S}_{\mathcal{W}}$ a radial geodesic embedding.*

1. *$M \cong \hat{N}$ if and only if $\hat{M} \cong N$.*
2. *There is a natural radial geodesic embedding $\hat{M} \hookrightarrow \mathcal{S}_{\hat{\mathcal{V}}}$ that is rigid if and only if $M \hookrightarrow \mathcal{S}_{\mathcal{V}}$ is rigid.*
3. *There is a natural axial geodesic embedding $\hat{N} \hookrightarrow \mathcal{S}_{\hat{\mathcal{W}}}$ that is rigid if and only if $N \hookrightarrow \mathcal{S}_{\mathcal{W}}$ is rigid.*

Proof. Part 1 is simply duality for spherical maps, as in the first paragraph of this section. We prove part 2, since part 3 is similar and can even be simplified by using parts 1 and 2. The vertex axiom for $\mathcal{S}_{\hat{\mathcal{V}}}$ follows immediately from Lemma 8.3. The main point for the rest of the proof is that

$$\{\sigma \in \mathcal{S}_{\hat{\mathcal{V}}} \mid \sigma \preceq \hat{\alpha}\} \quad = \quad \{\hat{\alpha} - \sigma \mid \sigma \in \mathcal{S}_{\mathcal{V}} \text{ and } \sigma \preceq \hat{\alpha}\}.$$

In other words, $\{\sigma \in \mathcal{S}_{\mathcal{V}} \mid \sigma \preceq \hat{\alpha}\}$ lies in the topological boundary of $\mathbb{R}^3 \smallsetminus \langle \mathcal{V} \rangle$. Every elbow geodesic $[\nu, \omega]$ in M borders precisely two regions of M_∞ because M_∞ is triconnected (apply Proposition 7.2 along with Theorem 8.4, or Proposition 13.1 along with Corollary 14.4). If these two regions A and B are bounded, then the maximal points $\hat{A}$ and $\hat{B}$ in them (Lemma 8.3) connect via straight line segments in $\mathcal{S}_{\mathcal{V}}$ to $\nu \vee \omega$ by the region axiom (see the proof of Lemma 8.3). Furthermore, one of these segments must transform into an orthogonal ray in $\mathcal{S}_{\hat{\mathcal{V}}}$ via $\sigma \mapsto \hat{\alpha} - \sigma$, because $\nu \vee \omega = \hat{A} \wedge \hat{B}$ (by the region axiom: for each $u \in \{x, y, z\}$ one of the two vectors ν and ω shares its u-coordinate with one of the vectors $\hat{A}$ and $\hat{B}$). Thus every bounded edge of $\hat{M}_\infty$ is an elbow geodesic in $\mathcal{S}_{\hat{\mathcal{V}}}$.

If $[\nu, \omega]$ borders a bounded region A and an unbounded region, then $\nu_u = \omega_u = 0$ for some $u \in \{x, y, z\}$. The vector $\nu \vee \omega$ shares both of its nonzero coordinates with $\hat{A}$ in this case, so $\hat{A}$ is the unique maximal point of $\mathcal{S}_{\mathcal{V}}$ preceded by $\nu \vee \omega$. This forces the negative ray $\hat{A} - U$ passing through $\hat{A}$ and $\nu \vee \omega$ to transform

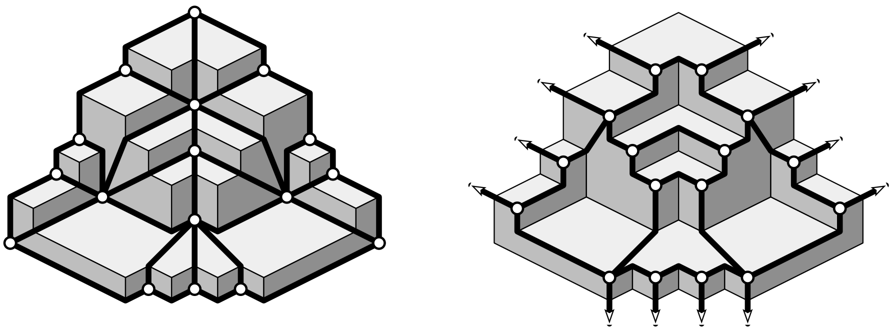


Figure 10: Duality for geodesic embeddings

into an unbounded orthogonal ray $\hat{\alpha} - (\hat{A} - U) = (\hat{\alpha} - \hat{A}) + U$ in $\mathcal{S}_{\hat{\mathcal{V}}}$. Thus the unbounded edges of $\hat{M}_\infty$ are unbounded orthogonal rays in $\mathcal{S}_{\hat{\mathcal{V}}}$.

Now we show that every orthogonal ray in $\mathcal{S}_{\hat{\mathcal{V}}}$ is contained in an edge of $\hat{M}$. Equivalently, for each maximal point $\hat{A} \in \mathcal{S}_{\mathcal{V}}$ and $U \in \{X, Y, Z\}$, there is an elbow geodesic $[\nu, \omega] \in M$ such that $\hat{A} - \nu\vee\omega \in U$. By symmetry, set $U = Z$. Choose the edge $[\nu, \omega] \subset A$ so that $\nu_x = \hat{A}_x$ and $\omega_y = \hat{A}_y$; such an edge exists by orthogonality property (iii) and Proposition 13.1.

For the statement about rigidity, use the rigid region axiom: $\nu\vee\omega \preceq \hat{A} \Leftrightarrow \nu \preceq \hat{A}$ and $\omega \preceq \hat{A}$, and this occurs precisely when $\nu \in A$ and $\omega \in A$. Thus, when $[\nu, \omega] \subset \mathcal{S}_{\mathcal{V}}$ is a rigid geodesic, $\nu\vee\omega$ precedes the maximal points in exactly two regions of M: the regions A and B containing both ν and ω. That $\nu\vee\omega = \hat{A}\wedge\hat{B}$ was shown above. □

Example 15.2 Figure 10 illustrates Theorem 15.1 for a particular (nonrigid) geodesic embedding. Turning the picture upside-down yields two pictures of the Alexander dual grid surface, with the radial embedding appearing the right way out and the axial embedding backwards. □

Considering the Theorem 8.4, the duality result here essentially falls under the umbrella of duality for resolutions of monomial ideals [Mil00a, Section 4.2]. Although the above proof can be simplified greatly by applying duality for resolutions, particularly in concert with [Mil00a, Proposition 3.20], it seemed appropriate to keep Theorem 15.1 as self-contained as possible.

16 Open problems

16.1 Planar maps supporting resolutions of nonartinian ideals

Problem 16.1 *Characterize those planar maps that support minimal resolutions of trivariate monomial ideals.*

It would be nice to clean up the statement of Theorem B by proving $3 \Rightarrow 2$. Unfortunately, the most direct proof attempt fails.

To be precise, suppose the planar map N supports a minimal cellular free resolution of some ideal $I_{\mathcal{V}}$. Following the procedure in the proof of Proposition 7.2, the underlying graph H of N has paths analogous to orthogonal flows. However, since $\mathcal{V}$ may not be axial, these flows do not reach axial vertices, but instead reach unbounded orthogonal rays. As an abstract graph, we can define $\overline{H}$ by adding three new vertices $\dot{x}, \dot{y}, \dot{z}$ to H, and then connecting each vertex $\nu \in N$ to $\dot{u}$ if the orthogonal ray U_ν is unbounded. Judging from Theorems A and B, one might expect that $\overline{H}$ is planar and triconnected, supporting a minimal free resolution of some artinian approximation to $I_{\mathcal{V}}$. But although $\overline{H}$ is obviously triconnected, it need not be planar!

Take, for instance, the set $\mathcal{V}$ from Example 9.2, without the axial vectors. Deleting $\dot{x}, \dot{y}, \dot{z}$ from the nonplanar map M there yields a planar map N supporting a minimal free resolution of $I_{\mathcal{V}}$, but reconnecting to $\dot{x}, \dot{y}, \dot{z}$ as above returns M again.

16.2 Nonrigid geodesic embeddings

Conjecture 16.2 *Let the axial grid surface $\mathcal{S}_{\mathcal{V}}$ have vertices $\nu_1, \nu_2, \nu_3, \nu_4 \in \mathcal{V}$, each with no coordinate zero, such that $[\nu_i, \nu_j]$ is an elbow geodesic whenever $i \neq j$. Then the ideal $I_{\mathcal{V}}$ possesses a nonplanar cellular minimal free resolution.*

Thus relatively minor violation of rigidity for $\mathcal{S}_{\mathcal{V}}$ not only implies nonuniqueness of minimal cellular resolutions (cf. Corollary 9.1), but should even imply nonplanarity. Intuitively, it should be possible to construct K_5 out of elbow geodesics (two of which cross) using two of the ν_i and $\dot{x}, \dot{y}, \dot{z}$, as in Example 9.2. Given such a configuration, one is tempted to "fill in" the resulting 1-skeleton to form a cell complex minimally resolving $I_{\mathcal{V}}$. This is, in fact, how Example 9.2 was constructed.

16.3 Orthogonal coloring to rigid embedding

Conjecture 16.3 (Felsner's Theorem [Fel02]) *Every orthogonal coloring on a planar map is induced by a rigid embedding, as in Proposition 13.1.*

This converse to Proposition 13.1 reduces easily to the axial case, via Lemma 8.2 and Remark 14.5. Since the axial case was proved by Felsner [Fel02] in response to seeing the conjecture here, this is in fact no longer an open problem. But see [Fel02] for Felsner's open questions regarding the set of orthogonal colorings.

Besides its applications to the questions discussed in the next subsection, the motivation behind formulating the above statement was that it reduces Theorem 5.1 to verifying that every planar map there has an orthogonal coloring. This actually follows by the same induction used for rigid embeddings, but the details become much simpler; moreover, Felsner already proved existence of orthogonal colorings in [Fel01]. Thus we get a substantially more palatable proof of Theorem 5.1 via [Fel01, Fel02].

16.4 The Scarf stratification

Part of the motivation for the results presented here stems from the desire to understand not just how to assign a minimal free resolution to any particular monomial ideal, but to understand the collection of all minimal free or injective resolutions of monomial ideals. Some recent results, notably those in [GPW00], aim to classify monomial ideals according to whether their minimal resolutions are in some sense isomorphic, by ascertaining what data determines the minimal resolutions. Other results, such as those in [MSY00], raise the question of which deformations (as in Section 10) preserve minimal resolutions. This latter idea originated from [BPS98, PS98], which was in turn based upon work of H. Scarf on related classifications for integer programs.

It seems that the most universal approach for minimal resolutions of monomial ideals should combine the two types of classification above. Heuristically, the question becomes, 'What are all possible ways of deforming continuously between monomial ideals having "nearby" isomorphism classes of resolutions?' Ideally one would like a 'fine moduli space' for minimal resolutions, in the sense that arcs in that space correspond to families of monomial ideals whose minimal resolutions deform continuously. Of course, there are only finitely many deformation classes of minimal resolutions, so the space should interact well with the poset of monomial ideals under deformation.

One tempting candidate for 'fine moduli space' can be defined as follows. Giving r monomials in n variables is the same as giving an $r\times n$ matrix of exponents. The **generic** monomial ideals partition an open dense subset of the nonnegative orthant in this matrix space, by [BPS98, MSY00]. Taking the cells formed by intersections of the closures of the generic loci yields a decomposition that we propose to call the **Scarf stratification**. It should be a rational polyhedral fan, if life is fair; but at least there should be a subdividing fan with finitely many maximal cones such that the maximal Scarf cells are unions of maximal cones.

Any classifying space such as the Scarf stratification will feel rather more like an algebraic stack than a fine moduli space, because even if it classifies deformations of minimal resolutions, the actual *set* of isomorphism classes of minimal resolutions would be a quotient by some finite group (containing the symmetric group on the variables, at least) of the set of strata: two monomial ideals differing by a permutation of the variables might be far from each other in the stratification.

Whatever the correct space of minimal resolutions ends up being, the methods introduced here can be applied to elucidate its combinatorial structure, for $n = 3$. Note that Felsner's theorem (Conjecture 16.3) classifies the maximal strata for the case of three variables—that is, the generic trivariate monomial ideals—by [MSY00]:

Corollary 16.4 *Deformation classes of generic artinian trivariate monomial ideals correspond bijectively to orthogonally colored axial planar triangulations.*

The triangulations referred to here are orthogonally colored triangulations of the simplex *with new vertices on the boundary*, so they are not Schnyder normal colorings. Corollary 16.4 and Corollary 9.1 prompt the following:

Question 16.5 *In terms of monomial ideals, what do deformation classes of axial rigid embeddings correspond to in general, for non-triangulations?*

For instance, do they correspond to certain Scarf strata? If the Scarf stratification is a rational polyhedral fan, what linear equations define these cones? We note that arbitrary geodesic embeddings have considerably more freedom than do rigid embeddings, from the point of view of deformations (hence the adjective 'rigid'). The possible application of the material in this paper to the classification of deformations of minimal resolutions was one of the motivations for stating as many results as possible in the context of arbitrary geodesic embeddings.

Our final remark concerns the bias in this paper toward artinian monomial ideals. Combinatorial considerations such as triconnectivity notwithstanding, the bias also makes sense algebraically. Briefly, the homological characterization of genericity for arbitrary monomial ideals [MSY00, Theorem 1.5 and Remark 1.7] is a statement about graded injective resolutions under deformation; and any result concerning $\mathbb{Z}^3$-graded injective resolutions of arbitrary monomial ideals has an equivalent statement in terms of $\mathbb{Z}^3$-graded free resolutions of artinian monomial ideals, by the duality results in [Mil00a]. We emphasize that the theory surrounding injective resolutions played a crucial role in properly formulating the graph-theoretic results in Part I, as well as the algebraic results in Part II (the exposition in Sections 10–12 is based on [Mil00b, Theorem 5.60]). Moreover, concentrating on injective resolutions (equivalently, artinian monomial ideals) should ease the nonplanarity difficulties raised in Section 16.1, by applying [Mil00a, Theorem 4.5.5 and Example 4.8.5], which says how to recover free resolutions from injective resolutions.

References

[BPS98] Dave Bayer, Irena Peeva, and Bernd Sturmfels, *Monomial resolutions*, Math. Res. Lett. 5 (1998), no. 1-2, 31–46.

[BS98] Dave Bayer and Bernd Sturmfels, *Cellular resolutions of monomial modules*, J. Reine Angew. Math. 502 (1998), 123–140.

[BT93] Graham Brightwell and William T. Trotter, *The order dimension of convex polytopes*, SIAM J. Discrete Math. 6 (1993), no. 2, 230–245.

[Fel01] Stefan Felsner, *Convex drawings of planar graphs and the order dimension of 3-polytopes*, Order 18 (2001), no. 1, 19–37.

[Fel02] Stefan Felsner, *Geodesic embeddings and planar graphs*, Preprint, 2002.

[GPW00] Vesselin Gasharov, Irena Peeva, and Volkmar Welker, *The LCM-lattice in monomial resolutions*, Math. Res. Lett. 6 (1999), no. 5–6, 521–532.

[Hoc77] Melvin Hochster, *Cohen-Macaulay rings, combinatorics, and simplicial complexes*, Ring theory, II (Proc. Second Conf., Univ. Oklahoma, Norman, Okla., 1975) (B. R. McDonald and R. Morris, eds.), Lect. Notes in Pure and Appl. Math., no. 26, Dekker, New York, 1977, pp. 171–223. Lecture Notes in Pure and Appl. Math., Vol. 26.

[Mil00a] Ezra Miller, *The Alexander duality functors and local duality with monomial support*, J. Algebra 231 (2000), 180–234.

[Mil00b] Ezra Miller, *Resolutions and duality for monomial ideals*, Ph.D. thesis, University of California at Berkeley, 2000.

[MS99] Ezra Miller and Bernd Sturmfels, *Monomial ideals and planar graphs*, Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (M. Fossorier, H. Imai, S. Lin, and A. Poli, eds.), Springer Lecture Notes in Computer Science, no. 1719, Springer Verlag, 1999, Proceedings of AAECC-13 (Honolulu, Nov. 1999), pp. 19–28.

[MSY00] Ezra Miller, Bernd Sturmfels, and Kohji Yanagawa, *Generic and cogeneric monomial ideals*, J. Symbolic Comp. 29 (2000), 691–708.

[PS98] Irena Peeva and Bernd Sturmfels, *Generic lattice ideals*, J. Amer. Math. Soc. 11 (1998), no. 2, 363–373.

[Roz70] I. Z. Rozenknop, *Polynomial ideals that are generated by monomials*, Moskov. Oblast. Ped. Inst. Učen. Zap. 282 (1970), 151–159, (in Russian).

[RW01] Victor Reiner and Volkmar Welker, *On the linear syzygies of a Stanley–Reisner ideal*, Math. Scand. 89 (2001), no. 1, 117–132.

[Tro92] William T. Trotter, *Combinatorics and partially prdered sets: Dimension theory*, Johns Hopkins University Press, Baltimore, MD, 1992.

[Zie95] Günter M. Ziegler, *Lectures on polytopes*, Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995.

Ezra Miller
Department of Mathematics
M.I.T.
Cambridge, Massachusetts
ezra@math.mit.edu

Complex Structure on the Smooth Dual of $GL(n)$

Jacek Brodzki[1] and Roger Plymen

Received: December 10, 2001
Revised: April 24, 2002

Communicated by Peter Schneider

Abstract. Let G denote the p-adic group $GL(n)$, let $\Pi(G)$ denote the smooth dual of G, let $\Pi(\Omega)$ denote a Bernstein component of $\Pi(G)$ and let $\mathcal{H}(\Omega)$ denote a Bernstein ideal in the Hecke algebra $\mathcal{H}(G)$. With the aid of Langlands parameters, we equip $\Pi(\Omega)$ with the structure of complex algebraic variety, and prove that the periodic cyclic homology of $\mathcal{H}(\Omega)$ is isomorphic to the de Rham cohomology of $\Pi(\Omega)$. We show how the structure of the variety $\Pi(\Omega)$ is related to Xi's affirmation of a conjecture of Lusztig for $GL(n, \mathbb{C})$. The smooth dual $\Pi(G)$ admits a deformation retraction onto the tempered dual $\Pi^t(G)$.

2000 Mathematics Subject Classification: 46L80 22E50 46L87 11S37
Keywords and Phrases: Langlands correspondence, p-adic $GL(n)$, Baum-Connes map, smooth dual, tempered dual.

Introduction

The use of unramified quasicharacters to create a complex structure is well established in number theory. The group of unramified quasicharacters of the idele class group of a global field admits a complex structure: this complex structure provides the background for the functional equation of the zeta integral $Z(\omega, \Phi)$, see [39, Theorem 2, p. 121].
Let now G be a reductive p-adic group and let M be a Levi subgroup of G. Let $\Pi^{sc}(M)$ denote the set of equivalence classes of irreducible supercuspidal representations of M. Harish-Chandra creates a complex structure on the set $\Pi^{sc}(M)$ by using unramified quasicharacters of M [16, p.84]. This complex structure provides the background for the Harish-Chandra functional equations [16, p. 91].
Bernstein considered the set $\Omega(G)$ of all conjugacy classes of pairs (M, σ) where M is a Levi subgroup of G and σ is an irreducible supercuspidal representation of M. Making use of unramified quasicharacters of M, Bernstein gave the set

[1]Partially supported by a Leverhulme Trust Fellowship

$\Omega(G)$ the structure of a complex algebraic variety. Each irreducible component Ω of $\Omega(G)$ has the structure of a complex affine algebraic variety [5].
Let $\Pi(G)$ denote the set of equivalence classes of irreducible smooth representations of G. We will call $\Pi(G)$ the smooth dual of G. Bernstein defines the *infinitesimal character* from $\Pi(G)$ to $\Omega(G)$:

$$inf.ch. : \Pi(G) \to \Omega(G).$$

The infinitesimal character is a finite-to-one map from the set $\Pi(G)$ to the variety $\Omega(G)$.
Let F be a nonarchimedean local field and from now on let $G = GL(n) = GL(n, F)$. Let now W_F be the Weil group of the local field F, then W_F admits unramified quasicharacters, namely those which are trivial on the inertia subgroup I_F. Making use of the unramified quasicharacters of W_F, we introduced in [8] a complex structure on the set of Langlands parameters for $GL(n)$. In view of the local Langlands correspondence for $GL(n)$ this creates, by transport of structure, a complex structure on the smooth dual of $GL(n)$.
In Section 1 of this article, we describe in detail the complex structure on the set of L-parameters for $GL(n)$. We prove that the smooth dual $\Pi(GL(n))$ has the structure of complex manifold. The local L-factors $L(s, \pi)$ then appear as complex valued functions of several complex variables. We illustrate this with the local L-factors attached to the unramified principal series of $GL(n)$.
The complex structure on $\Pi(GL(n))$ is well adapted to the periodic cyclic homology of the Hecke algebra $\mathcal{H}(GL(n))$. The identical structure arises in the work of Xi on Lusztig's conjecture [40]. Let W be the extended affine Weyl group associated to $GL(n, \mathbb{C})$, and let J be the associated based ring (asymptotic algebra) [27, 40]. Xi confirms Lusztig's conjecture and proves that $J \otimes_{\mathbb{Z}} \mathbb{C}$ is Morita equivalent to the coordinate ring of the complex algebraic variety $\widetilde{(\mathbb{C}^\times)^n / S_n}$, the *extended* quotient by the symmetric group S_n of the complex n-dimensional torus $(\mathbb{C}^\times)^n$. In Section 2 we describe the theorem of Xi on the structure of the based ring J.
So the structure of extended quotient, which runs through our work, occurs in the work of Xi *at the level of algebras.* The link with our work is now provided by the theorem of Baum and Nistor [3, 4]

$$\mathrm{HP}_*(\mathcal{H}(n, q)) \simeq \mathrm{HP}_*(J)$$

where $\mathcal{H}(n, q)$ is the associated extended affine Hecke algebra.
Let Ω be a component in the Bernstein variety $\Omega(GL(n))$, and let $\mathcal{H}(G) = \bigoplus \mathcal{H}(\Omega)$ be the Bernstein decomposition of the Hecke algebra.
Let

$$\Pi(\Omega) = (inf.ch.)^{-1}\Omega.$$

Then $\Pi(\Omega)$ is a smooth complex algebraic variety with finitely many irreducible components. We have the following Bernstein decomposition of $\Pi(G)$:

$$\Pi(G) = \bigsqcup \Pi(\Omega).$$

Let M be a compact C^∞ manifold. Then $C^\infty(M)$ is a Fréchet algebra, and we have Connes' fundamental theorem [14, Theorem 2, p. 208]:

$$\mathrm{HP}_*(C^\infty(M)) \cong \mathrm{H}^*(M;\mathbb{C}).$$

Now the ideal $\mathcal{H}(\Omega)$ is a purely algebraic object, and, in computing its periodic cyclic homology, we would hope to find an algebraic variety to play the role of the manifold M. This algebraic variety is $\Pi(\Omega)$.

Theorem 0.1. *Let Ω be a component in the Bernstein variety $\Omega(G)$. Then the periodic cyclic homology of $\mathcal{H}(\Omega)$ is isomorphic to the periodised de Rham cohomology of $\Pi(\Omega)$:*

$$\mathrm{HP}_*(\mathcal{H}(\Omega)) \cong \mathrm{H}^*(\Pi(\Omega);\mathbb{C}).$$

This theorem constitutes the main result of Section 3, which is then used to show that the periodic cyclic homology of the Hecke algebra of $GL(n)$ is isomorphic to the periodic cyclic homology of the Schwartz algebra of $GL(n)$. We also provide an explicit numerical formula for the dimension of the periodic cyclic homology of $\mathcal{H}(\Omega)$ in terms of certain natural number invariants attached to Ω.

The smooth dual $\Pi(GL(n))$ has a natural stratification-by-dimension. We compare this stratification with the Schneider-Zink stratification [34]. Stratification-by-dimension is finer than the Schneider-Zink stratification, see Section 3.

A *scheme* X is a topological space, called the *support* of X and denoted $|X|$, together with a sheaf $\mathcal{O}_X$ of rings on X, such that the pair $(|X|, \mathcal{O}_X)$ is locally affine, see [15, p. 21]. The smooth dual $\Pi(G)$ determines a reduced scheme, see [18, Prop. 2.6]. If S is the reduced scheme determined by the Bernstein variety $\Omega(G)$, then $\Pi(G)$ is a *scheme over* S, i.e. a scheme together with a morphism $\Pi(G) \to S$. This morphism is the q-projection introduced in [8]:

$$\pi_q : \Pi(G) \to S.$$

In Section 4 we give a detailed description of the q-projection and prove that the q-projection is a finite morphism.

From the point of view of noncommutative geometry it is natural to seek the spaces which underlie the noncommutative algebras $\mathcal{H}(G)$ and $\mathcal{S}(G)$. The space which underlies the Hecke algebra $\mathcal{H}(G)$ is the complex manifold $\Pi(G)$. The space which underlies the Schwartz algebra is the Harish-Chandra parameter space, which is a disjoint union of compact orbifolds. In Section 5 we construct a deformation retraction of the smooth dual onto the tempered dual. We view this deformation retraction as a geometric counterpart of the Baum-Connes assembly map for $GL(n)$.

In Section 6 we track the fate of supercuspidal representations of G through the diagram which appears in Section 5. In particular, the index map μ manifests itself as an example of Ahn reciprocity.

We would like to thank Paul Baum, Alain Connes, Jean-Francois Dat and Nigel Higson for many valuable conversations. Jacek Brodzki was supported in part by a Leverhulme Trust Fellowship. This article was completed while Roger Plymen was at IHES, France.

1. The complex structure on the smooth dual of $GL(n)$

The field F is a nonarchimedean local field, so that F is a finite extension of $\mathbb{Q}_p$, for some prime p or F is a finite extension of the function field $\mathbb{F}_p((x))$. The residue field k_F of F is the quotient $\mathbf{o}_F/\mathbf{m}_F$ of the ring of integers $\mathbf{o}_F$ by its unique maximal ideal $\mathbf{m}_F$. Let q be the cardinality of k_F.

The essence of local class field theory, see [29, p.300], is a pair of maps

$$(d : G \longrightarrow \widehat{\mathbb{Z}}, v : F^\times \longrightarrow \mathbb{Z})$$

where G is a profinite group, $\widehat{\mathbb{Z}}$ is the profinite completion of $\mathbb{Z}$, and v is the valuation.

Let $\overline{F}$ be a separable algebraic closure of F. Then the absolute Galois group $G(\overline{F}|F)$ is the projective limit of the finite Galois groups $G(E|F)$ taken over the finite extensions E of F in $\overline{F}$. Let $\widetilde{F}$ be the maximal unramified extension of F. The map d is in this case the projection map

$$d : G(\overline{F}|F) \longrightarrow G(\tilde{F}|F) \cong \hat{\mathbb{Z}}$$

The group $G(\tilde{F}|F)$ is procyclic. It has a single topological generator: the Frobenius automorphism ϕ_F of $\tilde{F}|F$. The Weil group W_F is by definition the pre-image of $< \phi_F >$ in $G(\overline{F} \mid F)$. We thus have the surjective map

$$d : W_F \longrightarrow \mathbb{Z}$$

The pre-image of 0 is the inertia group I_F. In other words we have the following short exact sequence

$$1 \to I_F \to W_F \to \mathbb{Z} \to 0$$

The group I_F is given the profinite topology induced by $G(\overline{F}|F)$. The topology on the Weil group W_F is dictated by the above short exact sequence. The Weil group W_F is a locally compact group with maximal compact subgroup I_F. The map

$$W_F \longrightarrow G(\tilde{F}|F)$$

is a continuous homomorphism with dense image.

A detailed account of the Weil group for local fields may be found in [37]. For a topological group G we denote by G^{ab} the quotient $G^{\mathrm{ab}} = G/G^c$ of G by the closure G^c of the commutator subgroup of G. Thus G^{ab} is the maximal abelian Hausdorff quotient of G. The local reciprocity laws [29, p.320]

$$r_{E|F} : G(E|F)^{ab} \cong F^\times/N_{E|F}E^\times$$

now create an isomorphism [30, p.69]:

$$r_F : W_F^{ab} \cong F^\times$$

We have $W_F = \sqcup \Phi^n I_F, n \in \mathbb{Z}$. The Weil group is a locally compact, totally disconnected group, whose maximal compact subgroup is I_F. This subgroup is also open. There are three models for the Weil-Deligne group.

One model is the crossed product $W_F \ltimes \mathbb{C}$, where the Weil group acts on $\mathbb{C}$ by $w \cdot x = \|w\| x$, for all $w \in W_F$ and $x \in \mathbb{C}$.
The action of W_F on $\mathbb{C}$ extends to an action of W_F on $SL(2,\mathbb{C})$. The semidirect product $W_F \ltimes SL(2,\mathbb{C})$ is then isomorphic to the direct product $W_F \times SL(2,\mathbb{C})$, see [22, p.278]. Then a complex representation of $W_F \times SL(2,\mathbb{C})$ is determined by its restriction to $W_F \times SU(2)$, where $SU(2)$ is the standard compact Lie group.
From now on, we shall use this model for the Weil-Deligne group:

$$\mathcal{L}_F = W_F \times SU(2).$$

Definition 1.1. An L-parameter is a continuous homomorphism

$$\phi : \mathcal{L}_F \to GL(n,\mathbb{C})$$

such that $\phi(w)$ is semisimple for all $w \in W_F$. Two L-parameters are equivalent if they are conjugate under $GL(n,\mathbb{C})$. The set of equivalence classes of L-parameters is denoted $\Phi(G)$.

Definition 1.2. A representation of G on a complex vector space V is *smooth* if the stabilizer of each vector in V is an open subgroup of G. The set of equivalence classes of irreducible smooth representations of G is the *smooth dual* $\Pi(G)$ of G.

Theorem 1.3. *Local Langlands Correspondence for $GL(n)$. There is a natural bijection between $\Phi(GL(n))$ and $\Pi(GL(n))$.*

The naturality of the bijection involves compatibility of the L-factors and ϵ-factors attached to the two types of objects.
The local Langlands conjecture for $GL(n)$ was proved by Laumon, Rapoport and Stuhler [25] when F has positive characteristic and by Harris-Taylor [17] and Henniart [19] when F has characteristic zero.
We recall that a *matrix coefficient* of a representation ρ of a group G on a vector space V is a function on G of the form $f(g) = \langle \rho(g)v, w\rangle$, where $v \in V$, $w \in V^*$, and V^* denotes the dual space of V. The inner product is given by the duality between V and V^*. A representation ρ of G is called *supercuspidal* if and only if the support of every matrix coefficient is compact modulo the centre of G.
Let $\tau_j = \mathrm{spin}(j)$ denote the $(2j+1)$-dimensional complex irreducible representation of the compact Lie group $SU(2)$, $j = 0, 1/2, 1, 3/2, 2, \ldots$.
For $GL(n)$ the local Langlands correspondence works in the following way.

- Let ρ be an irreducible representation of the Weil group W_F. Then $\pi_F(\rho \otimes 1)$ is an irreducible supercuspidal representation of $GL(n)$, and every irreducible supercuspidal representation of $GL(n)$ arises in this way. If $\det(\rho)$ is a unitary character, then $\pi_F(\rho\otimes 1)$ has unitary central character, and so is pre-unitary.
- We have $\pi_F(\rho \otimes \mathrm{spin}(j)) = Q(\Delta)$, the Langlands quotient associated to the segment $\{|\ |^{-(j-1)/2}\pi_F(\rho), \ldots, |\ |^{(j-1)/2}\pi_F(\rho)\}$. If $\det(\rho)$ is unitary,

then $Q(\Delta)$ is in the discrete series. In particular, if $\rho = 1$ then $\pi_F(1 \otimes \mathrm{spin}(j))$ is the Steinberg representation $St(2j+1)$ of $GL(2j+1)$.
- If ϕ is an L-parameter for $GL(n)$ then $\phi = \phi_1 \oplus \ldots \oplus \phi_m$ where $\phi_j = \rho_j \otimes \mathrm{spin}(j)$. Then $\pi_F(\rho)$ is the Langlands quotient $Q(\Delta_1, \ldots, \Delta_m)$. If $\det(\rho_j)$ is a unitary character for each j, then $\pi_F(\phi)$ is a tempered representation of $GL(n)$.

This correspondence creates, as in [23, p. 381], a natural bijection

$$\pi_F : \Phi(GL(n)) \to \Pi(GL(n)).$$

A quasi-character $\psi : W_F \to \mathbb{C}^\times$ is *unramified* if ψ is trivial on the inertia group I_F. Recall the short exact sequence

$$0 \to I_F \to W_F \xrightarrow{d} \mathbb{Z} \to 0$$

Then $\psi(w) = z^{d(w)}$ for some $z \in \mathbb{C}^\times$. Note that ψ is not a *Galois* representation unless z has finite order in the complex torus $\mathbb{C}^\times$, see [37]. Let $\Psi(W_F)$ denote the group of all unramified quasi-characters of W_F. Then

$$\begin{array}{ccc} \Psi(W_F) & \simeq & \mathbb{C}^\times \\ \psi & \mapsto & z \end{array}$$

Each L-parameter $\phi : \mathcal{L}_F \to GL(n, \mathbb{C})$ is of the form $\phi_1 \oplus \cdots \oplus \phi_m$ with each ϕ_j irreducible. Each irreducible L-parameter is of the form $\rho \otimes \mathrm{spin}(j)$ with ρ an irreducible representation of the Weil group W_F.

Definition 1.4. The orbit $\mathcal{O}(\phi) \subset \Phi_F(G)$ is defined as follows

$$\mathcal{O}(\phi) = \{\bigoplus_{r=1}^{m} \psi_r \phi_r \mid \psi_r \in \Psi(W_F), 1 \le r \le m\}$$

where each ψ_r is an unramified quasi-character of W_F.

Definition 1.5. Let $\det \phi_r$ be a unitary character, $1 \le r \le m$ and let $\phi = \phi_1 \oplus \ldots \oplus \phi_m$. The compact orbit $\mathcal{O}^t(\phi) \subset \Phi^t(G)$ is defined as follows:

$$\mathcal{O}^t(\phi) = \{\bigoplus_{r=1}^{m} \psi_r \phi_r \mid \psi_r \in \Psi^t(W_F), 1 \le r \le m\}$$

where each ψ_r is an unramified unitary character of W_F.

We note that $I_F \times SU(2) \subset W_F \times SU(2)$ and in fact $I_F \times SU(2)$ is the maximal compact subgroup of $\mathcal{L}_F$. Now let ϕ be an L-parameter. Moving (if necessary) to another point in the orbit $\mathcal{O}(\phi)$ we can write ϕ in the canonical form

$$\phi = \phi_1 \oplus \ldots \oplus \phi_1 \oplus \ldots \oplus \phi_k \oplus \ldots \oplus \phi_k$$

where ϕ_1 is repeated l_1 times, …, ϕ_k is repeated l_k times, and the representations

$$\phi_j|_{I_F \times SU(2)}$$

are irreducible and pairwise inequivalent, $1 \le j \le k$. We will now write $k = k(\phi)$. This natural number is an invariant of the orbit $\mathcal{O}(\phi)$. We have

$$\mathcal{O}(\phi) = \mathrm{Sym}^{l_1}\mathbb{C}^\times \times \ldots \times \mathrm{Sym}^{l_k}\mathbb{C}^\times$$

the product of symmetric products of $\mathbb{C}^\times$.

Theorem 1.6. *The set $\Phi(GL(n))$ has the structure of complex algebraic variety. Each irreducible component $\mathcal{O}(\phi)$ is isomorphic to the product of a complex affine space and a complex torus*

$$\mathcal{O}(\phi) = \mathbb{A}^l \times (\mathbb{C}^\times)^k$$

where $k = k(\phi)$.

Proof. Let $Y = \mathbb{V}(x_1y_1 - 1, \ldots, x_ny_n - 1) \subset \mathbb{C}^{2n}$. Then Y is a Zariski-closed set in $\mathbb{C}^{2n}$, and so is an affine complex algebraic variety. Let $X = (\mathbb{C}^\times)^n$. Set $\alpha : Y \to X, \alpha(x_1, y_1, \ldots, x_n, y_n) = (x_1, \ldots, x_n)$ and $\beta : X \to Y, \beta(x_1, \ldots, x_n) = (x_1, x_1^{-1}, \ldots, x_n, x_n^{-1})$. So X can be embedded in affine space $\mathbb{C}^{2n}$ as a Zariski-closed subset. Therefore X is an affine algebraic variety, as in [36, p.50].
Let $A = \mathbb{C}[X]$ be the coordinate ring of X. This is the restriction to X of polynomials on $\mathbb{C}^{2n}$, and so $A = \mathbb{C}[X] = \mathbb{C}[x_1, x_1^{-1}, \ldots, x_n, x_n^{-1}]$, the ring of Laurent polynomials in n variables $x_1, \ldots, x_n$. Let S_n be the symmetric group, and let Z denote the quotient variety X/S_n. The variety Z is an affine complex algebraic variety.
The coordinate ring of Z is

$$\mathbb{C}[Z] \simeq \mathbb{C}[x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}]^{S_n}.$$

Let σ_i, $i = 1, \ldots, n$ be the elementary symmetric polynomials in n variables. Then from the last isomorphism we have

$$\begin{aligned}
\mathbb{C}[Z] &\simeq \mathbb{C}[x_1, \ldots, x_n]^{S_n} \otimes \mathbb{C}[\sigma_n^{-1}] \\
&\simeq \mathbb{C}[\sigma_1, \ldots, \sigma_n] \otimes \mathbb{C}[\sigma_n^{-1}] \\
&\simeq \mathbb{C}[\sigma_1, \ldots, \sigma_{n-1}] \otimes \mathbb{C}[\sigma_n, \sigma_n^{-1}] \\
&\simeq \mathbb{C}[\mathbb{A}^{n-1}] \otimes \mathbb{C}[\mathbb{A} - \{0\}] \\
&\simeq \mathbb{C}[\mathbb{A}^{n-1} \times (\mathbb{A} - \{0\})]
\end{aligned}$$

where $\mathbb{A}^n$ denotes complex affine n-space. The coordinate ring of the quotient variety $\mathbb{C}^{\times n}/S_n$ is isomorphic to the coordinate ring of $\mathbb{A}^{n-1} \times (\mathbb{A} - \{0\})$. Now the categories of affine algebraic varieties and of finitely generated reduced $\mathbb{C}$-algebras are equivalent, see [36, p.26]. Therefore the variety $\mathbb{C}^{\times n}/S_n$ is isomorphic to the variety $\mathbb{A}^{n-1} \times (\mathbb{A} - \{0\})$.
Consider $\mathbb{A} - \{0\} = \mathbb{V}(f)$ where $f(x) = x_1x_2 - 1$. Then $\partial f/\partial x_1 = x_2 \neq 0$ and $\partial f/\partial x_2 = x_1 \neq 0$ on the variety $\mathbb{V}(f)$. So $\mathbb{A} - \{0\}$ is smooth. Then $\mathbb{A}^{n-1} \times (\mathbb{A} - \{0\})$ is smooth. Therefore the quotient variety $\mathbb{C}^{\times n}/S_n$ is a smooth complex affine algebraic variety of dimension n. Now each orbit $\mathcal{O}(\phi)$ is a product of symmetric products of $\mathbb{C}^\times$. Therefore each orbit $\mathcal{O}(\phi)$ is a smooth complex affine algebraic variety. We have

$$\mathcal{O}(\phi) = \mathrm{Sym}^{l_1}\mathbb{C}^\times \times \ldots \times \mathrm{Sym}^{l_k}\mathbb{C}^\times = \mathbb{A}^l \times (\mathbb{C}^\times)^k$$

where $l = l_1 + \ldots + l_k - k$ and $k = k(\phi)$. $\square$

We now transport the complex structure from $\Phi(GL(n))$ to $\Pi(GL(n))$ via the local Langlands correspondence. This leads to the next result.

Theorem 1.7. *The smooth dual* $\Pi(GL(n))$ *has a natural complex structure. Each irreducible component is a smooth complex affine algebraic variety.*

The smooth dual $\Pi(GL(n))$ has countably many irreducible components of each dimension d with $1 \leq d \leq n$. The irreducible supercuspidal representations of $GL(n)$ arrange themselves into the 1-dimensional tori.
It follows from Theorems 1.6 and 1.7 that the smooth dual $\Pi(GL(n))$ is a complex manifold. Then $\mathbb{C} \times \Pi(GL(n))$ is a complex manifold. So the local L-factor $L(s, \pi_v)$ and the local ϵ-factor $\epsilon(s, \pi_v)$ are functions of *several complex variables*:

$$L : \mathbb{C} \times \Pi(GL(n)) \longrightarrow \mathbb{C}$$

$$\epsilon : \mathbb{C} \times \Pi(GL(n)) \longrightarrow \mathbb{C}.$$

Example 1.8. Unramified representations. Let $\psi_1, \ldots, \psi_n$ be unramified quasicharacters of the Weil group W_F. Then we have

$$\psi_j(w) = z_j^{d(w)}$$

with $z_j \in \mathbb{C}^\times$ for all $1 \leq j \leq n$. Let ϕ be the L-parameter given by $\psi_1 \oplus \ldots \oplus \psi_n$. Then the image $\pi_F(\phi)$ of ϕ under the local Langlands correspondence π_F is an unramified principal series representation.
For the local L-factors $L(s, \pi)$ see [23, p. 377]. The local L-factor attached to such an unramified representation of $GL(n)$ is given by

$$L(s, \pi_F(\phi)) = \prod_{j=1}^{n} (1 - z_j q^{-s})^{-1}.$$

This exhibits the local L-factor as a function on the complex manifold $\mathbb{C} \times \mathrm{Sym}^n \mathbb{C}^\times$.

2. The structure of the based ring J

Let W be the extended affine Weyl group associated to $GL(n, \mathbb{C})$. For each two-sided cell $\mathbf{c}$ of W we have a corresponding partition λ of n. Let μ be the dual partition of λ. Let u be a unipotent element in $GL(n, \mathbb{C})$ whose Jordan blocks are determined by the partition μ. Let the distinct parts of the dual partition μ be $\mu_1, \ldots, \mu_p$ with μ_r repeated n_r times, $1 \leq r \leq p$.
Let $C_G(u)$ be the centralizer of u in $G = GL(n, \mathbb{C})$. Then the maximal reductive subgroup $F_\mathbf{c}$ of $C_G(u)$ is isomorphic to $GL(n_1, \mathbb{C}) \times GL(n_2, \mathbb{C}) \times \cdots \times GL(n_p, \mathbb{C})$.
Following Lusztig [27] and Xi [40, 1.5] let J be the free $\mathbb{Z}$-module with basis $\{t_w \mid w \in W\}$. The multiplication $t_w t_u = \sum_{v \in W} \gamma_{w,u,v} t_v$ defines an associative ring structure on J. The ring J is the based ring of W. For each two-sided cell $\mathbf{c}$ of W the $\mathbb{Z}$-submodule $J_\mathbf{c}$ of J, spanned by all t_w, $w \in \mathbf{c}$, is a two-sided ideal of J. The ring $J_\mathbf{c}$ is the based ring of the two-sided cell $\mathbf{c}$. Let $|Y|$ be the

number of left cells contained in $\mathbf{c}$. The Lusztig conjecture says that there is a ring isomorphism

$$J_{\mathbf{c}} \simeq M_{|Y|}(R_{F_{\mathbf{c}}}), \quad t_w \mapsto \pi(w)$$

where $R_{F_{\mathbf{c}}}$ is the rational representation ring of $F_{\mathbf{c}}$. This conjecture for $GL(n,\mathbb{C})$ has been proved by Xi [40, 1.5, 4.1, 8.2].

Since $F_{\mathbf{c}}$ is isomorphic to a direct product of the general linear groups $GL(n_i,\mathbb{C})$ $(1 \le i \le p)$ we see that $R_{F_{\mathbf{c}}}$ is isomorphic to the tensor product over $\mathbb{Z}$ of the representation rings $R_{GL(n_i,\mathbb{C})}$, $1 \le i \le p$. For the ring $R_{GL(n,\mathbb{C})}$ we have

$$R_{GL(n,\mathbb{C})} \simeq \mathbb{Z}[X_1, X_2, \ldots, X_n][X_n^{-1}]$$

where the elements $X_1, X_2, \ldots, X_n, X_n^{-1}$ are described in [40, 4.2][6, IX.125]. Then

$$\begin{array}{rcl} R_{GL(n,\mathbb{C})} & \simeq & \mathbb{Z}[\sigma_1, \ldots, \sigma_n, \sigma_n^{-1}] \\ & \simeq & \mathbb{Z}[x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}]^{S_n} \end{array}$$

We have

$$R_{GL(n,\mathbb{C})} \otimes_{\mathbb{Z}} \mathbb{C} \simeq \mathbb{C}[\mathrm{Sym}^n \mathbb{C}^\times]$$

and

$$R_{F_{\mathbf{c}}} \otimes_{\mathbb{Z}} \mathbb{C} \simeq \mathbb{C}[\mathrm{Sym}^{n_1}\mathbb{C}^\times \times \cdots \times \mathrm{Sym}^{n_p}\mathbb{C}^\times]$$

We recall the *extended quotient*. Let the finite group Γ act on the space X. Let $\widetilde{X} = \{(x,\gamma) : \gamma x = x\}$, let Γ act on $\widehat{X}$ by $\gamma_1(x,\gamma) = (\gamma_1 x, \gamma_1\gamma\gamma_1^{-1})$. Then $\widetilde{X}/\Gamma$ is the extended quotient of X by Γ, and we have

$$\widetilde{X}/\Gamma = \bigsqcup X^\gamma / Z(\gamma)$$

where one γ is chosen in each Γ-conjugacy class.

There is a canonical projection $\widetilde{X}/\Gamma \to X/\Gamma$.

Let $\gamma \in S_n$ have cycle type μ, let $X = (\mathbb{C}^\times)^n$. Then

$$\begin{array}{rcl} X^\gamma & \simeq & (C^\times)^{n_1} \times \cdots \times (C^\times)^{n_p} \\ Z(\gamma) & \simeq & (\mathbb{Z}/\mu_1\mathbb{Z}) \wr S_{n_1} \times \cdots \times (\mathbb{Z}/\mu_p\mathbb{Z}) \wr S_{n_p} \\ X^\gamma/Z(\gamma) & \simeq & \mathrm{Sym}^{n_1}\mathbb{C}^\times \times \cdots \times \mathrm{Sym}^{n_p}\mathbb{C}^\times \end{array}$$

and so

$$R_{F_{\mathbf{c}}} \otimes_{\mathbb{Z}} \mathbb{C} \simeq \mathbb{C}[X^\gamma/Z(\gamma)]$$

Then

$$J \otimes_{\mathbb{Z}} \mathbb{C} = \oplus_{\mathbf{c}}(J_{\mathbf{c}} \otimes_{\mathbb{Z}} \mathbb{C}) \sim \oplus_{\mathbf{c}}(R_{F_{\mathbf{c}}} \otimes_{\mathbb{Z}} \mathbb{C}) \simeq \mathbb{C}[\widetilde{X}/S_n]$$

The algebra $J \otimes_{\mathbb{Z}} \mathbb{C}$ is Morita equivalent to a reduced, finitely generated, commutative unital $\mathbb{C}$-algebra, namely the coordinate ring of the extended quotient $\widetilde{X}/S_n$.

3. Periodic cyclic homology of the Hecke algebra

The Bernstein variety $\Omega(G)$ of G is the set of G-conjugacy classes of pairs (M,σ), where M is a Levi (i.e. block-diagonal) subgroup of G, and σ is an irreducible supercuspidal representation of M. Each irreducible smooth representation of G is a subquotient of an induced representation $i_{GM}\sigma$. The pair (M,σ) is unique up to conjugacy. This creates a finite-to-one map, the infinitesimal character, from $\Pi(G)$ onto $\Omega(G)$.

Let $\Omega(G)$ be the Bernstein variety of G. Each point in $\Omega(G)$ is a conjugacy class of cuspidal pairs (M,σ). A quasicharacter $\psi : M \longrightarrow \mathbb{C}^\times$ is *unramified* if ψ is trivial on M°. The group of unramified quasicharacters of M is denoted $\Psi(M)$. We have $\Psi(M) \cong (\mathbb{C}^\times)^\ell$ where ℓ is the parabolic rank of M. The group $\Psi(M)$ now creates orbits: the orbit of (M,σ) is $\{(M,\psi\otimes\sigma) : \psi \in \Psi(M)\}$. Denote this orbit by D, and set $\Omega = D/W(M,D)$, where $W(M)$ is the Weyl group of M and $W(M,D)$ is the subgroup of $W(M)$ which leaves D globally invariant. The orbit D has the structure of a complex torus, and so Ω is a complex algebraic variety. We view Ω as a component in the algebraic variety $\Omega(G)$.

The Bernstein variety $\Omega(G)$ is the disjoint union of ordinary quotients. We now replace the ordinary quotient by the extended quotient to create a new variety $\Omega^+(G)$. So we have

$$\Omega(G) = \bigsqcup D/W(M,D) \;\text{ and }\; \Omega^+(G) = \bigsqcup \widetilde{D}/W(M,D)$$

Let Ω be a component in the Bernstein variety $\Omega(GL(n))$, and let $\mathcal{H}(G) = \bigoplus \mathcal{H}(\Omega)$ be the Bernstein decomposition of the Hecke algebra.

Let

$$\Pi(\Omega) = (inf.ch.)^{-1}\Omega.$$

Then $\Pi(\Omega)$ is a smooth complex algebraic variety with finitely many irreducible components. We have the following Bernstein decomposition of $\Pi(G)$:

$$\Pi(G) = \bigsqcup \Pi(\Omega).$$

Let M be a compact C^∞ manifold. Then $C^\infty(M)$ is a Fréchet algebra, and we have Connes' fundamental theorem [14, Theorem 2, p. 208]:

$$\mathrm{HP}_*(C^\infty(M)) \cong \mathrm{H}^*(M;\mathbb{C}).$$

Now the ideal $\mathcal{H}(\Omega)$ is a purely algebraic object, and, in computing its periodic cyclic homology, we would hope to find an algebraic variety to play the role of the manifold M. This algebraic variety is $\Pi(\Omega)$.

Theorem 3.1. *Let Ω be a component in the Bernstein variety $\Omega(G)$. Then the periodic cyclic homology of $\mathcal{H}(\Omega)$ is isomorphic to the periodised de Rham cohomology of $\Pi(\Omega)$:*

$$\mathrm{HP}_*(\mathcal{H}(\Omega)) \cong \mathrm{H}^*(\Pi(\Omega);\mathbb{C}).$$

Proof. We can think of Ω as a vector $(\tau_1,\ldots,\tau_r)$ of irreducible supercuspidal representations of smaller general linear groups, the entries of this vector being only determined up to tensoring with unramified quasicharacters and permutation. If the vector is equivalent to $(\sigma_1, \ldots, \sigma_1, \ldots, \sigma_r, \ldots, \sigma_r)$ with σ_j repeated e_j times, $1 \leq j \leq r$, and $\sigma_1, \ldots, \sigma_r$ are pairwise distinct, then we say that Ω has *exponents* $e_1, \ldots, e_r$.

Then there is a Morita equivalence

$$\mathcal{H}(\Omega) \sim \mathcal{H}(e_1, q_1) \otimes \ldots \otimes \mathcal{H}(e_r, q_r)$$

where $q_1, \ldots, q_r$ are natural number invariants attached to Ω.

This result is due to Bushnell-Kutzko [11, 12, 13]. We describe the steps in the proof. Let (ρ, W) be an irreducible smooth representation of the compact open subgroup K of G. As in [12, 4.2], the pair (K, ρ) is an Ω-type in G if and only if, for $(\pi, V) \in \Pi(G)$, we have inf.ch.$(\pi) \in \Omega$ if and only if π contains ρ. The existence of an Ω-type in $GL(n)$, for each component Ω in $\Omega(GL(n))$, is established in [13, 1.1]. So let (K, ρ) be an Ω-type in $GL(n)$. As in [12, 2.9], let

$$e_\rho(x) = (\text{vol}K)^{-1}(\dim \rho) \ \text{Trace}_W(\rho(x^{-1}))$$

for $x \in K$ and 0 otherwise.

Then e_ρ is an idempotent in the Hecke algebra $\mathcal{H}(G)$. Then we have

$$\mathcal{H}(\Omega) \cong \mathcal{H}(G) * e_\rho * \mathcal{H}(G)$$

as in [12, 4.3] and the two-sided ideal $\mathcal{H}(G) * e_\rho * \mathcal{H}(G)$ is Morita equivalent to $e_\rho * \mathcal{H}(G) * e_\rho$. Now let $\mathcal{H}(K, \rho)$ be the endomorphism-valued Hecke algebra attached to the semisimple type (K, ρ). By [12, 2.12] we have a canonical isomorphism of unital $\mathbb{C}$-algebras :

$$\mathcal{H}(G, \rho) \otimes_{\mathbb{C}} \text{End}_{\mathbb{C}} W \cong e_\rho * \mathcal{H}(G) * e_\rho$$

so that $e_\rho * \mathcal{H}(G) * e_\rho$ is Morita equivalent to $\mathcal{H}(G, \rho)$. Now we quote the main theorem for semisimple types in $GL(n)$ [13, 1.5]: there is an isomorphism of unital $\mathbb{C}$-algebras

$$\mathcal{H}(G, \rho) \cong \mathcal{H}(G_1, \rho_1) \otimes \ldots \otimes \mathcal{H}(G_r, \rho_r)$$

The factors $\mathcal{H}(G_i, \rho_i)$ are (extended) affine Hecke algebras whose structure is given explicitly in [11, 5.6.6]. This structure is in terms of generators and relations [11, 5.4.6]. So let $\mathcal{H}(e, q)$ denote the affine Hecke algebra associated to the affine Weyl group $\mathbb{Z}^e \rtimes S_e$. Putting all this together we obtain a Morita equivalence

$$\mathcal{H}(\Omega) \sim \mathcal{H}(e_1, q_1) \otimes \ldots \otimes \mathcal{H}(e_r, q_r)$$

The natural numbers $q_1, \ldots, q_r$ are specified in [11, 5.6.6]. They are the cardinalities of the residue fields of certain extension fields $E_1/F, \ldots, E_r/F$.

Using the Künneth formula the calculation of $\text{HP}_*(\mathcal{H}(\Omega))$ is reduced to that of the affine Hecke algebra $\mathcal{H}(e, q)$. Baum and Nistor demonstrate the spectral invariance of periodic cyclic homology in the class of finite type algebras [3, 4]. Now $\mathcal{H}(e, q)$ is the Iwahori-Hecke algebra associated to the extended affine

Weyl group $\mathbb{Z}^e \rtimes S_e$, and let J be the asymptotic Hecke algebra (based ring) associated to $\mathbb{Z}^e \rtimes S_e$. According to [3, 4], Lusztig's morphisms $\phi_q : \mathcal{H}(e,q) \to J$ induce isomorphisms

$$(\phi_q)_* : \mathrm{HP}_*(\mathcal{H}(e,q)) \to \mathrm{HP}_*(J)$$

for all $q \in \mathbb{C}^\times$ that are not proper roots of unity. At this point we can back track and deduce that

$$\mathrm{HP}_*(\mathcal{H}(e,q)) \simeq \mathrm{HP}_*(J) \simeq \mathrm{HP}_*(\mathcal{H}_1)$$

and use the fact that $\mathcal{H}(e,1) \simeq \mathbb{C}[\mathbb{Z}^e \rtimes S_e]$. It is much more illuminating to quote Xi's proof of the Lusztig conjecture for the based ring J, see Section 2. Then we have

$$\mathrm{HP}_*(\mathcal{H}(e,q)) \simeq \mathrm{HP}_*(J) \simeq \mathrm{HP}_*(\mathbb{C}[\widetilde{(\mathbb{C}^\times)^e/S_e}]) \simeq \mathrm{H}^*(\widetilde{(\mathbb{C}^\times)^e/S_e};\mathbb{C}).$$

If Ω has exponents $e_1, \ldots, e_r$ then $e_1 + \ldots + e_r = d(\Omega) = \dim_{\mathbb{C}} \Omega$, and $W(\Omega)$ is a product of symmetric groups:

$$W(\Omega) = S_{e_1} \times \ldots \times S_{e_r}$$

We have

$$\begin{array}{rcl}
\mathrm{HP}_*(\mathcal{H}(\Omega)) & \simeq & \mathrm{HP}_*(\mathcal{H}(e_1,q_1) \otimes \cdots \otimes \mathcal{H}(e_r,q_r)) \\
& \simeq & \mathrm{HP}_*(\mathcal{H}(e_1,q_1)) \otimes \cdots \otimes \mathrm{HP}_*(\mathcal{H}(e_r,q_r)) \\
& \simeq & \mathrm{H}^*(\widetilde{(\mathbb{C}^\times)^{e_1}/S_{e_1}};\mathbb{C}) \otimes \cdots \otimes \mathrm{H}^*(\widetilde{(\mathbb{C}^\times)^{e_r}/S_{e_r}};\mathbb{C})
\end{array}$$

Now the extended quotient is multiplicative, i.e.

$$\widetilde{(\mathbb{C}^\times)^{d(\Omega)}/W(\Omega)} = \widetilde{(\mathbb{C}^\times)^{e_1}/S_{e_1}} \times \cdots \times \widetilde{(\mathbb{C}^\times)^{e_r}/S_{e_r}}$$

which implies that

$$\mathrm{HP}_*(\mathcal{H}(\Omega)) = \mathrm{H}^*(\widetilde{(\mathbb{C}^\times)^{d(\Omega)}/W(\Omega)};\mathbb{C})$$

Recall that

$$\begin{array}{rcl}
\Omega & = & (\mathbb{C}^\times)^{d(\Omega)}/W(\Omega) \\
\Omega^+ & = & \widetilde{(\mathbb{C}^\times)^{d(\Omega)}/W(\Omega)}
\end{array}$$

and by [8, p. 217] we have $\Pi(\Omega) \simeq \Omega^+$. It now follows that

$$\mathrm{HP}_*(\mathcal{H}(\Omega)) \simeq \mathrm{H}^*(\Pi(\Omega);\mathbb{C})$$

$\square$

Lemma 3.2. *Let Ω be a component in the variety $\Omega(G)$ and let Ω have exponents $\{e_1, \ldots, e_r\}$. Then for $j = 0, 1$ we have*

$$\dim_{\mathbb{C}} \mathrm{HP}_j\, \mathcal{H}(\Omega) = 2^{r-1}\beta(e_1)\cdots\beta(e_r)$$

where

$$\beta(e) = \sum_{|\lambda|=e} 2^{\alpha(\lambda)-1}$$

and where $\alpha(\lambda)$ is the number of unequal parts of λ. Here $|\lambda|$ is the weight of λ, i.e. the sum of the parts of λ so that λ is a partition of e.

Proof. Suppose first that Ω has the single exponent e. By Theorem 3.1 the periodic cyclic homology of $\mathcal{H}(\Omega)$ is isomorphic to the periodised de Rham cohomology of the extended quotient of $(\mathbb{C}^\times)^e$ by the symmetric group S_e. The components in this extended quotient correspond to the partitions of e. In fact, if $\alpha(\lambda)$ is the number of unequal parts in the partition λ then the corresponding component is homotopy equivalent to the compact torus of dimension $\alpha(\lambda)$. We now proceed by induction, using the fact that the extended quotient is multiplicative and the Künneth formula. $\square$

Theorem 3.1, combined with the calculation in [7], now leads to the next result.

Theorem 3.3. *The inclusion $\mathcal{H}(G) \longrightarrow \mathcal{S}(G)$ induces an isomorphism at the level of periodic cyclic homology:*

$$\mathrm{HP}_*(\mathcal{H}(G)) \simeq \mathrm{HP}_*(\mathcal{S}(G)).$$

Remark 3.4. We now consider further the disjoint union

$$\Phi(\Omega) = \mathcal{O}(\phi_1) \sqcup \cdots \sqcup \mathcal{O}(\phi_r) \simeq \Omega^+$$

If we apply the local Langlands correspondence π_F then we obtain

$$\Pi(\Omega) = \pi_F(\mathcal{O}(\phi_1)) \sqcup \cdots \sqcup \pi_F(\mathcal{O}(\phi_r)) \simeq \Omega^+$$

This partition of $\Pi(\Omega)$ is *identical* to that in Schneider-Zink [34, p. 198], modulo notational differences. In their notation, for each $\mathcal{P} \in \mathcal{B}$ there is a natural map

$$Q_{\mathcal{P}} : X_{nr}(N_{\mathcal{P}}) \to \mathrm{Irr}(\Omega)$$

such that

$$\mathrm{Irr}(\Omega) = \bigsqcup_{\mathcal{P} \in \mathcal{B}} \mathrm{im}(Q_{\mathcal{P}}).$$

In fact this is a special stratification of $\mathrm{Irr}(\Omega)$ in the precise sense of their article [34, p.198].
Let

$$Z_{\mathcal{P}} = \bigcup_{\mathcal{P}' \leq \mathcal{P}} \mathrm{im}(Q_{\mathcal{P}})$$

Then $Z_{\mathcal{P}}$ is a Jacobson closed set, in fact $Z_{\mathcal{P}} = V(J_{\mathcal{P}})$, where $J_{\mathcal{P}}$ is a certain 2-sided ideal [34, p.198]. We note that the set $Z_{\mathcal{P}}$ is also closed in the topology of the present article: each component in Ω^+ is equipped with the classical (analytic) topology.
Issues of stratification play a dominant role in [34]. The stratification of the tempered dual $\Pi^t(GL(n))$ arises from their construction of *tempered* K-types, see [34, p. 162, p. 189]. In the context of the present article, there is a natural stratification-by-dimension as follows. Let $1 \leq k \leq n$ and define

$$k\text{-stratum} = \{\mathcal{O}(\phi) \mid \dim_{\mathbb{C}} \mathcal{O}(\phi) \leq k\}$$

If $\pi_F(\mathcal{O}(\phi))$ is the complexification of the component $\Theta \subset \Pi^t(G)$ then we have

$$\dim_{\mathbb{R}} \Theta = \dim_{\mathbb{C}} \mathcal{O}(\phi).$$

The partial order in [34] on the components Θ transfers to a partial order on complex orbits $\mathcal{O}(\phi)$. This partial order originates in the opposite of the natural partial order on partitions, and the partitions manifest themselves in terms of Langlands parameters. For example, let

$$\begin{aligned}\phi &= \rho\otimes\mathrm{spin}(j_1)\oplus\cdots\oplus\rho\otimes\mathrm{spin}(j_r)\\ \phi' &= \rho\otimes\mathrm{spin}(j_1')\oplus\cdots\oplus\rho\otimes\mathrm{spin}(j_r')\end{aligned}$$

Let $\lambda_1 = 2j_1+1,\ldots,\lambda_r = 2j_r+1$, $\mu_1 = 2j_1'+1,\ldots,\mu_r = 2j_r'+1$ and define partitions as follows

$$\begin{aligned}\lambda &= (\lambda_1,\ldots,\lambda_r), \quad \lambda_1\geq\lambda_2\geq\ldots\\ \mu &= (\mu_1,\ldots,\mu_r), \quad \mu_1\geq\mu_2\geq\ldots\end{aligned}$$

The natural partial order on partitions is: $\lambda\leq\mu$ if and only if

$$\lambda_1+\cdots+\lambda_i\leq\mu_1+\cdots+\mu_i$$

for all $i\geq 1$, see [28, p.6]. Let $l(\lambda)$ be the length of λ, that is the number of parts in λ. Then $\dim_{\mathbb{C}}\mathcal{O}(\phi) = l(\lambda)$. Let λ', μ' be the dual partitions as in [28]. Then we have [28, 1.11] $\lambda\geq\mu$ if and only if $\mu'\geq\lambda'$. Note that $l(\lambda) = \lambda_1'$, $l(\mu) = \mu_1'$. Then

$$\Theta_\lambda\leq\Theta_\mu\Leftrightarrow\lambda\geq\mu\Leftrightarrow\mu'\geq\lambda'\Rightarrow\lambda_1'\leq\mu_1'$$

So if $\Theta_\lambda\leq\Theta_\mu$ then $\dim_{\mathbb{R}}\Theta_\lambda\leq\dim_{\mathbb{R}}\Theta_\mu$, similarly $\mathcal{O}(\phi)\leq\mathcal{O}(\phi')$ implies $\dim_{\mathbb{C}}\mathcal{O}(\phi)\leq\dim_{\mathbb{C}}\mathcal{O}(\phi')$. Stratification-by-dimension is finer than the Schneider-Zink stratification [34].

Let now R denote the ring of all regular functions on $\Pi(G)$. The ring R is a commutative, reduced, unital ring over $\mathbb{C}$ which is not finitely generated. We will call R the *extended centre* of G. It is natural to believe that the extended centre R of G is the centre of an 'extended category' made from smooth G-modules. The work of Schneider-Zink [34, p. 201] contains various results in this direction.

4. The q-projection

Let Ω be a component in the Bernstein variety. This component is an ordinary quotient D/Γ. We now consider the extended quotient $\widetilde{D}/\Gamma = \bigsqcup D^\gamma/Z_\gamma$, where D is the complex torus $\mathbb{C}^{\times m}$. Let γ be a permutation of n letters with cycle type

$$\gamma = (1\ldots\alpha_1)\cdots(1\ldots\alpha_r)$$

where $\alpha_1+\cdots+\alpha_r = m$. On the fixed set D^γ the map π_q, by definition, sends the element $(z_1,\ldots,z_1,\ldots,z_r,\ldots,z_r)$ where z_j is repeated α_j times, $1\leq j\leq r$, to the element

$$(q^{(\alpha_1-1)/2}z_1,\ldots,q^{(1-\alpha_1)/2}z_1,\ldots,q^{(\alpha_r-1)/2}z_r,\ldots,q^{(1-\alpha_r)/2}z_r)$$

The map π_q induces a map from D^γ/Z_γ to D/Γ, and so a map, still denoted π_q, from the extended quotient $\widetilde{D}/\Gamma$ to the ordinary quotient D/Γ. This creates a map π_q from the extended Bernstein variety to the Bernstein variety:

$$\pi_q : \Omega^+(G) \longrightarrow \Omega(G).$$

Definition 4.1. The map π_q is called the *q-projection.*

The q-projection π_q occurs in the following commutative diagram [8]:

$$\begin{array}{ccc} \Phi(G) & \longrightarrow & \Pi(G) \\ \alpha \big\downarrow & & \big\downarrow \text{inf. ch.} \\ \Omega^+(G) & \xrightarrow{\pi_q} & \Omega(G) \end{array}$$

Let A, B be commutative rings with $A \subset B, 1 \in A$. Then the element $x \in B$ is *integral* over A if there exist $a_1, \ldots, a_n \in A$ such that

$$x^n + a_1 x^{n-1} + \ldots + a_n = 0.$$

Then B is *integral* over A if each $x \in B$ is integral over A. Let X, Y be affine varieties, $f : X \longrightarrow Y$ a regular map such that $f(X)$ is dense in Y. Then the pull-back $f^{\#}$ defines an isomorphic inclusion $\mathbb{C}[Y] \longrightarrow \mathbb{C}[X]$. We view $\mathbb{C}[Y]$ as a subring of $\mathbb{C}[X]$ by means of $f^{\#}$. Then f is a *finite* map if $\mathbb{C}[X]$ is integral over $\mathbb{C}[Y]$, see [35]. This implies that the pre-image $F^{-1}(y)$ of each point $y \in Y$ is a finite set, and that, as y moves in Y, the points in $F^{-1}(y)$ may merge together but not disappear. The map $\mathbb{A}^1 - \{0\} \longrightarrow \mathbb{A}^1$ is the classic example of a map which is *not* finite.

Lemma 4.2. *Let X be a component in the extended variety $\Omega^+(G)$. Then the q-projection π_q is a finite map from X onto its image $\pi_q(X)$.*

Proof. Note that the fixed-point set D^γ is a complex torus of dimension r, that $\pi_q(D^\gamma)$ is a torus of dimension r and that we have an isomorphism of affine varieties $D^\gamma \cong \pi_q(D^\gamma)$. Let $X = D^\gamma/Z_\gamma, Y = \pi_q(D^\gamma)/\Gamma$ where Z_γ is the Γ-centralizer of γ. Now each of X and Y is a quotient of the variety D^γ by a finite group, hence X, Y are affine varieties [35, p.31]. We have $D^\gamma \longrightarrow X \longrightarrow Y$ and $\mathbb{C}[Y] \longrightarrow \mathbb{C}[X] \longrightarrow \mathbb{C}[D^\gamma]$. According to [35, p.61], $\mathbb{C}[D^\gamma]$ is integral over $\mathbb{C}[Y]$ since $Y = D^\gamma/\Gamma$. Therefore the subring $\mathbb{C}[X]$ is integral over $\mathbb{C}[Y]$. So the map $\pi_q : X \longrightarrow Y$ is finite. $\square$

Example 4.3. $GL(2)$. Let T be the diagonal subgroup of $G = GL(2)$ and let Ω be the component in $\Omega(G)$ containing the cuspidal pair $(T, 1)$. Then $\sigma \in \Pi(GL(2))$ is *arithmetically unramified* if $inf.ch.\sigma \in \Omega$. If $\pi_F(\phi) = \sigma$ then ϕ is a 2-dimensional representation of $\mathcal{L}_F$ and there are two possibilities:
ϕ is reducible, $\phi = \psi_1 \oplus \psi_2$ with ψ_1, ψ_2 unramified quasicharacters of W_F. So $\psi_j(w) = z_j^{d(w)}, z_j \in \mathbb{C}^\times, j = 1, 2$. We have $\pi_F(\phi) = Q(\psi_1, \psi_2)$ where ψ_1 does not precede ψ_2. In particular we obtain the 1-dimensional representations of G as follows:

$$\pi_F(|\ |^{1/2}\psi \oplus |\ |^{-1/2}\psi) = Q(|\ |^{1/2}\psi, |\ |^{-1/2}\psi) = \psi \circ det.$$

ϕ is irreducible, $\phi = \psi \otimes \mathrm{spin}(1/2)$. Then $\pi_F(\phi) = Q(\Delta)$ with $\Delta = \{|\ |^{-1/2}\psi, |\ |^{1/2}\psi\}$ so $\pi_F(\phi) = \psi \otimes St(2)$ where $St(2)$ is the Steinberg representation of $GL(2)$.
The orbit of $(T,1)$ is $D = (\mathbb{C}^\times)^2$, and $W(T,D) = \mathbb{Z}/2\mathbb{Z}$. Then $\Omega \cong (\mathbb{C}^\times)^2/\mathbb{Z}/2\mathbb{Z} \cong \mathrm{Sym}^2\,\mathbb{C}^\times$. The extended quotient is $\Omega^+ = \mathrm{Sym}^2\ \mathbb{C}^\times \sqcup \mathbb{C}^\times$. The $q-$projection works as follows:

$$\pi_q : \{z_1, z_2\} \mapsto \{z_1, z_2\}$$

$$\pi_q : z \mapsto \{q^{1/2}z, q^{-1/2}z\}$$

where q is the cardinality of the residue field of F.

Let $A = \mathcal{H}(GL(2)//I)$ be the Iwahori-Hecke algebra of $GL(2)$. This is a finite type algebra. Following [21, p. 327], denote by $\mathrm{Prim}_n(A) \subset \mathrm{Prim}(A)$ the set of primitive ideals $B \subset A$ which are kernels of irreducible representations of A of dimension n. Set $X_1 = \mathrm{Prim}_1(A)$, $X_2 = \mathrm{Prim}_1(A) \sqcup \mathrm{Prim}_2(A) = \mathrm{Prim}(A)$. Then X_1 and X_2 are closed sets in $\mathrm{Prim}(A)$ defining an increasing filtration of $\mathrm{Prim}(A)$. Now A is Morita equivalent to the Bernstein ideal $\mathcal{H}(\Omega)$, and $\Pi(\Omega) \simeq \mathrm{Prim}(A)$.
Let $\phi_1 = 1 \otimes \mathrm{spin}(1/2)$, $\phi_2 = 1 \otimes 1 \oplus 1 \otimes 1$. The 1-dimensional representations of $GL(2)$ determine 1-dimensional representations of $\mathcal{H}(G//I)$ and so lie in X_1. The L-parameters of the 1-dimensional representations of $GL(2)$ do *not* lie in the 1-dimensional orbit $\mathcal{O}(\phi_1)$: they lie in the 2-dimensional orbit $\mathcal{O}(\phi_2)$. The Kazhdan-Nistor-Schneider stratification [21] does *not* coincide with stratification-by-dimension.

Example 4.4. $GL(3)$. In the above example, the q-projection is stratified-injective, i.e. injective on each orbit type. This is not so in general, as shown by the next example. Let T be the diagonal subgroup of $GL(3)$ and let Ω be the component containing the cuspidal pair $(T,1)$. Then $\Omega = \mathrm{Sym}^3\,\mathbb{C}^\times$ and

$$\Omega^+ = \mathrm{Sym}^3\,\mathbb{C}^\times \sqcup (\mathbb{C}^\times)^2 \sqcup \mathbb{C}^\times$$

.
The map π_q works as follows:

$$\{z_1, z_2, z_3\} \mapsto \{z_1, z_2, z_3\}$$
$$(z, w, w) \mapsto \{z, q^{1/2}w, q^{-1/2}w\}$$
$$(z, z, z) \mapsto \{qz, z, q^{-1}z\}.$$

Consider the L-parameter

$$\phi = \psi_1 \otimes 1 \oplus \psi_2 \otimes \mathrm{spin}(1/2) \in \Phi(GL(3)).$$

If $\psi(w) = z^{d(w)}$ then we will write $\psi = z$. With this understood, let

$$\phi_1 = q \otimes 1 \oplus q^{-1/2} \otimes \mathrm{spin}(1/2)$$

$$\phi_2 = q^{-1} \otimes 1 \oplus q^{1/2} \otimes \mathrm{spin}(1/2).$$

Then $\alpha(\phi_1), \alpha(\phi_2)$ are distinct points in the same stratum of the extended quotient, but their image under the q-projection π_q is the single point $\{q^{-1}, 1, q\} \in \operatorname{Sym}^3 \mathbb{C}^\times$.
Let

$$\phi_3 = 1 \otimes \operatorname{spin}(3/2)$$
$$\phi_4 = q^{-1} \otimes 1 \oplus 1 \otimes 1 \oplus q \otimes 1.$$

Then the distinct L-parameters $\phi_1, \phi_2, \phi_3, \phi_4$ all have the same image under the q-projection π_q.

5. The diagram

In this section we create a diagram which incorporates several major results. The following diagram serves as a framework for the whole article:

$$\begin{tikzcd}
K_*^{\mathrm{top}}(G) \arrow[rr, "\mu"] \arrow[d, "\mathrm{ch}"'] & & K_*(C_r^*(G)) \arrow[d, "\mathrm{ch}"] \\
\mathrm{H}_*(G; \beta G) \arrow[r] \arrow[d, dotted] & \mathrm{HP}_*(\mathcal{H}(G)) \arrow[r, "\imath_*"] \arrow[d] & \mathrm{HP}_*(\mathcal{S}(G)) \arrow[d] \\
\mathrm{H}_c^*(\Phi(G); \mathbb{C}) \arrow[r] & \mathrm{H}_c^*(\Pi(G); \mathbb{C}) \arrow[r] & \mathrm{H}_c^*(\Pi^t(G); \mathbb{C})
\end{tikzcd}$$

The Baum-Connes assembly map μ is an isomorphism [1, 24]. The map

$$\mathrm{H}_*(G; \beta G) \to \mathrm{HP}_*(\mathcal{H}(G))$$

is an isomorphism [20, 33]. The map $\imath_*$ is an isomorphism by Theorem 3.3. The right hand Chern character is constructed in [9] and is an isomorphism after tensoring over $\mathbb{Z}$ with $\mathbb{C}$ [9, Theorem 3]. The Chern character on the left hand side of the diagram is the unique map for which the top half of the diagram is commutative.
In the diagram, $\mathrm{H}_c^*(\Pi^t(G); \mathbb{C})$ denotes the (periodised) compactly supported de Rham cohomology of the tempered dual $\Pi^t(G)$, and $\mathrm{H}_c^*(\Pi(G); \mathbb{C})$ denotes the (periodised) de Rham cohomology supported on finitely many components of the smooth dual $\Pi(G)$. The map

$$\mathrm{HP}_*(\mathcal{S}(G)) \to \mathrm{H}_c^*(\Pi^t(G); \mathbb{C})$$

is constructed in [7] and is an isomorphism [7, Theorem 7].
The map

$$\mathrm{H}_c^*(\Pi(G); \mathbb{C}) \to \mathrm{H}_c^*(\Pi^t(G); \mathbb{C})$$

is constructed in the following way. Given an L-parameter $\phi : \mathcal{L}_F \to GL(n, \mathbb{C})$ we have

$$\phi = \phi_1 \oplus \ldots \oplus \phi_m$$

with each ϕ_j an irreducible representation. We have $\phi_j = \rho_j \otimes \operatorname{spin}(j)$ where each ρ_j is an irreducible representation of the Weil group W_F. We shall assume that $\det \rho_j$ is a unitary character. Let $\mathcal{O}(\phi)$ be the orbit of ϕ as in Definition 1.4. The map $\mathcal{O}(\phi) \to \mathcal{O}^t(\phi)$ is now defined as follows

$$\psi_1\phi_1 \oplus \ldots \oplus \psi_m\phi_m \mapsto |\psi_1|^{-1} \cdot \psi_1\phi_1 \oplus \ldots \oplus |\psi_m|^{-1} \cdot \psi_m\phi_m.$$

This map is a deformation retraction of the complex orbit $\mathcal{O}(\phi)$ onto the compact orbit $\mathcal{O}^t(\phi)$. Since $\Phi(G)$ is a disjoint union of such complex orbits this formula determines, via the local Langlands correspondence for $GL(n)$, a deformation retraction of $\Pi(G)$ onto the tempered dual $\Pi^t(GL(n))$, which implies that the induced map on cohomology is an isomorphism.
The map

$$\mathrm{H}^*_c(\Phi(G);\mathbb{C}) \to \mathrm{H}^*_c(\Pi(G);\mathbb{C})$$

is an isomorphism, induced by the local Langlands correspondence π_F.
The map

$$\mathrm{HP}_*(\mathcal{H}(G)) \to \mathrm{H}^*_c(\Pi(G);\mathbb{C})$$

is an isomorphism by Theorem 3.1.
There is at present no direct definition of the map

$$\mathrm{H}_*(G;\beta G) \to \mathrm{H}^*_c(\Phi(G);\mathbb{C}).$$

Suppose for the moment that F has characteristic 0 and has residue field of characteristic p. An irreducible representation ρ of the Weil group W_F is called wildly ramified if $\dim \rho$ is a power of p and $\rho \not\simeq \rho \otimes \psi$ for any unramified quasicharacter $\psi \neq 1$ of W_F. We write $\Phi^{wr}_m(F)$ for the set of equivalence classes of such representations of dimension p^m. An irreducible supercuspidal representation π of $GL(n)$ is wildly ramified if n is a power of p and $\pi \not\simeq \pi \otimes (\psi \circ \det)$ for any unramified quasicharacter $\psi \neq 1$ of $F^\times$. We write $\Pi^{wr}_m(F)$ for the set of equivalence classes of such representations of $GL(p^m, F)$. In this case Bushnell-Henniart [10] construct, for each m, a canonical bijection

$$\pi_{F,m} : \Phi^{wr}_m(F) \to \Pi^{wr}_m(F).$$

Now the maximal simple type (J, λ) of an irreducible supercuspidal representation determines an element in the chamber homology of the affine building [2, 6.7]. The construction of Bushnell-Henniart therefore determines a map from a *subspace* of $\mathrm{H}^{\mathrm{even}}_c(\Phi(G);\mathbb{C})$ to a *subspace* of $\mathrm{H}_0(G;\beta G)$.
In the context of the above diagram the Baum-Connes map has a geometric counterpart: it is induced by the deformation retraction of $\Pi(GL(n))$ onto the tempered dual $\Pi^t(GL(n))$.

6. Supercuspidal representations of $GL(n)$

In this section we track the fate of supercuspidal representations of $GL(n)$ through the diagram constructed in the previous Section. Let ρ be an irreducible n-dimensional complex representation of the Weil group W_F such that $\det \rho$ is a unitary character and let $\phi = \rho \otimes 1$. Then ϕ is the L-parameter for a pre-unitary supercuspidal representation ω of $GL(n)$. Let $\mathcal{O}(\phi)$ be the orbit of ϕ and $\mathcal{O}^t(\phi)$ be the compact orbit of ϕ. Then $\mathcal{O}(\phi)$ is a component in the

Bernstein variety isomorphic to $\mathbb{C}^\times$ and $\mathcal{O}^t(\phi)$ is a component in the tempered dual, isomorphic to $\mathbb{T}$. The L-parameter ϕ now determines the following data.

6.1. Let (J, λ) be a maximal simple type for ω in the sense of Bushnell and Kutzko [11, chapter 6]. Then J is a compact open subgroup of G and λ is a smooth irreducible complex representation of J.
We will write

$$\mathbb{T} = \{\psi \otimes \omega : \psi \in \Psi^t(G)\}$$

where $\Psi^t(G)$ denotes the group of unramified unitary characters of G.

Theorem 6.1. *Let K be a maximal compact subgroup of G containing J and form the induced representation $W = Ind_J^K(\lambda)$. We then have*

$$\ell^2(G \times_K W) \simeq Ind_K^G(W) \simeq Ind_J^G(\lambda) \simeq \int_{\mathbb{T}} \pi d\pi.$$

Proof. The supercuspidal representation ω contains λ and, modulo unramified unitary twist, is the only irreducible unitary representation with this property [11, 6.2.3]. Now the Ahn reciprocity theorem expresses Ind_J^G as a direct integral [26, p.58]:

$$\mathrm{Ind}_J^G(\lambda) = \int n(\pi, \lambda)\pi d\pi$$

where $d\pi$ is Plancherel measure and $n(\pi, \lambda)$ is the multiplicity of λ in $\pi|_J$. But the Hecke algebra of a maximal simple type is commutative (a Laurent polynomial ring). Therefore $\omega|_J$ contains λ with multiplicity 1 (thanks to C. Bushnell for this remark). We then have $n(\psi \otimes \omega, \lambda) = 1$ for all $\psi \in \Psi^t(G)$. We note that Plancherel measure induces Haar measure on $\mathbb{T}$, see [31].
The affine building of G is defined as follows [38, p. 49]:

$$\beta G = \mathbb{R} \times \beta SL(n)$$

where $g \in G$ acts on the affine line $\mathbb{R}$ via $t \mapsto t + val(det(g))$. Let $G^\circ = \{g \in G : val(det(g)) = 0\}$. We use the standard model for $\beta SL(n)$ in terms of equivalence classes of $\mathbf{o}_F$ - lattices in the n-dimensional F-vector space V. Then the vertices of $\beta SL(n)$ are in bijection with the maximal compact subgroups of G°, see [32, 9.3]. Let $P \in \beta G$ be the vertex for which the isotropy subgroup is $K = GL(n, \mathbf{o}_F)$. Then the G-orbit of P is the set of all vertices in βG and the discrete space G/K can be identified with the set of vertices in the affine building βG. Now the base space of the associated vector bundle $G \times_K W$ is the discrete coset space G/K, and the Hilbert space of ℓ^2-sections of this homogeneous vector bundle is a realization of the induced representation $\mathrm{Ind}_K^G(W)$. □

The $C_0(\beta G)$-module structure is defined as follows. Let $f \in C_0(\beta G)$, $s \in \ell^2(G \times_K W)$ and define

$$(fs)(v) = f(v)s(v)$$

for each vertex $v \in \beta G$. We proceed to construct a K-cycle in degree 0. This K-cycle is

$$(C_0(\beta G), \ell^2(G \times_K W) \oplus 0, 0)$$

interpreted as a $\mathbb{Z}/2\mathbb{Z}$-graded module. This triple satisfies the properties of a (pre)-Fredholm module [14, IV] and so creates an element in $K_0^{\mathrm{top}}(G)$. By Theorem 5.1 this generator creates a free $C(\mathbb{T})$-module of rank 1, and so provides a generator in $K_0(C_r^*(G))$.

6.2. The Hecke algebra of the maximal simple type (J, λ) is commutative (the Laurent polynomials in one complex variable). The periodic cyclic homology of this algebra is generated by 1 in degree zero and dz/z in degree 1.
The corresponding summand of the Schwartz algebra $\mathcal{S}(G)$ is Morita equivalent to the Fréchet algebra $C^\infty(\mathbb{T})$. By an elementary application of Connes' theorem [14, Theorem 2, p. 208], the periodic cyclic homology of this Fréchet algebra is generated by 1 in degree 0 and $d\theta$ in degree 1.

6.3. The corresponding component in the Bernstein variety is a copy of $\mathbb{C}^\times$. The cohomology of $\mathbb{C}^\times$ is generated by 1 in degree 0 and $d\theta$ in degree 1.
The corresponding component in the tempered dual is the circle $\mathbb{T}$. The cohomology of $\mathbb{T}$ is generated by 1 in degree 0 and $d\theta$ in degree 1.

References

[1] P. Baum, N. Higson, R. J. Plymen, A proof of the Baum-Connes conjecture for p-adic $GL(n)$, C. R. Acad. Sci. Paris 325 (1997) 171-176.

[2] P. Baum, N. Higson, R. J. Plymen, Representation theory of p-adic groups: a view from operator algebras, Proc. Symp. Pure Math. 68 (2000) 111 – 149.

[3] P. Baum, V. Nistor, The periodic cyclic homology of Iwahori-Hecke algebras, C. R. Acad. Sci. Paris 332 (2001) 1 –6.

[4] P. Baum, V. Nistor, The periodic cyclic homology of Iwahori-Hecke algebras, preprint 2001.

[5] J. Bernstein, Representations of p-adic groups, Notes by K.E. Rumelhart, Harvard University, 1992.

[6] N. Bourbaki, Groupes et algèbres de Lie, Chapitre 9, Masson, Paris 1982.

[7] J. Brodzki, R.J. Plymen, Periodic cyclic homology of certain nuclear algebras, C. R. Acad. Sci. Paris., 329 (1999), 671–676.

[8] J. Brodzki, R. J. Plymen, Geometry of the smooth dual of $GL(n)$, C. R. Acad. Sci. Paris., 331 (2000), 213–218.

[9] J. Brodzki, R.J. Plymen, Chern character for the Schwartz algebra of p-adic $GL(n)$, Bull. London. Math. Soc. 34 (2002), 219-228

[10] C.J. Bushnell and G. Henniart, Local tame lifting for $GL(n)$ II: wildly ramified supercuspidals, Astérisque 254, SMF 1999.

[11] C.J. Bushnell and P.C. Kutzko, The admissible dual of $GL(N)$ via compact open subgroups, Annals of Math. Studies 129, Princeton University Press, Princeton, 1993.

[12] C.J. Bushnell and P.C. Kutzko, Smooth representations of p-adic reductive groups: Structure theory via types, Proc. London Math. Soc 77(1998) 582 – 634.

[13] C.J. Bushnell and P.C. Kutzko, Semisimple types for $GL(N)$, Compositio Math. 119 (1999) 53 –97.

[14] A. Connes, Noncommutative Geometry, Academic Press, New York 1994.

[15] D. Eisenbud, J. Harris, The geometry of schemes, Graduate Text 197, Springer, Berlin 2000.

[16] Harish-Chandra, Collected papers Volume IV, Springer 1984.

[17] M. Harris, R. Taylor, On the geometry and cohomology of some simple Shimura varieties, Ann. Math. Study 151, Princeton University Press 2001.

[18] R. Hartshorne, Algebraic geometry, Springer, Berlin 1977.

[19] G. Henniart, Une preuve simple des conjectures de Langlands pour $GL(n)$ sur un corps p-adique, Invent. Math. 139 (2000) 439 –455.

[20] N. Higson, V. Nistor, Cyclic homology of totally disconnected groups acting on buildings, J. Functional Analysis 141 (1996), 466–485.

[21] D. Kazhdan, V. Nistor, P. Schneider, Hochschild and cyclic homology of finite type algebras, Sel. Math. 4 (1998), 321-359.

[22] A. W. Knapp, Introduction to the Langlands program, Proc. Symp. Pure Math. 61 (1997) 245–302.

[23] S. S. Kudla, The local Langlands correspondence, Proc. Symp. Pure Math. 55 (1994) 365 –391.

[24] V. Lafforgue, Une démonstration de la conjecture de Baum-Connes pour les groupes réductifs sur un corps p-adique et pour certains groupes discrets possédant la proriété (T), C. R. Acad. Sci. Paris 327 (1998), 439–444.

[25] G. Laumon, M. Rapoport, U. Stuhler, $\mathcal{D}$-elliptic sheaves and the Langlands correspondence, Invent. Math. 113 (1993) 217 –338.

[26] R.L. Lipsman, Group representations, Lecture Notes in Math. 388, Springer, 1974.

[27] G. Lusztig, Cells in affine Weyl groups II, J. Algebra 109 (1987), 536-548.

[28] I. G. Macdonald, Symmetric functions and Hall polynomials, Oxford Math. Monograph, Oxford 1979.

[29] J. Neukirch, Algebraic Number Theory, Springer, Berlin, 1999.

[30] J. Neukirch, Local class field theory, Springer, Berlin 1986.

[31] R.J. Plymen, Reduced C^*-algebra of the p-adic group $GL(n)\,II$, J. Functional Analysis, to appear.

[32] M. Ronan, Lectures on buildings, Perspectives in Math. 7, Academic Press, 1989.

[33] P. Schneider, The cyclic homology of p-adic reductive groups, J. reine angew. Math. 475 (1996), 39-54.

[34] P. Schneider, E.-W. Zink, K-types for the tempered components of a p-adic general linear group, J. reine angew. Math., 517 (1999) 161-208.

[35] I. R. Shafarevich, Basic algebraic geometry 1, Springer, Berlin, 1994.

[36] K. E. Smith, L. Kahanpää, P. Kekäläinen, W. Traves, An invitation to algebraic geometry, Universitext, Springer, Berlin, 2000.

[37] J. Tate, Number theoretic background, Proc. Symp. Pure Math. 33 (1979) part 2, 3 - 26.

[38] J. Tits, Reductive groups over local fields, Proc. Symp. Pure Math. 33 (1979) part 1, 29 – 69.
[39] A. Weil, Basic number theory, Classics in Math., Springer 1995.
[40] N. Xi, The based ring of two-sided cells of affine Weyl groups of type $\tilde{A}_{n-1}$, AMS Memoir 749 (2002).

Jacek Brodzki
Faculty of Mathematical Studies
University of Southampton
Southampton SO17 1BJ, U.K.
j.brodzki@maths.soton.ac.uk

Roger Plymen
Department of Mathematics
University of Manchester
Manchester M13 9PL, U.K.
roger@maths.man.ac.uk

Equivariant Fiber Polytopes

To the memory of Rodica Simion.

Victor Reiner

Received: February 15, 2002

Communicated by Günter M. Ziegler

Abstract. The equivariant generalization of Billera and Sturmfels' fiber polytope construction is described. This gives a new relation between the associahedron and cyclohedron, a different natural construction for the type B permutohedron, and leads to a family of order-preserving maps between the face lattice of the type B permutohedron and that of the cyclohedron

2000 Mathematics Subject Classification: 52B12, 52B15
Keywords and Phrases: fiber polytope, associahedron, cyclohedron, equivariant

Contents

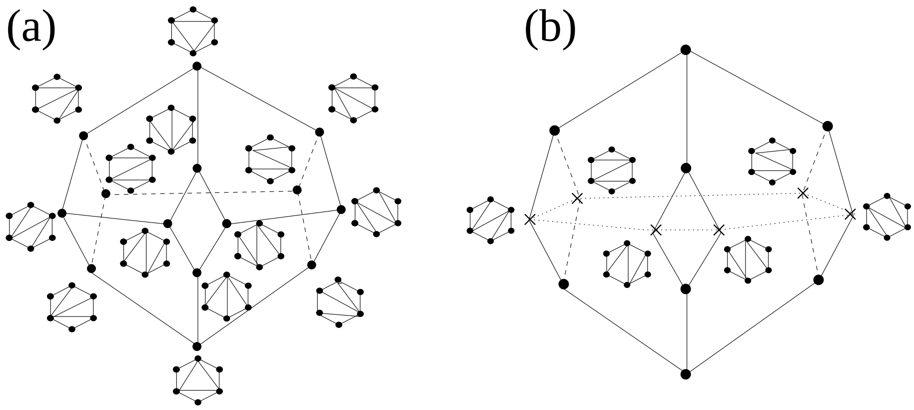


FIGURE 1. The 3-dimensional associahedron, containing the 2-dimensional cyclohedron as a planar slice.

1. Introduction

Much of this paper is motivated by a new relationship between two families of convex polytopes which have appeared in diverse places within topology, geometry, and combinatorics [4, 10, 12, 13, 16]: the *associahedra* (or *Stasheff polytopes*), and the *cyclohedra* (or *type B associahedra*).

There is already a well-known relation between these two families: any face in a cyclohedron is combinatorially isomorphic to a Cartesian product of a lower-dimensional cyclohedron along with a collection of lower-dimensional associahedra [16, §3.3], [6, Proposition 3.2.1]. Our point of departure is a different relationship, illustrated in Figure 1. The 3-dimensional associahedron is depicted in Figure 1(a), with its vertices indexed by triangulations of a (centrally symmetric) hexagon. Figure 1(b) shows inside it a hexagonal slice representing the 2-dimensional cyclohedron, with vertices indexed by the subset of triangulations possessing central symmetry – this slice is the invariant subpolytope for a reflection symmetry acting on the associahedron, which swaps two vertices if they correspond to triangulations that differ by 180° rotation.

Our main result (Theorem 2.10) asserts a similar relationship generally when one considers subdivisions of a polytope which are invariant under some finite group of symmetries. The theory of fiber polytopes introduced by Billera and Sturmfels [1] shows that whenever one has a linear surjection of convex polytopes $P \xrightarrow{\pi} Q$, there is an associated convex polytope $\Sigma(P \xrightarrow{\pi} Q)$, called the *fiber polytope*. This fiber polytope has dimension $\dim P - \dim Q$, and its faces correspond (roughly) to the subdivisions of Q by cells which are projections of families of faces of P. Our result says that when the projection $P \xrightarrow{\pi} Q$ is G-equivariant for some finite group G acting as symmetries on both P and Q, then G acts as a group of symmetries on $\Sigma(P \xrightarrow{\pi} Q)$, and the G-invariant subpolytope $\Sigma^G(P \stackrel{\pi}{\to} Q)$ (the *equivariant fiber polytope*) is a polytope of dimension $\dim P^G - \dim Q^G$ whose faces correspond to those subdivisions which are G-invariant.

The paper is structured as follows. Section 2 proves the main results, with Subsection 2.1 containing the technical details needed to generalize fiber polytopes to the equivariant setting. Theorem 2.10 on the existence and dimension of the equivariant fiber polytope is deduced in Subsection 2.2. The special case of equivariant secondary polytopes is discussed in Subsection 2.3.

Section 3 contains some low-dimensional examples that are easily visualized, while Section 4 gives some general examples as applications. In particular, Example 4.1 explains the above relation between associahedra and cyclohedra, and Example 4.2 explains how the *type B permutohedron* (that is, the zonotope generated by the root system of type B) occurs as an equivariant fiber polytope. In Section 4.3 we answer a question of R. Simion, by exhibiting a family of natural maps between the face lattices of the B_n-permutohedron and n-dimensional cyclohedron.

Section 5 lists some remarks and open questions.

2. Equivariant polytope bundles

2.1. Group actions on polytope bundles. For convenience, we work with the same notation as in [1, §1] in working out the equivariant versions of the same results.

Let $\mathcal{B} \to Q$ be a *polytope bundle*, that is, Q is a convex polytope in $\mathbb{R}^d$, and for each x in Q, the set $\mathcal{B}_x$ is a convex polytope in $\mathbb{R}^n$, such that the graph $\bigcup\{\mathcal{B}_x \times x : x \in Q\}$ is a bounded Borel subset of $\mathbb{R}^{n+d}$. We further assume that we have a finite group G acting linearly on both $\mathbb{R}^d$ and $\mathbb{R}^n$.

Definition 2.1. Say that $\mathcal{B} \to Q$ is a *G-equivariant polytope bundle* if

- G acts as symmetries of Q, i.e. $g(Q) = Q$ for all g in G. In particular, without loss of generality, the centroid of Q is the origin 0.
- for every x in Q and g in G one has $\mathcal{B}_{g(x)} = g(\mathcal{B}_x)$.

Alternatively, equivariance of a polytope bundle is equivalent to G-invariance of Q, along with G-invariance of $\mathcal{B}$ with respect to a natural G-action on polytope bundles: given $\mathcal{B} \to Q$ and g in G, let $g(\mathcal{B} \to Q)$ be the polytope bundle defined by $g(\mathcal{B} \to Q)_x := g(\mathcal{B}_{g^{-1}x})$.

The linear action of G on $\mathbb{R}^n$ induces a *contragredient* action on the dual space $(\mathbb{R}^n)^*$ of functionals: $g(\psi)(x) := \psi(g^{-1}(x))$ for $g \in G, y \in \mathbb{R}^n, \psi \in (\mathbb{R}^n)^*$.

A *section* γ of $\mathcal{B} \to Q$ is a choice of $\gamma(x) \in \mathcal{B}_x$ for each x. If $\mathcal{B} \to Q$ is equivariant, there is a G-action on sections defined by $g(\gamma)(x) := g\gamma(g^{-1}x)$.

For any of these G-actions, define the *averaging (or Reynolds) operator*

$$\pi_G = \frac{1}{|G|} \sum_{g \in G} g$$

which is an idempotent projector onto the subset (or subspace) of G-invariants, e.g. π maps $\mathbb{R}^n \to (\mathbb{R}^n)^G$ and maps $(\mathbb{R}^n)^* \to ((\mathbb{R}^n)^*)^G$.

It turns out that much of the reason that fiber polytopes interact well with finite group actions boils down to π_G being a *linear* operator which is a *convex* combination of the group operations g in G.

Recall that for a polytope bundle $\mathcal{B} \to Q$, the Minkowski integral $\int_Q \mathcal{B}$ is the subset of $\mathbb{R}^n$ consisting of all integrals $\int_Q \gamma$ of measurable sections γ, and that this is a non-empty compact, convex subset of $\mathbb{R}^n$.

PROPOSITION 2.2. *Integration commutes with the G-action on sections of a G-equivariant polytope bundle $\mathcal{B} \to Q$:*

$$\int_Q g\gamma = g\left(\int_Q \gamma\right)$$

for all g in G.

Proof.

$$\begin{aligned}\int_Q g\gamma &= \int_Q g(\gamma(g^{-1}x))dx \\ &= g\left(\int_Q \gamma(g^{-1}x)dx\right) \\ &= g\left(\int_Q \gamma(u)du\right)\end{aligned}$$

Here the second equality uses linearity of g and linearity of integration. The third equality comes from the change of variable $x = g(u)$, using the fact that the Jacobian determinant for this change of variable is $\det(g)$, which must be ± 1, since g is an element of finite order in $GL(\mathbb{R}^n)$. $\square$

COROLLARY 2.3. *For any G-equvariant polytope bundle $\mathcal{B} \to Q$, the group G acts on the convex set $\int_Q \mathcal{B}$.*
Furthermore, one has

$$\begin{aligned}\left(\int_Q \mathcal{B}\right)^G &:= \left(\int_Q \mathcal{B}\right) \cap (\mathbb{R}^n)^G \\ &= \pi_G\left(\int_Q \mathcal{B}\right) \\ &= \left\{\int_Q \gamma : G\text{-equivariant, measurable sections } \gamma\right\},\end{aligned} \tag{2.1}$$

Proof. Proposition 2.2 implies the first assertion.
For the second equality in (2.1), we claim more generally that

$$C \cap (\mathbb{R}^n)^G = \pi_G(C)$$

for *any* convex G-invariant subset $C \subset \mathbb{R}^n$. To see this, note that the left-hand side is contained in the right because of idempotence of π_G. The right-hand side is obviously contained in $(\mathbb{R}^n)^G$. It also lies in C since for any x in C, the convex combination $\pi_G(x) = \frac{1}{|G|}\sum_{g \in G} g(x)$ will also lie in C.

For the third equality in (2.1), note that the right-hand side is contained in the left since the integral $\int_Q \gamma$ of any G-equivariant section will be a G-invariant point of $\mathbb{R}^n$ by Proposition 2.2. Conversely, a typical point on the left-hand side is $\int_Q \gamma$ where γ is a section such that $g(\int_Q \gamma) = \int_Q \gamma$ for all $g \in G$, and one can check using Proposition 2.2 that the G-equivariant section $\pi_G \gamma$ has the same integral:

$$\begin{aligned} \int_Q \pi_G \gamma &= \frac{1}{|G|} \sum_{g \in G} \int_Q g\gamma \\ &= \frac{1}{|G|} \sum_{g \in G} g \left(\int_Q \gamma \right) \\ &= \int_Q \gamma \end{aligned}$$

□

We wish to interpret faces of $\left(\int_Q \mathcal{B}\right)^G$ in terms of face bundles. Recall that a *face bundle* $\mathcal{F} \to Q$ of $\mathcal{B} \to Q$ is a polytope bundle in which $\mathcal{F}_x$ is a face of $\mathcal{B}_x$ for every x in Q. A *coherent face bundle* of $\mathcal{B} \to Q$ is one of the form $\mathcal{B}^\psi \to Q$ having $\mathcal{B}^\psi_x := (\mathcal{B}_x)^\psi$, where $\psi \in (\mathbb{R}^n)^*$ is any linear functional and P^ψ denotes the face of P on which the functional ψ is maximized.
When $\mathcal{B} \to Q$ is G-equivariant, the G-action on bundles restricts to a G-action on face bundles, and as before, a face bundle is G-equivariant if and only if it is invariant under the G-action. The next proposition points out the compatibility between the G-action on face bundles and the G-action on functionals.

Proposition 2.4. *For any G-equivariant polytope bundle $\mathcal{B} \to Q$ and any functional g in $(\mathbb{R}^n)^*$, the face bundle $\mathcal{B}^{g\psi} \to Q$ coincides with the bundle $g(\mathcal{B}^\psi \to Q)$.*
Consequently, $\mathcal{B}^\psi$ is a G-equivariant (coherent) face bundle if and only if $\mathcal{B}^\psi = \mathcal{B}^{\pi_G \psi}$.

Proof. For the first assertion, we compute

$$\begin{aligned} g(\mathcal{B}^\psi \to Q)_x &:= g(\mathcal{B}^\psi_{g^{-1}x}) \\ &= g(\{y \in \mathcal{B}_{g^{-1}x} : \psi(y) \text{ is maximized}\}) \\ &= \{y' \in g\mathcal{B}_{g^{-1}x} : \psi(g^{-1}y') \text{ is maximized}\}) \\ &= \{y' \in \mathcal{B}_x : g(\psi)(y') \text{ is maximized}\}) \\ &= \mathcal{B}^{g\psi}_x. \end{aligned}$$

For the second, note that

$$\begin{aligned} \mathcal{B}^\psi \text{ is } G\text{-equivariant} &\Leftrightarrow g(\mathcal{B}^\psi) = \mathcal{B}^\psi \quad \forall g \in G \\ &\Leftrightarrow \mathcal{B}^{g\psi}_x = \mathcal{B}^\psi_x \quad \forall g \in G,\, x \in Q \\ &\Leftrightarrow \mathcal{B}^{\pi_G\psi}_x = \mathcal{B}^\psi_x \quad \forall x \in Q \end{aligned}$$

where the last equality uses the fact that if for every g in G the functional $g\psi$ maximizes on the same face of the polytope $\mathcal{B}_x$, then the convex combination $\pi_G\psi$ will also maximize on this face. $\square$

We recall also these key results from [1].

PROPOSITION 2.5. [1, Prop. 1.2]. *The Minkowski integral commutes with taking faces (face bundles) in the following sense*

$$\left(\int_Q \mathcal{B}\right)^{\psi} = \int_Q \mathcal{B}^{\psi} \quad \forall \psi \in (\mathbb{R}^n)^*. \quad \square$$

THEOREM 2.6. [1, Thm. 1.3, Cor. 1.4] *If $\mathcal{B} \to Q$ is piecewise-linear, then $\int_Q \mathcal{B}$ is a convex polytope. Furthermore, the map $\mathcal{B}^{\psi} \mapsto \int_Q \mathcal{B}^{\psi}$ induces an isomorphism from its face lattice to the poset of coherent face bundles of $\mathcal{B} \to Q$ ordered by inclusion.* $\square$

Here is the equivariant generalization.

THEOREM 2.7. *Let $\mathcal{B} \to Q$ be any G-equivariant piecewise-linear polytope bundle. Then $\left(\int_Q \mathcal{B}\right)^G$ is a convex polytope whose face lattice is isomorphic to the poset of G-equivariant coherent face bundles of $\mathcal{B} \to Q$ ordered by inclusion.*

Proof. We use the following well-known fact about face lattices of affine images of polytopes:

LEMMA 2.8. [1, Lemma 2.2] *For any affine surjection of polytopes $\hat{P} \xrightarrow{f} P$, the map sending a face F of P to the face $f^{-1}(F)$ of $\hat{P}$ embeds the face lattice of P as the subposet of faces of $\hat{P}$ of the form $P^{\psi \circ f}$ for some ψ in $(\mathbb{R}^n)^*$.* $\square$

Applying this lemma to the surjection $\int_Q \mathcal{B} \xrightarrow{\pi_g} (\int_Q \mathcal{B})^G$, we conclude that $(\int_Q \mathcal{B})^G$ has face poset isomorphic to the subposet of faces of $\int_Q \mathcal{B}$ consisting of all faces of the form

$$\begin{aligned}
&\left(\int_Q \mathcal{B}\right)^{\psi \circ \pi_G} \quad \text{for } \psi \in (\mathbb{R}^n)^* \\
&= \left(\int_Q \mathcal{B}\right)^{\pi_G \psi} \quad \text{for } \psi \in (\mathbb{R}^n)^* \\
&= \int_Q \mathcal{B}^{\pi_G \psi} \text{ for } \psi \in (\mathbb{R}^n)^*
\end{aligned}$$

where the last equality uses Proposition 2.5. By Proposition 2.4, the set of faces $\int_Q \mathcal{B}^{\pi_G\psi}$ of $\int_Q \mathcal{B}$ for $\psi \in (\mathbb{R}^n)^*$ is exactly the same as the subset of faces $\int_Q \mathcal{B}^{\psi}$ for $\psi \in (\mathbb{R}^n)^*$ with $\mathcal{B}^{\psi} \to Q$ being G-equivariant. By Theorem 2.6, the inclusion order on these faces is the same as the inclusion order on the set of G-equivariant coherent face bundles of $\mathcal{B} \to Q$. $\square$

2.2. Equivariant fiber polytopes. We apply Theorem 2.7 to the situation of a G-equivariant projection of polytopes.
Let $P \xrightarrow{\pi} Q$ be a linear surjection of convex polytopes, with

$$P \subset \mathbb{R}^n, \quad \dim(P) = n$$
$$Q \subset \mathbb{R}^d, \quad \dim(Q) = n$$

Recall from [1] that this gives rise to a polytope bundle $\mathcal{B} \to Q$ via $x \mapsto \mathcal{B}_x := \pi^{-1}(x)$, and in this setting, the fiber polytope defined by

$$\Sigma(P \xrightarrow{\pi} Q) := \int_Q \mathcal{B}.$$

is a full $(n-d)$-dimensional polytope living in the fiber $\ker \pi$ of the map π over the centroid of Q.
For the equivariant set-up, we further assume that G is a finite group with linear G-actions on $\mathbb{R}^n$ and $\mathbb{R}^d$ which have G acting as symmetries of P, Q, and also that π is G-equivariant: $g(\pi(y)) = \pi(g(y))$ for all $y \in \mathbb{R}^n$. In particular, this implies that P, Q both have centroids at the origin. It is then easy to check that $\mathcal{B} \to Q$ defined as above is G-equivariant.

Definition 2.9. Define the *equivariant fiber polytope* by

$$\Sigma^G(P \xrightarrow{\pi} Q) := \frac{1}{\mathrm{vol}(Q)} \left(\int_Q \mathcal{B} \right)^G = \pi_G \Sigma(P \xrightarrow{\pi} Q).$$

Theorem 2.10. *The equivariant fiber polytope* $\Sigma^G(P \xrightarrow{\pi} Q)$ *is a full-dimensional polytope inside* $\dim\ \ker(\pi) \cap (\mathbb{R}^n)^G$*, and therefore has the same dimension as this space, namely*

$$\dim\,(\mathbb{R}^n)^G - \dim\,(\mathbb{R}^d)^G$$
$$(=\dim P^G - \dim Q^G).$$

Its face lattice is isomorphic to the poset of all G-equivariant π-coherent subdivisions of Q ordered by refinement.

Proof. If R is any full-dimensional polytope in $\mathbb{R}^r$ containing the origin in its interior, then its intersection $R \cap V$ with any linear subspace V has $\dim\,(R \cap V) = \dim V$. Since $\Sigma^G(P \xrightarrow{\pi} Q) = \Sigma(P \xrightarrow{\pi} Q) \cap (\mathbb{R}^n)^G$, this proves the first assertion.
To see that $\dim\ \ker(\pi) \cap (\mathbb{R}^n)^G = \dim\,(\mathbb{R}^n)^G - \dim\,(\mathbb{R}^d)^G$, note that

$$(\mathbb{R}^n)^G / (\ker(\pi) \cap (\mathbb{R}^n)^G) \cong (\mathbb{R}^d)^G$$

since G-equivariance of π implies that the surjection $\mathbb{R}^n \xrightarrow{\pi} \mathbb{R}^d$ restricts to a surjection $(\mathbb{R}^n)^G \xrightarrow{\pi} (\mathbb{R}^d)^G$.
The last assertion comes from the interpretation of coherent face bundles of $\mathcal{B} \to Q$ as π-coherent subdivisions, just as in [1, Thm 1.3]. $\square$

Recall that the fiber polytope $\Sigma(P \xrightarrow{\pi} Q)$ has an expression [1, Thm 1.5] as a finite Minkowski sum

$$\Sigma(P \xrightarrow{\pi} Q) = \sum_i \frac{\mathrm{vol}(\sigma_i)}{\mathrm{vol}(Q)} \pi^{-1}(x_i).$$

where the x_i are the centroids of the chambers (maximal cells) σ_i in the cell decomposition of Q induced by the projection of faces of P. From this one immediately deduces a similar expression for $\Sigma^G(P \xrightarrow{\pi} Q)$ by applying the averaging operator π_G:

$$\Sigma^G(P \xrightarrow{\pi} Q) = \sum_i \frac{\mathrm{vol}(\sigma_i)}{\mathrm{vol}(Q)} \pi_G(\pi^{-1}(x_i))$$

Similarly, one can obtain a (redundant) set of vertex coordinates for $\Sigma^G(P \xrightarrow{\pi} Q)$ by applying π_G to the vertex coordinates for $\Sigma(P \xrightarrow{\pi} Q)$ given in [1, Cor 2.6]. On the other hand, identifying an irredundant subset of these vertices is not so simple. One might expect that vertices of $\Sigma^G(P \xrightarrow{\pi} Q)$ correspond to *tight* G-invariant π-coherent subdivisions (see [1, §2] for the definition of tightness). However, Example 3.1 below shows that this is not the case. Rather, vertices of $\Sigma^G(P \xrightarrow{\pi} Q)$ correspond to G-invariant π-coherent subdivisions satisfying the weaker condition that they cannot be further refined while retaining both G-invariance and π-coherence.

2.3. Equivariant secondary polytopes. We specialize Theorem 2.10 to the situation where P is an $(n-1)$-dimensional simplex.
Let $\mathcal{A} := \{a_1, \ldots, a_n\}$ be the images of the vertices of P under the map π, so that $Q = \pi(P)$ is the convex hull of the point set $\mathcal{A}$, a d-dimensional polytope in $\mathbb{R}^d$. Note that not every point in $\mathcal{A}$ need be a vertex of Q, but we assume that the group G of symmetries acting linearly on $\mathbb{R}^d$ not only preserves Q, but also the set $\mathcal{A}$, i.e. $g\mathcal{A} = \mathcal{A}$ for all $g \in G$. There is a well-defined notion of a *polytopal subdivision* of $\mathcal{A}$, and when such subdivisions are *coherent (or regular)*; see [1, §1], [8, Chap. 7]. Say that such a subdivision is *G-invariant* if G permutes the polytopal cells occurring in the subdivision, taking into account the labelling of cells by elements of $\mathcal{A}$.
We may assume without loss of generality (e.g. by choosing P to be a *regular* $(n-1)$-simplex, that there is a linear G-action on $\mathbb{R}^{n-1}$ which permutes the vertices of P in the same way that G permutes $\mathcal{A}$. In this setting, define the *equivariant secondary polytope*

$$\Sigma^G(\mathcal{A}) := \Sigma^G(P \xrightarrow{\pi} Q).$$

Let $\mathcal{A}/G$ denote the set of G-orbits of points in $\mathcal{A}$.

Corollary 2.11. $\Sigma^G(\mathcal{A})$ *is an* $(|\mathcal{A}/G| - \dim(\mathbb{R}^d)^G - 1)$*-dimensional polytope, whose face lattice is isomorphic to the poset of all G-invariant coherent polytopal subdivisions of Q ordered by refinement.*

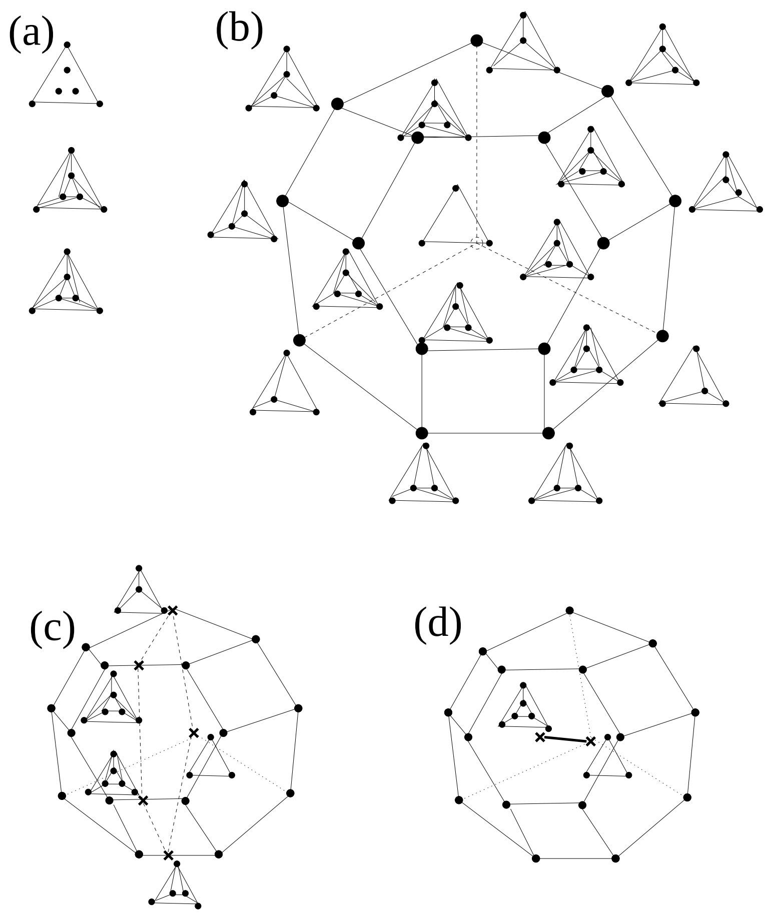


Figure 2. Equivariant secondary polytopes for subgroups G of the dihedral group D_6 acting on the point configuration $\mathcal{A}$ shown in (a).
(b) $G = 1$. (c) $G = C_2$. (d) $G = C_3$ or D_6.

Proof. Immediate from Theorem 2.10, once one notes that

$$\dim\,(\mathbb{R}^{n-1})^G = |\mathcal{A}/G| - 1$$

and that G-invariance and coherence of a polytopal subdivision correspond to G-equivariance and π-coherence with respect to the map $P \xrightarrow{\pi} Q$. □

3. Small, visualizable examples

3.1. The standard example. There is a classic example of a configuration $\mathcal{A}$ of 6 points in the plane $\mathbb{R}^2$, depicted at the top of Figure 2(a), which has up to two incoherent triangulations (shown below it), depending on the exact coordinates of its 6 points. If we choose coordinates so that $\mathcal{A}$ has the dihedral group D_6 as symmetries, then both of the triangulations shown in Figure 2(a) are incoherent, and hence do not correspond to vertices of the secondary polytope $\Sigma(\mathcal{A}) := \Sigma^1(\mathcal{A})$ depicted in (b) of the same figure[1] The only non-trivial proper subgroups of D_6 up to conjugacy are C_2 generated by a reflection symmetry, and C_3 generated by a three-fold rotation. Figures 2(c), (d) respectively depict as slices of $\Sigma(\mathcal{A})$ the equivariant secondary polytopes $\Sigma^{C_2}(\mathcal{A})$ (a pentagon) and $\Sigma^{C_3}(\mathcal{A})(= \Sigma^{D_6}(\mathcal{A})$ (a line segment) respectively. In both cases, the G-invariant coherent triangulations labelling their vertices are shown.

If instead we slightly perturb the coordinates of the three interior points of $\mathcal{A}$, so that the D_6-symmetry is destroyed, but still maintaining the C_3-symmetry, then something interesting happens in both $\Sigma(\mathcal{A})$ and $\Sigma^{C_3}(\mathcal{A})$. One of the two incoherent triangulations depicted in Figure 2(a) becomes coherent, and corresponds to a vertex which subdivides the "front" hexagon of $\Sigma(\mathcal{A})$ in (b) into 3 quadrangles. This new vertex also lies on the 1-dimensional slice $\Sigma^{C_3}(\mathcal{A})$, replacing one of its old endpoints.

Note that in Figure 2, some of the subdivisions labelling the vertices of $\Sigma^G(\mathcal{A})$ are not triangulations, that is, they are not *tight* π-coherent subdivisions of Q.

3.2. Triangulations of a regular hexagon. Consider the vertex set $\mathcal{A}$ of a regular hexagon. Its symmetry group is the dihedral group

$$D_{12} = \langle s, r : s^2 = r^3 = 1, srs = r^{-1} \rangle$$

where s is any reflection symmetry, and r is a rotation through $\frac{\pi}{3}$. In what follows, we will assume for the sake of definiteness that s is chosen in the conjugacy class of reflections whose reflection line passes through two vertices of the hexagon.

A list of representatives of the subgroups G of D_{12} up to conjugacy is given in the table below, along with the calculation of the dimension of the equivariant secondary polytope $\Sigma^G(\mathcal{A})$ in each case.

G	$\dim (\mathbb{R}^n)^G$ $(= \lvert\mathcal{A}/G\rvert - 1)$	$\dim (\mathbb{R}^d)^G$	$\dim \Sigma^G(\mathcal{A})$ $(= \dim (\mathbb{R}^n)^G - \dim (\mathbb{R}^d)^G)$
1	5	2	3
$\langle s \rangle$	3	1	2
$\langle r^3 \rangle$	2	0	2
$\langle sr \rangle$	2	1	1
$\langle r^2 \rangle$	1	0	1
$\langle r \rangle$	0	0	0
$\langle s, r^3 \rangle$	1	1	0
$\langle s, r \rangle (= D_{12})$	0	0	0

[1]For an on-line manipulable version of this secondary polytope, see Electronic Geometry Model No. 2000.09.033 at `http://www.eg-models.de`.

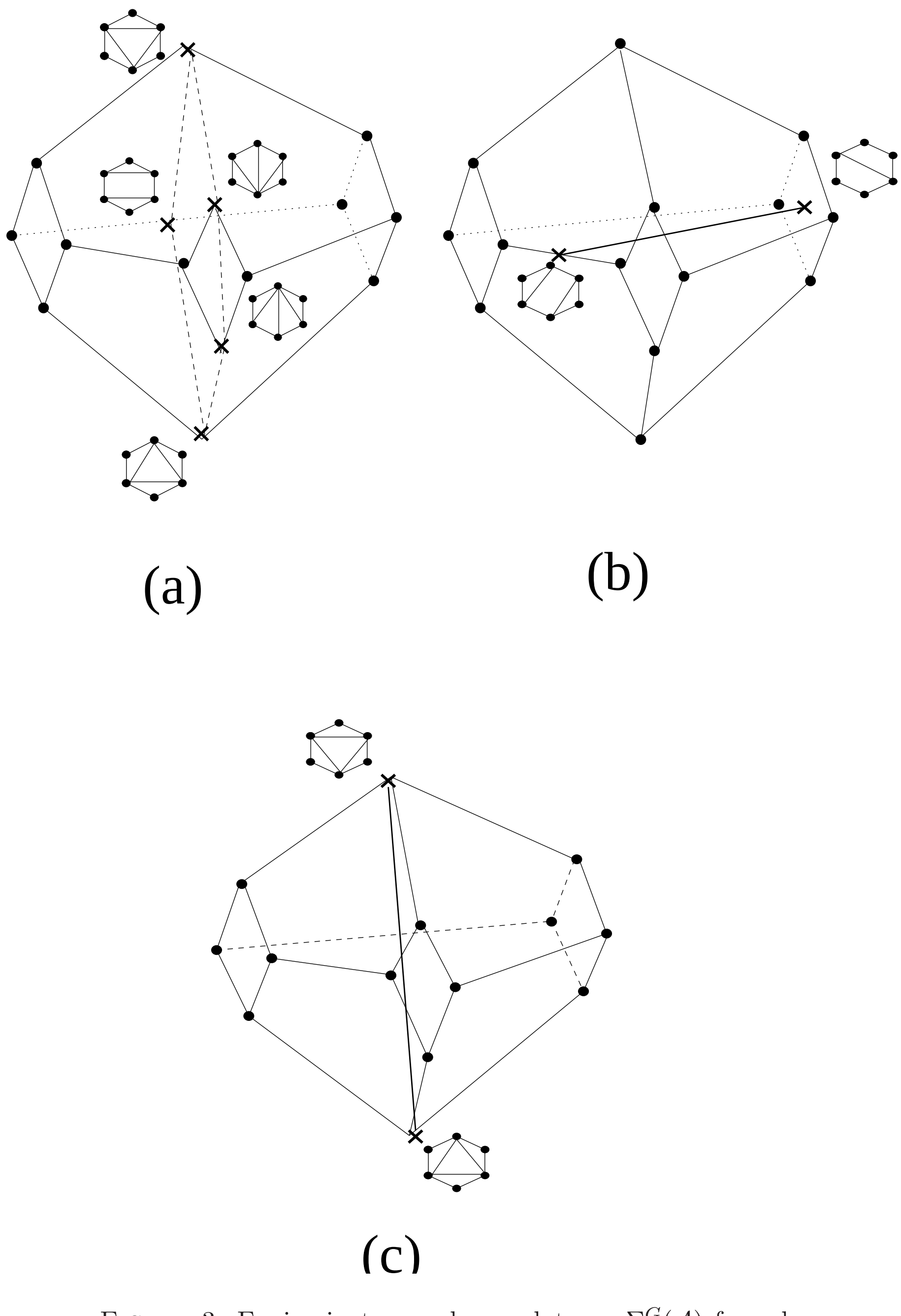


FIGURE 3. Equivariant secondary polytopes $\Sigma^G(\mathcal{A})$ for subgroups G of the dihedral group D_{12} acting on the set $\mathcal{A}$ of vertices of a regular hexagon.
(a) $G = \langle s \rangle$. (b) $G = \langle sr \rangle$. (c) $G = \langle r^2 \rangle$.

Depictions of these $\Sigma^G(\mathcal{A})$ as a slice of the 3-dimensional associahedron, in all the non-trivial cases where it has dimension at least 1, appear in Figures 1 and 3.

4. Application examples.

4.1. Cyclohedra and associahedra. The case $G = \langle r^3 \rangle$ in the previous example was discussed in the Introduction, and is a special case of a new construction for the cyclohedron.

The cyclohedron was introduced by Bott and Taubes [4], rediscovered independently and called the *type B associahedron* by Simion [16], and has been studied by several other authors [6, 7, 13]. The n-dimensonal cyclohedron can be thought of as the unique regular cell complex whose faces are indexed by centrally symmetric subdivisions of a centrally symmetric $2n$-gon. A proof that this cell complex is realized by a convex polytope appears in [13] and in [16, §2], and proceeds by a sequence of *shavings* (or *blow-ups*) of faces of an n-simplex, similar to the construction of the associahedron in [12].

Simion provided one of the original motivations for our work by asking whether a polytopal realization could be given along the lines of the fiber polytope construction. An answer is that it can be achieved as the equivariant secondary polytope $\Sigma^{C_2}(\mathcal{A})$, where $\mathcal{A}$ is the set of vertices of a centrally-symmetric $2n$-gon, on which C_2 acts antipodally. In other words, one has the following proposition.

Proposition 4.1. *The $(n-1)$-dimensional cyclohedron embeds naturally in the $(2n-3)$-dimensional associahedron, namely as the inclusion*

$$\Sigma^{C_2}(\mathcal{A}) \hookrightarrow \Sigma(\mathcal{A})$$

where $\mathcal{A}$ is the set of vertices of a centrally symmetric $2n$-gon. $\square$

It turns out that *all* equivariant secondary polytopes for a regular polygon are either associahedra or cyclohedra.

Proposition 4.2. *Let $\mathcal{A}$ be the vertex set of a regular n-gon, and G a non-trivial subgroup of its dihedral symmetry group D_{2n}.*
Then the combinatorial type of the equivariant secondary polytope $\Sigma^G(\mathcal{A})$ is either that of an associahedron or cyclohedron, depending upon whether G contains reflections or not (that is, whether G is dihedral or cyclic).

Proof. Assume G contains some reflections, so $G \cong D_{2m}$ for some m dividing n. Choose a fundamental domain for the action of G on the regular n-gon Q_n consisting of a sector between two adjacent reflection lines. Label the vertices in $\mathcal{A}$ which lie in this (closed) sector consecutively as $v_1, \ldots, v_r$ (here r is approximately $\frac{n}{2m}$, but its exact value depends upon which conjugacy classes of reflections in D_{2m} are represented among the reflections in G).

The assertion then follows from the claim that there is an isomorphism between the poset of G-invariant polygonal subdivisions of Q_n and the poset of all polygonal subdivisions of an $(r+1)$-gon Q_{r+1} labelled with vertices $w_1, w_2, w_3, \ldots, w_r, w$: the isomorphism sends a G-invariant subdivision σ of

Q_n to the unique subdivision τ of Q_{r+1} having an edge connecting w_i, w_j if and only if v_i, v_j are connected by an edge in in σ, and with an edge between w_i, w if and only if σ has v_i connected by an edge to some vertex of $\mathcal{A}$ outside the fundamental sector. The fact that this map is a bijection requires some straightforward geometric argumentation, which we omit. However, once one knows that it is a bijection, it is easy to see that both it and its inverse are order-preserving, since the refinement partial order on subdivisions of a polygon can be defined by the inclusion ordering of their edge sets.

Now assume G contains only rotations, so $G \cong C_m$ for some m dividing n, and let $k = \frac{n}{m}$. Label the vertices of Q_n consecutively in m groups of size k by

$$1_1, 2_1, \ldots, k_1, 1_2, 2_2, \ldots, k_2, \ldots 1_m, 2_m, \ldots, k_m.$$

Note that any C_m-invariant subdivision of Q_n is completely determined by the set of interior edges connecting vertices in the first two groups $1_1, 2_1, \ldots, k_1, 1_2, 2_2, \ldots, k_2$. This leads to an isomorphism between the poset of G-invariant polygonal subdivisions of Q_n and the poset of centrally symmetric polygonal subdivisions of a centrally symmetric $2k$-gon Q_{2k}: label the vertices of Q_{2k} using the same scheme, and send a G-invariant subdivision σ of Q_n to the unique subdivision τ of Q_{2k} whose interior diagonals have exactly the same endpoint labels as the interior diagonals of σ involving vertices in the first two groups in Q_{2n}. Again we omit the straightforward geometric details involved in checking that this is a bijection. $\square$

4.2. The type B permutohedron. Given any finite reflection group W, form the *zonotope* which is the Minkowski sum of any collection of line segments which contains exactly one line segment perpendicular to each of the reflecting hyperplanes for a reflection in W. Call this zonotope the *W-permutohedron*. It is known that the vertices of this zonotope are indexed by the elements of W, and its 1-skeleton is isomorphic to the (undirected) *Cayley graph* for W with respect to a natural set of Coxeter generators. Explicit descriptions of the facial structure of W-permutohedra when W is one of the classical reflection groups of type $A, B(=C)$, or D may be found in [14].

In the case where $W = A_{n-1}$ is the symmetric group on n letters, the A_{n-1}-permutohedron is usually known simply as the *permutohedron*. It can be constructed [1, Example 5.4] as the equivariant fiber polytope $\Sigma(P \xrightarrow{\pi} Q)$ where $P = [0,1]^n$ is the unit n-cube, $\pi : \mathbb{R}^n \to \mathbb{R}^1$ is the linear map sending $e_i \mapsto 1$ for each standard basis vector e_i, and Q is the line segment $[0, n] = \pi(P)$. This is a special case of a *monotone path polytope* [1, §5]: the vertices correspond to edge paths in the n-cube P which are monotone with respect to the functional π. In this case there is an obvious bijection between such paths and permutations of $\{1, 2 \ldots, n\}$; one simply reads off the parallelism class of the edges in the edge paths.

In the case where $W = B_n$, something similar works using the equivariant fiber polytope construction. Let P be the unit $2n$-cube in $\mathbb{R}^{2n}$ with standard basis vectors labelled $\{e_{+i}, e_{-i} : i = 1, 2 \ldots, n\}$. Consider the linear map

$\pi : \mathbb{R}^{2n} \to \mathbb{R}^1$ sending $e_{\pm i} \mapsto \pm 1$, and let $Q = \pi(P) = [-n, n]$. Let the generator of the group C_2 of order 2 act on $\mathbb{R}^{2n}$ by swapping e_{+i}, e_{-i}, and let it act on $\mathbb{R}^1$ by -1. Then C_2 acts as symmetries of both P, Q, and the map π is C_2-equivariant. It is then straightforward to check that the equivariant fiber polytope $\Sigma^{C_2}(P \xrightarrow{\pi} Q)$ is combinatorially isomorphic to the B_n-permutohedron. Its vertices correspond to the C_2-invariant monotone edge paths in the $2n$-cube P, which biject with signed permutations again by reading off the parallelism class of the edges in the edge path.

4.3. Maps from the type B permutohedron to the cyclohedron. There is a well-known set-map from the symmetric group S_{n+1} to triangulations of a convex $(n+2)$-gon, or to equivalent objects such as binary trees - see [18], [3, §9], [17, §1.3]. This map has several pleasant properties, including the fact that it extends to a map from faces of the A_n-permutohedron to faces of the n-dimensional associahedron. Simion [16, §4.2] asked whether there is an analogous map between the B_n-permutohedron and n-dimensional cyclohedron.

In fact, there is a whole family of such maps. To explain this, we first further explicate [3, Remark 9.14] on how to view the map in type A as a consequence of some theory of iterated fiber polytopes [2].

Given any tower $P \xrightarrow{\pi} Q \xrightarrow{\rho} R$ of linear surjections of polytopes, it turns out that π restricts to a surjection $\Sigma(P \xrightarrow{\rho\circ\pi} R) \xrightarrow{\pi} \Sigma(Q \xrightarrow{\rho} R)$ and so one can form the *iterated fiber polytope*

$$\Sigma(P \xrightarrow{\pi} Q \xrightarrow{\rho} R) := \Sigma\left(\Sigma(P \xrightarrow{\rho\circ\pi} R) \xrightarrow{\pi} \Sigma(Q \xrightarrow{\rho} R)\right).$$

Both $\Sigma(P \xrightarrow{\pi} Q \xrightarrow{\rho} R)$ and $\Sigma(P \xrightarrow{\pi} Q)$ live in the vector space $\ker(\pi)$, and [2, Theorem 2.1] says that the normal fan of $\Sigma(P \xrightarrow{\pi} Q \xrightarrow{\rho} R)$ refines that of $\Sigma(P \xrightarrow{\pi} Q)$ (or equivalently, the latter is a Minkowski summand of the former). This implies the existence of an order-preserving map from faces of $\Sigma(P \xrightarrow{\pi} Q \xrightarrow{\rho} R)$ to faces of $\Sigma(P \xrightarrow{\pi} Q)$, corresponding to the map on their normal cones which sends a cone in the finer fan to the unique cone containing it in the coarser fan.

We apply this to the tower of projections

$$\Delta^{n+1} \xrightarrow{\pi} Q_{n+2} \xrightarrow{\rho} I$$

in which Δ^{n+1} is an $(n+1)$-dimensional simplex whose vertices map canonically to the vertices of a convex $(n+2)$-gon Q_{n+2}, which then projects onto a 1-dimensional interval I. In [2, §4] it is shown that $\Sigma(\Delta^{n+1} \xrightarrow{\rho\circ\pi} I)$ and $\Sigma(\Delta^{n+1} \xrightarrow{\pi} Q_{n+2} \xrightarrow{\rho} I)$, are combinatorially isomorphic (but not affinely equivalent) to the n-cube and to the A_{n-1}-permutohedron, respectively. Since $\Sigma(\Delta^{n+1} \xrightarrow{\pi} Q_{n+2})$ is the $(n-1)$-dimensional associahedron, in this case the above general theory gives a map from the faces of the permutohedron to the faces of the associahedron, which can be checked to coincide with the usual one.

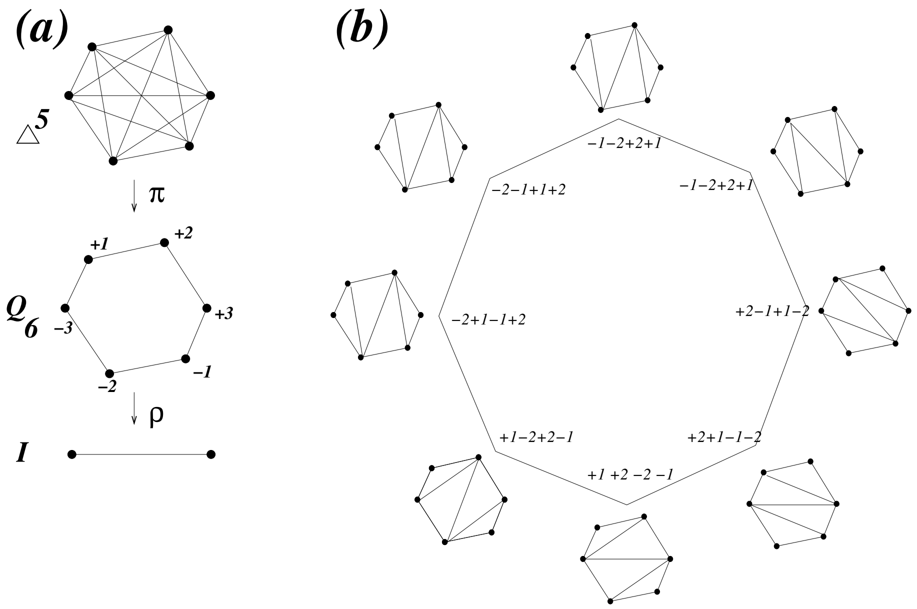


FIGURE 4. (a) The tower of projections (4.1) for $n = 2$ (the 5-simplex Δ^5 is shown only in a 2-dimensional projection). (b) One of the $4(= 2^2)$ maps from vertices of the B_2-permutohedron to vertices of the 2-dimensional cyclohedron. The B_2-permutohedron is shown with vertices labelled both by a signed permutation and by the centrally-symmetric hexagon triangulation which is their image under the map.

Now suppose we instead apply this theory to the following tower of C_2-equivariant projections

$$\Delta^{2n+1} \xrightarrow{\pi} Q_{2n+2} \xrightarrow{\rho} I \tag{4.1}$$

in which Q_{2n+2} is a centrally-symmetric $(2n+2)$-gon, Δ^{2n+1} is a $(2n+1)$-dimensional (regular) simplex, and I is an interval. We assume that the vertices of Q_{2n+2} are labelled in cyclic order as

$$+1, +2, \ldots, +(n+1), -1, -2, \ldots, -(n+1)$$

and that the C_2-actions on $\Delta^{2n+1}, Q_{2n+2}, I$ are chosen so that the projections are equivariant, i.e. C_2 swaps the two vertices of I, and exchanges the pairs of vertices of Δ^{2n+1} which map to the vertices labelled $+i, -i$ of Q_{2n+2}. We further assume that the map ρ is generic in the sense that it takes on distinct values on the different vertices of Q_{2n+2}, and hence gives a linear ordering of these vertices, which we will assume orders the vertices labelled $-(n+1), +(n+1)$ first and last, respectively. The map we eventually define will depend on this ordering.

Note that $\Sigma(\Delta^{2n+1} \stackrel{\rho\circ\pi}{\to} I)$ will still be a combinatorial $2n$-cube, and by Corollary 2.3, it will carry a C_2-action that makes the map π onto the interval $\Sigma(Q^{2n+2} \stackrel{\rho}{\to} I)$ equivariant. We still have from [2, §4], that the iterated fiber polytope $\Sigma(\Delta^{2n+1} \stackrel{\pi}{\to} Q_{2n+2} \stackrel{\rho}{\to} I)$ is combinatorially an A_{2n-1}-permutohedron, and by Corollary 2.3 will carry a C_2-action. With some work, for which we omit the details, one can check that the C_2-action on faces corresponds to the same C_2-action as in Example 4.2. Hence C_2-invariant faces under this action are identified with the faces of the B_n-permutohedron, so that the C_2-invariant subpolytope $\Sigma^{C_2}(\Delta^{2n+1} \stackrel{\pi}{\to} Q_{2n+2} \stackrel{\rho}{\to} I)$ is combinatorially isomorphic to the B_n-permutohedron.

On the other hand, $\Sigma^{C_2}(\Delta^{2n+1} \stackrel{\pi}{\to} Q_{2n+2})$ is the n-dimensional cyclohedron, by Example 4.1. Since the normal fan of $\Sigma^G(P \stackrel{\pi}{\to} Q \stackrel{\rho}{\to} R)$ refines that of $\Sigma^G(P \stackrel{\pi}{\to} Q)$ (by restricting this refinement of fans from the non-equivariant setting to the invariant subspace $(\mathbb{R}^n)^G$), we obtain the existence of an order-preserving map between their face lattices as desired.

To be explicit about this map and its dependence on the ordering of the vertices of Q_{2n+2} by ρ, it suffices to describe its effect on vertices. A vertex of the n-dimensional permutohedron is indexed by a signed permutation, that is a sequence $w = w_1 w_2 \dots w_n$ where $w_i \in \{\pm 1, \dots, \pm n\}$ containing exactly one value from each pair $\pm i$. To obtain a centrally-symmetric triangulation of Q_{2n+2} from w, we associate to w a sequence of $2n+1$ polygonal paths $\gamma_0, \gamma_1, \dots, \gamma_{2n}$ visiting only vertices of Q_{2n+2}, and let the triangulation be the one whose edges are the union of these paths. Each γ_i is a section of the map $Q_{2n+2} \stackrel{\rho}{\to} I$, and hence completely specified by the set of vertices of Q_{2n+2} it visits (although this implicitly requires knowledge of the fixed ordering of vertices of Q_{2n+2} by ρ). Set γ_0 to be the path visiting vertices $-(n+1), +1, +2, \dots, +(n+1)$, that is, γ_0 is half of the boundary of the polygon Q_{2n+2}. Then inductively define γ_i to be the unique path obtained from γ_{i-1} by reading the i^{th} value in the sequence $\hat{w} := w_1 w_2 \dots w_n - w_n \dots - w_2 - w_1$ and either removing this value from the list of visited vertices when it is positive, or adding it when it is negative. Note that the palindromic nature of $\hat{w}$ insures that the associated triangulation is centrally symmetric. Some examples of the map are shown in Figure 4.

How many maps have we defined in this way? The map depends only the ordering of the vertices of Q_{2n+2} by ρ. Any such ordering starts and ends with $-(n+1), +(n+1)$, and in between is a shuffle of the usual integer order on the positive vertices $+1, \dots, +n$ with the usual order on the negative vertices. Since Q_{2n+2} is centrally-symmetric and ρ is linear, the order is determined by knowing its first half. Hence it can be parametrized by the set $S \subset \{1, 2, \dots, n\}$ giving the positions in the first half of the order where the negative vertices occur. This means there are 2^n such maps.

5. Remarks/Questions

Remark 5.1. One might expect a relation between

$$\Sigma(P^G \xrightarrow{\pi} Q^G)$$
$$\Sigma^G(P \xrightarrow{\pi} Q)$$

since both are full-dimensional polytopes embedded in the subspace $(\mathbb{R}^n)^G \cap \ker \pi$. The case $G = C_2$ in Example 3.1 already shows that they are not isomorphic: here $\Sigma(P^G \xrightarrow{\pi} Q^G) = P^G$ is a triangle, while $\Sigma^G(P \xrightarrow{\pi} Q)$ is a pentagon.

Neither is there an inclusion in either direction, as illustrated by the following example. Let $P \xrightarrow{\pi} Q$ be the canonical projection of a regular 3-simplex onto a square Q, with an equivariant C_2-action that reflects the square across one of its diagonals. Label the vertices of P by v_1, v_2, v_3, v_4 in such a way that the C_2-action swaps $\pi(v_1), \pi(v_2)$ and fixes $\pi(v_3), \pi(v_4)$. Then both polytopes in question are 1-dimensional intervals, and one can calculate directly that

$$\Sigma(P^G \xrightarrow{\pi} Q^G) = \left[\frac{1}{3}x + \frac{2}{3}y,\ \frac{2}{3}x + \frac{1}{3}y\right]$$
$$\Sigma^G(P \xrightarrow{\pi} Q) = \left[0 \cdot x + \ 1 \cdot y,\ \frac{1}{2}x + \frac{1}{2}y\right] \left(= \Sigma(P \xrightarrow{\pi} Q)\right)$$

where $x = \frac{v_1+v_2}{2}, y = \frac{v_3+v_4}{2}$.

The distinction between the two relates to weighted averages[2] We know from Corollary 2.3 that $\Sigma^G(P \xrightarrow{\pi} Q)$ is the set of all average values of G-equivariant sections γ of $P \xrightarrow{\pi} Q$, while $\Sigma(P^G \xrightarrow{\pi} Q^G)$ consists of average values of sections $\bar{\gamma}$ of $P^G \xrightarrow{\pi} Q^G$. Using Fubini's Theorem, one can show that the average value over Q of a G-equivariant section γ is the same as the *weighted* average value of an appropriately defined section $\bar{\gamma}$ of $P^G \xrightarrow{\pi} Q^G$, obtained by integrating γ over fibers $\pi_G^{-1}(x)$, in which the weight at a point x in Q^G is equal to the volume of the fiber $\pi_G^{-1}(x)$. If these fiber volumes are not constant, the weighted average and the average need not coincide.

One might still ask whether there is a relation between their associated normal fans living in $\ker(\pi)^{*G}$, e.g. one refining the other so that the one polytope is a Minkowski summand of the other. But as far as we know there is no a priori reason for such a relation. Using the notation of [2], one has that

$$\begin{aligned}\mathcal{N}\Sigma^G(P \xrightarrow{\pi} Q) &:= \mathcal{N}\pi_G\Sigma(P \xrightarrow{\pi} Q)\\ &= \mathcal{N}\Sigma(P \xrightarrow{\pi} Q) \cap \operatorname{im}(\pi_G^*)\\ &= (\operatorname{proj}_{\ker(\pi)^*}\mathcal{N}P) \cap \ker(\pi)^{*G}\end{aligned}$$

[2]Thanks to John Baxter for an enlightening conversation in this regard.

whereas

$$\begin{aligned}\mathcal{N}\Sigma(P^G \xrightarrow{\pi} Q^G) &:= \mathrm{proj}_{\ker(\pi)^*}\mathcal{N}P^G \\ &= \mathrm{proj}_{\ker(\pi)^*}\mathcal{N}\pi_G P \\ &= \mathrm{proj}_{\ker(\pi)^*}(\mathcal{N}P \cap \mathbb{R}^{n*G}).\end{aligned}$$

Remark 5.2. The fiber polytope $\Sigma(P \xrightarrow{\pi} Q)$ has a toric interpretation given by Kapranov, Sturmfels and Zelevinsky [11]. The associated toric variety $X_{\Sigma(P \xrightarrow{\pi} Q)}$ is the *Chow quotient* X_P/T of the toric variety X_P by the subtorus T defined by the kernel of π.

One might then expect in the G-equivariant setting that there is a more general interpretation for $X_{\Sigma^G(P \xrightarrow{\pi} Q)}$, perhaps relating it to the G-invariant subvariety $(X_P/T)^G$ for an induced G-action on X_P/T. However, even in the very special case where $Q = \{0\}$, so that

$$\begin{aligned}\Sigma(P \xrightarrow{\pi} Q) &= P \\ \Sigma^G(P \xrightarrow{\pi} Q) &= P^G \\ X_P/T &= X_P\end{aligned}$$

the general relation between the G-invariant subvariety X_P^G and the toric variety X_{P^G} seems not to be trivial- see [9, Theorem 2] for a special case. We leave the problem of interpreting $\Sigma^G(P \xrightarrow{\pi} Q)$ torically to the real experts.

Remark 5.3. Fomin and Zelevinsky [7] recently introduced a family of simplicial spheres associated to each finite (crystallographic) root system, whose facial structure coincides with the associahedron in type A and with the cyclohedron in type B. They and Chapoton subsequently proved [5] that these spheres can be realized as the boundaries of simplicial convex polytopes.

One might wonder whether they could be realized as the boundaries of equivariant fiber polytopes, as is true in the case of types A and B. However we do not see how to do this for their spheres in the case of type D, whose 1-skeleton is described in [7, Prop. 3.16].

Remark 5.4. In type A, the 1-skeleton of the permutohedron and associahedron each have an acyclic orientation making them the Hasse diagrams for the weak Bruhat order on S_n and the Tamari lattice, respectively. Both of these partial orders are self-dual lattices, and the map between them mentioned in Section 4.3 enjoys some very pleasant properties [3, §9].

There is a similar acyclic orientation for the 1-skeleton of the type B permutohedron, and Simion [16, §4.1] asked whether there are corresponding well-behaved acyclic orientations and partial orders for the 1-skeleton of the cyclohedron. She proposed two such orders, one of which is self-dual and has some nice properties explored in [15], but neither of which is a lattice.

Because the maps introduced in Section 4.3 are surjective, they can be used to transfer the acyclic orientation from the 1-skeleton of the type B permutohedron to an orientation of the 1-skeleton of the cyclohedron, which may or may not be acyclic. Do any of these induced orientations end up being acyclic, and are their partial orders well-behaved in any sense?

Remark 5.5. It was pointed out in the proof of Proposition 2.3 that whenever a finite group G acts linearly on a convex polytope P, there is induced a linear surjection $P \overset{\pi_G}{\to} P^G$. Does the "coinvariant polytope" $\Sigma(P \overset{\pi_G}{\to} P^G)$ enjoy any nice properties or interpretations?

Acknowledgements

The author thanks Mike Matsko for helpful comments, and Nirit Sandman for making available her work [15].

References

[1] L.J. Billera and B. Sturmfels, Fiber polytopes, *Ann. of Math.* 135 (1992), 527–549.

[2] L.J. Billera and B. Sturmfels, Iterated fiber polytopes, *Mathematika*, 41 (1994), 348–363.

[3] A. Björner and M. Wachs, Shellable nonpure complexes and posets, II, *Trans. Amer. Math. Soc.* 349 (1997), 3945–3975.

[4] R. Bott, C. Taubes, On the self-linking of knots, Topology and physics *J. Math. Phys.* 35 (1994), 5247–5287.

[5] F. Chapoton, S.V. Fomin and A.V. Zelevinsky, Polytopal realizations of generalized associahedra, preprint 2002, Los Alamos archive `math.CO/0202004`.

[6] S. Devadoss, A space of cyclohedra, preprint 2001.

[7] S.V. Fomin and A.V. Zelevinsky, Y-systems and generalized associahedra, preprint 2002, Los Alamos archive `hep-th/0111053`.

[8] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky, Discriminants, resultants, and multidimensional determinants, Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1994.

[9] H.A. Jorge, Smith-type inequalities for a polytope with a solvable group of symmetries, *Adv. Math.* 152 (2000), 134–158.

[10] M.M. Kapranov and M. Saito, Hidden Stasheff polytopes in algebraic K-theory and in the space of Morse functions, Higher homotopy structures in topology and mathematical physics (Poughkeepsie, NY, 1996), 191–225, *Contemp. Math.* 227, Amer. Math. Soc., Providence, RI, 1999.

[11] M.M. Kapranov, B. Sturmfels, and A.V. Zelevinsky, Quotients of toric varieties, *Math. Ann.* 290 (1991), 643–655.

[12] C.W. Lee, The associahedron and triangulations of the n-gon. *European J. Combin.* 10 (1989), 551–560.

[13] M. Markl, Martin, Simplex, associahedron, and cyclohedron, in "Higher homotopy structures in topology and mathematical physics" (Poughkeepsie, NY, 1996) *Contemp. Math.* 227, 235-265 Amer. Math. Soc., Providence, RI, 1999.

[14] V. Reiner and G.M. Ziegler, Coxeter-associahedra, *Mathematika* 41 (1994), 364–393.
[15] N. Sandman, Honors Senior Thesis, George Washington University, 2000.
[16] R. Simion, A type B associahedron, *Adv. Applied Math.*, to appear.
[17] R.P. Stanley, Enumerative combinatorics, Vol. 1, *Cambridge Studies in Advanced Mathematics* 49. Cambridge University Press, Cambridge, 1997
[18] A. Tonks, Relating the associahedron and the permutohedron. "Operads: Proceedings of Renaissance Conferences" (Hartford, CT/Luminy, 1995), 33–36, *Contemp. Math.* 202, Amer. Math. Soc., Providence, RI, 1997.

Victor Reiner
University of Minnesota
Minneapolis, MN 55455, USA
reiner@math.umn.edu

On the Uniqueness Problem of Bivariant Chern Classes

Shoji Yokura[1]

Received: February 14, 2002
Revised: May 31, 2002

Communicated by Joachim Cuntz

Abstract. In this paper we show that the bivariant Chern class $\gamma : \mathbb{F} \to \mathbb{H}$ for morphisms from possibly singular varieties to nonsingular varieties are uniquely determined, which therefore implies that the Brasselet bivariant Chern class is unique for cellular morphisms with nonsingular target varieties. Similarly we can see that the Grothendieck transformation $\tau : \mathbb{K}_{\mathrm{alg}} \to \mathbb{H}_{\mathbb{Q}}$ constructed by Fulton and MacPherson is also unique for morphisms with nonsingular target varieties.

2000 Mathematics Subject Classification: 14C17, 14F99, 55N35
Keywords and Phrases: Bivariant theory; Bivariant Chern class; Chern-Schwartz-MacPherson class; Constructible function

§1 Introduction

In [FM, Part I] W. Fulton and R. MacPherson developed the so-called *Bivariant Theories*, which are simultaneous generalizations of covariant functors and contravariant functors. They are equipped with three operations of *product, pushforward, pullback*, and they are supposed to satisfy seven kinds of axioms. A transformation from one bivariant theory to another bivariant theory, preserving these three operations, is called a *Grothendieck transformation*, which is a generalization of ordinary natural transformations.

The *Chern-Schwartz-MacPherson class* is the unique natural transformation $c_* : F \to H_*$ from the covariant functor F of constructible functions to the integral homology covariant functor H_*, satisfying the normalization condition that the value $c_*(\mathbb{1}_X)$ of the characteristic function $\mathbb{1}_X$ of a nonsingular variety X is equal to the Poincaré dual of the total Chern class $c(TX)$ of the tangent

[1]Partially supported by Grant-in-Aid for Scientific Research (C) (No.12640081), the Japanese Ministry of Education, Science, Sports and Culture

bundle TX of X. The existence of this transformation was conjectured by Deligne and Grothendieck, and was proved by MacPherson (see [M] and also [BS], [Sc]).

In [FM, Part I, §10.4] Fulton and MacPherson conjectured (or posed as a question) the existence of a bivariant Chern class, i.e., a Grothendieck transformation $\gamma : \mathbb{F} \to \mathbb{H}$ from the bivariant theory $\mathbb{F}$ of constructible functions to the bivariant homology theory $\mathbb{H}$, satisfying the normalization condition that for a morphism from a nonsingular variety X to a point the value $\gamma(1\!\!1_X)$ of the characteristic function $1\!\!1_X$ of X is equal to the Poincaré dual of the total Chern class of X. The bivariant Chern class specializes to the original Chern-Schwartz-MacPherson class, i.e., when restricted to morphisms to a point it becomes the Chern-Schwartz-MacPherson class. As applications of the bivariant Chern class, for example, one obtains the Verdier-Riemann-Roch for Chern class and the Verdier's specialization of Chern classes [V].

In [B] J.-P. Brasselet has solved the conjecture affirmatively in the category of complex analytic varieties and cellular analytic maps. Any analytic map is "conjecturally" cellular and indeed no example of a non-cellular analytic map has been found so far. In this sense the condition of "cellularness" could be dropped. For example, it follows from a result of Teissier [T] that an analytic map to a smooth curve is cellular (see [Z2, 2.2.5 Lemme]). In [S] C. Sabbah gave another construction of bivariant Chern classes, using the notions of *bivariant cycle, relative local Euler obstruction, morphisme sans éclatement en codimension 0* (see [S] or [Z1, Z2] for more details). And in [Z1] (and [Z2]) J. Zhou showed that for a morphism from a variety to a smooth curve these two bivariant Chern classes due to Brasselet and Sabbah are identical. However, the uniqueness of bivariant Chern classes still remains as an open problem.

In [FM, Part II] Fulton and MacPherson constructed a Grothendieck transformation $\tau : \mathbb{K}_{\mathrm{alg}} \to \mathbb{H}_{\mathbb{Q}}$, which is a bivariant-theoretic version of Baum-Fulton-MacPherson's Riemann-Roch $\tau^{\mathrm{BFM}} : \mathbf{K}_0 \to H_{*\mathbb{Q}}$ constructed in [BFM]. The uniqueness problem of this Grothendieck transformation remains open.

As remarked in [FM, Part I, §10.9: Uniqueness Questions], there are few uniqueness theorems available concerning Grothendieck transformations.

In this paper we show that the bivariant Chern class for morphisms with nonsingular target varieties is unique if it exists. Therefore it follows that the Brasselet bivariant Chern class is unique for cellular morphisms with nonsingular target varieties, thus it gives another proof of Zhou's result mentioned above. Our method also implies that the above Grothendieck transformation $\tau : \mathbb{K}_{\mathrm{alg}} \to \mathbb{H}_{\mathbb{Q}}$ constructed by Fulton and MacPherson is unique for morphisms with nonsingular target varieties.

The author would like to thank Jean-Paul Brasselet and the referee for useful comments and suggestions.

§2 Bivariant constructible functions and bivariant homology theory

For a general reference for the bivariant theory, see Fulton-MacPherson's book [FM]. In this section, we recall some basic ingredients, needed in this paper, of the bivariant theory of constructible functions and bivariant homology theory. For a morphism $f : X \to Y$ the bivariant theory $\mathbb{F}(X \xrightarrow{f} Y)$ of constructible functions consists of all the constructible functions on X which satisfy the *local Euler condition with respect to* f, i.e., the condition that for any point $x \in X$ and for any local embedding $(X, x) \to (\mathbf{C}^N, 0)$ the following equality holds

$$\alpha(x) = \chi\big(B_\epsilon \cap f^{-1}(z); \alpha\big),$$

where B_ϵ is a sufficiently small open ball of the origin 0 with radius ϵ and z is any point close to $f(x)$ (see [B], [FM], [S], [Z1]). The three operations on $\mathbb{F}$ are defined as follows:

(i): the product operation

$$\bullet : \mathbb{F}(X \xrightarrow{f} Y) \otimes \mathbb{F}(Y \xrightarrow{g} Z) \to \mathbb{F}(X \xrightarrow{gf} Z)$$

is defined by:

$$\alpha \bullet \beta := \alpha \cdot f^*\beta.$$

(ii): the pushforward operation

$$f_* : \mathbb{F}(X \xrightarrow{gf} Z) \to \mathbb{F}(Y \xrightarrow{g} Z)$$

is the pushforward

$$(f_*\alpha)(y) := \chi(f^{-1}(y); \alpha) = \int_{f^{-1}(y)} c_*(\alpha|_{f^{-1}(y)}).$$

(iii): For a fiber square

$$\begin{array}{ccc} X' & \xrightarrow{g'} & X \\ {\scriptstyle f'}\downarrow & & \downarrow{\scriptstyle f} \\ Y' & \xrightarrow{g} & Y, \end{array}$$

the pullback operation

$$g^* : \mathbb{F}(X \xrightarrow{f} Y) \to \mathbb{F}(X' \xrightarrow{f'} Y')$$

is the functional pullback

$$({g'}^*\alpha)(x') := \alpha(g'(x')).$$

These operations satisfy the seven axioms listed in [FM, Part I, §2.2] and it is also known that these three operations are well-defined (e.g., see [BY], [FM], [S], [Z1]). Note that $\mathbb{F}(X \xrightarrow{\mathrm{id}_X} X)$ consists of all locally constant functions and $\mathbb{F}(X \to pt) = F(X)$.
Let $\mathbb{H}$ be the Fulton-MacPherson bivariant homology theory, constructed from the cohomology theory. For a morphism $f : X \to Y$, choose a morphism $\phi : X \to M$ to a smooth manifold M of real dimension n such that $\Phi := (f, \phi) : X \to Y \times M$ is a closed embedding. Of course, the morphism $\phi : X \to M$ can be already an embedding. Then the i-th bivariant homology group $\mathbb{H}^i(X \xrightarrow{f} Y)$ is defined by

$$\mathbb{H}^i(X \xrightarrow{f} Y) := H^{i+n}(Y \times M, (Y \times M) \setminus X_\phi),$$

where X_ϕ is defined to be the image of the morphism $\Phi = (f, \phi)$. The definition is independent of the choice of ϕ, i.e., for any other morphism $\phi' : X \to M'$ to a smooth manifold M' of real dimension n' there is an isomorphism

$$H^{i+n}(Y \times M, (Y \times M) \setminus X_\phi) \cong H^{i+n'}(Y \times M', (Y \times M') \setminus X_{\phi'}).$$

See [FM, §3.1] for more details of $\mathbb{H}$.
A bivariant Chern class is a Grothendieck transformation from the bivariant theory $\mathbb{F}$ of constructible functions to the bivariant homology theory $\mathbb{H}$

$$\gamma : \mathbb{F} \to \mathbb{H}$$

satisfying the normalization condition that for a nonsingular variety X and for the map $\pi : X \to pt$ to a point pt

$$\gamma(\mathbb{1}_\pi) = c(TX) \cap [X]$$

where $\mathbb{1}_\pi = \mathbb{1}_X \in F(X) = \mathbb{F}(X \xrightarrow{\pi} pt)$. Here the Grothendieck transformation $\gamma : \mathbb{F} \to \mathbb{H}$ preserves the three operations of product, pushforwad and pullback, i.e.,
(i) $\gamma(\alpha \bullet \beta) = \gamma(\alpha) \bullet \gamma(\beta)$,
(ii) $\gamma(f_*\alpha) = f_*\gamma(\alpha)$ and
(iii) $\gamma(g^*\alpha) = g^*\gamma(\alpha)$.

Theorem (2.1). *(Brasselet's Theorem [B]) For the category of analytic varieties with cellular morphisms there exists a bivariant Chern class $\gamma : \mathbb{F} \to \mathbb{H}$.*

§3 Uniqueness of the bivariant Chern class

First we take a bit closer look at the definition of the bivariant homology theory. As seen above, for *any* morphism $\phi : X \to M$ such that $(f, \phi) : X \to Y \times M$ is a closed embedding (or simply, for any closed embedding $\phi : X \to M$), we have the isomorphism

$$\mathbb{H}(X \xrightarrow{f} Y) \cong H^*(Y \times M, (Y \times M) \setminus X).$$

This isomorphism is thought to be a "realization isomorphism with respect to the embedding $X \to Y \times M$" of the group $\mathbb{H}(X \xrightarrow{f} Y)$. We denote this isomorphism by $\Re_{X \hookrightarrow Y \times M}$, emphasizing the embedding $\Phi : X \to Y \times M$. In particular, for a morphism $f : X \to pt$ to a point pt, the bivariant homology group $\mathbb{H}(X \xrightarrow{f} pt)$ is considered to be the homology group $H_*(X)$, since for any embedding of X into any manifold N we have the Alexander duality isomorphism

$$H^*(N, N \setminus X) \cong H_*(X),$$

which shall be denoted by $\mathcal{A}_{X \hookrightarrow N}$, again indicating the embedding $X \hookrightarrow N$. Note that the Alexander isomorphism is given by taking the cap product with the fundamental class, i.e.,$\mathcal{A}_{X \hookrightarrow N}(a) = a \cap [N]$. Therefore $\mathbb{H}(X \xrightarrow{f} pt) = H_*(X)$ and $\Re_{X \hookrightarrow N} = \big(\mathcal{A}_{X \hookrightarrow N}\big)^{-1}$. In particular, if X is nonsingular, the Alexander duality isomorphism is the Poincaré duality isomorphism via Thom isomorphism, denoted by $\mathcal{P}_X$:

$$\mathcal{A}_{X \hookrightarrow X} = \mathcal{P}_X : H^*(X) \cong H_*(X).$$

With these notation, it follows from the definition of the bivariant product in $\mathbb{H}$ [FM, Part I, §3.1.7] that the bivariant product

$$\bullet : \mathbb{H}(X \xrightarrow{f} Y) \otimes \mathbb{H}(Y \xrightarrow{g} Z) \to \mathbb{H}(X \xrightarrow{gf} Z)$$

is described as follows: consider the following commutative diagram where the rows are closed embedding and the verticals are the projections with M and N being manifolds:

$$\begin{tikzcd} X \arrow[r] & Y \times M \arrow[r] \arrow[d] & Z \times N \times M \arrow[d, "p"] \\ & Y \arrow[r] & Z \times N \arrow[d] \\ & & Z \end{tikzcd}$$

Then for $\alpha \in \mathbb{H}(X \xrightarrow{f} Y)$ and $\beta \in \mathbb{H}(Y \xrightarrow{g} Z)$

$$\alpha \bullet \beta := \Big(\Re_{X \hookrightarrow Z \times N \times M}\Big)^{-1} \Bigg(\Re_{X \hookrightarrow Y \times M}(\alpha) \cdot p^* \Re_{Y \hookrightarrow Z \times N}(\beta)\Bigg), \tag{3.1}$$

where the center dot $\cdot$ is the product defined by [FM, §3.1.7 (1), p.36]. The well-definedness of the bivariant homology product given in Fulton-MacPherson's book [FM] means that the above description (3.1) is independent of the choices of M and N, i.e., the realization isomorphisms. This viewpoint becomes a crucial one in our proof.

Remark (3.2). Suppose that $\gamma : \mathbb{F} \to \mathbb{H}$ is a bivariant Chern class. For a morphism $f : X \to Y$, we denote the homomorphim $\gamma : \mathbb{F}(X \xrightarrow{f} Y) \to \mathbb{H}(X \xrightarrow{f} Y)$ by $\gamma_{X\to Y}$. Then, for any variety X we see that the homomorphism $\gamma_{X\to pt} : F(X) = \mathbb{F}(X \to pt) \to \mathbb{H}(X \to pt) = H_*(X)$ is nothing but the Chern-Schwartz-MacPherson class homomorphism $c_* : F(X) \to H_*(X)$, because $\gamma_{X\to pt}$ is a natural transformation satisfying the normalization condtion and thus it has to be the Chern-Schwartz-MacPherson class $c_* : F(X) \to H_*(X)$ since it is unique.

Let $\gamma : \mathbb{F} \to \mathbb{H}$ be a bivariant Chern class and let $\alpha \in \mathbb{F}(X \xrightarrow{f} Y)$. Then we have

$$\gamma_{X\to pt}(\alpha) = \gamma_{X\to Y}(\alpha) \bullet \gamma_{Y\to pt}(1\!\!1_Y),$$

Therefore it follows from Remark (3.2) that we have

$$c_*(\alpha) = \gamma_{X\to Y}(\alpha) \bullet c_*(1\!\!1_Y). \tag{3.3}$$

Furthermore, for any constructible function $\beta \in F(Y)$, we have

$$c_*(\alpha \cdot f^*\beta) = \gamma_{X\to Y}(\alpha) \bullet c_*(\beta).$$

The uniqueness of $\gamma : \mathbb{F} \to \mathbb{H}$, therefore, follows if we can show that $\omega \in \mathbb{H}(X \xrightarrow{f} Y)$ and $\omega \bullet c_*(\beta) = 0$ for any $\beta \in F(Y)$ automatically implies that $\omega = 0$.

Heuristically or very loosely speaking, the bivariant Chern class $\gamma_{X\to Y}(\alpha)$ of the bivariant constructible function $\alpha \in \mathbb{F}(X \xrightarrow{f} Y)$ could be or should be "described" as a "quotient"

$$\gamma_{X\to Y}(\alpha) := \frac{c_*(\alpha)}{c_*(1\!\!1_Y)}$$

in a reasonable way. Otherwise it would be an interesting problem to see if there is a reasonable bivariant homology theory so that this "quotient" is well-defined. We hope to come back to this problem in a different paper.

However, in the case of morphisms whose target varieties are nonsingular, the above argument gives us the uniqueness of the bivariant Chern class and furthermore we can describe the above "quotient" $\frac{c_*(\alpha)}{c_*(\mathbf{1}_Y)}$ explicitly.

Theorem (3.4). *Let $\gamma : \mathbb{F} \to \mathbb{H}$ be a bivariant Chern class. Then it is unique, when restricted to morphisms whose target varieties are nonsingular.*

Explicitly, for a morphism $f : X \to Y$ with Y being nonsingular and for any bivariant constructible function $\alpha \in \mathbb{F}(X \xrightarrow{f} Y)$ the bivariant Chern class $\gamma_{X\to Y}(\alpha)$ is expressed by

$$\gamma_{X\to Y}(\alpha) = f^*s(TY) \cap c_*(\alpha)$$

where $s(TY)$ is the total Segre class of the tangent bundle TY, i.e., $s(TY) = c(TY)^{-1}$ the inverse of the total Chern class $c(TY)$.

Corollary (3.5). *The Brasselet bivariant Chern classes, defined on cellular morphisms with nonsingular target varieties, are unique.*

Thus in particular, we get the following

Corollary (3.6). *(Zhou's theorem [Z1, Z2]) For a morphism $f : X \to S$ with S being a smooth curve, Brasselet's bivariant Chern class and Sabbah's bivariant Chern class are the same.*

Proof of Theorem (3.4). First, the hypothesis that the target variety Y is nonsingular implies that we have

$$\mathbb{H}(X \xrightarrow{f} Y) = \mathbb{H}(X \to pt) = H_*(X).$$

This turns out to be a key fact. Let $\gamma : \mathbb{F} \to \mathbb{H}$ be a bivariant Chern class. Then it follows from (3.3) that we have

$$c_*(\alpha) = \gamma_{X \to Y}(\alpha) \bullet c_*(1\!\!1_Y).$$

To consider the product, we look at the following commutative diagram with $j : X \to Y \times M$ being an embedding and $p : Y \times M \to Y$ the projection such that $f = p \circ j$:

$$\begin{tikzcd}
X \arrow[r, "j"] & Y \times M \arrow[r, "\mathrm{id}"] \arrow[d, "p"'] & Y \times M \arrow[d, "p"] \\
& Y \arrow[r, "\mathrm{id}"] & Y \arrow[d] \\
& & pt
\end{tikzcd}$$

Hence we have, via the realization isomorphisms $\Re_{X \hookrightarrow Y \times M}$, that

$$c_*(\alpha) = \Big(\Re_{X \hookrightarrow Y \times M}\Big)^{-1} \Bigg(\Re_{X \hookrightarrow Y \times M}(\gamma_{X \to Y}(\alpha)) \cdot p^* \Re_{Y \hookrightarrow Y}(c_*(Y))\Bigg).$$

Here it should be noted that the realization isomorphism $\Re_{X \hookrightarrow Y \times M}$ functions as two kinds of realization isomorphism: the first one is

$$\Re_{X \hookrightarrow Y \times M} : \mathbb{H}(X \to pt) = H_*(X) \cong H^*(Y \times M, (Y \times M) \setminus X)$$

and the second one is

$$\Re_{X \hookrightarrow Y \times M} : \mathbb{H}(X \xrightarrow{f} Y) = H_*(X) \cong H^*(Y \times M, (Y \times M) \setminus X).$$

Since $\mathrm{id} : Y \times M \to Y \times M$ is the identity, it follows from the definition of the product $\cdot$ that it is nothing but the usual cup product, thus we have

$$c_*(\alpha) = \Big(\Re_{X \hookrightarrow Y \times M}\Big)^{-1} \Bigg(\Re_{X \hookrightarrow Y \times M}(\gamma_{X \to Y}(\alpha)) \cup p^* \Re_{Y \hookrightarrow Y}(c_*(Y))\Bigg).$$

Since Y is nonsingular, $c_*(Y) = c(TY) \cap [Y]$ and $\Re_{Y \hookrightarrow Y} : H_*(Y) = \mathbb{H}(Y \to pt) \cong H^*(Y)$ which is the inverse of the Poincaré duality isomorphism $\mathcal{P}_Y$, we have $\Re_{Y \hookrightarrow Y}(c_*(Y)) = c(TY)$. Therefore we get that

$$c_*(\alpha) = \Big(\Re_{X \hookrightarrow Y \times M}\Big)^{-1}\left(\Re_{X \hookrightarrow Y \times M}(\gamma_{X \to Y}(\alpha)) \cup p^* c(TY)\right).$$

Which implies that

$$\Re_{X \hookrightarrow Y \times M}(c_*(\alpha)) = \Re_{X \hookrightarrow Y \times M}(\gamma_{X \to Y}(\alpha)) \cup p^* c(TY).$$

Thus we get that

$$\Re_{X \hookrightarrow Y \times M}(\gamma_{X \to Y}(\alpha)) = \Re_{X \hookrightarrow Y \times M}(c_*(\alpha)) \cup p^* s(TY),$$

which implies that

$$\gamma_{X \to Y}(\alpha) = \Big(\Re_{X \hookrightarrow Y \times M}\Big)^{-1}\left(\Re_{X \hookrightarrow Y \times M}(c_*(\alpha)) \cup p^* s(TY)\right).$$

Furthermore this can be simplified more as follows. Since, as we observe in the previous section,

$$\Big(\Re_{X \hookrightarrow Y \times M}\Big)^{-1}(a) = a \cap [Y \times M],$$

we get

$$\gamma_{X \to Y}(\alpha) = \left(\Re_{X \hookrightarrow Y \times M}(c_*(\alpha)) \cup p^* s(TY)\right) \cap [Y \times M].$$

Then it follows from the equation [F, §19.1, (8), p.371] that we get the following:

$$\begin{aligned}
\gamma_{X \to Y}(\alpha) &= \left(\Re_{X \hookrightarrow Y \times M}(c_*(\alpha)) \cup p^* s(TY)\right) \cap [Y \times M] \\
&= j^* p^* s(TY) \cap \left(\Re_{X \hookrightarrow Y \times M}(c_*(\alpha)) \cap [Y \times M]\right) \\
&= j^* p^* s(TY) \cap c_*(\alpha) \\
&= f^* s(TY) \cap c_*(\alpha).
\end{aligned}$$

□

By the same argument as above, we can show the following:

Theorem (3.7). *The Grothendieck transformation*

$$\tau : \mathbb{K}_{\mathrm{alg}} \to \mathbb{H}_{\mathbb{Q}}$$

constructed in [FM, Part II] is unique on morphisms with nonsingular target varieties. And the bivariant class $\tau_{X\to Y}(\alpha)$ for a bivariant coherent sheaf $\alpha \in \mathbb{K}_{\mathrm{alg}}(X \to Y)$ is given by

$$\tau_{X\to Y}(\alpha) = \frac{1}{f^* td(TY)} \cap \tau^{\mathrm{BFM}}(\alpha)$$

where $\tau^{\mathrm{BFM}} : \mathbf{K}_0 \to H_{*\mathbb{Q}}$ is the Baum-Fulton-MacPherson's Riemann-Roch and $td(TY)$ is the total Todd class of the tangent bundle.

Remark (3.8). In the case when the target variety Y is singular, the above argument does not work at all. Thus the target variety being nonsingular is essential (cf. [Y]). However, "modulo resolution" the uniqueness holds. Namely, by taking any resolution of singularities $\pi : \widetilde{Y} \to Y$, for any bivariant constructible function $\alpha \in \mathbb{F}(X \to Y)$, the pullback $\pi^*\gamma(\alpha)$ is uniquely determined; i.e., suppose that we have two bivariant Chern classes $\gamma, \gamma' : \mathbb{F} \to \mathbb{H}$, then for any resolution $\pi : \widetilde{Y} \to Y$ we have

$$\pi^*\gamma(\alpha) = \pi^*\gamma'(\alpha).$$

It is the same for the Grothendieck transformation $\tau : \mathbb{K}_{\mathrm{alg}} \to \mathbb{H}_{\mathbb{Q}}$, i.e.,

$$\pi^*\tau(\alpha) = \pi^*\tau'(\alpha).$$

References

[BFM] P. Baum, W. Fulton and R. MacPherson, *Riemann-Roch for singular varieties*, Publ. Math. I.H.E.S. **45** (1975), 101–145.

[B] J.-P. Brasselet, *Existence des classes de Chern en théorie bivariante*, Astérisque **101-102** (1981), 7–22.

[BS] J.-P. Brasselet and M.-H. Schwartz, *Sur les classes de Chern d'une ensemble analytique complexe*, Astérisque **82–83** (1981), 93–148.

[BY] J.-P. Brasselet and S. Yokura, *Remarks on bivariant constructible functions*, Adv. Stud. Pure Math. **29** (2000), 53–77.

[FM] W. Fulton and R. MacPherson, *Categorical frameworks for the study of singular spaces*, Memoirs of Amer. Math.Soc. **243** (1981).

[G1] V. Ginzburg, *$\mathfrak{G}$-Modules, Springer's Representations and Bivariant Chern Classes*, Adv. in Maths. **61** (1986), 1–48.

[G2] ———, *Geometric methods in the representation theory of Hecke algebras and quantum groups, in "Representation theories and algebraic geometry (Montreal, PQ, 1997)" (ed. by A. Broer and A. Daigneault)* (1998), Kluwer Acad. Publ., Dordrecht, 127–183.

[M] R. MacPherson, *Chern classes for singular algebraic varieties*, Ann. of Math. **100** (1974), 423–432.

[S] C. Sabbah, *Espaces conormaux bivariants*, Thèse, Université Paris VII (1986).

[Sc] M.-H. Schwartz, *Classes caractéristiques définies par une stratification d'une variété analytique complexe*, C. R. Acad. Sci. Paris **t. 260** (1965), 3262-3264, 3535–3537.

[T] B. Teissier, *Sur la triangulation des morphismes sous-analytiques*, Publ. Math. I.H.E.S. **70** (1989), 169–198.

[V] J.-L. Verdier, *Spécialisation des classes de Chern*, Astérisque 82-83 (1981).

[Y] S. Yokura, *Remarks on Ginzburg's bivariant Chern classes*, to appear in Proc. Amer. Math. Soc. (2002).

[Z1] J. Zhou, *Classes de Chern en théorie bivariante*, in Thèse, Université Aix-Marseille II (1995).

[Z2] ______, *Classes de Chern bivariantes*, preprint (1999).

Shoji Yokura
Department of Mathematics
and Computer Science
Faculty of Science
University of Kagoshima
21-35 Korimoto 1-chome
Kagoshima 890-0065
Japan
yokura@sci.kagoshima-u.ac.jp

Homology Stability for Unitary Groups

B. Mirzaii, W. van der Kallen

Received: January 29, 2002

Communicated by Ulf Rehmann

Abstract. In this paper the homology stability for unitary groups over a ring with finite unitary stable rank is established. First we develop a 'nerve theorem' on the homotopy type of a poset in terms of a cover by subposets, where the cover is itself indexed by a poset. We use the nerve theorem to show that a poset of sequences of isotropic vectors is highly connected, as conjectured by Charney in the eighties. Homology stability of symplectic groups and orthogonal groups appear as a special case of our results.

2000 Mathematics Subject Classification: Primary 19G99; Secondary 11E70, 18G30, 19B10.
Keywords and Phrases: Poset, acyclicity, unitary groups, homology stability.

1. Introduction

Interest in homological stability problems in algebraic K-theory started with Quillen, who used it in [15] to study the higher K-groups of a ring of integers. As a result of stability he proved that these groups are finitely generated (see also [7]). After that there has been considerable interest in homological stability for general linear groups. The most general results in this direction are due to the second author [20] and Suslin [19].

Parallel to this, similar questions for other classical groups such as orthogonal and symplectic groups have been studied. For work in this direction, see [23], [1], [5], [12], [13]. The most general result is due to Charney [5]. She proved the homology stability for orthogonal and symplectic groups over a Dedekind domain. Panin in [13] proved a similar result but with a different method and with better range of stability.

Our goal in this paper is to prove that homology stabilizes of the unitary groups over rings with finite unitary stable rank. To do so we prove that the poset of isotropic unimodular sequences is highly connected. Recall that Panin in [12] had already sketched how one can do this for a finite dimensional affine algebra over an infinite field, in the case of symplectic and orthogonal

groups. However, while the assumption about the infinite field provides a significant simplification, it excludes cases of primary interest, namely rings that are finitely generated over the integers.
Our approach is as follows. We first extend a theorem of Quillen [16, Thm 9.1] which was his main tool to prove that certain posets are highly connected. We use it to develop a quantitative analogue for posets of the nerve theorem, which expresses the homotopy type of a space in terms of the the nerve of a suitable cover. In our situation both the elements of the cover and the nerve are replaced with posets. We work with posets of ordered sequences 'satisfying the chain condition', as this is a good replacement for simplicial complexes in the presence of group actions. (Alternatively one might try to work with barycentric subdivisions of a simplicial complex.) The new nerve theorem allows us to exploit the higher connectivity of the poset of unimodular sequences due to the second author. The higher connectivity of the poset of isotropic unimodular sequences follows inductively. We conclude with the homology stability theorem.

2. Preliminaries

Recall that a topological space X is (-1)-*connected* if it is non-empty, 0-*connected* if it is non-empty and path connected, 1-*connected* if it is non-empty and simply connected. In general for $n \geq 1$, X is called n-*connected* if X is nonempty, X is 0-connected and $\pi_i(X, x) = 0$ for every base point $x \in X$ and $1 \leq i \leq n$. For $n \geq -1$ a space X is called n-*acyclic* if it is nonempty and $\tilde{H}_i(X, \mathbb{Z}) = 0$ for $0 \leq i \leq n$. For $n < -1$ the conditions of n-connectedness and n-acyclicness are vacuous.

Theorem 2.1 (Hurewicz). *For $n \geq 0$, a topological space X is n-connected if and only if the reduced homology groups $\tilde{H}_i(X, \mathbb{Z})$ are trivial for $0 \leq i \leq n$ and X is 1-connected if $n \geq 1$.*

Proof. See [25], Chap. IV, Corollaries 7.7 and 7.8. □

Let X be a partially ordered set or briefly a *poset*. Consider the simplicial complex associated to X, that is the simplicial complex where vertices or 0-simplices are the elements of X and the k-simplices are the $(k+1)$-tuples $(x_0, \dots, x_k)$ of elements of X with $x_0 < \cdots < x_k$. We denote it again by X. We denote the geometric realization of X by $|X|$ and we consider it with the weak topology. It is well known that $|X|$ is a CW-complex [11]. By a *morphism* or *map* of posets $f : X \to Y$ we mean an order-preserving map i. e. if $x \leq x'$ then $f(x) \leq f(x')$. Such a map induces a continuous map $|f| : |X| \to |Y|$.

Remark 2.2. If K is a simplicial complex and X the partially ordered set of simplices of K, then the space $|X|$ is the barycentric subdivision of K. Thus every simplicial complex, with weak topology, is homeomorphic to the geometric realization of some, and in fact many, posets. Furthermore since it is well known that any CW-complex is homotopy equivalent to a simplicial complex, it follows that any interesting homotopy type is realized as the geometric realization of a poset.

Proposition 2.3. *Let X and Y be posets.*
(i) *(*Segal [17]*) If $f, g : X \to Y$ are maps of posets such that $f(x) \le g(x)$ for all $x \in X$, then $|f|$ and $|g|$ are homotopic.*
(ii) *If the poset X has a minimal or maximal element then $|X|$ is contractible.*
(iii) *If X^{op} denotes the opposite poset of X, i. e. with opposite ordering, then $|X^{op}| \simeq |X|$.*

Proof. (i) Consider the poset $I = \{0, 1 : 0 < 1\}$ and define the poset map $h : I \times X \to Y$ as $h(0, x) = f(x)$, $h(1, x) = g(x)$. Since $|I| \simeq [0, 1]$, we have $|h| : [0, 1] \times |X| \to |Y|$ with $|h|(0, x) = |f|(x)$ and $|h|(1, x) = |g|(x)$. This shows that $|f|$ and $|g|$ are homotopic.
(ii) Suppose X has a maximal element z. Consider the map $f : X \to X$ with $f(x) = z$ for every $x \in X$. Clearly for every $x \in X$, $\mathrm{id}_X(x) \le f(x)$. This shows that id_X and the constant map f are homotopic. So X is contractible. If X has a minimal element the proof is similar.
(iii). This is natural and easy. $\square$

The construction $X \mapsto |X|$ allows us to assign topological concepts to posets. For example we define the homology groups of a poset X to be those of $|X|$, we call X *n-connected* or *contractible* if $|X|$ is n-connected or contractible etc. Note that X is connected if and only if X is connected as a poset. By the *dimension* of a poset X, we mean the dimension of the space $|X|$, or equivalently the supremum of the integers n such that there is a chain $x_0 < \cdots < x_n$ in X. By convention the empty set has dimension -1.
Let X be a poset and $x \in X$. Define $Link_X^+(x) := \{u \in X : u > x\}$ and $Link_X^-(x) := \{u \in X : u < x\}$. Given a map $f : X \to Y$ of posets and an element $y \in Y$, define subposets f/y and $y\backslash f$ of X as follows

$$f/y := \{x \in X : f(x) \le y\} \qquad y\backslash f := \{x \in X : f(x) \ge y\}.$$

In fact $f/y = f^{-1}(Y_{\le y})$ and $y\backslash f = f^{-1}(Y_{\ge y})$ where $Y_{\le y} = \{z \in Y : z \le y\}$ and $Y_{\ge y} = \{z \in Y : z \ge y\}$. Note that by 2.3 (ii), $Y_{\le y}$ and $Y_{\ge y}$ are contractible. If $\mathrm{id}_Y : Y \to Y$ is the identity map, then $\mathrm{id}_Y/y = Y_{\le y}$ and $y\backslash\mathrm{id}_Y = Y_{\ge y}$.
Let $\mathcal{F} : X \to \underline{\mathrm{Ab}}$ be a functor from a poset X, regarded as a category in the usual way, to the category of abelian groups. We define the homology groups $H_i(X, \mathcal{F})$ of X with coefficient $\mathcal{F}$ to be the homology of the complex $C_*(X, \mathcal{F})$ given by

$$C_n(X, \mathcal{F}) = \bigoplus_{x_0 < \cdots < x_n} \mathcal{F}(x_0)$$

where the direct sum is taken over all n-simplices in X, with differential $\partial_n = \Sigma_{i=0}^n (-1)^i d_i^n$ where $d_i^n : C_n(X, \mathcal{F}) \to C_{n-1}(X, \mathcal{F})$ and d_i^n takes the $(x_0 < \cdots < x_n)$-component of $C_n(X, \mathcal{F})$ to the $(x_0 < \cdots < \widehat{x_i} < \cdots < x_n)$-component of $C_{n-1}(X, \mathcal{F})$ via $d_i^n = id_{\mathcal{F}(x_0)}$ if $i > 0$ and $d_0^n : \mathcal{F}(x_0) \to \mathcal{F}(x_1)$. In particular, for the empty set we have $H_i(\varnothing, \mathcal{F}) = 0$ for $i \ge 0$.

Let $\mathcal{F}$ be the constant functor $\mathbb{Z}$. Then the homology groups with this coefficient coincide with the integral homology of $|X|$, that is $H_k(X, \mathbb{Z}) = H_k(|X|, \mathbb{Z})$ for all $k \in \mathbb{Z}$, [6, App. II]. Let $\tilde{H}_i(X, \mathbb{Z})$ denote the reduced integral homology of the poset X, that is $\tilde{H}_i(X, \mathbb{Z}) = \ker\{H_i(X, \mathbb{Z}) \to H_i(pt, \mathbb{Z})\}$ if $X \neq \varnothing$ and $\tilde{H}_i(\varnothing, \mathbb{Z}) = \begin{cases} \mathbb{Z} & \text{if } i = -1 \\ 0 & \text{if } i \neq -1 \end{cases}$. So $\tilde{H}_i(X, \mathbb{Z}) = H_i(X, \mathbb{Z})$ for $i \geq 1$ and for $i = 0$ we have the exact sequence

$$0 \to \tilde{H}_0(X, \mathbb{Z}) \to H_0(X, \mathbb{Z}) \to \mathbb{Z} \to \tilde{H}_{-1}(X, \mathbb{Z}) \to 0$$

where $\mathbb{Z}$ is identified with the group $H_0(pt, \mathbb{Z})$. Notice that $H_0(X, \mathbb{Z})$ is identified with the free abelian group generated by the connected components of X.

A *local system* of abelian groups on a space (resp. poset) X is a functor $\mathcal{F}$ from the groupoid of X (resp. X viewed as a category), to the category of abelian groups which is morphism-inverting, i. e. such that the map $\mathcal{F}(x) \to \mathcal{F}(x')$ associated to a path from x to x' (resp. $x \leq x'$) is an isomorphism. Clearly, a local system $\mathcal{F}$ on a path connected space (resp. 0-connected poset) is determined, up to canonical isomorphism, by the following data: if $x \in X$ is a base point, it suffices to be given the group $\mathcal{F}(x)$ and an action of $\pi_1(X, x)$ on $\mathcal{F}(x)$.

The homology groups $H_k(X, \mathcal{F})$ of a space X with a local system $\mathcal{F}$ are a generalization of the ordinary homology groups. In fact if X is a 0-connected space and if $\mathcal{F}$ is a constant local system on X, then $H_k(X, \mathcal{F}) \simeq H_k(X, \mathcal{F}(x_0))$ for every $x_0 \in X$ [25, Chap. VI, 2.1].

Let X be a poset and $\mathcal{F}$ a local system on $|X|$. Then the restriction of $\mathcal{F}$ to X is a local system on X. Considering $\mathcal{F}$ as a functor from X to the category of abelian groups, we can define $H_k(X, \mathcal{F})$ as in the above. Conversely if $\mathcal{F}$ is a local system on the poset X, then there is a unique local system, up to isomorphism, on $|X|$ such that the restriction to X is $\mathcal{F}$ [25, Chap. VI, Thm 1.12], [14, I, Prop. 1]. We denote both local systems by $\mathcal{F}$.

Theorem 2.4. *Let X be a poset and $\mathcal{F}$ a local system on X. Then the homology groups $H_k(|X|, \mathcal{F})$ are isomorphic with the homology groups $H_k(X, \mathcal{F})$.*

Proof. See [25, Chap. VI, Thm. 4.8] or [14, I, p. 91]. □

Theorem 2.5. *Let X be a path connected space with a base point x and let $\mathcal{F}$ be a local system on X. Then the inclusion $\{x\} \hookrightarrow X$ induces an isomorphism $\mathcal{F}(x)/G \xrightarrow{\simeq} H_0(X, \mathcal{F})$ where G is the subgroup of $\mathcal{F}(x)$ generated by all the elements of the form $a - \beta a$ with $a \in \mathcal{F}(x)$, $\beta \in \pi_1(X, x)$.*

Proof. See [25], Chap. VI, Thm. 2.8* and Thm. 3.2. □

We need the following interesting and well known lemma about the covering spaces of the space $|X|$, where X is a poset (or more generally a simplicial set). For a definition of a covering space, useful for our purpose, and some more information, see [18, Chap. 2].

Lemma 2.6. *For a poset X the category of the covering spaces of the space $|X|$ is equivalent to the category $\mathcal{L}_S(X)$, the category of functors $\mathcal{F} : X \to \underline{\mathrm{Set}}$, where $\underline{\mathrm{Set}}$ is the category of sets, such that $\mathcal{F}(x) \to \mathcal{F}(x')$ is a bijection for every relation $x \leq x'$.*

Proof. See [16, Section 7] or [14, I, p. 90]. □

3. Homology and homotopy of posets

Theorem 3.1. *Let $f : X \to Y$ be a map of posets. Then there is a first quadrant spectral sequence*

$$E^2_{p,q} = H_p(Y, y \mapsto H_q(f/y, \mathbb{Z})) \Rightarrow H_{p+q}(X, \mathbb{Z}).$$

The spectral sequence is functorial, in the sense that if there is a commutative diagram of posets

$$\begin{array}{ccc} X' & \xrightarrow{f'} & Y' \\ \Big\downarrow{\scriptstyle g_X} & & \Big\downarrow{\scriptstyle g_Y} \\ X & \xrightarrow{f} & Y \end{array}$$

then there is a natural map from the spectral sequence arising from f' to the spectral sequence arising from f. Moreover the map $g_{X*} : H_i(X', \mathbb{Z}) \to H_i(X, \mathbb{Z})$ is compatible with this natural map.

Proof. Let $C_{*,*}(f)$ be the double complex such that $C_{p,q}(f)$ is the free abelian group generated by the set $\{(x_0 < \cdots < x_q, f(x_q) < y_0 < \cdots < y_p) : x_i \in X,\ y_i \in Y\}$. The first spectral sequence of this double complex has as E^1-term $E^1_{p,q}(\mathrm{I}) = H_q(C_{p,*}(f)) = \bigoplus_{y_0 < \cdots < y_p} H_q(f/y_0, \mathbb{Z})$. By the general theory of double complexes (see for example [24, Chap. 5]), we know that $E^2_{p,q}(\mathrm{I})$ is the homology of the chain complex $E^1_{*,q}(\mathrm{I}) = C_*(Y, \mathcal{G}_q)$ where $\mathcal{G}_q : Y \to \underline{\mathrm{Ab}}$, $\mathcal{G}_q(y) = H_q(f/y, \mathbb{Z})$ and hence $E^2_{p,q}(\mathrm{I}) = H_p(Y, \mathcal{G}_q) = H_p(Y, y \mapsto H_q(f/y, \mathbb{Z}))$. The second spectral sequence has as E^1-term $E^1_{p,q}(\mathrm{II}) = H_q(C_{*,p}(f)) = \bigoplus_{f(x_p) < y_0 < \cdots < y_q} H_q(f(x_p)\backslash \mathrm{id}_Y, \mathbb{Z})$. But by 2.3 (ii), $f(x_p)\backslash \mathrm{id}_Y = Y_{\geq f(x_p)}$ is contractible, so $E^1_{*,0}(\mathrm{II}) = C_*(X^{op}, \mathbb{Z})$ and $E^1_{*,q}(\mathrm{II}) = 0$ for $q > 0$. Hence $H_i(\mathrm{Tot}(C_{*,*}(f))) \simeq H_i(X^{op}, \mathbb{Z}) \simeq H_i(X, \mathbb{Z})$. This completes the proof of existence and convergence of the spectral sequence. The functorial behavior of the spectral sequence follows from the functorial behavior of the spectral sequence of a filtration [24, 5.5.1] and the fact that the first and the second spectral sequences of the double complex arise from some filtrations. □

Remark 3.2. The above spectral sequence is a special case of a more general Theorem [6, App. II]. The above proof is taken from [9, Chap. I] where the functorial behavior of the spectral sequence is more visible. For more details see [9].

Definition 3.3. Let X be a poset. A map $\mathrm{ht}_X : X \to \mathbb{Z}_{\geq 0}$ is called height function if it is a strictly increasing map.

Example 3.4. The height function $\mathrm{ht}_X(x) = 1 + \dim(Link_X^-(x))$ is the usual one considered in [16], [9] and [5].

Lemma 3.5. *Let X be a poset such that $Link_X^+(x)$ is $(n - \mathrm{ht}_X(x) - 2)$-acyclic, for every $x \in X$, where ht_X is a height function on X. Let $\mathcal{F} : X \to \underline{\mathrm{Ab}}$ be a functor such that $\mathcal{F}(x) = 0$ for all $x \in X$ with $\mathrm{ht}_X(x) \geq m$, where $m \geq 1$. Then $H_k(X, \mathcal{F}) = 0$ for $k \leq n - m$.*

Proof. First consider the case of a functor $\mathcal{F}$ such that $\mathcal{F}(x) = 0$ if $\mathrm{ht}_X(x) \neq m - 1$. Then $C_k(X, \mathcal{F}) = \bigoplus_{\substack{x_0 < \cdots < x_k \\ \mathrm{ht}_X(x_0) = m-1}} \mathcal{F}(x_0)$. Clearly $0 = d_0^k = \mathcal{F}(x_0 < x_1) = \mathcal{F}(x_0) \to \mathcal{F}(x_1)$. Thus $\partial_k = \Sigma_{i=1}^k (-1)^i d_i^k$. Define $C_{-1}(Link_X^+(x_0), \mathcal{F}(x_0)) = \mathcal{F}(x_0)$ and complete the singular complex of $Link_X^+(x_0)$ with coefficient in $\mathcal{F}(x_0)$ to

$$\cdots \to C_0(Link_X^+(x_0), \mathcal{F}(x_0)) \xrightarrow{\varepsilon} C_{-1}(Link_X^+(x_0), \mathcal{F}(x_0)) \to 0$$

where $\varepsilon((g_i)) = \Sigma_i g_i$. Then

$$\begin{aligned} C_k(X, \mathcal{F}) &= \bigoplus_{\mathrm{ht}_X(x_0) = m-1} \Big(\bigoplus_{\substack{x_1 < \cdots < x_k \\ x_0 < x_1}} \mathcal{F}(x_0) \Big) \\ &= \bigoplus_{\mathrm{ht}_X(x_0) = m-1} C_{k-1}(Link_X^+(x_0), \mathcal{F}(x_0)). \end{aligned}$$

The complex $C_{k-1}(Link_X^+(x_0), \mathcal{F}(x_0))$ is the standard complex for computing the reduced homology of $Link_X^+(x_0)$ with constant coefficient $\mathcal{F}(x_0)$. So

$$H_k(X, \mathcal{F}) = \bigoplus_{\mathrm{ht}_X(x) = m-1} \tilde{H}_{k-1}(Link_X^+(x), \mathcal{F}(x)).$$

If $\mathrm{ht}_X(x_0) = m - 1$ then $Link_X^+(x_0)$ is $(n - (m-1) - 2)$-acyclic, and by the universal coefficient theorem [18, Chap. 5, Thm. 8], $\tilde{H}_{k-1}(Link_X^+(x_0), \mathcal{F}(x_0)) = 0$ for $-1 \leq k - 1 \leq n - (m-1) - 2$. This shows that $H_k(X, \mathcal{F}) = 0$ for $0 \leq k \leq n - m$. To prove the lemma in general, we argue by induction on m. If $m = 1$ then for $\mathrm{ht}_X(x) \geq 1$, $\mathcal{F}(x) = 0$. So the lemma follows from the special case above. Suppose $m \geq 2$. Define $\mathcal{F}_0$ and $\mathcal{F}_1$ to be the functors

$$\mathcal{F}_0(x) = \begin{cases} \mathcal{F}(x) & \text{if } \mathrm{ht}_X(x) < m - 1 \\ 0 & \text{if } \mathrm{ht}_X(x) \geq m - 1 \end{cases}, \quad \mathcal{F}_1(x) = \begin{cases} \mathcal{F}(x) & \text{if } \mathrm{ht}_X(x) = m - 1 \\ 0 & \text{if } \mathrm{ht}_X(x) \neq m - 1 \end{cases}$$

respectively. Then there is a short exact sequence $0 \to \mathcal{F}_1 \to \mathcal{F} \to \mathcal{F}_0 \to 0$. By the above discussion, $H_k(X, \mathcal{F}_1) = 0$ for $0 \leq k \leq n - m$ and by induction for $m - 1$, we have $H_k(X, \mathcal{F}_0) = 0$ for $k \leq n - (m - 1)$. By the long exact sequence for the above short exact sequence of functors it is easy to see that $H_k(X, \mathcal{F}) = 0$ for $0 \leq k \leq n - m$. □

Theorem 3.6. *Let $f : X \to Y$ be a map of posets and ht_Y a height function on Y. Assume for every $y \in Y$, that $Link_Y^+(y)$ is $(n - \mathrm{ht}_Y(y) - 2)$-acyclic and f/y*

is $(\mathrm{ht}_Y(y)-1)$*-acyclic. Then* $f_*: H_k(X,\mathbb{Z}) \to H_k(Y,\mathbb{Z})$ *is an isomorphism for* $0 \le k \le n-1$.

Proof. By theorem 3.1, we have the first quadrant spectral sequence

$$E^2_{p,q} = H_p(Y, y \mapsto H_q(f/y,\mathbb{Z})) \Rightarrow H_{p+q}(X,\mathbb{Z}).$$

Since $H_q(f/y,\mathbb{Z}) = 0$ for $0 < q \le \mathrm{ht}_Y(y)-1$, the functor $\mathcal{G}_q : Y \to \underline{\mathrm{Ab}}$, $\mathcal{G}_q(y) = H_q(f/y,\mathbb{Z})$ is trivial for $\mathrm{ht}_Y(y) \ge q+1$, $q > 0$. By lemma 3.5, $H_p(Y,\mathcal{G}_q) = 0$ for $p \le n-(q+1)$. Hence $E^2_{p,q} = 0$ for $p+q \le n-1$, $q > 0$. If $q = 0$, by writing the long exact sequence for the short exact sequence $0 \to \tilde{H}_0(f/y,\mathbb{Z}) \to H_0(f/y,\mathbb{Z}) \to \mathbb{Z} \to 0$, valid because f/y is nonempty, we have

$$\cdots \to H_n(Y,\mathbb{Z}) \to H_{n-1}(Y, y \mapsto \tilde{H}_0(f/y,\mathbb{Z})) \to E^2_{n-1,0} \to$$
$$\cdots \to H_1(Y,\mathbb{Z}) \to H_0(Y, y \mapsto \tilde{H}_0(f/y,\mathbb{Z})) \to E^2_{0,0} \to H_0(Y,\mathbb{Z}) \to 0.$$

If $\mathrm{ht}_Y(y) \ge 1$, then $\tilde{H}_0(f/y,\mathbb{Z}) = 0$. By lemma 3.5, $H_k(Y, y \mapsto \tilde{H}_0(f/y,\mathbb{Z})) = 0$ for $0 \le k \le n-1$. Thus

$$E^2_{p,q} = \begin{cases} H_p(Y,\mathbb{Z}) & \text{if } q = 0,\ 0 \le p \le n-1 \\ 0 & \text{if } p+q \le n-1,\ q > 0 \end{cases}.$$

This shows that $E^2_{p,q} \simeq \cdots \simeq E^\infty_{p,q}$ for $0 \le p+q \le n-1$. Therefore $H_k(X,\mathbb{Z}) \simeq H_k(Y,\mathbb{Z})$ for $0 \le k \le n-1$. Now consider the commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \big\downarrow{\scriptstyle f} & & \big\downarrow{\scriptstyle \mathrm{id}_Y} \\ Y & \xrightarrow{\mathrm{id}_Y} & Y \end{array}.$$

By functoriality of the spectral sequence 3.1, and the above calculation we get the diagram

$$\begin{array}{ccc} H_k(Y, y \mapsto H_0(f/y,\mathbb{Z})) & \xrightarrow{\simeq} & H_k(X,\mathbb{Z}) \\ \big\downarrow{\scriptstyle \mathrm{id}_{Y*}} & & \big\downarrow{\scriptstyle f_*} \\ H_k(Y, y \mapsto H_0(\mathrm{id}_Y/y,\mathbb{Z})) & \xrightarrow{\simeq} & H_k(Y,\mathbb{Z}) \end{array}.$$

Since $\mathrm{id}_Y/y = Y_{\le y}$ is contractible, we have $H_k(Y, y \mapsto H_0(\mathrm{id}_Y/y,\mathbb{Z})) = H_k(Y,\mathbb{Z})$. The map id_{Y*} is an isomorphism for $0 \le k \le n-1$, from the above long exact sequence. This shows that f_* is an isomorphism for $0 \le k \le n-1$. $\square$

Lemma 3.7. *Let X be a 0-connected poset. Then X is 1-connected if and only if for every local system $\mathcal{F}$ on X and every $x \in X$, the map $\mathcal{F}(x) \to H_0(X,\mathcal{F})$, induced from the inclusion $\{x\} \hookrightarrow X$, is an isomorphism (or equivalently, every local system on X is a isomorphic with a constant local system).*

Proof. If X is 1-connected then by theorem 2.5 and the connectedness of X, one has $\mathcal{F}(x) \xrightarrow{\simeq} H_0(X,\mathcal{F})$ for every $x \in X$. Now let every local system on X be isomorphic with a constant local system. Let $\mathcal{F} : X \to \underline{\mathrm{Set}}$ be in $\mathcal{L}_S(X)$.

Define the functor $\mathcal{G} : X \to \underline{\mathrm{Ab}}$ where $\mathcal{G}(x)$ is the free abelian group generated by $\mathcal{F}(x)$. Clearly $\mathcal{G}$ is a local system and so it is constant system. It follows that $\mathcal{F}$ is isomorphic to a constant functor. So by lemma 2.6, any connected covering space of $|X|$ is isomorphic to $|X|$. This shows that the universal covering of $|X|$, is $|X|$. Note that the universal covering of a connected simplicial simplex exists and is simply connected [18, Chap. 2, Cor. 14 and 15]. Therefore X is 1-connected. $\square$

Theorem 3.8. *Let $f : X \to Y$ be a map of posets and ht_Y a height function on Y. Assume for every $y \in Y$, that $Link_Y^+(y)$ is $(n - \mathrm{ht}_Y(y) - 2)$-connected and f/y is $(\mathrm{ht}_Y(y) - 1)$-connected. Then X is $(n-1)$-connected if and only if Y is $(n-1)$-connected.*

Proof. By 2.1 and 3.6 we may assume $n \geq 2$. So it is enough to prove that X is 1-connected if and only if Y is 1-connected. Let $\mathcal{F} : X \to \underline{\mathrm{Ab}}$ be a local system. Define the functor $\mathcal{G} : Y \to \underline{\mathrm{Ab}}$ with

$$\mathcal{G}(y) = \begin{cases} H_0(f/y, \mathcal{F}) & \text{if } \mathrm{ht}_Y(y) \neq 0 \\ H_0(Link_Y^+(y), y' \mapsto H_0(f/y', \mathcal{F})) & \text{if } \mathrm{ht}_Y(y) = 0 \end{cases}.$$

We prove that $\mathcal{G}$ is a local system. If $\mathrm{ht}_Y(y) \geq 2$ then f/y is 1-connected and by 3.6, $\mathcal{F}|_{f/y}$ is a constant system, so by 3.7, $H_0(f/y, \mathcal{F}) \simeq \mathcal{F}(x)$ for every $x \in f/y$. If $\mathrm{ht}_Y(y) = 1$, then f/y is 0-connected and $Link_Y^+(y)$ is nonempty. Choose $y' \in Y$ such that $y < y'$. Now f/y' is 1-connected and so $\mathcal{F}|_{f/y'}$ is a constant system on f/y'. But $f/y \subset f/y'$, so $\mathcal{F}|_{f/y}$ is a constant system. Since f/y is 0-connected, by 2.5 and the fact that we mentioned before theorem 2.4, $H_0(f/y, \mathcal{F}) \simeq \mathcal{F}(x)$ for every $x \in f/y$. Now let $\mathrm{ht}_Y(y) = 0$. Then $Link_Y^+(y)$ is 0-connected, f/y is nonempty and for every $y' \in Link_Y^+(y)$, $H_0(f/y', \mathcal{F}) \simeq H_0((f/y)^\circ, \mathcal{F})$ where $(f/y)^\circ$ is a component of f/y, which we fix. This shows that the local system $\mathcal{F}' : Link_Y^+(y) \to \underline{\mathrm{Ab}}$ with $y' \mapsto H_0(f/y', \mathcal{F})$ is isomorphic to a constant system, so $H_0(Link_Y^+(y), y' \mapsto H_0(f/y', \mathcal{F})) = H_0(Link_Y^+(y), \mathcal{F}') \simeq \mathcal{F}'(y') \simeq \mathcal{F}(x)$ for every $x \in f/y'$. Therefore $\mathcal{G}$ is a local system.

If Y is 1-connected, by 3.7, $\mathcal{G}$ is a constant system. But it is easy to see that $\mathcal{F} \simeq \mathcal{G} \circ f$. Therefore $\mathcal{F}$ is a constant system. Since X is connected by our homology calculation, by 3.7 we conclude that X is 1-connected. Now let X be 1-connected. If $\mathcal{E}$ is a local system on Y, then $f^*\mathcal{E} := \mathcal{E} \circ f$ is a local system on X. So it is a constant local system. As above we can construct a local system $\mathcal{G}'$ on Y from $\mathcal{F}' := \mathcal{E} \circ f$. This gives a natural transformation from $\mathcal{G}'$ to $\mathcal{E}$ which is an isomorphism. Since $\mathcal{E} \circ f$ is constant, by 2.5 and 3.7 and an argument as above one sees that $\mathcal{G}'$ is constant. Therefore $\mathcal{E}$ is isomorphic to a constant local system and 3.7 shows that Y is 1-connected. $\square$

Remark 3.9. In the proof of the above theorem 3.8 we showed in fact that: Let $f : X \to Y$ be a map of posets and ht_Y a height function on Y and $n \geq 2$. Assume for every $y \in Y$, that $Link_Y^+(y)$ is $(n - \mathrm{ht}_Y(y) - 2)$-connected and f/y

is $(\mathrm{ht}_Y(y)-1)$-connected. Then $f^*:\mathcal{L}_S(Y)\to\mathcal{L}_S(X)$, with $\mathcal{E}\mapsto\mathcal{E}\circ f$ is an equivalence of categories.

Remark 3.10. Theorem 3.8 is a generalization of a theorem of Quillen [16, Thm. 9.1]. We proved that the converse of that theorem is also valid. Our proof is similar in outline to the proof by Quillen. Furthermore, lemma 3.5 is a generalized version of lemma 1.3 from [5]. With more restrictions, Maazen, in [9, Chap. II] gave an easier proof of Quillen's theorem.

4. Homology and homotopy of posets of sequences

Let V be a nonempty set. We denote by $\mathcal{O}(V)$ the poset of finite ordered sequences of distinct elements of V, the length of each sequence being at least one. The partial ordering on $\mathcal{O}(V)$ is defined by refinement: $(v_1,\dots,v_m)\le(w_1,\dots,w_n)$ if and only if there is a strictly increasing map $\phi:\{1,\dots,m\}\to\{1,\dots,n\}$ such that $v_i=w_{\phi(i)}$, in other words, if $(v_1,\dots,v_m)$ is an order preserving subsequence of $(w_1,\dots,w_n)$. If $v=(v_1,\dots,v_m)$ we denote by $|v|$ the length of v, that is $|v|=m$. If $v=(v_1,\dots,v_m)$ and $w=(w_1,\dots,w_n)$, we write $(v_1,\dots,v_m,w_1,\dots,w_n)$ as vw. For $v\in F$, but for such v only, we define F_v to be the set of $w\in F$ such that $wv\in F$. Note that $(F_v)_w=F_{wv}$. A subset F of $\mathcal{O}(V)$ is said to satisfy the *chain condition* if $v\in F$ whenever $w\in F$, $v\in\mathcal{O}(V)$ and $v\le w$. The subposets of $\mathcal{O}(V)$ which satisfy the chain condition are extensively studied in [9], [20] and [4]. In this section we will study them some more.

Let $F\subseteq\mathcal{O}(V)$. For a nonempty set S we define the poset $F\langle S\rangle$ as

$$F\langle S\rangle:=\{((v_1,s_1),\dots,(v_r,s_r))\in\mathcal{O}(V\times S):(v_1,\dots,v_r)\in F\}.$$

Assume $s_0\in S$ and consider the injective poset map $l_{s_0}:F\to F\langle S\rangle$ with $(v_1,\dots,v_r)\mapsto((v_1,s_0),\dots,(v_r,s_0))$. We have clearly a projection $p:F\langle S\rangle\to F$ with $((v_1,s_1),\dots,(v_r,s_r))\mapsto(v_1,\dots,v_r)$ such that $p\circ l_{s_0}=\mathrm{id}_F$.

Lemma 4.1. *Suppose $F\subseteq\mathcal{O}(V)$ satisfies the chain condition and S is a nonempty set. Assume for every $v\in F$, that F_v is $(n-|v|)$-connected.*
(i) *If $s_0\in S$ then $(l_{s_0})_*:H_k(F,\mathbb{Z})\to H_k(F\langle S\rangle,\mathbb{Z})$ is an isomorphism for $0\le k\le n$.*
(ii) *If F is $\min\{1,n-1\}$-connected, then $(l_{s_0})_*:\pi_k(F,v)\to\pi_k(F\langle S\rangle,l_{s_0}(v))$ is an isomorphism for $0\le k\le n$.*

Proof. This follows by [4, Prop. 1.6] from the fact that $p\circ l_{s_0}=\mathrm{id}_F$. □

Lemma 4.2. *Let $F\subseteq\mathcal{O}(V)$ satisfies the chain condition. Then $|Link_F^-(v)|\simeq S^{|v|-2}$ for every $v\in F$.*

Proof. Let $v=(v_1,\dots,v_n)$. By definition $Link_F^-(v)=\{w\in F:w<v\}=\{(v_{i_1},\dots,v_{i_k}):k<n,i_1<\dots<i_k\}$. Hence $|Link_F^-(v)|$ is isomorphic to the barycentric subdivision of the boundary of the standard simplex Δ_{n-1}. It is well known that $\partial\Delta_{n-1}\simeq S^{n-2}$, hence $|Link_F^-(v)|\simeq S^{|v|-2}$. □

Theorem 4.3 (Nerve Theorem for Posets). *Let V and T be two nonempty sets, $F \subseteq \mathcal{O}(V)$ and $X \subseteq \mathcal{O}(T)$. Assume $X = \bigcup_{v \in F} X_v$ such that if $v \leq w$ in F, then $X_w \subseteq X_v$. Let F, X and X_v, for every $v \in F$, satisfy the chain condition. Also assume*
(i) *for every $v \in F$, X_v is $(l - |v| + 1)$-acyclic (resp. $(l - |v| + 1)$-connected),*
(ii) *for every $x \in X$, $\mathcal{A}_x := \{v \in F : x \in X_v\}$ is $(l - |x| + 1)$-acyclic (resp. $(l - |x| + 1)$-connected).*
Then $H_k(F, \mathbb{Z}) \simeq H_k(X, \mathbb{Z})$ for $0 \leq k \leq l$ (resp. F is l-connected if and only if X is l-connected).

Proof. Let $F_{\leq l+2} = \{v \in F : |v| \leq l + 2\}$ and let $i : F_{\leq l+2} \to F$ be the inclusion. Clearly $|F_{\leq l+2}|$ is the $(l + 1)$-skeleton of $|F|$, if we consider $|F|$ as a cell complex whose k-cells are the $|F_{\leq v}|$ with $|v| = k + 1$. It is well known that $i_* : H_k(F_{\leq l+2}, \mathbb{Z}) \to H_k(F, \mathbb{Z})$ and $i_* : \pi_k(F_{\leq l+2}, v) \to \pi_k(F, v)$ are isomorphisms for $0 \leq k \leq l$ (see [25], Chap. II, corollary 2.14, and [25], Chap. II, Corollary 3.10 and Chap. IV lemma 7.12.) So it is enough to prove the theorem for $F_{\leq l+2}$ and $X_{\leq l+2}$. Thus assume $F = F_{\leq l+2}$ and $X = X_{\leq l+2}$. We define $Z \subseteq X \times F$ as $Z = \{(x, v) : x \in X_v\}$. Consider the projections

$$f : Z \to F,\ (x, v) \mapsto v \qquad , \qquad g : Z \to X,\ (x, v) \mapsto x.$$

First we prove that $f^{-1}(v) \sim v\backslash f$ and $g^{-1}(x) \sim x\backslash g$, where $\sim$ means homotopy equivalence. By definition $v\backslash f = \{(x, w) : w \geq v,\ x \in X_w\}$. Define $\phi : v\backslash f \to f^{-1}(v)$, $(x, w) \mapsto (x, v)$. Consider the inclusion $j : f^{-1}(v) \to v\backslash f$. Clearly $\phi \circ j(x, v) = \phi(x, v) = (x, v)$ and $j \circ \phi(x, w) = j(x, v) = (x, v) \leq (x, w)$. So by 2.3(ii), $v\backslash f$ and $f^{-1}(v)$ are homotopy equivalent. Similarly $x\backslash g \sim g^{-1}(x)$.
Now we prove that the maps $f^{op} : Z^{op} \to Y^{op}$ and $g^{op} : Z^{op} \to X^{op}$ satisfy the conditions of 3.6. First $f^{op} : Z^{op} \to Y^{op}$; define the height function $\mathrm{ht}_{F^{op}}$ on F^{op} as $\mathrm{ht}_{F^{op}}(v) = l + 2 - |v|$. It is easy to see that $f^{op}/v \simeq v\backslash f \sim f^{-1}(v) \simeq X_v$. Hence f^{op}/v is $(l - |v| + 1)$-acyclic (resp. $(l - |v| + 1)$-connected). But $l-|v|+1 = (l+2-|v|)-1 = \mathrm{ht}_{F^{op}}(v)-1$, so f^{op}/v is $(\mathrm{ht}_{F^{op}}(v)-1)$-acyclic (resp. $(\mathrm{ht}_{F^{op}}(v) - 1)$-connected). Let $n := l + 1$. Clearly $Link^+_{F^{op}}(v) = Link^-_F(v)$. By lemma 4.2, $|Link^-_F(v)|$ is $(|v|-3)$-connected. But $|v|-3 = l+1-(l+2-|v|)-2 = n - \mathrm{ht}_{F^{op}}(v) - 2$. Thus $Link^+_{F^{op}}(v)$ is $(n - \mathrm{ht}_{F^{op}}(v) - 2)$-acyclic (resp. $(n - \mathrm{ht}_{F^{op}}(v) - 2)$-connected). Therefore by theorem 3.6, $f_* : H_i(Z, \mathbb{Z}) \to H_i(F, \mathbb{Z})$ is an isomorphism for $0 \leq i \leq l$ (resp. by 3.8, F is l-connected if and only if Z is l-connected). Now consider $g^{op} : Z^{op} \to X^{op}$. We saw in the above that $g^{op}/x \simeq x\backslash g \sim g^{-1}(x)$ and $g^{-1}(x) = \{(x, v) : x \in X_v\} \simeq \{v \in F : x \in X_v\}$. It is similar to the case of f^{op} to see that g^{op} satisfies the conditions of theorem 3.6, hence $g_* : H_i(Z, \mathbb{Z}) \to H_i(X, \mathbb{Z})$ is an isomorphism for $0 \leq i \leq l$ (resp. by 3.8, X is l-connected if and only if Z is l-connected). This completes the proof. $\square$

Let K be a simplicial complex and $\{K_i\}_{i \in I}$ a family of subcomplexes such that $K = \bigcup_{i \in I} K_i$. The *nerve* of this family of subcomplexes of K is the simplicial complex $\mathcal{N}(K)$ on the vertex set I so that a finite subset $\sigma \subseteq I$ is in $\mathcal{N}(K)$ if and only if $\bigcap_{i \in \sigma} K_i \neq \varnothing$. The nerve $\mathcal{N}(K)$ of K, with the inclusion relation, is a poset. As we already said we can consider a simplicial complex as a poset of its simplices.

COROLLARY 4.4 (Nerve Theorem). *Let K be a simplicial complex and $\{K_i\}_{i\in I}$ a family of subcomplexes such that $K = \bigcup_{i\in I} K_i$. Suppose every nonempty finite intersection $\bigcap_{j=1}^{t} K_{i_j}$ is $(l-t+1)$-acyclic (resp. $(l-t+1)$connected). Then $H_k(K,\mathbb{Z}) \simeq H_k(\mathcal{N}(K),\mathbb{Z})$ for $0 \le k \le l$ (resp. K is l-connected if and only if $\mathcal{N}(K)$ is l-connected).*

Proof. Let V be the set of vertices of K. We give a total ordering to V and I. Put $F = \{(i_1,\dots,i_r) : i_1 < \cdots < i_r \text{ and } \bigcap_{j=1}^{r} K_{i_j} \neq \varnothing\} \subseteq \mathcal{O}(I)$, $X = \{(x_1,\dots,x_t) : x_1 < \cdots < x_t \text{ and } \{x_1,\dots,x_t\} \text{ is a simplex in } K\} \subseteq \mathcal{O}(V)$ and for every $(i_1,\dots,i_r) \in F$, put $X_{(i_1,\dots,i_r)} = \{(x_1,\dots,x_t) \in X : \{x_1,\dots,x_t\} \in \bigcap_{j=1}^{r} K_{i_j}\}$. It is not difficult to see that $F \simeq \mathcal{N}(K)$ and $X \simeq K$. Also one should notice that $\mathcal{A}_x := \{v \in F : x \in X_v\}$ is contractible for $x \in X$. We leave the details to interested readers. □

Remark 4.5. In [7], a special case of the theorem 4.3 is proved. The nerve theorem for a simplicial complex 4.4, in the stated generality, is proved for the first time in [3], see also [2, p. 1850]. For more information about different types of nerve theorem and more references about them see [2, p. 1850].

LEMMA 4.6. *Let $F \subseteq \mathcal{O}(V)$ satisfy the chain condition and let $\mathcal{G} : F^{op} \to \underline{\mathrm{Ab}}$ be a functor. Then the natural map $\psi : \bigoplus_{v\in F,\ |v|=1} \mathcal{G}(v) \to H_0(F^{op},\mathcal{G})$ is surjective.*

Proof. By definition $C_0(F^{op},\mathcal{G}) = \bigoplus_{v\in F^{op}} \mathcal{G}(v)$, $C_1(F^{op},\mathcal{G}) = \bigoplus_{v<v'\in F^{op}} \mathcal{G}(v)$ and we have the chain complex

$$\cdots \to C_1(F^{op},\mathcal{G}) \xrightarrow{\partial_1} C_0(F^{op},\mathcal{G}) \to 0,$$

where $\partial_1 = d_0^1 - d_1^1$. Again by definition $H_0(F^{op},\mathcal{G}) = C_0(F^{op},\mathcal{G})/\partial_1$. Now let $w \in F$ and $|w| \ge 2$. Then there is a $v \in F$, $v \le w$, with $|v| = 1$. So $w < v$ in F^{op}, and we have the component $\partial_1|_{\mathcal{G}(w)} : \mathcal{G}(w) \to \mathcal{G}(w) \oplus \mathcal{G}(v)$, $x \mapsto d_0^1(x) - d_1^1(x) = d_0^1(x) - x$. This shows that $\mathcal{G}(w) \subseteq \mathrm{im}\partial_1 + \mathrm{im}\psi$. Therefore $H_0(F^{op},\mathcal{G})$ is generated by the groups $\mathcal{G}(v)$ with $|v| = 1$. □

THEOREM 4.7. *Let V and T be two nonempty sets, $F \subseteq \mathcal{O}(V)$ and $X \subseteq \mathcal{O}(T)$. Assume $X = \bigcup_{v\in F} X_v$ such that if $v \le w$ in F, then $X_w \subseteq X_v$ and let F, X and X_v, for every $v \in F$, satisfy the chain condition. Also assume*
(i) *for every $v \in F$, X_v is $\min\{l-1, l-|v|+1\}$-connected,*
(ii) *for every $x \in X$, $\mathcal{A}_x := \{v \in F : x \in X_v\}$ is $(l-|x|+1)$-connected,*
(iii) *F is l-connected.*
Then X is $(l-1)$-connected and the natural map

$$\bigoplus_{v\in F,\ |v|=1} (i_v)_* : \bigoplus_{v\in F,\ |v|=1} H_l(X_v,\mathbb{Z}) \to H_l(X,\mathbb{Z})$$

is surjective, where $i_v : X_v \to X$ is the inclusion. Moreover, if for every v with $|v| = 1$, there is an l-connected Y_v with $X_v \subseteq Y_v \subseteq X$, then X is also l-connected.

Proof. If $l = -1$, then everything is easy. If $l = 0$, then for v of length one, X_v is nonempty, so X is nonempty. This shows that X is (-1)-connected. Also, every connected component of X intersects at least one X_w and therefore also contains a connected component of an X_v with $|v| = 1$. This gives the surjectivity of the homomorphism

$$\bigoplus_{v\in F,\ |v|=1} (i_v)_* : \bigoplus_{v\in F,\ |v|=1} H_0(X_v, \mathbb{Z}) \to H_0(X, \mathbb{Z}).$$

Now assume that, for every v of length one, $X_v \subseteq Y_v$ where Y_v is connected. We prove, in a combinatorial way, that X is connected. Let $x, y \in X$, $x \in X_{(v_1)}$ and $y \in X_{(v_2)}$ where $(v_1), (v_2) \in F$. Since F is connected, there is a sequence $(w_1), \dots, (w_r) \in F$ such that they give a path, in F, from (v_1) to (v_2), that is

$$\begin{array}{ccccccccc} (v_1) & & & & (w_1) & & (w_r) & & & & (v_2) \\ & \diagdown & & \diagup & & \cdots & & \diagdown & & \diagup & \\ & & (v_1, w_1) & & & & & & (w_r, v_2) & & \end{array}\ .$$

Since $Y_{(v_1)}$ is connected, $x \in X_{(v_1)} \subseteq Y_{(v_1)}$ and $X_{(v_1,w_1)} \neq \varnothing$, there is an element $x_1 \in X_{(v_1,w_1)}$ such that there is a path, in $Y_{(v_1)}$, from x to x_1. Now $x_1 \in Y_{(w_1)}$. Similarly we can find $x_2 \in X_{(w_1,w_2)}$ such that there is a path, in $Y_{(w_1)}$, from x_1 to x_2. Now $x_2 \in Y_{(w_2)}$. Repeating this process finitely many times, we find a path from x to y. So X is connected.
Hence we assume that $l \geq 1$. As we said in the proof of theorem 4.3, we can assume that $F = F_{\leq l+2}$ and $X = X_{\leq l+2}$ and we define Z, f and g as we defined them there. Define the height function $\mathrm{ht}_{F^{op}}$ on F^{op} as $\mathrm{ht}_{F^{op}}(v) = l + 2 - |v|$. As we proved in the proof of theorem 4.3, $f^{op}/v \simeq v\backslash f \sim f^{-1}(v) \simeq X_v$. Thus f^{op}/v is $(\mathrm{ht}_{F^{op}}(v) - 1)$-connected if $|v| > 1$ and it is $(\mathrm{ht}_{F^{op}}(v) - 2)$-connected if $|v| = 1$ and also $|Link^{+}_{F^{op}}(v)|$ is $(l+1-\mathrm{ht}_{F^{op}}(v)-2)$-connected. By theorem 3.1, we have the first quadrant spectral sequence

$$E^2_{p,q} = H_p(F^{op}, v \mapsto H_q(f^{op}/v, \mathbb{Z})) \Rightarrow H_{p+q}(Z^{op}, \mathbb{Z}).$$

For $0 < q \leq \mathrm{ht}_{F^{op}}(v) - 2$, $H_q(f^{op}/v, \mathbb{Z}) = 0$. Define $\mathcal{G}_q : F^{op} \to \underline{\mathrm{Ab}}$, $\mathcal{G}_q(v) = H_q(f^{op}/v, \mathbb{Z})$. Then $\mathcal{G}_q(v) = 0$ for $\mathrm{ht}_{F^{op}}(v) \geq q + 2$, $q > 0$. By lemma 3.5, $H_p(F^{op}, \mathcal{G}_q) = 0$ for $p \leq l + 1 - (q + 2)$. Therefore $E^2_{p,q} = 0$ for $p + q \leq l - 1$, $q > 0$. If $q = 0$, arguing similarly to the proof of theorem 3.6, we get $E^2_{p,0} = 0$ if $0 < p \leq l - 1$ and $E^2_{0,0} = \mathbb{Z}$. Also by the fact that F^{op} is l-connected we get the surjective homomorphism $H_l(F^{op}, v \mapsto \tilde{H}_0(f^{op}/v, \mathbb{Z})) \twoheadrightarrow E^2_{l,0}$. Since $l \geq 1$, $\tilde{H}_0(f^{op}/v, \mathbb{Z}) = 0$ for all $v \in F^{op}$ with $\mathrm{ht}_{F^{op}}(v) \geq 1$ and so $H_l(F^{op}, v \mapsto \tilde{H}_0(f^{op}/v, \mathbb{Z})) = 0$ by lemma 3.5. Therefore $E^2_{l,0} = 0$. Let

$$\mathcal{G}'_q : F^{op} \to \underline{\mathrm{Ab}}, \quad \mathcal{G}'_q(v) = \begin{cases} 0 & \text{if } \mathrm{ht}_{F^{op}}(v) < l + 1 \\ H_q(f^{op}/v, \mathbb{Z}) & \text{if } \mathrm{ht}_{F^{op}}(v) = l + 1 \end{cases}$$

and

$$\mathcal{G''}_q : F^{op} \to \underline{\mathrm{Ab}}, \quad \mathcal{G''}_q(v) = \begin{cases} H_q(f^{op}/v, \mathbb{Z}) & \text{if } \mathrm{ht}_{F^{op}}(v) < l+1 \\ 0 & \text{if } \mathrm{ht}_{F^{op}}(v) = l+1. \end{cases}$$

Then we have the short exact sequence $0 \to \mathcal{G'}_q \to \mathcal{G}_q \to \mathcal{G''}_q \to 0$ and the associated long exact sequence

$$\cdots \to H_{l-q}(F^{op}, \mathcal{G'}_q) \to H_{l-q}(F^{op}, \mathcal{G}_q) \to$$
$$H_{l-q}(F^{op}, \mathcal{G''}_q) \to H_{l-q-1}(F^{op}, \mathcal{G'}_q) \to \cdots .$$

If $q > 0$, then $\mathcal{G''}_q(v) = 0$ for $0 < q \leq \mathrm{ht}_{F^{op}}(v) - 1$ and so by lemma 3.5, $H_p(F^{op}, \mathcal{G''}_q) = 0$ for $p + q \leq l$, $q > 0$. Also if $|v| = 1$ then $H_0(f^{op}/v, \mathbb{Z}) = 0$ for $0 < q \leq \mathrm{ht}_{F^{op}}(v) - 2 = l - 1$. This shows $\mathcal{G'}_q = 0$ for $0 < q \leq l-1$. From the long exact sequence and the above calculation we get

$$E^2_{p,q} = \begin{cases} \mathbb{Z} & \text{if } p = q = 0 \\ 0 & \text{if } 0 < p + q \leq l, \; q \neq l. \end{cases}$$

$$\begin{array}{r|cccccc} l+1 & * & & & & & E^2_{p,q} \\ l & * & * & & & & \\ & 0 & 0 & * & & & \\ & \vdots & \vdots & \ddots & \ddots & & \\ & 0 & \cdot & \cdot & \ddots & * & \\ 0 & \mathbb{Z} & 0 & \cdots & 0 & 0 & * \\ \hline & 0 & 1 & & & l & l+1 \end{array}$$

Thus for $0 \leq p+q \leq l$, $q \neq l$, $E^2_{p,q} \simeq \cdots \simeq E^{\infty}_{p,q}$ and there exist an integer r such that $E^2_{0,l} \twoheadrightarrow \cdots \twoheadrightarrow E^r_{0,l} \simeq E^{r+1}_{0,l} \simeq \cdots \simeq E^{\infty}_{0,l}$. Hence we get a surjective map $H_0(F^{op}, v \mapsto H_l(f^{op}/v, \mathbb{Z})) \twoheadrightarrow H_l(Z^{op}, \mathbb{Z})$. By lemma 4.6, we have a surjective map $\bigoplus_{v \in F, \, |v|=1} H_l(f^{op}/v, \mathbb{Z}) \twoheadrightarrow H_l(Z^{op}, \mathbb{Z})$.

Now consider the map $g^{op} : Z^{op} \to X^{op}$ and define the height function $\mathrm{ht}_{X^{op}}(x) = l + 2 - |x|$ on X^{op}. Arguing similarly to the proof of theorem 4.3 one sees that $g_* : H_k(Z, \mathbb{Z}) \to H_k(X, \mathbb{Z})$ is an isomorphism for $0 \leq k \leq l$. Therefore we get a surjective map $\bigoplus_{v \in F, \, |v|=1} H_l(X_v, \mathbb{Z}) \twoheadrightarrow H_l(X, \mathbb{Z})$. We call it ψ. We prove that this map is the same map that we claimed. For v of length one consider the commutative diagram of posets

$$\begin{tikzcd} & \{v\}^{op} \arrow[r] & F^{op} \\ f^{-1}(\{v\})^{op} \arrow[ur] \arrow[r] \arrow[dr] & Z^{op} \arrow[ur] \arrow[dr] & \\ & X^{op}_v \arrow[r] & X^{op} \end{tikzcd}$$

By functoriality of the spectral sequence for the above diagram and lemma 4.6 we get the commutative diagram

$$\begin{array}{ccc} H_l(f_v^{op}/v,\mathbb{Z}) & \xrightarrow{(j_v)_*} & \bigoplus_{v\in F,|v|=1} H_l(f^{op}/v,\mathbb{Z}) \\ \downarrow & & \downarrow \\ H_0(\{v\}^{op}, v\mapsto H_l(f_v^{op}/v,\mathbb{Z})) & \longrightarrow & H_0(F^{op}, v\mapsto H_l(f^{op}/v,\mathbb{Z})) \\ \downarrow & & \downarrow \\ H_l(f^{-1}(v)^{op},\mathbb{Z}) & \longrightarrow & H_l(Z^{op},\mathbb{Z}) \\ \downarrow & & \downarrow \\ H_l(X_v^{op},\mathbb{Z}) & \xrightarrow{(i_v^{op})_*} & H_l(X^{op},\mathbb{Z}) \end{array}$$

where $j_v : f_v^{op}/v \to f^{op}/v$ is the inclusion which is a homotopy equivalence as we already mentioned. It is not difficult to see that the composition of homomorphisms in the left column of the above diagram induces the identity map from $H_l(X_v,\mathbb{Z})$, the composition of homomorphisms in the right column of above diagram induces the surjective map ψ and the last row induces the homomorphism $(i_v)_*$. This show that $(i_v)_* = \psi|_{H_l(X_v,\mathbb{Z})}$. This completes the proof of surjectiveness.
Now let for v of length one $X_v \subseteq Y_v$ where Y_v is l-connected. Then we have the commutative diagram

$$\begin{array}{ccc} H_l(X_v,\mathbb{Z}) & \xrightarrow{(i_v)_*} & H_l(X,\mathbb{Z}) \\ & \searrow \quad \nearrow & \\ & H_l(Y_v,\mathbb{Z}) & \end{array}.$$

By the assumption $H_l(Y_v,\mathbb{Z})$ is trivial and this shows that $(i_v)_*$ is the zero map. Hence by the surjectivity, $H_l(X,\mathbb{Z})$ is trivial. If $l \geq 2$, the nerve theorem 4.3 says that X is simply connected and by the Hurewicz theorem 2.1, X is l-connected. So the only case that is left is when $l = 1$. By theorem 3.8, X is 1-connected if and only if Z is 1-connected. So it is enough to prove that Z^{op} is 1-connected. Note that as we said, we can assume that $F = F_{\leq 3}$ and $X = X_{\leq 3}$. Suppose $\mathcal{F}$ is a local system on Z^{op}. Define the functor $\mathcal{G} : F^{op} \to \underline{\mathrm{Ab}}$, as

$$\mathcal{G}(y) = \begin{cases} H_0(f^{op}/v,\mathcal{F}) & \text{if } |v| = 1,2 \\ H_0(Link^+_{F^{op}}(v), v' \mapsto H_0(f^{op}/v',\mathcal{F})) & \text{if } |v| = 3 \end{cases}.$$

We prove that $\mathcal{G}$ is a local system on F^{op}. Put $Z_w := g^{-1}(Y_w)$ for $|w| = 1$. If $|v| = 1,2$, then f^{op}/v is 0-connected and $f^{op}/v \subseteq Z_w^{op}$, where $w \leq v$, $|w| = 1$. By remark 3.9 we can assume that $\mathcal{F} = \mathcal{E}\circ g^{op}$ where $\mathcal{E}$ is a local system on X^{op}. Then $\mathcal{F}|_{Z_w^{op}} = \mathcal{E}|_{Y_w^{op}} \circ g^{op}|_{Z_w^{op}}$. Since Y_w^{op} is 1-connected, $\mathcal{E}|_{Y_w^{op}}$ is a constant local system. This shows that $\mathcal{F}|_{Z_w^{op}}$ is a constant local system. So $\mathcal{F}|_{f^{op}/v}$ is a constant local system and since f^{op}/v is 0-connected we have $H_0(f^{op}/v,\mathbb{Z}) \simeq \mathcal{F}(x)$, for every $x \in f^{op}/v$. If $|v| = 3$, with an argument similar to the proof of the theorem 3.8 and the above discussion one can get $\mathcal{G}(v) \simeq \mathcal{F}(x)$ for every $x \in f^{op}/v$. This shows that $\mathcal{G}$ is a local system on F^{op}. Hence it is a constant local system, because F^{op} is 1-connected. It is easy to see that $\mathcal{F} \simeq \mathcal{G}\circ f$. Therefore $\mathcal{F}$ is a constant system. Since X is connected by our homology calculation, by 3.7 we conclude that X is 1-connected. This completes the proof. $\square$

5. Posets of unimodular sequences

Let R be an associative ring with unit. A vector $(r_1, \ldots, r_n) \in R^n$ is called unimodular if there exist $s_1, \ldots, s_n \in R$ such that $\Sigma_{i=1}^n r_i s_i = 1$, or equivalently if the submodule generated by this vector is a free summand of the left R-module R^n. We denote the standard basis of R^n by $e_1, \ldots, e_n$. If $n \leq m$, we assume that R^n is the submodule of R^m generated by $e_1, \ldots, e_n \in R^m$.
We say that a ring R satisfies the *stable range condition* (S_m), if $m \geq 1$ is an integer so that for every unimodular vector $(r_0, r_1, \ldots, r_m) \in R^{m+1}$, there exist $t_1, \ldots, t_m$ in R such that $(r_1 + r_0 t_1, \ldots, r_m + r_0 t_m) \in R^m$ is unimodular. We say that R has *stable rank* m, we denote it with $\mathrm{sr}(R) = m$, if m is the least number such that (S_m) holds. If such a number does not exist we say that $\mathrm{sr}(R) = \infty$.
Let $\mathcal{U}(R^n)$ denote the subposet of $\mathcal{O}(R^n)$ consisting of unimodular sequences. Recall that a sequence of vectors $v_1, \ldots, v_k$ in R^n is called unimodular when $v_1, \ldots, v_k$ is basis of a free direct summand of R^n. Note that if $(v_1, \ldots, v_k) \in \mathcal{O}(R^n)$ and if $n \leq m$, it is the same to say that $(v_1, \ldots, v_k)$ is unimodular as a sequence of vectors in R^n or as a sequence of vectors in R^m. We call an element $(v_1, \ldots, v_k)$ of $\mathcal{U}(R^n)$ a *k-frame.*

Theorem 5.1 (Van der Kallen). *Let R be a ring with $\mathrm{sr}(R) < \infty$ and $n \leq m+1$. Let δ be 0 or 1. Then*
(i) *$\mathcal{O}(R^n + \delta e_{n+1}) \cap \mathcal{U}(R^m)$ is $(n - \mathrm{sr}(R) - 1)$-connected.*
(ii) *$\mathcal{O}(R^n + \delta e_{n+1}) \cap \mathcal{U}(R^m)_v$ is $(n - \mathrm{sr}(R) - |v| - 1)$-connected for all $v \in \mathcal{U}(R^m)$.*

Proof. See [20, Thm. 2.6]. $\square$

Example 5.2. Let R be a ring with $\mathrm{sr}(R) < \infty$. Let $n \geq \mathrm{sr}(R) + k + 1$ and assume $(v_1, \ldots, v_k) \in \mathcal{U}(R^{2n})$. Set $W = e_2 + \Sigma_{i=2}^n R e_{2i}$. Renumbering the basis one gets by theorem 5.1 that the poset $F := \mathcal{O}(W) \cap \mathcal{U}(R^{2n})_{(v_1, \ldots, v_k)}$ is $((n-1) - \mathrm{sr}(R) - k - 1)$-connected. Since $n \geq \mathrm{sr}(R) + k + 1$, it follows that F is not empty. This shows that there is $v \in W$ such that $(v, v_1, \ldots, v_k) \in \mathcal{U}(R^{2n})$. We will need such result in the next section but with a different method we can prove a sharper result. Compare this with lemma 5.4.

An $n \times k$-matrix B with $n < k$ is called unimodular if B has a right inverse. If B is an $n \times k$-matrix and $C \in GL_k(R)$, then B is unimodular if and only if CB is unimodular. A matrix of the form $\begin{pmatrix} 1 & u \\ 0 & B \end{pmatrix}$, where u is a row vector with coordinates in R, is unimodular if and only if the matrix B is unimodular.
We say that the ring R satisfies the *stable range condition* (S_k^n) if for every $n \times (n+k)$-matrix B, there exists a vector $r = (r_1, \ldots, r_{n+k-1})$ such that

$B\begin{pmatrix} 1 & r \\ 0 & I_{n+k-1} \end{pmatrix} = \begin{pmatrix} u & B' \end{pmatrix}$, where the $n \times (n+k-1)$-matrix B' is unimodular and u is the first column of the matrix B. Note that (S_k^1) is the same as (S_k).

Theorem 5.3 (Vaserstein). *For every $k \geq 1$ and $n \geq 1$, a ring R satisfies* (S_k) *if and only if it satisfies* (S_k^n).

Proof. The definition of (S_k^n) and the proof of this theorem is similar to the theorem [22, Thm. 3′] of Vaserstein. □

Lemma 5.4. *Let R be ring with $\mathrm{sr}(R) < \infty$ and let $n \geq \mathrm{sr}(R)+k$. Then for every $(v_1, \dots, v_k) \in \mathcal{U}(R^{2n})$ there is a $v \in e_2 + \Sigma_{i=2}^n Re_{2i}$ such that $(v, v_1, \dots, v_k) \in \mathcal{U}(R^{2n})$.*

Proof. There is a permutation matrix $A \in GL_{2n}(R)$ such that $(e_2 + \Sigma_{i=2}^n Re_{2i})A = e_1 + \Sigma_{j=n+2}^{2n} Re_j$. Let $w_i = v_i A$ for $i = 1, \dots, k$. So $(w_1, \dots, w_k) \in \mathcal{U}(R^{2n})$. Consider the $k \times 2n$-matrix B whose i-th row is the vector w_i. By theorem 5.3 there exists a vector $r = (r_2, \dots, r_{2n})$ such that $B\begin{pmatrix} 1 & r \\ 0 & I_{2n-1} \end{pmatrix} = \begin{pmatrix} u_1 & B_1 \end{pmatrix}$, where the $k \times (2n-1)$-matrix B_1 is unimodular and u_1 is the first column of the matrix B. Now let $s = (s_3, \dots, s_{2n})$ such that $B_1\begin{pmatrix} 1 & s \\ 0 & I_{2n_2} \end{pmatrix} = \begin{pmatrix} u_2 & B_2 \end{pmatrix}$, where the $k \times (2n-2)$-matrix B_2 is unimodular and u_2 is the first column of the matrix B_1. Now clearly

$$B\begin{pmatrix} 1 & r \\ 0 & I_{2n-1} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & s \\ 0 & 0 & I_{2n-2} \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & B_2 \end{pmatrix}.$$

By continuing this process, n times, we find a $2n \times 2n$ matrix C of the form

$$\begin{pmatrix} 1 & * & * & * & \\ & \ddots & * & * & \mathrm{N} \\ & & 1 & * & * \\ & 0 & & 1 & * \\ & & & & I_{n-1} \end{pmatrix}$$

where N is an $(n-1)\times(n-1)$ matrix and $BC = (L \mid M)$ where L is a $k \times (n+1)$ matrix and M is a unimodular $k \times (n-1)$ matrix. Now let $t = (t_{n+2}, \dots, t_{2n}) = -(\text{first row of } N)$. Then

$$\begin{pmatrix} 1 & 0 & \dots & 0 & t_{n+2} & \dots & t_{2n} \\ & & & B & & & \end{pmatrix} C = \begin{pmatrix} 1 & * & * & 0 & \dots & 0 \\ * & * & * & & M & \end{pmatrix}.$$

Since M is unimodular the right hand side of the above equality is unimodular. This shows that the matrix

$$\begin{pmatrix} 1 & 0 & \dots & 0 & t_{n+2} & \dots & t_{2n} \\ & & & B & & & \end{pmatrix}$$

is unimodular. Put $w = (1, 0, \dots, 0, t_{n+2}, \dots, t_{2n})$. Then $(w, w_1, \dots, w_k) \in \mathcal{U}(R^{2n})$. Now $v = wA^{-1}$ is the one that we are looking for. $\square$

6. HYPERBOLIC SPACES AND SOME POSETS

Let there be an involution on R, that is an automorphism of the additive group of R, $R \to R$ with $r \mapsto \overline{r}$, such that $\overline{\overline{r}} = r$ and $\overline{rs} = \overline{s}\,\overline{r}$. Let ϵ be an element in the center of R such that $\epsilon\overline{\epsilon} = 1$. Set $R_\epsilon := \{r - \epsilon\overline{r} : r \in R\}$ and $R^\epsilon := \{r \in R : \epsilon\overline{r} = -r\}$ and observe that $R_\epsilon \subseteq R^\epsilon$. A *form parameter* relative to the involution and ϵ is a subgroup Λ of $(R, +)$ such that $R_\epsilon \subseteq \Lambda \subseteq R^\epsilon$ and $\overline{r}\Lambda r \subseteq \Lambda$, for all $r \in R$. Notice that R_ϵ and R^ϵ are form parameters. We denote them by $\Lambda_{\min}$ and $\Lambda_{\max}$, respectively. If there is an s in the center of R such that $s + \overline{s} \in R^*$, in particular if $2 \in R^*$, then $\Lambda_{\min} = \Lambda_{\max}$.
Let $e_{i,j}(r)$ be the $2n \times 2n$-matrix with $r \in R$ in the (i,j) place and zero elsewhere. Consider $Q_n = \Sigma_{i=1}^n e_{2i-1,2i}(1) \in M_{2n}(R)$ and $F_n = Q_n + \epsilon\ {}^tQ_n = \Sigma_{i=1}^n (e_{2i-1,2i}(1) + e_{2i,2i-1}(\epsilon)) \in GL_{2n}(R)$. Define the bilinear map $h : R^{2n} \times R^{2n} \to R$ by $h(x,y) = \Sigma_{i=1}^n (\overline{x_{2i-1}} y_{2i} + \epsilon\overline{x_{2i}} y_{2i-1})$ and $q : R^{2n} \to R/\Lambda$ by $q(x) = \Sigma_{i=1}^n \overline{x_{2i-1}} x_{2i} \bmod \Lambda$, where $x = (x_1, \dots, x_{2n})$, $y = (y_1, \dots, y_{2n})$ and $\overline{x} = (\overline{x_1}, \dots, \overline{x_{2n}})$. The triple (R^{2n}, h, q) is called a *hyperbolic space*. By definition the *unitary group* relative Λ is the group

$$U_{2n}^\epsilon(R, \Lambda) := \{A \in GL_{2n}(R) : h(xA, yA) = h(x,y), q(xA) = q(x), x, y \in R\}.$$

For more general definitions and the properties of these spaces and groups see [8].

Example 6.1. (i) Let $\Lambda = \Lambda_{\max} = R$. Then $U_{2n}^\epsilon(R, \Lambda) = \{A \in GL_{2n}(R) : h(xA, yA) = h(x,y) \text{ for all } x, y \in R^{2n}\} = \{A \in GL_{2n}(R) : {}^t\overline{A} F_n A = F_n\}$. In particular if $\epsilon = -1$ and if the involution is the identity map id_R, then $\Lambda_{\max} = R$. In This case $U_{2n}^\epsilon(R, \Lambda_{\max}) := Sp_{2n}(R)$ is the usual symplectic group. Note that R is commutative in this case.
(ii) Let $\Lambda = \Lambda_{\min} = 0$. Then $U_{2n}^\epsilon(R, \Lambda) = \{A \in GL_{2n}(R) : q(xA) = q(x) \text{ for all } x \in R^{2n}\}$. In particular if $\epsilon = 1$ and if the involution is the identity map id_R, then $\Lambda_{\min} = 0$. In this case $U_{2n}^\epsilon(R, \Lambda_{\min}) := O_{2n}(R)$ is the usual orthogonal group. As in the symplectic case, R is necessarily commutative.
(iii) Let $\epsilon = -1$ and the involution is not the identity map id_R. If $\Lambda = \Lambda_{\max}$ then $U_{2n}^\epsilon(R, \Lambda) := U_{2n}(R)$ is the classical unitary group corresponding to the involution.

Let σ be the permutation of the set of natural numbers given by $\sigma(2i) = 2i - 1$ and $\sigma(2i-1) = 2i$. For $1 \le i, j \le 2n$, $i \ne j$, and every $r \in R$ define

$$E_{i,j}(r) = \begin{cases} I_{2n} + e_{i,j}(r) & \text{if } i = 2k-1, j = \sigma(i), r \in \Lambda \\ I_{2n} + e_{i,j}(r) & \text{if } i = 2k, j = \sigma(i), \overline{r} \in \Lambda \\ I_{2n} + e_{i,j}(r) + e_{\sigma(j),\sigma(i)}(-\overline{r}) & \text{if } i + j = 2k,\ i \ne j \\ I_{2n} + e_{i,j}(r) + e_{\sigma(j),\sigma(i)}(-\epsilon^{-1}\overline{r}) & \text{if } i \ne \sigma(j),\ i = 2k-1,\ j = 2l \\ I_{2n} + e_{i,j}(r) + e_{\sigma(j),\sigma(i)}(\epsilon\overline{r}) & \text{if } i \ne \sigma(j),\ i = 2k,\ j = 2l-1 \end{cases}$$

where I_{2n} is the identity element of $GL_{2n}(R)$. It is easy to see that $E_{i,j}(r) \in U^{\epsilon}_{2n}(R,\Lambda)$. Let $EU^{\epsilon}_{2n}(R,\Lambda)$ be the group generated by the $E_{i,j}(r)$, $r \in R$. We call it *elementary unitary group.*

A nonzero vector $x \in R^{2n}$ is called isotropic if $q(x) = 0$. This shows automatically that if x is isotropic then $h(x,x) = 0$. We say that a subset S of R^{2n} is isotropic if for every $x \in S$, $q(x) = 0$ and for every $x, y \in S$, $h(x,y) = 0$. If $h(x,y) = 0$, then we say that x is perpendicular to y. We denote by $\langle S\rangle$ the submodule of R^{2n} generated by S, and by $\langle S\rangle^{\perp}$ the submodule consisting of all the elements of R^{2n} which are perpendicular to all the elements of S.

From now, we fix an involution, an ϵ, a form parameter Λ and we consider the triple (R^{2n}, h, q) as defined above.

Definition 6.2 (Transitivity condition). Let $r \in R$ and define $C^{\epsilon}_r(R^{2n},\Lambda) = \{x \in \mathrm{Um}(R^{2n}) : q(x) = r \bmod \Lambda\}$, where $\mathrm{Um}(R^{2n})$ is the set of all unimodular vectors of R^{2n}. We say that R satisfies the transitivity condition (T_n), if $EU^{\epsilon}_{2n}(R,\Lambda)$ acts transitively on $C^{\epsilon}_r(R^{2n},\Lambda)$, for every $r \in R$. It is easy to see that $e_1 + re_2 \in C^{\epsilon}_r(R^{2n},\Lambda)$.

Definition 6.3 (Unitary stable range). We say that a ring R satisfies the unitary stable range condition (US_m) if R satisfies the conditions (S_m) and (T_{m+1}). We say that R has unitary stable rank m, we denote it with $\mathrm{usr}(R)$, if m is the least number such that (US_m) is satisfied. If such a number does not exist we say that $\mathrm{usr}(R) = \infty$. Clearly $\mathrm{sr}(R) \le \mathrm{usr}(R)$.

Remark 6.4. Our definition of unitary stable range is a little different than the one in [8]. In fact if (USR_{m+1}) satisfied then, by [8, Chap. VI, Thm. 4.7.1], (US_m) is satisfied where (USR_{m+1}) is the unitary stable range as defined in [8, Chap. VI, 4.6]. In comparison with the absolute stable rank $\mathrm{asr}(R)$ from [10], we have that if $m \ge \mathrm{asr}(R) + 1$ or if the involution is the identity map (so R is commutative) and $m \ge \mathrm{asr}(R)$ then (US_m) is satisfied [10, 8.1].

Example 6.5. Let R be a commutative Noetherian ring where the dimension d of the maximal spectrum $\mathrm{Mspec}(R)$ is finite. If A is a finite R-algebra then $\mathrm{usr}(A) \le d + 1$ (see [21, Thm. 2.8], [8, Thm. 6.1.4]). In particular if R is local ring or more generally a semilocal ring then $\mathrm{usr}(R) = 1$ [8, 6.1.3].

Lemma 6.6. *Let R be a ring with $\mathrm{usr}(R) < \infty$. Assume $n \ge \mathrm{usr}(R) + k$ and $(v_1, \dots, v_k) \in \mathcal{U}(R^{2n})$. Then there is a hyperbolic basis $\{x_1, y_1, \dots, x_n, y_n\}$ of R^{2n} such that $v_1, \dots, v_k \in \langle x_1, y_1, \dots, x_k, y_k\rangle$.*

Proof. The proof is by induction on k. If $k = 1$, by definition of unitary stable range there is an $E \in EU^{\epsilon}_{2n}(R,\Lambda)$ such that $v_1E = e_1 + re_2$. So the base of the induction is true. Let $k \ge 2$ and assume the induction hypothesis. Arguing as in the base of the induction we can assume that $v_1 = (1, r, 0, \dots, 0)$, $r \in R$. Let $W = e_2 + \Sigma^n_{i=2} Re_{2i}$. By lemma 5.4, choose $w \in W$ so that $(w, v_1, \dots, v_k) \in \mathcal{U}(R^{2n})$. Then $(w, v_1 - rw, v_2, \dots, v_k) \in \mathcal{U}(R^{2n})$. But $(w, v_1 - rw)$ is a hyperbolic pair, so there is an $E \in EU^{\epsilon}_{2n}(R,\Lambda)$ such that $wE = e_{2n-1}, (v_1 - rw)E = e_{2n}$ by [8, Chap. VI, Thm. 4.7.1]. Let

$(wE, (v_1-rw)E, v_2E, \dots, v_kE) =: (w_0, w_1, \dots, w_k)$ where $w_i = (r_{i,1}, \dots, r_{i,2n})$. Put $u_i = w_i - r_{i,2n-1}e_{2n-1} - r_{i,2n}e_{2n}$ for $2 \le i \le k$. Then $(u_2, \dots, u_k) \in \mathcal{U}(R^{2n-2})$. Now by induction there is a hyperbolic basis $\{a_2, b_2, \dots, a_n, b_n\}$ of R^{2n-2} such that $u_i \in \langle a_2, b_2, \dots, a_k, b_k\rangle$. Let $a_1 = e_{2n-1}$ and $b_1 = e_{2n}$. Then $w_i \in \langle a_1, b_1, \dots, a_k, b_k\rangle$. But $v_1E = w_1 + rwE = e_{2n} + re_{2n-1}$, $v_iE = w_i$ for $2 \le i \le k$ and considering $x_i = a_iE^{-1}$, $y_i = b_iE^{-1}$, one sees that $v_1, \dots, v_k \in \langle x_1, y_1, \dots, x_k, y_k\rangle$. □

Definition 6.7. Let $Z_n = \{x \in R^{2n} : q(x) = 0\}$. We define the poset $\mathcal{U}'(R^{2n})$ as $\mathcal{U}'(R^{2n}) := \mathcal{O}(Z_n) \cap \mathcal{U}(R^{2n})$.

Lemma 6.8. *Let R be a ring with* $\mathrm{sr}(R) < \infty$ *and* $n \le m$. *Then*
(i) $\mathcal{O}(R^{2n}) \cap \mathcal{U}'(R^{2m})$ *is* $(n - \mathrm{sr}(R) - 1)$*-connected,*
(ii) $\mathcal{O}(R^{2n}) \cap \mathcal{U}'(R^{2m})_v$ *is* $(n - \mathrm{sr}(R) - |v| - 1)$*-connected for every* $v \in \mathcal{U}'(R^{2m})$,
(iii) $\mathcal{O}(R^{2n}) \cap \mathcal{U}'(R^{2m}) \cap \mathcal{U}(R^{2m})_v$ *is* $(n - \mathrm{sr}(R) - |v| - 1)$*-connected for every* $v \in \mathcal{U}(R^{2m})$.

Proof. Let $W = \langle e_2, e_4, \dots, e_{2n}\rangle$ and $F := \mathcal{O}(R^{2n}) \cap \mathcal{U}'(R^{2m})$. It is easy to see that $\mathcal{O}(W) \cap F = \mathcal{O}(W) \cap \mathcal{U}(R^{2m})$ and $\mathcal{O}(W) \cap F_u = \mathcal{O}(W) \cap \mathcal{U}(R^{2m})_u$ for every $u \in \mathcal{U}'(R^{2m})$. By theorem 5.1, the poset $\mathcal{O}(W) \cap F$ is $(n - \mathrm{sr}(R) - 1)$-connected and the poset $\mathcal{O}(W) \cap F_u$ is $(n - \mathrm{sr}(R) - |u| - 1)$-connected for every $u \in F$. It follows from lemma [20, 2.13 (i)] that F is $(n - \mathrm{sr}(R) - 1)$-connected. The proof of (ii) and (iii) is similar to the proof of (i). □

Lemma 6.9. *Let R be a ring with* $\mathrm{usr}(R) < \infty$ *and let* $(v_1, \dots, v_k) \in \mathcal{U}'(R^{2n})$. *If* $n \ge \mathrm{usr}(R) + k$ *then* $\mathcal{O}(\langle v_1, \dots, v_k\rangle^{\perp}) \cap \mathcal{U}'(R^{2n})_{(v_1,\dots,v_k)}$ *is* $(n - \mathrm{usr}(R) - k - 1)$*-connected.*

Proof. By lemma 6.6 there is a hyperbolic basis $\{x_1, y_1, \dots, x_n, y_n\}$ of R^{2n} such that $v_1, \dots, v_k \in \langle x_1, y_1, \dots, x_k, y_k\rangle$. Let $W = \langle x_{k+1}, y_{k+1}, \dots, x_n, y_n\rangle \simeq R^{2(n-k)}$ and $F := \mathcal{O}(\langle v_1, \dots, v_k\rangle^{\perp}) \cap \mathcal{U}'(R^{2n})_{(v_1,\dots,v_k)}$. It is easy to see that $\mathcal{O}(W) \cap F = \mathcal{O}(W) \cap \mathcal{U}'(R^{2n})$. Let $V = \langle v_1, \dots, v_k\rangle$, then $\langle x_1, y_1, \dots, x_k, y_k\rangle = V \oplus P$ where P is a (finitely generated) projective module. Consider $(u_1, \dots, u_l) \in F \backslash \mathcal{O}(W)$ and let $u_i = x_i + y_i$ where $x_i \in V$ and $y_i \in P \oplus W$. One should notice that $(u_1 - x_1, \dots, u_l - x_l) \in \mathcal{U}(R^{2n})$ and not necessarily in $\mathcal{U}'(R^{2n})$. It is not difficult to see that $\mathcal{O}(W) \cap F_{(u_1,\dots,u_l)} = \mathcal{O}(W) \cap \mathcal{U}'(R^{2n}) \cap \mathcal{U}(R^{2n})_{(u_1-x_1,\dots,u_l-x_l)}$. By lemma 6.8, $\mathcal{O}(W) \cap F$ is $(n - k - \mathrm{usr}(R) - 1)$-connected and $\mathcal{O}(W) \cap F_u$ is $(n - k - \mathrm{usr}(R) - |u| - 1)$-connected for every $u \in F \backslash \mathcal{O}(W)$. It follows from lemma [20, 2.13 (i)] that F is $(n - \mathrm{usr}(R) - k - 1)$-connected. □

7. Posets of isotropic and hyperbolic unimodular sequences

Let $\mathcal{IU}(R^{2n})$ be the set of sequences $(x_1, \dots, x_k)$, $x_i \in R^{2n}$, such that $x_1, \dots, x_k$ form a basis for an isotropic direct summand of R^{2n}. Let $\mathcal{HU}(R^{2n})$ be the set of sequences $((x_1, y_1), \dots, (x_k, y_k))$ such that $(x_1, \dots, x_k)$, $(y_1, \dots, y_k) \in \mathcal{IU}(R^{2n})$, $h(x_i, y_j) = \delta_{i,j}$, where $\delta_{i,j}$ is the Kronecker delta. We call $\mathcal{IU}(R^{2n})$ and $\mathcal{HU}(R^{2n})$ the poset of isotropic unimodular sequences and the poset of hyperbolic unimodular sequences, respectively. For $1 \le k \le n$, let $\mathcal{IU}(R^{2n}, k)$ and $\mathcal{HU}(R^{2n}, k)$ be the set of all elements of length k of $\mathcal{IU}(R^{2n})$ and $\mathcal{HU}(R^{2n})$ respectively. We call the elements of $\mathcal{IU}(R^{2n}, k)$ and $\mathcal{HU}(R^{2n}, k)$ the isotropic k-frames and the hyperbolic k-frames, respectively. Define the poset $\mathcal{MU}(R^{2n})$ as the set of $((x_1, y_1), \dots, (x_k, y_k)) \in \mathcal{O}(R^{2n} \times R^{2n})$ such that, (i) $(x_1, \dots, x_k) \in \mathcal{IU}(R^{2n})$, (ii) for each i, either $y_i = 0$ or $(x_j, y_i) = \delta_{ji}$, (iii) $\langle y_1, \dots, y_k\rangle$ is isotropic. We identify $\mathcal{IU}(R^{2n})$ with $\mathcal{MU}(R^{2n}) \cap \mathcal{O}(R^{2n} \times \{0\})$ and $\mathcal{HU}(R^{2n})$ with $\mathcal{MU}(R^{2n}) \cap \mathcal{O}(R^{2n} \times (R^{2n} \backslash \{0\}))$.

Lemma 7.1. *Let R be a ring with $\operatorname{usr}(R) < \infty$. If $n \ge \operatorname{usr}(R) + k$ then $EU_{2n}^{\epsilon}(R, \Lambda)$ acts transitively on $\mathcal{IU}(R^{2n}, k)$ and $\mathcal{HU}(R^{2n}, k)$.*

Proof. The proof is by induction on k. If $k = 1$, by definition $EU_{2n}^{\epsilon}(R, \Lambda)$ acts transitively on $\mathcal{IU}(R^{2n}, 1)$ and by [8, Chap. VI, Thm. 4.7.1] the group $EU_{2n}^{\epsilon}(R, \Lambda)$ acts transitively on $\mathcal{HU}(R^{2n}, 1)$. The rest is an easy induction and the fact that for every isotropic k-frame $(x_1, \dots, x_k)$ there is an isotropic k-frame $(y_1, \dots, y_k)$ such that $((x_1, y_1), \dots, (x_k, y_k))$ is a hyperbolic k-frame [8, Chap. I, Cor. 3.7.4]. □

Lemma 7.2. *Let R be a ring with $\operatorname{usr}(R) < \infty$, and let $n \ge \operatorname{usr}(R) + k$. Let $((x_1, y_1), \dots, (x_k, y_k)) \in \mathcal{HU}(R^{2n})$, $(x_1, \dots, x_k) \in \mathcal{IU}(R^{2n})$ and $V = \langle x_1, \dots, x_k\rangle$. Then*
(i) $\mathcal{IU}(R^{2n})_{(x_1,\dots,x_k)} \simeq \mathcal{IU}(R^{2(n-k)})\langle V\rangle$,
(ii) $\mathcal{HU}(R^{2n}) \cap \mathcal{MU}(R^{2n})_{((x_1,0),\dots,(x_k,0))} \simeq \mathcal{HU}(R^{2n})_{((x_1,y_1),\dots,(x_k,y_k))}\langle V \times V\rangle$,
(iii) $\mathcal{HU}(R^{2n})_{((x_1,y_1),\dots,(x_k,y_k))} \simeq \mathcal{HU}(R^{2(n-k)})$.

Proof. See [5], the proof of lemma 3.4 and the proof of Thm. 3.2. □

For a real number l, by $\lfloor l \rfloor$ we mean the largest integer n with $n \le l$.

Theorem 7.3. *The poset $\mathcal{IU}(R^{2n})$ is $\lfloor \frac{n-\operatorname{usr}(R)-2}{2} \rfloor$-connected and $\mathcal{IU}(R^{2n})_x$ is $\lfloor \frac{n-\operatorname{usr}(R)-|x|-2}{2} \rfloor$-connected for every $x \in \mathcal{IU}(R^{2n})$.*

Proof. If $n \le \operatorname{usr}(R)$, the result is clear, so let $n > \operatorname{usr}(R)$. Let $X_v = \mathcal{IU}(R^{2n}) \cap \mathcal{U}'(R^{2n})_v \cap \mathcal{O}(\langle v\rangle^{\perp})$, for every $v \in \mathcal{U}'(R^{2n})$, and put $X := \bigcup_{v\in F} X_v$ where $F = \mathcal{U}'(R^{2n})$. It follows from lemma 7.1 that $\mathcal{IU}(R^{2n})_{\le n-\operatorname{usr}(R)} \subseteq X$. So to treat $\mathcal{IU}(R^{2n})$, it is enough to prove that X is $\lfloor \frac{n-\operatorname{usr}(R)-2}{2} \rfloor$-connected. First we prove that X_v is $\lfloor \frac{n-\operatorname{usr}(R)-|v|-2}{2} \rfloor$-connected for every $v \in F$. The proof is by descending induction on $|v|$. If $|v| > n-\operatorname{usr}(R)$, then $\lfloor \frac{n-\operatorname{usr}(R)-|v|-2}{2} \rfloor < -1$. In this case there is nothing to prove. If $n - \operatorname{usr}(R) - 1 \le |v| \le n - \operatorname{usr}(R)$, then $\lfloor \frac{n-\operatorname{usr}(R)-|v|-2}{2} \rfloor = -1$, so we must prove that X_v is nonempty. This follows from lemma 6.6. Now assume $|v| \le n - \operatorname{usr}(R) - 2$ and assume by induction that X_w is $\lfloor \frac{n-\operatorname{usr}(R)-|w|-2}{2} \rfloor$-connected for every w, with $|w| > |v|$. Let $l = \lfloor \frac{n-\operatorname{usr}(R)-|v|-2}{2} \rfloor$, and observe that $n - |v| - \operatorname{usr}(R) \ge l + 2$. Put $T_w = \mathcal{IU}(R^{2n}) \cap \mathcal{U}'(R^{2n})_{wv} \cap \mathcal{O}(\langle wv\rangle^{\perp})$ where $w \in G_v = \mathcal{U}'(R^{2n})_v \cap \mathcal{O}(\langle v\rangle^{\perp})$ and put $T := \bigcup_{w\in G_v} T_w$. It follows by lemma 6.6 that $(X_v)_{\le n-|v|-\operatorname{usr}(R)} \subseteq T$. So it is enough to prove that T is l-connected. The poset G_v is l-connected by lemma

6.9. By induction, T_w is $\lfloor\frac{n-\mathrm{usr}(R)-|v|-|w|-2}{2}\rfloor$-connected. But $\min\{l-1, l-|w|+1\} \le \lfloor\frac{n-\mathrm{usr}(R)-|v|-|w|-2}{2}\rfloor$, so T_w is $\min\{l-1, l-|w|+1\}$-connected. For every $y \in T$, $\mathcal{A}_y = \{w \in G_v : y \in T_w\}$ is isomorphic to $\mathcal{U}'(R^{2n})_{vy} \cap \mathcal{O}(\langle vy\rangle^{\perp})$ so by lemma 6.9, it is $(l-|y|+1)$-connected. Let $w \in G_v$ with $|w| = 1$. For every $z \in T_w$ we have $wz \in X_v$, so T_w is contained in a cone, call it C_w, inside X_v. Put $C(T_w) = T_w \cup (C_w)_{\le n-|v|-\mathrm{usr}(R)}$. Thus $C(T_w) \subseteq T$. The poset $C(T_w)$ is l-connected because $C(T_w)_{\le n-|v|-\mathrm{usr}(R)} = (C_w)_{\le n-|v|-\mathrm{usr}(R)}$. Now by theorems 5.1 and 4.7 , T is l-connected. In other words, we have now shown that X_v is $\lfloor\frac{n-\mathrm{usr}(R)-|v|-2}{2}\rfloor$-connected. By knowing this one can prove, in a similar way, that X is $\lfloor\frac{n-\mathrm{usr}(R)-2}{2}\rfloor$-connected. (Just pretend that $|v| = 0$.)
Now consider the poset $\mathcal{IU}(R^{2n})_x$ for an $x = (x_1, \dots, x_k) \in \mathcal{IU}(R^{2n})$. The proof is by induction on n. If $n = 1$, everything is easy. Similarly, we may assume $n - \mathrm{usr}(R) - |x| \ge 0$. Let $l = \lfloor\frac{n-\mathrm{usr}(R)-|x|-2}{2}\rfloor$. By lemma 7.2, $\mathcal{IU}(R^{2n})_x \simeq \mathcal{IU}(R^{2(n-|x|)})\langle V\rangle$, where $V = \langle x_1, \dots, x_k\rangle$. In the above we proved that $\mathcal{IU}(R^{2(n-|x|)})$ is l-connected and by induction, the poset $\mathcal{IU}(R^{2(n-|x|)})_y$ is $\lfloor\frac{n-|x|-\mathrm{usr}(R)-|y|-2}{2}\rfloor$-connected for every $y \in \mathcal{IU}(R^{2(n-|x|)})$. But $l-|y| \le \lfloor\frac{n-|x|-\mathrm{usr}(R)-|y|-2}{2}\rfloor$. So $\mathcal{IU}(R^{2(n-|x|)})\langle V\rangle$ is l-connected by lemma 4.1. Therefore $\mathcal{IU}(R^{2n})_x$ is l-connected. $\square$

THEOREM 7.4. *The poset $\mathcal{HU}(R^{2n})$ is $\lfloor\frac{n-\mathrm{usr}(R)-3}{2}\rfloor$-connected and $\mathcal{HU}(R^{2n})_x$ is $\lfloor\frac{n-\mathrm{usr}(R)-|x|-3}{2}\rfloor$-connected for every $x \in \mathcal{HU}(R^{2n})$.*

Proof. The proof is by induction on n. If $n = 1$, then everything is trivial. Let $F = \mathcal{IU}(R^{2n})$ and $X_v = \mathcal{HU}(R^{2n}) \cap \mathcal{MU}(R^{2n})_v$, for every $v \in F$. Put $X := \bigcup_{v\in F} X_v$. It follows from lemma 7.1 that $\mathcal{HU}(R^{2n})_{\le n-\mathrm{usr}(R)} \subseteq X$. Thus to treat $\mathcal{HU}(R^{2n})$, it is enough to prove that X is $\lfloor\frac{n-\mathrm{usr}(R)-3}{2}\rfloor$-connected, and we may assume $n \ge \mathrm{usr}(R)+1$. Take $l = \lfloor\frac{n-\mathrm{usr}(R)-3}{2}\rfloor$ and $V = \langle v_1, \dots, v_k\rangle$, where $v = (v_1, \dots, v_k)$. By lemma 7.2, there is an isomorphism $X_v \simeq \mathcal{HU}(R^{2(n-|v|)})\langle V \times V\rangle$, if $n \ge \mathrm{usr}(R) + |v|$. By induction $\mathcal{HU}(R^{2(n-|v|)})$ is $\lfloor\frac{n-|v|-\mathrm{usr}(R)-3}{2}\rfloor$-connected and again by induction $\mathcal{HU}(R^{2(n-|v|)})_y$ is $\lfloor\frac{n-|v|-\mathrm{usr}(R)-|y|-3}{2}\rfloor$-connected for every $y \in \mathcal{HU}(R^{2(n-|v|)})$. So by lemma 4.1 , X_v is $\lfloor\frac{n-|v|-\mathrm{usr}(R)-3}{2}\rfloor$-connected. Thus the poset X_v is $\min\{l-1, l-|v|+1\}$-connected. Let $x = ((x_1, y_1), \dots, (x_k, y_k))$. It is easy to see that $\mathcal{A}_x = \{v \in F : x \in X_v\} \simeq \mathcal{IU}(R^{2n})_{(x_1,\dots,x_k)}$. By the above theorem 7.3, $\mathcal{A}_x$ is $\lfloor\frac{n-\mathrm{usr}(R)-k-2}{2}\rfloor$-connected. But $l-|x|+1 \le \lfloor\frac{n-\mathrm{usr}(R)-k-2}{2}\rfloor$, so $\mathcal{A}_x$ is $(l-|x|+1)$-connected. Let $v = (v_1) \in F$, $|v| = 1$, and let $D_v := \mathcal{HU}(R^{2n})_{(v_1,w_1)} \simeq \mathcal{HU}(R^{2(n-1)})$ where $w_1 \in R^{2n}$ is a hyperbolic dual of $v_1 \in R^{2n}$. Then $D_v \subseteq X_v$ and D_v is contained in a cone, call it C_v, inside $\mathcal{HU}(R^{2n})$. Take $C(D_v) := D_v \cup (C_v)_{\le n-\mathrm{usr}(R)}$. By induction D_v is $\lfloor\frac{n-1-\mathrm{usr}(R)-3}{2}\rfloor$-connected and so $(l-1)$-connected. Let $Y_v = X_v \cup C(D_v)$. By the Mayer-Vietoris theorem and the fact that $C(D_v)$ is l-connected, we get the exact sequence

$$\tilde{H}_l(D_v, \mathbb{Z}) \stackrel{(i_v)_*}{\rightarrow} \tilde{H}_l(X_v, \mathbb{Z}) \rightarrow \tilde{H}_l(Y_v, \mathbb{Z}) \rightarrow 0.$$

where $i_v : D_v \rightarrow X_v$ is the inclusion. By induction $(D_v)_w$ is $\lfloor \frac{n-1-\mathrm{usr}(R)-|w|-3}{2} \rfloor$-connected and so $(l-|w|)$-connected, for $w \in D_v$. By lemma 4.1(i) and lemma 7.2, $(i_v)_*$ is an isomorphism, and by exactness of the above sequence we get $\tilde{H}_l(Y_v, \mathbb{Z}) = 0$. If $l \geq 1$ by the Van Kampen theorem $\pi_1(Y_v, x) \simeq \pi_1(X_v, x)/N$ where $x \in D_v$ and N is the normal subgroup generated by the image of the map $(i_v)_* : \pi_1(D_v, x) \rightarrow \pi_1(X_v, x)$. Now by lemma 4.1(ii), $\pi_1(Y_v, x)$ is trivial. Thus by the Hurewicz theorem 2.1, Y_v is l-connected. By having all this we can apply theorem 4.7 and so X is l-connected. The fact that $\mathcal{HU}(R^{2n})_x$ is $\lfloor \frac{n-\mathrm{usr}(R)-|x|-3}{2} \rfloor$-connected follows from the above and lemma 7.2. □

Remark 7.5. One can define a more generalized version of hyperbolic space $H(P) = P \oplus P^*$ where P is a finitely generated projective module. Charney in [5, 2.10] introduced the posets $\mathcal{IU}(P)$, $\mathcal{HU}(P)$ and conjectured that if P contains a free summand of rank on rank n then $\mathcal{IU}(P)$ and $\mathcal{HU}(P)$ are in fact highly connected. We leave it as exercise to the interested reader to prove this conjecture using the theorems 7.3 and 7.4 as in the proof of lemma 6.8. In fact one can prove that if P contains a free summand of rank n then $\mathcal{IU}(P)$ is $\lfloor \frac{n-\mathrm{usr}(R)-2}{2} \rfloor$-connected and $\mathcal{HU}(P)$ is $\lfloor \frac{n-\mathrm{usr}(R)-3}{2} \rfloor$-connected. Also, by assuming the high connectivity of the $\mathcal{IU}(R^{2n})$, Charney proved that $\mathcal{HU}(R^{2n})$ is highly connected. Our proof is different and relies on our theory, but we use ideas from her paper, such as the lemma 7.2 and her lemma 4.1, which is a modified version of work of Maazen [9].

8. Homology stability

From theorem 7.4 one can get the homology stability of unitary groups. The approach is well known.

Remark 8.1. To prove homology stability of this type one only needs high acyclicity of the corresponding poset, not high connectivity. But usually this type of posets are also highly connected. Here we also proved the high connectivity. In particular we wished to confirm the conjecture of Charney [5, 2.10], albeit with different bounds (see 7.5).

Theorem 8.2. *Let R be a ring with $\mathrm{usr}(R) < \infty$ and let the action of the unitary group on the Abelian group A is trivial. Then the homomorphism Inc_* : $H_k(U^\epsilon_{2n}(R, \Lambda), A) \rightarrow H_k(U^\epsilon_{2n+2}(R, \Lambda), A)$ is surjective for $n \geq 2k + \mathrm{usr}(R) + 2$ and injective for $n \geq 2k + \mathrm{usr}(R) + 3$.*

Proof. See [5, Section 4] and theorem 7.4. □

Remark 8.3. With the result of the previous section one also can prove homology stability of the unitary groups with twisted coefficients. For more information in this direction see [20, §5] and [5, 4.2].

References

[1] Betley, S. Homological stability for $O_{n,n}$ over semi-local rings. Glasgow Math. J. 32 (1990), no. 2, 255–259.

[2] Björner, A. Topological methods. Handbook of combinatorics, Vol. 1, 2, 1819–1872, Elsevier, Amsterdam, (1995).

[3] Björner, A.; Lovász, L.; Vrećica, S. T.; Živaljević, R. T. Chessboard complexes and matching complexes. J. London Math. Soc. (2) 49 (1994), no. 1, 25–39.

[4] Charney R. On the problem of homology stability for congruence subgroups. Comm. in Algebra 12 (47) (1984), 2081–2123.

[5] Charney, R. A generalization of a theorem of Vogtmann. J. Pure Appl. Algebra 44 (1987), 107–125.

[6] Gabriel, P.; Zisman, M. Calculus of Fractions and Homotopy Theory. Band 35 Springer-Verlag New York, (1967).

[7] Grayson, D. R. Finite generation of K-groups of a curve over a finite field (after Daniel Quillen). Algebraic K-theory, Part I, 69–90, Lecture Notes in Math., 966, (1982).

[8] Knus, M. A. Quadratic and Hermitian forms over rings. Grundlehren der Mathematischen Wissenschaften , 294. Springer-Verlag, Berlin, 1991.

[9] Maazen, H. Homology stability for the general linear group. Thesis, Utrecht, (1979).

[10] Magurn, B.; Van der Kallen, W.; Vaserstein, L. Absolute stable rank and Witt cancellation for noncommutative rings, Invent. Math. 19 (1988), 525–542.

[11] Milnor, J. The geometric realization of a semi-simplicial complex. Ann. of Math. (2) 65 (1957), 357–362.

[12] Panin, I. A. Homological stabilization for the orthogonal and symplectic groups. (Russian) Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 160 (1987), Anal. Teor. Chisel i Teor. Funktsii. 8, 222–228, 301–302 translation in J. Soviet Math. 52 (1990), no. 3, 3165–3170.

[13] Panin, I. A. On stabilization for orthogonal and symplectic algebraic K-theory. (Russian) Algebra i Analiz 1 (1989), no. 3, 172–195 translation in Leningrad Math. J. 1 (1990), no. 3, 741–764.

[14] Quillen, D. Higher algebraic K-theory. I. Algebraic K-theory, I: Higher K-theories, pp. 85–147. Lecture Notes in Math., Vol. 341, Springer, Berlin 1973.

[15] Quillen, D. Finite generation of the groups K_i of rings of algebraic integers. Algebraic K-theory, I: Higher K-theories, pp. 179–198. Lecture Notes in Math., Vol. 341, Springer, Berlin, 1973.

[16] Quillen, D. Homotopy properties of the poset of nontrivial p-subgroups of a group. Adv. in Math. 28, no. 2 (1978), 101–128.

[17] Segal, G. Classifying spaces and spectral sequences. I.H.E.S. Publ. Math. No. 34 (1968), 105–112.

[18] Spanier, E. H. Algebraic Topology. McGram Hill (1966).

[19] Suslin, A. A. Stability in algebraic K-theory. Algebraic K-theory, Part I (Oberwolfach, 1980), pp. 304–333, Lecture Notes in Math., 966, Springer, Berlin-New York, 1982.

[20] Van der Kallen, W. Homology stability for linear groups. Invent. Math. 60 (1980), 269–295.

[21] Vaserstein, L. N. Stabilization of unitary and orthogonal groups over a ring with involution. (Russian) Mat. Sb. 81 (123) (1970) 328–351, translation in Math. USSR Sbornik 10 (1970), no. 3, 307–326.

[22] Vaserstein, L. N. The stable range of rings and the dimension of topological spaces. (Russian) Funkcional. Anal. i Prilozen. 5 (1971) no. 2, 17–27, translation in Functional Anal. Appl. 5 (1971), 102–110.

[23] Vogtmann, K. Spherical posets and homology stability for $\mathrm{O}_{n,n}$. Topology 20 (1981), no. 2, 119–132.

[24] Weibel, C. A. An Introduction to Homological Algebra. Cambridge Studies in Advanced Mathematics, 38. Cambridge University Press, Cambridge, (1994).

[25] Whitehead, G. W. Elements of Homotopy Theory. Grad. Texts in Math. 61, Springer-Verlag, (1978).

Behrooz Mirzaii
Department of Mathematics
Utrecht University
P.O.Box 80.010
3508 TA Utrecht
The Netherlands.
mirzaii@math.uu.nl

Wilberd van der Kallen
Department of Mathematics
Utrecht University
P.O.Box 80.010
3508 TA Utrecht
The Netherlands.
vdkallen@math.uu.nl

The Ground State Energy of Relativistic One-Electron Atoms According to Jansen and Hess

Raymond Brummelhuis, Heinz Siedentop,
and Edgardo Stockmeyer

Received: August 16, 2001
Revised: June 15, 2002

Communicated by Alfred K. Louis

Abstract. Jansen and Heß – correcting an earlier paper of Douglas and Kroll – have derived a (pseudo-)relativistic energy expression which is very successful in describing heavy atoms. It is an approximate no-pair Hamiltonian in the Furry picture. We show that their energy in the one-particle Coulomb case, and thus the resulting self-adjoint Hamiltonian and its spectrum, is bounded from below for $\alpha Z \leq 1.006$.

2000 Mathematics Subject Classification: 81Q10, 81V45

1 Introduction

The energy of relativistic electrons in the electric field of a nucleus of charge Ze is described by the Dirac Operator

$$D_\gamma = c\boldsymbol{\alpha} \cdot \frac{\hbar}{i}\nabla + mc^2\beta - \frac{\gamma}{|\mathfrak{x}|} \tag{1}$$

with $\gamma = Ze^2$ and α, β the four Dirac matrices. The constant m is the mass of the electron, c is the velocity of light, and $\hbar$ is the rationalized Planck constant which we both take equal to one by a suitable choice of units. This operator describes both electrons and positrons. In low energy processes as, e.g., in quantum chemistry, there occur, however, only electrons. Brown and Ravenhall [2] proposed to project the positrons out and to use the electronic degrees of freedom only. They originally took the electrons and positrons given by the free Dirac operator D_0. Later it was observed that it might be suitable to define electrons directly by their external field (Furry picture). (See Sucher [17] for a review.) This strategy, however, meets immediate difficulties, since the projection $\chi_{(0,\infty)}(D_\gamma)$ is much harder to find for positive γ than for $\gamma = 0$. To handle this problem Douglas and Kroll [4] used an approximate Foldy-Wouthuysen transform to decouple the positive and negative spectral subspaces of D_γ. Their approximation is perturbative of second order in the coupling constant γ. Jansen and Heß [11] — correcting a sign mistake in [4] — wrote down pseudo-relativistic one- and multi-particle operators to describe the energy which were successfully used to describe heavy relativistic atoms (see, e.g., [12]).

This derivation yields the operator (see [11], Equation (17))

$$H_D^{\mathrm{ext}} = \beta e + E + \frac{1}{2}\left[W, O\right], \tag{2}$$

where

$$\begin{aligned} e(p) &:= \sqrt{\mathfrak{p}^2 + m^2}, & (3)\\ E &:= A(V + RVR)A, & (4)\\ O &:= \beta A[R, V]A, & (5)\\ A(\mathfrak{p}) &:= \left(\frac{e(p) + m}{2e(p)}\right)^{\frac{1}{2}}, & (6)\\ R(\mathfrak{p}) &:= \frac{\alpha \cdot \mathfrak{p}}{e(p) + m}, & (7)\\ W(\mathfrak{p}, \mathfrak{p}') &= \beta \frac{O(\mathfrak{p}, \mathfrak{p}')}{e(p) + e(p')}. & (8) \end{aligned}$$

(Note that we write p for $|\mathfrak{p}|$.) Here V is the external potential which in the case at hand is the Coulomb potential, and in configuration space it is multiplication by $-\gamma/|\mathfrak{x}|$.

This operator – which acts on four spinors – is then sandwiched by the projection onto the first two components, namely $(1+\beta)/2$. The resulting upper left corner matrix operator $J_\gamma : C_0^\infty(\mathbb{R}^3) \otimes \mathbb{C}^2 \to L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$ is

$$J_\gamma := B_\gamma + \gamma^2 \tilde{K} = e - (\gamma/(2\pi^2))K + \gamma^2 \tilde{K}. \tag{9}$$

with

$$K(\mathfrak{p}, \mathfrak{p}') = \frac{(e(p) + m)(e(p') + m) + (\mathfrak{p} \cdot \boldsymbol{\sigma})(\mathfrak{p}' \cdot \boldsymbol{\sigma})}{n(p)|\mathfrak{p} - \mathfrak{p}'|^2 n(p')} \tag{10}$$

where $n(p) := (2e(p)(e(p) + m))^{1/2}$, i.e., B_γ is the Brown-Ravenhall operator [2]. (See also Bethe and Salpeter[1] and Evans et al. [5]).

The last summand in (9) is given by the kernel

$$\tilde{K}(\mathfrak{p}, \mathfrak{p}') = -\frac{1}{2} \int d\mathfrak{p}'' [W(\mathfrak{p}, \mathfrak{p}'')P(\mathfrak{p}'', \mathfrak{p}') + P(\mathfrak{p}, \mathfrak{p}'')W(\mathfrak{p}'', \mathfrak{p}')] \tag{11}$$

with

$$P(\mathfrak{p},\mathfrak{p}') = \frac{\boldsymbol{\sigma}\cdot\mathfrak{p}(e(p')+m)-(e(p)+m)\boldsymbol{\sigma}\cdot\ gp'}{2\pi^2 n(p)|\mathfrak{p}-\mathfrak{p}'|^2 n(p')} \tag{12}$$

and

$$W(\mathfrak{p},\mathfrak{p}') = \frac{P(\mathfrak{p},\mathfrak{p}')}{e(p)+e(p')}. \tag{13}$$

Introducing $b(p) := p/n(p)$ and $a(p) := ((e(p)+m)/2e(p))^{1/2}$ we get more explicitly

$$\begin{aligned}
&\tilde{K}(\mathfrak{p},\mathfrak{p}')\\
&\quad = \frac{1}{2(2\pi^2)^2}\int d\mathfrak{p}'' \frac{1}{|\mathfrak{p}-\mathfrak{p}''|^2|\mathfrak{p}''-\mathfrak{p}'|^2}\left(\frac{1}{e(p)+e(p'')}+\frac{1}{e(p'')+e(p')}\right)\\
&\quad \big[\,(\omega_{\mathfrak{p}}\cdot\boldsymbol{\sigma})\,(\omega_{\mathfrak{p}'}\cdot\boldsymbol{\sigma})\,b(p)a(p'')^2 b(p') - (\omega_{\mathfrak{p}}\cdot\boldsymbol{\sigma})\,(\omega_{\mathfrak{p}''}\cdot\boldsymbol{\sigma})\,b(p)b(p'')a(p'')a(p')\\
&\qquad\qquad + a(p)b(p'')^2 a(p') - (\omega_{\mathfrak{p}''}\cdot\boldsymbol{\sigma})\,(\omega_{\mathfrak{p}'}\cdot\boldsymbol{\sigma})\,a(p)b(p'')a(p'')b(p')\big].
\end{aligned} \tag{14}$$

(For later use we name the expression in the first line of the integrand in (14) C and the four terms in the square bracket $T_1,...,T_4$.)
The corresponding energy in a state $u\in C_0^\infty(\mathbb{R}^3)\otimes\mathbb{C}^2$ is

$$\mathcal{J}(u) := (u, J_\gamma u) = \mathcal{B}(u) + \gamma^2(u,\tilde{K}u) \tag{15}$$

with

$$\mathcal{B}(u) = \int_{\mathbb{R}^3} d\mathfrak{p}\, e(p)|u(\mathfrak{p})|^2 - \frac{\gamma}{2\pi^2}\int_{\mathbb{R}^3} d\mathfrak{p}\int_{\mathbb{R}^3} d\mathfrak{p}' u(\mathfrak{p})^* K(\mathfrak{p},\mathfrak{p}')u(\mathfrak{p}') \tag{16}$$

It is the quadratic form $\mathcal{J}$ which is our prime interest.
Throughout the paper we will use the following constants $\gamma_c := 4\pi(\pi^2+4-\sqrt{-\pi^4+24\pi^2-16})/(\pi^2-4)^2)$, $\gamma_c^B := 2/(\pi/2+2/\pi)$, and $d_\gamma := 1-\gamma-4\sqrt{2}(3+\sqrt{2})\gamma^2$. Our goal is to show

Theorem 1. *For all nonnegative masses m the following holds:*

1. *If $\gamma\in[0,\gamma_c]$ then $\mathcal{J}$ is bounded from bellow, i.e., there exist a constant $c\in\mathbb{R}$ such that for all $u\in C_0^\infty(\mathbb{R}^3)\otimes\mathbb{C}^2$*

$$\mathcal{J}(u)\geq -cm\|u\|^2.$$

2. *If $\gamma>\gamma_c$, then $\mathcal{J}(u)$ is unbounded from below.*
3. *If $\gamma\in[0,\gamma_c^B]$ then*

$$\mathcal{J}(u)\geq d_\gamma m\|u\|^2.$$

Note that $\gamma_c\approx 1.006077340$. Because $\gamma=\alpha Z$ where α is the Sommerfeld fine structure constant which has the physical value of about $1/137$ and Z is the atomic number, this allows for the treatment of all known elements.

It also means that the method is applicable for all αZ where the Coulomb-Dirac operator can be defined in a natural way through form methods (Nenciu [15]). — Note, in particular, that the energy is bounded from below, even if $\gamma_c > \gamma > 1$ although the perturbative derivation of the symmetric operator H_D^{ext} is questionable in this case.
We would like to remark that the lower bound can most likely be improved for positive masses. In fact, we conjecture that the energy is positive for all sub-critical γ. However, this is outside the scope of this work.
According to Friedrichs our theorem has the following immediate consequence:

Corollary 1. *The symmetric operator J_γ has a unique self-adjoint extension whose form domain contains $C_0^\infty(\mathbb{R}^3) \otimes \mathbb{C}^2$ for $\gamma \in [0, \gamma_c]$.*

In fact for $\gamma < \gamma_c$, since the potential turns out to be form bounded with relative bound less than one, the self-adjoint operator defined has form domain $H^{1/2}(\mathbb{R}^3) \otimes \mathbb{C}^2$.
The structure of the paper is as follow: in Section 2 using spherical symmetry we decompose the operator in angular momentum channels. In Section 3 we prove the positivity of the massless operators. Since these operators are homogeneous under dilation an obvious tool to use is the Mellin transform, a method that previously has been used with success to obtain tight estimates on critical coupling constant (see, e.g., [3]). In Section 4 we find that the difference between the massless and the massive operator is bounded. Finally, some useful identities are given in the Appendix.

2 Partial Wave Analysis of the Energy

To obtain a sharp estimate for the potential energy we decompose the operator as direct sum on invariant subspaces. Because of the rotational symmetry of the problem one might suspect that the angular momenta are conserved quantities. Indeed, as a somewhat lengthy calculation shows, the total angular momentum $\mathfrak{J} = \frac{1}{2}(\mathfrak{x} \times \mathfrak{p} + \boldsymbol{\sigma})$ commutes with H^{ext}. In fact we can largely follow a strategy carried out by Hardekopf and Sucher [9] and Evans et al. [5] in somewhat simpler contexts.
We begin by observing that those of the spherical spinors

$$\Omega_{l,m,s}(\omega) := \begin{cases} \begin{pmatrix} \sqrt{\frac{l+s+m}{2(l+s)}}\, Y_{l,m-\frac{1}{2}}(\omega) \\ \sqrt{\frac{l+s-m}{2(l+s)}}\, Y_{l,m+\frac{1}{2}}(\omega) \end{pmatrix} & s = \frac{1}{2} \\ \begin{pmatrix} -\sqrt{\frac{l+s-m+1}{2(l+s)+2}}\, Y_{l,m-\frac{1}{2}}(\omega) \\ \sqrt{\frac{l+s+m+1}{2(l+s)+2}}\, Y_{l,m+\frac{1}{2}}(\omega) \end{pmatrix} & s = -\frac{1}{2} \end{cases} \tag{17}$$

with $l = 0, 1, 2, \ldots$ and $m = -l - \frac{1}{2}, \ldots, l + \frac{1}{2}$, that do not vanish, form an orthonormal basis of $L^2(S^2) \otimes \mathbb{C}^2$. Here $Y_{l,k}$ are normalized spherical harmonics on the unit sphere S^2 (see, e.g., [14], p. 421) with the convention that $Y_{l,k} = 0$, if $|k| > l$. We denote the corresponding index set by I, i.e., $I := \{(l,m,s) | l \in \mathbb{N}_0, m = -l - \frac{1}{2}, ..., l + \frac{1}{2}, s = \pm\frac{1}{2}, \Omega_{l,m,s} \neq 0\}$. Thus any $u \in L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$ can be written as

$$u(\mathfrak{p}) = \sum_{(l,m,s)\in I} p^{-1} f_{l,m,s}(p) \Omega_{l,m,s}(\omega_{\mathfrak{p}}) \tag{18}$$

where $p = |\mathfrak{p}|$, $\omega_{\mathfrak{p}} = \mathfrak{p}/p$, and

$$\sum_{(l,m,s)\in I} \int_0^\infty |f_{l,m,s}(p)|^2 dp = \int_{\mathbb{R}^3} |u(\mathfrak{p})|^2 d\mathfrak{p}.$$

We now remind the reader that the expansion of the Coulomb potential in spherical harmonics is given by

$$\frac{1}{|\mathfrak{p} - \mathfrak{p}'|^2} = \frac{2\pi}{pp'} \sum_{l=0}^{\infty} \sum_{m=-l}^{l} q_l(p/p') Y_{l,m}(\omega_{\mathfrak{p}}) \overline{Y}_{l,m}(\omega_{\mathfrak{p}'}) \tag{19}$$

where $q_l(x) := Q_l((x + 1/x)/2)$; Q_l are Legendre functions of the second kind, i.e.,

$$Q_l(z) = \tfrac{1}{2} \int_{-1}^{1} \tfrac{P_l(t)}{z-t} dt \tag{20}$$

where the P_l are Legendre polynomials. [See Stegun [16] for the notation and some properties of these special functions.]
Inserting the expansion (18) and (19) into (15) yields

$$\mathcal{J}(u) = \sum_{(l,m,s)\in I} \mathcal{J}_{l,s}(f_{l,m,s})$$

with

$$\begin{aligned} \mathcal{J}_{l,s}(f) := \int_0^\infty e(p)|f(p)|^2 dp - \frac{\gamma}{\pi} \int_0^\infty \int_0^\infty \overline{f(p)} k_{l,s}(p,p') f(p) dp dp' \\ + \gamma^2 \int_0^\infty dp \int_0^\infty dp' \overline{f(p)} \tilde{k}_{l,s}(p,p') f(p') \end{aligned} \tag{21}$$

and

$$k_{l,s}(p',p) = \frac{(e(p')+m) q_l(\frac{p'}{p})(e(p)+m) + p' q_{l+2s}(\frac{p'}{p}) p}{n(p')n(p)} \tag{22}$$

and

$$\begin{aligned} \tilde{k}_{l,s}(p,p') = \frac{1}{2\pi^2} \int_0^\infty dp'' \left(\frac{1}{e(p)+e(p'')} + \frac{1}{e(p'')+e(p')} \right) \\ \Big[q_{l+2s}(\frac{p}{p''}) q_{l+2s}(\frac{p'}{p''}) b(p) a(p'')^2 b(p') - q_{l+2s}(\frac{p}{p''}) q_l(\frac{p''}{p'}) b(p) b(p'') a(p'') a(p') \\ + q_l(\frac{p}{p''}) q_l(\frac{p'}{p''}) a(p) b(p'')^2 a(p') - q_l(\frac{p}{p''}) q_{l+2s}(\frac{p''}{p'}) a(p) b(p'') a(p'') b(p') \Big]. \end{aligned} \tag{23}$$

The Legendre functions of the second kind appear here for exactly the same reasons as in the treatment of the Schrödinger equation for the hydrogen atom in momentum space (Flügge [6], Problem 77).] To obtain (21), we also use that $(\omega_{\mathfrak{p}} \cdot \boldsymbol{\sigma})\Omega_{l,m,s}(\omega_{\mathfrak{p}}) = -\Omega_{l+2s,m,-s}(\omega_{\mathfrak{p}})$ (see, e.g., Greiner [8], p. 171, (12)). The operators $h_{l,s}$ defined by the sesquilinear form (21) via the equation $(f, h_{l,s}f) = \mathcal{J}_{l,s}(f)$ are reducing the operator H^{ext} on the corresponding angular momentum subspaces.

3 The Massless Operators and Their Positivity

To proceed, we will first consider the massless operators. The lower bound in the massive case will be a corollary of the positivity of the massless one. The energy in angular momentum channel (l, m, s) in the massless case can be read of from (14) and is given by

$$\mathcal{J}_{l,s}(f) := \mathcal{B}_{l,s}(f) + \gamma^2 \int_0^\infty dp \int_0^\infty dp' \overline{f(p)} \tilde{k}_{l,s}(p,p') f(p') \tag{24}$$

with

$$\begin{aligned} &\mathcal{B}_{l,s}(f) \\ &= \int_0^\infty p|f(p)|^2 dp - \frac{\gamma}{2\pi} \int_0^\infty dp \int_0^\infty dp' \overline{f(p)} \left(q_l(\frac{p}{p'}) + q_{l+2s}(\frac{p}{p'}) \right) f(p') \end{aligned} \tag{25}$$

and

$$\begin{aligned} \tilde{k}_{l,s}(p,p') = \frac{1}{8\pi^2} \int_0^\infty dp'' &\left(\frac{1}{p+p''} + \frac{1}{p''+p'} \right) \\ &\Big[q_{l+2s}(\frac{p}{p''}) q_{l+2s}(\frac{p''}{p'}) - q_{l+2s}(\frac{p}{p''}) q_l(\frac{p''}{p'}) \\ &\quad + q_l(\frac{p}{p''}) q_l(\frac{p''}{p'}) - q_l(\frac{p}{p''}) q_{l+2s}(\frac{p''}{p'}) \Big]. \end{aligned} \tag{26}$$

Using the simplifications of Appendix A, Formulae (57) and (59) we get

$$\begin{aligned} \tilde{k}_{l,s}(p,p') = \frac{1}{8\pi^2} \int_0^\infty \frac{dp''}{p''} &\Big(q_l(\frac{p}{p''}) q_l(\frac{p''}{p'}) - q_{l+2s}(\frac{p}{p''}) q_l(\frac{p''}{p'}) \\ &\quad - q_l(\frac{p}{p''}) q_{l+2s}(\frac{p''}{p'}) + q_{l+2s}(\frac{p}{p''}) q_{l+2s}(\frac{p''}{p'}) \Big). \end{aligned} \tag{27}$$

Since the operator in question is homogeneous of degree minus one we Mellin transform (see Appendix B) the quadratic form $\varepsilon_{l,s}$. If we write this form as a functional $\mathcal{J}^\#_{l,s}$ of the Mellin transformed radial functions $f^\#$, we get

$$\mathcal{J}^\#_{l,s}(f^\#) = \mathcal{B}^\#_{l,s}(f^\#) + \frac{1}{2}\left(\frac{\gamma}{2\pi}\right)^2 \int_{-\infty}^\infty dt \left| f^\#(t+i/2) \right|^2 F^\#(t) \tag{28}$$

where $\mathcal{B}^{\#}_{l,s}$ is the Brown-Ravenhall energy in angular momentum channel (l,s) in Mellin space, i.e.,

$$\mathcal{B}^{\#}_{l,s}(g) := \int_{-\infty}^{\infty} dt\, |g\,(t+i/2)|^2 \left[1 - \frac{\gamma}{2}(V_l(t) + V_{l+2s}(t))\right] \tag{29}$$

with

$$V_l(t) = \sqrt{\frac{2}{\pi}} q_l^{\#}\,(t-i/2) = \frac{1}{2}\left|\frac{\Gamma\left(\frac{l+1-it}{2}\right)}{\Gamma\left(\frac{l+2-it}{2}\right)}\right|^2 \tag{30}$$

(see Tix [19] [note also the factor $\sqrt{2/\pi}$ which is different from Tix's original formula]) and

$$F^{\#}(t) = \sqrt{2\pi}\left(q_l^{\#}(t-i/2) - q_{l+2s}^{\#}(t-i/2)\right)^2. \tag{31}$$

Formulae (29), (30), and (31) are obtained from (24), (25), and (27) using the fact that the occurring integrals can be read as a Mellin convolution which is turned by the Mellin transform into a product (see Appendix B, Formulae (61) and (63)).
Note that V_l is the Coulomb potential after Fourier transform, partial wave analysis, and Mellin transform.

3.1 Positivity of the Brown-Ravenhall Energy

To warm up for the minimization of $\mathcal{J}^{\#}_{l,s}$ we start with $\mathcal{B}^{\#}_{l,s}$ only. To this end we first note

Lemma 1. *We have*

$$V_{l+1}(t) \le V_{l+1}(0) \le V_l(0). \tag{32}$$

Note, that this is similar to Lemma 2 in [5].

Proof. First note that $q_0 \ge q_1 \ge q_2 ...$ which follows from the integral representation in [21], Chapter XV, Section 32, p. 334. This implies

$$\begin{aligned}\left|q_{l+1}^{\#}(t-i/2)\right| =& \frac{1}{\sqrt{2\pi}}\left|\int_0^{\infty} q_{l+1}(p) p^{-it}\frac{dp}{p}\right| \le \frac{1}{\sqrt{2\pi}}\int_0^{\infty} q_{l+1}(p)\frac{dp}{p} \\ \le& \frac{1}{\sqrt{2\pi}}\int_0^{\infty} q_l(p)\frac{dp}{p},\end{aligned} \tag{33}$$

which implies the lemma. ∎

Theorem 2. *For all $u \in C_0^{\infty}(\mathbb{R}^3) \otimes \mathbb{C}^2$ and $m = 0$ we have $\mathcal{B} \ge 0$ if and only if $\gamma \le \gamma_c^B$.*

Proof. Note that

$$V_l(t) + V_{l+2s}(t) \le V_0(0) + V_1(0) = \frac{\pi}{2} + \frac{2}{\pi}. \tag{34}$$

Thus

$$\mathcal{B}^{\#}_{l,s}(g) \ge \int_{-\infty}^{\infty} dt\, |g\,(t+i/2)|^2 \left(1 - \frac{\gamma}{2}\left(\frac{\pi}{2} + \frac{2}{\pi}\right)\right) \tag{35}$$

which implies that the energy is nonnegative if $\gamma \le 2/(\pi/2 + 2/\pi)$. □

We remark that Theorem 2 was proved by Evans et al. [5]. However, since g can be localized at $t = 0$, our method shows that Inequality (35) is sharp, i.e., the present proof shows also the sharpness of γ_c^B, a result of Hundertmark et al. [10] obtained by different means.
Since — according to Tix [19] — the difference of the massive and massless Brown-Ravenhall operators is bounded, Theorem 2 shows also that the energy in the massive case is bounded from below under the same condition on γ as in the massless case.

3.2 The Jansen-Hess Energy

We now wish to treat the full relativistic energy according to Jansen and Heß as given in (28) through (31). From these equations it is obvious that the energy is positive, if the coupling constant γ does not exceed γ_c^B, since the additional energy term is non-negative. However, as can be expected, the critical coupling constant is in fact bigger, i.e. we want to prove Theorem 1 in the massless case.

Lemma 2. *For all $u \in C_0^\infty(\mathbb{R}^3) \otimes \mathbb{C}^2$, $m = 0$, and $\gamma \le \gamma_c$ we have $(u, \mathcal{J}u) \ge 0$. Moreover, if $\gamma > \gamma_c$, then $\mathcal{J}$ is not bounded from bellow.*

Proof. We write the energy density in Mellin space as given in Equations (28) through (31) as

$$j_{l,s}(t) := 1 - \frac{\gamma}{2}(V_l(t) + V_{l+2s}(t)) + \frac{\gamma^2}{8}(V_l(t) - V_{l+2s}(t))^2. \tag{36}$$

As in the case of the Brown-Ravenhall energy we want to show that $j_{l,s}$ attains its minimum for $l = 0$ and $t = 0$.
First we note, that $j_{l,s}(t) = j_{l+2s,-s}(t)$ which means that we can restrict the following to $s = 1/2$, i.e., to $j_{l,1/2}$.
Next we show that it is monotone decreasing in l. For $\gamma \le 4/\pi$ we have

$$\begin{aligned} 0 &\le 1 - \frac{\gamma}{2}V_0(0) \le 1 - \frac{\gamma}{2}V_l(t) \le 1 - \frac{\gamma}{4}V_l(t) - \frac{\gamma}{4}V_{l+2}(t) \\ &\le 1 + \frac{\gamma}{2}V_{l+1}(t) - \frac{\gamma}{4}V_l(t) - \frac{\gamma}{4}V_{l+2}(t) \end{aligned} \tag{37}$$

where use successively (64), (32), Lemma 6 in Appendix C, and the positivity of the V_l. Inequality (37) is – after multiplication by $\gamma((V_l(t)-V_{l+2}(t))/2$ – identical with the desired monotonicity inequality

$$j_{l+1,1/2}(t) \geq j_{l,1/2}(t). \tag{38}$$

For later purposes we note that functions $j_{l,1/2}$ are symmetric about the origin. Next we will show that the energy density has its absolute minimum at the origin: to this end we simply show that the derivative of $j_{0,1/2}$ is nonnegative on the positive axis, if $\gamma \leq 2/(\pi/2+2/\pi)$ which is bigger than $4/\pi$. Since

$$|V_0(t)-V_1(t)| \leq \int_0^\infty (q_0(x)-q_1(x))\frac{dx}{x} = V_0(0)-V_1(0) = \frac{\pi}{2}-\frac{2}{\pi}$$

we have

$$-1+\frac{\gamma}{2}(V_0(t)-V_1(t)) \leq 0 \tag{39}$$

and obviously we have

$$-1-\frac{\gamma}{2}(V_0(t)-V_1(t)) \leq 0. \tag{40}$$

Thus the derivative of the energy $j_{0,1/2}$ is

$$\begin{aligned} j'_{0,1/2}(t) &= \frac{\gamma}{2}[-V_0'(t)-V_1'(t)+\frac{\gamma}{2}(V_0(t)-V_1(t))(V_0'(t)-V_1'(t))] \\ &= \frac{\gamma}{2}\{V_0'(t)[-1+\frac{\gamma}{2}(V_0(t)-V_1(t))]+V_1'(t)[-1-\frac{\gamma}{2}(V_0(t)-V_1(t))]\} \geq 0, \end{aligned} \tag{41}$$

since V_0 and V_1 are symmetrically decreasing about the origin (see Appendix C).

Finally, the polynomial

$$j_{0,1/2}(0) = 1-\frac{\gamma}{2}\left(\frac{\pi}{2}+\frac{2}{\pi}\right)+\frac{\gamma^2}{8}\left(\frac{\pi}{2}-\frac{2}{\pi}\right)^2$$

is nonnegative for $\gamma \leq \gamma_c$ as defined in the hypothesis. Thus, we have

$$j_{l,s}(t) \geq j_{0,1/2}(0) \geq 0.$$

□

4 Lower Bound on the Energy According to Jansen and Hess

To distinguish the massive and the massless expressions we will indicate in this section the dependence their on the mass m by a superscript m, if it seems appropriate.

The goal of this section is to show Theorem 1 for the massive case. We proceed by enunciating the following lemmata.

Lemma 3 (Tix [18, 20]). *For all* $u \in C_0^\infty(\mathbb{R}^3) \otimes \mathbb{C}^2$, $m \geq 0$, *and* $\gamma \leq \gamma_c^B$ *then*

$$\mathcal{B}(u) \geq m(1-\gamma).$$

Lemma 4 (Tix [19]). *The expression* $|\mathcal{B}^m(u) - \mathcal{B}^0(u)|$ *is bounded for* $u \in C_0^\infty(\mathbb{R}^3) \otimes \mathbb{C}^2$.

Lemma 5. *For all* $m \geq 0$ *and for all* $u \in C_0^\infty(\mathbb{R}^3) \otimes \mathbb{C}^2$ *we have*

$$|\tilde{K}^m(u) - \tilde{K}^0(u)| \leq md\|u\|^2 \tag{42}$$

where $d := \sqrt{2}(12 + 2^{5/2})$.

We note that the first part of Theorem 1 follows from Lemmata 2, 4, and 5. The third part is a consequence of Lemmata 3 and 5.

Proof. First we remark that

$$\sup\{|\mathcal{J}^m(u) - \mathcal{J}^0(u)| \mid \|u\| = 1\} = m \sup\{|\mathcal{J}^1(u) - \mathcal{J}^0(u)| \mid \|u\| = 1\}.$$

Then it is enough to start bounding $|(u, \tilde{K}^1 u) - (u, \tilde{K}^0 u)|$: By the mean value theorem we have

$$|\tilde{K}^1(\mathfrak{p}, \mathfrak{p}') - \tilde{K}^0(\mathfrak{p}, \mathfrak{p}')| \leq \lambda |D(\mu, \mathfrak{p}, \mathfrak{p}')| \tag{43}$$

for some $\mu \in (0, \lambda)$ where $\lambda \in (0,1)$ is a deformation parameter and $D(\mu, \mathfrak{p}, \mathfrak{p}')$ is the derivative of $\tilde{K}^\mu(\mathfrak{p}, \mathfrak{p}')$ with respect to μ. Computing the derivative yields

$$|D(\mu, \mathfrak{p}, \mathfrak{p}')| = \left| \int d\mathfrak{p}'' F(\mu, \mathfrak{p}, \mathfrak{p}'', \mathfrak{p}') \right| \tag{44}$$

with

$$\begin{aligned} &F(\mu, \mathfrak{p}, \mathfrak{p}'', \mathfrak{p}') \\ &\quad := \frac{1}{2(2\pi^2)^2} \left(\frac{\partial C}{\partial \lambda}(T_1 + ... + T_4) + C \frac{\partial (T_1 + ... + T_4)}{\partial \lambda} \right)(\mu, \mathfrak{p}, \mathfrak{p}'', \mathfrak{p}') \end{aligned} \tag{45}$$

where C and $T_1, ..., T_4$ are defined right below (14). Note that $a(p)^2 \leq 1$ and $b(p)^2 \leq 1/2$, i.e., by the definition $T_1, ..., T_4 \leq 1/2$. Furthermore we note that

$$\begin{aligned} \frac{\partial C}{\partial \lambda} = \frac{-\lambda}{E(\mathfrak{p}'')} &\frac{1}{|\mathfrak{p} - \mathfrak{p}''|^2 |\mathfrak{p}'' - \mathfrak{p}'|^2} \\ &\left(\frac{1}{(E(\mathfrak{p}) + E(\mathfrak{p}''))E(\mathfrak{p})} + \frac{1}{(E(\mathfrak{p}'') + E(\mathfrak{p}'))E(\mathfrak{p}')} \right). \end{aligned} \tag{46}$$

First we treat $\frac{\partial C}{\partial \lambda}(T_1 + ... + T_4)$. We get using the above estimates on T_1 through T_4 and (46)

$$\left| \frac{\partial C}{\partial \lambda}(T_1 + ... + T_4)(\mu, \mathfrak{p}, \mathfrak{p}'', \mathfrak{p}') \right| \leq \frac{2}{p''} \frac{1}{|\mathfrak{p} - \mathfrak{p}''|^2 |\mathfrak{p}'' - \mathfrak{p}'|^2} \left(\frac{1}{p + p''} + \frac{1}{p'' + p'} \right) \tag{47}$$

Next we treat $C\frac{\partial(T_1+\ldots+T_4)}{\partial\lambda}$. To this end we note

$$\left|\frac{\partial a}{\partial\lambda}(p)\right| = \frac{p^2}{4E(\mathfrak{p})^3}\sqrt{\frac{2E(\mathfrak{p})}{E(\mathfrak{p})+\lambda}} \leq \frac{\sqrt{2}}{4p} \tag{48}$$

and

$$\left|\frac{\partial b}{\partial\lambda}(p)\right| = p\sqrt{\frac{E(\mathfrak{p})+\lambda}{8E(\mathfrak{p})^5}} \leq \frac{1}{2p}. \tag{49}$$

Thus

$$\begin{aligned}&|C\frac{\partial(T_1+\ldots+T_4)}{\partial\lambda}(\mu,\mathfrak{p},\mathfrak{p}'',\mathfrak{p}')|\\ &\qquad\leq \frac{3}{2^{3/2}}\frac{1}{|\mathfrak{p}-\mathfrak{p}''|^2|\mathfrak{p}''-\mathfrak{p}'|^2}\left(\frac{1}{p+p''}+\frac{1}{p''+p'}\right)\left(\frac{1}{p}+\frac{2}{p''}+\frac{1}{p'}\right).\end{aligned} \tag{50}$$

We now bound the integral operator $\tilde{K}^1 - \tilde{K}^0$ by a multiplication operator: First pick $\alpha \in \mathbb{R}$. Then we have — using the symmetry of $F(\mu,\mathfrak{p},\mathfrak{p}'',\mathfrak{p}')$ in $\mathfrak{p}$ and $\mathfrak{p}'$ for fixed $\mathfrak{p}''$ —

$$\begin{aligned}|(u,(\tilde{K}^1-\tilde{K}^0)u)| &= \left|\int d\mathfrak{p}''\int d\mathfrak{p}\int d\mathfrak{p}' u(\mathfrak{p})^* F(\mu,\mathfrak{p},\mathfrak{p}'',\mathfrak{p}')u(\mathfrak{p}')\right|\\ &\leq \int d\mathfrak{p}''\int d\mathfrak{p}|u(\mathfrak{p})|^2\int d\mathfrak{p}'\left|\frac{p}{p'}\right|^\alpha |F(\mu,\mathfrak{p},\mathfrak{p}'',\mathfrak{p}')|\end{aligned} \tag{51}$$

where we used the Schwarz inequality in the measure $d\mathfrak{p}d\mathfrak{p}'$ in the last step for fixed $\mathfrak{p}''$. Now using the estimates (47) and (50) and collecting similar terms yields

$$\begin{aligned}|(u,(\tilde{K}^1-\tilde{K}^0)u)| &\leq \frac{1}{2^{5/2}(2\pi^2)^2}\int d\mathfrak{p}|u(\mathfrak{p})|^2\int d\mathfrak{p}''\int d\mathfrak{p}'\left|\frac{p}{p'}\right|^\alpha\\ &\quad\frac{1}{|\mathfrak{p}-\mathfrak{p}''|^2|\mathfrak{p}''-\mathfrak{p}'|^2}\left(\frac{1}{p+p''}+\frac{1}{p''+p'}\right)\left(\frac{3}{p}+\frac{2^{\frac{5}{2}}+6}{p''}+\frac{3}{p'}\right)\end{aligned} \tag{52}$$

where we claim the last line to be bounded by $32(12+2^{5/2})\pi^4$, i.e.,

$$|(u,(\tilde{K}^1-\tilde{K}^0)u)| \leq \sqrt{2}\int d\mathfrak{p}|u(\mathfrak{p})|^2(12+2^{5/2}). \tag{53}$$

To show the above bound we break the integral into three parts

$$\begin{aligned}I &:= \int d\mathfrak{p}''\int d\mathfrak{p}'\left(\frac{p}{p'}\right)^\alpha\frac{1}{|\mathfrak{p}-\mathfrak{p}''|^2|\mathfrak{p}''-\mathfrak{p}'|^2}\left(\frac{1}{p+p''}+\frac{1}{p''+p'}\right)\frac{1}{p},\\ I'' &:= \int d\mathfrak{p}''\int d\mathfrak{p}'\left(\frac{p}{p'}\right)^\alpha\frac{1}{|\mathfrak{p}-\mathfrak{p}''|^2|\mathfrak{p}''-\mathfrak{p}'|^2}\left(\frac{1}{p+p''}+\frac{1}{p''+p'}\right)\frac{1}{p''},\\ I' &:= \int d\mathfrak{p}''\int d\mathfrak{p}'\left(\frac{p}{p'}\right)^\alpha\frac{1}{|\mathfrak{p}-\mathfrak{p}''|^2|\mathfrak{p}''-\mathfrak{p}'|^2}\left(\frac{1}{p+p''}+\frac{1}{p''+p'}\right)\frac{1}{p'}.\end{aligned} \tag{54}$$

We will also use the following integral (see [13], p.124)

$$\Upsilon(\beta) := \int_{\mathbb{R}^3} d\mathfrak{p} \frac{1}{|\mathfrak{e}-\mathfrak{p}|^2} \frac{1}{p^\beta} = \pi^2 \frac{\Gamma(\frac{\beta-1}{2})\Gamma(1-\frac{\beta-1}{2})}{\Gamma(2-\frac{\beta}{2})\Gamma(\frac{\beta}{2})}, \tag{55}$$

where $\mathfrak{e}$ is an (arbitrary) unit vector in $\mathbb{R}^3$ and $\beta \in (1,3)$. We observe that each of the integrals in (54) do not depend on the value of p (what becomes evident after substitution of $\mathfrak{p}' \to p\mathfrak{p}'$ and $\mathfrak{p}'' \to p\mathfrak{p}''$). So picking $p = 1$ and doing $\mathfrak{p}' \to p''\mathfrak{p}'$ in each integral in (54) we find

$$\begin{aligned}
I &= \int d\mathfrak{p}'' \int d\mathfrak{p}' \left(\frac{1}{p'p''}\right)^\alpha \frac{1}{|\mathfrak{u}-\mathfrak{p}''|^2|\mathfrak{u}''-\mathfrak{p}'|^2} \left\{\frac{p''}{1+p''} + \frac{1}{1+p'}\right\} \leq 2\Upsilon(\alpha)^2, \\
I'' &= \int d\mathfrak{p}'' \int d\mathfrak{p}' \left(\frac{1}{p'p''}\right)^\alpha \frac{1}{|\mathfrak{u}-\mathfrak{p}''|^2|\mathfrak{u}''-\mathfrak{p}'|^2} \left\{\frac{1}{1+p''} + \frac{1}{p''(1+p')}\right\} \\
&\quad \leq \Upsilon(\alpha)^2 + \Upsilon(\alpha)\Upsilon(\alpha+1), \\
I' &= \int d\mathfrak{p}'' \int d\mathfrak{p}' \left(\frac{1}{p'p''}\right)^\alpha \frac{1}{|\mathfrak{u}-\mathfrak{p}''|^2|\mathfrak{u}''-\mathfrak{p}'|^2} \left\{\frac{1}{p'(1+p'')} + \frac{1}{p'p''(1+p')}\right\} \\
&\quad \leq 2\Upsilon(\alpha+1)^2,
\end{aligned} \tag{56}$$

We choose $\alpha = 3/2$ and using (55) we obtain the same bound for each integral, namely $32\pi^4$. Equation (53) proves Lemma 5 and follows by using the latter bound in (52). $\square$

A Some Useful Integral Identities

Suppose $f(x) = f(1/x)$ and suppose $f(x)/(1+x)$ is integrable on $(0,\infty)$. Then

$$\int_0^\infty \frac{f(x)}{1+x} dx = \int_0^1 \frac{f(x)}{x} dx = \frac{1}{2}\int_0^\infty \frac{f(x)}{x} dx \tag{57}$$

To show (57) we split the first integral

$$\begin{aligned}
&\int_0^1 \frac{dx}{x} f(x) \frac{x}{1+x} + \int_1^\infty \frac{dx}{x} f(x) \frac{x}{1+x} = \int_0^1 \frac{dx}{x} f(x) \\
&\quad = \int_0^\infty \frac{dx}{x} f(x) - \int_1^\infty \frac{dx}{x} f(x) = \int_0^\infty \frac{dx}{x} f(x) - \int_0^1 \frac{dx}{x} f(x).
\end{aligned} \tag{58}$$

where we used the invariance under inversion of f for the first and third equality. Next we wish to simplify the kernel $j_{l,s}$. To this end we use again the abbreviation $q_l(x) := Q_l\left(\frac12\left(x+\frac1x\right)\right)$ as in (26). We claim

$$\begin{aligned}
&I(p,p')\\
&:= \int_0^\infty \frac{dp''}{p''}\left(q_l(\frac{p}{p''})q_m(\frac{p''}{p'}) + q_m(\frac{p}{p''})q_l(\frac{p''}{p'})\right)\left(\frac{p''}{p+p''}+\frac{p''}{p''+p'}\right) \qquad (59)\\
&= \int_0^\infty \frac{dp''}{p''}\left(q_l(\frac{p}{p''})q_m(\frac{p''}{p'}) + q_m(\frac{p}{p''})q_l(\frac{p''}{p'})\right)
\end{aligned}$$

To prove this we take the integral with the complete first factor times the first summand of the second factor –we name I_1– and the integral over the complete first factor times the second summand of the second factor, I_2. In I_1 we substitute $p'' \to pp''$ whereas in I_2 we substitute $p'' \to p'p''$. This yields using (57)

$$\begin{aligned}
I(p,p') = I_1 + I_2 = \frac12 \int_0^\infty \frac{dp''}{p''}\Bigg[q_l\left(p''\right)q_m\left(\frac{p''p}{p'}\right)\\
+ q_m(p'')q_l\left(p''\frac{p}{p'}\right) + q_l\left(\frac{p''p'}{p}\right)q_m(p'') + q_m\left(p''\frac{p'}{p}\right)q_l(p'')\Bigg]. \qquad (60)
\end{aligned}$$

Undoing the substitutions yields the desired result.

B The Mellin Transform

The Mellin transform is a unitary map from $L^2(\mathbb{R}^+)$ to $L^2(\mathbb{R})$ given by the formula

$$f^\#(s) := \frac{1}{\sqrt{2\pi}}\int_0^\infty f(p)p^{-\frac12 - is}dp.$$

The Mellin convolution of two function f and g is defined as

$$(f \star g)(p) = \int_0^\infty f\left(\frac{p}{q}\right)g(q)\frac{dq}{q}. \qquad (61)$$

If $f \in C_0^\infty(\mathbb{R}^+)$, then $f^\#$ extends to an entire function, and we have

$$(p^\alpha f)^\#(s) = f^\#(s+i\alpha). \qquad (62)$$

We also have

$$(f\star g)^\#(s) = \sqrt{2\pi}f^\#(s)g^\#(s). \qquad (63)$$

Both, (62) and (63), can be verified by direct computation.

C Some Properties Related to the Partial Wave Analysis of the Coulomb potential in Mellin Space

We first remark the follow property on the difference of V_l and V_{l+2}.

Lemma 6. *For $l = 0, 1, 2, \ldots$ and $t \in \mathbb{R}$ we have $V_{l+2}(t) < V_l(t)$.*

Proof. From the definition of V_l in (30) we see that the claim is equivalent to

$$\left|\frac{\Gamma\left(\frac{l+1-it}{2}\right)}{\Gamma\left(\frac{l+2-it}{2}\right)}\right|^2 > \left|\frac{\Gamma\left(\frac{l+3-it}{2}\right)}{\Gamma\left(\frac{l+4-it}{2}\right)}\right|^2.$$

This, however, can be easily verified using the functional equation $\Gamma(x+1) = x\Gamma(x)$ of the Gamma function in the numerator and denominator of the right hand side with $x = (l+1-it)/2$ and $x = (l+2-it)/2$. □

From the definition of the V_l and from Formulae 8.332.2 and 8.333.3 in [7] one finds V_0 and V_1 in terms of the hyperbolic tangent and cotangent:

$$V_0(t) = \frac{\mathfrak{Tg}(\pi t/2)}{t} \tag{64}$$

$$V_1(t) = \frac{t}{1+t^2}\mathfrak{Ctg}(\pi t/2). \tag{65}$$

Moreover, both of these functions are decreasing symmetricly about the origin.

Acknowledgment: We thank Dr. Doris Jakubaßa-Amundsen for careful reading of the manuscript and pointing out several mistakes. The work has been partially supported by the European Union through its Training, Research, and Mobility program, grant FMRX-CT 96-0001960001, the Volkswagen Foundation through a cooperation grant, and the Deutsche Forschungsgemeinschaft (Schwerpunktprogramm 464 "Theorie relativistischer Effekte in der Chemie und Physik schwerer Elemente"). E. S. acknowledges partial support of the project by FONDECYT (Chile, project 2000004), CONICYT (Chile), and Fundación Andes for support through a doctoral fellowship.

References

[1] Hans A. Bethe and Edwin E. Salpeter. Quantum mechanics of one- and two-electron atoms. In S. Flügge, editor, *Handbuch der Physik, XXXV*, pages 88–436. Springer, Berlin, 1 edition, 1957.

[2] G. E. Brown and D. G. Ravenhall. On the interaction of two electrons. *Proc. Roy. Soc. London Ser. A.*, 208:552–559, 1951.

[3] Raymond Brummelhuis, Norbert Röhrl, and Heinz Siedentop. Stability of the relativistic electron-positron field of atoms in Hartree-Fock approximation: Heavy elements. *Doc. Math., J. DMV*, 6:1–8, 2001.

[4] Marvin Douglas and Norman M. Kroll. Quantum electrodynamical corrections to the fine structure of helium. *Annals of Physics*, 82:89–155, 1974.

[5] William Desmond Evans, Peter Perry, and Heinz Siedentop. The spectrum of relativistic one-electron atoms according to Bethe and Salpeter. *Commun. Math. Phys.*, 178(3):733–746, July 1996.

[6] Siegfried Flügge. *Practical Quantum Mechanics I*, volume 177 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1 edition, 1982.

[7] I.S. Gradshteyn and I.M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, 4th edition, 1980.

[8] Walter Greiner. *Relativistic Quantum Mechanics*, volume 3 of *Theoretical Physics – Text and Excercise Books*. Springer, Berlin, 1 edition, 1990.

[9] G. Hardekopf and J. Sucher. Relativistic wave equations in momentum space. *Phys. Rev. A*, 30(2):703–711, August 1984.

[10] Dirk Hundertmark, Norbert Röhrl, and Heinz Siedentop. The sharp bound on the stability of the relativistic electron-positron field in Hartree-Fock approximation. *Commun. Math. Phys.*, 211(3):629–642, May 2000.

[11] Georg Jansen and Bernd A. Heß. Revision of the Douglas-Kroll transformation. *Physical Review A*, 39(11):6016–6017, June 1989.

[12] V. Kellö, A. J. Sadlej, and B. A. Hess. Relativistic effects on electric properties of many-electron systems in spin-averaged Douglas-Kroll and Pauli approximations. *Journal of Chemical Physics*, 105(5):1995–2003, August 1996.

[13] Elliott H. Lieb and Michael Loss. *Analysis*. Number 14 in Graduate Studies in Mathematics. American Mathematical Society, Providence, 1 edition, 1996.

[14] Albert Messiah. *Mécanique Quantique*, volume 1. Dunod, Paris, 2 edition, 1969.

[15] G. Nenciu. Self-adjointness and invariance of the essential spectrum for Dirac operators defined as quadratic forms. *Commun. Math. Phys.*, 48:235–247, 1976.

[16] Irene A. Stegun. Legendre functions. In Milton Abramowitz and Irene A. Stegun, editors, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, chapter 8, pages 331–353. Dover Publications, New York, 1965.

[17] J. Sucher. Foundations of the relativistic theory of many-electron atoms. *Phys. Rev. A*, 22(2):348–362, August 1980.

[18] C. Tix. Lower bound for the ground state energy of the no-pair Hamiltonian. *Phys. Lett. B*, 405(3-4):293–296, 1997.

[19] C. Tix. Self-adjointness and spectral properties of a pseudo-relativistic Hamiltonian due to Brown and Ravenhall. *Preprint*, mp-arc: 97-441, 1997.

[20] C. Tix. Strict positivity of a relativistic Hamiltonian due to Brown and Ravenhall. *Bull. London Math. Soc.*, 30(3):283–290, 1998.

[21] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis; An Introduction to the General Theory of Infinite Processes and of Analytic Functions, with an Account of the Principal Transcendental Functions.* Cambridge University Press, Cambridge, 4 edition, 1927.

Raymond Brummelhuis
Birkbeck College
University of London
School of Economics,
Mathematics and Statistics
Gresse Street
London W1T 1LL
United Kingdom
r.brummelhuis@statistics.bbk.ac.uk

Heinz Siedentop
Mathematik
Theresienstr. 39
80333 München
Germany
h.s@lmu.de

Edgardo Stockmeyer
Pontificia Universidad
Católica de Chile
Departamento de Física
Casilla 306
Santiago 22
Chile
estockme@maxwell.fis.puc.cl

Families of p-Divisible Groups with Constant Newton Polygon

Frans Oort and Thomas Zink

Received: May 2, 2002
Revised: September 19, 2002

Communicated by Peter Schneider

Abstract. Let X be a p-divisible group with constant Newton polygon over a normal Noetherian scheme S. We prove that there exists an isogeny $X \to Y$ such that Y admits a slope filtration. In case S is regular this was proved by N. Katz for dim $S = 1$ and by T. Zink for dim $S \geq 1$.

2000 Mathematics Subject Classification: 14L05, 14F30

Introduction

In this paper we work over base fields, and over base schemes over $\mathbb{F}_p$, i.e. we work entirely in characteristic p. We study p-divisible groups X over a base scheme S (and, colloquially, a p-divisible group over a base scheme of positive dimension will be called a "family of p-divisible groups"), such that the Newton polygon of a fiber X_s is independent of $s \in S$. We call X a p-divisible group with constant Newton polygon.

A p-divisible group over a field has a slope filtration, see [Z1], Corollary 13; for the definition of a slope filtration, see Definition 1.1. Over a base of positive dimension a slope filtration can only exist if the Newton polygon is constant. In Example 4.1 we show that even in this case there are p-divisible groups which do not admit a slope filtration.

The main result of this paper has as a corollary that *for a p-divisible group with constant Newton polygon over a normal base up to isogeny a slope filtration does exist*, see Corollary 2.2.

We have access to this kind of questions by the definition of a *completely slope divisible* p-divisible group, see Definition 1.2, which implies a structure finer than a slope filtration. The main theorem of this paper, Theorem 2.1, says that over a *normal* base this structure on a p-divisible group exists up to isogeny.

In [Z1], Theorem 7, this was shown to be true over a regular base. In 4.2 we show that without the condition "normal" the conclusion of the theorem does not hold.

Here is a motivation for this kind definition and of results:

- A p-divisible group over an algebraically closed field is isogeneous with a p-divisible group which can be defined over a finite field.
- A p-divisible group over an algebraically closed field is completely slope divisible, if and only if it is isomorphic with a direct sum of isoclinic p-divisible groups which can be defined over a finite field, see 1.5.

We see that a completely slope divisible p-divisible group comes "as close as possible" to a constant one, in fact up to extensions of p-divisible groups annihilated by an inseparable extension of the base, and up to monodromy.

From Theorem 2.1 we deduce constancy results which generalize results of Katz [K] and more recently of de Jong and Oort [JO]. In particular we prove, Corollary 3.4 below:

Let R be a Henselian local ring with residue field k. Let h be a natural number. Then there exists a constant c with the following property. Let X and Y be isoclinic p-divisible groups over $S = \operatorname{Spec} R$ whose heights are smaller than h. Let $\psi : X_k \to Y_k$ be a homomorphism. Then $p^c\psi$ lifts to a homomorphism $X \to Y$.

1 Completely slope divisible p-divisible groups

In this section we present basic definitions and methods already used in the introduction.

Let S be a scheme over $\mathbb{F}_p$. Let $\mathrm{Frob} : S \to S$ be the absolute Frobenius morphism. For a scheme G/S we write:

$$G^{(p)} \quad = \quad G \times_{S,\mathrm{Frob}} S.$$

We denote by $\mathrm{Fr} = \mathrm{Fr}_G : G \to G^{(p)}$ the Frobenius morphism relative to S. If G is a finite locally free commutative group scheme we write $\mathrm{Ver} = \mathrm{Ver}_G : G^{(p)} \to G$ for the "Verschiebung".

Let X be a p-divisible group over S. We denote by $X(n)$ the kernel of the multiplication by $p^n : X \to X$. This is a finite, locally free group scheme which has rank p^{nh} if X is of height h.

Let $s = \operatorname{Spec} k$ the spectrum of a field of characteristic p. Let X be a p-divisible group over s. Let $\lambda \geq 0$ be a rational number. We call X isoclinic of slope λ, if there exists integers $r \geq 0$, $s > 0$ such that $\lambda = r/s$, and a p-divisible group Y over S, which is isogeneous to X such that

$$p^{-r}\,\mathrm{Fr}^s : Y \to Y^{(p^s)}$$

is an isomorphism.

A p-divisible group X over S is called isoclinic of slope λ, if for each point $s \in S$ the group X_s is isoclinic of slope λ.

1.1 Definition. *Let X/S be a p-divisible group over a scheme S. A filtration*

$$0 = X_0 \subset X_1 \subset X_2 \subset \ldots \subset X_m = X$$

consisting of p-divisible groups contained in X is called a slope filtration *of X if there exists rational numbers $\lambda_1, \ldots, \lambda_m$ satisfying $1 \geq \lambda_1 > \ldots > \lambda_m \geq 0$ such that every subquotient X_i/X_{i-1}, $1 < i \leq m$, is isoclinic of slope λ_i.*

A p-divisible group X over a field admits a slope filtration, see [Z1], Corollary 13. The slopes λ_i and the heights of X_i/X_{i-1} depend only on X. The height of X_i/X_{i-1} is called the multiplicity of λ_i.

Over connected base scheme S of positive dimension a slope filtration of X can only exist if the slopes of X_s and their multiplicities are independent of $s \in S$. In this case we say that X is a family of p-divisible groups with constant Newton polygon. Even if the Newton polygon is constant a slope filtration in general does not exist, see Example 4.1 below.

1.2 Definition. *Let $s > 0$ and $r_1, \ldots, r_m$ be integers such that $s \geq r_1 > r_2 > \ldots > r_m \geq 0$. A p-divisible group Y over a scheme S is said to be* completely slope divisible *with respect to these integers if Y has a filtration by p-divisible subgroups:*

$$0 = Y_0 \subset Y_1 \subset \ldots \subset Y_m = Y$$

such that the following properties hold:

- *The quasi-isogenies*
$$p^{-r_i}\,\mathrm{Fr}^s : Y_i \to Y_i^{(p^s)}$$
are isogenies for $i = 1, \ldots, m$.
- *The induced morphisms:*
$$p^{-r_i}\,\mathrm{Fr}^s : Y_i/Y_{i-1} \to (Y_i/Y_{i-1})^{(p^s)}$$
are isomorphisms.

Note that the last condition implies that Y_i/Y_{i-1} is isoclinic of slope $\lambda_i := r_i/s$. A filtration described in this definition is a slope filtration in the sense of the previous definition.

Remark. Note that we do not require s and r_i to be relatively prime. If Y is as in the definition, and $t \in \mathbb{Z}_{>0}$, it is also completely slope divisible with respect to $t{\cdot}s \geq t{\cdot}r_1 > t{\cdot}r_2 > \ldots > t{\cdot}r_m \geq 0$.

We note that the filtration Y_i of Y is uniquely determined, if it exists. Indeed, consider the isogeny $\Phi = p^{-r_m}\,\mathrm{Fr}^s : Y \to Y^{(p^s)}$. Then Y_m/Y_{m-1} is necessarily the Φ-étale part Y^Φ of Y, see [Z1] respectively 1.6 below. This proves the uniqueness by induction.

We will say that a p-divisible group is completely slope divisible if it is completely slope divisible with respect to some set integers and inequalities $s \geq r_1 > r_2 > \ldots > r_m \geq 0$.

REMARK. A p-divisible group Y over a field K is completely slope divisible iff $Y \otimes_K L$ is completely slope divisible for some field $L \supset K$. - PROOF. The slope filtration on Y/K exists. We have $\mathrm{Ker}(p^r_Y) \subset \mathrm{Ker}(\mathrm{Fr}^s_Y)$ iff $\mathrm{Ker}(p^r_{Y_L}) \subset \mathrm{Ker}(\mathrm{Fr}^s_{Y_L})$, and the same for equalities. This proves that the conditions in the definition for completely slope divisibility hold over K iff they hold over $L \supset K$.

1.3 PROPOSITION. *Let Y be a completely slope divisible p-divisible group over a perfect scheme S. Then Y is isomorphic to a direct sum of isoclinic and completely slope divisible p-divisible groups.*

PROOF. With the notation of Definition 1.2 we set $\Phi = p^{-r_m}\,\mathrm{Fr}^s$. Let $Y(n) = \mathrm{Spec}\,\mathcal{A}(n)$ and let $Y(n)^\Phi = \mathrm{Spec}\,\mathcal{L}(n)$ be the Φ-étale part (see Corollary 1.7 below). Then Φ induces a Frob^s-linear endomorphism Φ^* of $\mathcal{A}(n)$ and $\mathcal{L}(n)$ which is by definition bijective on $\mathcal{L}(n)$ and nilpotent on the quotient $\mathcal{A}(n)/\mathcal{L}(n)$. One verifies (compare [Z1], page 84) that there is a unique Φ^*-equivariant section of the inclusion $\mathcal{L}(n) \to \mathcal{A}(n)$. This shows that Y_m/Y_{m-1} is a direct factor of $Y = Y_m$. The result follows by induction on m. *Q.E.D.*

Although not needed, we give a characterization of completely slope divisible p-divisible groups over a field. If X is a p-divisible group over a field K, and $k \supset K$ is an algebraic closure, then X is completely slope divisible if and only if X_k is completely slope divisible; hence it suffices to give a characterization over an algebraically closed field.

Convention: We will work with the covariant Dieudonné module of a p-divisible group over a perfect field ([Z2], [Me]). We write V, respectively F for Verschiebung, respectively Frobenius on Dieudonné modules. Let K be a perfect field, and let $W(K)$ be its ring of Witt vectors. A Dieudonné module M over K is the Dieudonné module of an isoclinic p-divisible group of slope r/s, iff there exists a $W(K)$-submodule $M' \subset M$ such that M/M' is annihilated by a power of p, and such that $p^{-r}V^s(M') = M'$.

For later use we introduce the p-divisible group $G_{m,n}$ for coprime positive integers m and n. Its Dieudonné module is generated by one element, which is stable under $F^m - V^n$. $G_{m,n}$ is isoclinic of slope $\lambda = m/(m+n)$ (in the terminology of this paper). The height of $G_{m,n}$ is $h = m+n$, and this p-divisible group is completely slope divisible with respect to $h = m+n \geq m$. The group $G_{m,n}$ has dimension m, and its Serre dual has dimension n.

We have $G_{1,0} = \mu_{p^\infty}$, and $G_{0,1} = \underline{\mathbb{Q}_p/\mathbb{Z}_p}$.

1.4 Proposition. *Let k be an algebraically closed field. An isoclinic p-divisible group Y over k is completely slope divisible iff it can be defined over a finite field; i.e. iff there exists a p-divisible group Y' over some $\mathbb{F}_q$ and an isomorphism $Y \cong Y' \otimes_{\mathbb{F}_q} k$.*

Proof. Assume that Y is slope divisible with respect to $s \geq r \geq 0$. Let M be the covariant Dieudonné module of Y. We set $\Phi = p^{-r}V^s$. By assumption this is a semilinear automorphism of M. By a theorem of Dieudonné (see 1.6 below) M has a basis of Φ-invariant vectors. Hence $M = M_0 \otimes_{W(\mathbb{F}_{p^s})} W(k)$ where $M_0 \subset M$ is the subgroup of Φ-invariant vectors. Then M_0 is the Dieudonné module of a p-divisible group over $\mathbb{F}_{p^s}$ such that $Y \cong Y' \otimes_{\mathbb{F}_q} k$.

Conversely assume that Y is isoclinic over a finite field k of slope r/s. Let M be the Dieudonné module of Y. By definition there is a finitely generated $W(k)$-submodule $M' \subset M \otimes \mathbb{Q}$ such that $p^m M' \subset M \subset M'$ for some natural number m, and such that $p^{-r}V^s(M') = M'$. Then $\Phi = p^{-r}V^s$ is an automorphism of the finite set $M'/p^m M'$. Hence some power Φ^t acts trivially on this set. This implies that $\Phi^t(M) = M$. We obtain that $p^{-rt}F^{st}$ induces an automorphism of Y. Therefore Y is completely slope divisible. *Q.E.D.*

1.5 Corollary. *Let Y be a p-divisible group over an algebraically closed field k. This p-divisible group is completely slope divisible iff $Y \cong \oplus Y_i$ such that every Y_i is isoclinic, and can be defined over a finite field.*

Proof. Indeed this follows from 1.3 and 1.4. *Q.E.D.*

1.6 The Φ-étale part.

For further use, we recall a notion explained and used in [Z1], Section 2. This method goes back to Hasse and Witt, see [HW], and to Dieudonné, see [D], Proposition 5 on page 233. It can be formulated and proved for locally free sheaves, and it has a corollary for finite flat group schemes.

Let V be a finite dimensional vector space over a separably closed field k of characteristic p. Let $f : V \to V$ be a Frob^s-linear endomorphism. The set $C_V = \{x \in V \mid f(x) = x\}$ is a vector space over $\mathbb{F}_{p^s}$. Then $V^f = C_V \otimes_{\mathbb{F}_{p^s}} k$ is a subspace of V. The endomorphism f acts as a Frob^s-linear automorphism on V^f and acts nilpotently on the quotient V/V^f. This follows essentially form Dieudonné loc.cit.. Moreover if k is any field of characteristic p we have still unique exact sequence of k-vector spaces

$$0 \to V^f \to V \to V/V^f \to 0$$

such that f acts as a a Frob^s-linear automorphism on V^f and acts nilpotently on the quotient V/V^f.

This can be applied in the following situation: Let S be a scheme over $\mathbb{F}_p$. Let G be a locally free group scheme over S endowed with a homomorphism

$$\Phi : G \to G^{(p^s)}.$$

In case $S = \mathrm{Spec}(K)$, where K is a field we consider the affine algebra A of G. The Φ induces a Frob^s-linear endomorphism $f : A \to A$. The vector subspace A^f inherits the structure of a bigebra. We obtain a finite group scheme $G^\Phi = \mathrm{Spec}\, A^f$, which is called the Φ-étale part of G. Moreover we have an exact sequence of group schemes:

$$0 \to G^{\Phi-\mathrm{nil}} \longrightarrow G \longrightarrow G^\Phi \to 0.$$

The morphism Φ induces an isomorphism $G^\Phi \to (G^\Phi)^{(p^s)}$, and acts nilpotently on the kernel $G^{\Phi-\mathrm{nil}}$.

Let now S be an arbitrary scheme over $\mathbb{F}_p$. Then we can expect a Φ-étale part only in the case where the rank of G_x^Φ is independent of $x \in S$:

1.7 Corollary. *Let $G \to S$ be a finite, locally free group scheme; let $\Phi : G \to G^{(p^t)}$ be a homomorphism. Assume that the function*

$$S \to \mathbb{Z}, \quad \text{defined by} \quad x \mapsto \mathrm{rank}((G_x)^{\Phi_x})$$

is constant. Then there exists an exact sequence

$$0 \to G^{\Phi-\mathrm{nil}} \longrightarrow G \longrightarrow G^\Phi \to 0$$

such that Φ is nilpotent on $G^{\Phi-\mathrm{nil}}$ and an isomorphism on the Φ-étale part G^Φ.

The prove is based on another proposition which we use in section 3. Let $\mathcal{M}$ be a finitely generated, locally free $\mathcal{O}_S$-module. Let $t \in \mathbb{Z}_{>0}$

$$f : \mathcal{O}_S \otimes_{\mathrm{Frob}_S^t, S} \mathcal{M} = \mathcal{M}^{(p^t)} \to \mathcal{M}$$

be a morphism of $\mathcal{O}_S$-modules. To every morphism $T \to S$ we associate

$$C_\mathcal{M}(T) = \{x \in \Gamma(T, \mathcal{M}_T) \mid f(1 \otimes x) = x\}.$$

1.8 Proposition *(*see [Z1], Proposition 3*). The functor $C_\mathcal{M}$ is represented by a scheme that is étale and affine over S. Suppose S to be connected; the scheme $C_\mathcal{M}$ is finite over S iff for each geometric point $\eta \to S$ the cardinality of $C_\mathcal{M}(\eta)$ is the same.*

Let X be a p-divisible group over a field K. Suppose $\Phi : X \to X^{(p^t)}$ is a homomorphism. Then the Φ-étale part X^Φ is the inductive limit of $X(n)^\Phi$. This is a p-divisible group.

1.9 Corollary. *Let X be a p-divisible group over S. Assume that for each geometric point $\eta \to S$ the height of the Φ-étale part of X_η is the same. Then a p-divisible group X^Φ exists and commutes with arbitrary base change. There is an exact sequence of p-divisible groups:*

$$0 \to X^{\Phi-\mathrm{nil}} \to X \to X^\Phi \to 0.$$

The following proposition can be deduced from proposition 1.8.

1.10 Corollary. *Assume that $G \to S$ is a finite, locally free group scheme over a connected base scheme S. Let $\Phi : G \xrightarrow{\sim} G^{(q)}$, $q = p^s$ be an isomorphism. Then there exists a finite étale morphism $T \to S$, and a morphism $T \to \operatorname{Spec}(\mathbb{F}_q)$, such that G_T is obtained by base change from a finite group scheme H over $\mathbb{F}_q$:*

$$H \otimes_{\operatorname{Spec}\mathbb{F}_q} T \xrightarrow{\sim} G_T.$$

Moreover Φ is induced from the identity on H.

Remark. If S is a scheme over $\bar{\mathbb{F}}_p$ the Corollary says in particular that G_T is obtained by base change from a finite group scheme over $\bar{\mathbb{F}}_p$. In this case we call G_T constant (compare [K], (2.7)). This should not be confused with the étale group scheme associated to a finite Abelian group A. We will discuss "constant" p-divisible groups, see Section 3 below.

2 The main result: slope filtrations

In this section we show:

2.1 Theorem. *Let h be a natural number. Then there exists a natural number $N(h)$ with the following property. Let S be an integral, normal Noetherian scheme. Let X be a p-divisible group over S of height h with constant Newton polygon. Then there is a completely slope divisible p-divisible group Y over S, and an isogeny:*

$$\varphi : X \to Y \quad \textit{over} \quad S \quad \textit{with} \quad \deg(\varphi) \le N(h).$$

In Example 4.2 we see that the condition "S is normal" is essential. By this theorem we see that a slope filtration exists up to isogeny:

2.2 Corollary. *Let X be a p-divisible group with constant Newton polygon over an integral, normal Noetherian scheme S. There exists an isogeny $\varphi : X \to Y$, such that Y over S admits a slope filtration.*

Q.E.D.

2.3 Proposition. *Let S be an integral scheme with function field $K = \kappa(S)$. Let X be a p-divisible group over S with constant Newton polygon, such that X_K is completely slope divisible with respect to the integers $s \ge r_1 > r_2 > \ldots > r_m \ge 0$. Then X is completely slope divisible with respect to the same integers.*

Proof. The quasi-isogeny $\Phi = p^{-r_m}\operatorname{Fr}^s : X \to X^{(p^s)}$ is an isogeny, because this is true over the general point. Over any geometric point $\eta \to S$ the Φ-étale part of X_η has the same height by constancy of the Newton polygon. Hence the Φ-étale part of X exists by Corollary 1.9. We obtain an exact sequence:

$$0 \to X^{\Phi-\mathrm{nil}} \to X \to X^{\Phi} \to 0.$$

Assuming an induction hypothesis on $X^{\Phi-\mathrm{nil}}$ gives the result. *Q.E.D.*

A basic tool in the following proofs is the moduli scheme of isogenies of degree d of a p-divisible group (compare [RZ] 2.22): Let X be a p-divisible group over a scheme S, and let d be a natural number. Then we define the following functor $\mathcal{M}$ on the category of S-schemes T. A point of $\mathcal{M}(T)$ consists of a p-divisible group Z over T and an isogeny $\alpha : X_T \to Z$ of degree d up to isomorphism. *The functor $\mathcal{M}$ is representable by a projective scheme over S.* Indeed, to each finite, locally free subgroup scheme $G \subset X_T$ there is a unique isogeny α with kernel G. Let n be a natural number such that $p^n \geq d$. Then G is a finite, locally free subgroup scheme on $X(n)_T$. We set $X(n) = \operatorname{Spec}_S \mathcal{A}$. The affine algebra of G is a quotient of the locally free sheaf $\mathcal{A}_T$. Hence we obtain a point of the Grassmannian of $\mathcal{A}$. This proves that $\mathcal{M}$ is representable as a closed subscheme of this Grassmannian.

2.4 LEMMA. *For every $h \in \mathbb{Z}_{>0}$ there exists a number $N(h) \in \mathbb{Z}$ with the following property. Let S be an integral Noetherian scheme. Let X be a p-divisible group of height h over S with constant Newton polygon. There is a non-empty open subset $U \subset S$, and a projective morphism $\pi : S^{\sim} \to S$ of integral schemes which induces an isomorphism $\pi : \pi^{-1}(U) \to U$ such that there exist a completely slope divisible p-divisible group Y over $S^{\sim}$, and an isogeny $X_{S^{\sim}} \to Y$, whose degree is bounded by $N(h)$.*

PROOF. Let K be the function field of S. We know by [Z1], Prop. 12, that there is a completely slope divisible p-divisible group Y^0 over K, and an isogeny $\beta^0 : X_K \to Y^0$, whose degree is bounded by a constant which depends only on the height of X. The kernel of this isogeny is a finite group scheme $G^0 \subset X_K(n)$, for some n. Let $\bar{G}$ be the scheme-theoretic image of G^0 in $X(n)$, see EGA, I.9.5.3. Then $\bar{G}$ is flat over some nonempty open set $U \subset S$, and inherits there the structure of a finite, locally free group scheme $G \subset X(n)_U$. We form the p-divisible group $Z = X_U/G$. By construction there are integers $s \geq r_m$, such that

$$p^{-r_m} \operatorname{Fr}^s : Z_K \to Z_K^{(p^s)}$$

is an isogeny, and r_m/s is a smallest slope in the Newton polygon of X.
Therefore $\Phi = p^{-r_m} \operatorname{Fr}^s : Z \to Z^{(p^s)}$ is an isogeny too. As in the proof of the last proposition the constancy of the Newton polygon implies that the Φ-étale part Z^{Φ} exists. We obtain an exact sequence of p-divisible groups on U:

$$0 \to Z^{\Phi-\text{nil}} \to Z \to Z^{\Phi} \to 0.$$

By induction we find a non-empty open subset $V \subset U$ and a completely slope divisible p-divisible group Y_{m-1} which is isogeneous to $Z_V^{\Phi-\text{nil}}$. Taking the push-out of the last exact sequence by the isogeny $Z_V^{\Phi-\text{nil}} \to Y_{m-1}$ we find a completely slope divisible p-divisible group Y over V which is isogeneous to X_V.

Let d be the degree of the isogeny $\rho : X_V \to Y$. We consider the moduli scheme $\mathcal{M}$ of isogenies of degree d of X defined above. The isogeny $X_V \to Y$ defines an S-morphism $V \to \mathcal{M}$. The scheme-theoretic image $S^{\sim}$ of V is an integral scheme, which is projective over S. Moreover the morphism $\pi : S^{\sim} \to S$ induces an isomorphism $\pi^{-1}(V) \to V$. The closed immersion $S^{\sim} \to \mathcal{M}$ corresponds to an isogeny $\rho^{\sim} : X_{S^{\sim}} \to Y^{\sim}$ to a p-divisible group $Y^{\sim}$ on $S^{\sim}$. Moreover the restriction of $\rho^{\sim}$ to V is ρ. Since $Y^{\sim}$ has constant Newton polygon, and since $Y^{\sim}$ is completely slope divisible in the generic point of $S^{\sim}$ it is completely slope divisible by Proposition 2.3. *Q.E.D.*

2.5 Lemma. *Let k be an algebraically closed field of characteristic p. Let $s \geq r_1 > r_2 > \ldots > r_m \geq 0$ and $d > 0$ be integers. Let X be a p-divisible group over k. Then there are up to isomorphism only finitely many isogenies $X \to Z$ of degree d to a p-divisible group Z, which is completely slope divisible with respect to $s \geq r_1 > r_2 > \ldots > r_m \geq 0$.*

Proof. It suffices to show this in case also X is completely slope divisible with respect to $s \geq r_1 > r_2 > \ldots > r_m \geq 0$. Then X and Z are a direct product of isoclinic slope divisible groups. Therefore we assume that we are in the isoclinic case $m = 1$.

In this case we consider the contravariant Dieudonné modules M of X, and N of Z. Let σ be the Frobenius on $W(k)$. Then $N \subset M$ is a submodule such that $\operatorname{length} M/N = \log_p d$. By assumption $\Phi = p^{-r_1}F^s$ induces a σ^s-linear automorphism of M respectively N. Let C_N respectively C_M be the invariants of Φ acting of N respectively M. Hence C_N is a $W(\mathbb{F}_{p^s})$-submodule of N such that $W(k) \otimes_{W(\mathbb{F}_{p^s})} C_N = N$ (e.g. [Z2] 6.26). The same holds for M. We see that C_N is a $W(\mathbb{F}_{p^s})$-submodule of C_M, such that $\operatorname{length} C_M/C_N = \log_p d$. Since there are only finitely many such submodules, the assertion follows. *Q.E.D.*

2.6 Lemma. *Let $f : T \to S$ be a proper morphism of schemes such that $f_*\mathcal{O}_T = \mathcal{O}_S$. Let $g : T \to M$ be a morphism of schemes. We assume that for any point $\xi \in S$ the set-theoretic image of the fiber T_ξ by g is a single point in M. Then there is a unique morphism $h : S \to M$ such that $hf = g$.*

Proof. For $\xi \in S$ we set $h(\xi) = f(T_\xi)$. This defines a set-theoretic map $h : S \to M$. If $U \subset M$ is an open neighborhood of $h(\xi)$ then $g^{-1}(U)$ is an open neighborhood of T_ξ. Since f is closed we find an open neighborhood V of ξ with $f^{-1}(V) \subset g^{-1}(U)$. Hence $h(V) \subset U$; we see that h is continuous. Then $h_*\mathcal{O}_S = h_*f_*\mathcal{O}_T = g_*\mathcal{O}_T$. We obtain a morphism of ringed spaces $\mathcal{O}_M \to g_*\mathcal{O}_T = h_*\mathcal{O}_S$. *Q.E.D.*

Theorem 2.1 follows from the following technical variant which is useful if we do not know that the normalization is finite. We will need that later on.

2.7 Proposition. *Let h be a natural number. Then there exists a natural number $N(h)$ with the following property. Let S be an integral Noetherian scheme. Let X be a p-divisible group over S of height h with constant Newton* *polygon. Then there is a finite birational morphism $T \to S$, a completely slope divisible p-divisible group Y over T, and an isogeny:*

$$X_T \to Y$$

over T whose degree is smaller than $N(h)$.

PROOF. Consider the proper birational map $\pi : S^{\sim} \to S$, and the isogeny $\rho : X_{S^{\sim}} \to Y$ given by Lemma 2.4. Take the Stein factorization $S^{\sim} \to T \to S$. It is enough to find over T an isogeny to a completely slope divisible p-divisible group. Therefore we assume $S = T$, i.e. $\pi_* \mathcal{O}_{S^{\sim}} = \mathcal{O}_S$.

Let $\mathcal{M} \to S$ be the moduli scheme of isogenies of X of degree $d =$ degree ρ. We will show that the S-morphism $g : S^{\sim} \to \mathcal{M}$ defined by Y factors through $S \to \mathcal{M}$.

Let $\xi \in S$. We write $S^{\sim}_{\xi} = S^{\sim} \times_S \operatorname{Spec}(\kappa(\xi))$. By Lemma 2.6 it is suffices to show that the set-theoretic image of $S^{\sim}_{\xi}$ by g is a single point of $\mathcal{M}$. Clearly $\mathcal{M}_{\xi}$ classifies isogenies starting at X_{ξ} of degree d. Over the algebraic closure $\bar{\xi}$, by Lemma 2.5, there are only finitely many isogenies of $X_{\bar{\xi}} \to Z$ of degree d to a completely slope divisible group (for fixed s and r_i). This shows that the image of $S^{\sim}_{\xi}(\bar{\xi}) \to \mathcal{M}_{\xi}(\bar{\xi})$ is finite. Since $S^{\sim}_{\xi}$ is connected, see [EGA], III^1.4.3.1, the image of $S^{\sim}_{\xi}$ is a single point of $\mathcal{M}_{\xi}$.

Hence we have the desired factorization $S \to \mathcal{M}$. It defines an isogeny $X \to Z$ over S. Finally Z is completely slope divisible since it is completely slope divisible in the general point of S, and because its Newton polygon is constant.

Q.E.D. 2.7 & 2.1

3 CONSTANCY RESULTS

Let T be a scheme over $\bar{\mathbb{F}}_p$. We study the question if a p-divisible group X over T is constant up to isogeny, i.e. there exists a p-divisible group Y over $\bar{\mathbb{F}}_p$ such that X is isogeneous to $Y \times_{\bar{\mathbb{F}}_p} T$.

3.1 PROPOSITION. *Let S be a Noetherian integral normal scheme over $\bar{\mathbb{F}}_p$. Let K be the function field of S and let $\bar{K}$ be an algebraic closure of K. We denote by $L \subset \bar{K}$ the maximal unramified extension of K with respect to S. Let T be the normalization of S in L.*
Let X be an isoclinic p-divisible group over S. Then there is a p-divisible group X_0 over $\bar{\mathbb{F}}_p$ and an isogeny $X \times_S T \to X_0 \times_{\operatorname{Spec} \bar{\mathbb{F}}_p} T$ such that the degree of this isogeny is smaller than an integer which depends only on the height of X.

PROOF: We use Theorem 2.1: there exists an isogeny $\varphi : X \to Y$, where Y over S is completely slope divisible. There are natural numbers r and s, such that

$$\Phi = p^{-r} \operatorname{Fr}^s : Y \to Y^{(p^s)}$$

is an isomorphism. Applying Corollary 1.10 to $Y(n)$ and Φ we obtain finite group schemes $X_0(n)$ over $\mathbb{F}_{p^s}$ and isomorphisms

$$Y(n)_T \cong X_0(n) \times_{\mathbb{F}_{p^s}} T$$

The inductive limit of the group schemes $X_0(n)$ is a p-divisible group X_0 over $\mathbb{F}_{p^s}$. It is isogeneous to X over T. *Q.E.D.*

3.2 Corollary. *Let S and T be as in the proposition. Let $T^{\mathrm{perf}} \to T$ be the perfect hull of T. Let X be a p-divisible group over S with constant Newton polygon. Then there is a p-divisible group X_0 over $\bar{\mathbb{F}}_p$ and an isogeny $X_0 \times_{\operatorname{Spec} \bar{\mathbb{F}}_p} T^{\mathrm{perf}} \to X \times_S T^{\mathrm{perf}}$, whose degree is smaller than an integer which depends only on the height of X.*

Proof. This follows using Proposition 1.3. *Q.E.D.*

Finally we prove constancy results without the normality condition.

3.3 Proposition. *Let R be a strictly Henselian reduced local ring over $\bar{\mathbb{F}}_p$. Let X be an isoclinic p-divisible group over $S = \operatorname{Spec} R$. Then there is a p-divisible group X_0 over $\bar{\mathbb{F}}_p$ and an isogeny $X_0 \times_{\operatorname{Spec} \bar{\mathbb{F}}_p} S \to X$, whose degree is smaller than an integer which depends only on the height of X.*

3.4 Corollary. *To each natural number h there is a natural number c with the following property: Let R be a Henselian reduced local ring over $\mathbb{F}_p$ with residue field k. Let X and Y be isoclinic p-divisible groups over $S = \operatorname{Spec} R$ whose heights are smaller than h. Let $\psi : X_k \to Y_k$ be a homomorphism. Then $p^c\psi$ lifts to a homomorphism $X \to Y$.*

A proof of the proposition, and of the corollary will be given later.

Remark. In case the R considered in the previous proposition, or in the previous corollary, is not reduced, but satisfies all other properties, the conclusions still hold, except that the integer bounding the degree of the isogeny, respectively the integer c, depend on h and on R.

If R is strictly Henselian the corollary follows from Proposition 3.3. Indeed, assume that X and Y are isogeneous to constant p-divisible groups X_0 and Y_0 by isogenies which are bounded by a constant which depends only on h. The corollary follows because:

$$\operatorname{Hom}((X_0)_k, (Y_0)_k) = \operatorname{Hom}((X_0)_R, (Y_0)_R).$$

Conversely the corollary implies the proposition since by Proposition 3.1 over a separably closed field an isoclinic p-divisible group is isogeneous to a constant p-divisible group.

Remark. Assume Corollary 3.4. An isoclinic slope divisible p-divisible group Y over k can be lifted to an isoclinic slope divisible p-divisible group over R. Indeed the étale schemes associated by 1.8 to the affine algebra of $Y(n)$ and the isomorphism $p^{-r}\operatorname{Fr}^s$ lift to R. Hence the categories of isoclinic p-divisible groups up to isogeny over R respectively k are equivalent.

3.5 LEMMA. *Consider a commutative diagram of rings over $\mathbb{F}_p$:*

$$\begin{array}{ccc} R & \rightarrow & A \\ \downarrow & & \downarrow \\ R_0 & \rightarrow & A_0. \end{array}$$

Assume that $R \rightarrow R_0$ is a surjection with nilpotent kernel $\mathfrak{a}$, and that $A \rightarrow A_0$ is a surjection with nilpotent kernel $\mathfrak{b}$. Moreover let $R \rightarrow A$ be a monomorphism. Let X and Y be p-divisible groups over R. Let $\varphi_0 : X_{R_0} \rightarrow Y_{R_0}$ be a morphism of the p-divisible groups obtained by base change. Applying base change with respect to $R_0 \rightarrow A_0$ we obtain a morphism $\psi_0 : X_{A_0} \rightarrow Y_{A_0}$.
If ψ_0 lifts to a morphism $\psi : X_A \rightarrow Y_A$, then φ_0 lifts to a morphism $\varphi : X \rightarrow Y$.

PROOF. By rigidity, liftings of homomorphisms of p-divisible groups are unique. Therefore we may replace R_0 by its image in A_0 and assume that $R_0 \rightarrow A_0$ is injective. Then we obtain $\mathfrak{a} = \mathfrak{b} \cap R$.

Let n be a natural number such that $\mathfrak{b}^n = 0$. We argue by induction on n. If $n = 0$, we have $\mathfrak{b} = 0$ and therefore $\mathfrak{a} = 0$. In this case there is nothing to prove. If $n > 0$ we consider the commutative diagram:

$$\begin{array}{ccc} R & \rightarrow & A \\ \downarrow & & \downarrow \\ R/(\mathfrak{b}^{n-1} \cap R) & \rightarrow & A/\mathfrak{b}^{n-1} \\ \downarrow & & \downarrow \\ R_0 & \rightarrow & A_0. \end{array}$$

We apply the induction hypothesis to the lower square. Hence it is enough to show the lemma for the upper square. We assume therefore without loss of generality that $\mathfrak{a}^2 = 0,\ \mathfrak{b}^2 = 0$.

Let D_X and D_Y be the crystals associated to X and Y by Messing [Me]. The values $D_X(R)$ respectively $D_Y(R)$ are finitely generated projective R-modules which are endowed with the Hodge filtration $Fil_X \subset D_X(R)$ respectively $Fil_Y \subset D_Y(R)$. We put on $\mathfrak{a}$ respectively $\mathfrak{b}$ the trivial divided power structure. Then φ_0 induces a map $D_X(R) \rightarrow D_Y(R)$. By the criterion of Grothendieck and Messing φ_0 lifts to a homomorphism over R, iff $D(\varphi_0)(Fil_X) \subset Fil_Y$.

Since the construction of the crystal commutes with base change, see [Me], Chapt. IV, 2.4.4, we have canonical isomorphisms:

$$\begin{aligned} D_{X_A}(A) &= A \otimes_R D_X(R), \qquad & Fil_{X_A} &= A \otimes_R Fil_X, \\ D_{Y_A}(A) &= A \otimes_R D_Y(R), \qquad & Fil_{Y_A} &= A \otimes_R Fil_Y. \end{aligned}$$

Since ψ_0 lifts we have $id_A \otimes D(\varphi)(A \otimes_R Fil_X) \subset A \otimes_R Fil_Y$. Since $R \rightarrow A$ is injective this implies $D(\varphi_0)(Fil_X) \subset Fil_Y$. *Q.E.D.*

PROOF of Proposition 3.3. We begin with the case where R is an integral domain. By Proposition 2.7 there is a finite ring extension $R \rightarrow A$ such that A is contained in the quotient field of R, and such that there is an isogeny

$X_A \to Y$ to a completely slope divisible p-divisible group Y over A. The degree of this isogeny is smaller than a constant which depends only on the height of X. Since A is a product of local rings we may assume without loss of generality that A is local. The ring A is a strictly Henselian local ring, see [EGA] IV 18.5.10, and has therefore no non-trivial finite étale coverings. The argument of the proof of Proposition 3.1 shows that Y is obtained by base change from a p-divisible group X_0 over $\bar{\mathbb{F}}_p$. Therefore we find an isogeny

$$\varphi : (X_0)_A \to X_A$$

Let us denote by k the common residue field of A and R. Then φ induces an isogeny $\varphi : (X_0)_k \to X_k$. The last lemma shows that φ_0 lifts to an isogeny $\hat{\varphi} : (X_0)_{\hat{R}} \to X_{\hat{R}}$ over the completion $\hat{R}$ of R.

We apply the following fact:

Claim. Consider a fiber product of rings:

$$\begin{array}{ccc} R & \to & A_1 \\ \downarrow & & \downarrow \\ A_2 & \to & B. \end{array}$$

Let X and Y be p-divisible groups over R. Let $\psi_i : X_{A_i} \to Y_{A_i}$ for $i = 1, 2$ be two homomorphisms of p-divisible groups which agree over B. Then there is a unique homomorphism $\psi : X \to Y$ which induces ψ_1 and ψ_2.

In our concrete situation we consider the diagram:

$$\begin{array}{ccc} R & \to & A \\ \downarrow & & \downarrow \\ \hat{R} & \to & \hat{A} = \hat{R} \otimes_R A. \end{array}$$

The morphisms $\hat{\varphi}$ and φ agree over $\hat{A}$ because they agree over the residue field k. This proves the case of an integral domain R.

In particular we have shown the Corollary 3.4 in the case where R is a strictly Henselian integral domain. To show the corollary in the reduced case we consider the minimal prime ideals $\mathfrak{p}_1, \ldots \mathfrak{p}_s$ of R. Let $\psi : X_k \to Y_k$ be a homomorphism. Then $p^c\psi$ lifts to a homomorphims over each of the rings $R/\mathfrak{p}_i$, for $i = 1, \ldots, s$. But then we obtain a homomorphism over R using the Claim above. This proves the Corollary 3.4 and hence the Proposition 3.3 in the case where R is reduced and strictly Henselian.

If R is not reduced one applies standard deformation theory to $R \to R_{\mathrm{red}}$, [Z2], 4.47. *Q.E.D.*

Proof of Corollary 3.4. Consider the diagram:

$$\begin{array}{ccc} R & \to & R^{\mathrm{sh}} \\ \downarrow & & \downarrow \\ \hat{R} & \to & \hat{R}^{\mathrm{sh}}. \end{array}$$

The upper index "sh" denotes the strict henselization. Using the fact that the categories of finite étale coverings of R, k, respectively $\hat{R}$ are equivalent it is easy to see that the last diagram is a fiber product. We have already proved that $p^c\psi$ lifts to a homomorphism over R^{sh}. Applying Lemma 3.5 to the following diagram we see that $p^c\psi$ lifts to $\hat{R}$. This is enough to prove the corollary (compare the Claim above).

$$\begin{array}{ccc} R/\mathfrak{m}^n & \rightarrow & R^{\mathrm{sh}}/\mathfrak{m}^n R^{\mathrm{sh}} \\ \downarrow & & \downarrow \\ R/\mathfrak{m} & \rightarrow & R^{\mathrm{sh}}/\mathfrak{m}R^{\mathrm{sh}} \end{array}$$

In this diagram n is a positive integer and $\mathfrak{m}$ is the maximal ideal of R. *Q.E.D.*

3.6 Corollary. *Let R be a strictly Henselian reduced local ring over $\bar{\mathbb{F}}_p$. Let R^{perf} be the perfect hull of R. Let X be a p-divisible group over $S = \operatorname{Spec} R$ with constant Newton polygon. We set $S^{\mathrm{perf}} = \operatorname{Spec}(R^{\mathrm{perf}})$. Then there is a p-divisible group X_0 over $\bar{\mathbb{F}}_p$ and an isogeny $X_0 \times_{\operatorname{Spec} \bar{\mathbb{F}}_p} S^{\mathrm{perf}} \rightarrow X \times_S S^{\mathrm{perf}}$ such that the degree of this isogeny is bounded by an integer which depends only on the height of X.*

Proof: This follows using Proposition 1.3. *Q.E.D.*

4 Examples

In this section we use the p-divisible groups either $Z = G_{1,n}$, with $n \geq 1$, or $Z = G_{m,1}$, with $m \geq 1$ as building blocks for our examples. These have the property to be iso-simple, they are defined over $\mathbb{F}_p$, they contain a unique subgroup scheme $N \subset Z$ isomorphic with α_p, and $Z/N \cong Z$. Indeed, for $Z = G_{1,n}$ we have an exact sequence of sheaves:

$$0 \rightarrow \alpha_p \rightarrow Z \xrightarrow{\mathrm{Fr}} Z \rightarrow 0.$$

For $Z = G_{m,1}$ we have the exact sequence

$$0 \rightarrow \alpha_p \rightarrow Z \xrightarrow{\mathrm{Ver}} Z \rightarrow 0.$$

Moreover every such Z has the following property: if $Z_K \rightarrow Z'$ is an isogeny over some field K, and k an algebraic closed field containing K, then $Z_k \cong Z'_k$.

4.1 Example. *In this example we produce a p-divisible group X with constant Newton polygon over a regular base scheme which does not admit a slope filtration.*

Choose Z_1 and Z_2 as above, with $\mathrm{slope}(Z_1) = \lambda_1 > \lambda_2 = \mathrm{slope}(Z_2)$; e.g. $Z_1 = G_{1,1}$ and $Z_2 = G_{1,2}$. We choose $R = K[t]$, where K is a field. We write $S = \operatorname{Spec}(R)$ and $\mathcal{Z}_i = Z_i \times S$ for $i = 1, 2$. We define

$$(id, t) : \alpha_p \rightarrow \alpha_p \times \alpha_p \cong N_1 \times N_2; \quad \text{this defines} \quad \psi : \alpha_p \times S \rightarrow \mathcal{Z}_1 \times \mathcal{Z}_2.$$

Claim: *$\mathcal{X} := (\mathcal{Z}_1 \times \mathcal{Z}_2)/\psi(\alpha_p \times S)$ is a p-divisible group over S which does not admit a slope filtration.*

Indeed, for the generic point we do have slope filtration, where $X = \mathcal{X} \otimes K(t)$, and $0 \subset X_1 \subset X$ is given by: X_1 is the image of

$$\xi_K : (\mathcal{Z}_1 \otimes K(t) \to (\mathcal{Z}_1 \times \mathcal{Z}_2) \otimes K(t) \to X).$$

However the inclusion ξ_K extends uniquely a homomorphism $\xi : \mathcal{Z}_1 \to \mathcal{X}$, which is not injective at $t \mapsto 0$. This proves the claim.

4.2 Example. *In this example we construct a p-divisible group $\mathcal{X}$ with constant Newton polygon over a base scheme S* which is not normal*, such that there is no isogeny $\phi : \mathcal{X} \to \mathcal{Y}$ to a completely slope divisible p-divisible group. (i.e. we show the condition that S is* normal *in Theorem 2.1 is necessary).*

We start again with the exact sequences over $\mathbb{F}_p$:

$$\begin{array}{lllll} 0 & \to \alpha_p & \to G_{2,1} & \xrightarrow{\text{Ver}} G_{2,1} & \to 0, \\ 0 & \to \alpha_p & \to G_{1,2} & \xrightarrow{\text{Fr}} G_{1,2} & \to 0. \end{array} \tag{1}$$

We fix an algebraically closed field k. We write $T = \mathbb{P}^1_k$, and:

$$Z_1 = G_{2,1} \times_{\mathbb{F}_p} T, \quad Z_2 = G_{1,2} \times_{\mathbb{F}_p} T, \quad Z = Z_1 \times Z_2, \quad A = \alpha_p \times_{\mathbb{F}_p} T.$$

By base change we obtain sequences of sheaves on the projective line $T = \mathbb{P}^1_k$:

$$\begin{array}{lllll} 0 & \to A & \to Z_1 & \xrightarrow{\text{Ver}} Z_1 & \to 0, \\ 0 & \to A & \to Z_2 & \xrightarrow{\text{Fr}} Z_2 & \to 0. \end{array}$$

Lemma. *Consider $Z \to T = \mathbb{P}^1_k$ as above. Let $\beta : Z \to Y$ be an isogeny to a completely slope divisible p-divisible group Y over T. Then $Y = Y_1 \times Y_2$ is a product of two p-divisible groups and $\beta = \beta_1 \times \beta_2$ is the product of two isogenies $\beta_i : Z_i \to Y_i$.*

Proof. The statement is clear if we replace the base T by a perfect field, see Proposition 1.3. In our case we show first that the kernel of the morphisms $Z_i \to Y$, $i = 1, 2$ induced by β are representable by a finite, locally free group schemes G_i. Indeed, let $\mathcal{G}_i$ the kernel in the sense of f.p.p.f sheaves. Let us denote by G the kernel of the isogeny β. Choose a number n such that p^n annihilates G. Then p^n annihilates $\mathcal{G}_i$. Therefore $\mathcal{G}_i$ coincides with the kernel of the morphism of finite group schemes $Z_i(n) \to Y(n)$. Hence $\mathcal{G}_i$ is representable by a finite group scheme G_i. We prove that G_i is locally free. It suffices to verify that the rank of G_i in any geometric point η of T is the same. But we have seen that over η the p-divisible group Y splits into a product $Y_\eta = (Y_\eta)_1 \times (Y_\eta)_2$. This implies that

$$G_\eta = (G_1)_\eta \times (G_2)_\eta.$$

We conclude that the ranks of $(G_i)_\eta$ are independent of η since G is locally free.

We define p-divisible groups $Y_i = Z_i/G_i$. We obtain a homomorphism of p-divisible groups $Y_1 \times Y_2 \to Y$ which is an isomorphism over each geometric point η. Therefore this is an isomorphism. *Q.E.D.*

Next we construct a p-divisible group X on $T = \mathbb{P}^1_k$. Let $\mathcal{L}$ be a line bundle on $\mathbb{P}^1_k$. We consider the associated vector group

$$\mathbf{V}(\mathcal{L})(T') = \Gamma(T', \mathcal{L}'_T),$$

where $T' \to T = \mathbb{P}^1_k$ is a scheme and $\mathcal{L}'_T$ is the pull-back. The kernel of the Frobenius morphism $\mathrm{Fr} : \mathbf{V}(\mathcal{L}) \to \mathbf{V}(\mathcal{L})^{(p)}$ is a finite, locally free group scheme $\alpha_p(\mathcal{L})$ which is locally isomorphic to α_p. We set $A(-1) = \alpha_p(\mathcal{O}_{\mathbb{P}^1_k}(-1))$. There are up to multiplication by an element of k^* unique homomorphisms ι_0 respectively $\iota_\infty : \mathcal{O}_{\mathbb{P}^1_k}(-1) \to \mathcal{O}_{\mathbb{P}^1_k}$ whose unique zeroes are $0 \in \mathbb{P}^1_k$ respectively $\infty \in \mathbb{P}^1_k$. This induces homomorphisms of finite group schemes $\iota_0 : A(-1) \to A$ respectively $\iota_\infty : A(-1) \to A$ which are isomorphisms outside 0 respectively outside ∞. We consider the embeddings

$$A(-1) \quad \xrightarrow{(\iota_0, \iota_\infty)} \quad A \times A \quad \subset \quad Z_1 \times Z_2 = Z.$$

We define $X = Z/A(-1)$:

$$\psi : Z \longrightarrow Z/A(-1) = X.$$

Note that

$$\psi_0 : G_{2,1} \times G_{1,2} = Z_0 \longrightarrow G_{2,1} \times (G_{1,2}/\alpha_p) = X_0,$$

and

$$\psi_\infty : G_{2,1} \times G_{1,2} = Z_\infty \longrightarrow (G_{2,1}/\alpha_p) \times G_{1,2} = X_\infty.$$

We consider the quotient space $S = \mathbb{P}^1_k/\{0, \infty\}$, by identifying 0 and ∞ into a normal crossing at $P \in S$, i.e.

$$\mathcal{O}_{S,P} = \{f \in \mathcal{O}_{T,0} \cap \mathcal{O}_{T,\infty} \mid f(0) = f(\infty)\};$$

S is a nodal curve, and

$$\mathbb{P}^1_k = T \longrightarrow S, \quad 0 \mapsto P, \quad \infty \mapsto P,$$

is the normalization morphism.

A finite, locally free scheme $\mathcal{G}$ over S is the same thing as a finite, locally free scheme G over $\mathbb{P}^1_k$ endowed with an isomorphism

$$G_0 \cong G_\infty.$$

It follows that the category of p-divisible groups $\mathcal{Y}$ over S is equivalent to the category of pairs $(Y, \gamma_{\mathcal{Y}})$, where Y is a p-divisible group on $\mathbb{P}^1_k$ and $\gamma_{\mathcal{Y}}$ is an isomorphism

$$\gamma_{\mathcal{Y}} : Y_0 \cong Y_\infty$$

of the fibers of Y over $0 \in \mathbb{P}^1_k$ and $\infty \in \mathbb{P}^1_k$. We call $\gamma_{\mathcal{Y}}$ the gluing datum of $\mathcal{Y}$. We construct a p-divisible $\mathcal{X}$ over S by defining a gluing datum on the p-divisible group X. In fact, the exact sequences in (1) give:

$$X_0 = G_{2,1} \times (G_{1,2}/\alpha_p) \quad \cong \quad (G_{2,1}/\alpha_p) \times G_{1,2} = X_\infty;$$

this gluing datum provides a p-divisible group $\mathcal{X}$ over S.

Claim: *This p-divisible group $\mathcal{X} \to S$ satisfies the property mentioned in the example.*

Let us assume that there exists an isogeny $\phi : \mathcal{X} \to \mathcal{Y}$ to a completely slope divisible p-divisible group $\mathcal{Y}$ over S. We set $Y = \mathcal{Y} \times_S T$ and consider the induced isogeny

$$\phi \times_S T = \varphi : X \to Y.$$

and the induced isogeny:

$$\beta = \varphi{\cdot}\psi : (Z \xrightarrow{\psi} X \xrightarrow{\varphi} Y).$$

By the lemma we see that

$$\beta = \beta_1 \times \beta_2 : Z \longrightarrow Y_1 \times Y_2 = Y.$$

Note that

$$\varphi_\infty = \phi_P = \varphi_0 : X_\infty = \mathcal{X}_P = X_0 \quad \longrightarrow \quad Y_\infty = \mathcal{Y}_P = Y_0;$$

these p-divisible groups both have a splitting into isoclinic summands:

$$X_\infty = \mathcal{X}_P = X_0 = X' \times X'', \quad Y_\infty = \mathcal{Y}_P = Y_0 = Y' \times Y'',$$

and

$$\phi_P = \varphi' \times \varphi'' : X' \times X'' \quad \longrightarrow \quad Y' \times Y''$$

is in diagonal form. On the one hand we conclude from

$$((\beta_1)_0 : (Z_1)_0 \longrightarrow (Y_1)_0) = \Big((Z_1)_0 \xrightarrow{\sim} X' \to Y'\Big)$$

that $\deg(\beta_1) = \deg(\beta_1)_0 = \deg(\varphi')$; on the other hand

$$((\beta_1)_\infty : (Z_1)_\infty \longrightarrow (Y_1)_\infty) = ((Z_1)_\infty \to (G_{2,1}/\alpha_p) = X' \to Y');$$

hence

$$\deg(\beta_1) = \deg(\beta_1)_\infty = p{\cdot}\deg(\varphi').$$

We see that the assumption that the isogeny $\phi : \mathcal{X} \to \mathcal{Y}$ to a completely slope divisible $\mathcal{Y} \to S$ would exist leads to a contradiction. This finishes the description and the proof of Example 4.2. *Q.E.D.*

4.3 Example. *For every positive integer d there exists a scheme S' of dimension d, a point $P' \in S'$ such that S' is regular outside P', and a p-divisible group $\mathcal{X}' \to S'$ which does not admit an isogeny to a completely slope divisible group over S'.*

This follows directly form the previous example. Indeed choose T as in the previous example, and let $T' \to T$ smooth and surjective with T' of dimension d. Pull back X/T to X'/T'; choose geometric points $0'$ and $\infty' \in T'$ above 0 and $\infty \in T$; construct S' by "identifying $0'$ and ∞'": outside $P' \in S'$, this scheme is $T' \backslash \{0', \infty'\}$, and the local ring of $P' \in S'$ is the set of pairs of elements in the local rings of $0'$ and ∞' having the same residue value. We can descend $X' \to T'$ to $\mathcal{X}' \to S'$, and this has the desired property.

References

[D] J. Dieudonné: *Lie groups and Lie hyperalgebras over a field of characteristic* $p > 0$. II. Amer. Journ. Math. 77 (1955), 218 - 244.

[G] A. Grothendieck: *Groupes de Barsotti-Tate et cristaux de Dieudonné.* Sém. Math. Sup. 45, Presses de l'Univ. de Montreal, 1970.

[EGA] A. Grothendieck & J. Dieudonné: *Éléments de Géometrie Algébrique.*
I: *Le language des schémas.*
III[1]: *Étude cohomologique des faisceaux cohérents.*
IV[4]: *Étude locale des schémas et des morphismes de schémas.*
Publ. Math. IHES No 4, 11, 32; 1960, 1961, 1967.

[HW] H. Hasse & E. Witt: *Zyklische unverzweigte Erweiterungskörper vom Primzahlgrade p über einem algebraischen Funktionenkörper der Charactersitik p.* Monatshefte für Math. und Physik 43 (1936), 477 - 492.

[J] A. J. de Jong: *Homomorphisms of Barsotti-Tate groups and crystals in positive characteristics.* Invent. Math. 134 (1998) 301-333, Erratum 138 (1999) 225.

[JO] A. J. de Jong & F. Oort: *Purity of the stratification by Newton polygons.* J. Amer. Math. Soc. 13 (2000), 209-241.

[K] N. M. Katz: *Slope filtration of F–crystals.* Astérisque 63 (1979), 113 - 164.

[Ma] Yu. I. Manin: *The theory of commutative formal groups over fields of finite characteristic.* Usp. Math. 18 (1963), 3-90; Russ. Math. Surveys 18 (1963), 1-80.

[Me] W. Messing: *The crystals associated to Barsotti-Tate groups.* Lect. Notes Math. 264, Springer - Verlag 1972.

[FO] F. Oort: *Commutative group schemes.* Lect. Notes Math. 15, Springer - Verlag 1966.

[RZ] M. Rapoport & Th. Zink: *Period spaces for p–divisible groups.* Annals of Mathematics Studies 141, Princeton 1996.

[Z1] Th. Zink: *On the slope filtration.* Duke Math. J. Vol.109 (2001), 79-95.

[Z2] Th. Zink: *Cartiertheorie kommutativer formaler Gruppen.* Teubner Texte zur Mathematik 68, Leipzig 1984.

Frans Oort
Mathematisch Instituut
P.O. Box. 80.010
NL - 3508 TA Utrecht
The Netherlands
oort@math.uu.nl

Thomas Zink
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
D-33501 Bielefeld
Deutschland
zink@mathematik.uni-bielefeld.de

The Gersten Conjecture for Witt Groups in the Equicharacteristic Case

Paul Balmer, Stefan Gille,
Ivan Panin and Charles Walter

Received: May 22, 2002
Revised: November 1, 2002

Communicated by Ulf Rehmann

Abstract. We prove the Gersten conjecture for Witt groups in the equicharacteristic case, that is for regular local rings containing a field of characteristic not 2.

2000 Mathematics Subject Classification: 11E81, 18E30, 19G12
Keywords and Phrases: Witt group, Gersten conjecture, equicharacteristic, triangulated categories

0. Introduction

The *Witt group* is the classical invariant classifying symmetric spaces, up to isometry and modulo metabolic spaces, see for instance [12] for rings and [11] for schemes. The *Gersten conjecture for Witt groups*, stated by Pardon in 1982 [16], claims the existence and the exactness of a complex $\mathrm{GWC_a}(R)$:

$$0 \to \mathrm{W}(R) \to \mathrm{W}(K) \to \bigoplus_{x \in X^{(1)}} \mathrm{W}(\kappa(x)) \to \ldots \to \bigoplus_{x \in X^{(n-1)}} \mathrm{W}(\kappa(x)) \to \mathrm{W}(\kappa(\mathfrak{m})) \to 0$$

where $(R, \mathfrak{m})$ is an n-dimensional regular local ring in which 2 is invertible; we denote by $X = \mathrm{Spec}(R)$ the spectrum of R, by $X^{(p)}$ the primes of height p, by $\kappa(x)$ the residue field at a point $x \in X$, and by $K = \kappa(0)$ the field of fractions of R. We call $\mathrm{GWC_a}(R)$ *an augmented Gersten-Witt complex.* In [5] Balmer and Walter constructed a *Gersten-Witt complex*

$$\mathrm{GWC}(X) := \quad \ldots \to 0 \longrightarrow \bigoplus_{x \in X^{(0)}} \mathrm{W}(\kappa(x)) \longrightarrow \ldots \longrightarrow \bigoplus_{x \in X^{(p)}} \mathrm{W}(\kappa(x)) \to \ldots$$

for general regular schemes X, not necessarily local or essentially of finite type, as part of the so-called *Gersten-Witt spectral sequence.* We will recall these constructions in Section 3. The augmented Gersten-Witt complex that we consider here is simply their complex $\mathrm{GWC}(R)$ augmented by the natural map $\mathrm{W}(R) \to \mathrm{W}(K)$. Our main result is Theorem 6.1 below, which says:

Theorem. *The augmented Gersten-Witt complex* $\mathrm{GWC_a}(R)$ *is exact for any equicharacteristic regular local ring R, i.e. for R regular local containing some field k.*

If the field k can be taken infinite with R essentially smooth over k, this has already been proven by Balmer [4] and independently by Pardon [17]. Here we extend this result first to essentially smooth local algebras over finite ground fields k. Then we extend it to regular local algebras which are not essentially of finite type to obtain the above Theorem, following a method introduced by Panin [15] to prove the equicharacteristic Gersten conjecture in K-theory.
Although these strategies have already been used for other theories, their application to Witt theory has not been rapid. For instance the Gersten conjecture for K-theory was proven by Quillen 30 years before the analogue for Witt groups. The most significant problem was that until recently [2] [3] [4] [7] it had not been established that Witt groups were part of a *cohomology theory with supports* in the sense of Colliot-Thélène, Hoobler and Kahn [6]. It is basically this observation which led to the proof of the conjecture for essentially smooth local algebras over infinite ground fields by means of a geometric proof whose roots reach back to Ojanguren's pioneering article [13].
Let us explain the general strategies to

(1) Go from infinite ground fields to any ground field.
(2) Go from essentially smooth local algebras to any regular local algebra.

The strategy for proving (1) is seemingly due to Colliot-Thélène, *cf.* [13], p. 115. One considers infinite towers of finite field extensions $k \subset F_1 \subset F_2 \subset \ldots$; the result holds "at the limit" by assumption, hence holds for some finite extension, and finally it holds for k itself by a transfer argument.
The strategy for proving (2) is due to Panin [15] and relies on results of Popescu [18] [19] which imply that any equicharacteristic regular local ring R is the filtered colimit of essentially smooth local algebras over some field $k \subset R$. There is usually no hope of getting this limit to commute with Gersten-type complexes because the morphisms in the colimit may be pretty wild. Panin's trick consists in finding a statement in terms of Zariski cohomology which is equivalent to the considered Gersten conjecture (he did it for K-theory) and then using a theorem of Grothendieck [1] asserting that the colimit and the cohomology commute.
We follow these strategies for Witt groups. The main difference between the usual cohomology theories (such as K-theory) and Witt groups is that the latter depend not only on a scheme or a category but also on a duality functor $E \mapsto E^*$ and biduality isomorphisms $\varpi_E : E \cong E^{**}$. Most schemes and categories which one studies this way come equipped with numerous choices for $({}^*, \varpi)$. For instance one can twist the duality functor for vector bundles by a line bundle, one can use shifted dualities for chain complexes, and one can change the sign of the biduality isomorphisms. When one wishes to apply a geometric argument with a pullback or a pushforward along a map $\pi : Y \to X$ one has to worry about which dualities on X and Y correspond for the construction in

question. Pushforwards (or transfers) in particular are not yet widely available (although some of the authors are working on it). Nevertheless, things have reached the point where one understands enough to construct the Gersten-Witt complex (Balmer-Walter [5]) and to treat pushforwards along a closed embedding $\mathrm{Spec}(R/fR) \hookrightarrow \mathrm{Spec}\, R$ of spectra of regular local rings (Gille [7]). This allows us to carry out (1) and (2).
Another reason we write this paper is that Panin's strategy for (2) is still quite new and has not yet been assimilated by the community. We hope that our exposition of this method will aid the process of digestion.
Note that our proof of the Gersten conjecture is independent of Ojanguren's and Panin's work [14] and hence we get a new proof of their main theorem, namely the purity theorem for equicharacteristic regular local rings. Nevertheless, the present work is more a generalization than a simplification of [14] since the various pieces of our proof (geometric presentation lemmas, transfers, and Panin's trick) are of similar complexity.
Apart from the ideas described in this introduction, our basic technical device is the recourse to *triangular Witt groups* [2] [3], namely Witt groups of suitable derived categories.

We would like to thank Winfried Scharlau and the *Sonderforschungsbereich* of the University of Münster for their precious support and for the one week workshop where this article was started.
The third author thanks very much for the support the TMR Network ERB FMRX CT-97-0107, the grant of the year 2002 of the "Support Fund of National Science" at the Russian Academy of Science, the grant INTAS-99-00817, and the RFFI-grant 00-01-00116.

1. Notations

Convention 1.1. Each time we consider the Witt group of a scheme X or of a category $\mathcal{A}$, we implicitly assume that 2 is invertible, *i.e.* that $1/2$ is in the ring of global sections $\Gamma(X, \mathcal{O}_X)$, respectively that $\mathcal{A}$ is $\mathbb{Z}[1/2]$-linear. Of course, this has nothing to do with "tensoring outside with $\mathbb{Z}[1/2]$" and our Witt groups might very well have non-trivial 2-torsion.

Let X be a noetherian scheme with structure sheaf $\mathcal{O}_X$, and let Z be a closed subset. For a complex P of quasi-coherent $\mathcal{O}_X$-modules we define the (*homological*) *support* of P to be

$$\mathrm{supph}(P) := \bigcup_{i\in\mathbb{Z}} \mathrm{supp}\,\big(H_i(P)\big).$$

We denote by $\mathcal{M}_X$ the category of quasi-coherent $\mathcal{O}_X$-modules and by $\mathcal{P}_X$ the category of locally free $\mathcal{O}_X$-modules of finite rank. We denote by $D^b(\mathcal{E})$ the bounded derived category of an exact category $\mathcal{E}$. Let $D^b_{\mathrm{coh}}(\mathcal{M}_X)$ be the full subcategory of $D^b(\mathcal{M}_X)$ of complexes whose homology modules are coherent, and let $D^b_{\mathrm{coh},\,Z}(\mathcal{M}_X)$ be the full subcategory of $D^b_{\mathrm{coh}}(\mathcal{M}_X)$ of those complexes whose homological support is contained in Z. The symbol $D^b_Z(\mathcal{P}_X)$ has an analogous self-explanatory meaning. For any positive integer $p \geq 0$ we set

$$D^b_Z(\mathcal{P}_X)^{(p)} := \bigcup_{\substack{\operatorname{codim} \mathrm{W} = p \\ \mathrm{W} \subset Z}} D^b_{\mathrm{W}}(\mathcal{P}_X).$$

The category $D^b_{\mathrm{coh},\,Z}(\mathcal{M}_X)^{(p)}$ is defined similarly.

Remark 1.2. We shall use here the following standard abbreviations:

(1) When $Z = X$, we drop its mention, as in $D^b(\mathcal{P}_X)$ to mean $D^b_X(\mathcal{P}_X)$.
(2) In the affine case, $X = \operatorname{Spec}(R)$, we drop "Spec", as in $D^b_{\mathrm{coh}}(\mathcal{M}_R)^{(p)}$ which stands for $D^b_{\mathrm{coh}}(\mathcal{M}_{\operatorname{Spec}(R)})^{(p)}$.
(3) If $X = \operatorname{Spec}(R)$ and $Z = V(I)$ is defined by an ideal $I \subset R$, we replace "Z" by "I", and even further we abbreviate $D^b_f(\mathcal{P}_R)$ instead of $D^b_{fR}(\mathcal{P}_R)$ where $f \in R$.

2. Triangulated categories with duality and their Witt groups

When not mentioned, the reference for this section is [2].

A *triangulated category with duality* is a triple $(\mathcal{K}, \sharp, \varpi)$, where $\mathcal{K}$ is a triangulated category, $\sharp : \mathcal{K} \to \mathcal{K}$ is a δ-exact contravariant functor ($\delta = \pm 1$) and $\varpi : \mathrm{id}_{\mathcal{K}} \to \sharp\sharp$ is an isomorphism of functors such that $\mathrm{id}_{M^\sharp} = (\varpi_M)^\sharp \cdot \varpi_{M^\sharp}$ and $\varpi_M[1] = \varpi_{M[1]}$.

Triangular Witt groups. We can associate to a triangulated category with duality a series of Witt groups $\mathrm{W}^n(\mathcal{K})$, for $n \in \mathbb{Z}$. The group $\mathrm{W}^n(\mathcal{K})$ classifies the n-symmetric spaces modulo Witt equivalence. Here an *n-symmetric space* is a pair (P, ϕ) with $P \in \mathcal{K}$ and with $\phi : P \xrightarrow{\sim} P^\sharp[n]$ an isomorphism such that $\phi^\sharp[n]\,\varpi_P = (-1)^{\frac{n(n+1)}{2}} \delta^n \phi$. The isometry classes of n-symmetric spaces form a monoid with the orthogonal sum as addition. Dividing this monoid by the submonoid of *neutral* n-symmetric spaces (see [2] Definition 2.12) gives the n-th Witt group of $\mathcal{K}$. These groups are 4-periodic, *i.e.* $\mathrm{W}^n(\mathcal{K}) = \mathrm{W}^{n+4}(\mathcal{K})$. The class of (P, ϕ) in the Witt group is written $[P, \phi]$.

Derived Witt groups of schemes. Let X be a scheme and $Z \subset X$ a closed subset. The derived functor of $\mathcal{H}\mathrm{om}_{\mathcal{O}_X}(\,-\,, \mathcal{O}_X)$ is then a duality on $D^b_Z(\mathcal{P}_X)$ making it a triangulated category with 1-exact duality. We denote the corresponding triangular Witt groups by $\mathrm{W}^n_Z(X)$ and call them the *derived Witt groups of X with support in Z*. The abbreviations introduced in 1.2 also apply to this notation, like $\mathrm{W}^n(X)$ for $\mathrm{W}^n_X(X)$. The comparison with the classical Witt group of the scheme X defined by Knebusch [11] is given by the following fact ([3] Theorem 4.7): The natural functor $\mathcal{P}_X \to D^b(\mathcal{P}_X)$ induces an isomorphism $\mathrm{W}(X) \xrightarrow{\simeq} \mathrm{W}^0(X)$.

The cone construction and the localization long exact sequence. The main theorem of triangular Witt theory is the localization theorem. Let

$$0 \to \mathcal{J} \longrightarrow \mathcal{K} \xrightarrow{q} \mathcal{K}/\mathcal{J} \to 0 \tag{2.1}$$

be an exact sequence of triangulated categories with duality, *i.e.* a localization with $\mathcal{J}^\sharp \subset \mathcal{J}$. Let z be an element of $\mathrm{W}^0(\mathcal{K}/\mathcal{J})$. Then there exists a symmetric morphism $\psi : P \to P^\sharp$ (*i.e.* $\psi^\sharp \varpi_P = \psi$) such that $z = [q(P), q(\psi)]$. In particular $C := \operatorname{cone} \psi$ belongs to $\mathcal{J}$. By [2] Theorem 2.6, there is a commutative diagram

$$\begin{tikzcd}
P \arrow[r, "\psi"] \arrow[d, "\delta\varpi_P"'] & P^\sharp \arrow[r] \arrow[d, "="'] & C \arrow[r] \arrow[d, "\simeq"', "\phi"] & P[1] \arrow[d, "\delta\varpi_P[1]"] \\
P^{\sharp\sharp} \arrow[r, "\delta\psi^\sharp"'] & P^\sharp \arrow[r] & C^\sharp[1] \arrow[r] & P^{\sharp\sharp}[1]
\end{tikzcd}$$

such that the upper and the lower rows are exact triangles dual to each other and such that $\phi^\sharp[1]\varpi_C = -\delta\phi$. The last property means that (C, ϕ) is a 1-symmetric space, *i.e.* represents an element of $\mathrm{W}^1(\mathcal{J})$. The isometry class of (C, ϕ) is uniquely determined by the isometry class of (P, ψ). We get a morphism $\mathrm{W}^0(\mathcal{K}) \longrightarrow \mathrm{W}^1(\mathcal{J})$ sending $z = [q(P), q(\psi)]$ to $[C, \phi]$. In the same manner we can define morphisms $\partial : \mathrm{W}^n(\mathcal{K}) \longrightarrow \mathrm{W}^{n+1}(\mathcal{J})$ fitting in a long exact sequence, the *localization sequence* associated to the exact sequence (2.1) of triangulated categories with duality:

$$\ldots \longrightarrow \mathrm{W}^n(\mathcal{K}) \longrightarrow \mathrm{W}^n(\mathcal{K}/\mathcal{J}) \xrightarrow{\partial} \mathrm{W}^{n+1}(\mathcal{J}) \longrightarrow \mathrm{W}^{n+1}(\mathcal{K}) \longrightarrow \ldots$$

3. Gersten-Witt spectral sequences and complexes

We review the Gersten-Witt spectral sequence, which was introduced by Balmer and Walter [5] for regular schemes, and generalized by Gille [7] to Gorenstein schemes of finite Krull dimension.

The construction: Let X be a regular scheme of finite Krull dimension and $Z \subset X$ a closed subset. Then $D^b_Z(\mathcal{P}_X)$ has a filtration

$$D^b_Z(\mathcal{P}_X) = D^0_Z \supset D^1_Z \supset \cdots \supset D^{\dim X}_Z \supset D^{\dim X + 1}_Z \simeq 0$$

where we have written $D^p_Z := D^b_Z(\mathcal{P}_X)^{(p)}$. The localization exact sequences

$$\ldots \longrightarrow \mathrm{W}^i(D^{p+1}_Z) \longrightarrow \mathrm{W}^i(D^p_Z) \longrightarrow \mathrm{W}^i(D^p_Z/D^{p+1}_Z) \longrightarrow \mathrm{W}^{i+1}(D^{p+1}_Z) \longrightarrow \ldots,$$

can be organized into an exact couple, giving rise to a convergent spectral sequence:

$$E_1^{p,q}(X, Z) = \mathrm{W}^{p+q}(D^p_Z/D^{p+1}_Z) \quad \Longrightarrow \quad \mathrm{W}^{p+q}_Z(X).$$

This is the *Gersten-Witt spectral sequence* for X with supports in Z.
Using the 4-periodicity of Witt groups, as well as [5] Theorem 7.2 and [7] Theorem 3.14, one sees that the E_1-page is zero everywhere except for the

lines with $q \equiv 0 \pmod 4$, which are all the same and which vanish outside the interval $\operatorname{codim}_X Z \le p \le \dim X$. So the information of the E_1-page is essentially given by the complex

$$0 \to E_1^{\operatorname{codim} Z,0}(X,Z) \to E_1^{\operatorname{codim} Z+1,0}(X,Z) \to \cdots \to E_1^{\dim X,0}(X,Z) \to 0.$$

Definition 3.1. Let X be a regular scheme of finite Krull dimension and $Z \subset X$ a closed subset. Then we define the complex $\mathrm{GWC}^\bullet(X,Z) := E_1^{\bullet,0}(X,Z)$. In other words, we have

$$\mathrm{GWC}^p(X,Z) = \mathrm{W}^p\left(D_Z^p/D_Z^{p+1}\right)$$

where $D_Z^p = D_Z^b(\mathcal{P}_X)^{(p)}$, and the differential $d^p = d_1^{p,0}$ is the composition

$$\mathrm{W}^p\left(D_Z^p/D_Z^{p+1}\right) \xrightarrow{\partial} \mathrm{W}^{p+1}\left(D_Z^{p+1}\right) \longrightarrow \mathrm{W}^{p+1}\left(D_Z^{p+1}/D_Z^{p+2}\right)$$

where ∂ is the connecting homomorphism of the localization long exact sequence and where the second homomorphism is the natural one. When $X = Z$ we write $E_1^{p,q}(X)$ instead of $E_1^{p,q}(X,X)$, and similarly for $\mathrm{GWC}^\bullet(X)$. We adopt the notation $\mathrm{GWC}^\bullet(X)$ to avoid confusion with the Grothendieck-Witt group $\mathrm{GW}(X)$.
Adding in the edge morphism $E^0(X) \to E_1^{0,0}(X)$ of the spectral sequence gives the *augmented Gersten-Witt complex* of X

$$\mathrm{GWC_a}(X): \quad 0 \longrightarrow \mathrm{W}(X) \longrightarrow \mathrm{GWC}^0(X) \longrightarrow \mathrm{GWC}^1(X) \longrightarrow \cdots$$

The E_2-page of the spectral sequence has $E_2^{p,0}(X) = H^p(\mathrm{GWC}^\bullet(X))$. From this we deduce the following result which will be used in the proofs of Theorems 4.4 and 6.1.

Lemma 3.2. *The following hold true:*

(1) *Let X be a regular scheme. Assume that $H^i(\mathrm{GWC}^\bullet(X)) = 0$ for all $i \ge 1$. Then the augmented Gersten-Witt complex for X is exact.*
(2) *Let R be a regular local ring. Assume only that $H^i(\mathrm{GWC}^\bullet(R)) = 0$ for all $i \ge 4$. Then the Gersten conjecture for Witt groups holds for R.*

Proof. We start with (1). The hypothesis implies that $E_2^{p,q}(X) = 0$ for all $p \ne 0$. It follows that the spectral sequence degenerates at E_2, and so the edge morphisms give isomorphisms $E^q(X) \xrightarrow{\sim} E_2^{0,q}(X)$ for all q. For $q = 0$ this means that the natural map $\mathrm{W}(X) \to H^0(\mathrm{GWC}^\bullet(X))$ is an isomorphism. This is what we needed to show.
For (2), recall from above that E_1 is concentrated in the lines $q \equiv 0 \bmod 4$. Therefore the spectral sequence degenerates again at $E_2 = E_5$ because no nonzero higher differentials can occur. Observe that for $p = 1,2,3$, the homology $H^p(\mathrm{GWC}^\bullet(R))$ is simply $E_2^{p,0}$ and the latter is isomorphic to $\mathrm{W}^p(R)$ by the convergence of the spectral sequence. Now, when R is local, we have $\mathrm{W}^p(R) = 0$ for $p = 1,2,3$ by [4] Theorem 5.6. So we can apply (1). ☐

Dévissage: The localization morphisms $\operatorname{Spec}\mathcal{O}_{X,x} \longrightarrow X$ induce an isomorphism (see [5] Proposition 7.1 or [7] Theorem 3.12)

$$\mathrm{GWC}^p(X,Z) \xrightarrow{\simeq} \bigoplus_{x\in X^{(p)}\cap Z} \mathrm{W}^p_{\mathfrak{m}_x}(\mathcal{O}_{X,x}), \tag{3.1}$$

where $X^{(p)}$ is the set of points of codimension p and $(\mathcal{O}_{X,x}, \mathfrak{m}_x)$ is the local ring at $x \in X$.

Let $(S,\mathfrak{n})$ be a regular local ring and $\ell = S/\mathfrak{n}$. Then we have an isomorphism $\mathrm{W}(\ell) \xrightarrow{\sim} \mathrm{W}^{\dim S}_{\mathfrak{n}}(S)$ which depends on the choice of local parameters (see [5] Theorem 6.1) and hence

$$\mathrm{GWC}^p(X,Z) \;\simeq\; \bigoplus_{x\in X^{(p)}\cap Z} \mathrm{W}(\kappa(x))$$

for any regular scheme X with closed subset Z. It follows that our Gersten-Witt complex has the form announced in the Introduction.

There are two ways of having a complex independent of choices. Either work as we do here in Definition 3.1 with the underlying complex before *dévissage*, or twist the dualities on the residue fields, see [5]. The latter means that one can consider for each $x \in X^{(p)}$ the Witt group $\mathrm{W}(\kappa(x), \omega_{x_p/X})$ with twisted coefficients in the one-dimensional $\kappa(x)$-vector space $\omega_{x_p/X} := \mathrm{Ext}^p(\kappa(x), \mathcal{O}_{X,x}) = \Lambda^p(\mathfrak{m}_x/\mathfrak{m}_x^2)^*$, and we have then a canonical isomorphism $\mathrm{W}(\kappa(x), \omega_{x_p/X}) \cong \mathrm{W}^p_{\mathfrak{m}_x}(\mathcal{O}_{X,x})$ and thus a canonical Gersten-Witt complex with

$$\mathrm{GWC}^p(X,Z) \;\cong\; \bigoplus_{x\in X^{(p)}\cap Z} \mathrm{W}(\kappa(x), \omega_{x_p/X}).$$

After all, this *dévissage* is only relevant for cognitive reasons since it relates the terms of the Gersten-Witt complex with quadratic forms over the residue fields. But we will see in the sequel that our initial canonical definition $\mathrm{GWC}^p(X,Z) = \mathrm{W}^p\left(D^p_Z/D^{p+1}_Z\right)$ is more convenient to handle.

Another construction of a Gersten Witt spectral sequence has been given by Gille [7], Section 3. Let Y be a Gorenstein scheme of finite Krull dimension and $Z \subset Y$ a closed subset. Then the derived functor of $\mathcal{H}\mathrm{om}_{\mathcal{O}_Y}(-,\mathcal{O}_Y)$ is a duality on $D^b_{\mathrm{coh},\,Z}(\mathcal{M}_Y)$ making it a triangulated category with 1-exact duality. Following [7] we denote the associated so called *coherent Witt groups* by $\tilde{\mathrm{W}}^i_Z(Y)$. On the triangulated category $D^b_{\mathrm{coh}}(\mathcal{M}_Y)$ we have also a finite filtration

$$D^b_{\mathrm{coh},\,Z}(\mathcal{M}_Y) = D^0_Z \supset D^1_Z \supset D^2_Z \supset \ldots \supset D^{\dim Y}_Z\,,$$

where $D^p_Z := D^b_{\mathrm{coh},\,Z}(\mathcal{M}_Y)^{(p)}$. As above this gives us long exact sequences

$$\ldots \longrightarrow \mathrm{W}^i(D^{p+1}_Z) \longrightarrow \mathrm{W}^i(D^p_Z) \longrightarrow \mathrm{W}^i(D^p_Z/D^{p+1}_Z) \longrightarrow \mathrm{W}^{i+1}(D^{p+1}_Z) \longrightarrow \ldots\,,$$

and hence by Massey's method of exact couples a convergent spectral sequence

$$\tilde{E}^{p,q}(Y,Z) := \mathrm{W}^{p+q}(D^{p+q}_Z/D^{p+q+1}_Z) \implies \tilde{\mathrm{W}}^{p+q}_Z(Y)\,.$$

If now Y is regular we have equivalences $D^b_Z(\mathcal{P}_Y)^{(p)} \xrightarrow{\simeq} D^b_{\mathrm{coh},\,Z}(\mathcal{M}_Y)^{(p)}$ which are duality preserving and hence give isomorphisms $\mathrm{W}^i(D^b_Z(\mathcal{P}_Y)^{(p)}) \simeq \mathrm{W}^i(D^b_{\mathrm{coh},\,Z}(\mathcal{M}_Y)^{(p)})$. We get then from the functorial properties of the localization sequence an isomorphism of spectral sequences $E_1^{p,q}(Y,Z) \xrightarrow{\simeq} \tilde{E}_1^{p,q}(Y,Z)$. Hence in the regular case both constructions lead to the same result.

One advantage of this "coherent" approach is the following. Let $Y = \operatorname{Spec} R$ with R a Gorenstein ring of finite Krull dimension and let the closed subset Z be defined by a regular element f of R, *i.e.* $Z = \operatorname{Spec} R/Rf$. We set $\bar{D}^p := D^b_{\mathrm{coh}}(\mathcal{M}_Z)^{(p)}$ and as before $D^p_Z := D^b_{\mathrm{coh},\,Z}(\mathcal{M}_Y)^{(p)}$. The natural morphism $\alpha : Z \hookrightarrow Y$ induces a pushforward functor $\alpha_* : \bar{D}^p \to D^{p+1}_Z$ for any $p \in \mathbb{N}$. This functor shifts the duality structure by 1 (*cf.* [7], Theorem 4.2), *i.e.* it induces morphisms $\mathrm{W}^i(\bar{D}^p) \to \mathrm{W}^{i+1}(D^{p+1}_Z)$ for all $i \in \mathbb{Z}$ and $p \in \mathbb{N}$. From the functoriality of the localization sequence (*cf.* [7] Theorem 2.9) we get commutative diagrams with exact rows

$$\begin{tikzcd}
\cdots\ \mathrm{W}^i(\bar{D}^{p+1}) \arrow[r] \arrow[d] & \mathrm{W}^i(\bar{D}^p) \arrow[r] \arrow[d] & \mathrm{W}^i(\bar{D}^p/\bar{D}^{p+1}) \arrow[r] \arrow[d] & \mathrm{W}^{i+1}(\bar{D}^{p+1})\ \cdots \arrow[d] \\
\cdots\ \mathrm{W}^{i+1}(D^{p+2}_Z) \arrow[r] & \mathrm{W}^{i+1}(D^{p+1}_Z) \arrow[r] & \mathrm{W}^{i+1}(D^{p+1}_Z/D^{p+2}_Z) \arrow[r] & \mathrm{W}^{i+2}(D^{p+2}_Z)\ \cdots
\end{tikzcd}$$

(*cf.* [7], diagram on the bottom of p. 130). In particular we have a morphism of spectral sequences $\tilde{\alpha}_* : \tilde{E}_1^{p,q}(Z) \longrightarrow \tilde{E}_1^{p+1,q}(Y,Z)$ which is an isomorphism as shown in [7], Section 4.2.3.

If now R is regular local and f a regular parameter, *i.e.* R/Rf is regular too, the identification above gives the following

Lemma 3.3. *Let R be a regular local ring and f a regular parameter. Then we have an isomorphism of spectral sequences*

$$E_r^{p,q}(R/fR) \xrightarrow{\simeq} E_r^{p+1,q}(R,fR).$$

In particular, we have isomorphisms of complexes

$$\mathrm{GWC}^\bullet(R/fR) \xrightarrow{\simeq} \mathrm{GWC}^{\bullet+1}(R,fR)\,.$$

4. A reformulation of the conjecture

Let X be a regular scheme and Z a closed subset. From Definition 3.1 and from the dévissage formula (3.1), we immediately obtain a degree-wise split short exact sequence

$$0 \to \mathrm{GWC}^\bullet(X,Z) \longrightarrow \mathrm{GWC}^\bullet(X) \longrightarrow \mathrm{GWC}^\bullet(X \setminus Z) \to 0.$$

We will consider below the long exact cohomology sequence of this short exact sequence of complexes in the case $X = \operatorname{Spec} R$, for R a regular local ring, and Z is defined by a regular parameter f.

Definition 4.1. Recall from the introduction that the Gersten conjecture asserts that for a regular local ring R the Gersten complex $\mathrm{GWC}^\bullet(R)$ is an exact resolution of $\mathrm{W}(R)$.
We denote by $\mathcal{W}$ the *Witt sheaf*, *i.e.* the sheafification of the presheaf on the Zariski site

$$U \longmapsto \mathrm{W}(U) \qquad U \subset X \qquad \text{open.}$$

We have also a Gersten-Witt complex $\mathcal{GWC}^\bullet$ of sheaves on any regular scheme of finite Krull dimension X. The definition of this complex in degree $p \geq 0$ is:

$$U \longmapsto \mathrm{GWC}^p(U) \qquad U \subset X \qquad \text{open.}$$

Lemma 4.2. *Let X be a regular scheme of finite Krull dimension and assume that the Gersten conjecture holds for all local rings $\mathcal{O}_{X,x}$ of X. Then for all i we have*

$$\mathrm{H}^i_{\mathrm{Zar}}(X, \mathcal{W}) \simeq H^i(\Gamma(X, \mathcal{GWC}^\bullet)) = H^i(\mathrm{GWC}^\bullet(X)).$$

Proof. Note that $\mathcal{GWC}^p$ is a flabby sheaf and that $(\mathcal{GWC}^p)_x \simeq \mathrm{GWC}^p(\mathcal{O}_{X,x})$ for all points x of the scheme X. Since the natural morphism $\mathrm{W}(\mathcal{O}_{X,x}) \longrightarrow \mathcal{W}_x$ is an isomorphism for all $x \in X$ it follows that if the Gersten conjecture is true for every local ring of the regular scheme X, then $\mathcal{GWC}^\bullet$ is a flabby resolution of $\mathcal{W}$ on X. □

Definition 4.3. Let $\mathcal{C}$ be a class of regular local rings. We say that $\mathcal{C}$ is *nepotistic* if the following holds: whenever R belongs to $\mathcal{C}$, so do $R_\mathfrak{p}$ for all $\mathfrak{p} \in \operatorname{Spec} R$ and R/fR for all regular parameters $f \in R$.

The main result of this section is the following:

Theorem 4.4. *If $\mathcal{C}$ is a nepotistic class of regular local rings* (*see 4.3*), *then the following conditions are equivalent*:

(i) *The Gersten conjecture for Witt groups is true for any $R \in \mathcal{C}$.*
(ii) *For any $R \in \mathcal{C}$ and for any regular parameter $f \in R$, we have for all $i \geq 1$ that $\mathrm{H}^i_{\mathrm{Zar}}(\operatorname{Spec} R_f, \mathcal{W}) = 0$. (When R is a field, this condition is empty and thus always true.)*

Proof. The short exact sequence $0 \to \mathrm{GWC}^\bullet(R, fR) \to \mathrm{GWC}^\bullet(R) \to \mathrm{GWC}^\bullet(R_f) \to 0$ of Gersten-Witt complexes gives rise to a long exact sequence of cohomology

$$0 \to H^0(\mathrm{GWC}^\bullet(R)) \longrightarrow H^0(\mathrm{GWC}^\bullet(R_f)) \xrightarrow{\delta} H^1(\mathrm{GWC}^\bullet(R, fR)) \to \cdots$$

(i) ⇒ (ii). Let $R \in \mathcal{C}$, and let $f \in R$ be a regular parameter. Then the Gersten conjecture for Witt groups holds for R and R/fR, and so we have $H^i(\mathrm{GWC}^\bullet(R)) = 0$ and $H^i(\mathrm{GWC}^\bullet(R/fR)) = 0$ for all $i \geq 1$. Because of the isomorphism of Lemma 3.3, we get $H^i(\mathrm{GWC}^\bullet(R, fR)) = 0$ for all $i \geq 2$. It now follows from the long exact sequence that $H^i(\mathrm{GWC}^\bullet(R_f)) = 0$ for all $i \geq 1$. The local rings of $\operatorname{Spec} R_f$ are the $R_\mathfrak{p}$ with $f \notin \mathfrak{p}$, so they are all in $\mathcal{C}$. So we also have $\mathrm{H}^i_{\mathrm{Zar}}(\operatorname{Spec} R_f, \mathcal{W}) = 0$ for all $i \geq 1$ by Lemma 4.2.

(ii) ⇒ (i). We will prove the Gersten conjecture for Witt groups for $R \in \mathcal{C}$ by induction on $n = \dim R$. For $n = 0$ the ring R is a field, and this is trivial.
So suppose $n \geq 1$ and that the Gersten conjecture for Witt groups is true for all $S \in \mathcal{C}$ with $\dim S < n$. Let $f \in \mathfrak{m} \setminus \mathfrak{m}^2$ be a regular parameter. The local rings of R_f are the local rings $R_\mathfrak{p}$ for primes with $f \notin \mathfrak{p}$, and they satisfy $R_\mathfrak{p} \in \mathcal{C}$ and $\dim R_\mathfrak{p} < n$. So the Gersten conjecture holds for all local rings of $\operatorname{Spec} R_f$, and so by (ii) and Lemma 4.2 we have $H^i(\mathrm{GWC}^\bullet(R_f)) = 0$ for all $i \geq 1$. We also have $R/fR \in \mathcal{C}$ with $\dim R/fR = n-1$, so we get $H^i(\mathrm{GWC}^\bullet(R/fR)) = 0$ for all $i \geq 1$. The identification of Lemma 3.3 now gives us $H^i(\mathrm{GWC}^\bullet(R, fR)) = 0$ for all $i \geq 2$. So from the long exact sequence we get $H^i(\mathrm{GWC}^\bullet(R)) = 0$ for all $i \geq 2$ and the Gersten conjecture for Witt groups holds for R by Lemma 3.2, part (ii). □

5. A pairing, a trace map and a projection formula

We recall here some techniques of Gille and Nenashev [8] that we shall use below. For more details and a more general point of view see [8].

The pairing $\mathrm{W}(k) \times \mathrm{W}^i(D^b(\mathcal{P}_R)) \longrightarrow \mathrm{W}^i(D^b(\mathcal{P}_R))$. Let k be a field of characteristic not 2, and let R be a regular k-algebra of Krull dimension n. Denote the duality on k-mod by $V^* = \mathrm{Hom}_k(V,k)$ and that on $D^b(\mathcal{P}_R)$ by $F^\sharp := \mathrm{Hom}_R(F,R)$. Let (V,φ) be a nondegenerate symmetric bilinear space over k. Then $V \otimes_k - :\ D^b(\mathcal{P}_R) \longrightarrow D^b(\mathcal{P}_R)$ is an exact functor, and the system of isomorphisms between $V \otimes_k F^\sharp$ and $(V \otimes_k F)^\sharp \cong V^* \otimes_k F^\sharp$ given by $\varphi \otimes 1_{F^\sharp}$ makes $V \otimes_k -$ duality-preserving. Actually, the duality-preserving functor is formally the pair $(V \otimes_k -\,,\ \varphi \otimes 1)$, and we will abbreviate it as $(V,\varphi) \otimes_k -$. Moreover, if we let $D^p_R = D^b(\mathcal{P}_R)^{(p)}$ be the subcategory of complexes of homological support of codimension at least p, then $(V,\varphi) \otimes_k -$ is compatible with the filtration

$$D^b(\mathcal{P}_R) = D^0_R \supset D^1_R \supset \cdots \supset D^{n+1}_R \simeq 0.$$

Hence the maps $\mathrm{W}^i(D^p_R) \longrightarrow \mathrm{W}^i(D^p_R)$ and $\mathrm{W}^i(D^p_R/D^{p+1}_R) \longrightarrow \mathrm{W}^i(D^p_R/D^{p+1}_R)$ induced by $(V,\varphi)\otimes_k -$ are compatible with the localization exact sequences and induce endomorphisms of the Gersten-Witt exact couple and spectral sequence for R. These endomorphisms depend only on the Witt class of (V,φ), and they are compatible with the orthogonal direct sum $(V,\varphi) \perp (W,\psi)$ and tensor product $(V \otimes W, \varphi \otimes \psi)$ of symmetric bilinear spaces over k. This gives us the following result.

Lemma 5.1. *If R is a regular k-algebra of finite Krull dimension, then the pairing makes the Gersten-Witt spectral sequence $E^{p,q}_r(R)$ into a spectral sequence of $\mathrm{W}(k)$-modules.* □

Base change. Let ℓ/k be a separable algebraic field extension. Denote by π the projection $\pi : \operatorname{Spec}(\ell \otimes_k R) \to \operatorname{Spec} R$. Then $\ell \otimes_k R$ is a regular ℓ-algebra of Krull dimension n (see [9], and in particular Prop. 6.7.4, p. 146 for regularity). Moreover, $D^b(\mathcal{P}_{\ell\otimes_k R})$ has a duality $(-)^\flat$ given by $G^\flat := \operatorname{Hom}_{\ell\otimes_k R}(G, \ell \otimes_k R)$, and the exact functor $\ell \otimes_k - : D^b(\mathcal{P}_R) \longrightarrow D^b(\mathcal{P}_{\ell\otimes_k R})$ is naturally duality-preserving. It is compatible with the filtration of $D^b(\mathcal{P}_R)$ and the corresponding filtration

$$D^b(\mathcal{P}_{\ell\otimes_k R}) = D^0_{\ell\otimes_k R} \supset D^1_{\ell\otimes_k R} \supset \cdots \supset D^{n+1}_{\ell\otimes_k R} = 0.$$

of $D^b(\mathcal{P}_{\ell\otimes_k R})$, so it induces a morphism of spectral sequences $\pi^* : E_r^{p,q}(R) \longrightarrow E_r^{p,q}(\ell \otimes_k R)$.

A trace map. If ℓ/k is finite and separable, then choose $0 \neq \tau \in \operatorname{Hom}_k(\ell, k)$ and extend it to an R-linear map $\tau_R : \ell \otimes_k R \longrightarrow R$ by setting $\tau_R(l \otimes a) = \tau(l)a$. Let $\pi_* : \mathcal{P}_{\ell\otimes_k R} \to \mathcal{P}_R$ be the restriction-of-scalars functor. The natural homomorphism of R-modules

$$\pi_*(G^\flat) = \operatorname{Hom}_{\ell\otimes_k R}(G, \ell \otimes_k R) \xrightarrow{\tau_{R*}} \operatorname{Hom}_R(G, R) = (\pi_* G)^\sharp$$

sending $f \mapsto \tau_R \circ f$ is an isomorphism for any G in $\mathcal{P}_{\ell\otimes_k R}$, and it makes π_* into a duality-preserving exact functor $\pi_* : D^b(\mathcal{P}_{\ell\otimes_k R}) \to D^b(\mathcal{P}_R)$. Actually, the duality-preserving functor is formally the pair (π_*, τ_{R*}), and we will abbreviate it as $\operatorname{Tr}^\tau_{\ell\otimes_k R/R}$. Since $\ell \otimes_k R$ is flat and finite over R, the restriction-of-scalars functor preserves the codimension of the support of the homology modules, and so π_* is compatible with the filtrations on the two derived categories. So we again get a morphism of spectral sequences $\operatorname{Tr}^\tau_{\ell\otimes_k R/R} : E_r^{p,q}(\ell \otimes_k R) \longrightarrow E_r^{p,q}(R)$.

Remark 5.2. For $R = k$ and $i = 0$ our $\operatorname{Tr}^\tau_{\ell/k}$ is just the Scharlau transfer $\tau_* : \mathrm{W}(\ell) \longrightarrow \mathrm{W}(k)$ (*cf.* [20] Section 2.5).

Let (U, ψ) be a nondegenerate symmetric bilinear space over ℓ. The following diagram of duality-preserving functors commutes up to isomorphism of duality-preserving functors (see [5] §4 for the definition):

$$\begin{tikzcd} D^b(\mathcal{P}_R) \arrow[r, "\ell \otimes_k -"] \arrow[d, "\operatorname{Tr}^\tau_{\ell/k}(U,\psi) \otimes_k -"'] & D^b(\mathcal{P}_{\ell\otimes_k R}) \arrow[d, "(U,\psi) \otimes_\ell -"] \\ D^b(\mathcal{P}_R) & D^b(\mathcal{P}_{\ell\otimes_k R}) . \arrow[l, "\operatorname{Tr}^\tau_{\ell\otimes_k R/R}"'] \end{tikzcd}$$

The induced maps on derived Witt groups, exact couples, and spectral sequences are then the same (*cf.* [5] Lemma 4.1(b)). This gives us a projection formula (*cf.* [8] Theorem 4.1):

Theorem 5.3. *Let R be a regular k-algebra of finite Krull dimension, let ℓ/k be a finite separable extension of fields, and let $\pi : \operatorname{Spec}(\ell \otimes_k R) \to \operatorname{Spec} R$ be* *the projection. Let $E_r^{p,q}(R)$ and $E_r^{p,q}(\ell \otimes_k R)$ be the two Gersten-Witt spectral sequences. Then*

$$\mathrm{Tr}^{\tau}_{\ell\otimes_k R/R}(u \cdot \pi^*(x)) = \mathrm{Tr}^{\tau}_{\ell/k}(u) \cdot x,$$

for all $u \in \mathrm{W}(\ell)$ and all $x \in E_r^{p,q}(R)$. $\square$

Odd-degree extensions. If ℓ/k is of odd degree, then there exists a $\tau \in \mathrm{Hom}_k(\ell, k)$ such that $\mathrm{Tr}^{\tau}_{\ell/k}(1_{\mathrm{W}(\ell)}) = 1_{\mathrm{W}(k)}$ ([20] Lemma 2.5.8). We then have $\mathrm{Tr}^{\tau}_{\ell\otimes_k R/R}(\pi^* x) = x$ for all $x \in E_r^{p,q}(R)$. In other words:

Corollary 5.4. *If ℓ/k is separable of odd degree, then $\pi^* : E_r^{p,q}(R) \to E_r^{p,q}(\ell \otimes_k R)$ is a split monomorphism of spectral sequences. In particular $H^i(\mathrm{GWC}^\bullet(R)) \to H^i(\mathrm{GWC}^\bullet(\ell \otimes_k R))$ is a split monomorphism for every i.* $\square$

The base change maps for separable algebraic extensions commute with filtered colimits.

Corollary 5.5. *If ℓ/k is a filtered colimit of separable finite extensions of odd degree, then $\pi^* : E_r^{p,q}(R) \to E_r^{p,q}(\ell \otimes_k R)$ is a filtered colimit of split monomorphisms of spectral sequences. In particular $H^i(\mathrm{GWC}^\bullet(R)) \to H^i(\mathrm{GWC}^\bullet(\ell \otimes_k R))$ is a monomorphism for every i.* $\square$

6. The equicharacteristic case of the Gersten conjecture for Witt groups

We are now ready to prove the main result of the paper.

Theorem 6.1. *Let R be an equicharacteristic regular local ring, i.e. R contains some field* (*of characteristic not* 2). *Then the Gersten conjecture for Witt groups is true for R.*

Fix the following notation: $\mathfrak{m}$ is the maximal ideal of R. We prove the theorem in two steps.

Step 1. Assume that R is essentially smooth over some field k.

When k is an infinite field, this is a special case of [4] Theorem 4.3 which states the following. If S is a semilocal ring essentially smooth over a field ℓ (*i.e.* S is the semi-localization of a smooth scheme over ℓ), and if the field ℓ is infinite, then the Gersten conjecture for Witt groups is true for S, *i.e.* $H^i(\mathrm{GWC}(S)) = 0$ for all $i \geq 1$.

Assume now k is a finite field and hence perfect. Fix an odd prime s. For any $n \geq 0$ let ℓ_n be the unique extension of the finite field k of degree s^n, and let $\ell = \bigcup_{n=0}^{\infty} \ell_n$. The ℓ-algebra $\ell \otimes_k R$ is integral over R and hence semilocal. It is further essentially smooth over the infinite field ℓ ([10] Prop. 10.1.b) and hence by the result above the Gersten conjecture is true for $\ell \otimes_k R$. Using Corollary 5.5, we see that the same is true for R.

Step 2. General equicharacteristic R.

For this case we use the following result:

Theorem 6.2. *Let R be an equicharacteristic regular local ring, and let $f \in \mathfrak{m} \setminus \mathfrak{m}^2$ be a regular parameter. Then*:

(1) *There exist a perfect field k contained in R and a filtered system of pairs (R_j, f_j) such that each R_j is an essentially smooth local k-algebra, each f_j is a regular parameter in R_j, and such that $R = \operatorname{colim} R_j$ and $R_f = \operatorname{colim}(R_j)_{f_j}$, and the morphisms $R_j \to R$ are local.*

(2) *In addition the natural maps*

$$\operatorname{colim}_j \ \mathrm{H}^i_{\mathrm{Zar}}(\operatorname{Spec}(R_j)_{f_j}, \mathcal{W}) \longrightarrow \mathrm{H}^i_{\mathrm{Zar}}(\operatorname{Spec} R_f, \mathcal{W})$$

are isomorphisms for all $i \geq 0$.

Proof. The first part is a consequence of Popescu's Theorem [18] [19] (see also [15] §3), while the second part follows from [15] Theorem 6.6, which was inspired by the étale analogue of this result: [1] Exposé VII, Théorème 5.7. □

Let $\mathcal{C}_{\mathrm{eq}}$ be the class of all equicharacteristic regular local rings, and let $\mathcal{C}_{\mathrm{sm}}$ be the subclass of regular local rings that are essentially smooth over a field. Both $\mathcal{C}_{\mathrm{eq}}$ and $\mathcal{C}_{\mathrm{sm}}$ are nepotistic (Definition 4.3). The class $\mathcal{C}_{\mathrm{sm}}$ satisfies condition (i) of Theorem 4.4 by the first step of our proof, and we wish to show that $\mathcal{C}_{\mathrm{eq}}$ satisfies the same condition. But conditions (i) and (ii) of Theorem 4.4 are equivalent, so it is enough to show that $\mathcal{C}_{\mathrm{eq}}$ satisfies condition (ii) of Theorem 4.4, knowing that $\mathcal{C}_{\mathrm{sm}}$ satisfies the same condition.

Let R be in $\mathcal{C}_{\mathrm{eq}}$, and let f be a regular parameter of R. By Theorem 6.2 there exist a perfect subfield $k \subset R$ and a filtered system (R_j, f_j) of essentially smooth local k-algebras R_j plus regular parameters $f_j \in R_j$ such that $R_f = \operatorname{colim}(R_j)_{f_j}$. Since the R_j are in $\mathcal{C}_{\mathrm{sm}}$, we have $\mathrm{H}^i_{\mathrm{Zar}}(\operatorname{Spec}(R_j)_{f_j}, \mathcal{W}) = 0$ for all $i \geq 1$ and all j because $\mathcal{C}_{\mathrm{sm}}$ satisfies condition (ii) of Theorem 4.4. Then by Theorem 6.2 (2) we also have $\mathrm{H}^i_{\mathrm{Zar}}(\operatorname{Spec} R_f, \mathcal{W}) = 0$ for $i \geq 1$. This is condition (ii) of Theorem 4.4 for the class $\mathcal{C}_{\mathrm{eq}}$, so we have completed the proof.

References

[1] M. Artin, A. Grothendieck, J.-L. Verdier, *Théorie des Topos et Cohomologie Étale des Schémas (SGA 4). Tome 2*, Lecture Notes in Math., 270, Springer-Verlag, 1972.

[2] P. Balmer, *Triangular Witt groups. Part I: The 12-term localization exact sequence*, K-Theory 19 (2000), 311–363.

[3] P. Balmer, *Triangular Witt groups. Part II: From usual to derived*, Math. Z. 236 (2001), 351–382.

[4] P. Balmer, *Witt cohomology, Mayer-Vietoris, homotopy invariance, and the Gersten conjecture*, K-Theory 23 (2001), 15–30.

[5] P. Balmer, C. Walter, *A Gersten-Witt spectral sequence for regular schemes*, Ann. Scient. Éc. Norm. Sup. (4) 35 (2002), 127–152.

[6] J.-L. Colliot-Thélène, R.T. Hoobler, B. Kahn, *The Bloch-Ogus-Gabber Theorem*, Algebraic K-theory (Toronto, 1996), 31–94, Fields Inst. Commun., 16, Amer. Math. Soc., Providence, 1997.

[7] S. Gille, *On Witt groups with support*, Math. Ann. 322 (2002), 103–137.

[8] S. Gille, A. Nenashev, *Pairings in triangular Witt theory* (2001) to appear in J. Algebra.

[9] A. Grothendieck, *Élements de géométrie algébrique* IV_2 , Publ. Math. IHES 24 (1965).

[10] R. Hartshorne, *Algebraic geometry*, Springer-Verlag, 1977.

[11] M. Knebusch, *Symmetric bilinear forms over algebraic varieties*, Conference on Quadratic Forms (Kingston, 1976), 103–283, Queen's Papers in Pure and Appl. Math., No. 46, Queen's Univ., Kingston, 1977.

[12] J. Milnor, D. Husemoller, *Symmetric Bilinear Forms*, Springer-Verlag, 1973.

[13] M. Ojanguren, *Quadratic forms over regular rings*, J. Indian Math. Soc. (N.S.) 44 (1980), 109–116.

[14] M. Ojanguren, I. Panin, *A purity theorem for the Witt group*, Ann. Scient. Éc. Norm. Sup. (4) 32 (1999), 71–86.

[15] I. Panin, *The equi-characteristic case of the Gersten conjecture*, preprint (2000), available on the server: `http://www.math.uiuc.edu/K-theory/0389/`.

[16] W. Pardon, *A "Gersten conjecture" for Witt groups*, Algebraic K-theory, Part II (Oberwolfach, 1980), pp. 300–315, Lecture Notes in Math., 967, Springer-Verlag, 1982.

[17] W. Pardon, *The filtered Gersten-Witt resolution for regular schemes*, preprint, 2000, available at `http://www.math.uiuc.edu/K-theory/0419/`

[18] D. Popescu, *General Néron desingularization*, Nagoya Math. J. 100 (1985), 97–126.

[19] D. Popescu, *General Néron desingularization and approximation*, Nagoya Math. J. 104 (1986), 85–115.

[20] W. Scharlau, *Quadratic and hermitian forms*, Springer-Verlag, 1985.

Paul Balmer
D-Math
ETH Zentrum
8092 Zürich
Switzerland
balmer@math.ethz.ch
www.math.ethz.ch/~balmer

Stefan Gille
SFB 478
Universität Münster
Hittorfstrasse 27
48149 Münster
Germany
gilles@math.uni-muenster.de

Ivan Panin
Laboratory of Algebra
St-Petersburg Department of
Steklov Mathematical Institute
Fontanka 27
St-Petersburg 191011
Russia
panin@pdmi.ras.ru

Charles Walter
Laboratoire Dieudonné
Université de Nice
– Sophia Antipolis
06108 Nice Cedex 02
France
walter@math.unice.fr
www-math.unice.fr/~walter

L^2-Invariants of Locally Symmetric Spaces

Martin Olbrich

Received: August 9, 2001

Communicated by Ursula Hamenstädt

Abstract. Let $X = G/K$ be a Riemannian symmetric space of the noncompact type, $\Gamma \subset G$ a discrete, torsion-free, cocompact subgroup, and let $Y = \Gamma\backslash X$ be the corresponding locally symmetric space. In this paper we explain how the Harish-Chandra Plancherel Theorem for $L^2(G)$ and results on $(\mathfrak{g}, K)$-cohomology can be used in order to compute the L^2-Betti numbers, the Novikov-Shubin invariants, and the L^2-torsion of Y in a uniform way thus completing results previously obtained by Borel, Lott, Mathai, Hess and Schick, Lohoue and Mehdi. It turns out that the behaviour of these invariants is essentially determined by the fundamental rank $m = \mathrm{rk}_{\mathbb{C}} G - \mathrm{rk}_{\mathbb{C}} K$ of G. In particular, we show the nonvanishing of the L^2-torsion of Y whenever $m = 1$.

2000 Mathematics Subject Classification: 58J35, 57R19, 22E46
Keywords and Phrases: locally symmetric spaces, L^2-cohomology, Novikov-Shubin invariants, L^2-torsion, relative Lie algebra cohomology

1 Introduction

During the last two decades L^2-invariants have proved to be a powerful tool in the topology of compact manifolds (see [15] for an overview). Although they can be defined in purely combinatorial terms we are interested here in their equivalent analytic versions: They are spectral invariants of the p-form Laplacians of the universal cover of the manifold. For particular nice manifolds these might be computable. Indeed, the aim of the present paper is to extract from the representation theoretic work of Harish-Chandra [9] and Borel-Wallach [2] information on the spectral decomposition of the form Laplacians on Riemannian symmetric spaces of the non-compact type which is sufficiently explicit in order to compute the spectral invariants of interest. We try to do this in a rather detailed way which, we hope, keeps the paper readable for nonspecialists in harmonic analysis. [1]

Let $X \to Y$ be the universal cover of a compact Riemannian manifold. Set $\Gamma := \pi_1(Y)$. The form Laplacian $\Delta_p = d^*d + dd^*$ defines a non-negative, elliptic, self-adjoint operator acting on $L^2(X, \Lambda^p T^*X)$, the square integrable p-forms on X. By Δ_p' and $\Delta_p^c = (d^*d)'$ we denote the restriction of Δ_p to the orthogonal complement of its kernel and the coclosed forms in this orthogonal complement, respectively. We consider the corresponding heat kernels $e^{-t\Delta_p^*}(x,x') := (P_* e^{-t\Delta_p})(x,x')$, $x, x' \in X$, for $* = \emptyset, \prime$ or c. Here P_* denotes the orthogonal projection to the corresponding subspace. The local traces $\operatorname{tr} e^{-t\Delta_p^*}(x,x)$ are Γ-invariant functions on X. For a thorough discussion of the following definitions we refer to [14], [17], [16], and [15].

We set

$$\operatorname{Tr}_\Gamma e^{-t\Delta_p^*} := \int_F \operatorname{tr} e^{-t\Delta_p^*}(x,x)\, dx \ ,$$

where $F \subset X$ is a fundamental domain of the action of Γ on X and dx is the Riemannian volume element of X. Then the L^2-Betti numbers are given by

$$b_p^{(2)}(Y) := \lim_{t\to\infty} \operatorname{Tr}_\Gamma e^{-t\Delta_p} \in [0,\infty) \ .$$

They are equal to the von Neumann dimension of $\ker \Delta_p$ viewed as Hilbert $\mathcal{N}(\Gamma)$-module, where $\mathcal{N}(\Gamma)$ is the group von Neumann algebra of Γ. If the spectrum of Δ_p has no gap around 0 the Novikov-Shubin invariants of Y are defined by

$$\tilde{\alpha}_p(Y) := \sup\{\beta \mid \operatorname{Tr}_\Gamma e^{-t\Delta_p'} = O(t^{-\frac{\beta}{2}}) \text{ as } t \to \infty\} \in [0,\infty] \ .$$

It measures the asymptotic behaviour of the spectral density function of Δ_p at 0. In case of a gap around 0 we set $\tilde{\alpha}_p(Y) := \infty^+$. Replacing Δ_p' by Δ_p^c we obtain the analogously defined Novikov-Shubin invariants $\alpha_p(Y)$ of d^*d. Using the action of the exterior differential d on the Hodge decomposition of L^2-forms we obtain

$$\tilde{\alpha}_p(Y) = \min\{\alpha_p(Y), \alpha_{p-1}(Y)\} \ . \tag{1}$$

Finally, if $\alpha_p(Y) > 0$ for all p (or, more generally, if X is of determinant class (see [15])), then the L^2-torsion of Y is defined by

$$\rho^{(2)}(Y) := \frac{1}{2}\sum (-1)^{p+1} p \, \log \det\nolimits_\Gamma(\Delta_p') = \frac{1}{2}\sum (-1)^p \log \det\nolimits_\Gamma(\Delta_p^c) \ ,$$

where for $* = \prime, c$

$$\begin{aligned} -\log \det\nolimits_\Gamma(\Delta_p^*) \quad &:= \quad \frac{d}{ds}_{|s=0} \left(\frac{1}{\Gamma(s)} \int_0^\varepsilon \operatorname{Tr}_\Gamma e^{-t\Delta_p^*} t^{s-1} dt \right) \\ &\qquad + \int_\varepsilon^\infty \operatorname{Tr}_\Gamma e^{-t\Delta_p^*} t^{-1} dt \end{aligned}$$

[1] *Note added in proof*: In the meanwhile the interested reader can find a discussion without proofs of the results of the present paper in the recent monograph [16] which gives a comprehensive treatment of the theory of L^2-invariants.

for any $\varepsilon > 0$, where the first integral is considered as a meromorphic function in s. By Poincaré duality $\rho^{(2)}(Y) = 0$ for even dimensional manifolds Y.
From now on let $X = G/K$ be a Riemannian symmetric space of the noncompact type. Here G is a real, connected, linear, semisimple Lie group without compact factors, and $K \subset G$ is a maximal compact subgroup. It is the universal cover of compact locally symmetric spaces of the form $Y = \Gamma\backslash X$, where $\Gamma \cong \pi_1(Y)$ can be identified with a discrete, torsion-free, cocompact subgroup of G. It will be convenient to consider also the compact dual X^d of X. X^d is defined as follows: Let $\mathfrak{g}$, $\mathfrak{k}$ be the Lie algebras of G, K. Then we have the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, and $\mathfrak{g}^d := \mathfrak{k} \oplus i\mathfrak{p}$ is another subalgebra of the complexification of $\mathfrak{g}$. Let G^d be the corresponding analytic subgroup of the complexification $G_{\mathbb{C}}$ of G. Then G^d is a compact group, and $X^d = G^d/K$. We normalize the Riemannian metric on X^d such that multiplication by i becomes an isometry $T_{eK}X \cong \mathfrak{p} \to i\mathfrak{p} \cong T_{eK}X^d$. In this paper we are going to prove the following theorem:

Theorem 1.1 *Let $n = \dim Y$ and $m = m(X) := \mathrm{rk}_{\mathbb{C}}G - \mathrm{rk}_{\mathbb{C}}K$ be the fundamental rank of G. Let $\chi(Y)$ be the Euler characteristic of Y. Then*

(a) $b_p^{(2)}(Y) \neq 0 \Leftrightarrow m = 0$ *and* $p = \frac{n}{2}$.
In particular, $b_{\frac{n}{2}}^{(2)}(Y) = (-1)^{\frac{n}{2}}\chi(Y) = \dfrac{\mathrm{vol}(Y)}{\mathrm{vol}(X^d)}\chi(X^d)$.

(b) $\alpha_p(Y) \neq \infty^+ \Leftrightarrow m > 0$ *and* $p \in [\frac{n-m}{2}, \frac{n+m}{2} - 1]$.
In this range $\alpha_p(Y) = m$.

(c) $\rho^{(2)}(Y) \neq 0 \Leftrightarrow m = 1$.

Note that $n - m$ is always even and positive. Part (a) of the theorem was known for a long time, at least since Borel's paper [1]. For special cases see also [4], [5]. For the convenience of the reader we include a proof here. In fact, we prove a stronger statement which should have been known to the experts although we were not able to find it in the literature:

Proposition 1.2 *The discrete spectrum of Δ_p on $L^2(X, \Lambda^p T^*X)$ is empty unless $m(X) = 0$ and $p = \frac{n}{2}$. In this case 0 is the only eigenvalue of Δ_p.*

The strategy of the proof of (b) can already be found in Lott's paper [14], Section VII. To be more precise, Equation (1) implies the slightly weaker result

$$\tilde{\alpha}_p(Y) = \begin{cases} \infty^+ & p \notin [\frac{n-m}{2}, \frac{n+m}{2}] \text{ or } m = 0 \\ m & p \in [\frac{n-m}{2}, \frac{n+m}{2}] \text{ and } m \neq 0 \end{cases} . \tag{2}$$

Lott proved the first line of (2) and that $\tilde{\alpha}_p(Y)$ is finite and independent of p for the remaining values of p. He indicated how one should be able to compute the precise value of $\tilde{\alpha}_p(Y)$. But he finished the computation in the real rank one case, only. In addition, already Borel [1] showed that the range of the differential d_p of the L^2-de-Rham-complex is not closed for $p \in [\frac{n-m}{2}, \frac{n+m}{2} - 1]$. Theorem 1.1 (b) can be considered as a quantitative refinement of this result. After the present paper was written I was informed by S. Mehdi that there is a joint paper of him with N. Lohoue [13] which has recently appeared in print and which contains a proof of (2). The interested reader will also find there more information concerning the material presented here in Sections 2 and 3. But he should be aware that in that paper the range of finiteness of $\tilde{\alpha}_p(Y)$ is constantly misprinted and that the proof of the second line of (2) as written down there is not quite complete (it is not mentioned that it is important to know that $p_\xi(0) > 0$, see Equation (14) below).

The main motivation to do the present work was to obtain part (c) of the theorem. In our locally homogeneous situation we have for any $x \in X$

$$\mathrm{Tr}_\Gamma e^{-t\Delta_p^*} = \mathrm{vol}(Y) \cdot \mathrm{tr}\, e^{-t\Delta_p^*}(x,x)$$

and thus

$$\rho^{(2)}(Y) = \mathrm{vol}(Y) \cdot T^{(2)}(X)$$

for a certain real number $T^{(2)}(X)$. Note that in contrast to $\rho^{(2)}(Y)$ the number $T^{(2)}(X)$ depends on the normalization of the invariant metric on X (of course only via the volume form). A well-known symmetry argument ([18], Proposition 2.1) yields that $T^{(2)}(X) = 0$ whenever $m \neq 1$. Lott [14] (see also [17]) showed that $T^{(2)}(H^n) \neq 0$ for $n = 3, 5, 7$, where H^n is the real hyperbolic space (his values for $T^{(2)}(H^n)$ for $n = 5, 7$ were not correct). This led to the conjecture that $T^{(2)}(H^n) \neq 0$ for all odd n which was open until the work of Hess-Schick [11] who found a trick in order to control the sign of $\log \det_\Gamma(\Delta_p^c)$ in terms of p and n. So they were able to show that there is a *positive* rational number q_n such that

$$T^{(2)}(H^n) = \left(-\frac{1}{\pi}\right)^{\frac{n-1}{2}} q_n \ . \tag{3}$$

Here the metric on H^n is normalized to have sectional curvature -1. Along the same lines Hess [10] obtained

$$T^{(2)}(SL(3,\mathbb{R})/SO(3)) \neq 0 \ . \tag{4}$$

Let us introduce

$$Q_n = \frac{2q_n}{(\frac{n-1}{2})!}$$

and rewrite (3) as

$$T^{(2)}(H^n) = (-1)^{\frac{n-1}{2}} \frac{\pi Q_n}{\mathrm{vol}(S^n)} \ . \tag{5}$$

Recall that S^n is the compact dual of H^n. The rational number Q_n which does not depend on the normalization of the metric has a nice interpretation in terms of Weyl's dimension polynomial for finite-dimensional representations, see Proposition 5.3. But its significance remains to be clarified further, and it

seems to be difficult to write down a practical formula valid for all odd n. One has $Q_3 = \frac{1}{3}$, $Q_5 = \frac{31}{45}$, $Q_7 = \frac{221}{210}$. For further information see [11].
We will reduce Theorem 1.1 (c) to (4) and the positivity of Q_n. Let X be an arbitrary symmetric space satisfying $m(X) = 1$. By the classification of simple Lie groups $X = X_1 \times X_0$, where $m(X_0) = 0$ and $X_1 = SL(3,\mathbb{R})/SO(3)$ or $X_1 = X_{p,q} := SO(p,q)^0/SO(p) \times SO(q)$ for p, q odd. (Here as throughout the paper a superscript 0 denotes the connected component of the identity.) Note that a corresponding decomposition of Y does not necessarily exist. We show

Proposition 1.3

(a) $T^{(2)}(X_{p,q}) = (-1)^{\frac{pq-1}{2}} \chi(X^d_{p-1,q-1}) \dfrac{\pi Q_{p+q-1}}{\mathrm{vol}(X^d_{p,q})}$.

(b) If $m(X) = 1$, *then* $T^{(2)}(X) = \dfrac{(-1)^{\frac{n_0}{2}} \chi(X^d_0)}{\mathrm{vol}(X^d_0)} T^{(2)}(X_1)$.
Here $n_0 = \dim X_0$.

Note that $X^d_{p-1,q-1} = SO(p+q-2)/SO(p-1) \times SO(q-1)$ and for $p, q > 1$

$$\chi(X^d_{p-1,q-1}) = 2 \binom{\frac{p+q-2}{2}}{\frac{p-1}{2}} .$$

In fact, it is a classical result that $\chi(X^d) > 0$ whenever $m(X) = 0$ (compare Theorem 1.1 (a)). It is equal to the quotient of the orders of certain Weyl groups (see Section 5). Now Theorem 1.1 (c) follows from Proposition 1.3, (4) and the positivity of Q_n.
We are also able to identify the missing constant in (4).

Proposition 1.4 *If* $X = SL(3,\mathbb{R})/SO(3)$, *then*

$$T^{(2)}(X) = \frac{2\pi}{3\mathrm{vol}(X^d)} .$$

If the invariant metric on X *is induced from twice the trace form of the standard representation of* $\mathfrak{sl}(3,\mathbb{R})$, *then* $\mathrm{vol}(X^d) = 4\pi^3$, *and we have*

$$T^{(2)}(X) = \frac{1}{6\pi^2} .$$

In particular, we see that $(-1)^{\frac{n-1}{2}} T^{(2)}(X)$ is positive for all X with $m(X) = 1$. Proposition 5.3 provides a uniform formula for the L^2-torsion of all these spaces.
Acknowledgements: I am grateful to Wolfgang Lück and Thomas Schick for inspiring discussions which have provided me with a sufficient amount of motivation and of knowledge on L^2-invariants in order to perform the computations which led to the results of the present paper. I am also indebted to Wolfgang Lück for giving me the opportunity to report on them at the Oberwolfach conference "L^2-methods and K-theory", September 1999. In addition, I benefited from discussions with J. Lott, P. Pansu, E. Hess and U. Bunke.

2 The Harish-Chandra Plancherel Theorem

We want to understand the action of the Laplacian and of the corresponding heat kernels on $L^2(X, \Lambda^p T^*X)$. Since the Laplacian coincides (up to the sign) with the action of the Casimir operator Ω of G (Kuga's Lemma [2], Thm. 2.5.) it is certainly enough to understand the "decomposition" of $L^2(X, \Lambda^p T^*X)$ into irreducible unitary representations of G. There is an isomorphism of homogeneous vector bundles $\Lambda^p T^*X \cong G \times_K \Lambda^p \mathfrak{p}^*$, and, hence, of G-representations

$$L^2(X, \Lambda^p T^*X) \cong [L^2(G) \otimes \Lambda^p \mathfrak{p}^*]^K \ .$$

Thus our task consists of two steps: First to understand $L^2(G)$ as a representation of $G \times G$ which is accomplished by the Harish-Chandra Plancherel Theorem recalled in the present section and, second, to understand spaces of the form $[V_\pi \otimes \Lambda^p \mathfrak{p}^*]^K$, where (π, V_π) is an irreducible unitary representation of G occurring in the Plancherel decomposition. For general G, the second step will resist a naive approach. However, if $\pi(\Omega) = 0$, then the space $[V_\pi \otimes \Lambda^p \mathfrak{p}^*]^K$ has cohomological meaning, and the theory of relative $(\mathfrak{g}, K)$-cohomology as recalled in the next section will provide a sufficient amount of information.

An irreducible unitary representation (π, V_π) of G is called a representation of the discrete series if there is a G-invariant embedding $V_\pi \hookrightarrow L^2(G)$. Let $\hat{G}_d$ denote the set of equivalence classes of discrete series representations of G. Then we have

Theorem 2.1 (Harish-Chandra [7]) *$\hat{G}_d$ is non-empty if and only if $m(X) = 0$.*

Note that $m(X) = 0$ means that G has a compact Cartan subgroup. The Plancherel Theorem provides a decomposition of $L^2(G)$ which is indexed by discrete series representations of certain subgroups $M \subset G$ which we are going to define now.

Let $\mathfrak{a}_0 \subset \mathfrak{p}$ be a maximal abelian subspace. It induces a root space decomposition

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\alpha \in \Delta(\mathfrak{g}, \mathfrak{a}_0)} \mathfrak{g}_\alpha \ .$$

Choose a decomposition $\Delta(\mathfrak{g}, \mathfrak{a}_0) = \Delta^+ \ \dot{\cup}\ -\Delta^+$ into positive and negative roots, and let $\Pi \subset \Delta^+$ be the subset of simple roots. For any subset $F \subset \Pi$ we define

$$\mathfrak{a}_F = \{H \in \mathfrak{a}_0 \mid \alpha(H) = 0 \text{ for all } \alpha \in \Pi\} \ , \qquad A_F = \exp(\mathfrak{a}_F) \ ,$$
$$\mathfrak{n}_F = \bigoplus_{\{\alpha \in \Delta^+ \mid \alpha_{|\mathfrak{a}_F} \neq 0\}} \mathfrak{g}_\alpha \ , \qquad N_F = \exp(\mathfrak{n}_F) \ .$$

Furthermore, there is a unique (possibly disconnected) subgroup $M_F \subset G$ with Lie algebra $\mathfrak{m}_F$ such that $M_F A_F$ is the centralizer of $\mathfrak{a}_F$ in G and $\mathfrak{m}_F$

is orthogonal to $\mathfrak{a}_F$ with respect to any invariant bilinear form on $\mathfrak{g}$. M_F is a reductive subgroup with compact center. The corresponding parabolic subgroup $P_F := M_F A_F N_F$ is called a standard parabolic. P_F is called cuspidal if M_F has a compact Cartan subgroup. If for two subsets $F, I \subset \Pi$ the spaces $\mathfrak{a}_F$ and $\mathfrak{a}_I$ are conjugated by an element of K (thus by an element of the Weyl group $W(\mathfrak{g}, \mathfrak{a}_0)$) we call P_F and P_I associate. (In many cases this already implies $F = I$.) The assignment $P_F \mapsto A_F T$, where T is a compact Cartan subgroup of M_F, gives a one-to-one correspondence between association classes of cuspidal parabolic subgroups and conjugacy classes of Cartan subgroups of G.

For two subsets $F \subset I \subset \Pi$ we have $P_F \subset P_I$, $M_F \subset M_I$, $A_F \supset A_I$, $N_F \supset N_I$. If F is understood we will often suppress the subscript F.

For illustration let us consider the two extreme cases. The minimal parabolic arises for $F = \emptyset$. Since $M_\emptyset$ is compact $P_\emptyset$ is always cuspidal. For $F = \Pi$ we have $P = M = G$, and G is cuspidal iff $m(X) = 0$. For any cuspidal parabolic subgroup we have $\dim A \geq m(X)$, and there is exactly one association class of cuspidal parabolic subgroups, called fundamental, with $\dim A = m(X)$.

Theorem 2.1 also holds in the context of such reductive groups like M. Thus a parabolic $P = MAN$ is cuspidal iff M has a non-empty discrete series $\hat{M}_d$. Let $\mathfrak{a}^*_{\mathbb{C}}$ be the complexified dual of the Lie algebra $\mathfrak{a}$ of A. For a discrete series representation (ξ, W_ξ) of M and $\nu \in \mathfrak{a}^*_{\mathbb{C}}$ we form the induced representation $(\pi_{\xi,\nu}, H^{\xi,\nu})$ by

$$\begin{aligned} H^{\xi,\nu} &= \left\{ f : G \to W_\xi \mid \begin{array}{l} f(gman) = a^{-(\nu+\rho_{\mathfrak{a}})}\xi(m)^{-1} f(g) \text{ for all} \\ g \in G, man \in MAN,\ f_{|K} \in L^2(K, W_\xi) \end{array} \right\}, \\ & \quad (\pi_{\xi,\nu}(g)f)(x) = f(g^{-1}x) \ . \end{aligned}$$

Here $\rho_{\mathfrak{a}} = \dfrac{1}{2} \sum_{\alpha \in \Delta^+} \alpha_{|\mathfrak{a}}$. If $\nu \in i\mathfrak{a}^*$, then $\pi_{\xi,\nu}$ is unitary. An invariant bilinear form on $\mathfrak{g}$ induces corresponding forms on $\mathfrak{m}$ and $\mathfrak{a}^*_{\mathbb{C}}$ and determines Casimir operators Ω and Ω_M of G and M, respectively. Then we have

$$\pi_{\xi,\nu}(\Omega) = \langle \nu, \nu \rangle - \langle \rho_{\mathfrak{a}}, \rho_{\mathfrak{a}} \rangle + \xi(\Omega_M) \ . \tag{6}$$

Note that $\xi(\Omega_M)$ is a non-negative real scalar.

Let $\mathcal{C}(G) \subset L^2(G)$ be the Harish-Chandra Schwartz space (for a definition see e.g. [19], 7.1.2). It is stable under the left and right regular representations l and r of G. Let $\mathcal{C}(G)_{K\times K} = \{f \in \mathcal{C}(G) \mid \dim \operatorname{span}\{l_{k_1} r_{k_2} f \mid k_1, k_2 \in K\} < \infty\}$ be the subspace of Schwartz functions which are K-finite from the left and the right. Note that $\mathcal{C}(G)_{K\times K}$ is dense in $L^2(G)$. Suppose that $\nu \in i\mathfrak{a}^*$. Then for $f \in \mathcal{C}(G)$

$$\pi_{\xi,\nu}(f) := \int_G f(g) \pi_{\xi,\nu}(g)\, dg$$

is a well-defined trace class operator on $H^{\xi,\nu}$ which has finite rank if $f \in \mathcal{C}(G)_{K\times K}$. Note that the map $f \mapsto \pi_{\xi,\nu}(f)$ intertwines the G-actions in the following way: $\pi_{\xi,\nu}(l_x r_y f) = \pi_{\xi,\nu}(x) \pi_{\xi,\nu}(f) \pi_{\xi,\nu}(y^{-1})$ for $x, y \in G$.

The Harish-Chandra Plancherel Theorem can now be formulated as follows:

Theorem 2.2 (Harish-Chandra [9]) *For each cuspidal parabolic subgroup as constructed above and any discrete series representation ξ of the corresponding group M there exists an explicitly computable analytic function $p_\xi : i\mathfrak{a}^* \to [0,\infty)$ of polynomial growth (the Plancherel density) such that for any $f \in \mathcal{C}(G)_{K\times K}$ and $g \in G$*

$$f(g) = \sum_P \sum_{\xi\in\hat{M}_d} \int_{\mathfrak{a}^*} \mathrm{Tr}(\pi_{\xi,i\nu}(f)\pi_{\xi,i\nu}(g^{-1}))\, p_\xi(i\nu)d\nu \ .$$

Here the first sum runs over a set of representatives $P = P_F$ of association classes of cuspidal parabolic subgroups of G.

For more details on the Plancherel Theorem and the structure theory behind it the interested reader may consult the textbooks [12], [19], [20].

Note that the Plancherel measures $p_\xi(i\nu)d\nu$ depend on the normalization of the Haar measure dg. In the remainder of the paper we use the following one. Let dx be the Riemannian volume form of $X = G/K$ and dk be the Haar measure of K with total mass one. Then $\int_G f(g)\, dg = \int_X \tilde{f}(x)\, dx$, where $\tilde{f}(gK) = \int_K f(gk)\, dk$. We normalize the invariant bilinear form on $\mathfrak{g}$ such that its restriction to $\mathfrak{p} \cong T_{eK}X$ coincides with the Riemannian metric of X. Let $d\nu$ be the Lebesgue measure corresponding to the induced form on $\mathfrak{a}^*$. By these choices p_ξ is uniquely determined.

We are now able to give a kind of spectral expansion of $\mathrm{Tr}_\Gamma e^{-t\Delta_p^*}$.

Corollary 2.3

$$\begin{aligned} \mathrm{Tr}_\Gamma e^{-t\Delta_p} &= \mathrm{vol}(Y) \sum_P \sum_{\xi\in\hat{M}_d} \int_{\mathfrak{a}^*} e^{-t(\|\nu\|^2+\|\rho_\mathfrak{a}\|^2-\xi(\Omega_M))} \\ &\qquad \dim[H^{\xi,i\nu}\otimes\Lambda^p\mathfrak{p}^*]^K p_\xi(i\nu)d\nu \qquad (7) \\ &= \mathrm{vol}(Y) \sum_P \sum_{\xi\in\hat{M}_d} \int_{\mathfrak{a}^*} e^{-t(\|\nu\|^2+\|\rho_\mathfrak{a}\|^2-\xi(\Omega_M))} \\ &\qquad \dim[W_\xi\otimes\Lambda^p\mathfrak{p}^*]^{K_M} p_\xi(i\nu)d\nu \ . \qquad (8) \end{aligned}$$

Here $K_M := K\cap M$ denotes the maximal compact subgroup of M. There are only finitely many pairs (P,ξ) with $[H^{\xi,i\nu}\otimes\Lambda^p\mathfrak{p}^*]^K \cong [W_\xi\otimes\Lambda^p\mathfrak{p}^*]^{K_M} \neq \{0\}$.

Proof. We define $k_t \in [\mathcal{C}(G)\otimes\mathrm{End}(\Lambda^p\mathfrak{p}^*)]^{K\times K}$ by $k_t(g) := e^{-t\Delta_p}(eK, gK)\circ g$. For $f \in L^2(X,\Lambda^pT^*X) \cong [L^2(G)\otimes\Lambda^p\mathfrak{p}^*]^K$ we have $e^{-t\Delta_p}f(g_0) = \int_G k_t(g)f(g_0g)\, dg$. In addition, $\mathrm{tr}\, e^{-t\Delta_p}(x,x) = \mathrm{tr}\, k_t(e)$ for any $x \in X$. We consider $\pi_{\xi,i\nu}(k_t)$ as an operator acting on $H^{\xi,i\nu}\otimes\Lambda^p\mathfrak{p}^*$. Using Kuga's

Lemma, Equation (6), and the $K \times K$-invariance of k_t one derives that $\pi_{\xi,i\nu}(k_t) = e^{-t(\|v\|^2+\|\rho_{\mathfrak{a}}\|^2-\xi(\Omega_M))}P$, where P is the orthogonal projection onto the subspace of K-invariants in $H^{\xi,i\nu} \otimes \Lambda^p\mathfrak{p}^*$. The Plancherel formula now yields

$$\operatorname{tr} k_t(e) = \sum_P \sum_{\xi \in \hat{M}_d} \int_{\mathfrak{a}^*} e^{-t(\|v\|^2+\|\rho_{\mathfrak{a}}\|^2-\xi(\Omega_M))} \dim[H^{\xi,i\nu} \otimes \Lambda^p\mathfrak{p}^*]^K \, p_\xi(i\nu) d\nu \ .$$

This proves (7). Since $H^{\xi,i\nu} \cong L^2(K \times_{K_M} W_\xi)$ as a representation of K Equation (8) follows by Frobenius reciprocity. The last assertion is a consequence of the Blattner formula (see e.g. [19], 6.5.4) for the K_M-types of discrete series representations of M. $\square$

3 $(\mathfrak{g}, K)$-cohomology

If (π, V_π) is a representation of G on a complete locally convex Hausdorff topological vector space, then we can form its subspace $V_{\pi,K}$ consisting of all K-finite smooth vectors of V_π. $V_{\pi,K}$ becomes a simultaneous module under $\mathfrak{g}$ and K, where both actions satisfy the obvious compatibility conditions. Such a module is called a $(\mathfrak{g}, K)$-module (see [2], 0.2).

We are interested in the functor of $(\mathfrak{g}, K)$-cohomology $V \mapsto H^*(\mathfrak{g}, K, V)$ which goes from the category of $(\mathfrak{g}, K)$-modules to the category of vector spaces. It is the right derived functor of the left exact functor taking $(\mathfrak{g}, K)$-invariants. $H^*(\mathfrak{g}, K, V)$ can be computed using the standard relative Lie algebra cohomology complex $([V \otimes \Lambda^*\mathfrak{p}^*]^K, d)$, where

$$\begin{aligned} d\omega(X_0, \dots, X_p) &= \sum_{i=0}^p (-1)^i \pi(X_i)\omega(X_0, \dots, \hat{X}_i, \dots, X_p) \ , \\ &\qquad \omega \in [V \otimes \Lambda^p\mathfrak{p}^*]^K, \ X_i \in \mathfrak{p} \ . \end{aligned}$$

Note that for $V = C^\infty(G)_K$ this complex is isomorphic to the de Rham complex of the symmetric space X.

Let $\mathcal{Z}(\mathfrak{g})$ be the center of the universal enveloping algebra of $\mathfrak{g}$. If (π, V) is an irreducible $(\mathfrak{g}, K)$-module, then any $z \in \mathcal{Z}(\mathfrak{g})$ acts by a scalar $\chi_\pi(z)$ on V. The homomorphism $\chi_\pi : \mathcal{Z}(\mathfrak{g}) \to \mathbb{C}$ is called the infinitesimal character of V. The following basic result can be considered as an algebraic version of Hodge theory.

Proposition 3.1 ([2], II.3.1. and I.5.3.) *Let (π, V_π) be an irreducible unitary representation of G, and let (τ, F) be an irreducible finite-dimensional representation of G. Then*

$$H^p(\mathfrak{g}, K, V_{\pi,K} \otimes F) = \begin{cases} [V_\pi \otimes F \otimes \Lambda^p\mathfrak{p}^*]^K & \pi(\Omega) = \tau(\Omega) \\ \{0\} & \pi(\Omega) \neq \tau(\Omega) \end{cases} \ .$$

If $H^p(\mathfrak{g}, K, V_{\pi,K} \otimes F) \neq \{0\}$, then $\chi_\pi = \chi_{\tilde{\tau}}$, where $\tilde{\tau}$ is the dual representation of τ.

The cohomology groups $H^p(\mathfrak{g}, K, V_{\pi,K} \otimes F)$ for the representations $\pi = \pi_{\xi,i\nu}$ occurring in the Plancherel Theorem have been computed in [2]. We shall need the following information.

Proposition 3.2 ([2],II.5.3., III.5.1., and III.3.3.)

(a) *Let (τ, F) be an irreducible finite-dimensional representation of G and $\pi \in \hat{G}_d$ with $\chi_\pi = \chi_{\tilde{\tau}}$. Then*

$$\dim H^p(\mathfrak{g}, K, V_{\pi,K} \otimes F) = \begin{cases} 1 & p = \frac{n}{2} \\ 0 & \textit{otherwise} \end{cases} .$$

(b) *Let $(\pi_{\xi,\nu}, H^{\xi,\nu})$ be a representation occurring in the Plancherel Theorem. Then*

$$H^*(\mathfrak{g}, K, H_K^{\xi,i\nu}) = \{0\}$$

unless P is fundamental, $\nu = 0$ and ξ belongs to a certain non-empty finite subset $\Xi \subset \hat{M}_d$. If P is fundamental and $\xi \in \Xi$, then

$$\dim H^p(\mathfrak{g}, K, H_K^{\xi,0}) = \begin{cases} \binom{m}{p-\frac{n-m}{2}} & p \in [\frac{n-m}{2}, \frac{n+m}{2}] \\ 0 & \textit{otherwise} \end{cases} .$$

Choose a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ and a system of positive roots. Via the Harish-Chandra isomorphism any infinitesimal character $\chi_\pi : \mathcal{Z}(\mathfrak{g}) \to \mathbb{C}$ is given by an element $\Lambda_\pi \in \mathfrak{h}^*_{\mathbb{C}}$, which is uniquely determined up to the action of the Weyl group $W(\mathfrak{g}, \mathfrak{h})$ ([19], 3.2.4.). The set of infinitesimal characters of discrete series representations coincides (in case $m = 0$) with the set of infinitesimal characters of finite-dimensional representations, which are of the form $\mu_\tau + \rho_{\mathfrak{g}}$, μ_τ and $\rho_{\mathfrak{g}}$ being the highest weight of τ and the half-sum of positive roots, respectively. In the following we will represent infinitesimal characters of discrete series representations by elements of this form. We introduce a partial order on $\mathfrak{h}^*_{\mathbb{C}}$ by saying that $\mu > \nu$, if $\mu - \nu$ is a sum of (not necessarily distinct) positive roots. Then a careful examination of the proof of Proposition II.5.3. in [2], which only rests on some basic knowledge of the possible K-types occurring in discrete series representations, shows that slightly more than Proposition 3.2 (a) is true.

Proposition 3.3 *Let (τ, F) be an irreducible finite-dimensional representation of G and $\pi \in \hat{G}_d$ with $\Lambda_\pi \not< \Lambda_{\tilde{\tau}}$. Then*

$$\dim[V_{\pi,K} \otimes F \otimes \Lambda^p \mathfrak{p}^*]^K = \begin{cases} 1 & p = \frac{n}{2}, \chi_\pi = \chi_{\tilde{\tau}} \\ 0 & \textit{otherwise} \end{cases} .$$

4 L^2-Betti numbers and Novikov-Shubin invariants

In this section we shall prove parts (a) and (b) of Theorem 1.1 as well as Proposition 1.2.
Let $L^2(X, \Lambda^p T^*X)_d$ be the discrete subspace of $L^2(X, \Lambda^p T^*X)$, i.e., the direct sum of the L^2-eigenspaces of the Laplacian. The Plancherel Theorem in particular says that as a representation of G

$$L^2(X, \Lambda^p T^*X)_d \cong \bigoplus_{\pi \in \hat{G}_d} V_{\tilde{\pi}} \otimes [V_\pi \otimes \Lambda^p \mathfrak{p}^*]^K .$$

Let $(\tau_0, \mathbb{C})$ be the trivial representation of G. Let $\pi \in \hat{G}_d$. Then $\Lambda_\pi \not< \Lambda_{\tau_0} = \rho_{\mathfrak{g}}$. Proposition 3.3 now yields

$$\dim[V_\pi \otimes \Lambda^p \mathfrak{p}^*]^K = \begin{cases} 1 & p = \frac{n}{2}, \chi_\pi = \chi_{\tau_0} \\ 0 & \text{otherwise} \end{cases} . \tag{9}$$

Since in case $m = 0$ discrete series representations with infinitesimal character χ_{τ_0} always exist (see e.g. [12], Thm. 9.20 or [19], Thm. 6.8.2) and $\tau_0(\Omega) = 0$ this implies Proposition 1.2. In particular, $b_p^{(2)}(Y)$ is non-zero exactly when $m = 0$ and $p = \frac{n}{2}$. Using that the L^2-Euler characteristic coincides with the usual Euler characteristic and applying Hirzebruch proportionality we obtain

$$b_{\frac{n}{2}}^{(2)}(Y) = (-1)^{\frac{n}{2}} \chi(Y) = \frac{\mathrm{vol}(Y)}{\mathrm{vol}(X^d)} \chi(X^d) . \tag{10}$$

We will give an alternative, purely analytic proof of that formula in the sequel of Corollary 5.2. This finishes the proof of part (a) of Theorem 1.1.
We now turn to part (b). In order to compute $\alpha_p(Y)$ we need an expression for $\mathrm{Tr}_\Gamma e^{-t\Delta_p^c}$.

Proposition 4.1 *For any triple (P, ξ, ν) appearing in (7) let $B^p(\xi, \nu) = d([H^{\xi, i\nu} \otimes \Lambda^{p-1}\mathfrak{p}^*]^K)$ be the space of coboundaries in the relative Lie algebra cohomology complex and $b^p(\xi, \nu)$ be its dimension. Then*

$$\mathrm{Tr}_\Gamma e^{-t\Delta_p^c} = \mathrm{vol}(Y) \sum_{P \neq G} \sum_{\xi \in \hat{M}_d} \int_{\mathfrak{a}^*} e^{-t(\|\nu\|^2 + \|\rho_{\mathfrak{a}}\|^2 - \xi(\Omega_M))} b^{p+1}(\xi, \nu)\, p_\xi(i\nu) d\nu . \tag{11}$$

Here the first sum runs over a set of representatives $P = P_F$ of association classes of proper cuspidal parabolic subgroups of G. If $m > 0$, P is fundamental, $\xi \in \Xi$ (see Proposition 3.2 (b)), and $\nu \neq 0$, then

$$b^{p+1}(\xi, \nu) = \begin{cases} \binom{m-1}{p - \frac{n-m}{2}} & p \in [\frac{n-m}{2}, \frac{n+m}{2} - 1] \\ 0 & \textit{otherwise} \end{cases} . \tag{12}$$

Proof. We proceed exactly as in the proof of Corollary 2.3. The kernel $e^{-t\Delta_p^c}$ determines a function $k_t^c \in [\mathcal{C}(G) \otimes \mathrm{End}(\Lambda^p \mathfrak{p}^*)]^{K\times K}$. Then one computes that $\pi_{\xi,i\nu}(k_t^c) = e^{-t(\|\nu\|^2+\|\rho_\mathfrak{a}\|^2-\xi(\Omega_M))}P^c$, where P^c is the projection to the orthogonal complement in $[H^{\xi,i\nu} \otimes \Lambda^p\mathfrak{p}^*]^K$ of the space of cocycles in the relative Lie algebra cohomology complex. The dimension of that complement is equal to $b^{p+1}(\xi,\nu)$. If $m = 0$, $P = G$, and $\xi \in \hat{G}_d$, then $b^{p+1}(\xi) = 0$ for all p by (9). This proves (11).

Let now $m > 0$, P be fundamental, $\xi \in \Xi$, and $\nu \in \mathfrak{a}^*$. In this case $H^p(\mathfrak{g}, K, H_K^{\xi,0}) \neq \{0\}$ for some p. Hence Proposition 3.1 implies that $\pi_{\xi,0}(\Omega) = 0$ and

$$h^p(\xi) := \dim[H^{\xi,i\nu} \otimes \Lambda^p\mathfrak{p}^*]^K = \dim[H^{\xi,0} \otimes \Lambda^p\mathfrak{p}^*]^K = \dim H^p(\mathfrak{g}, K, H_K^{\xi,0}) \ .$$

By Proposition 3.2 (b) we have $h^p(\xi) = \binom{m}{p-\frac{n-m}{2}}$. On the other hand, $\dim H^p(\mathfrak{g}, K, H_K^{\xi,i\nu}) = 0$ for $\nu \neq 0$ implies that $h^p(\xi) = b^p(\xi,\nu) + b^{p+1}(\xi,\nu)$. (12) now follows inductively. □

If $H^p(\mathfrak{g}, K, H_K^{\xi,i\nu}) = \{0\}$ for all $\nu \in \mathfrak{a}^*$, then by Proposition 3.1 $\dim[H^{\xi,i\nu} \otimes \Lambda^p\mathfrak{p}^*]^K = 0$ (which is independent of ν) or $\inf\limits_{\nu\in\mathfrak{a}^*}(\|\nu\|^2 + \|\rho_\mathfrak{a}\|^2 - \xi(\Omega_M)) > 0$. Now Proposition 4.1 implies that the spectrum of Δ_p^c has a gap around zero, which means $\alpha_p(Y) = \infty^+$, unless $m > 0$ and $p \in [\frac{n-m}{2}, \frac{n+m}{2} - 1]$. In the latter case we obtain for some $c > 0$

$$\mathrm{Tr}_\Gamma e^{-t\Delta_p^c} = \mathrm{vol}(Y) \binom{m-1}{p-\frac{n-m}{2}} \sum_{\xi\in\Xi} \int_{\mathfrak{a}^*} e^{-t\|\nu\|^2}\, p_\xi(i\nu) d\nu + O(e^{-ct})$$

as $t \to \infty$, where $\mathfrak{a}$ corresponds to a fundamental parabolic subgroup and thus has dimension m.

Let $P = MAN$ be fundamental and $\xi \in \hat{M}_d$. Then $\dim \mathfrak{n} =: 2u$ is even. Choose a compact Cartan subgroup $T \subset M$ with Lie algebra $\mathfrak{t}$ and a system $\Delta^+(\mathfrak{m},\mathfrak{t})$ of positive roots. Considering the pair $(\mathfrak{m},\mathfrak{t})$ instead of $(\mathfrak{g},\mathfrak{h})$ we can define $\Lambda_\xi \in \mathfrak{t}_\mathbb{C}^*$ in the same way as at the end of Section 3. Then $\mathfrak{h} := \mathfrak{t} \oplus \mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$. Let Φ^+ be a system of positive roots for $(\mathfrak{g},\mathfrak{h})$ containing $\Delta^+(\mathfrak{m},\mathfrak{t})$. Then there exists a positive constant c_X depending only on the normalization of the volume form dx such that

$$p_\xi(\nu) = c_X(-1)^u \prod_{\alpha\in\Phi^+} \frac{\langle\alpha, \Lambda_\xi + \nu\rangle}{\langle\alpha, \rho_\mathfrak{g}\rangle} \tag{13}$$

(see [9], Thm. 24.1, [20], Thm. 13.5.1 or [12], Thm. 13.11). In particular, p_ξ is an even polynomial of degree $\dim \mathfrak{n}$. The factor $(-1)^u$ makes it nonnegative on $i\mathfrak{a}^*$.

The element $\Lambda_\xi + \nu$ gives the infinitesimal character of $\pi_{\xi,\nu}$: $\Lambda_{\pi_{\xi,\nu}} = \Lambda_\xi + \nu$ ([12], Prop. 8.22). Let now $\xi \in \Xi$. Then Propositions 3.2 (b) and 3.1 imply that $\pi_{\xi,0}$ has the same infinitesimal character as the trivial representation. It follows that Λ_ξ is conjugated in $\mathfrak{h}_{\mathbb{C}}^*$ by an element of the Weyl group $W(\mathfrak{g},\mathfrak{h})$ to $\rho_{\mathfrak{g}}$. By (13) we obtain $p_\xi(0) = \pm c_X \neq 0$. On the other hand $p_\xi(0) \geq 0$, hence

$$p_\xi(0) > 0 \ . \tag{14}$$

We decompose $p_\xi(\nu) = \sum_{k=0}^{u} p_{\xi,2k}(\nu)$ into homogeneous polynomials. Set $q_{\xi,k} = \int_{\|\nu\|=1} p_{\xi,2k}(\nu)\, d\nu$. Then $q_{\xi,0} > 0$ and

$$\begin{aligned} \int_{\mathfrak{a}^*} e^{-t\|\nu\|^2} p_\xi(i\nu) d\nu &= \sum q_{\xi,k} \int_0^\infty e^{-tr^2} r^{m-1+2k} dr \\ &= \sum t^{-(\frac{m}{2}+k)} q_{\xi,k} \int_0^\infty e^{-y^2} y^{m-1+2k} dy \ . \end{aligned}$$

Thus for $p \in [\frac{n-m}{2}, \frac{n+m}{2} - 1]$ the leading term of $\mathrm{Tr}_\Gamma e^{-t\Delta_p^c}$ as $t \to \infty$ is a non-zero multiple of $t^{-\frac{m}{2}}$. This completes the proof of Theorem 1.1 (b).

5 L^2-torsion

For even dimensional manifolds the L^2-torsion vanishes. Thus we may assume that m is odd, in particular $m \geq 1$. Then $\Delta_p' = \Delta_p$. We first want to compute

$$k_X(t) := \frac{1}{2\mathrm{vol}(Y)} \sum_{p=0}^{n} (-1)^p p \, \mathrm{Tr}_\Gamma e^{-t\Delta_p} \ .$$

Then

$$T^{(2)}(X) = \frac{d}{ds}_{|s=0} \left(\frac{1}{\Gamma(s)} \int_0^\varepsilon k_X(t) t^{s-1} dt \right) + \int_\varepsilon^\infty k_X(t) t^{-1} dt \ .$$

Let $P = MAN$ be a parabolic subgroup appearing in (8). Set $K_M = K \cap M$ and $\mathfrak{p}_{\mathfrak{m}} = \mathfrak{p} \cap \mathfrak{m}$. Then an elementary calculation in the representation ring $R(K_M)$ of K_M yields

$$\sum_{p=0}^{n} (-1)^p p \, \Lambda^p \mathfrak{p}^* = 0 \ , \qquad \text{if } \dim \mathfrak{a} \geq 2 \ ,$$

while for $\dim \mathfrak{a} = 1$ we have

$$\begin{aligned} \sum_{p=0}^{n} (-1)^p p \, \Lambda^p \mathfrak{p}^* &= \sum_{p=0}^{n-1} (-1)^{p+1} \Lambda^p (\mathfrak{p}_{\mathfrak{m}}^* \oplus \mathfrak{n}^*) \\ &= \sum_{l=0}^{\dim \mathfrak{n}} (-1)^{l+1} (\Lambda^{ev} \mathfrak{p}_{\mathfrak{m}}^* - \Lambda^{odd} \mathfrak{p}_{\mathfrak{m}}^*) \otimes \Lambda^l \mathfrak{n} \end{aligned}$$

(see [18], Prop. 2.1 and Lemma 2.3). It follows from (8) that $k_X(t) \equiv 0$ for $m > 1$, hence $\rho^{(2)}(Y) = 0$.
From now on let $m = 1$. Let $P = MAN$ be a fundamental parabolic subgroup of G. Then (8) gives

$$k_X(t) = \frac{1}{2}\sum_{l=0}^{\dim \mathfrak{n}}(-1)^{l+1}\sum_{\xi\in\hat{M}_d}\dim\,[W_\xi\otimes(\Lambda^{ev}\mathfrak{p}_{\mathfrak{m}}^*-\Lambda^{odd}\mathfrak{p}_{\mathfrak{m}}^*)\otimes\Lambda^l\mathfrak{n}^*]^{K_M}$$
$$\int_{\mathfrak{a}^*} e^{-t(\|\nu\|^2+\|\rho_{\mathfrak{a}}\|^2-\xi(\Omega_M))}\,p_\xi(i\nu)d\nu\ . \qquad (15)$$

Now $X = X_1 \times X_0$, $X_1 = G_1/K_1$, $X_0 = G_0/K_0$, $m(X_0) = 0$ as explained in the introduction. Although (15) can be evaluated directly for general X with $m(X) = 1$ we prefer to reduce the computation to the irreducible case $X = X_1$. In order to compute $T^{(2)}(X)$ it is sufficient to compare $\rho^{(2)}(Y)$ with $\mathrm{vol}(Y)$ for one particular $Y = \Gamma\backslash X$. If we choose Γ of the form $\Gamma_1\times\Gamma_0$, where $\Gamma_0\subset G_0$ and $\Gamma_1\subset G_1$, then

$$\rho^{(2)}(Y) = \chi(Y_0)\rho^{(2)}(Y_1)\ ,$$

where $Y_1 = \Gamma_1\backslash X_1$, $Y_0 = \Gamma_0\backslash X_0$. Applying Hirzebruch proportionality we obtain the assertion of Proposition 1.3 (b)

$$\begin{aligned} T^{(2)}(X) &= \frac{\rho^{(2)}(Y)}{\mathrm{vol}(Y_0)\mathrm{vol}(Y_1)} = \frac{\chi(Y_0)}{\mathrm{vol}(Y_0)}T^{(2)}(X_1)\\ &= \frac{(-1)^{\frac{n_0}{2}}\chi(X_0^d)}{\mathrm{vol}(X_0^d)}T^{(2)}(X_1)\ . \end{aligned} \qquad (16)$$

It remains to deal with the case $X = X_1$. We can assume that $G = SO(p,q)^0$, $p \le q$ odd, or $G = SL(3,\mathbb{R})$. Then $M \cong SO(p-1,q-1)$, $\mathfrak{n}\cong\mathbb{R}^{p+q-2}$ or $M \cong {}^0GL(2,\mathbb{R}) := \{A\in GL(2,\mathbb{R})\mid |\det A| = 1\}$, $\mathfrak{n}\cong\mathbb{R}^2$, respectively, and M acts on $\mathfrak{n}$ via the standard representation. Note that M is not connected unless $G = SO(1,q)^0$. The M^0-representations $\Lambda^l\mathfrak{n}^*\otimes\mathbb{C}$ are irreducible unless $G = SO(p,q)$ and $l = u = \frac{1}{2}\dim\mathfrak{n}$. In the latter case $\Lambda^u\mathfrak{n}^*\otimes\mathbb{C}$ decomposes into two irreducible components $\Lambda^+\mathfrak{n}\oplus\Lambda^-\mathfrak{n}$. Since compact Cartan subgroups of M are connected the discrete series representations of M are induced from discrete series representations of M^0: $W_\xi = \mathrm{Ind}_{M^0}^M(W_{\xi^0})$, $\xi_0\in(\hat{M}^0)_d$ (see [19], 6.9 and 8.7.1). As representations of K_M we have $W_\xi\cong\mathrm{Ind}_{K_M^0}^{K_M}(W_{\xi^0})$. By Frobenius reciprocity we obtain

$$\begin{aligned}&\dim[W_\xi\otimes(\Lambda^{ev}\mathfrak{p}_{\mathfrak{m}}^*-\Lambda^{odd}\mathfrak{p}_{\mathfrak{m}}^*)\otimes\Lambda^l\mathfrak{n}^*]^{K_M}\\ &\quad = \dim[W_{\xi_0}\otimes(\Lambda^{ev}\mathfrak{p}_{\mathfrak{m}}^*-\Lambda^{odd}\mathfrak{p}_{\mathfrak{m}}^*)\otimes\Lambda^l\mathfrak{n}^*]^{K_M^0}\ .\end{aligned}$$

Note that the infinitesimal characters χ_ξ and χ_{ξ_0} coincide. By $\chi(\mathfrak{m},K_M^0,.)$ we denote the Euler characteristic of relative Lie algebra cohomology. Set

$v = \frac{1}{2}\dim \mathfrak{p}_{\mathfrak{m}}$. Applying Propositions 3.1 and 3.2 (a) to M^0 instead of G we obtain

$$\begin{aligned}
&\dim[W_{\xi_0} \otimes (\Lambda^{ev}\mathfrak{p}_{\mathfrak{m}}^* - \Lambda^{odd}\mathfrak{p}_{\mathfrak{m}}^*) \otimes \Lambda^*\mathfrak{n}^*]^{K_M^0} \\
&= \chi(\mathfrak{m}, K_M^0, W_{\xi_0, K_M^0} \otimes \Lambda^*\mathfrak{n}^*) \\
&= \begin{cases} (-1)^v & \chi_\xi = \chi_{\Lambda^*\mathfrak{n}} \\ 0 & \text{otherwise} \end{cases} .
\end{aligned}$$

Here $\Lambda^*\mathfrak{n}^*$, $* = l, +, -$, denotes an irreducible component of $\Lambda^l\mathfrak{n}^* \otimes \mathbb{C}$.
In all cases under consideration the set $\{\alpha_{|\mathfrak{a}} \mid \alpha \in \Delta^+, \alpha_{|\mathfrak{a}} \neq 0\}$ consists of a single element $\alpha_0 \in \mathfrak{a}^*$. It follows that $\rho_{\mathfrak{a}} = u\alpha_0$. Moreover, Ω_M acts on $\Lambda^l\mathfrak{n}$ as $l(2u-l)\|\alpha_0\|^2\mathrm{id}$ (compare [18], Lemma 2.5).
In order to evaluate (15) further we have to determine the constant c_X in formula (13). This can be done in complete generality. So for a moment we drop the assumptions $m = 1$, $X = X_1$.

Lemma 5.1 *Let $P = MAN \subset G$ be fundamental. We retain the notation introduced before (13). Set $W_A = \{k \in K | \mathrm{Ad}(k)\mathfrak{a} \subset \mathfrak{a}\}/K_M$, $S_A^d = \exp(i\mathfrak{a})K \subset X^d$, and let $\Phi_{\mathfrak{k}}^+$ be a positive root system for $(\mathfrak{k}, \mathfrak{t})$ with corresponding half sum $\rho_{\mathfrak{k}}$. Then*

$$c_X = \frac{1}{|W_A|(2\pi)^{\frac{n+m}{2}}} \frac{\prod_{\alpha\in\Phi^+}\langle\alpha, \rho_{\mathfrak{g}}\rangle}{\prod_{\alpha\in\Phi_{\mathfrak{k}}^+}\langle\alpha, \rho_{\mathfrak{k}}\rangle} \tag{17}$$

$$= \frac{1}{|W_A|} \frac{\mathrm{vol}(S_A^d)}{(2\pi)^m} \frac{1}{\mathrm{vol}(X^d)} . \tag{18}$$

Proof. Formula (17) is a combination of [8], Thm. 37.1, with [9], Cor. 23.1, Thm. 24.1 and Thm. 27.3. In order to apply these results correctly one has to take into account that Harish-Chandra's and our normalizations of the measures dg and $d\nu$ all of them starting from a fixed invariant bilinear form on $\mathfrak{g}$ differ by the factors $2^{\frac{n-\dim \mathfrak{a}_0}{2}}$ ([8], Section 7 and Lemma 37.2) and $(2\pi)^m$, respectively. On the other hand we have

$$\prod_{\alpha\in\Phi_{\mathfrak{k}}^+}\langle\alpha, \rho_{\mathfrak{k}}\rangle = (2\pi)^{\frac{\dim K/T}{2}} \frac{\mathrm{vol}(T)}{\mathrm{vol}(K)} \tag{19}$$

(see e.g. [8], Lemma 37.4). Here the volumes are the Riemannian ones corresponding to the invariant bilinear form $\langle ., . \rangle$. Formula (19) holds for any pair (K, T) of a connected compact Lie group and a maximal torus. Applying it also to the pair (G^d, H^d), where H^d is the maximal torus of G^d with Lie algebra $\mathfrak{t} \oplus i\mathfrak{a}$, we obtain

$$\frac{\prod_{\alpha\in\Phi^+}\langle\alpha, \rho_{\mathfrak{g}}\rangle}{\prod_{\alpha\in\Phi_{\mathfrak{k}}^+}\langle\alpha, \rho_{\mathfrak{k}}\rangle} = (2\pi)^{\frac{n-m}{2}} \frac{\mathrm{vol}(K)\mathrm{vol}(H^d)}{\mathrm{vol}(G^d)\mathrm{vol}(T)} = (2\pi)^{\frac{n-m}{2}} \frac{\mathrm{vol}(S_A^d)}{\mathrm{vol}(X^d)} . \tag{20}$$

The second equality follows from the fact that the map $H^d/T \to X^d$, $hT \mapsto hK$, is an isometric embedding with image S^d_A. This proves (18). $\square$

In particular, specializing (13) and (18) to the case $m = 0$ we obtain as a consequence of Weyl's dimension formula

Corollary 5.2 *Let π be a discrete series representation of G having the same infinitesimal character as the finite-dimensional representation τ, then*

$$p_\pi = \frac{\dim \tau}{\mathrm{vol}(X^d)} \ .$$

Let us now give the promised analytic proof of (10). For a fixed finite-dimensional representation τ of G (which still is assumed to be connected) there are exactly $|W(\mathfrak{g},\mathfrak{t})|/|W(\mathfrak{k},\mathfrak{t})|$ equivalence classes of discrete series representations with infinitesimal character χ_τ (see [19], Thm. 8.7.1, or [12], Thm. 12.21). But this quotient of orders of Weyl groups is equal to $\chi(X^d)$ (see e.g. [3]). By (7), (9) and Corollary 5.2 we obtain

$$b^{(2)}_{\frac{n}{2}} = \mathrm{vol}(Y) \sum_{\pi\in\hat{G}_d, \chi_\pi=\chi_{\tau_0}} p_\pi = \mathrm{vol}(Y)\chi(X^d)\frac{1}{\mathrm{vol}(X^d)} \ .$$

We return to the evaluation of $k_X(t)$ for $m = 1$, $X = X_1$. The polynomial p_ξ only depends on the infinitesimal character of ξ. By $\Lambda_* \in i\mathfrak{t}^*$, $* = l, +, -$, we denote the infinitesimal character of the irreducible M-representation $\Lambda^*\mathfrak{n}$. We set

$$p_l(\nu) := \begin{cases} \prod_{\alpha\in\Phi^+} \frac{\langle\alpha,\Lambda_l+\nu\rangle}{\langle\alpha,\rho_\mathfrak{g}\rangle} & G = SL(3,\mathbb{R}) \text{ or } l \neq u \\ \prod_{\alpha\in\Phi^+} \frac{\langle\alpha,\Lambda_++\nu\rangle}{\langle\alpha,\rho_\mathfrak{g}\rangle} + \prod_{\alpha\in\Phi^+} \frac{\langle\alpha,\Lambda_-+\nu\rangle}{\langle\alpha,\rho_\mathfrak{g}\rangle} & \\ \quad = 2\prod_{\alpha\in\Phi^+} \frac{\langle\alpha,\Lambda_++\nu\rangle}{\langle\alpha,\rho_\mathfrak{g}\rangle} & \text{otherwise} \end{cases} .$$

For fixed infinitesimal character there are $|W(\mathfrak{m},\mathfrak{t})|/|W_{K_M}|$ equivalence classes of discrete series representations, where $W_{K_M} = \{k \in K_M \mid \mathrm{Ad}(k)\mathfrak{t} \subset \mathfrak{t}\}/T$. Note that there is an embedding $W(\mathfrak{k}_\mathfrak{m},\mathfrak{t}) \hookrightarrow W_{K_M}$ which becomes an isomorphism if K_M is connected. Furthermore, $|W_A| = 1$ except for $G = SO(1,q)^0$, where $|W_A| = 2$. In any case $|W_{K_M}||W_A| = 2|W(\mathfrak{k}_\mathfrak{m},\mathfrak{t})|$. Let X^d_M be the compact dual of $X_M = M/K_M = M^0/K^0_M$. Then $|W(\mathfrak{m},\mathfrak{t})|/|W_{K_M}||W_A| = \frac{1}{2}\chi(X^d_M)$. In addition, $u + v = \frac{n-1}{2}$ and

$$\frac{\mathrm{vol}(S^d_A)}{2\pi} = \frac{1}{\|\alpha_0\|} \ .$$

Summarizing the above discussion we obtain

$$k_X(t) = (-1)^{\frac{n-1}{2}} \frac{\chi(X^d_M)}{4\|\alpha_0\|\mathrm{vol}(X^d)} \sum_{l=0}^{2u} (-1)^{l+1} k_l(t) \ ,$$

where

$$k_l(t) = \int_{\mathfrak{a}^*} e^{-t(\|\nu\|^2+(u-l)^2\|\alpha_0\|^2)} p_l(i\nu) d\nu \ .$$

For an even polynomial P and $c \geq 0$ set

$$k_{P,c}(t) = \int_{-\infty}^{\infty} e^{-t(y^2+c^2)} P(iy) dy \ .$$

Then (compare [6], Lemma 2 and Lemma 3, [17], Lemma 6.4, [11], p.332)

$$\begin{aligned}
&\frac{d}{ds}_{|s=0}\left(\frac{1}{\Gamma(s)}\int_0^{\varepsilon} k_{P,c}(t)t^{s-1}dt\right) + \int_{\varepsilon}^{\infty} k_{P,c}(t)t^{-1}dt \\
&= \left(\int_0^{\varepsilon} k_{P,c}(t)t^{s-1}dt\right)_{|s=0} + \int_{\varepsilon}^{\infty} k_{P,c}(t)t^{-1}dt \\
&= -2\pi \int_0^c P(y)\, dy \ .
\end{aligned}$$

Since $p_l = p_{2u-l}$ we obtain

Proposition 5.3 *Let $X = G/K$ with $m(X) = 1$, $P = MAN \subset G$ a fundamental parabolic subgroup, $u = \frac{1}{2}\dim \mathfrak{n}$, and $X_M = M/K_M$. Then*

$$T^{(2)}(X) = (-1)^{\frac{n-1}{2}} \chi(X_M^d) \frac{\pi Q_X}{\mathrm{vol}(X^d)} \ ,$$

where

$$Q_X = \sum_{l=0}^{u-1} (-1)^l \int_0^{u-l} p_l(y \cdot \alpha_0)\, dy \ .$$

In fact, we have proved the proposition for irreducible X, but (16) shows that it holds in the general case, too.
Let us first discuss the case $G = SO(p,q)^0$, $X = X_{p,q}$, $p \leq q$ odd. Then $n = pq$ and $X_M = X_{p-1,q-1}$. Furthermore, the polynomials p_l depend only on the complexification of G, i.e., on $p+q$. Thus $Q_{X_{p,q}} = Q_{H^{p+q-1}}$. This proves Proposition 1.3 (a). We emphasize again that $Q_{H^{p+q-1}}$ is a positive rational number [11]. In fact, Hess-Schick showed that

$$(-1)^l \int_0^{u-l} p_l(y \cdot \alpha_0)\, dy > 0 \ , \qquad l = 0, \dots, u-1 \ .$$

Let $X = SL(3,\mathbb{R})/SO(3)$. Then $n = 5$ and $u = 1$. We find that

$$Q_X = \int_0^1 y^2 dy = \frac{1}{3} \ , \quad X_M^d = S^2, \ \chi(X_M^d) = 2 \ .$$

Using for instance (20) also $\mathrm{vol}(X^d)$ can be easily computed. This proves Proposition 1.4 and finishes the proof of Theorem 1.1.

References

[1] A. Borel. The L^2-cohomology of negatively curved Riemannian symmetric spaces. *Ann. Acad. Sci. Fenn. Ser. A I Math.* 10 (1985), 95–105.

[2] A. Borel and N. Wallach. *Continuous Cohomology, Discrete Subgroups, and Representations of Reductive Groups.* Princeton University Press, 1980.

[3] R. Bott. The index theorem for homogeneous differential operators. In *Differential and Combinatorial Topology. A Symposion in Honour of Marston Morse*, pages 167–186. Princeton University Press, 1965.

[4] J. Dodziuk. L^2-harmonic forms on rotationally symmetric Riemannian manifolds. *Proc. Amer. Math. Soc.* 77 (1979), 395–400.

[5] H. Donnelly. The differential form spectrum of hyperbolic space. *manuscripta math.* 33 (1981), 365–385.

[6] D. Fried. Analytic torsion and closed geodesics on hyperbolic manifolds. *Invent. Math.* 84 (1986), 523–540.

[7] Harish-Chandra. Discrete series for semisimple Lie groups II. *Acta Math.* 116 (1966), 1–111.

[8] Harish-Chandra. Harmonic analysis on real reductive groups I. The theory of the constant term. *J. Funct. Anal.* 19 (1975), 104–204.

[9] Harish-Chandra. Harmonic analysis on real reductive groups III. The Maass-Selberg relations and the Plancherel formula. *Ann. of Math.* 104 (1976), 117–201.

[10] E. Hess. *Simpliziales Volumen und L^2-Invarianten bei asphärischen Mannigfaltigkeiten.* Dissertation Universität Mainz, 1998. Available at wwwmath.uni-muenster.de/u/lueck/publ.

[11] E. Hess and T. Schick. L^2-torsion of hyperbolic manifolds. *manuscripta math.* 97 (1998), 329–334.

[12] A. W. Knapp. *Representation Theory of Semisimple Lie Groups. An Overview Based on Examples.* Princeton University Press, 1986.

[13] N. Lohoue and S. Mehdi. The Novikov-Shubin invariants for locally symmetric spaces. *J. Math. Pures Appl.* 79, 2 (2000), 111–140.

[14] J. Lott. Heat kernels on covering spaces and topological invariants. *J. Differential Geom.* 35 (1992), 471–510.

[15] W. Lück. L^2-invariants of regular coverings of compact manifolds and CW-complexes. In *Handbook of geometric topology*, pages 735–817. Elsevier, Amsterdam, 2002.

[16] W. Lück. *L^2-Invariants: Theory and Applications to Geometry and K-Theory.* Springer Verlag, 2002.

[17] V. Mathai. L^2-analytic torsion. *J. Funct. Anal.* 107 (1992), 369–386.

[18] H. Moscovici and R. Stanton. R-torsion and zeta functions for locally symmetric manifolds. *Invent. Math.* 105 (1991), 185–216.

[19] N. R. Wallach. *Real Reductive Groups.* Academic Press, 1988.

[20] N. R. Wallach. *Real Reductive Groups II.* Academic Press, 1992.

Martin Olbrich
Mathematisches Institut
Universität Göttingen
Bunsenstr. 3-5
37073 Göttingen
Germany
olbrich@uni-math.gwdg.de

The Farrell Cohomology of SP$(p-1,\mathbb{Z})$

Cornelia Busch

Received: March 24, 2002

Communicated by Günter M. Ziegler

Abstract. Let p be an odd prime with odd relative class number h^-. In this article we compute the Farrell cohomology of $\mathrm{Sp}(p-1,\mathbb{Z})$, the first p-rank one case. This allows us to determine the p-period of the Farrell cohomology of $\mathrm{Sp}(p-1,\mathbb{Z})$, which is $2y$, where $p-1=2^r y$, y odd. The p-primary part of the Farrell cohomology of $\mathrm{Sp}(p-1,\mathbb{Z})$ is given by the Farrell cohomology of the normalizers of the subgroups of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$. We use the fact that for odd primes p with h^- odd a relation exists between representations of $\mathbb{Z}/p\mathbb{Z}$ in $\mathrm{Sp}(p-1,\mathbb{Z})$ and some representations of $\mathbb{Z}/p\mathbb{Z}$ in $\mathrm{U}((p-1)/2)$.

2000 Mathematics Subject Classification: 20G10
Keywords and Phrases: Cohomology theory

1 Introduction

We define a homomorphism

$$\begin{array}{cccc} \phi: & \mathrm{U}(n) & \longrightarrow & \mathrm{Sp}(2n,\mathbb{R}) \\ & X=A+iB & \longmapsto & \begin{pmatrix} A & B \\ -B & A \end{pmatrix} =: \phi(X) \end{array}$$

where A and B are real matrices. Then ϕ is injective and maps $\mathrm{U}(n)$ on a maximal compact subgroup of $\mathrm{Sp}(2n,\mathbb{R})$. This homomorphism allows to consider each representation

$$\widetilde{\rho}:\mathbb{Z}/p\mathbb{Z}\longrightarrow \mathrm{U}((p-1)/2)$$

as a representation

$$\phi\circ\widetilde{\rho}:\mathbb{Z}/p\mathbb{Z}\longrightarrow \mathrm{Sp}(p-1,\mathbb{R}).$$

In an article of Busch [6] it is determined which properties $\widetilde{\rho}$ has to fulfil for $\phi\circ\widetilde{\rho}$ to be conjugate in $\mathrm{Sp}(p-1,\mathbb{R})$ to a representation

$$\rho:\mathbb{Z}/p\mathbb{Z}\longrightarrow \mathrm{Sp}(p-1,\mathbb{Z}).$$

Theorem 2.2. *Let $X \in \mathrm{U}((p-1)/2)$ be of odd prime order p. We define $\phi : \mathrm{U}((p-1)/2) \to \mathrm{Sp}(p-1,\mathbb{R})$ as above. Then $\phi(X) \in \mathrm{Sp}(p-1,\mathbb{R})$ is conjugate to $Y \in \mathrm{Sp}(p-1,\mathbb{Z})$ if and only if the eigenvalues $\lambda_1,\dots,\lambda_{(p-1)/2}$ of X are such that*

$$\left\{\lambda_1,\dots,\lambda_{(p-1)/2},\overline{\lambda}_1,\dots,\overline{\lambda}_{(p-1)/2}\right\}$$

is a complete set of primitive p-th roots of unity.

The proof of Theorem 2.2 involves the theory of cyclotomic fields. For the p-primary component of the Farrell cohomology of $\mathrm{Sp}(p-1,\mathbb{Z})$, the following holds:

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})_{(p)} \cong \prod_{P\in\mathfrak{P}} \widehat{\mathrm{H}}^*(N(P),\mathbb{Z})_{(p)}$$

where $\mathfrak{P}$ is a set of representatives for the conjugacy classes of subgroups of order p of $\mathrm{Sp}(p-1,\mathbb{Z})$ and $N(P)$ denotes the normalizer of $P \in \mathfrak{P}$. This property also holds if we consider $\mathrm{GL}(p-1,\mathbb{Z})$ instead of the symplectic group. This fact was used by Ash in [1] to compute the Farrell cohomology of $\mathrm{GL}(n,\mathbb{Z})$ with coefficients in $\mathbb{F}_p$ for $p-1 \leqslant n < 2p-2$. Moreover, we have

$$\widehat{\mathrm{H}}^*(N(P),\mathbb{Z})_{(p)} \cong \left(\widehat{\mathrm{H}}^*(C(P),\mathbb{Z})_{(p)}\right)^{N(P)/C(P)}$$

where $C(P)$ is the centralizer of P. We will determine the structure of $C(P)$ and of $N(P)/C(P)$. After that we will compute the number of conjugacy classes of those subgroups for which $N(P)/C(P)$ has a given structure. Here again arithmetical questions are involved. In the articles of Brown [2] and Sjerve and Yang [9] is shown that the number of conjugacy classes of elements of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$ is $2^{(p-1)/2}h^-$ where h^- denotes the relative class number of the cyclotomic field $\mathbb{Q}(\xi)$, ξ a primitive p-th root of unity. If h^- is odd, each conjugacy class of matrices of order p in $\mathrm{Sp}(p-1,\mathbb{R})$ that lifts to $\mathrm{Sp}(p-1,\mathbb{Z})$ splits into h^- conjugacy classes in $\mathrm{Sp}(p-1,\mathbb{Z})$. The main results in this article are

Theorem 3.7. *Let p be an odd prime for which h^- is odd. Then*

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})_{(p)} \cong \prod_{\substack{k\mid p-1\\ k\ odd}} \left(\prod_{1}^{\widetilde{\mathcal{K}}_k} \mathbb{Z}/p\mathbb{Z}[x^k,x^{-k}]\right),$$

where $\widetilde{\mathcal{K}}_k$ denotes the number of conjugacy classes of subgroups of order p of $\mathrm{Sp}(p-1,\mathbb{Z})$ for which $|N/C| = k$. Moreover $\widetilde{\mathcal{K}}_k \geqslant \mathcal{K}_k$, where $\mathcal{K}_k$ is the number of conjugacy classes of subgroups of $\mathrm{U}((p-1)/2)$ with $|N/C| = k$. As usual N denotes the normalizer and C the centralizer of the corresponding subgroup.

Theorem 3.8. *Let p be an odd prime for which h^- is odd and let y be such that $p-1 = 2^r y$ and y is odd. Then the period of $\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})_{(p)}$ is $2y$.*

Corresponding results have been shown for other groups, for example $\mathrm{GL}(n,\mathbb{Z})$ in the p-rank one case [1], the mapping class group [8] and the outerautomorphism group of the free group in the p-rank one case [7].
This article presents results of my doctoral thesis, which I wrote at the ETH Zürich under the supervision of G. Mislin. I thank G. Mislin for the suggestion of this interesting subject.

2 The symplectic group

2.1 Definition

Let R be a commutative ring with 1. The general linear group $\mathrm{GL}(n,R)$ is defined to be the multiplicative group of invertible $n\times n$-matrices over R.

Definition. The symplectic group $\mathrm{Sp}(2n,R)$ over the ring R is the subgroup of matrices $Y\in\mathrm{GL}(2n,R)$ that satisfy

$$Y^{\mathrm{T}}JY=J:=\begin{pmatrix}0 & I_n\\ -I_n & 0\end{pmatrix}$$

where I_n is the $n\times n$-identity matrix.

It is the group of isometries of the skew-symmetric bilinear form

$$\begin{array}{llll}\langle\ ,\ \rangle: & R^{2n}\times R^{2n} & \longrightarrow & R\\ & (x,y) & \longmapsto & \langle x,y\rangle:=x^{\mathrm{T}}Jy.\end{array}$$

It follows from a result of Bürgisser [4] that elements of odd prime order p exist in $\mathrm{Sp}(2n,\mathbb{Z})$ if and only if $2n\geqslant p-1$.

Proposition 2.1. *The eigenvalues of a matrix $Y\in\mathrm{Sp}(p-1,\mathbb{Z})$ of odd prime order p are the primitive p-th roots of unity, hence the zeros of the polynomial*

$$m(x)=x^{p-1}+\cdots+x+1.$$

Proof. If λ is an eigenvalue of Y, we have $\lambda=1$ or $\lambda=\xi$, a primitive p-th root of unity, and the characteristic polynomial of Y divides x^p-1 and has integer coefficients. Since $m(x)$ is irreducible over $\mathbb{Q}$, the claim follows. □

2.2 A relation between $\mathrm{U}\left(\frac{p-1}{2}\right)$ and $\mathrm{Sp}(p-1,\mathbb{Z})$

Let $X\in\mathrm{U}(n)$, i.e., $X\in\mathrm{GL}(n,\mathbb{C})$ and $X^*X=I_n$ where $X^*=\overline{X}^{\mathrm{T}}$ and I_n is the $n\times n$-identity matrix. We can write $X=A+iB$ with $A,B\in\mathrm{M}(n,\mathbb{R})$, the ring of real $n\times n$-matrices. We now define the following map

$$\begin{array}{llll}\phi: & \mathrm{U}(n) & \longrightarrow & \mathrm{Sp}(2n,\mathbb{R})\\ & X=A+iB & \longmapsto & \begin{pmatrix}A & B\\ -B & A\end{pmatrix}=:\phi(X).\end{array}$$

The map ϕ is an injective homomorphism. Moreover, it is well-known that ϕ maps $\mathrm{U}(n)$ on a maximal compact subgroup of $\mathrm{Sp}(2n,\mathbb{R})$.

Theorem 2.2. *Let $X \in \mathrm{U}((p-1)/2)$ be of odd prime order p. We define $\phi : \mathrm{U}((p-1)/2) \to \mathrm{Sp}(p-1,\mathbb{R})$ as above. Then $\phi(X) \in \mathrm{Sp}(p-1,\mathbb{R})$ is conjugate to $Y \in \mathrm{Sp}(p-1,\mathbb{Z})$ if and only if the eigenvalues $\lambda_1,\dots,\lambda_{(p-1)/2}$ of X are such that*

$$\left\{\lambda_1,\dots,\lambda_{(p-1)/2},\overline{\lambda}_1,\dots,\overline{\lambda}_{(p-1)/2}\right\}$$

is a complete set of primitive p-th roots of unity.

Proof. See [5] or [6]. □

In the proof of Theorem 2.2 we used the following facts. For a primitive p-th root of unity ξ, we consider the cyclotomic field $\mathbb{Q}(\xi)$. It is well-known that $\mathbb{Q}(\xi+\xi^{-1})$ is the maximal real subfield of $\mathbb{Q}(\xi)$, and that $\mathbb{Z}[\xi]$ and $\mathbb{Z}[\xi+\xi^{-1}]$ are the rings of integers of $\mathbb{Q}(\xi)$ and $\mathbb{Q}(\xi+\xi^{-1})$ respectively. Let $(\mathfrak{a},a)$ denote a pair where $\mathfrak{a} \subseteq \mathbb{Z}[\xi]$ and $a \in \mathbb{Z}[\xi]$ are chosen such that $\mathfrak{a} \neq 0$ is an ideal in $\mathbb{Z}[\xi]$ and $\mathfrak{a}\overline{\mathfrak{a}} = (a)$, a principal ideal. Here $\overline{\mathfrak{a}}$ denotes the complex conjugate of $\mathfrak{a}$. We define an equivalence relation on the set of those pairs by $(\mathfrak{a},a) \sim (\mathfrak{b},b)$ if and only if $\lambda,\mu \in \mathbb{Z}[\xi]\setminus\{0\}$ exist such that $\lambda\mathfrak{a} = \mu\mathfrak{b}$ and $\lambda\overline{\lambda}a = \mu\overline{\mu}b$. We denote by $[\mathfrak{a},a]$ the equivalence class of the pair $(\mathfrak{a},a)$ and by $\mathcal{P}$ the set of equivalence classes $[\mathfrak{a},a]$.

Let $\mathcal{S}_p$ denote the set of conjugacy classes of elements of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$. Sjerve and Yang have shown in [9] that a bijection exists between $\mathcal{P}$ and $\mathcal{S}_p$. If $Y \in \mathrm{Sp}(p-1,\mathbb{Z})$ is a matrix of order p, then the equivalence class $[\mathfrak{a},a] \in \mathcal{P}$ corresponding to the conjugacy class of Y in $\mathrm{Sp}(p-1,\mathbb{Z})$ can be determined in the following way. Let $\alpha = (\alpha_1,\dots,\alpha_{p-1})^{\mathrm{T}}$ be an eigenvector of Y corresponding to the eigenvalue $\xi = e^{i2\pi/p}$, that is $Y\alpha = \xi\alpha$. Then $\alpha_1,\dots,\alpha_{p-1}$ is a basis of an ideal $\mathfrak{a} \subseteq \mathbb{Z}[\xi]$. Sjerve and Yang [9] proved that this ideal $\mathfrak{a}$ has the property $[\mathfrak{a},a] \in \mathcal{P}$. Let h and h^+ be the class numbers of $\mathbb{Q}(\xi)$ and $\mathbb{Q}(\xi+\xi^{-1})$ respectively. Then $h^- := h/h^+$ denotes the relative class number. Sjerve and Yang [9] showed that the number of conjugacy classes of matrices of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$ is $h^-\,2^{(p-1)/2}$. The number of conjugacy classes in $\mathrm{U}((p-1)/2)$ of unitary matrices that satisfy the condition in Theorem 2.2 is $2^{(p-1)/2}$.

Let $\mathcal{U}_p$ denote the set of conjugacy classes of matrices in $\mathrm{U}((p-1)/2)$ that satisfy the condition on the eigenvalues that is given in Theorem 2.2. A consequence of Theorem 2.2 is that it is possible to define a map

$$\Psi : \mathcal{S}_p \longrightarrow \mathcal{U}_p$$

and that this map is surjective. Therefore the map

$$\psi : \mathcal{P} \longrightarrow \mathcal{U}_p$$

is surjective either.

For a given choice of the ideal $\mathfrak{a}$ (for example $\mathfrak{a} = \mathbb{Z}[\xi]$), we denote by $\mathcal{P}_{\mathfrak{a}}$ the set of those classes $[\mathfrak{a},a] \in \mathcal{P}$, where $\mathfrak{a}$ corresponds to our choice. If the restriction

$$\psi|_{\mathcal{P}_{\mathfrak{a}}} : \mathcal{P}_{\mathfrak{a}} \longrightarrow \mathcal{U}_p$$

is surjective each conjugacy class in $\mathcal{U}_p$ of matrices that satisfy Theorem 2.2 yields h^- conjugacy classes in $\mathrm{Sp}(p-1,\mathbb{Z})$. In general $\psi|_{\mathcal{P}_\mathfrak{a}}$ is not surjective. It is a result of Busch, [5], [6], that $\psi|_{\mathcal{P}_\mathfrak{a}}$ is surjective if h^- is odd. If h^- is even and h^+ is odd, we have no surjectivity of $\psi|_{\mathcal{P}_\mathfrak{a}}$. This happens for example for the primes 29 and 113.

2.3 Subgroups of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$

It follows from Theorem 2.2 that a mapping exists that sends the conjugacy classes of matrices $Y \in \mathrm{Sp}(p-1,\mathbb{Z})$ of odd prime order p onto the conjugacy classes of matrices X in $\mathrm{U}((p-1)/2)$ that satisfy the condition on the eigenvalues described in Theorem 2.2. This mapping is surjective.
It is clear that $\det X = e^{l2\pi i/p}$ for some $1 \leqslant l \leqslant p$. If $X \in \mathrm{U}((p-1)/2)$ satisfies the condition on the eigenvalues, then so does X^k, $k = 1,\ldots,p-1$. If $\det X = e^{l2\pi i/p}$ for some $1 \leqslant l \leqslant p-1$, then

$$\left\{\det X,\ldots,\det X^{p-1}\right\} = \left\{e^{i2\pi/p},\ldots,e^{i(p-1)2\pi/p}\right\}$$

and the X^k are in different conjugacy classes. If $\det X = 1$, it is possible that some k exists such that X and X^k are in the same conjugacy class. In this section we will analyse when and how many times this happens. The number of conjugacy classes of matrices $X \in \mathrm{U}((p-1)/2)$ that satisfy the condition required in Theorem 2.2 is $2^{(p-1)/2}$. Herewith we will be able to compute the number of conjugacy classes of subgroups of matrices of order p in $\mathrm{U}((p-1)/2)$. We remember that the number of conjugacy classes of matrices of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$ is $2^{(p-1)/2}h^-$. If $h^- = 1$, a bijection exists between the conjugacy classes of matrices of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$ and the conjugacy classes of matrices of order p in $\mathrm{U}((p-1)/2)$ that satisfy the condition required in Theorem 2.2.
Let $X \in \mathrm{U}((p-1)/2)$ with $X^p = 1$, $X \neq 1$. Then X generates a subgroup S of order p in $\mathrm{U}((p-1)/2)$. If $\det X = 1$, it is possible that X is conjugate to $X' \in S$ with $X \neq X'$. Two matrices in $\mathrm{U}((p-1)/2)$ are conjugate to each other if and only if they have the same eigenvalues. The set of eigenvalues of X is

$$\left\{e^{ig_1 2\pi/p},\ldots,e^{ig_{(p-1)/2}2\pi/p}\right\}$$

where $1 \leqslant g_l \leqslant p-1$ for $l = 1,\ldots,\frac{p-1}{2}$ and for all $l \neq j$, $l,j = 1,\ldots,(p-1)/2$, $g_l \neq p-g_j$ and $g_l \neq g_j$. From now on we consider the g_j as elements of $(\mathbb{Z}/p\mathbb{Z})^*$. The matrix X is conjugate to X^κ for some κ if the eigenvalues of X and X^κ are the same. This is equivalent to

$$\{g_1,\ldots,g_{(p-1)/2}\} = \{\kappa g_1,\ldots,\kappa g_{(p-1)/2}\} \subset (\mathbb{Z}/p\mathbb{Z})^*$$

where g_j and κg_j, $j = 1,\ldots,(p-1)/2$, denote the corresponding congruence classes.

We introduce some notation that will be used in the whole section. Let

$$G := \{g_1, \dots, g_{(p-1)/2}\} \subset (\mathbb{Z}/p\mathbb{Z})^*,$$
$$\kappa G := \{\kappa g_1, \dots, \kappa g_{(p-1)/2}\} \subset (\mathbb{Z}/p\mathbb{Z})^*$$

for some $\kappa \in (\mathbb{Z}/p\mathbb{Z})^*$. Let x be a generator of the multiplicative cyclic group $(\mathbb{Z}/p\mathbb{Z})^*$ and let K be a subgroup of $(\mathbb{Z}/p\mathbb{Z})^*$ with $|K| = k$. Then K is cyclic and k divides $p-1$. Let $m := (p-1)/k$, then x^m generates K.
First we prove the following proposition.

Proposition 2.3. *Let $G \subset (\mathbb{Z}/p\mathbb{Z})^*$ be a subset with $|G| = (p-1)/2$. The following are equivalent.*

i) For all $g_j, g_l \in G$, $g_j \neq -g_l$ and $\kappa \in (\mathbb{Z}/p\mathbb{Z})^*$ exists with $\kappa G = G$, $\kappa \neq 1$.

ii) An integer $h \in \mathbb{N}$, $1 \leqslant h \leqslant (p-1)/2$, and $n_j \in (\mathbb{Z}/p\mathbb{Z})^*$, $j = 1, \dots, h$, exist with

$$G = \bigcup_{j=1}^{h} n_j K$$

where

- *$K \subset (\mathbb{Z}/p\mathbb{Z})^*$ is the subgroup generated by κ,*
- *the order of K is odd,*
- *for $\kappa' \in K$ and all $j, l = 1, \dots, h$, $n_j \neq -n_l \kappa'$,*
- *and for all $j = 2, \dots, h$, $n_j \notin K$.*

Then we will analyse the uniqueness of this decomposition of G. This will enable us to determine the number of $G \subset (\mathbb{Z}/p\mathbb{Z})^*$ with $|G| = (p-1)/2$ and $G = \kappa G$ for some $1 \neq \kappa \in (\mathbb{Z}/p\mathbb{Z})^*$. Herewith we will determine the number of conjugacy classes of subgroups of order p in $\mathrm{U}((p-1)/2)$ whose group elements satisfy the condition of Theorem 2.2.

Definition. Let $\kappa \in (\mathbb{Z}/p\mathbb{Z})^*$ and let K be the subgroup of $(\mathbb{Z}/p\mathbb{Z})^*$ generated by κ. Let $G \subset (\mathbb{Z}/p\mathbb{Z})^*$ be a subset with $|G| = (p-1)/2$. We say that K decomposes G if G, κ and K fulfil the conditions of Proposition 2.3.

So K decomposes G if the order of the group K is odd and G is a disjoint union of cosets $n_1 K, \dots, n_h K$ of K in $(\mathbb{Z}/p\mathbb{Z})^*$ for which for all n_j, n_l, $j, l = 1, \dots, h$, holds $n_j K \neq -n_l K$.

Lemma 2.4. *Let $G \subset (\mathbb{Z}/p\mathbb{Z})^*$ with $|G| = (p-1)/2$. Then $1 \neq \kappa \in (\mathbb{Z}/p\mathbb{Z})^*$ exists with $\kappa G = G$ if and only if $1 \leqslant h \leqslant (p-1)/2$ and $n_j \in (\mathbb{Z}/p\mathbb{Z})^*$, $j = 1, \dots, h$, exist with*

$$G = \bigcup_{j=1}^{h} n_j K$$

where $n_j \notin K$ for $j = 2, \dots, h$, and K is the subgroup of $(\mathbb{Z}/p\mathbb{Z})^*$ that is generated by κ.

Proof. $\underline{\Leftarrow}$: Let $\kappa^l \in K$. Then

$$\kappa^l G = \kappa^l \bigcup_{j=1}^{h} n_j K = \bigcup_{j=1}^{h} n_j \kappa^l K = \bigcup_{j=1}^{h} n_j K = G.$$

$\underline{\Rightarrow}$: Without loss of generality we assume that $1 \in G$. If $1 \notin G$, $\lambda \in (\mathbb{Z}/p\mathbb{Z})^*$ exists with $1 \in \lambda G$ because $(\mathbb{Z}/p\mathbb{Z})^*$ is a multiplicative group. Of course $\kappa\lambda G = \lambda G$. Moreover, it is easy to see that if λG is a union of cosets of K, this is also true for G. The equation $\kappa G = G$ implies that $KG = G$. If $1 \in G$, then $K \subseteq G$ since $KG = G$. If $K = G$, we have finished the proof. If $K \neq G$, we consider $G'_1 = G \setminus K$. For all $\kappa^l \in K$ we have $\kappa^l K = K$ and

$$\kappa^l G'_1 = \kappa^l (G \setminus K) = G \setminus K = G'_1.$$

Now $\lambda_1 \in (\mathbb{Z}/p\mathbb{Z})^*$ exists with $1 \in \lambda_1 G'_1 =: G_1$. Then $G = K \cup \lambda_1^{-1} G_1$ and we can repeat the construction on G_1 instead of G. This procedure finishes after $h := (p-1)/2k$ steps. Let $n_1 := 1$ and for $j = 2, \ldots, h$ let $n_j := n_{j-1}\lambda_{j-1}^{-1}$. Then $G = \bigcup_{j=1}^{h} n_j K$. □

Let $G = \{g_1, \ldots, g_{(p-1)/2}\} \subset (\mathbb{Z}/p\mathbb{Z})^*$ with $|G| = (p-1)/2$ and $\kappa G = G$ for some $\kappa \in (\mathbb{Z}/p\mathbb{Z})^*$ with $\kappa \neq 1$, $\kappa^k = 1$. The following lemma will give an answer to the question when G satisfies the conditions $g_l \neq g_j$, $g_l \neq -g_j$ for all $j \neq l$ with $j, l = 1, \ldots, \frac{p-1}{2}$.

Lemma 2.5. *Let $G = \bigcup_{j=1}^{h} n_j K \subset (\mathbb{Z}/p\mathbb{Z})^*$ be defined like in Lemma 2.4. Then for all $g_j, g_l \in G$ holds $g_j \neq -g_l$ if and only if $-1 \notin K$ and for all $\kappa \in K$ and all $j, l = 1, \ldots, h$ holds $n_j \neq -n_l \kappa$.*

Proof. $\underline{\Rightarrow}$: Suppose $-1 \in K$. Then $-1 = \kappa^l$ for some l and $n_1 = -n_1 \kappa^l$. But then we have found $g_1 := n_1 \in G$ and $g_2 := n_1 \kappa^l \in G$ with $g_1 = -g_2$.
$\underline{\Leftarrow}$: Suppose $g_j, g_l \in G$ exist with $g_j = -g_l$. Let $g_j = n_j \kappa^j$, $g_l = n_l \kappa^l$. Then $n_j \kappa^j = -n_l \kappa^l$, and we have found $\kappa^{j-l} \in K$ with $n_l = -n_j \kappa^{j-l}$. □

Which subgroups $K \subseteq (\mathbb{Z}/p\mathbb{Z})^*$ satisfy the condition $-1 \notin K$?

Lemma 2.6. *Let $K \subseteq (\mathbb{Z}/p\mathbb{Z})^*$ be a subgroup of order k. Then $-1 \notin K$ if and only if k is odd.*

Proof. The group $(\mathbb{Z}/p\mathbb{Z})^*$ is cyclic of order $p-1$ and K is a cyclic group. Let x be a generator of K, then $x^k = 1$. If k is even, $k/2 \in \mathbb{Z}$ and $x^{k/2} \in K$. But then $(x^{k/2})^2 = x^k = 1$ and therefore $x^{k/2} = -1 \in K$ since -1 is the element of order 2 in $(\mathbb{Z}/p\mathbb{Z})^*$. On the other hand if $-1 \in K$, then K contains an element of order 2. But then k is even, since the order of any element of K divides the order of K. □

Proof of Proposition 2.3. A subgroup K decomposes a set G as required in Lemma 2.5 if and only if the order of K is odd. Moreover, the order of K divides $p-1$. Now Proposition 2.3 follows from Lemma 2.4 and Lemma 2.5. □

We did not yet analyse the uniqueness of the decomposition of a set G. It is evident that the n_j can be permuted and multiplied with any $\kappa^l \in K$, but we will see that K and h are not uniquely determined. The next lemma states that if K decomposes G then so does any nontrivial subgroup of K.

Lemma 2.7. *Let $G = \bigcup_{j=1}^h n_j K \subset (\mathbb{Z}/p\mathbb{Z})^*$, $|G| = (p-1)/2$, be such that K decomposes G (Proposition 2.3). Let $|K| = k$ be not a prime and let $K' \neq K$ be a nontrivial subgroup of K. Then K' decomposes G.*

Proof. Since K' is a subgroup of K, K can be written as a union of cosets of K' in K. Moreover, G is a union of cosets of K in $(\mathbb{Z}/p\mathbb{Z})^*$. Therefore

$$G = \bigcup_{j=1}^{h} n_j K = \bigcup_{i=1}^{h'} n'_i K'.$$

Since K decomposes G, we have $n_l K \neq -n_j K$ for all $l, j = 1, \ldots, h$. This implies that $n'_l K' \neq -n'_i K'$ for all $i, l = 1, \ldots, h'$. So K' decomposes G. □

Our next aim is to determine the number of sets G. Therefore we consider for a given G the group K with $|K|$ maximal and K decomposes G.

Lemma 2.8. *Let $K \subset (\mathbb{Z}/p\mathbb{Z})^*$ be a nontrivial subgroup of odd order k. Then $2^{(p-1)/2k}$ different sets G exist such that K decomposes G and $|G| = (p-1)/2$.*

Proof. The order of $K \subset (\mathbb{Z}/p\mathbb{Z})^*$ is odd. Then it follows from Lemma 2.6 that $-1 \notin K$. Consider the cosets $n_j K$ of K in $(\mathbb{Z}/p\mathbb{Z})^*$. Since $-1 \notin K$, we have $n_j K \neq -n_j K$. So n_j, $j = 1, \ldots, (p-1)/2k$, exist such that

$$(\mathbb{Z}/p\mathbb{Z})^* = \bigcup_{j=1}^{(p-1)/2k} (n_j K \cup -n_j K).$$

The group K decomposes G if and only if G is a union of cosets of K and $m_j K \subseteq G$ implies that $-m_j K \not\subseteq G$ for $m_j = \pm n_j$, $j = 1, \ldots, (p-1)/2k$. Therefore $2^{(p-1)/2k}$ sets G exist such that K decomposes G. □

Definition. Let $K \subset (\mathbb{Z}/p\mathbb{Z})^*$ be a group of odd order k. We define $\mathcal{N}_k$ to be the number of $G \subset (\mathbb{Z}/p\mathbb{Z})^*$ such that K decomposes G but any K' with $K \subset K' \subset (\mathbb{Z}/p\mathbb{Z})^*$, $K \neq K'$, does not decompose G.

To determine $\mathcal{N}_k$ we have to subtract the number $\mathcal{N}_{k'}$ from $2^{(p-1)/2k}$ for each odd $k' \neq k$ with $k|k'$, $k'|p-1$. The integer k' is the order of the group K' with $K \subset K'$. Therefore we get a recursive formula

$$\mathcal{N}_k = 2^{(p-1)/2k} - \sum_{\substack{k' \text{ odd},\ k'>k \\ k|k',\ k'|p-1}} \mathcal{N}_{k'}.$$

Now it remains to determine $\mathcal{N}_y$. Let $y \in \mathbb{Z}$ be such that $p-1 = 2^r y$ and y is odd. Then

$$\mathcal{N}_y = 2^{(p-1)/2y} = 2^{2^{r-1}}.$$

Let $p-1 = 2^r p_1^{r_1} \dots p_l^{r_l}$ be a factorisation of $p-1$ into primes where $p_1, \dots, p_l$ are odd and $p_i \neq p_j$ for all $i \neq j$ with $i, j = 1, \dots, l$. Since $p-1$ is even, $r \geqslant 1$. Let K be of order $k = p_1^{s_1} \dots p_l^{s_l}$ where $0 \leqslant s_j \leqslant r_j$ for $j = 1, \dots, l$. Let x be a generator of $(\mathbb{Z}/p\mathbb{Z})^*$. Then K is generated by x^m, $m = 2^r p_1^{r_1-s_1} \dots p_l^{r_l-s_l}$. If $k' = p_1^{t_1} \dots p_l^{t_l}$ where $s_j \leqslant t_j \leqslant r_j$ for $j = 1, \dots, l$, then K is a proper subgroup of K' of order k' if $s_j < t_j$ for some $1 \leqslant j \leqslant l$. Herewith $-1 + \prod_{j=1}^{l}(r_j - s_j + 1)$ groups K' exist such that K is a proper subgroup of K'. So the number of sets G that are decomposed by K and for which no $K' \supsetneq K$ exists such that K' decomposes G is

$$\mathcal{N}_k = 2^{(p-1)/2k} - \sum_{y \in T_k} \mathcal{N}_y$$

where

$$T_k := \left\{ y \in \mathbb{N} \;\middle|\; y \text{ odd},\ k|y,\ y \neq k \text{ and } y|p-1 \right\}.$$

Now we have to determine the number of sets G that satisfy the conditions of Proposition 2.3. Let this be the number $\mathcal{N}_G$. One easily sees that

$$\mathcal{N}_G = \sum_{\substack{K \subset (\mathbb{Z}/p\mathbb{Z})^* \\ |K| \neq 1 \\ |K| \text{ odd}}} \mathcal{N}_{|K|} = \sum_{\substack{k|p-1 \\ k \neq 1 \\ k \text{ odd}}} \mathcal{N}_k.$$

Now let $G \subset (\mathbb{Z}/p\mathbb{Z})^*$ with $|G| = (p-1)/2$, such that for all $g_i, g_j \in G$, $g_i \neq -g_j$. Let $\mathcal{N}_1$ be the number of sets G for which no $\kappa \in (\mathbb{Z}/p\mathbb{Z})^*$, $\kappa \neq 1$, exists such that $\kappa G = G$. Then

$$\mathcal{N}_1 = 2^{(p-1)/2} - \mathcal{N}_G = 2^{(p-1)/2} - \sum_{\substack{1 \neq k | p-1 \\ k \text{ odd}}} \mathcal{N}_k.$$

We have seen that each set G corresponds to the set of eigenvalues of a matrix in $\mathrm{U}((p-1)/2)$ that satisfies Theorem 2.2.

Definition. We define a matrix $X_G \in \mathrm{U}\left(\frac{p-1}{2}\right)$ with the eigenvalues

$$\left\{ e^{i g_1 2\pi/p}, \dots, e^{i g_{(p-1)/2} 2\pi/p} \right\}$$

where $G = \{g_1, \dots, g_{(p-1)/2}\} \subset (\mathbb{Z}/p\mathbb{Z})^*$. We used the same notation for the elements of $(\mathbb{Z}/p\mathbb{Z})^*$ and their representatives in $\mathbb{Z}$.

Let the maximal order of K that decomposes G be k. Then G yields k elements of the group generated by X_G. As a result we have:

Proposition 2.9. *The number of conjugacy classes of subgroups of order p in* U$((p-1)/2)$ *whose group elements satisfy the necessary and sufficient condition is*

$$\mathcal{K}(p) = \frac{1}{p-1} \sum_{\substack{k \text{ } odd \\ k|p-1}} k\mathcal{N}_k.$$

3 The Farrell cohomology

3.1 An introduction to Farrell cohomology

An introduction to the Farrell cohomology can be found in the book of Brown [3]. The Farrell cohomology is a complete cohomology for groups with finite virtual cohomological dimension (vcd). It is a generalisation of the Tate cohomology for finite groups. If G is finite, the Farrell cohomology and the Tate cohomology of G coincide. It is well-known that the groups $\mathrm{Sp}(2n, \mathbb{Z})$ have finite vcd.

Definition. An elementary abelian p-group of rank $r \geqslant 0$ is a group that is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^r$.

It is well-known that $\widehat{\mathrm{H}}^i(G,\mathbb{Z})$ is a torsion group for every $i \in \mathbb{Z}$. We write $\widehat{\mathrm{H}}^i(G,\mathbb{Z})_{(p)}$ for the p-primary part of this torsion group, i.e., the subgroup of elements of order some power of p. We will use the following theorem.

Theorem 3.1. *Let G be a group such that $\operatorname{vcd} G < \infty$ and let p be a prime. Suppose that every elementary abelian p-subgroup of G has rank $\leqslant 1$. Then*

$$\widehat{\mathrm{H}}^*(G,\mathbb{Z})_{(p)} \cong \prod_{P \in \mathfrak{P}} \widehat{\mathrm{H}}^*(N(P),\mathbb{Z})_{(p)}$$

where $\mathfrak{P}$ is a set of representatives for the conjugacy classes of subgroups of G of order p and $N(P)$ denotes the normalizer of P.

Proof. See Brown's book [3]. □

We also have

$$\widehat{\mathrm{H}}^*(G,\mathbb{Z}) \cong \prod_{p} \widehat{\mathrm{H}}^*(G,\mathbb{Z})_{(p)}$$

where p ranges over the primes such that G has p-torsion.
A group G of finite virtual cohomological dimension is said to have periodic cohomology if for some $d \neq 0$ there is an element $u \in \widehat{\mathrm{H}}^d(G,\mathbb{Z})$ that is invertible in the ring $\widehat{\mathrm{H}}^*(G,\mathbb{Z})$. Cup product with u then gives a periodicity isomorphism $\widehat{\mathrm{H}}^i(G,M) \cong \widehat{\mathrm{H}}^{i+d}(G,M)$ for any G-module M and any $i \in \mathbb{Z}$. Similarly we say that G has p-periodic cohomology if the p-primary component $\widehat{\mathrm{H}}^*(G,\mathbb{Z})_{(p)}$, which is itself a ring, contains an invertible element of non-zero degree d. Then we have

$$\widehat{\mathrm{H}}^i(G,M)_{(p)} \cong \widehat{\mathrm{H}}^{i+d}(G,M)_{(p)},$$

and the smallest positive d that satisfies this condition is called the p-period of G.

Proposition 3.2. *The following are equivalent:*

i) G has p-periodic cohomology.

ii) Every elementary abelian p-subgroup of G has rank $\leqslant 1$.

Proof. See Brown's book [3]. □

3.2 Normalizers of subgroups of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$

In order to use Theorem 3.1, we have to analyse the structure of the normalizers of subgroups of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$. We already analysed the conjugacy classes of subgroups of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$. Let N be the normalizer and let C be the centralizer of such a subgroup. Then we have a short exact sequence

$$1 \longrightarrow C \longrightarrow N \longrightarrow N/C \longrightarrow 1.$$

Moreover, it follows from the discussion in the paper of Brown [2] that for p an odd prime

$$C \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}/2p\mathbb{Z},$$

and therefore N is a finite group. We will use the following proposition.

Proposition 3.3. *Let*

$$1 \longrightarrow U \longrightarrow G \longrightarrow Q \longrightarrow 1$$

be a short exact sequence with Q a finite group of order prime to p. Then

$$\widehat{\mathrm{H}}^*(G,\mathbb{Z})_{(p)} \cong \left(\widehat{\mathrm{H}}^*(U,\mathbb{Z})_{(p)}\right)^Q.$$

Proof. See Brown [3], the Hochschild-Serre spectral sequence. □

Applying this to our case, we get

$$\widehat{\mathrm{H}}^*(N,\mathbb{Z})_{(p)} \cong \left(\widehat{\mathrm{H}}^*(C,\mathbb{Z})_{(p)}\right)^{N/C}.$$

Therefore we have to determine N/C and its action on $C \cong \mathbb{Z}/2p\mathbb{Z}$. From now on, if we consider subgroups or elements of order p in $\mathrm{U}((p-1)/2)$, we mean those that satisfy the condition of Theorem 2.2. In what follows we assume that p is an odd prime for which $h^- = 1$, because in this case we have a bijection between the conjugacy classes of subgroups of order p in $\mathrm{U}((p-1)/2)$ and those in $\mathrm{Sp}(p-1,\mathbb{Z})$. Therefore, in order to determine the structure of the conjugacy classes of subgroups of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$, we can consider the corresponding conjugacy classes in $\mathrm{U}((p-1)/2)$. We have already seen that in a subgroup of $\mathrm{U}((p-1)/2)$ of order p different elements can be in the same conjugacy class. Let $\mathcal{N}_k$ be the number of conjugacy classes of elements of order p in $\mathrm{U}((p-1)/2)$ where k powers of one element are in the same conjugacy class. Let $\mathcal{K}_k$ be the number of conjugacy classes of subgroups of $\mathrm{U}((p-1)/2)$ with $|N/C| = k$, where N denotes the normalizer and C the centralizer of this subgroup. Then the number $\mathcal{K}(p)$ of conjugacy classes of subgroups of order p in $\mathrm{U}((p-1)/2)$ is

$$\mathcal{K}(p) = \sum_{\substack{k|p-1,\\ k \text{ odd}}} \mathcal{K}_k.$$

If $|N/C| = k$, then

$$N/C \cong \mathbb{Z}/k\mathbb{Z} \subseteq \mathbb{Z}/(p-1)\mathbb{Z} \cong \mathrm{Aut}(\mathbb{Z}/2p\mathbb{Z})$$

where $k|p-1$ and k is odd. This means that N/C is isomorphic to a subgroup of $\mathrm{Aut}(\mathbb{Z}/p\mathbb{Z})$. So we get the short exact sequence

$$1 \longrightarrow \mathbb{Z}/2p\mathbb{Z} \longrightarrow N \longrightarrow \mathbb{Z}/k\mathbb{Z} \longrightarrow 1.$$

Moreover, we have an injection $\mathbb{Z}/p\mathbb{Z} \hookrightarrow \mathbb{Z}/2p\mathbb{Z} \hookrightarrow N$. Applying the proposition to this case yields

$$\widehat{\mathrm{H}}^*(N, \mathbb{Z})_{(p)} \cong \left(\widehat{\mathrm{H}}^*(\mathbb{Z}/2p\mathbb{Z}, \mathbb{Z})_{(p)}\right)^{\mathbb{Z}/k\mathbb{Z}}.$$

The action of $\mathbb{Z}/k\mathbb{Z}$ on $\mathbb{Z}/2p\mathbb{Z}$ is given by the action of $\mathbb{Z}/k\mathbb{Z}$ as a subgroup of the group of automorphisms of $\mathbb{Z}/p\mathbb{Z} \subset \mathbb{Z}/2p\mathbb{Z}$.

Lemma 3.4. *The Farrell cohomology of* $\mathbb{Z}/l\mathbb{Z}$ *is*

$$\widehat{\mathrm{H}}^*(\mathbb{Z}/l\mathbb{Z}, \mathbb{Z}) = \mathbb{Z}/l\mathbb{Z}\,[x, x^{-1}]$$

where $\deg x = 2$, $x \in \widehat{\mathrm{H}}^2(\mathbb{Z}/l\mathbb{Z}, \mathbb{Z})$, *and* $\langle x \rangle \cong \mathbb{Z}/l\mathbb{Z}$.

Proof. See Brown's book [3]. For finite groups the Farrell cohomology and the Tate cohomology coincide. □

Proposition 3.5. *Let p be an odd prime and let $k \in \mathbb{Z}$ divide $p-1$. Then*

$$\left(\widehat{\mathrm{H}}^*(\mathbb{Z}/2p\mathbb{Z}, \mathbb{Z})_{(p)}\right)^{\mathbb{Z}/k\mathbb{Z}} \cong \mathbb{Z}/p\mathbb{Z}[x^k, x^{-k}]$$

where $x \in \widehat{\mathrm{H}}^2(\mathbb{Z}/2p\mathbb{Z}, \mathbb{Z})$.

Proof. For an odd prime p

$$\widehat{\mathrm{H}}^*(\mathbb{Z}/2p\mathbb{Z}, \mathbb{Z})_{(p)} = \left(\mathbb{Z}/2p\mathbb{Z}[x, x^{-1}]\right)_{(p)} = \mathbb{Z}/p\mathbb{Z}[x, x^{-1}].$$

We have to consider the action of $\mathbb{Z}/k\mathbb{Z}$ on $\mathbb{Z}/p\mathbb{Z}[x, x^{-1}]$. We have $px = 0$ and $x \in \widehat{\mathrm{H}}^2(\mathbb{Z}/2p\mathbb{Z}, \mathbb{Z})$. The action is given by $x \mapsto qx$ with q such that $(q, p) = 1$,

$q^k \equiv 1 \pmod p$ and k is the smallest number such that this is fulfilled. The action of $\mathbb{Z}/k\mathbb{Z}$ on

$$\widehat{\mathrm{H}}^{2m}(\mathbb{Z}/2p\mathbb{Z},\mathbb{Z})_{(p)} \cong (\langle x^m \rangle) \cong \mathbb{Z}/p\mathbb{Z}$$

is given by

$$x^m \mapsto q^m x^m.$$

The $\mathbb{Z}/k\mathbb{Z}$-invariants of $\widehat{\mathrm{H}}^*(\mathbb{Z}/2p\mathbb{Z},\mathbb{Z})_{(p)}$ are the $x^m \in \widehat{\mathrm{H}}^{2m}(\mathbb{Z}/2p\mathbb{Z},\mathbb{Z})_{(p)}$ with $x^m \mapsto x^m$, or equivalently $q^m \equiv 1 \pmod p$. Herewith we get

$$\begin{aligned}\widehat{\mathrm{H}}^*(N,\mathbb{Z})_{(p)} &\cong \left(\widehat{\mathrm{H}}^*(\mathbb{Z}/2p\mathbb{Z},\mathbb{Z})_{(p)}\right)^{\mathbb{Z}/k\mathbb{Z}} \cong \left(\mathbb{Z}/p\mathbb{Z}[x,x^{-1}]\right)^{\mathbb{Z}/k\mathbb{Z}}\\ &\cong \mathbb{Z}/p\mathbb{Z}[x^k,x^{-k}].\end{aligned}$$

□

Proposition 3.6. *Let p be an odd prime for which $h^- = 1$. Then*

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})_{(p)} \cong \prod_{\substack{k|p-1\\ k \ odd}} \left(\prod_1^{\mathcal{K}_k} \mathbb{Z}/p\mathbb{Z}[x^k,x^{-k}]\right),$$

where $\mathcal{K}_k$ is the number of conjugacy classes of subgroups of $\mathrm{U}((p-1)/2)$ *with $|N/C| = k$. As usual N denotes the normalizer and C the centralizer of this subgroup.*

Proof. Let p be a prime with $h^- = 1$. Then a bijection exists between the conjugacy classes of matrices of order p in $\mathrm{U}((p-1)/2)$ that satisfy the conditions of Theorem 2.2 and the conjugacy classes of matrices of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$. Now this proposition follows from Theorem 3.1. □

Now it remains to determine $\mathcal{K}_k$, the number of conjugacy classes of subgroups of $\mathrm{U}((p-1)/2)$ of order p with $N/C \cong \mathbb{Z}/k\mathbb{Z}$. Therefore we need $\mathcal{N}_k$, the number of conjugacy classes of elements $X \in \mathrm{U}((p-1)/2)$ of order p for which $1 = j_1 < \cdots < j_k < p$ exist such that the X^{j_l}, $l = 1,\ldots,k$, are in the same conjugacy class than X and k is maximal. One such class yields k elements in a group for which $|N/C| = k$ and therefore

$$\mathcal{K}_k = k\mathcal{N}_k \frac{1}{p-1}.$$

We recall the formula for $\mathcal{N}_k$:

$$\mathcal{N}_k = 2^{\frac{p-1}{2k}} - \sum_{\substack{k' \text{ odd},\ k'>k\\ k|k',\ k'|p-1}} \mathcal{N}_{k'}.$$

Now we have everything we need to compute the p-primary part of the Farrell cohomology of $\mathrm{Sp}(p-1,\mathbb{Z})$ for some examples of primes with $h^- = 1$.

3.3 Examples with $3 \leqslant p \leqslant 19$

$p=3$: It is $\mathrm{Sp}(2,\mathbb{Z}) = \mathrm{SL}(2,\mathbb{Z})$. One conjugacy class exists with $N = C$. Therefore

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(2,\mathbb{Z}),\mathbb{Z})_{(3)} \cong \mathbb{Z}/3\mathbb{Z}[x,x^{-1}],$$

and $\mathrm{Sp}(2,\mathbb{Z})$ has 3-period 2.

$p=5$: One conjugacy class exists with $N=C$. Therefore

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(4,\mathbb{Z}),\mathbb{Z})_{(5)} \cong \mathbb{Z}/5\mathbb{Z}[x,x^{-1}],$$

and $\mathrm{Sp}(4,\mathbb{Z})$ has 5-period 2.

$p=7$: One conjugacy class exists with $N/C \cong \mathbb{Z}/3\mathbb{Z}$, and one class exists with $N=C$. Therefore

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(6,\mathbb{Z}),\mathbb{Z})_{(7)} \cong \mathbb{Z}/7\mathbb{Z}[x^3,x^{-3}] \times \mathbb{Z}/7\mathbb{Z}[x,x^{-1}],$$

and $\mathrm{Sp}(6,\mathbb{Z})$ has 7-period 6.

$p=11$: One conjugacy class exists with $N/C \cong \mathbb{Z}/5\mathbb{Z}$ and 3 classes exist with $N=C$. Therefore

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(10,\mathbb{Z}),\mathbb{Z})_{(11)} \cong \mathbb{Z}/11\mathbb{Z}[x^5,x^{-5}] \times \prod_1^3 \mathbb{Z}/11\mathbb{Z}[x,x^{-1}],$$

and $\mathrm{Sp}(10,\mathbb{Z})$ has 11-period 10.

$p=13$: One conjugacy class exists with $N/C \cong \mathbb{Z}/3\mathbb{Z}$ and 5 classes exist with $N=C$. Therefore

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(12,\mathbb{Z}),\mathbb{Z})_{(13)} \cong \mathbb{Z}/13\mathbb{Z}[x^3,x^{-3}] \times \prod_1^5 \mathbb{Z}/13\mathbb{Z}[x,x^{-1}],$$

and $\mathrm{Sp}(12,\mathbb{Z})$ has 13-period 6.

$p=17$: 16 conjugacy classes exist with $N=C$. Therefore

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(16,\mathbb{Z}),\mathbb{Z})_{(17)} \cong \prod_1^{16} \mathbb{Z}/17\mathbb{Z}[x,x^{-1}],$$

and $\mathrm{Sp}(16,\mathbb{Z})$ has 17-period 2.

$p=19$: One conjugacy class exists with $N/C \cong \mathbb{Z}/9\mathbb{Z}$, one class exists with $N/C \cong \mathbb{Z}/3\mathbb{Z}$, and 28 classes exist with $N=C$.

$$\begin{aligned}\widehat{\mathrm{H}}^*(\mathrm{Sp}(18,\mathbb{Z}),\mathbb{Z})_{(19)} \cong{}& \mathbb{Z}/19\mathbb{Z}[x^9,x^{-9}] \times \mathbb{Z}/19\mathbb{Z}[x^3,x^{-3}]\\ &\times \prod_1^{28} \mathbb{Z}/19\mathbb{Z}[x,x^{-1}],\end{aligned}$$

and $\mathrm{Sp}(18,\mathbb{Z})$ has 19-period 18.

3.4 The p-primary part of the Farrell cohomology of Sp($p-1,\mathbb{Z}$)

Let p be an odd prime and let ξ be a primitive p-th root of unity. Let h^- be the relative class number of the cyclotomic field $\mathbb{Q}(\xi)$. In this section we compute $\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})_{(p)}$ and its period for any odd prime p for which h^- is odd.

Theorem 3.7. *Let p be an odd prime for which h^- is odd. Then*

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})_{(p)} \cong \prod_{\substack{k|p-1\\ k \text{ odd}}} \left(\prod_{1}^{\widetilde{\mathcal{K}}_k} \mathbb{Z}/p\mathbb{Z}[x^k, x^{-k}] \right),$$

where $\widetilde{\mathcal{K}}_k$ denotes the number of conjugacy classes of subgroups of order p of $\mathrm{Sp}(p-1,\mathbb{Z})$ *for which $|N/C| = k$. Moreover $\widetilde{\mathcal{K}}_k \geqslant \mathcal{K}_k$, where $\mathcal{K}_k$ is the number of conjugacy classes of subgroups of* $\mathrm{U}((p-1)/2)$ *with $|N/C| = k$. As usual N denotes the normalizer and C the centralizer of the corresponding subgroup.*

Proof. We have seen in Section 2.2 that if h^- is odd, a bijection exists between the conjugacy classes of matrices of order p in $\mathrm{U}((p-1)/2)$ that satisfy the conditions of Theorem 2.2 and the conjugacy classes of matrices of order p in $\mathrm{Sp}(p-1,\mathbb{Z})$ that correspond to the equivalence classes $[\mathbb{Z}[\xi], u] \in \mathcal{P}$. Each conjugacy class of subgroups of order p in $\mathrm{U}((p-1)/2)$ whose group elements satisfy the condition required in Theorem 2.2 yields at least one conjugacy class in $\mathrm{Sp}(p-1,\mathbb{Z})$. This implies that the p-primary part of the Farrell cohomology of $\mathrm{Sp}(p-1,\mathbb{Z})$ is a product

$$\prod_{\substack{k|p-1,\\ k \text{ odd}}} \left(\prod_{1}^{\widetilde{\mathcal{K}}_k} \mathbb{Z}/p\mathbb{Z}[x^k, x^{-k}] \right)$$

where $\widetilde{\mathcal{K}}_k$ denotes the number of conjugacy classes of subgroups of order p of $\mathrm{Sp}(p-1,\mathbb{Z})$ that satisfy $|N/C| = k$. Let $\mathcal{K}_k$ be the number of such subgroups in $\mathrm{U}((p-1)/2)$. Because h^- is odd, each such subgroup gives at least one such subgroup of $\mathrm{Sp}(p-1,\mathbb{Z})$. Therefore, if h^- is odd, $\widetilde{\mathcal{K}}_k \geqslant \mathcal{K}_k$. If h^- is even, it may be possible that no subgroup of $\mathrm{Sp}(p-1,\mathbb{Z})$ of order p exists for which $|N/C| = k$. □

Theorem 3.8. *Let p be an odd prime for which h^- is odd and let y be such that $p-1 = 2^r y$ and y is odd. Then the period of $\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})_{(p)}$ is $2y$.*

Proof. By Theorem 3.7 we know that the p-primary part of the Farrell cohomology of $\mathrm{Sp}(p-1,\mathbb{Z})$ is

$$\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})_{(p)} \cong \prod_{\substack{k|p-1\\ k \text{ odd}}} \left(\prod_{1}^{\widetilde{\mathcal{K}}_k} \mathbb{Z}/p\mathbb{Z}[x^k, x^{-k}] \right).$$

Moreover, $\widetilde{\mathcal{K}}_k \geqslant 1$ and the period of $\mathbb{Z}/p\mathbb{Z}[x^k, x^{-k}]$ is $2k$. Herewith the period of the p-primary part of the Farrell cohomology is $2y$. $\square$

If p is a prime for which h^- is even, the p-period of $\widehat{\mathrm{H}}^*(\mathrm{Sp}(p-1,\mathbb{Z}),\mathbb{Z})$ is $2z$ where z is odd and divides $p-1$. The period is not necessarily $2y$ because there may be no subgroup of order p in which y elements are conjugate in $\mathrm{Sp}(p-1,\mathbb{Z})$ even if we know that they are conjugate in $\mathrm{Sp}(p-1,\mathbb{R})$.

References

[1] A. Ash, *Farrell cohomology of $GL(n,\mathbb{Z})$*, Israel J. of Math. 67 (1989), 327-336.

[2] K. S. Brown, *Euler characteristics of discrete groups and G-spaces*, Invent. Math. 27 (1974), 229-264.

[3] K. S. Brown, *Cohomology of Groups*, GTM, vol. 87, Springer, 1982.

[4] B. Bürgisser, *Elements of finite order in symplectic groups*, Arch. Math. 39 (1982), 501-509.

[5] C. Busch, *Symplectic characteristic classes and the Farrell cohomology of* $\mathrm{Sp}(p-1,\mathbb{Z})$, Diss. ETH No. 13506, ETH Zürich (2000).

[6] C. Busch, *Symplectic characteristic classes*, L'Ens. Math. t. 47 (2001), 115-130.

[7] H. H. Glover, G. Mislin, *On the p-primary cohomology of $Out(F_n)$ in the p-rank one case*, J. Pure Appl. Algebra 153 (2000), 45-63.

[8] H. H. Glover, G. Mislin, Y. Xia, *On the Farrell cohomology of mapping class groups*, Invent. Math. 109 (1992), 535-545.

[9] D. Sjerve and Q. Yang, *Conjugacy classes of p-torsion in* $\mathrm{Sp}_{p-1}(\mathbb{Z})$, J. Algebra 195, No. 2 (1997), 580-603.

Cornelia Busch
Katholische Universität Eichstätt
- Ingolstadt
Mathematisch-Geographische Fakultät
Ostenstr. 26–28
D-85072 Eichstätt
Cornelia.Busch@ku-eichstaett.de

On the Classification of Simple Inductive Limit C^*-Algebras, I: The Reduction Theorem

Dedicated to Professor Ronald G. Douglas on the occasion of his sixtieth birthday

Guihua Gong[1]

Received: October 4, 2002

Communicated by Joachim Cuntz

Abstract. Suppose that

$$A = \lim_{n\to\infty} (A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$$

is a simple C^*-algebra, where $X_{n,i}$ are compact metrizable spaces of uniformly bounded dimensions (this restriction can be relaxed to a condition of very slow dimension growth). It is proved in this article that A can be written as an inductive limit of direct sums of matrix algebras over certain special 3-dimensional spaces. As a consequence it is shown that this class of inductive limit C^*-algebras is classified by the Elliott invariant — consisting of the ordered K-group and the tracial state space — in a subsequent paper joint with G. Elliott and L. Li (Part II of this series). (Note that the C^*-algebras in this class do not enjoy the real rank zero property.)

[1]This material is based upon work supported by, or in part by, the U.S. Army Research Office under grant number DAAD19-00-1-0152. The research is also partially supported by NSF grant DMS 9401515, 9622250, 9970840 and 0200739.

Contents

0 Introduction

In this article and the subsequent article [EGL], we will classify all the unital simple C^*-algebras A, which can be written as the inductive limit of a sequence

$$\bigoplus_{i=1}^{t_1} P_{1,i}M_{[1,i]}(C(X_{1,i}))P_{1,i} \xrightarrow{\phi_{1,2}} \bigoplus_{i=1}^{t_2} P_{2,i}M_{[2,i]}(C(X_{2,i}))P_{2,i} \xrightarrow{\phi_{2,3}} \cdots,$$

where $X_{n,i}$ are compact metrizable spaces with $\sup\{\dim X_{n,i}\}_{n,i} < +\infty$, $[n,i]$ and t_n are positive integers, and $P_{n,i} \in M_{[n,i]}(C(X_{n,i}))$ are projections. The invariant consists of the ordered K-group and the space of traces on the algebra. The main result in the present article is that a C^*-algebra A as above can be written in another way as an inductive limit so that all the spaces $X_{n,i}$ appearing are certain special simplicial complexes of dimension at most three. Then, in [EGL], the classification theorem will be proved by assuming the C^*-algebras are such special inductive limits.

In the special case that the groups $K_*(C(X_{n,i}))$ are torsion free, the C^*-algebra A can be written as an inductive limit of direct sums of matrix algebras over $C(S^1)$(i.e., one can replace $X_{n,i}$ by S^1). Combining this result with [Ell2], without [EGL]–the part II of this series—, we can still obtain the classification theorem for this special case, which is a generalization of the result of Li for the case that $\dim(X_{n,i}) = 1$ (see [Li1-3]).

The theory of C^*-algebras can be regarded as noncommutative topology, and has broad applications in different areas of mathematics and physics (e.g., the study of foliated spaces, manifolds with group actions; see [Con]).

One extreme class of C^*-algebras is the class of commutative C^*-algebras, which corresponds to the category of ordinary locally compact Hausdorff topological spaces. The other extreme, which is of great importance, is the class of simple C^*-algebras, which must be considered to be highly noncommutative. For example, the (reduced) foliation C^*-algebra of a foliated space is simple if and only if every leaf is dense in the total space; the cross product C^*-algebra, for a $\mathbb{Z}$ action on a space X, is simple if and only if the action is minimal.

Even though the commutative C^*-algebras and the simple C^*-algebras are opposite extremes, remarkably, many (unital or nonunital) simple C^*-algebras (including the foliation C^*-algebra of a Kronecker foliation, see [EE]) have been proved to be inductive limits of direct sums of matrix algebras over commutative C^*-algebras, i.e., to be of the form $\lim_{n\to\infty}(A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$. (Note that the only commutative C^*-algebras, or matrix algebras over commutative C^*-algebras, which are simple are the very trivial ones, $\mathbb{C}$ or $M_k(\mathbb{C})$.) In general, it is a conjecture that any stably finite, simple, separable, amenable C^*-algebra is an inductive limit of subalgebras of matrix algebras over commutative C^*-algebras. This conjecture would be analogous to the result of Connes and Haagerup that any amenable von Neumann algebra is generated by an upward directed family of sub von Neumann algebras of type I.

The sweeping classification project of G. Elliott is aimed at the complete classification of simple, separable, amenable C^*-algebras in terms of a certain simple invariant, as we mentioned above, consisting of the ordered K-group and the space of traces on the algebra. Naturally, the class of inductive limit C^*-algebras $A = \lim\limits_{\to}(A_n = \bigoplus_{i=1}^{t_n} P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}, \phi_{n,m})$, considered in this article, is an essential ingredient of the project. Following Blackadar [Bl1], we will call such inductive limit algebras AH algebras.

The study of AH algebras has its roots in the theory of AF algebras (see [Br] and [Ell4]). But the modern classification theory of AH algebras was inspired by the seminal paper [Bl3] of B. Blackadar and was initiated by Elliott in [Ell5]. The real rank of a C^*-algebra is the noncommutative counterpart of the dimension of a topological space. Until recently, the only known possibilities for the real rank of a simple C^*-algebra were zero or one. It was proved in [DNNP] that any simple AH algebra

$$A = \lim_{\to}(A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$$

has real rank either zero or one, provided that $\sup\{\dim X_{n,i}\}_{n,i} < +\infty$. (Recently, Villadsen has found a simple C^*-algebra with real rank different from zero and one, see [Vi 2].)

For the case of simple C^*-algebras of real rank zero, the classification is quite successful and satisfactory, even though the problem is still not completely solved. Namely, on one hand, the remarkable result of Kirchberg [Kir] and Phillips [Phi1] completely classified all purely infinite, simple, separable, amenable C^*-algebras with the so called UCT property (see also [R] for an important earlier result). All purely infinite simple C^*-algebras are of real rank zero; see [Zh]. On the other hand, in [EG1-2] Elliott and the author completely classified all the stably finite, simple, real rank zero C^*-algebras which are AH algebras of the form $\lim\limits_{\to}(A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$ with $\dim(X_{n,i}) \leq 3$. It was proved by Dadarlat and the author that this class includes all simple real rank zero AH algebras with arbitrary but uniformly bounded dimensions for the spaces $X_{n,i}$ (see [D1-2], [G1-4] and [DG]).

In this article, the AH algebras considered are not assumed to have real rank zero. As pointed out above, they must have real rank either zero or one. In fact, in a strong sense, almost all of them have real rank one. The real rank zero C^*-algebras are the very special ones for which the space of traces (one part of the invariant mentioned above) is completely determined by the ordered K-group of the C^*-algebra (the other part of the invariant). Not surprisingly, the lack of the real rank zero property presents new essential difficulties. Presumably, dimension one noncommutative spaces are much richer and more complicated than dimension zero noncommutative spaces. In what follows, we would like to explain one of the main differences between the real rank zero case and the general case in the setting of simple AH algebras.

If $A = \lim\limits_{\rightarrow}(A_n = \bigoplus_{i=1}^{t_n} P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}, \phi_{n,m})$ is of real rank zero, then Elliott and the author proved a decomposition result (see Theorem 2.21 of [EG2]) which says that $\phi_{n,m}$ (for m large enough) can be approximately decomposed as a sum of two parts, $\phi_1 \oplus \phi_2$; one part, ϕ_1, having a very small support projection, and the other part, ϕ_2, factoring through a finite dimensional algebra.

In §4 of the present paper, we will prove a decomposition theorem which says that, for the simple AH algebra A above (with or without the real rank zero condition), $\phi_{n,m}$ (for m large enough) can be approximately decomposed as a sum of three parts, $\phi_1 \oplus \phi_2 \oplus \phi_3$: the part ϕ_1 having a very small support projection compared with the part ϕ_2; the part ϕ_2 factoring through a finite dimensional algebra; and the third part ϕ_3 factoring through a direct sum of matrix algebras over the interval $[0,1]$. (Note that, in the case of a real rank zero inductive limit, the part ϕ_3 does not appear. In the general case, though, the part ϕ_3 has a very large support projection compared with the part $\phi_1 \oplus \phi_2$.) With this decomposition theorem, we can often deal with the part $\phi_1 \oplus \phi_2$ by using the techniques developed in the classification of the real rank zero case (see [EG1-2], [G1-4], [D1-2] and [DG]).

This new decomposition theorem is much deeper. It reflects the real rank one (as opposed to real rank zero) property of the simple C^*-algebra. The special case of the decomposition result that the spaces $X_{n,i}$ are already supposed to be one-dimensional spaces is due to L. Li (see [Li3]). The proof for the case of higher dimensional spaces is essentially more difficult. In particular, as preparation, we need to prove certain combinatorial results (see §3) and also the following result (see §2): Any homomorphism from $C(X)$ to $M_k(C(Y))$ can be perturbed to a homomorphism whose maximum spectral multiplicity (for the definition of this terminology, see 1.2.4 below) is not larger than $\dim X + \dim Y$, provided that $X \neq \{pt\}$ and X is path connected.

The special simplicial complexes used in our main reduction theorem are the following spaces: $\{pt\}$, $[0,1]$, S^1, S^2, $\{T_{II,k}\}_{k=2}^{\infty}$, and $\{T_{III,k}\}_{k=2}^{\infty}$, where the spaces $T_{II,k}$ are two-dimensional connected simplicial complexes with $H^1(T_{II,k}) = 0$ and $H^2(T_{II,k}) = \mathbb{Z}/k$, and the spaces $T_{III,k}$ are three-dimensional connected simplicial complexes with $H^1(T_{III,k}) = 0 = H^2(T_{III,k})$ and $H^3(T_{III,k}) = \mathbb{Z}/k$. (See 4.2 of [EG2] for details.)

The spaces $T_{II,k}$ and $T_{III,k}$ are needed to produce the torsion part of the K-groups of the inductive limit C^*-algebras. Since the algebras $C(T_{II,k}), C(T_{III,k})$, and $C(S^2)$ are not stably generated (see [Lo]), difficulties occur in the construction of homomorphisms from these C^*-algebras, when we prove our main reduction theorem (and the isomorphism theorem in [EGL]). In the case of real rank zero algebras, this difficulty can be avoided by using unsuspended E-theory (see [D1-2] and [G1-4]) combined with a certain uniqueness theorem — Theorem 2.29 of [EG2], which only involves homomorphisms (instead of general completely positive linear maps). Roughly speaking, the trouble is that a completely positive linear $*$-contraction, which is an "almost homomorphism"—a G-δ multiplicative map (see 1.1.2 below for the definition of this concept) for sufficiently large G and sufficiently small δ—, may not be automatically close to a homomorphism. As we mentioned above, after we approximately decompose $\phi_{n,m}$ as $\phi_1 \oplus \phi_2 \oplus \phi_3$, we will deal with the part $\phi_1 \oplus \phi_2$, by using the results and techniques from the real rank zero case, in particular by using Theorem 1.6.9 below—a strengthened version of Theorem 2.29 of [EG2]. Therefore, we will consider the composition of the map $\phi_1 \oplus \phi_2$ and a homomorphism from a matrix algebra over $\{pt\}$, $[0,1]$, S^1, S^2, $\{T_{II,k}\}_{k=2}^{\infty}$, and $\{T_{III,k}\}_{k=2}^{\infty}$, to A_n. We need this composition to be close to a homomorphism, but $\phi_1 \oplus \phi_2$ is not supposed to be close to a homomorphism (it is close to the homomorphism $\phi_{n,m}$ in the case of real rank zero). To overcome the above difficulty, we prove a theorem in §5—a kind of uniqueness theorem, which may be roughly described as follows:

For any $\varepsilon > 0$, positive integer N, and finite set $F \subset A = M_k(C(X))$, where X is one of the spaces $\{pt\}, [0,1], T_{II,k}, T_{III,k}$, and S^2, there are a number $\delta > 0$, a finite set $G \subset A$, and a positive integer L, such that for any two G-δ multiplicative (see 1.1.2 below), completely positive, linear $*$-contractions $\phi, \psi : M_k(C(X)) \to B = M_l(C(Y))$ (where $\dim(Y) \leq N$), if they define the same map on the level of K-theory and also mod-p K-theory (this statement will be made precise in §5), then there are a homomorphism $\lambda : A \to M_L(B)$ with finite dimensional image and a unitary $u \in M_{L+1}(B)$ such that

$$\|(\phi \oplus \lambda)(f) - u(\psi \oplus \lambda)(f)u^*\| < \varepsilon$$

for all $f \in F$.

This result is quite nontrivial, and may be expected to have more general applications. Some similar results appear in the literature (e.g., [EGLP, 3.1.4], [D1, Thm A], [G4, 3.9]). But even for $*$-homomorphisms (which are G-δ multiplicative for any G and δ), all these results (except for contractible spaces) require that the number L, the size of the matrix, depends on the maps ϕ and ψ.

Note that the theorem stated above does not hold if one replaces X by S^1, even if both ϕ and ψ are $*$-homomorphisms. (Fortunately, we do not need the theorem for S^1 in this article, since $C(S^1)$ is stably generated. But on the other hand, the lack of such a theorem for S^1 causes a major difficulty in the formulation and the proof of the uniqueness theorem involving homomorphisms from $C(S^1)$ to $M_k(C(X))$, in [EGL]— part 2 of this series.)

With the above theorem, if a G-δ multiplicative, positive, linear $*$-contraction ϕ and a $*$-homomorphism (name it ψ) define the same map on the level of K-theory and mod-p K-theory, then $\phi\oplus\lambda$ is close to a $*$-homomorphism (e.g., $\mathrm{Ad}u\circ(\psi\oplus\lambda)$) for some $*$-homomorphism $\lambda : A \to M_L(B)$ with finite dimensional image. In particular, the size L of the $*$-homomorphism λ can be controlled. This is essential for the construction of $*$-homomorphisms from $A = M_k(C(X))$, where X is one of $T_{II,k}, T_{III,k}$, and S^2. In particular, once L is fixed, we can construct the decomposition of $\phi_{n,m}$ as $\phi_1 \oplus \phi_2 \oplus \phi_3$, as mentioned above, such that the supporting projection of the part ϕ_2 is larger than the supporting projection of the part ϕ_1 by the amplification of L times. Hence we can prove that, the composition of the map $\phi_1 \oplus \phi_2$ and a homomorphism from a matrix algebra over $T_{II,k}, T_{III,k}$, and S^2 to A_n, is close to a homomorphism (see Theorems 5.32a and 5.32b below for details).

The theorem is also true for a general finite CW complex X, provided that $K_1(C(X))$ is a torsion group.

(Note that for the space S^1 (or the spaces $\{pt\}, [0,1]$), we do not need such a theorem, since any G-δ multiplicative, positive, linear $*$-contraction from $M_k(C(S^1))$ will automatically be close to a $*$-homomorphism if G is sufficiently large and δ is sufficiently small.)

The above mentioned theorem and the decomposition theorem both play important roles in the proof of our main reduction theorem, and also in the proof of the isomorphism theorem in [EGL].

The main results of this article and [EGL] were announced in [G1] and in Elliott's lecture at the International Congress of Mathematicians in Zurich (see [Ell3]). Since then, several classes of simple inductive limit C^*-algebras have been classified (see [EGJS], [JS 1-2], and [Th1]). But all these later results involve only inductive limits of subhomogeneous algebras with 1-dimensional spectra. In particular, the K_0-groups have to be torsion free, since it is impossible to produce the torsion in K_0-group with one-dimensional spectra alone, even with subhomogeneous building blocks.

This article is organized as follows. In §1, we will introduce some notations, collect some known results, prove some preliminary results, and discuss some important preliminary ideas, which will be used in other sections. In particular, in §1.5, we will discuss the general strategy in the proof of the decomposition theorem, of which, the detailed proof will be given in in §2, §3 and §4. In §1.6, we will prove some uniqueness theorem and factorization theorem which are important in the proof of the main theorem. Even though the results in §1.6 are new, most of the methods are modification of known techniques from [EG2], [D2], [G4] and [DG]. In §2, we will prove the result about maximum spectral multiplicities, which will be used in §4 and other papers. In §3, we will prove certain results of a combinatorial nature. In §4, we will combine the results from §2, §3, and the results in [Li2], to prove the decomposition theorem. In §5, we will prove the result mentioned above concerning G-δ multiplicative maps. In §6, we will use §4, §5 and §1.6 to prove our main reduction theorem. Our main result can be generalized from the case of no dimension growth (i.e., $X_{n,i}$

have uniformly bounded dimensions) to the case of very slow dimension growth. Since the proof of this general case is much more tedious and complicated, we will deal with this generalization in [G5], which can be regarded as an appendix to this article.

Acknowledgements. The author would like to thank Professors M. Dadarlat, G. Elliott, L. Li, and H. Lin for helpful conversations. The author also like to thank G. Elliott, L. Li and H. Lin for reading the article and making suggestions to improve the readability of the article. In particular, L. Li suggested to the author to make the pictures (e.g., in 3.6, 3.10 and 6.3) to explain the ideas in the proof of some results; G. Elliott suggested to the author to write a subsection §1.5 to explain the general strategy for proving a decomposition theorem.

1 Preparation and some preliminary ideas

We will introduce some conventions, general assumptions, and preliminary results in this section.

1.1 General assumptions on inductive limits

1.1.1. If A and B are two C^*-algebras, we use $\mathrm{Map}(A,B)$ to denote the space of all linear, completely positive $*$-contractions from A to B. If both A and B are unital, then $\mathrm{Map}(A,B)_1$ will denote the subset of $\mathrm{Map}(A,B)$ consisting of unital maps. By word "map", we shall mean linear, completely positive $*$-contraction between C^*-algebras, or else we shall mean continuous map between topological spaces, which one will be clear from the context.

By a homomorphism between C^*-algebras, will be meant a $*$-homomorphism. Let $\mathrm{Hom}(A,B)$ denote the space of all homomorphisms from A to B. Similarly, if both A and B are unital, let $\mathrm{Hom}(A,B)_1$ denote the subset of $\mathrm{Hom}(A,B)$ consisting of unital homomorphisms.

Definition 1.1.2. Let $G \subset A$ be a finite set and $\delta > 0$. We shall say that $\phi \in \mathrm{Map}(A,B)$ is G-δ multiplicative if

$$\|\phi(ab) - \phi(a)\phi(b)\| < \delta$$

for all $a, b \in G$.

Sometimes, we use $\mathrm{Map}_{G-\delta}(A,B)$ to denote all the G-δ multiplicative maps.

1.1.3. In the notation for an inductive system $(A_n, \phi_{n,m})$, we understand that $\phi_{n,m} = \phi_{m-1,m} \circ \phi_{m-2,m-1} \cdots \circ \phi_{n,n+1}$, where all $\phi_{n,m} : A_n \to A_m$ are homomorphisms.

We shall assume that, for any summand A_n^i in the direct sum $A_n = \bigoplus_{i=1}^{t_n} A_n^i$, necessarily, $\phi_{n,n+1}(\mathbf{1}_{A_n^i}) \neq 0$, since, otherwise, we could simply delete A_n^i from A_n without changing the limit algebra.

1.1.4. If $A_n = \bigoplus_i A_n^i$ and $A_m = \bigoplus_j A_m^j$, we use $\phi_{n,m}^{i,j}$ to denote the partial map of $\phi_{n,m}$ from the i-th block A_n^i of A_n to the j-th block A_m^j of A_m.
In this article, we will assume that all inductive limit C^*-algebras are SIMPLE. That is, the limit algebra has no nontrivial proper closed two sided ideals. We will also assume that every inductive limit C^*-algebra $A = \varinjlim(A_n, \phi_{n,m})$ coming into consideration is different both from $M_k(\mathbb{C})$ (the matrix algebra over $\mathbb{C}$), and from $\mathcal{K}(H)$ (the algebra of all compact operators).
Since $A = \varinjlim(A_n = \bigoplus_i A_n^i, \phi_{n,m})$ is simple, by 5.3.2(b) of [DN], we may assume that $\phi_{n,m}^{i,j}(\mathbf{1}_{A_n^i}) \neq 0$ for any blocks A_n^i and A_m^j, where $n < m$.

1.1.5. To avoid certain counter examples (see [V]) of the main result of this article, we will restrict our attention, in this article, to inductive systems satisfying the following VERY SLOW DIMENSION GROWTH CONDITION. This is a strengthened form of the condition of slow dimension growth introduced in [BDR].
If $\varinjlim(A_n = \bigoplus_{i=1}^{t_n} P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}, \phi_{n,m})$ is a unital inductive limit system, the very slow dimension growth condition is

$$\lim_{n\to+\infty} \max_i \left\{ \frac{(\dim X_{n,i})^3}{\operatorname{rank}(P_{n,i})} \right\} = 0,$$

where $\dim(X_{n,i})$ denotes the (covering) dimension of $X_{n,i}$.
In this article, we will also study non-unital inductive limit algebras. The above formula must then be slightly modified. The very slow dimension growth condition in the non-unital case is that, for any summand
$A_n^i = P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}$ of a fixed A_n,

$$\lim_{m\to+\infty} \max_{i,j} \left\{ \frac{(\dim X_{m,j})^3}{\operatorname{rank}\phi_{n,m}^{i,j}(\mathbf{1}_{A_n^i})} \right\} = 0,$$

where $\phi_{n,m}^{i,j}$ is the partial map of $\phi_{n,m}$ from A_n^i to A_m^j.
(For a unital inductive limit, the two conditions above are equivalent. Of course, both conditions are only proposed for the simple case.)
If the set $\{\dim X_{n,i}\}$ is bounded, i.e, there is an M such that

$$\dim X_{n,i} \leq M$$

for all n and i, then the inductive system automatically satisfies the very slow dimension growth condition, as we already assume that the limit algebra is not $M_k(\mathbb{C})$ or $\mathcal{K}(H)$.

We will prove our main reduction theorem for the case of uniformly bounded dimensions in this article, since it is significantly simpler than the case of very slow dimension growth. The general case will be discussed in [G5]—an appendix of this article. But the decomposition theorem will be proved for the case of very slow dimension growth.
It must be noted that, without the above assumption on dimension growth, the main theorem of this article does not hold (see [Vi1]). We shall leave the following question open: can the above condition of very slow dimension growth be replaced by the similar (but weaker) condition of slow dimension growth (see [BDR]), in the main theorem of this article?

1.1.6. By 2.3 of [Bl1], in the inductive limit

$$A = \lim_{\rightarrow}(A_n = \bigoplus_{i=1}^{t_n} P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}, \phi_{n,m}),$$

one can always replace the compact metrizable spaces $X_{n,i}$ by finite simplicial complexes. Note that the replacement does not increase the dimensions of the spaces. Therefore, in this article, WE WILL ALWAYS ASSUME THAT ALL THE SPACES $X_{n,i}$ IN A GIVEN INDUCTIVE SYSTEM ARE FINITE SIMPLICIAL COMPLEXES. Also, WE WILL FURTHER ASSUME THAT ALL $X_{n,i}$ ARE PATH CONNECTED. Otherwise, we will separate different components into different direct summands. (Note that a finite simplicial complex has at most finitely many path connected components.)
By SIMPLICIAL COMPLEX we mean finite simplicial complex or polyhedron; see [St].

1.1.7.
(a) We use the notation $\#(\cdot)$ to denote the cardinal number of the set, if the argument is a finite set. Very often, the sets under consideration will be sets with multiplicity, and then we shall also count multiplicity when we use the notation $\#$.
(b) We shall use $a^{\sim k}$ to denote $\underbrace{a, \cdots, a}_{k\ copies}$. For example,

$$\{a^{\sim 3}, b^{\sim 2}\} = \{a, a, a, b, b\}.$$

(c) $\mathrm{int}(\cdot)$ is used to denote the integer part of a real number. We reserve the notation $[\cdot]$ for equivalence classes in possibly different contexts.
(d) For any metric space X, any $x_0 \in X$ and any $c > 0$, let $B_c(x_0) := \{x \in X \mid d(x, x_0) < c\}$ denote the open ball with radius c and centre x_0.
(e) Suppose that A is a C^*-algebra, $B \subset A$ is a subalgebra, $F \subset A$ is a (finite) subset and let $\varepsilon > 0$. If for each element $f \in F$, there is an element $g \in B$ such that $\|f - g\| < \varepsilon$, then we shall say that F is approximately contained in B to within ε, and denote this by $F \subset_\varepsilon B$.

(f) Let X be a compact metric space. For any $\delta > 0$, a finite set $\{x_1, x_2, \cdots x_n\}$ is said to be δ-dense in X, if for any $x \in X$, there is x_i such that $\mathrm{dist}(x, x_i) < \delta$.
(g) We shall use $\bullet$ to denote any possible positive integer. To save notation, $y, y', y'', \cdots$ or $a_1, a_2, \cdots$ may be used for finite sequences if we do not care how many terms are in the sequence. Similarly, $A_1 \cup A_2 \cup \cdots$ or $A_1 \cap A_2 \cap \cdots$ may be used for finite union or finite intersection. If there is a danger of confusion with infinite sequence, union, or intersection, we will write them as $a_1, a_2, \cdots, a_\bullet$, $A_1 \cup A_2 \cup \cdots \cup A_\bullet$, or $A_1 \cap A_2 \cap \cdots A_\bullet$.
(h) For $A = \bigoplus_{i=1}^t M_{k_i}(C(X_i))$, where X_i are path connected simplicial complexes, we use the notation $r(A)$ to denote $\bigoplus_{i=1}^t M_{k_i}(\mathbb{C})$, which could be considered to be a subalgebra of A consisting of t-tuples of constant functions from X_i to $M_{k_i}(\mathbb{C})$ $(i = 1, 2, \cdots, t)$. Fix a base point $x_i^0 \in X_i$ for each X_i, one can define a map $r : A \to r(A)$ by

$$r(f_1, f_2, \cdots, f_t) = (f_1(x_1^0), f_2(x_2^0), \cdots f_t(x_t^0)) \in r(A).$$

(i) For any two projections $p, q \in A$, we use the notation $[p] \leq [q]$ to denote that p is unitarily equivalent to a sub projection of q. And we use $p \sim q$ to denote that p is unitarily equivalent to q.

1.2 Spectrum and spectral variation of a homomorphism

1.2.1. Let Y be a compact metrizable space. Let $P \in M_{k_1}(C(Y))$ be a projection with $\mathrm{rank}(P) = k \leq k_1$. For each y, there is a unitary $u_y \in M_{k_1}(\mathbb{C})$ (depending on y) such that

$$P(y) = u_y \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} u_y^*,$$

where there are k 1's on the diagonal. If the unitary u_y can be chosen to be continuous in y, then P is called a trivial projection.
It is well known that any projection $P \in M_{k_1}(C(Y))$ is locally trivial. That is, for any $y_0 \in Y$, there is an open set $U_{y_0} \ni y_0$, and there is a continuous unitary-valued function

$$u : U_{y_0} \to M_{k_1}(\mathbb{C})$$

such that the above equation holds for $u(y)$ (in place of u_y) for any $y \in U_{y_0}$.
If P is trivial, then $PM_{k_1}(C(X))P \cong M_k(C(X))$.

1.2.2. Let X be a compact metrizable space and $\psi : C(X) \to PM_{k_1}(C(Y))P$

be a unital homomorphism. For any given point $y \in Y$, there are points $x_1(y), x_2(y), \cdots, x_k(y) \in X$, and a unitary $U_y \in M_{k_1}(\mathbb{C})$ such that

$$\psi(f)(y) = P(y)U_y \begin{pmatrix} f(x_1(y)) & & & & & \\ & \ddots & & & & \\ & & f(x_k(y)) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} U_y^* P(y) \in P(y)M_{k_1}(\mathbb{C})P(y)$$

for all $f \in C(X)$. Equivalently, there are k rank one orthogonal projections $p_1, p_2, \cdots, p_k$ with $\sum_{i=1}^k p_i = P(y)$ and $x_1(y), x_2(y), \cdots, x_k(y) \in X$, such that

$$\psi(f)(y) = \sum_{i=1}^k f(x_i(y))p_i, \qquad \forall\ f \in C(X).$$

Let us denote the set $\{x_1(y), x_2(y), \cdots, x_k(y)\}$, counting multiplicities, by $\mathrm{SP}\psi_y$. In other words, if a point is repeated in the diagonal of the above matrix, it is included with the same multiplicity in $\mathrm{SP}\psi_y$. We shall call $\mathrm{SP}\psi_y$ the spectrum of ψ at the point y. Let us define the spectrum of ψ, denoted by $\mathrm{SP}\psi$, to be the closed subset

$$\mathrm{SP}\psi := \overline{\bigcup_{y \in Y} \mathrm{SP}\psi_y} \subseteq X.$$

Alternatively, $\mathrm{SP}\psi$ is the complement of the spectrum of the kernel of ψ, considered as a closed ideal of $C(X)$. The map ψ can be factored as

$$C(X) \xrightarrow{i^*} C(\mathrm{SP}\psi) \xrightarrow{\psi_1} PM_{k_1}(C(Y))P$$

with ψ_1 an injective homomorphism, where i denotes the inclusion $\mathrm{SP}\psi \hookrightarrow X$. Also, if $A = PM_{k_1}(C(Y))P$, then we shall call the space Y the spectrum of the algebra A, and write $\mathrm{SP}A = Y (= \mathrm{SP}(\mathrm{id}))$.

1.2.3. In 1.2.2, if we group together all the repeated points in $\{x_1(y), x_2(y), \cdots, x_k(y)\}$, and sum their corresponding projections, we can write

$$\psi(f)(y) = \sum_{i=1}^l f(\lambda_i(y))P_i \qquad\qquad (l \le k),$$

where $\{\lambda_1(y), \lambda_2(y), \cdots, \lambda_l(y)\}$ is equal to $\{x_1(y), x_2(y), \cdots, x_k(y)\}$ as a set, but $\lambda_i(y) \neq \lambda_j(y)$ if $i \neq j$; and each P_i is the sum of the projections corresponding to $\lambda_i(y)$. If $\lambda_i(y)$ has multiplicity m (i.e., it appears m times in $\{x_1(y), x_2(y), \cdots, x_k(y)\}$), then $\mathrm{rank}(P_i) = m$.

DEFINITION 1.2.4. Let ψ, y, and P_i be as above. The MAXIMUM SPECTRAL MULTIPLICITY OF ψ AT THE POINT y is defined to be $\max_i(\mathrm{rank} P_i)$. The MAXIMUM SPECTRAL MULTIPLICITY OF ψ is defined to be the supremum of the maximum spectral multiplicities of ψ at the various points of Y.

The following result is the main theorem in §2, which says that we can make the homomorphism not to have too large spectral multiplicities, up to a small perturbation.

THEOREM 2.1. *Let X and Y be connected simplicial complexes and $X \neq \{pt\}$. Let $l = \dim(X) + \dim(Y)$. For any given finite set $G \subset C(X)$, any $\varepsilon > 0$, and any unital homomorphism $\phi : C(X) \to PM_\bullet(C(Y))P$, where $P \in M_\bullet(C(Y))$ is a projection, there is a unital homomorphism $\phi' : C(X) \to PM_\bullet(C(Y))P$ such that*
(1) $\|\phi(g) - \phi'(g)\| < \varepsilon$ for all $g \in G$;
(2) ϕ' has maximum spectral multiplicity at most l.

1.2.5. Set $P^k(X) = \underbrace{X \times X \times \cdots \times X}_{k} / \sim$, where the equivalence relation $\sim$ is defined by $(x_1, x_2, \cdots, x_k) \sim (x'_1, x'_2, \cdots, x'_k)$ if there is a permutation σ of $\{1, 2, \cdots, k\}$ such that $x_i = x'_{\sigma(i)}$, for each $1 \le i \le k$. A metric d on X can be extended to a metric on $P^k(X)$ by

$$d([x_1, x_2, \cdots, x_k], [x'_1, x'_2, \cdots, x'_k]) = \min_{\sigma} \max_{1 \le i \le k} d(x_i, x'_{\sigma(i)}),$$

where σ is taken from the set of all permutations, and $[x_1, \cdots, x_k]$ denotes the equivalence class in $P^k(X)$ of $(x_1, \cdots, x_k)$.

1.2.6. Let X be a metric space with metric d. Two k-tuples of (possibly repeating) points $\{x_1, x_2, \cdots, x_k\} \subset X$ and $\{x'_1, x'_2, \cdots, x'_k\} \subset X$ are said to BE PAIRED WITHIN η if there is a permutation σ such that

$$d(x_i, x'_{\sigma(i)}) < \eta, \qquad i = 1, 2, \cdots, k.$$

This is equivalent to the following. If one regards $(x_1, x_2, \cdots, x_k)$ and $(x'_1, x'_2, \cdots, x'_k)$ as two points in P^kX, then

$$d([x_1, x_2, \cdots, x_k], [x'_1, x'_2, \cdots, x'_k]) < \eta.$$

1.2.7. Let $\psi : C(X) \to PM_{k_1}(C(Y))P$ be a unital homomorphism as in 1.2.5. Then

$$\psi^* : y \mapsto \mathrm{SP}\psi_y$$

defines a map $Y \to P^kX$, if one regards $\mathrm{SP}\psi_y$ as an element of P^kX. This map is continuous. In term of this map and the metric d, let us define the spectral variation of ψ:

$$\mathrm{SPV}(\psi) := \text{diameter of the image of } \psi^*.$$

Definition 1.2.8. We shall call the projection P_i in 1.2.3 the spectral projection of ϕ at y with respect to the spectral element $\lambda_i(y)$. If $X_1 \subset X$ is a subset of X, we shall call

$$\sum_{\lambda_i(y)\in X_1} P_i$$

the spectral projection of ϕ at y corresponding to the subset X_1 (or with respect to the subset X_1).
In general, for an open set $U \subset X$, the spectral projection $P(y)$ of ϕ at y corresponding to U does not depend on y continuously. But the following lemma holds.

Lemma 1.2.9. *Let $U \subset X$ be an open subset. Let $\phi : C(X) \to M_\bullet(C(Y))$ be a homomorphism. Suppose that $W \subset Y$ is an open subset such that*

$$SP\phi_y \cap (\overline{U}\backslash U) = \emptyset, \quad \forall y \in W.$$

Then the function

$$y \mapsto \textit{ spectral projection of } \ \phi \textit{ at } y \textit{ correponding to } \ U$$

is a continuous function on W. Furthermore, if W is connected then $\#(SP\phi_y \cap U)$ (counting multiplicity) is the same for any $y \in W$, and the map $y \mapsto SP\phi_y \cap U \in P^l X$ is a continuous map on W, where $l = \#(SP\phi_y \cap U)$.

Proof: Let $P(y)$ denote the spectral projection of ϕ at y corresponding to the open set U. Fix $y_0 \in W$. Since $\mathrm{SP}\phi_{y_0}$ is a finite set, there is an open set $U_1 \subset \overline{U_1} \subset U(\subset X)$ such that $\mathrm{SP}\phi_{y_0} \cap U_1 = \mathrm{SP}\phi_{y_0} \cap U (= \mathrm{SP}\phi_{y_0} \cap \overline{U})$, or in other words, $\mathrm{SP}\phi_{y_0} \subset U_1 \cup (X\backslash\overline{U})$. Considering the open set $U_1 \cup (X\backslash\overline{U})$, by the continuity of the function

$$y \mapsto \mathrm{SP}\phi_y \in P^k X,$$

where $k = \mathrm{rank}(\phi(\mathbf{1}))$, there is an open set $W_1 \ni y_0$ such that

$$\mathrm{SP}\phi_y \subset U_1 \cup (X\backslash\overline{U}), \qquad \forall y \in W_1. \tag{1}$$

Let $\chi \in C(X)$ be a function satisfying

$$\chi(x) = \begin{cases} 1 & \text{if} \quad x \in \overline{U_1} \\ 0 & \text{if} \quad x \in X\backslash U \ . \end{cases}$$

Then from (1) and the definition of spectral projection it follows that

$$\phi(\chi)(y) = P(y), \ \forall y \in W_1.$$

In particular, $P(y)$ is continuous at y_0.
The additional part of the lemma follows from the continuity of $P(y)$ and the connectedness of W.

$\square$

In the above proof, we used the following fact, a consequence of the continuity of the map $y \mapsto \mathrm{SP}\phi_y$. We state it separately for our future use.

Lemma 1.2.10. *Let X be a finite simplicial complex, $X_1 \subset X$ be a closed subset, and $\phi : C(X) \to M_\bullet(C(Y))$ be a homomorphism. For any $y_0 \in Y$, if $SP\phi_{y_0} \cap X_1 = \emptyset$, then there is an open set $W \ni y_0$ such that $SP\phi_y \cap X_1 = \emptyset$ for any $y \in W$.*
Another equivalent statement is the following. Let $U \subset X$ be an open subset. For any $y_0 \in Y$, if $SP\phi_{y_0} \subset U$, then there is an open set $W \ni y_0$ such that $SP\phi_y \subset U$ for any $y \in W$.

1.2.11. In fact the above lemma is a consequence of the following more general principle: If $\phi : C(X) \to M_\bullet(\mathbb{C})$ is a homomorphism satisfying $\mathrm{SP}\phi \subset U$ for a certain open set U, then for any homomorphism $\psi : C(X) \to M_\bullet(\mathbb{C})$ which is close enough to ϕ, we have $\mathrm{SP}\psi \subset U$. We state it as the following lemma.

Lemma 1.2.12. *Let $F \subset C(X)$ be a finite set of elements which generate $C(X)$ as a C^*-algebra. For any $\varepsilon > 0$, there is a $\delta > 0$ such that if two homomorphisms ϕ, ψ : $C(X) \to M_\bullet(\mathbb{C})$ satisfy*

$$\|\phi(f) - \psi(f)\| < \delta, \qquad \forall f \in F,$$

then $SP\psi$ and $SP\phi$ can be paired within ε. In particular, $SP\psi \subset U$, where U is the open set defined by $U = \{x \in X \mid \exists x' \in SP\phi \text{ with } dist(x, x') < \varepsilon\}$.

1.2.13. For any C^* algebra A (usually we let $A = C(X)$ or $A = PM_k(C(X))P$), any homomorphism $\phi : A \to M_\bullet(C(Y))$, and any closed subset $Y_1 \subset Y$, denote by $\phi|_{Y_1}$ the following composition:

$$A \xrightarrow{\phi} M_\bullet(C(Y)) \overset{\text{restriction}}{\longrightarrow} M_\bullet(C(Y_1)).$$

(As usual, for a subset or subalgebra $A_1 \subset A$, $\phi|_{A_1}$ will be used to denote the restriction of ϕ to A_1. We believe that there will be no danger of confusion as the meaning will be clear from the context.)
The following trivial fact will be used frequently.

Lemma 1.2.14. *Let $Y_1, Y_2 \subset Y$ be two closed subsets. If $\phi_1 : A \to M_k(C(Y_1))$*

and $\phi_2 : A \to M_k(C(Y_2))$ are two homomorphisms with $\phi_1|_{Y_1\cap Y_2} = \phi_2|_{Y_1\cap Y_2}$, then for any $a \in A$, the matrix-valued function $y \mapsto \phi(a)(y)$, where

$$\phi(a)(y) = \begin{cases} \phi_1(a)(y) & \text{if} \quad y \in Y_1 \\ \phi_2(a)(y) & \text{if} \quad y \in Y_2 \ , \end{cases}$$

is a continuous function on $Y_1 \cup Y_2$ (i.e., it is an element of $M_k(C(Y_1 \cup Y_2))$. Furthermore, $a \mapsto \phi(a)$ defines a homomorphism $\phi : A \to M_k(C(Y_1 \cup Y_2))$.

1.2.15. Let X be a compact connected space and let Q be a projection of rank n in $M_N(C(X))$. The WEAK VARIATION OF A FINITE SET $F \subset QM_N(C(X))Q$ is defined by

$$\omega(F) = \sup_{\Pi_1,\Pi_2} \inf_{u\in U(n)} \max_{a\in F} \|u\Pi_1(a)u^* - \Pi_2(a)\|$$

where Π_1, Π_2 run through the set of irreducible representations of $QM_N(C(X))Q$ into $M_n(\mathbb{C})$.
Let X_i be compact connected spaces and $Q_i \in M_{n_i}(C(X_i))$ be projections. For a finite set $F \subset \bigoplus_i Q_iM_{n_i}(C(X_i))Q_i$, define the WEAK VARIATION $\omega(F)$ to be $\max_i \omega(\pi_i(F))$, where $\pi_i : \bigoplus_i Q_iM_{n_i}(C(X_i))Q_i \to Q_iM_{n_i}(C(X_i))Q_i$ is the natural project map onto the i-th block.
The set F is said to be WEAKLY APPROXIMATELY CONSTANT TO WITHIN ε if $\omega(F) < \varepsilon$. The other description of this concept can be found in [EG2, 1.4.11] (see also [D2, 1.3]).

1.2.16. Let $\phi : M_k(C(X)) \to PM_l(C(Y))P$ be a unital homomorphism. Set $\phi(e_{11}) = p$, where e_{11} is the canonical matrix unit corresponding to the upper left corner. Set

$$\phi_1 = \phi|_{e_{11}M_k(C(X))e_{11}} : C(X) \longrightarrow pM_l(C(Y))p.$$

Then $PM_l(C(Y))P$ can be identified with $pM_l(C(Y))p \otimes M_k$ in such a way that

$$\phi = \phi_1 \otimes \mathbf{1}_k.$$

Let us define

$$\mathrm{SP}\phi_y := \mathrm{SP}(\phi_1)_y,$$

$$\mathrm{SP}\phi := \mathrm{SP}\phi_1,$$

$$\mathrm{SPV}(\phi) := \mathrm{SPV}(\phi_1)\ .$$

Suppose that X and Y are connected. Let Q be a projection in $M_k(C(X))$ and $\phi : QM_k(C(X))Q \to PM_l(C(Y))P$ be a unital map. By the Dilation Lemma (2.13 of [EG2]; see Lemma 1.3.1 below), there are an n, a projection $P_1 \in M_n(C(Y))$, and a unital homomorphism

$$\tilde{\phi} : M_k(C(X)) \longrightarrow P_1M_n(C(Y))P_1$$

such that

$$\phi = \tilde{\phi}|_{QM_k(C(X))Q}.$$

(Note that this implies that P is a subprojection of P_1.) We define:

$$\mathrm{SP}\phi_y := \mathrm{SP}\tilde{\phi}_y,$$

$$\mathrm{SP}\phi := \mathrm{SP}\tilde{\phi},$$

$$\mathrm{SPV}(\phi) := \mathrm{SPV}(\tilde{\phi}) \ .$$

(Note that these definitions do not depend on the choice of the dilation $\tilde{\phi}$.)
The following lemma was essentially proved in [EG2, 3.27] (the additional part is [EG 1.4.13]).

Lemma 1.2.17. *Let X be a path connected compact metric space. Let $p_0, p_1, p_2, \cdots, p_n \in M_\bullet(C(Y))$ be mutually orthogonal projections such that rank$(p_i) \geq$ rank(p_0), $i = 1, 2, \cdots, n$. Let $\{x_1, x_2, \cdots, x_n\}$ be a $\frac{\delta}{2}$-dense subset of X. If a homomorphism $\phi : C(X) \to M_\bullet(C(Y))$ is defined by*

$$\phi(f) = \phi_0(f) \oplus \sum_{i=1}^{n} f(x_i)p_i,$$

where $\phi_0 : C(X) \to p_0 M_\bullet(C(Y))p_0$ is an arbitrary homomorphism, then SPV$(\phi) < \delta$. Consequently, if a finite set $F \subset C(X)$ satisfies the condition that $\|f(x) - f(x')\| < \varepsilon$, for any $f \in F$, whenever dist$(x, x') < \delta$, then $\phi(F)$ is weakly approximately constant to within ε.
(For convenience, we will call such a homomorphism $\psi : C(X) \to M_\bullet(C(Y))$, defined by $\psi(f) = \sum_{i=1}^{n} f(x_i)p_i$, a homomorphism defined by point evaluations on the set $\{x_1, x_2, \cdots, x_n\}$.)

Proof: For any two points $y, y' \in Y$, the sets $\mathrm{SP}\phi_y$ and $\mathrm{SP}\phi_{y'}$ have the following subset in common:

$$\{x_1^{\sim \mathrm{rank}(p_1)}, x_2^{\sim \mathrm{rank}(p_2)}, \cdots, x_n^{\sim \mathrm{rank}(p_n)}\}.$$

The remaining parts of $\mathrm{SP}\phi_y$ and $\mathrm{SP}\phi_{y'}$ are $\mathrm{SP}(\phi_0)_y$ and $\mathrm{SP}(\phi_0)_{y'}$, respectively, which have at most rank(p_0) elements.
It is easy to prove the following fact. For any $a, b \in X$, the sets $\{a, x_1, x_2, \cdots, x_n\}$ and $\{b, x_1, x_2, \cdots, x_n\}$ can be paired within δ. In fact, by path connectedness of X and $\frac{\delta}{2}$-density of the set $\{x_1, x_2, \cdots, x_n\}$, one can find a sequence

$$a, x_{j_1}, x_{j_2}, \cdots, x_{j_k}, b$$

beginning with a and ending with b such that each pair of consecutive terms has distance smaller than δ. So $\{a, x_{j_1}, \cdots, x_{j_{k-1}}, x_{j_k}\}$ can be paired with $\{x_{j_1}, x_{j_2}, \cdots, x_{j_k}, b\}$ $(= \{b, x_{j_1}, \cdots, x_{j_k}\})$ one by one to within δ. The other parts of the sets are identical, each element can be paired with itself.

Combining the above fact with the condition that $\mathrm{rank}(p_i) \geq \mathrm{rank}(p_0)$ for any i, we know that $\mathrm{SP}\phi_y$ and $\mathrm{SP}\phi_{y'}$ can be paired within δ. That is, $\mathrm{SPV}(\phi) < \delta$. The rest of the lemma is obvious. Namely, for any two points y, y', $\phi(f)(y)$ is approximately unitarily equivalent to $\phi(f)(y')$ to within ε, by the same unitary for all $f \in F$ (see [EG2, 1.4.13]).

$\square$

1.2.18. In the last part of the above lemma, one does not need ϕ_0 to be a homomorphism to guarantee $\phi(F)$ to be weakly approximately constant to within a small number. In fact, the following is true.
Suppose that all the notations are as in 1.2.17 except that the maps $\phi_0 : C(X) \to p_0 M_\bullet(C(Y))p_0$ and $\phi : C(X) \to M_\bullet(C(Y))$ are no longer homomorphisms. Suppose that for any $y \in Y$, there is a homomorphism $\psi_y : C(X) \to p_0(y)M_\bullet(\mathbb{C})p_0(y)$ such that

$$\|\phi_0(f)(y) - \psi_y(f)\| < \varepsilon, \ \forall f \in F.$$

Then the set $\phi(F)$ is weakly approximately constant to within 3ε. One can prove this claim as follows.
For any $y \in Y$, define a homomorphism $\phi_y \to M_\bullet(\mathbb{C})$ by $\phi_y(f) = \psi_y(f) \oplus \sum_{i=1}^{n} f(x_i)p_i$. Then for any two points $y, y' \in Y$, as same as in Lemma 1.2.17, $\mathrm{SP}(\phi_y)$ and $\mathrm{SP}(\phi_{y'})$ can be paired within δ. Therefore, $\phi_y(f)$ is approximately unitarily equivalent to $\phi_{y'}(f)$ to within ε, by the same unitary for all $f \in F$.
On the other hand,

$$\|\phi(f)(y) - \phi_y(f)\| < \varepsilon \quad \text{and} \quad \|\phi(f)(y') - \phi_{y'}(f)\| < \varepsilon, \qquad \forall f \in F.$$

Hence $\phi(f)(y)$ is approximately unitarily equivalent to $\phi(f)(y')$ to within 3ε, by the same unitary for all $f \in F$.

1.2.19. Suppose that $F \subset M_k(C(X))$ is a finite set and $\varepsilon > 0$. Let $F' \subset C(X)$ be the finite set consisting of all entries of elements in F and $\varepsilon' = \frac{\varepsilon}{k}$, where k is the order of the matrix algebra $M_k(C(X))$.
It is well known that, for any $k \times k$ matrix $a = (a_{ij}) \in M_k(B)$ with entries $a_{ij} \in B$, $\|a\| \leq k \max_{ij} \|a_{ij}\|$. This implies the following two facts.
Fact 1. If $\phi_1, \psi_1 \in \mathrm{Map}(C(X), B)$ are (complete positive) linear $*$-contraction (as the notation in 1.1.1) which satisfy

$$\|\phi_1(f) - \psi_1(f)\| < \varepsilon', \ \ \forall f \in F',$$

then $\phi := \phi_1 \otimes \mathrm{id}_k \in \mathrm{Map}(M_k(C(X)), M_k(B))$ and $\psi := \psi_1 \otimes \mathrm{id}_k \in \mathrm{Map}(M_k(C(X)), M_k(B))$ satisfy

$$\|\phi(f) - \psi(f)\| < \varepsilon, \ \ \forall f \in F.$$

Fact 2. Suppose that $\phi_1 \in \mathrm{Map}(C(X), M_\bullet(C(Y)))$ is a (complete positive) linear $*$-contraction. If $\phi_1(F')$ is weakly approximately constant to within ε', then $\phi_1 \otimes \mathrm{id}_k(F)$ is weakly approximately constant to within ε.

Suppose that a homomorphism $\phi_1 \in \mathrm{Hom}(C(X), B)$ has a decomposition described as follows. There exist mutually orthogonal projections $p_1, p_2 \in B$ with $p_1 + p_2 = \mathbf{1}_B$ and $\psi_1 \in \mathrm{Hom}(C(X), p_2 B p_2)$ such that

$$\|\phi_1(f) - p_1\phi_1(f)p_1 \oplus \psi_1(f)\| < \varepsilon', \quad \forall f \in F'.$$

Then there is a decomposition for $\phi := \phi_1 \otimes \mathrm{id}_k$:

$$\|\phi(f) - P_1\phi_1(f)P_1 \oplus \psi(f)\| < \varepsilon, \quad \forall f \in F,$$

where $\psi := \psi_1 \otimes \mathrm{id}_k$ and $P_1 = p_1 \otimes \mathbf{1}_k$.
In particular, if $B = M_\bullet(C(Y))$ and ψ_1 is described by

$$\psi_1(f)(y) = \sum f(\alpha_i(y))q_i(y), \quad \forall f \in C(X),$$

where $\sum q_i = p_2$ and $\alpha_i : Y \to X$ are continuous maps, then ψ can be described by

$$\psi(f)(y) = \sum q_i(y) \otimes f(\alpha_i(y)), \quad \forall f \in M_k(C(X)),$$

regarding $M_k(M_\bullet(C(Y)))$ as $M_\bullet(C(Y)) \otimes M_k$.
If α_i are constant maps, the homomorphism ψ_1 is called a homomorphism defined by point evaluations as in Lemma 1.2.17. In this case, we will also call the above ψ a homomorphism defined by point evaluations.
From the above, we know that to decompose a homomorphism $\phi \in \mathrm{Hom}(M_k(C(X)), M_\bullet(C(Y)))$, one only needs to decompose $\phi_1 := \phi|_{e_{11}M_k(C(X))e_{11}} \in Hom(C(X), \phi(e_{11})M_\bullet(C(Y))\phi(e_{11}))$.

1.3 Full matrix algebras, corners, and the dilation lemma

Some results in this article deal with a corner $QM_N(C(X))Q$ of the matrix algebra $M_N(C(X))$. But using the following lemma and some other techniques, we can reduce the problems to the case of a full matrix algebra $M_N(C(X))$. The following dilation lemma is Lemma 2.13 of [EG2].

Lemma 1.3.1. (cf. Lemma 2.13 of [EG2]) *Let X and Y be any connected finite CW complexes. If $\phi : QM_k(C(X))Q \to PM_n(C(Y))P$ is a unital homomorphism, then there are an n_1, a projection $P_1 \in M_{n_1}(C(Y))$, and a unital homomorphism $\tilde{\phi} : M_k(C(X)) \to P_1 M_{n_1}(C(Y))P_1$ with the property that $QM_k(C(X))Q$ and $PM_n(C(Y))P$ can be identified as corner subalgebras of $M_k(C(X))$ and $P_1M_{n_1}(C(Y))P_1$ respectively (i.e., Q and P can be considered to be subprojections of $\mathbf{1}_k$ and P_1, respectively) and, furthermore, in such a way that ϕ is the restriction of $\tilde{\phi}$.*
If $\phi_t : QM_k(C(X))Q \to PM_n(C(Y))P$, $(0 \le t \le 1)$ is a path of unital homomorphisms, then there are $P_1M_{n_1}(C(Y))P_1$ (as above) and a path of unital homomorphisms $\tilde{\phi}_t : M_k(C(X)) \to P_1M_{n_1}(C(Y))P_1$ such that $QM_k(C(X))Q$ and $PM_n(C(Y))P$ are corner subalgebras of $M_k(C(X))$ and $P_1M_{n_1}(C(Y))P_1$ respectively and ϕ_t is the restriction of $\tilde{\phi}_t$.

DEFINITION 1.3.2. Let A be a C^*-algebra. A sub-C^*-algebra $A_1 \subset A$ will be called a LIMIT CORNER SUBALGEBRA of A, if there is a sequence of increasing projections

$$P_1 \le P_2 \le \cdots \le P_n \le \cdots,$$

such that $A_1 = \overline{\bigcup_{n=1}^{\infty} P_n A P_n}$.
Using Lemma 1.3.1, it is routine to prove the following lemma.

LEMMA 1.3.3. (cf. 4.24 of [EG2]) *For any AH algebra* $A = \lim\limits_{\rightarrow}(A_n = \bigoplus_{i=1}^{t_n} P_{n,i} M_{[n,i]}(C(X_{n,i}))P_{n,i}, \phi_{n,m})$, *there is an inductive limit* $\tilde{A} = \lim\limits_{\rightarrow}(\tilde{A}_n = \bigoplus_{i=1}^{t_n} M_{\{n,i\}}(C(X_{n,i})), \tilde{\phi}_{n,m})$ *of full matrix algebras over* $\{X_{n,i}\}$, *such that* A *is isomorphic to a limit corner subalgebra of* $\tilde{A}$. *In particular, each* $P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}$ *is a corner subalgebra of* $M_{\{n,i\}}(C(X_{n,i}))$ *and* $\phi_{n,m}$ *is the restriction of* $\tilde{\phi}_{n,m}$ *on* $A_n = \bigoplus_{i=1}^{t_n} P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}$.
Furthermore, A *is stably isomorphic to* $\tilde{A}$.

REMARK 1.3.4. In the above lemma, in general, the homomorphisms $\tilde{\phi}_{n,m}$ cannot be chosen to be unital, even if all the homomorphisms $\phi_{n,m}$ are unital. If A is unital, then A can be chosen to be the cut-down of $\tilde{A}$ by a single projection (rather than a sequence of projections).

REMARK 1.3.5. In Lemma 1.3.3, if A is simple and satisfies the very slow dimension growth condition, then so is $\tilde{A}$. Hence when we consider the nonunital case, we may always assume that A is an inductive limit of direct sums of full matrix algebras over $C(X_{n,i})$ without loss of generality. Since the reduction theorem in this paper will be proved without the assumption of unitality, we may assume that the C^*-algebra A is an inductive limit of full matrix algebras over finite simplicial complexes. But even in this case, we still need to consider the cut-down $PM_l(C(X))P$ of $M_l(C(X))$ in some situations, since the image of a trivial projection may not be trivial.
In the proof of the decomposition theorem in §4, we will NOT assume that A is unital but we will assume that A is the inductive limit of full matrix algebras. Note that a projection in $M_\bullet(C(X))$ corresponds to a complex vector bundle over X. The following result is well known (see Chapter 8 of [Hu]). This result is often useful when we reduce the proof of a result involving the cut-down $PM_l(C(X))P$ to the special case of the full matrix algebra $M_l(C(X))$.

LEMMA 1.3.6. *Let* X *be a connected simplicial complex and* $P \in M_l(C(X))$ *be a non-zero projection. Let* $n = \mathrm{rank}(P) + \dim(X)$ *and* $m = 2\dim(X) + 1$.
Then P *is Murray-von Neumann equivalent to a subprojection of* $\mathbf{1}_n$, *and* $\mathbf{1}_n$ *is Murray-von Neumann equivalent to a subprojection of* $\underbrace{P \oplus P \oplus \cdots P}_{m}$, *where* $\mathbf{1}_n$ *is a trivial projection with rank* n. *Therefore,* $PM_l(C(X))P$ *can be identified*

as a corner subalgebra of $M_n(C(X))$, and $M_n(C(X))$ can be identified as a corner subalgebra of $M_m(PM_l(C(X))P)$.

1.4 Topological preliminaries

In this subsection, we will introduce some notations and results in the topology of simplicial complexes. We will also introduce a well known method for the construction of cross sections of a fibre bundle. The content of this subsection may be found in [St], [Hu] and [Wh].

1.4.1. Let X be a connected simplicial complex. Endow X with a metric d as follows.
For each n-simplex Δ, one can identify Δ with an n-simplex in $\mathbb{R}^n$ whose edges are of length 1 (of course the identification should preserve the affine structure of the simplices). (Such a simplex is the convex hull of $n+1$ points $\{x_0, x_1, \cdots, x_n\}$ in $\mathbb{R}^n$ with $\mathrm{dist}(x_i, x_j) = 1$ for any $i \neq j \in \{0, 1, \cdots n\}$.) Such an identification gives rise to a unique metric on Δ. The restriction of metric d of X to Δ is defined to be the above metric for any simplex $\Delta \subset X$. For any two points $x, y \in X$, $d(x, y)$ is defined to be the length of the shortest path connecting x and y. (The length is measured in individual simplexes, by breaking the path into small pieces.)
If X is not connected, denote by L the maximum of the diameters of all the connected components. Define $d(x, y) = L + 1$, if x and y are in different components. (Recall that all the simplicial complexes in this article are supposed to be finite.)

1.4.2. For a simplex Δ, by $\partial\Delta$, we denote the boundary of the simplex Δ, which is the union of all proper faces of Δ. Note that if Δ is a single point—zero dimensional simplex, then $\partial\Delta = \emptyset$. Obviously, $\dim(\partial\Delta) = \dim(\Delta) - 1$. (We use the standard convention that the dimension of the empty space is -1.) By interior(Δ), we denote $\Delta\backslash\partial\Delta$. Let X be a simplicial complex. Obviously, for each $x \in X$, there is a unique simplex Δ such that $x \in \mathrm{interior}(\Delta)$, which is the simplex Δ of lowest dimension with the condition that $x \in \Delta$. (Here we use the fact that if two different simplices of the same dimension intersect, then the intersection is a simplex of lower dimension.)
For any simplex Δ, define

$$\mathrm{Star}(\Delta) = \bigcup \{\mathrm{interior}(\Delta') \mid \Delta' \cap \Delta \neq \emptyset\}.$$

Then $\mathrm{Star}(\Delta)$ is an open set which covers Δ.
We will use the following two open covers of the simplicial complex X.
(a) For any vertex $x \in X$, let

$$W_x = \mathrm{Star}(\{x\}) (= \bigcup \{\mathit{interior}(\Delta) \mid x \in \Delta\}).$$

Obviously $\{W_x\}_{x \in \mathrm{Vertex}(X)}$ is an open cover parameterized by vertices of X.

In this open cover, the intersection $W_{x_1} \cap W_{x_2} \cap \cdots \cap W_{x_k}$ is nonempty if and only if $x_1, x_2, \cdots, x_k$ span a simplex of X.
(b) We denote the original simplicial structure of X by σ. Introduce a barycentric subdivision (X, τ) of (X, σ).
Then for each simplex Δ of (X, σ) (before subdivision), there is exactly one point $C_\Delta \in \mathrm{Vertex}(X, \tau)$—the barycenter of Δ, such that $C_\Delta \in \mathrm{interior}(\Delta)$. (Here interior($\Delta$) is clearly defined by referring Δ as a simplex of (X, σ).)
Define

$$U_\Delta = \mathrm{Star}_{(X,\tau)}(\{C_\Delta\}).$$

As in (a), $\{U_\Delta \mid \Delta \text{ is a simplex of } (X, \sigma)\}$ is an open cover. In fact, $U_\Delta \supset \mathrm{interior}(\Delta)$. This open cover is parameterized by simplices of (X, σ) (also by vertices of (X, τ), since there is a one to one correspondence between the vertices of (X, τ) and the simplices of (X, σ)).
This cover satisfies the following condition: The intersection $U_{\Delta_1} \cap U_{\Delta_2} \cap \cdots \cap U_{\Delta_k}$ is nonempty if and only if one can reorder the simplices, such that

$$\Delta_1 \subset \Delta_2 \subset \cdots \Delta_n.$$

One can verify that the open cover in (b) is a refinement of the open cover in (a).

1.4.3. It is well known that a simplicial complex X is locally contractible. That is, for any point $x \in X$ and an open neighborhood $U \ni x$, there is an open neighborhood $W \ni x$ with $W \subset U$ such that W can be contracted to a single point inside U. (One can prove this fact directly, using the metric in 1.4.1.)
One can endow different metrics on X, but all the metrics are required to induce the same topology as the one in 1.4.1.
Using the local contractibility and the compactness of X, one can prove the following fact.
For any simplicial complex X with a metric d (may be different from the metric in 1.4.1), there are $\delta_{X,d} > 0$ and a nondecreasing function $\rho : (0, \delta_{X,d}] \to \mathbb{R}^+$ such that the following are true.
(1) $\lim_{\delta \to 0^+} \rho(\delta) = 0$, and
(2) for any $\delta \in (0, \delta_{X,d}]$ and $x_0 \in X$, the ball $B_\delta(x_0)$ with radius δ and centre x_0 (see 1.1.7 (d) for the notation) can be contracted into a single point within the ball $B_{\rho(\delta)}(x_0)$. I.e., there is a continuous map $\alpha : B_\delta(x_0) \times [0, 1] \to B_{\rho(\delta)}(x_0)$ such that
(i) $\alpha(x, 0) = x$ for any $x \in B_\delta(x_0)$,
(ii) $\alpha(x, 1) = x_0$ for any $x \in B_\delta(x_0)$.
The following lemma is a consequence of the above fact.

LEMMA 1.4.4. *For any simplicial complex X with metric d, there are $\delta_{X,d} > 0$ and a nondecreasing function $\rho : (0, \delta_{X,d}] \to \mathbb{R}^+$ such that the following are true.*

(1) $\lim_{\delta\to 0^+}\rho(\delta)=0$*, and*
(2) for any ball $B_\delta(x_0)$ *with radius* $\delta\le\delta_{X,d}$*, any simplex* Δ *(not assumed to be a simplex in* X*), and any continuous map* $f:\partial\Delta\to B_\delta(x_0)$*, there is a continuous map* $g:\Delta\to B_{\rho(\delta)}(x_0)$ *such that* $g(y)=f(y)$ *for any* $y\in\partial\Delta$*.*

Proof: The simplex Δ can be identified with $\partial\Delta\times[0,1]/\partial\Delta\times\{1\}$ in such a way that $\partial\Delta$ is identified with $\partial\Delta\times\{0\}$. Define the map g by

$$g(y,t)=\alpha(f(y),t)\in B_{\rho(\delta)}(x_0),\forall y\in\partial\Delta,t\in[0,1],$$

where α is the map in 1.4.3.

$\square$

1.4.5. The following is a well known result in differential topology: Suppose that M is an m-dimensional smooth manifold, $N\subset M$ is an n-dimensional submanifold. If Y is an l-dimensional simplicial complex with $l<m-n$, then for any continuous map $f:Y\to M$ and any $\varepsilon>0$, there is a continuous map $g:Y\to M$ such that
(i) $g(Y)\cap N=\emptyset$ and
(ii) $\mathrm{dist}(g(y),f(y))<\varepsilon$, for any $y\in Y$.
There is an analogous result in the case of a simplicial complex M and a subcomplex N. Instead of the assumption that M is an m-dimensional smooth manifold, let us suppose that M has the PROPERTY $D(m)$: for each $x\in M$, there is a contractible open neighborhood $U_x\ni x$ such that $U_x\setminus\{x\}$ is $(m-2)$-connected, i.e.,

$$\pi_i(U_x\setminus\{x\})=0\quad\text{for any}\quad i\in\{0,1,\cdots m-2\}.$$

(We use the following convention: by $\pi_0(X)=0$, it will be meant that X is a path connected nonempty space.)
Note that $\mathbb{R}^m\backslash\{0\}$ is $(m-2)$-connected. Therefore, any m-dimensional manifold has property $D(m)$.
The following result is the relative version of Theorem 5.4.16 of [St] (see page 111 of [St]), which also holds according to the top of page 112 of [St].

Proposition 1.4.6. Suppose that M is a simplicial complex with property $D(m)$, and $N\subset M$ is a sub-simplicial complex. Suppose that Y is a simplicial complex of dimension $l<m-\dim(N)$, and suppose that $Y_1\subset Y$ is a sub-simplicial complex. Suppose that $f:Y\to M$ is a continuous map such that $f(Y_1)\cap N=\emptyset$. For any $\varepsilon>0$, there is a continuous map $f_1:Y\to M$ such that
(i) $f_1|_{Y_1}=f|_{Y_1}$,
(ii) $f_1(Y)\cap N=\emptyset$, and
(iii) $d(f(y),f_1(y))<\varepsilon$ for any $y\in Y$.

1.4.7. Let X,F be two simplicial complexes.

Let $\Gamma \subset Homeo(F)$ be a subgroup of the group of homeomorphisms of the space F.
Let us recall the definition of fibre bundle. A FIBRE BUNDLE OVER X WITH FIBRE F AND STRUCTURE GROUP Γ, is a simplicial complex M with a continuous surjection $p : M \to X$ such that the following is true. There is an open cover $\mathcal{U}$ of X, and associated to each $U \in \mathcal{U}$, there is a homeomorphism

$$t_U : p^{-1}(U) \to U \times F$$

(called a local trivialization of the bundle) such that
(1) Each t_U takes the fibre of $p^{-1}(U)$ at $x \in U$ to the fibre of $U \times F$ at the same point x—a trivialization of the restriction $p^{-1}(U)$ of the fibre bundle to U, i.e., the diagram

$$\begin{tikzcd} p^{-1}(U) \arrow[rr, "t_U"] \arrow[dr, "p"'] & & U \times F \arrow[dl, "p_1"] \\ & U & \end{tikzcd}$$

is commutative, where p_1 denotes the project map from the product $U \times F$ to the first factor U, and
(2) The given local trivializations differ fibre-wise only by homeomorphisms in the structure group Γ: for any $U, V \in \mathcal{U}$ and $x \in U \cap V$,

$$t_U \circ t_V^{-1}|_{\{x\}\times F} \in \Gamma \subset Homeo(F).$$

Furthermore, we will also suppose that THE METRIC d OF F IS INVARIANT UNDER THE ACTION OF ANY ELEMENT $g \in \Gamma$, i.e., $d(g(x), g(y)) = d(x, y)$ for any $x, y \in F$. We will see, the fibre bundles constructed in §2, satisfy this condition.
(Note that, if F is the vector space $\mathbb{R}^n$ with Euclidean metric d, then d is invariant under the action of $O(n) \subset Homeo(F)$, but not invariant under the action of $Gl(n) \subset Homeo(F)$.)
A subset $F_1 \subset F$ is called Γ-INVARIANT if for any $g \in \Gamma$, $g(F_1) \subset F_1$.

1.4.8. A CROSS SECTION of fibre bundle $p : M \to X$ is a continuous map $f : X \to M$ such that

$$(p \circ f)(x) = x \quad \text{for any} \quad x \in X.$$

The following Theorem is a consequence of Proposition 1.4.6. The proof is a standard argument often used in the construction of cross sections for fibre bundles (see [Wh]). (In the literature such an argument is often taken for granted.) We give it here for the convenience of the reader.

Theorem 1.4.9. *Suppose that $p : M \to X$ is a fibre bundle with fibre F. Suppose that $F_1 \subset F$ is a Γ-invariant sub-simplicial complex of F. Suppose that F has the property $D(m)$ and $\dim(X) < m - \dim(F_1)$. Then for any cross section $s : X \to M$ and any $\varepsilon > 0$, there is a cross section $s_1 : X \to M$ such that the following two statements are true:*
(1) $(t_U \circ s_1)(x) \notin \{x\} \times F_1$ for any $x \in U \in \mathcal{U}$, where $\mathcal{U}$ is the open cover of X and $\{t_U\}_{U\in\mathcal{U}}$ is the local trivialization of the fibre bundle. That is, $s_1(X)$ avoids F_1 in each fibre;
(2) $d(s_1(x), s(x)) < \varepsilon$ for any $x \in X$, where the distance is taken in the fibre F.

Proof: If the fibre bundle is trivial, then the cross sections of the bundle can be identified with maps from X to F. The conclusion follows immediately from 1.4.6.
For the general case, we will use the local trivializations.
For each open set $U \in \mathcal{U}$, using the trivialization

$$t_U : p^{-1}(U) \to U \times F,$$

each cross section f on U induces a continuous map $\tilde{f}_U : U \to F$ by

$$\tilde{f}_U(x) = p_2(t_U(f(x))),$$

where $p_2 : U \times F \to F$ is the projection onto the second factor.
Suppose that δ_F is as in 1.4.3 (see 1.4.4 also). That is, there is a nondecreasing function $\rho : (0, \delta_F] \to (0, \infty)$ such that $\lim_{\delta\to 0^+} \rho(\delta) = 0$ and such that any δ-ball $B_\delta(x)$ can be contracted to a single point within $B_{\rho(\delta)}(x)$.
Let $\dim(X) = n$. Choose a finite sequence of positive numbers

$$\varepsilon_n > \varepsilon_{n-1} > \varepsilon_{n-2} > \cdots > \varepsilon_1 > \varepsilon_0 > 0$$

as follows. Set $\varepsilon_n = \min\{\delta_F, \varepsilon\}$. Then choose ε_{n-1} to satisfy

$$\rho(3\varepsilon_{n-1}) < \frac{1}{3}\varepsilon_n \quad \text{and} \quad 3\varepsilon_{n-1} < \frac{1}{3}\varepsilon_n.$$

Once ε_l is defined, then choose ε_{l-1} to satisfy

$$\rho(3\varepsilon_{l-1}) < \frac{1}{3}\varepsilon_l \;\; \text{and} \;\; 3\varepsilon_{l-1} < \frac{1}{3}\varepsilon_l.$$

Repeat this procedure until we choose ε_0 to satisfy

$$\rho(3\varepsilon_0) < \frac{1}{3}\varepsilon_1 \;\; \text{and} \;\; 3\varepsilon_0 < \frac{1}{3}\varepsilon_1.$$

Let us refine the given simplicial complex structure on X in such a way that each simplex Δ is covered by an open set $U \in \mathcal{U}$ and that for any simplex Δ

and an open set $U \in \mathcal{U}$ which covers Δ, the map $\tilde{s}_U|_\Delta : \Delta \to F$, induced by the cross section s, satisfies

$$\text{diameter}(\tilde{s}_U(\Delta)) < \varepsilon_0.$$

(Since the metric on F is invariant under the action of any element in Γ, the above inequality holds or not does not depend on the choice of the open set U which covers Δ. In what follows, we will use this fact many times without saying so.)

We will apply Proposition 1.4.6 to each simplex of X from the lowest dimension to the highest dimension.

For any $l \in \{0, 1, \cdots, n\}$, let us denote the l-skeleton of X by $X^{(l)}$. So $X^{(n)} = X$, and $X^{(0)}$ is the set of vertices of X.

Step 1. Fix a vertex $x \in X^{(0)}$, and suppose that $x \in U \in \mathcal{U}$. Applying Proposition 1.4.6 to $\{x\}$ (in place of Y with $Y_1 = \emptyset$) and F (in place of M with $N = F_1$), there exists $\tilde{s}^0(x) \in F \setminus F_1$ such that

$$d(\tilde{s}^0(x), \tilde{s}_U(x)) < \varepsilon_0.$$

Any choice of $\tilde{s}^0(x)$ gives a cross section s^0 on $\{x\}$ by

$$s^0(x) = t_U^{-1}(x, \tilde{s}^0(x)),$$

where $(x, \tilde{s}^0(x)) \in \{x\} \times F \subset U \times F$. Defining s^0 on all vertices, we obtain a cross section s^0 on $X^{(0)}$ such that

$$d(\tilde{s}^0_U(x), \tilde{s}_U(x)) < \varepsilon_0$$

for each $x \in U \cap X^{(0)}$.

Step 2. Suppose that for $l < n = \dim(X)$, there is a cross section $s^l : X^{(l)} \to M$ such that for any $U \in \mathcal{U}$ and any $x \in X^{(l)} \cap U$, we have $\tilde{s}^l_U(x) \notin F_1$, and

$$(*) \qquad d(\tilde{s}^l_U(x), \tilde{s}_U(x)) < \varepsilon_l.$$

Let us define a cross section $s^{l+1} : X^{(l+1)} \to M$ as follows. We will work one by one on each $(l+1)$-simplex Δ.

First, we shall simply extend the cross section $s^l|_{\partial\Delta}$ to a cross section on Δ (see Substep 2.1 below). Then, apply Proposition 1.4.6 to perturb the cross section $s^l|_\Delta$ to avoid F_1 in each fibre (see Substep 2.2 below). Again, since Proposition 1.4.6 is only for maps (not for cross sections), we will use $\tilde{s}^l_U|\partial\Delta : \partial\Delta \to F$ to replace $s^l|\partial\Delta$, as in Step 1.

Substep 2.1. Let Δ be an $(l+1)$-simplex. Suppose that $\Delta \subset U \in \mathcal{U}$. Then $\tilde{s}^l_U|_{\partial\Delta} : \partial\Delta \to F$ is a continuous map. Since $(*)$ holds for any $x \in \partial\Delta$, and since

$$\text{diameter}(\tilde{s}_U(\Delta)) < \varepsilon_0,$$

we have

$$\text{diameter}(\tilde{s}^l_U(\partial\Delta)) < \varepsilon_l + \varepsilon_l + \varepsilon_0.$$

Let $\delta = \varepsilon_l + \varepsilon_l + \varepsilon_0 < \delta_F$. Then there is a $y \in F$ such that $\tilde{s}^l_U(\partial\Delta) \subset B_\delta(y)$. Since $\rho(\delta) \le \rho(3\varepsilon_l) < \frac{1}{3}\varepsilon_{l+1}$, by Lemma 1.4.4, $\tilde{s}^l_U : \partial\Delta \to F$ can be extended to a map (still denoted by $\tilde{s}^l_U$)

$$\tilde{s}^l_U : \Delta \to F,$$

such that $\tilde{s}^l_U(\Delta) \subset B_{\frac{1}{3}\varepsilon_{l+1}}(y)$. Consequently, the extended map $\tilde{s}^l_U$ also satisfies that $\mathrm{diameter}(\tilde{s}^l_U(\Delta)) < \frac{2}{3}\varepsilon_{l+1}$.

SUBSTEP 2.2. Note that $\tilde{s}^l_U(x) \notin F_1$ for any $x \in \partial\Delta$. Applying Proposition 1.4.6 to Δ (in place of Y with subcomplex $Y_1 = \partial\Delta$) and to F (in the place of M with subcomplex $N = F_1$), we obtain a continuous map

$$\tilde{s}^{l+1} : \Delta \to F$$

such that

(1) $\tilde{s}^{l+1}(x) \notin F_1$ for any $x \in \Delta$,

(2) $d(\tilde{s}^{l+1}(x), \tilde{s}^l_U(x)) < \varepsilon_0$, for any $x \in \Delta$, and

(3) $\tilde{s}^{l+1}|\partial\Delta = \tilde{s}^l_U|\partial\Delta$.

The map $\tilde{s}^{l+1}$ defines a cross section s^{l+1} by

$$s^{l+1}(x) = t_U^{-1}(x, \tilde{s}^{l+1}(x)).$$

After working out all the $(l+1)$-simplices, we obtain a cross section s^{l+1} on $X^{(l+1)}$—it is a continuous cross section because it is continuous on each $(l+1)$-simplex and $s^{l+1}|\partial\Delta = s^l|\partial\Delta$ from (3) above.

Recall that $\mathrm{diameter}(\tilde{s}_U(\Delta)) < \varepsilon_0$ and $\mathrm{diameter}(\tilde{s}^l_U(\Delta)) < \frac{2}{3}\varepsilon_{l+1}$. Combining these facts with $(*)$, we have

$$(**) \qquad d(\tilde{s}^l_U(x), \tilde{s}_U(x)) < \varepsilon_l + \frac{2}{3}\varepsilon_{l+1} + \varepsilon_0$$

for any $x \in \Delta$. Combining $(**)$ and (2) above, we have

$$d(\tilde{s}^{l+1}_U(x), \tilde{s}_U(x)) < \varepsilon_l + \frac{2}{3}\varepsilon_{l+1} + 2\varepsilon_0 < \varepsilon_{l+1}$$

for any $x \in X^{l+1} \cap U$. This is $(*)$ for $l+1$ (in place of l).

STEP 3. By mathematical induction, we can define s^l for each $l = 0, 1, \cdots n$ as the above. Let $s_1 = s^n$ to finish the proof.

$\square$

The following relative version of the theorem is also true.

COROLLARY 1.4.10. *Suppose that $p : M \to X$ is a fibre bundle with fibre F and F_1 is a Γ-invariant sub-simplicial complex of F. Suppose that F has the property $D(m)$ and that $\dim(X) < m - \dim(F_1)$. Suppose that $X_1 \subset X$ is a sub-simplicial complex. Suppose that the cross section $s : X \to M$ satisfies that $(t_U \circ s)(x) \notin \{x\} \times F_1$ for any $U \in \mathcal{U}$ and any $x \in X_1 \cap U$, where $\mathcal{U}$ and t_U*

are as in the definition of fibre bundle in 1.4.7. Then for any $\varepsilon > 0$, there is a cross section $s_1 : X \to M$ such that the following three statements are true:
(1) $(t_U \circ s_1)(x) \notin \{x\} \times F_1$, for any $x \in U$ and $U \in \mathcal{U}$;
(2) $d(s_1(x), s(x)) < \varepsilon$ for any $x \in X$, where the distance is taken inside the fibre F;
(3) $s_1|_{X_1} = s|_{X_1}$.

Proof: In the proof of Theorem 1.4.9, we have essentially proved this relative version. If fact, in Step 2, we proved that a cross section on a simplex Δ can be constructed within arbitrarily small distance of the original cross section such that
(1) it avoids F_1 in each fibre, and
(2) it agrees with the original cross section on $\partial\Delta$, provided that the original cross section avoids F_1 on $\partial\Delta$.
This is a local version of the Corollary. To prove the Corollary, one only needs to apply this local version, repeatedly, to the simplices Δ with $\Delta\backslash\partial\Delta \subset X\backslash X_1$, from the lowest dimension to the highest dimension.

$\square$

1.5 About the decomposition theorem

In this subsection, we will briefly discuss the main ideas in the proof of the decomposition theorem stated in §4. Mainly, we will review the ideas in the proofs of special cases already in the literature (see especially [EG2, Theorem 2.21]), point out the additional difficulties in our new setting, and discuss how to overcome these difficulties. This subsection could be skipped without any logical gap, but we do not encourage the reader to do so, except for the expert in the classification theory. By reading this subsection, the reader will get the overall picture of the proof. In particular, how §2, §3, and the results of [Li2] fit into the picture. We will also discuss some ideas in the proof of the combinatorial results of §3. This subsection may also be helpful for understanding the corresponding parts of [EG2], [Li3], and (perhaps) other papers. Even though the discussion in this subsection is sketched, the proof of Lemma 1.5.4 and Propositions 1.5.7 and 1.5.7' are complete. We will begin our discussion with some very elementary facts.

1.5.1. Let A and B be unital C^*-algebras, and $\phi : A \to B$, a unital homomorphism. If $P \in B$ is a projection which commutes with the image of ϕ, i.e., such that

$$P\phi(a) - \phi(a)P = 0, \ \forall a \in A,$$

then $\phi(a)$ can be decomposed into two mutually orthogonal parts $\phi(a)P = P\phi(a)P$ and $\phi(a)(\mathbf{1} - P) = (\mathbf{1} - P)\phi(a)(\mathbf{1} - P)$:

$$\phi(a) = P\phi(a)P + (\mathbf{1} - P)\phi(a)(\mathbf{1} - P).$$

1.5.2. In 1.5.1, let us consider the case that $A = C(X)$. Let $F \subset C(X)$ be a finite set. Let unital homomorphism $\phi : C(X) \to B$ and projection $P \in B$ be as in 1.5.1. Furthermore, suppose that there is a point $x_0 \in X$ such that $P\phi(f)P = \phi(f)P$ is approximately equal to $f(x_0)P$ to within ε on F:

$$\|\phi(f)P - f(x_0)P\| < \varepsilon, \quad \forall f \in F.$$

Then

$$\|\phi(f) - (\mathbf{1} - P)\phi(f)(\mathbf{1} - P) \oplus f(x_0)P\| < \varepsilon, \quad \forall f \in F.$$

More generally, if there are mutually orthogonal projections $P_1, P_2, \cdots, P_n \in B$, which commute with $\phi(C(X))$, and points $x_1, x_2, \cdots, x_n \in X$ such that

$$(*) \qquad \|\phi(f)P_i - f(x_i)P_i\| < \varepsilon, \quad \forall f \in F,\ i = 1, 2, \cdots n,$$

then

$$(**) \qquad \|\phi(f) - (\mathbf{1} - \sum_{i=1}^{n} P_i)\phi(f)(\mathbf{1} - \sum_{i=1}^{n} P_i) \oplus \sum_{i=1}^{n} f(x_i)P_i\| < \varepsilon,\ \forall f \in F.$$

Here, we used the following fact: the norm of the summation of a set of mutually orthogonal elements in a C^*-algebra is the maximum of the norms of all individual elements in the set. In this paper, this fact will be used many times without saying so.

Example 1.5.3. Let $F \subset C(X)$ be a finite set, and $\varepsilon > 0$. Choose $\eta > 0$ such that if $\mathrm{dist}(x, x') < \eta$, then $|f(x) - f(x')| < \varepsilon$ for any $f \in F$.
Let $x_1, x_2, \cdots, x_n \in X$ be distinct points, and $U_1 \ni x_1, U_2 \ni x_2, \cdots, U_n \ni x_n$ be mutually disjoint open neighborhoods with $U_i \subset B_\eta(x_i)$ $(= \{x \in X \mid \mathrm{dist}(x, x_i) < \eta\})$.
Consider the case that $B = M_\bullet(\mathbb{C})$ and let $\phi : C(X) \to M_\bullet(\mathbb{C})$ be a homomorphism. If $P_i,\ i = 1, 2, \cdots, n$ are the spectral projections corresponding to the open sets U_i (see Definition 1.2.4), then the projections P_i commute with $\phi(C(X))$ and satisfy (*) in 1.5.2. Therefore, the decomposition

$$(**) \qquad \|\phi(f) - (\mathbf{1} - \sum_{i=1}^{n} P_i)\phi(f)(\mathbf{1} - \sum_{i=1}^{n} P_i) \oplus \sum_{i=1}^{n} f(x_i)P_i\| < \varepsilon,$$

holds for all $f \in F$.
We remark that if $\#(\mathrm{SP}\phi \cap U_i)$ (counting multiplicities) is large, then, in the decomposition, $\mathrm{rank}(P_i)(= \#(\mathrm{SP}\phi \cap U_i))$ is large.
In the setting of 1.5.2, not only is (**) true for the original projections $P_1, P_2, \cdots, P_n$, but also it is true for any subprojections $p_1 \le P_1, p_2 \le P_2, \cdots, p_n \le P_n$, with ε replaced by 3ε. Namely, the following lemma holds.

Lemma 1.5.4. *Let X be a compact metrizable space, and write $A = C(X)$. Let $F \subset A$ be a finite set. Let B be a unital C^* -algebra, and $\phi : A \to B$ be a homomorphism. Let $\varepsilon > 0$. Suppose that there are mutually orthogonal projections $P_1, P_2, \cdots, P_n$ in B and points $x_1, x_2, \cdots, x_n$ in X such that $P_i\phi(f) = \phi(f)P_i \quad (i = 1, 2, \cdots, n)$ for any $f \in C(X)$ and such that*

$$(*) \qquad \|\phi(f)P_i - f(x_i)P_i\| < \varepsilon \quad (i = 1, 2, \cdots, n) \quad \text{for any} \;\; f \in F.$$

If $p_1, p_2, \cdots, p_n$ are subprojections of $P_1, P_2, \cdots, P_n$ respectively, then

$$\|\phi(f) - (\mathbf{1} - \sum_{i=1}^{n} p_i)\phi(f)(\mathbf{1} - \sum_{i=1}^{n} p_i) \oplus \sum_{i=1}^{n} f(x_i)p_i\| < 3\varepsilon,$$

for any $f \in F$.
(Notice that the condition that the projections P_i commute with $\phi(f)$ does not by itself imply that the p_i almost commute with $\phi(f)$, but this does follows if (*) holds.)
Different versions of this lemma have appeared in a number of papers (especially, [Cu], [GL], [EGLP]).

Proof: The proof is a straightforward calculation.
One verifies directly that

$$\|(\sum P_i)\phi(f)(\sum P_i) - \sum f(x_i)P_i\| < \varepsilon, \quad \forall f \in F.$$

Hence on multiplying by $\mathbf{1} - \sum p_i$ and $\sum p_i$ (one on each side),

$$\|(\mathbf{1} - \sum p_i)\phi(f)(\sum p_i)\| < \varepsilon, \text{ and } \|(\sum p_i)\phi(f)(\mathbf{1} - \sum p_i)\| < \varepsilon, \; \forall f \in F;$$

on multiplying by $\sum p_i$ on both sides,

$$\|(\sum p_i)\phi(f)(\sum p_i) - \sum f(x_i)p_i\| < \varepsilon, \quad \forall f \in F.$$

The desired conclusion follows from identity

$$\phi(f) = ((\mathbf{1} - \sum p_i) + \sum p_i)\phi(f)((\mathbf{1} - \sum p_i) + \sum p_i).$$

$\square$

Remark 1.5.5. One may wonder why we need the decomposition given in the preceding lemma. In fact, the decomposition (**) of 1.5.2, with the original projections, has a better estimation. Why do we need to use subprojections? The reason is as follows.
Suppose that the C^*-algebra $A = C(X)$, the finite set $F \subset A$, the points $x_1, x_2, \cdots, x_n \in X$, and the open sets $U_1 \ni x_1, U_2 \ni x_2, \cdots, U_n \ni x_n$ are as in 1.5.3. Let us consider the case $B = M_\bullet(C(Y))$ (instead of $M_\bullet(\mathbb{C})$ in 1.5.3), where Y is a simplicial complex.

Let $\phi:\ C(X)\to M_\bullet(C(Y))$ be a unital homomorphism. As in 1.5.3, let $P_i(y)$ denote the spectral projection of $\phi|_y$ corresponding to the open set U_i (see 1.2.8). Then for each $y\in Y$, we have the inequality (**) above,

$$\|\phi(f)(y)-(\mathbf{1}-\sum_{i=1}^{n}P_i(y))\phi(f)(y)(\mathbf{1}-\sum_{i=1}^{n}P_i(y))\oplus\sum_{i=1}^{n}f(x_i)P_i(y)\|<\varepsilon,\ \forall f\in F.$$

Unfortunately, $P_i(y)$ does not in general depend continuously on y, and so this estimation does not give rise to a decomposition for ϕ globally.

On the other hand, one can construct a globally defined continuous projection $p_i(y)$ which is a subprojection of $P_i(y)$ at each point y, and is such that $\mathrm{rank}(p_i)$ is not much smaller than $\min_{y\in Y}\mathrm{rank}(P_i(y))$ (more precisely, $\mathrm{rank}(p_i)\geq\min_{y\in Y}\mathrm{rank}(P_i(y))-\dim(Y)$), by using the continuous selection theorem of [DNNP] as 1.5.6 below.

Once this is done , then for each $y\in Y$, applying the lemma, we have

$$\|\phi(f)(y)-(\mathbf{1}-\sum_{i=1}^{n}p_i(y))\phi(f)(y)(\mathbf{1}-\sum_{i=1}^{n}p_i(y))\oplus\sum_{i=1}^{n}f(x_i)p_i(y)\|<3\varepsilon,\ \forall f\in F.$$

Since the projections $p_i(y)$ depend continuously on y, they define elements $p_i\in B$. We can then rewrite the preceding estimate as

$$\|\phi(f)-(\mathbf{1}-\sum_{i=1}^{n}p_i)\phi(f)(\mathbf{1}-\sum_{i=1}^{n}p_i)\oplus\sum_{i=1}^{n}f(x_i)p_i\|<3\varepsilon,\ \forall f\in F.$$

1.5.6. We would like to discuss how to construct the projections p_i referred to in 1.5.5, using the selection theorem [DNNP 3.2].

To guarantee p_i to have a large rank, we should assume that $P_i(y)$ has a large rank at every point y. So let us assume that for some positive integer k_i and for every point $y\in Y$,

$$\#(\mathrm{SP}\phi_y\cap U_i)\geq k_i,$$

equivalently, $\mathrm{rank}(P_i(y))\geq k_i$.

For the sake of simplicity, let us fix i and write U for U_i ($U\subset X$), P for P_i, k for k_i, and p for the desired projection p_i. So for every point $y\in Y$,

$$\#(\mathrm{SP}\phi_y\cap U)\geq k,$$

equivalently, $\mathrm{rank}P(y)\geq k$. Let us construct a projection $p(y)$, depending continuously on y, such that $\mathrm{rank}p(y)\geq k-\dim(Y)$ and $p(y)\leq P(y)$ for each $y\in Y$.

For each fixed $y_0\in Y$, since $\mathrm{SP}\phi_{y_0}\cap U$ is a finite set, one can choose an open set $U'\subset\overline{U'}\subset U$ such that $\mathrm{SP}\phi_{y_0}\cap U'=\mathrm{SP}\phi_{y_0}\cap U$. In particular, $\mathrm{SP}\phi_{y_0}\cap(\overline{U'}\backslash U')=\emptyset$. By Lemma 1.2.10, there is a connected open set $W\ni y_0$ in Y such that

$$\mathrm{SP}\phi_y\cap(\overline{U'}\backslash U')=\emptyset,\ \forall y\in W.$$

Let $P^W(y)$ be the spectral projection of ϕ_y corresponding to open set U'. By Lemma 1.2.9, this depends continuously on y, and so defines a continuous projection-valued function

$$P^W : W \to \text{ projections of } M_\bullet(\mathbb{C}).$$

Furthermore, $P^W(y) \le P(y)$ for any $y \in Y$ and, for each y in the (connected) subset W,

$$\mathrm{rank}(P^W(y)) = \#(\mathrm{SP}\phi_{y_0} \cap U') = \#(\mathrm{SP}\phi_{y_0} \cap U) \ge k.$$

Once we have the above locally defined continuous projection-valued functions $P^W(y)$, the existence of a globally defined continuous projection-valued function $p(y)$ follows from the following result.

Proposition ([DNNP 3.2]). *Let Y be a simplicial complex, and let k be a positive integer. Suppose that $\mathcal{W}$ is an open covering of Y such that for each $W \in \mathcal{W}$, there is a continuous projection-valued map $P^W : W \to M_\bullet(\mathbb{C})$ satisfying*

$$rank P^W(y) \ge k \qquad \textit{for all } \; y \in W.$$

Then there is a continuous projection-valued map $p : Y \to M_\bullet(\mathbb{C})$ such that for each $y \in Y$,

$$rank \;\; p(y) \ge k - dim(Y), and$$

$$p(y) \le \bigvee \{P^W(y); \quad W \in \mathcal{W}, y \in W\}.$$

Let $p(y)$ be as given in the preceding proposition with respect to P^W as defined above. Then as $P^W(y) \le P(y)$ for each W, $p(y) \le P(y)$ also holds .

Recall, we write U for U_i, P for P_i and p for p_i. So, we obtain a projection p_i such that $p_i(y)$ is a subprojection of $P_i(y)$ for every y. Since $P_i(y),\ i = 1, 2, \cdots, n$, are the spectral projections corresponding to $U_i,\ i = 1, 2, \cdots n$, which are mutually disjoint, the projections $P_i(y),\ i = 1, 2, \cdots, n$, are mutually orthogonal, and so are the projections $p_i,\ i = 1, 2, \cdots, n$. Combining this construction with Lemma 1.5.4, we have the following result.

Proposition 1.5.7. *Let X be a simplicial complex, and $F \subset C(X)$ a finite subset. Suppose that $\varepsilon > 0$ and $\eta > 0$ are as in 1.5.3, i.e., such that if $dist(x, x') < \eta$, then $|f(x) - f(x')| < \varepsilon$ for any $f \in F$.*

Suppose that $U_1, U_2, \cdots, U_n$ are disjoint open neighborhoods of (distinct) points $x_1, x_2, \cdots, x_n \in X$, respectively, such that $U_i \subset B_\eta(x_i)$ for all $1 \le i \le n$. Suppose that $\phi : C(X) \to M_\bullet(C(Y))$ is a unital homomorphism, where Y is a simplicial complex, such that

$$\#(SP\phi_y \cap U_i) \ge k_i \qquad \textit{for } 1 \le i \le n, \ \textit{and for all } y \in Y.$$

Then there are mutually orthogonal projections $p_1, p_2, \cdots, p_n \in M_\bullet(C(Y))$ *with* $rank(p_i) \geq k_i - dim(Y)$ *such that*

$$\|\phi(f) - p_0\phi(f)p_0 \oplus \sum_{i=1}^{n} f(x_i)p_i\| < 3\varepsilon \qquad \text{for all } f \in F,$$

where $p_0 = \mathbf{1} - \sum p_i$. *Consequently,*

$$rank(p_0) \leq (\#(SP\phi_y) - \sum_{i=1}^{n} k_i) + n \cdot dim(Y).$$

(Note that $\#(SP\phi_y)$ *is the order of the matrix algebra* $M_\bullet(C(Y))$*.)*
(In fact, the above is also true if one replaces $M_\bullet(C(Y))$ by $PM_\bullet(C(Y))P$, with the exact same proof.)

1.5.8. Proposition 1.5.7 is implicitly contained in the proof of the main decomposition theorem—Theorem 2.21 of [EG2], and explicitly stated as Theorem 2.3 of [Li3], for the case of dimension one.
To use 1.5.7 to decompose a partial map $\phi^{i,j}_{m,m'} : M_{[m,i]}(C(X_{m,i})) \rightarrow M_{[m',j]}(C(X_{m',j}))$ of the connecting homomorphism $\phi_{m,m'} : A_m \rightarrow A_{m'}$ in the inductive system $(A_m, \phi_{m,m'})$, we only need to write

$$\phi^{i,j}_{m,m'} = \phi \otimes \mathrm{id}_{[m,i]},$$

(see 1.2.9), and then decompose ϕ (cf. 1.2.19). In [EG2], we proved that such a map ϕ (for m' large enough) satisfies the condition in Proposition 1.5.7, if the inductive limit is of real rank zero. More precisely, we constructed mutually disjoint open sets $U_1, U_2, \cdots, U_n$, with small diameter, such that $\sum_{i=1}^n k_i$ is very large compared with $(\#(\mathrm{SP}\phi_y) - \sum_{i=1}^n k_i)$, where $k_i = \min_{y\in Y} \#(\mathrm{SP}\phi_y \cap U_i)$. (See the open sets W_i in the proof of Theorem 2.21 of [EG2].) Therefore, in the above decomposition, the part $\sum_{i=1}^n f(x_i)p_i$, which has rank at least $(\sum_{i=1}^n k_i) - n \cdot \dim(Y)$, has much larger size than the size of the part $p_0\phi(f)p_0$, which has rank at most $(\#(\mathrm{SP}\phi_y) - \sum_{i=1}^n k_i) + n \cdot \dim(Y)$, if $n \cdot \dim(Y)$ is very small compared with $\#(\mathrm{SP}\phi_y)$. (Notice that if ϕ is not unital, then $p_0 = \phi(\mathbf{1}) - \sum_{i=1}^n p_i$ and $\#(\mathrm{SP}\phi_y) = \mathrm{rank}\phi(\mathbf{1})$.) (We should mention that $n \cdot \dim(Y)$ is automatically small from the construction. This is a kind of technical detail, to which the reader should not pay much attention now. The number n depends only on η above, but $\#(\mathrm{SP}\phi_y)$ could be very large as m' (for $\phi_{m,m'}$) is large. In particular, it could be much larger than $\dim(Y)$ (note that $Y = X_{m',j}$), if the inductive limit has slow dimension growth.)
The above construction is not trivial. It depends heavily on the real rank zero property and Su's result concerning spectral variation (see [Su]).
What was proved by this construction in [EG2] is the decomposition theorem for the real rank zero case, as mentioned in the introduction.

1.5.9. For the case of a non real rank zero inductive system, we can not construct the mutually disjoint open set $\{U_i\}$ as described in 1.5.8. Notice that in the decomposition described in 1.5.8, the major part ψ defined by $\psi(f) := \sum_{i=1}^n f(x_i)p_i$ has property that

$$\mathrm{SP}\psi_y = \{x_1^{\sim \mathrm{rank}(p_1)}, x_2^{\sim \mathrm{rank}(p_2)}, \cdots, x_n^{\sim \mathrm{rank}(p_n)}\}.$$

That is, the spectrum consists of several fixed points $\{x_i\}_{i=1}^n (\subset X)$ with multiplicities. This kind of decomposition depends on the real rank zero property. For the decomposition of the simple inductive limit algebras, we are forced to allow the major part to have variable spectrum— $\mathrm{SP}\psi_y$ varies when y varies. The following results can be proved exactly the same as the way Proposition 1.5.7 is proved (see 1.5.6 and 1.5.4). (Proposition 1.5.7 is a special case of the following result by taking $\alpha_i(y) = x_i$, the constant maps, and $U_i(y) = U_i \ni x_i$, the fixed open sets.)

Proposition 1.5.7'. *Let X be a simplicial complex, and $F \subset C(X)$, a finite subset. Suppose that $\varepsilon > 0$ and $\eta > 0$ are as in 1.5.3, i.e., such that if dist$(x, x') < \eta$, then $|f(x) - f(x')| < \varepsilon$ for any $f \in F$.*
Suppose that $\alpha_1, \alpha_2, \cdots \alpha_n : Y \to X$ are continuous maps from a simplicial complex Y to X. Suppose that $U_1(y), U_2(y), \cdots, U_n(y)$ are mutually disjoint open sets satisfying $U_i \subset B_\eta(\alpha_i(y))$ and satisfying the following continuity condition:
For any $y_0 \in Y$ and closed set $F \subset U_i(y_0)$, there is an open set $W \ni y_0$ such that $F \subset U_i(y)$ for any $y \in W$.
Suppose that $\phi : C(X) \to M_\bullet(C(Y))$ is a unital homomorphism such that

$$\#(SP\phi_y \bigcap U_i(y)) \geq k_i, \qquad \forall i \in \{1, 2, \cdots, n\},\ y \in Y.$$

Then there are mutually orthogonal projections $p_1, p_2, \cdots, p_n \in M_\bullet(C(Y))$ with rank$(p_i) \geq k_i - dim(Y)$ such that

$$\|\phi(f)(y) - p_0(y)\phi(f)(y)p_0(y) \oplus \sum_{i=1}^n f(\alpha_i(y))p_i(y)\| < 3\varepsilon, \qquad \forall f \in F,$$

where $p_0 = \mathbf{1} - \sum p_i$. Consequently,

$$\mathit{rank}(p_0) \leq (\#(SP\phi_y) - \sum_{i=1}^n k_i) + n\ \mathit{dim}(Y).$$

(It is easy to see that the proof of Proposition 1.5.7 (see 1.5.6) can be generalized to this case. Notice that the above continuity condition for $U_i(y)$ assures that $\overline{U'} \subset U_i(y)$ for any $y \in W$, where U' is described in 1.5.6 corresponding to y_0 and $U = U_i(y_0)$. Then it will assure that $P^W(y) \leq P(y)$, where $P^W(y)$ is the locally defined continuous projection-valued function described in 1.5.6.)

1.5.10. In the above proposition, if one can choose the maps $\alpha_i : Y \to X$ factoring through the interval $[0,1]$—the best possible 1-dimensional space—as

$$\alpha_i : Y \xrightarrow{a_i} [0,1] \xrightarrow{b} X,$$

then the part ψ, defined by $\psi(f)(y) = \sum_{i=1}^{n} f(\alpha_i(y))p_i(y)$, factors through $C[0,1]$ as

$$C(X) \xrightarrow{b^*} C([0,1]) \xrightarrow{\psi'} (\bigoplus_{i=1}^{n} p_i)M_\bullet(C(Y))(\bigoplus_{i=1}^{n} p_i),$$

where b^* is induced by $b : [0,1] \to X$ and ψ' is defined by

$$\psi'(f)(y) = \sum_{i=1}^{n} f(a_i(y))p_i(y).$$

In particular, if

$$\text{(1)} \qquad \sum_{i=1}^{n} k_i >> (\#(\mathrm{SP}\phi_y) - \sum_{i=1}^{n} k_i),$$

where $k_i = \min_{y\in Y} \#(\mathrm{SP}\phi_y \cap U_i)$, then we obtain a decomposition with major part factoring through the interval algebras—direct sum of matrix algebras over interval.

1.5.11. The ideal approach for obtaining a decomposition of $\phi_{m,m'}$ with major part factoring through an interval algebra, is to reduce it to the setting of Proposition 1.5.7', that is, to construct continuous maps $\{\alpha_i\}$ (factoring through interval) and the mutually disjoint open sets $\{U_i\}$ as described in 1.5.7', such that (1) in 1.5.10 holds for homomorphism ϕ induced from the partial connecting homomorphism $\phi_{m,m'}^{i,j}$ described in 1.5.8.
Unfortunately, it seems impossible to realize such a construction globally.
A consequence of the property of α_i described in Proposition 1.5.7', is the following property of α_i, called property (Pairing):
Property (Pairing): For each $y \in Y$, there is a subset of $\mathrm{SP}\phi_y$, which can be paired with

$$\{\alpha_1(y)^{\sim k_1}, \alpha_2(y)^{\sim k_2}, \cdots, \alpha_n(y)^{\sim k_n}\}$$

to within η, counting multiplicities. (See 1.1.7 (b) for the notation $x^{\sim k}$.)
Even though one can not construct the continuous maps $\{\alpha_i\}$ (factoring through interval $[0,1]$ and open sets $\{U_i\}$ to satisfy the conditions in Proposition 1.5.7' together with the condition (1) in 1.5.10, for the connecting homomorphisms in the simple inductive limit, Li constructed the maps $\{\alpha_i\}$ to satisfy the above weaker property (Pairing) and (1) in 1.5.10.
In fact, Li proves the following lemma.

Lemma. *Suppose that* $\lim(A_m, \phi_{m,m'})$ *is a simple AH-inductive limit with slow dimension growth and with injective connecting homomorphisms. For any* $\eta > 0$*, and* A_m*, there are a* $\delta > 0$ *and an integer* $N > m$*, such that for any* $m' > N$*,* $SP(\phi^{i,j}_{m,m'})_y$ *can be paired with*

$$\Theta(y) = \{\alpha_1(y)^{\sim T_1}, \alpha_2(y)^{\sim T_2}, \cdots, \alpha_L(y)^{\sim T_L}\}$$

to within η*, counting multiplicities, for certain continuous maps* $\alpha_i : Y \to X$ *factoring through [0,1], where* $X = X_{m,i}$ *and* $Y = X_{m',j}$*. Furthermore, if we denote* $\frac{rank(\phi^{i,j}_{m,m'}(\mathbf{1}_{A^i_m}))}{rank(\mathbf{1}_{A^i_m})}$ *by* K*, then* $T_i \geq int(\delta K)$ *and therefore* $L < \frac{2}{\delta}$*, provided* $\delta K \geq 2$*, where* $int(\delta K)$ *is the integer part of* δK *as in 1.1.7.*
(This lemma is Theorem 2.19 (see also Remark 2.21) of [Li2]. The additional part about the size of T_i could be obtained by inspecting the proof of the theorem (see 2.16 and 2.18 of [Li2]).)
(Note that when we write $\phi^{i,j}_{m,m'} = \phi \otimes \mathrm{id}_{[m,i]}$, we have $\mathrm{SP}(\phi^{i,j}_{m,m'})_y = \mathrm{SP}\phi_y$.)
In particular, the condition (1) in 1.5.10 holds, since the right hand side is zero. Therefore the following theorem will be useful for our setting, which is the first theorem in §4.

Theorem 4.1. *Let* X *be a connected finite simplicial complex, and* $F \subset C(X)$ *be a finite set which generates* $C(X)$*. For any* $\varepsilon > 0$*, there is an* $\eta > 0$ *such that the following statement is true.*
Suppose that a unital homomorphism $\phi : C(X) \to PM_\bullet(C(Y))P$ *(*$rank(P) = K$*) (where* Y *is a finite simplicial complex) satisfies the following condition: There are* L *continuous maps*

$$a_1, a_2, \cdots, a_L : Y \longrightarrow X$$

such that for each $y \in Y$*,* $SP\phi_y$ *and* $\Theta(y)$ *can be paired within* η*, where*

$$\Theta(y) = \{a_1(y)^{\sim T_1}, a_2(y)^{\sim T_2}, \cdots, a_L(y)^{\sim T_L}\}$$

and $T_1, T_2, \cdots, T_L$ *are positive integers with*

$$T_1 + T_2 + \cdots + T_L = K = rank(P).$$

Let $T = 2^L(\dim X + \dim Y)^3$*. It follows that there are* L *mutually orthogonal projections* $p_1, p_2, \cdots, p_L \in PM_\bullet(C(Y))P$ *such that*
(i) $\|\phi(f)(y) - p_0(y)\phi(f)(y)p_0(y) \oplus \sum_{i=1}^L f(a_i(y))p_i(y)\| < \varepsilon$ *for any* $f \in F$ *and* $y \in Y$*, where* $p_0 = P - \sum_{i=1}^L p_i$*;*
(ii) $\|p_0(y)\phi(f)(y) - \phi(f)(y)p_0(y)\| < \varepsilon$ *for any* $f \in F$ *and* $y \in Y$*;*
(iii) $rank(p_i) \geq T_i - T$ *for* $1 \leq i \leq L$*, and hence* $rank(p_0) \leq LT$*.*
(In the above, η can be chosen to be any number satisfying that if $\mathrm{dist}(x, x') < 2\eta$, then $|f(x) - f(x')| < \frac{\varepsilon}{3}$, $\forall f \in F$.)
(Note that we can not make the number T in the above theorem as small as $\dim(Y)$, as in Proposition 1.5.7 or 1.5.7', for some technical difficulties. This

is the reason that we are forced to use the stronger condition of very slow dimension growth instead of slow dimension growth in our main decomposition theorem.)

1.5.12. The proof of the above theorem is much more difficult than that of Proposition 1.5.7 or 1.5.7'. In particular, Theorem 2.1 (see §1.2 above), and results in §3 are only for the purpose of proving the above theorem. With these results in hand, the proof of Theorem 4.1 will be given in 4.2–4.19. Then the main decomposition theorem described in the introduction—Theorems 4.35 and 4.37—will be proved based on Theorem 4.1 and the results from [Li2].
We would like to explain the difficulties, and how Theorem 2.1 and §3 will be used to over come the difficulties.
Now, our notations are as in Theorem 4.1 above.
Fix i. Let $U_i(y) = \{x \in X \mid \mathrm{dist}(x, a_i(y)) < \eta\}$, then from the condition of Theorem 4.1, we have

$$\#(\mathrm{SP}\phi_y \cap U_i(y)) \geq T_i.$$

Let $P_i(y)$ be the spectral projection corresponding to the open set $U_i(y)$. This is not a continuously defined projection. But using the same procedure in 1.5.6 (see 1.5.7 and 1.5.7'), one can construct a globally defined projection $p_i(y)$ such that $p_i(y) \leq P_i(y)$, and $\mathrm{rank}(p_i(y)) \geq T_i - \dim(Y)$.
But unfortunately, those $p_i(y)$ are not mutually orthogonal, since $U_i(y)$ are not mutually disjoint, and therefore $P_i(y)$ are not mutually orthogonal.

1.5.13. If we assume that the maximum spectral multiplicity of ϕ is at most Ω, then for each $y \in Y$, we can divide the set $\mathrm{SP}\phi_y$ (with multiplicity) into L mutually disjoint subsets $E_1, E_2, \cdots, E_L$ such that, for each $\lambda \in E_i$, $\mathrm{dist}(\lambda, a_i(y)) \leq \eta,\ i = 1, 2, \cdots, L$, and such that

$$T_i - \Omega < \#(E_i) < T_i + \Omega,\ i = 1, 2, \cdots, L,$$

counting multiplicity. By $\{E_i\}$ being mutually disjoint, we mean that if an element $\lambda \in \mathrm{SP}\phi_y$ has multiplicity k, then we put the entire k copies of λ into one of E_i, without separating them. (In the above, $\phi,\ a_i,\ T_i$, and L are all from Theorem 4.1.)
(Note that if we require that $\#(E_i) = T_i$, then we can not guarantee $\{E_i\}$ are mutually disjoint, because of spectral multiplicity.)
Then we can construct mutually disjoint open sets $U_1(y), U_2(y), \cdots, U_L(y)$ such that $E_i \subset U_i(y)$ and $U_i(y) \subset B_\eta(a_i(y))$. We can further assume that these open sets have mutually disjoint closures. That is $\overline{U_i(y)} \cap \overline{U_j(y)} = \emptyset$, for $i \neq j,\ i, j \in \{1, 2, \cdots, L\}$.
(The open sets from such construction usually do not satisfy the continuity condition in Proposition 1.5.7', so we can not apply Proposition 1.5.7'. We need to check the proof of it (e.g. the argument in 1.5.6) against our new setting.)

For each $y_0 \in Y$, there is an open set $W(y_0) \ni y_0$ such that

$$\mathrm{SP}\phi_y \subset U_1(y_0) \cup U_2(y_0) \cup \cdots \cup U_L(y_0), \ \forall y \in W(y_0).$$

As in 1.5.6, one can construct the mutually orthogonal locally defined continuous projection-valued functions

$$P_i^{W(y_0)}: \ W(y_0) \to \text{projection of } M_\bullet(\mathbb{C}), \ i = 1, 2, \cdots, L,$$

where $P_i^{W(y_0)}(y)$ $(y \in W(y_0))$ are the spectral projections of ϕ_y corresponding to open sets $U_i(y_0)$ (or $\mathrm{SP}\phi_y \cap U_i(y_0)$). Furthermore, $\mathrm{rank}P_i^{W(y_0)} = \#(E_i) > T_i - \Omega$.
(Note that we do not need to introduce the smaller open set U' as in 1.5.6, because it is automatically true that $\mathrm{SP}\phi_y \cap (\overline{U_i(y_0)} \backslash U_i(y_0)) = \emptyset$, as $\{\overline{U_i(y_0)}\}_{i=1}^{L}$ are mutually disjoint and $\mathrm{SP}\phi_y \subset U_1(y_0) \cup U_2(y_0) \cup \cdots \cup U_L(y_0)$.)
Theorem 2.1 guarantees that Ω is controlled, it is at most $\dim(X) + \dim(Y)$, which will be very small compared with T_i, in our future application.

1.5.14. There is a finite subcover $\mathcal{W} = \{W(y_j)\}_j$ of the open cover $\{W(y)\}_{y \in Y}$ of Y.
We can use the selection theorem [DNNP 3.2] (see 1.5.6 above) to construct global defined continuous projection valued functions $p_i(y)$, $i = 1, 2, \cdots, L$, of ranks at least $T_i - \Omega - \dim(Y)$, such that

$$(*) \qquad p_i(y) \leq \bigvee \{P_i^{W(y_j)}(y) \mid y \in W(y_j) \text{ and } W(y_j) \in \mathcal{W}\}.$$

For any $i_1 \neq i_2 \in \{1, 2, \cdots, L\}$, $W(y_j) \in \mathcal{W}$, and $y \in W(y_j)$, we have $P_{i_1}^{W(y_j)}(y) \perp P_{i_2}^{W(y_j)}(y)$.
Unfortunately, when $y \in W(y_{j_1}) \cap W(y_{j_2})$, we do not have

$$P_{i_1}^{W(y_{j_1})}(y) \perp P_{i_2}^{W(y_{j_2})}(y).$$

Therefore, one can not conclude that $p_{i_1}(y) \perp p_{i_2}(y)$ from the above (*).
(Notice that, in the above, $P_i^{W(y_0)}(y)$ is the spectral projection of ϕ_y with respect to the open set $U_i(y_0)$ (not $U_i(y)$), and in general, $U_{i_1}(y_{j_1}) \cap U_{i_2}(y_{j_2}) \neq \emptyset$ if $j_1 \neq j_2$. In Propositions 1.5.7 and 1.5.7', we do not have such problem, since $P_i^W(y)$ is the spectral projection of U', which is an open subset of U_i (does not depend on y) in the case of Proposition 1.5.7, or which is an open subset of $U_i(y_0) \cap U_i(y)$ in the case of Proposition 1.5.7'; see 1.5.6 and the explanation after Proposition 1.5.7' for more details.)

1.5.15. For any $W \in \mathcal{W}$ and $y \in W$, define $Q_i^W(y)$, $i = 1, 2, \cdots, L$, to be the spectral projections of ϕ at point y with respect to the open sets $\bigcap_{\{j:\ W(y_j) \cap W \neq \emptyset\}} U_i(y_j), i = 1, 2, \cdots, L$. These are subprojections of $P_i^W(y)$.

The advantage of using these projections is the following fact. For any $y \in W(y_{j_1}) \cap W(y_{j_2})$, we DO have

$$Q_{i_1}^{W(y_{j_1})}(y) \perp Q_{i_2}^{W(y_{j_2})}(y)$$

for any $i_1 \neq i_2 \in \{1, 2, \cdots, L\}$, because $U_{i_1}(y_{j_1}) \cap U_{i_2}(y_{j_1}) = \emptyset$, $\bigcap_{\{j:\ W(y_j)\cap W(y_{j_1})\neq\emptyset\}} U_{i_1}(y_j) \subset U_{i_1}(y_{j_1})$ and $\bigcap_{\{j:\ W(y_j)\cap W(y_{j_2})\neq\emptyset\}} U_{i_2}(y_j) \subset U_{i_2}(y_{j_1})$ (the second inclusion follows from $W(y_{j_1}) \cap W(y_{j_2}) \neq \emptyset$).
Then we apply the selection theorem to Q_i^W (instead of P_i^W) to find globally defined continuous projection-valued functions $p_i(y)$ such that

$$p_i(y) \leq \bigvee \{Q_i^{W(y_j)}(y) \mid y \in W(y_j) \text{ and } W(y_j) \in \mathcal{W}\},$$

and such that

$$\operatorname{rank}(p_i(y)) \geq \min_{y\in W\in\mathcal{W}} \{\operatorname{rank}(Q_i^W(y))\} - \dim(Y).$$

(The readers may notice that $Q_i^W(y)$ are not continuous on W, so one can not apply the selection theorem directly. But one can introduce the open subsets U' as in 1.5.6 for $\bigcap_{\{j:\ W(y_j)\cap W\neq\emptyset\}} U_i(y_j)$ (instead of U_i). We omit the details.)
This time, $p_{i_1}(y) \perp p_{i_2}(y)$ for any $i_1 \neq i_2 \in \{1, 2, \cdots, L\}$.
To guarantee $\operatorname{rank}(p_i(y))$ to be large—not too much smaller than T_i, $\min_{y\in W\in\mathcal{W}}\{\operatorname{rank}(Q_i^W(y))\}$ must be large.

1.5.16. Fixed $y_0 \in Y$ with $W(y_0) \in \mathcal{W}$. Recall from the definitions of $U_i(y_0)$ and $W(y_0)$ (see 1.5.13),

$$\mathrm{SP}\phi_y = \big(\mathrm{SP}\phi_y \cap U_1(y_0)\big) \cup \big(\mathrm{SP}\phi_y \cap U_2(y_0)\big) \cup \cdots \cup \big(\mathrm{SP}\phi_y \cap U_L(y_0)\big), \forall y \in W(y_0).$$

Define $E_i^{W(y_0)}(y) := \mathrm{SP}\phi_y \cap U_i(y_0)$. ($E_i^{W(y_0)}(y_0)$ is the set E_i in 1.5.13, with y_0 in place of y.) Then for each $y \in W \in \mathcal{W}$, $\{E_i^W(y)\}_{i=1}^L$ is a division of $\mathrm{SP}\phi_y$ (in the terminology in §3, it will be called a grouping of $\mathrm{SP}\phi_y$).
From 1.5.15,

$$\begin{aligned}\operatorname{rank}(Q_i^W(y)) &= \#\left(\mathrm{SP}\phi_y \cap \bigcap_{\{j:\ W(y_j)\cap W\neq\emptyset\}} U_i(y_j)\right)\\ &= \#\Big(\bigcap_{\{j:\ W(y_j)\cap W\neq\emptyset\}} E_i^{W(y_j)}(y)\Big).\end{aligned}$$

Roughly speaking, for our construction in 1.5.15 to work, we need the following condition.

Condition: For each $y \in Y$, the number $\#(\bigcap_{\{W:\ y\in W\in\mathcal{W}\}} E_i^W(y))$ is large—not too much smaller than T_i.

(This condition is a little weaker than $Q_i^W(y)$ to be large. But we are going to use some special open cover so that the above weaker condition will be enough. We are not going to discuss details here, and the reader does not need to pay much attention.)
But in 1.5.15, we only have $\#(E_i^W(y)) > T_i - \Omega$. To obtain the above condition, we need the combinatorial results in §3. We are going to discuss it now.

1.5.17. For an intersection to be large, it will be certainly natural to require the number of sets involved in the intersection be as small as possible. As in the setting of 1.5.16, we should require that for any $y \in Y$, the number of sets in $\mathcal{W}$ which cover y—$\#\{W \mid y \in W \in \mathcal{W}\}$— is not too large.
From the definition of covering dimension, we know that for any n-dimensional compact metrizable space Y and any finite cover $\mathcal{U}$ of Y, there is a refined cover $\mathcal{U}_1$ of $\mathcal{U}$ such that for any point $y \in Y$, there are at most $n+1$ open sets in $\mathcal{U}_1$ to cover the point y. In particular, for a simplicial complex Y, the construct of such open cover is given in 1.4.2 (a).
Let $\{W_y\}_{y\in \text{Vertex}(Y)}$ be the open cover of Y given in 1.4.2(a). Recall that, for any vertices $y_0, y_1, \cdots, y_k$, the intersection $\bigcap_{i=1}^k W_{y_i}$ is nonempty if and only if $y_0, y_1, \cdots, y_k$ share one simplex.
For any finite open cover, there is a finite open cover of the above form (for some refined simplicial complex structure), refining the given open cover.
Without loss of generality, we can assume the open cover $\mathcal{W} = \{W(y_j)\}$ in 1.5.14 is of the above form. Hence $\{y_j\}$ are vertices of the simplicial complex Y and $W(y_j)$ are the open sets W_{y_j} defined above. Then the condition in 1.5.16 becomes the following.
For any simplex Δ of Y with vertices $y_0, y_1, \cdots, y_k$,

$$\#\left(E_i^{W(y_0)}(y) \cap E_i^{W(y_1)}(y) \cap \cdots \cap E_i^{W(y_k)}(y)\right) \geq T_i - C$$

for any $y \in W(y_0) \cap W(y_1) \cap \cdots \cap W(y_k)$, where C is not too large. (In the proof of Theorem 4.1, the number C will be chosen to be $2^L\Omega(1+\dim(Y)(\dim(Y)+1))$, where Ω is the maximum spectral multiplicity which is bounded by $\dim(X)+\dim(Y)$, by Theorem 2.1.)

1.5.18. To make the discussion simpler, we suppose that the homomorphism ϕ has distinct spectrum at any point $y \in Y$. That is, the maximum spectral multiplicity of ϕ is one. (Of course, in the proof of Theorem 4.1 in §4, we will not make this assumption.)
If the simplicial structure is sufficiently refined, by the distinct property of the spectrum, we can assume the following holds: For any simplex Δ with $Z = \bigcup_{y\in \text{Vertex}(\Delta)} W(y)(\supset \Delta)$, there are continuous maps

$$\lambda_1, \lambda_2, \cdots, \lambda_K : \ Z \to X,$$

where $K = \text{rank}(P)$ as in Theorem 4.1, such that

$$\text{SP}\phi_y = \{\lambda_1(y), \lambda_2(y), \cdots, \lambda_K(y)\}, \ \forall y \in Y.$$

Then for any $y \in \mathrm{Vertex}(\Delta)$, the division (or grouping) $E_i^{W(y)}(y)$ of $\mathrm{SP}\phi_y$ gives rise to a division (or grouping) $E_i(y)$ of $\{\lambda_1, \lambda_2, \cdots, \lambda_K\}$. In the case of distinct spectrum, the condition in 1.5.13 concerning $\#(E_i)$ is $\#(E_i(y)) = T_i$. The condition at the end of 1.5.17 is

$$\#\left(\bigcap_{y \in \mathrm{Vertex}(\Delta)} E_i(y)\right) \geq T_i - C.$$

This is of course not true in general, unless we make some special arrangement. But with the following lemma, we can always subdivide (or refine) the simplicial complex and introduce the groupings of $\{\lambda_1, \lambda_2, \cdots, \lambda_K\}$ for all newly introduced vertices, to make the above true for any simplex of new simplicial structure (after subdivision).

The following formal definition of grouping is in 3.2 of §3.

Definition. Let $E = \{1, 2, \cdots, K\}$ be an index set. Let $T_1, T_2, \cdots, T_L$ be non negative integers with

$$T_1 + T_2 + \cdots + T_L = K.$$

A grouping of E of type $(T_1, T_2, \cdots, T_L)$ is a set of L mutually disjoint index sets $E_1, E_2, \cdots, E_L$ with

$$E = E_1 \cup E_2 \cup \cdots \cup E_L,$$

and $\#(E_j) = T_j$ for each $1 \leq j \leq L$.

Lemma. *Suppose that (Δ, σ) is a simplex, where σ is the standard simplicial structure of the simplex Δ. Suppose that for each vertex $x \in$ Vertex(Δ, σ), there is a grouping $E_1(x), E_2(x), \cdots, E_L(x)$ of E of type $(T_1, T_2, \cdots, T_L)$.*

It follows that there is a subdivision (Δ, τ) of (Δ, σ), and there is an extension of the definition of the groupings of E for Vertex(Δ, σ) to the groupings of E (of type $(T_1, T_2, \cdots, T_L)$) for Vertex(Δ, τ) ($\supset$ Vertex(Δ, σ)) such that:

(1) For each newly introduced vertex $x \in$ Vertex(X, τ),

$$E_j(x) \subset \bigcup_{y \in \, Vertex(\Delta,\sigma)} E_j(y), \qquad j = 1, 2, \cdots, L.$$

(2) For any simplex Δ_1 of (X, τ) (after subdivision),

$$\#(\bigcap_{x \in \, Vertex(\Delta_1)} E_j(x)) \geq T_j - \frac{\dim(\Delta)(\dim(\Delta)+1)}{2}, \qquad j = 1, 2, \cdots, L.$$

1.5.19. Condition (1) above is important for the following reason. In 1.5.13, when we define E_i as a subset of $\mathrm{SP}\phi_y$, we require that $\mathrm{dist}(\lambda, a_i(y)) < \eta$ for any $\lambda \in E_i$. This condition guarantees that the projection P_i^W in 1.5.13 (or Q_i^W in 1.5.16) satisfies that $\phi(f)(y)P_i^W(y)$ (or $\phi(f)(y)Q_i^W(y)$) is approximately equal to $f(a_i(y))P_i^W(y)$ (or $f(a_i(y))Q_i^W(y)$) to within ε, which is the condition (*) in Lemma 1.5.4.

We consider the grouping of E as the grouping of the spectral functions $\{\lambda_1, \lambda_2, \cdots, \lambda_K\}$ in 1.5.18. Then the condition (1) in the lemma implies the following fact. If for any vertex $y_0 \in \text{Vertex}(\Delta, \sigma)$ and any element $k \in E_i(y_0)$, we have $\text{dist}(\lambda_k(y), a_i(y)) < \eta$, $\forall y \in \Delta$, then for any newly introduced vertex $y_1 \in \text{Vertex}(\Delta, \tau)$ and any element $k' \in E_i(y_1)$ (here $E_i(y_1)$ is the set E_i for the newly introduced grouping for the vertex y_1), we still have $\text{dist}(\lambda_{k'}(y), a_i(y)) < \eta$, $\forall y \in \Delta$.

1.5.20. In fact, in the proof of Theorem 4.1, we need the relative version of the result: Suppose that there are a subdivision $(\partial\Delta, \tau')$ of the boundary $(\partial\Delta, \sigma)$ and groupings for all vertices in $\text{Vertex}(\partial\Delta, \tau')$ ($\supset \text{Vertex}(\Delta, \sigma)$), such that the above (1) holds for any vertex in $\text{Vertex}(\partial\Delta, \tau')$ and such that the above (2) holds for any simplex Δ_1 of $(\partial\Delta, \tau')$ with $\dim(\partial\Delta)$ in place of $\dim(\Delta)$. Then there is a subdivision (Δ, τ) of (Δ, σ), and groupings for all vertices in $\text{Vertex}(\Delta, \tau)$ such that the above (1) and (2) hold and in addition, the following holds: the restriction of (Δ, τ) onto the boundary $\partial\Delta$ is $(\partial\Delta, \tau')$ and the grouping associated to any vertex in $\text{Vertex}(\partial\Delta, \tau)(= \text{Vertex}(\partial\Delta, \tau'))$ is the same as the old one.

In §3, we will prove the above relative version. In fact, to prove the absolute version will automatically force us to prove the stronger one—the relative version.

Another complication comes from the multiplicities, since we can not assume the spectrum to be distinct. In this case, even the definition of grouping needs to be modified.

1.5.21. To give the readers some feeling about the lemma in 1.5.18, we shall discuss the special case of $\dim(\Delta) = 1$. That is, $\Delta = [0, 1]$, the interval.

In this case, we have two groupings for the end points, $\{E_1(0), E_2(0), \cdots E_L(0)\}$ and $\{E_1(1), E_2(1), \cdots E_L(1)\}$ of E of type $(T_1, T_2, \cdots, T_L)$. Then we need to introduce a sequence of points

$$0 = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = 1,$$

(this give rise to a subdivision of $\Delta = [0, 1]$) and define groupings $\{E_1(t_j), E_2(t_j), \cdots E_L(t_j)\}$ for $j = 1, 2, \cdots n-1$ such that conditions (1) and (2) in the lemma holds.

The condition (1) in the lemma means

$$E_i(t_j) \subset E_i(0) \cup E_i(1), \ \forall i \in \{1, 2, \cdots L\}, \ j \in \{1, 2, \cdots, n-1\}.$$

The condition (2) in the lemma means

$$\#(E_i(t_j) \cap E_i(t_{j+1})) \geq T_i - 1,$$

i.e., for any i, and any pair of adjacent points t_j, t_{j+1}, the set $E_i(t_{j+1})$ differs from the set $E_i(t_j)$ by at most one element.

Let us discuss how to make the above (2) hold for E_1. Suppose that $E_1(0)$ and $E_1(1)$ are different (otherwise, we do not need to do anything for them). We can modify $E_1(0)$ to obtain $E_1(t_1)$ as follows. Taking one element λ in $E_1(1)\backslash E_1(0)$ to replace an element μ in $E_1(0)\backslash E_1(1)$, and define it to be $E_1(t_1)$. So $E_1(t_1)$ contains λ but not μ. Since $\lambda \notin E_1(0)$, it must be in some $E_i(0),\ i > 1$. In the set $E_i(0)$, after we take out λ and put it into $E_0(t_1)$, E_i has one element less than it should have, so we can put μ in it, and call it $E_i(t_1)$. For $j \neq 1$ or i, $E_j(t_1)$ should be the same as $E_j(0)$. In such a way, we construct the grouping for t_1, which satisfies the above (2) for the pair $0, t_1$. Furthermore, compare to $\{E_i(0)\}_i$, the new grouping $\{E_i(t_1)\}_i$ is one step closer to the grouping $\{E_i(1)\}_i$. Repeating the above construction (e.g., for t_1 in place of 0) we can construct $E_i(t_2)$ and so on. Finally, we will reach the grouping at the other end point $1 \in [0,1]$.
If one does not require the condition (1), the above is the complete proof of the lemma for the one-dimensional case.

1.5.22. Since we require the condition (1), when we add an element $\lambda \in E_1(1)\backslash E_1(0)$ into $E_1(0)$ to define $E_1(t_1)$ (as in 1.5.21), we shall carefully choose the element $\mu \in E_1(0)\backslash E_1(1)$ to be replaced by λ. In §3, we shall prove the following assertion: For any $\lambda \in E_1(1)\backslash E_1(0)$, there is $\mu \in E_1(0)\backslash E_1(1)$ to satisfy the following condition: let $F = (E_1(0)\backslash\{\mu\}) \cup \{\lambda\}$; the set $E\backslash F$ can be grouped into $E_2', E_3', \cdots, E_L'$ ($\#(E_i') = T_i$), in such a way that

$$E_i \subset E_i(0) \cup E_i(1).$$

(Such an element μ is the element we should choose.) This is Lemma 3.9 with $E_i(0) \cup E_i(1)$ in place of H_i.
With the above ideas in mind, it should not be difficult (hopefully) to read the first part of §3, which does not involve multiplicity. The main step of §3 is contained in the proof of Lemma 3.11.

1.5.23. In the case with multiplicity, there are two possible ways of proceeding.
1. Define a grouping of

$$E = \{\lambda_1^{\sim w_1}, \lambda_2^{\sim w_2}, \cdots, \lambda_k^{\sim w_k}\}$$

to be a (set theoretical) partition of E as a disjoint union of L sets $E = E_1 \cup E_2 \cup \cdots E_L$. Using this definition, we have to allow that, at different vertices, the groupings may be of different types. That is, $\#(E_i)$ may be different for different vertices. (One can compare with $T_i - \Omega < \#(E_i) < T_i + \Omega$ in 1.5.13.)
2. Define a grouping of

$$E = \{\lambda_1^{\sim w_1}, \lambda_2^{\sim w_2}, \cdots, \lambda_k^{\sim w_k}\}$$

to be a collection of L subsets $E_1, E_2, \cdots, E_L$ with

$$E_j = \{\lambda_1^{\sim p_1^j}, \lambda_2^{\sim p_2^j}, \cdots, \lambda_k^{\sim p_k^j}\},$$

where $0 \leq p_i^j \leq w_j$, such that

$$\sum_{j=1}^{L} p_i^j = w_i, \qquad \text{for each } i = 1, 2, \cdots, k.$$

The grouping is called to be of type $(T_1, T_2, \cdots, T_L)$ if

$$\#(E_j) = \sum_{i=1}^{k} p_i^j = T_j, \qquad \text{for each } j = 1, 2, \cdots, L.$$

In this way, as will be seen, all the groupings corresponding to vertices of a simplicial complex (as in the proof of Theorem 4.1), may be chosen to be of the same type. But the conclusion (2) in the lemma should be modified. Instead of $\#(\bigcap_{x\in \mathrm{Vertex}(\Delta_1)} E_j(x))$, $j = 1, 2, \cdots, L$ to be big, we require $\#(\bigcap_{x\in \mathrm{Vertex}(\Delta_1)} \mathring{E}_j(x))$, $j = 1, 2, \cdots, L$, to be big, where for any set F with multiplicity, $\mathring{F}$ is the subset of F consisting of all such elements λ_i that $\{\lambda_i^{\sim w_i}\}$ are entirely inside F. (See 3.22 for detailed definition of $\mathring{F}$.)
In fact, either approaches can be carried out for our purpose. It turns out that the second approach is shorter and more elegant. Therefore we shall take this approach.
Even though, in our approach, it is allowed to separate some set $\{\lambda_i^{\sim w_i}\}$ into different groups E_i of the grouping, we should still group as many whole sets $\{\lambda_j^{\sim w_j}\}$ of the index set $\{\lambda_1^{\sim w_1}, \lambda_2^{\sim w_2}, \cdots, \lambda_n^{\sim w_n}\}$ as possible into the same set E_i of $\{E_1, E_2, \cdots, E_L\}$. Assumption 3.27 and Lemma 3.28 are for this purpose. (One needs to pay special attention to the definition and properties of G_I in 3.25.) Except this idea, all other parts of the proof are the same as the case of multiplicity one.

1.5.24. Once we have the combinatorial results in §3, and the explanations in 1.5.1–1.5.19, it will not be hard to understand the proof of Theorem 4.1, though there are some other small techniques, which will be clearly explained in the proof (see 4.2–4.19).

1.5.25. Combining Theorem 4.1 and the result of [Li2] (see the lemma stated in 1.5.11), we can obtain a decomposition $\phi_1 \oplus \psi$ of $\phi_{m,m'}$ (for m' large enough) such that the major part ψ factors through an interval algebra.
But to deal with the part ϕ_1, we should add to it, a relatively large (comparing with ϕ_1) homomorphism ϕ_2, which factors through a finite dimensional C^*-algebra—or which is defined by certain point evaluations (on a δ-dense subset of $X_{n,i}$ for some small number δ).
In [Li3], Li deals with this problem by another decomposition, taking such a homomorphism out of the part ψ. (She only proved the one dimensional case.) We take a different approach. Going back to the construction of the maps α_i in [Li2] (see the lemma inside 1.5.11 above), we can choose sufficiently many

of them to be constant maps (see Lemma 4.33 in §4 below). Therefore ψ automatically has such a part defined by point evaluations.
We believe that our approach is easier to understand than Li's approach, though the spirit is the same. Furthermore, our decomposition is a quantitative version (see Theorem 4.35), and is stronger than Li's theorem even in the case of one dimensional spaces. (This will be important in [EGL].)
We shall not use any result from [Li3]. But we encourage the reader to read the short article [Li2], on which our proof heavily depends.

1.6 Some uniqueness theorems and a factorization theorem

First, this subsection contains some uniqueness theorems. In general, a uniqueness theorem states that, under certain conditions, two maps $\phi, \psi : A \to B$ (homomorphisms or completely positive linear contractions between C^*-algebras A and B) are approximately unitarily equivalent to each other to within a given small number ε on a given finite set $F \subset A$, that is, there is a unitary $u \in B$ such that

$$\|\phi(f) - u\psi(f)u^*\| < \varepsilon, \quad \forall f \in F.$$

This subsection also contains a factorization theorem, which says that, there is a homomorphism (in the class of the so called unital simple embeddings) between matrix algebras over (perhaps higher dimensional) spaces, which must approximately factor through a sum of matrix algebras over the special spaces $\{pt\}, [0,1], S^1, \{T_{II,k}\}_{k=2}^{\infty}, \{T_{III,k}\}_{k=2}^{\infty}$, and S^2, by means of almost multiplicative maps.
We put these two kinds of results together into one subsection, since the proofs of them have some similarity. Also, in the proof of the factorization theorem, we use some uniqueness theorems of this same subsection.
Most of the results are modifications of some results in the literature [EG2], [D1-2], [G4] and [DG] (see [Phi], [GL], [Lin1-2] and [EGLP] also). One of the main theorems (Theorem 1.6.9) is a generalization of Theorem 2.29 of [EG2]—the main uniqueness theorem in the classification of real rank zero AH algebras. The proof given here is shorter than the proof given in [EG2]. Another main theorem—Theorem 1.6.26 (see also Corollary 1.6.29) is a refinement of Lemma 2.2 of [D2] (see also Lemma 3.13 and 3.14 of [G4]). Both Theorem 1.6.9 and Corollary 1.6.29 are important in the proof of our main results in §6.
The following well known result (see [Lo]) will be used frequently.

Lemma 1.6.1. *Suppose that $A = \bigoplus_{i=1}^{t} M_{k_i}(C(X_i))$, where $X_i = \{pt\}, [0,1]$, or S^1. For any finite set $F \subset A$ and any number $\varepsilon > 0$, there is a finite set $G \subset A$ and there is a number $\delta > 0$ such that if C is a C^*-algebra and $\phi \in Map(A, C)$ is a G-δ multiplicative map, then there is a homomorphism $\phi' \in Hom(A, C)$ satisfying*

$$\|\phi(f) - \phi'(f)\| < \varepsilon, \quad \forall f \in F.$$

The following result is essentially contained in [EG2] (see [EG2, 2.11]) and is stated as Theorem 1.2 of [D1] (see also [G4, 3.2 and 3.8]).

Lemma 1.6.2. ([D1, 1.2]) *Let X be a finite simplicial complex. For any finite subset $F \subset C(X)$ and any number $\varepsilon > 0$, there are a positive integer L, a unital homomorphism $\tau : C(X) \to M_L(C(X))$, and a unital homomorphism $\mu : C(X) \to M_{L+1}(C(X))$ with finite dimensional image such that*

$$\|diag(f, \tau(f)) - \mu(f)\| < \varepsilon, \quad \forall f \in F.$$

By the argument in 1.2.19, in the above lemma, the algebra $C(X)$ can be replaced by $M_n(C(X))$.

Lemma 1.6.3. *Let X be a finite simplicial complex and $A = M_n(C(X))$. For any finite subset $F \subset A$ and any number $\varepsilon > 0$, there are a positive integer L, a unital homomorphism $\tau : A \to M_L(A)(= M_{nL}(C(X)))$, and a unital homomorphism $\mu : A \to M_{L+1}(A)$ with finite dimensional image such that*

$$\|diag(f, \tau(f)) - \mu(f)\| < \varepsilon, \quad \forall f \in F.$$

Remark 1.6.4. In general, a unital homomorphism $\lambda : C(X) \to B$ with finite dimensional image is always of the form:

$$\lambda(f) = \sum f(x_i)p_i, \quad \forall f \in F,$$

where $\{x_i\}$ is a finite subset of X, and $\{p_i\} \subset B$ is a set of mutually orthogonal projections with $\sum p_i = \mathbf{1}_B$. A homomorphism $\lambda : A = M_n(C(X)) \to B$ with finite dimensional image is of the form

$$\lambda(f) = \sum p_i \otimes f(x_i), \quad \forall f \in M_n(C(X))$$

for a certain identification of $\lambda(\mathbf{1}_A)B\lambda(\mathbf{1}_A) \cong (\lambda(e_{11})B\lambda(e_{11})) \otimes M_n(\mathbb{C})$, where $\{p_i\}$ is a set of mutually orthogonal projections in $\lambda(e_{11})B\lambda(e_{11})$.
The following Lemma is essentially proved in [D1, Lemma 1.4] (see [G4, Theorem 3.9] also), using the idea from [Phi] and [GL].

Lemma 1.6.5. *Let X be a finite simplicial complex and $A = M_n(C(X))$. For a finite set $F \subset M_n(C(X))$, a positive number $\varepsilon > 0$ and a positive integer N, there are a finite set $G \subset M_n(C(X))$, a positive number $\delta > 0$ and a positive integer L, such that the following is true.*
For any unital C^*-algebra B, any $N+1$ completely positive G-δ multiplicative linear $*$-contraction $\phi_0, \phi_1, \cdots \phi_N \in Map_{G-\delta}(A, B)$, there are a homomorphism $\lambda \in Hom(A, M_L(B))$ with finite dimensional image and a unitary $u \in M_{L+1}(B)$ such that

$$\|diag(\phi_0(f), \lambda(f)) - u\,diag(\phi_N(f), \lambda(f))u^*\| < \varepsilon + \omega, \quad \forall f \in F,$$

where

$$\omega = \max_{f\in F}\max_{0\le j\le N-1}\|\phi_j(f)-\phi_{j+1}(f)\|.$$

Proof: If we allow the number L to depend on the maps $\{\phi_j\}$, then this is Lemma 1.4 of [D1]. In fact, in the proof of [D1, Lemma 1.4], the author proves this stronger version of the lemma. We will not repeat the entire proof in [D1], instead, we will only repeat the construction of G, δ in [D1] and at the same time choose the number L.

Apply Lemma 1.6.3 to $F \subset A$, $\frac{\varepsilon}{4} > 0$ to obtain the integer L_1, $\tau : A \to M_{L_1}(A)$ and $\mu : A \to M_{L_1+1}(A)$ as in Lemma 1.6.3. Then $D := \mu(A)$ is a finite dimensional C^*-subalgebra of $M_{L_1+1}(A)$. By Lemma 1.6.1, there are a finite set $F_1 \subset D(\subset M_{L_1+1}(A))$ and a positive number $\delta_1 > 0$ such that if C is any C^*-algebra and $\psi \in \mathrm{Map}(D,C)$ is any F_1-δ_1 multiplicative map, then there is a homomorphism $\psi' \in \mathrm{Hom}(D,C)$ such that

$$\|\psi'(f)-\psi(f)\| < \frac{\varepsilon}{4}, \quad \forall f \in \mu(F)\ (\subset D).$$

From 1.2.19, there exist a finite set $G \subset A$ and a positive number $\delta > 0$ such that if $\phi \in \mathrm{Map}(A,B)$ is G-δ multiplicative, then $\phi\otimes \mathrm{id}_{L_1+1} \in \mathrm{Map}(M_{L_1+1}(A), M_{L_1+1}(B))$ is F_1-δ_1 multiplicative.

Let $L := N(L_1+1)$. The proof of [D1, Lemma 1.4] proves that such G, δ and L are as desired. (Notice that the size of the homomorphism η on line 9 of page 122 of [D1] is the number L above.)

$\square$

In Lemma 1.6.5, if we further assume that $F \subset A$ is weakly approximately constant to within ε, then one can replace the homomorphism λ in Lemma 1.6.5 by an arbitrary homomorphism with finite dimensional image of sufficiently large size (with ε replaced by 5ε). One can even use two different homomorphisms (provided that the images of the matrix unit e_{11} under these two different homomorphisms are unitarily equivalent) for ϕ_0 and ϕ_N, i.e., instead of $\mathrm{diag}(\phi_0,\lambda)$ and $\mathrm{diag}(\phi_N,\lambda)$, one can use $\mathrm{diag}(\phi_0,\lambda_1)$ and $\mathrm{diag}(\phi_N,\lambda_2)$ in the estimation. Namely, we can prove the following result.

Corollary 1.6.6. *Let X be a finite simplicial complex and $A = M_n(C(X))$. Suppose that $\varepsilon > 0$ and that a finite set $F \subset M_n(C(X))$ is weakly approximately constant to within ε. Suppose that N is a positive integer. Then there are a finite set $G \subset M_n(C(X))$, a positive number $\delta > 0$, and a positive integer L such that the following is true.*

For any unital C^*-algebra B and projection $p \in B$, any $N+1$ completely positive G-δ multiplicative linear $*$-contractions $\phi_0, \phi_1, \cdots, \phi_N \in Map_{G-\delta}(A, pBp)$, any $\lambda^1, \lambda^2 \in Hom(A, (1-p)B(1-p))$ with finite dimensional images and with $\lambda^1(e_{11}) \sim \lambda^2(e_{11})$ (see 1.1.7(i)) and $[\lambda^1(e_{11})] \ge L\cdot[p]$, there is a unitary $u \in B$ such that

$$\|diag(\phi_0(f),\lambda^1(f)) - u\,diag(\phi_N(f),\lambda^2(f))u^*\| < 5\varepsilon+\omega, \quad \forall f \in F,$$

where

$$\omega = \max_{f\in F} \max_{0\le j\le N-1} \|\phi_j(f) - \phi_{j+1}(f)\|.$$

Proof: Suppose that L and $\lambda : A \to M_L(pBp)$ are as in Lemma 1.6.5 for the G-δ multiplicative maps $\phi_0, \phi_1, \cdots \phi_N \in \mathrm{Map}(A, pBp)$.
From 1.6.4, λ is of the following form

$$\lambda(f) = \sum_{i=1}^{s} p_i \otimes f(x_i), \quad \forall f \in M_n(C(X))$$

for a certain identification of $\lambda(\mathbf{1}_A)B\lambda(\mathbf{1}_A) \cong (\lambda(e_{11})B\lambda(e_{11})) \otimes M_n(\mathbb{C})$, where p_i, $i = 1, 2, \cdots, s$, are mutually orthogonal projections with $\sum_{i=1}^{s} p_i = \lambda(e_{11}) \in M_L(pBp)$, and $\{x_i\} \subset X$.
Fix a base point $x_0 \in X$. Since F is weakly approximately constant to within ε, for each i, $\{f(x_i)\}_{f\in F}$ is approximately unitarily equivalent to $\{f(x_0)\}_{f\in F}$ to within ε, one by one by the same unitary. I.e., for each i, there is a unitary $v \in M_n(\mathbb{C})$ such that $\|vf(x_i)v^* - f(x_0)\| < \varepsilon$, for any $f \in F$.
Define $new\lambda$ by $new\lambda(f) = \sum_{i=1}^{s} p_i \otimes f(x_0) = E \otimes f(x_0)$, where $E := \lambda(e_{11}) = \sum_{i=1}^{s} p_i$. Then $new\lambda$ is approximately unitarily equivalent to the old λ to within ε on F. Therefore, with this $new\lambda$, we still have

$$\|\mathrm{diag}(\phi_0(f), \lambda(f)) - u_1 \mathrm{diag}(\phi_N(f), \lambda(f))u_1\| < 3\varepsilon + \omega, \quad \forall f \in F, \tag{1}$$

for some unitary $u_1 \in M_{L+1}(pBp)$.
Since $\lambda^1(e_{11}) \sim \lambda^2(e_{11})$, without loss of generality, we can assume that $\lambda^1|_{M_n(\mathbb{C})} = \lambda^2|_{M_n(\mathbb{C})}$, where $M_n(\mathbb{C}) \subset M_n(C(X))$ $(= A)$. In particular, $\lambda^1(\mathbf{1}_A) = \lambda^2(\mathbf{1}_A)$ and $\lambda^1(e_{11}) = \lambda^2(e_{11})$. Denote $\lambda^1(e_{11})$ by E'. Similar to the case of λ, we can assume that

$$\lambda^1(f) = \sum_{i=1}^{s_1} q_i^1 \otimes f(x_i^1), \quad \forall f \in M_n(C(X)),$$

$$\lambda^2(f) = \sum_{i=1}^{s_2} q_i^2 \otimes f(x_i^2), \quad \forall f \in M_n(C(X))$$

for a certain identification of $\lambda^1(\mathbf{1}_A)B\lambda^1(\mathbf{1}_A) \cong (E'BE') \otimes M_n(\mathbb{C})$, where $\{q_i^1\}$ and $\{q_i^2\}$ are two sets of mutually orthogonal projections with $\sum_{i=1}^{s_1} q_i^1 = \sum_{i=1}^{s_2} q_i^2 = E' \in (1-p)B(1-p)$, and $\{x_i^1\}, \{x_i^2\} \subset X$.
Define $\tilde{\lambda} : A \to \lambda^1(\mathbf{1}_A)B\lambda^1(\mathbf{1}_A)$ by

$$\tilde{\lambda}(f) = E' \otimes f(x_0), \quad \forall f \in F.$$

Similar to the argument for λ, both λ^1 and λ^2 are approximately unitarily equivalent to $\tilde{\lambda}$ to within ε on F.

Since $[E] \leq L \cdot [p] \leq [E']$ $(= [\lambda^1(e_{11})])$, there is a sub-projection $E_1 \leq E'$ which is unitarily equivalent to E.
Write $\tilde{\lambda} = \mu_1 \oplus \mu_2$, where $\mu_1(f) = E_1 \otimes f(x_0)$ and $\mu_2(f) = (E' - E_1) \otimes f(x_0)$. Then μ_1 is unitarily equivalent to λ (strictly speaking, $new\lambda$). From (1), we have

$$\|\mathrm{diag}(\phi_0(f), \mu_1(f)) - u_2 \mathrm{diag}(\phi_N(f), \mu_1(f))u_2^*\| < 3\varepsilon + \omega, \quad \forall f \in F$$

for a unitary $u_2 \in (p \oplus E_1 \otimes \mathbf{1}_n)B(p \oplus E_1 \otimes \mathbf{1}_n)$. Notice that $E_1 \otimes \mathbf{1}_n \leq E' \otimes \mathbf{1}_n = \lambda^1(\mathbf{1}_A) \leq (1-p)$.
Therefore,

$$\|\mathrm{diag}(\phi_0(f), \tilde{\lambda}(f)) - u_3 \mathrm{diag}(\phi_N(f), \tilde{\lambda}(f))u_3^*\| < 3\varepsilon + \omega, \quad \forall f \in F,$$

where $u_3 := u_2 \oplus ((E' - E_1) \otimes \mathbf{1}_n) \in (p \oplus (E' \otimes \mathbf{1}_n))B(p \oplus (E' \otimes \mathbf{1}_n))$.
We already know that both λ^1 and λ^2 are approximately unitarily equivalent to $\tilde{\lambda}$ on F to within ε, so we have

$$\|\mathrm{diag}(\phi_0(f), \lambda^1(f)) - u \mathrm{diag}(\phi_N(f), \lambda^2(f))u^*\| < 5\varepsilon + \omega, \quad \forall f \in F,$$

for a unitary $u \in B$.

$\square$

Lemma 1.6.7. *Suppose that $A = M_k(C(X))$, and $F \subset A$ is weakly approximately constant to within ε. Suppose that A_1 is a C^*-algebra, and two homomorphisms ϕ and $\psi \in Hom(A, A_1)$ are homotopic to each other. There are a finite set $G \subset A_1$, a number $\delta > 0$, and a positive integer $L > 0$ such that the following is true.*
If B is a unital C^*-algebra, $p \in B$ is a projection, $\lambda_0 \in Map(A_1, pBp)$ is G-δ multiplicative, $\lambda_1 \in Hom(A_1, (\mathbf{1}-p)B(\mathbf{1}-p))$ is a homomorphism with finite dimensional image satisfying $[(\lambda_1 \circ \phi)(e_{11})] \geq L \cdot [p]$, and $\lambda \in Map(A_1, B)$ is defined by $\lambda = \lambda_0 \oplus \lambda_1$, then there is a unitary $u \in B$ such that

$$\|(\lambda \circ \phi)(f) - u(\lambda \circ \psi)(f)u^*\| < 6\varepsilon, \quad \forall f \in F.$$

Proof: Since ϕ is homotopic to ψ. There is a continuous path of homomorphisms $\phi_t, 0 \leq t \leq 1$, such that $\phi_0 = \phi$ and $\phi_1 = \psi$. Choose $0 = t_0 < t_1 < \cdots t_{N-1} < t_N = 1$ such that

$$\|\phi_{t_{j+1}}(f) - \phi_{t_j}(f)\| < \varepsilon, \quad \forall j \in \{0, 1, \cdots, N-1\} \text{ and } \forall f \in F.$$

Applying Corollary 1.6.6 to ε, $F \subset A$ (which is weakly approximately constant to within ε), and the number N from the above, there are $G_1 \subset A$ and $\delta > 0$ and L as in the Corollary 1.6.6.
The set $G := \bigcup_{j=0}^N \phi_{t_j}(G_1) \subset A_1$, $\delta > 0$ and number L are as desired.

Suppose that λ_0, λ_1 are the maps satisfying the conditions described in the lemma for G, δ, and L as chosen above. Choosing the sequence of G_1-δ multiplicative maps in Corollary 1.6.6 to be $\lambda_0 \circ \phi_{t_0} (= \lambda_0 \circ \phi), \lambda_0 \circ \phi_{t_1}, \cdots, \lambda_0 \circ \phi_{t_N} (= \lambda_0 \circ \psi)$, and the homomorphisms λ^1 and λ^2 (with finite dimensional images) to be $\lambda^1 = \lambda_1 \circ \phi$ and $\lambda^2 = \lambda_1 \circ \psi$, and using that $\lambda = \lambda_0 \oplus \lambda_1$, we have

$$\|(\lambda \circ \phi)(f) - u(\lambda \circ \psi)(f)u^*\| < 5\varepsilon + \omega, \quad \forall f \in F,$$

for a certain unitary $u \in B$, where

$$\omega = \max_{f \in F} \max_{0 \le j \le N-1} \|(\lambda_0 \circ \phi_{t_{j+1}})(f) - (\lambda_0 \circ \phi_{t_j})(f)\| < \varepsilon,$$

since λ_0 is a contraction—norm decreasing map. So the Lemma follows. (Note that if λ_0 is G-δ multiplicative, then $\lambda_0 \circ \phi_{t_j}$ is G_1-δ multiplicative. Also note that we have the condition that $[\lambda_1 \circ \phi(e_{11})] \ge L \cdot [p]$. Another condition $\lambda_1 \circ \phi(e_{11}) \sim \lambda_1 \circ \psi(e_{11})$ follows from the condition $\phi \sim_h \psi$.)

$\square$

The author is indebted to Professor G. Elliott for pointing out the proof of the following result to him.

LEMMA 1.6.8. *Suppose that C is a unital C^*-algebra, and $D \subset C$ is a finite dimensional C^*-subalgebra. For any finite set $F \subset C$ and any positive number $\varepsilon > 0$, there are a finite set $G \subset C$ and a number $\delta > 0$ such that if B is a unital C^*-algebra, $\lambda \in Map(C, B)$ is G-δ multiplicative, then there is a $\lambda' \in Map(C, B)$ satisfying the following conditions.*
1. $\lambda'|_D$ is a homomorphism.
2. $\|\lambda'(f) - \lambda(f)\| < \varepsilon, \quad \forall f \in F$.

Proof: Without loss of generality, we assume that $\|f\| \le 1$ for all $f \in F$.
By Kasparov's version of Stinespring Dilation Theorem, for the completely positive linear $*$-contraction $\lambda : C \to B$, there is a homomorphism $\phi : C \to M(B \otimes \mathcal{K})$ such that $\lambda(f) = p\phi(f)p$ for all $f \in F$, where $\mathcal{K}$ is the algebra of all compact operators on an infinite dimensional separable Hilbert space, $M(B \otimes \mathcal{K})$ is the multiplier algebra of $B \otimes \mathcal{K}$, and $p = \mathbf{1}_B \otimes e_{11} \in B \otimes \mathcal{K} \subset M(B \otimes \mathcal{K})$.
For the above λ and ϕ, it is straight forward to check that, for any fixed element $a \in C$, if $\|\lambda(a \cdot a^*) - \lambda(a)\lambda(a^*)\| < \delta$, then $\|p\phi(a)(1-p) \cdot (p\phi(a)(1-p))^*\| < \delta$. Therefore, if we choose the finite set G to satisfy that $G = G^*$, then the G-δ-multiplicativity of the map λ implies the following property of the dilation ϕ and the cutting down projection p:

$$(*) \qquad \|\phi(a) - (p\phi(a)p + (\mathbf{1} - p)\phi(a)(\mathbf{1} - p))\| < 2\sqrt{\delta}, \quad \forall a \in G,$$

where $\mathbf{1}$ is the unit of $M(B \otimes \mathcal{K})$.
By a well known perturbation technique (see [Gli] and [Br]), we have the following: If G contains all matrix units e_{ij} of each block of D and δ is small enough,

then the above condition (*) implies that there is a unitary $u \in M(B\otimes\mathcal{K})$ with $\|u-\mathbf{1}\|<\frac{\varepsilon}{2}$, such that

$$u\phi(D)u^* \subset pM(B\otimes\mathcal{K})p \oplus (\mathbf{1}-p)M(B\otimes\mathcal{K})(\mathbf{1}-p).$$

(One can obtain the above assertion by applying Lemma III.3.2 of [Da] (or even stronger result of [Ch]) with $\phi(D)$ and $pM(B\otimes\mathcal{K})p\oplus(\mathbf{1}-p)M(B\otimes\mathcal{K})(\mathbf{1}-p)$ in place of $\mathcal{U}$ and $\mathcal{B}$ in [Da, III.3.2], respectively.)
The map $\lambda' : C \to B$, defined by $\lambda'(f) = pu\phi(f)u^*p$, is as desired.

□

The following result can be considered to be a generalization of Theorem 2.29 of [EG 2].

Theorem 1.6.9. *Suppose that $A=\bigoplus_{i=1}^{s} M_{k_i}(C(X_i))$ and $F\subset A$ is weakly approximately constant to within ε. Suppose that C is a C^*-algebra, the homomorphisms ϕ and $\psi \in Hom(A,C)$ are homotopic to each other. There are a finite set $G\subset C$, a number $\delta>0$, and a positive integer $L>0$ such that the following is true.*
If B is a unital C^*-algebra, $p\in B$ is a projection, $\lambda_0\in Map(C,pBp)$ is G-δ multiplicative, $\lambda_1\in Hom(C,(\mathbf{1}-p)B(\mathbf{1}-p))$ is a homomorphism with finite dimensional image satisfying that for each $i\in\{1,2,\cdots,s\}$, $[(\lambda_1\circ\phi)(e_{11}^i)]\geq L\cdot[p]$, where e_{11}^i is the matrix unit (of upper left corner) of the i-th block, $M_{k_i}(C(X_i))$, of A, then there is a unitary $u\in B$ such that

$$\|(\lambda\circ\phi)(f)-u(\lambda\circ\psi)(f)u^*\|<8\varepsilon, \quad \forall f\in F,$$

where $\lambda\in Map(A_1,B)$ is defined by $\lambda=\lambda_0\oplus\lambda_1$.

Proof: Let ϕ_t be the homotopy between ϕ and ψ. It is well known that there is a unitary path $u_t\in C$ such that

$$\phi_t(\mathbf{1}_{A^i}) = u_t\phi_0(\mathbf{1}_{A^i})u_t^*,$$

for all blocks $A^i=M_{k_i}(C(X_i))$. Therefore, without loss of generality, we assume that $\phi(\mathbf{1}_{A^i})=\psi(\mathbf{1}_{A^i})$, and that $\phi|_{A^i}$ is homotopic to $\psi|_{A^i}$ within the corner $\phi(\mathbf{1}_{A^i})C\phi(\mathbf{1}_{A^i})$.
Apply Lemma 1.6.7 to $\phi|_{A^i}$, $\psi|_{A^i}$ and $\pi_i(F)$, where π_i is the quotient map from A to A^i, to obtain G_1 ($\subset C$), δ_1 and L as G, δ and L in Lemma 1.6.7. For convenience, without loss of generality, we assume that $\|g\|\leq 1$ for all $g\in G_1$. Let $E_i=\phi(\mathbf{1}_{A^i})$. Consider the finite dimensional subalgebra $D:=\mathbb{C}\cdot E_1\oplus\mathbb{C}\cdot E_2\oplus\cdots\oplus\mathbb{C}\cdot E_s\subset C$. Applying Lemma 1.6.8, there are $G\subset C$ with $G\supset G_1$ and $\delta>0$ with $\delta<\frac{\delta_1}{3}$ such that if $\lambda_0\in \mathrm{Map}(C,pBp)$ is G-δ multiplicative, then there is another map $\lambda_0'\in\mathrm{Map}(C,pBp)$ satisfying the following conditions.
1. The restriction $\lambda_0'|_D$ is a homomorphism.
2. $\|\lambda_0'(f)-\lambda_0(f)\|<\min(\frac{\delta_1}{3},\varepsilon), \quad \forall f\in G_1\cup\phi(F)\cup\psi(F)$.

As a consequence we also have
3. λ_0' is G_1-δ_1 multiplicative.
The condition 1 above yields that $\{\lambda_0'(E_i)\}_{i=1}^s$ are mutually orthogonal projections.
Such G, δ and L are as desired.
Suppose that λ_0 and λ_0' are as above. Set $\lambda' = \lambda_0' \oplus \lambda_1$. From Lemma 1.6.7 and the ways G_1, δ_1 and L are chosen, there are unitaries $u_i \in \lambda'(E_i)B\lambda'(E_i)$ such that

$$\|(\lambda' \circ \phi|_{A^i})(f) - u_i(\lambda' \circ \psi|_{A^i})(f)u_i^*\| < 6\varepsilon, \quad \forall f \in F_i.$$

Then the unitary $u = \bigoplus_i u_i \oplus (\mathbf{1} - \sum_i \lambda'(E_i))$ satisfies

$$\|(\lambda' \circ \phi)(f) - u(\lambda' \circ \psi)(f)u^*\| < 6\varepsilon, \quad \forall f \in F.$$

Hence

$$\|(\lambda \circ \phi)(f) - u(\lambda \circ \psi)(f)u^*\| < 6\varepsilon + 2\varepsilon = 8\varepsilon, \quad \forall f \in F.$$

$\square$

Remark 1.6.10. The version of Theorem 2.29 of [EG2] with A being a direct sum of full matrix algebras is a direct consequence of the above theorem and Corollary 2.24 of [EG2] (see [EG2, Theorem 2.21] also). In order to obtain the general version of Theorem 2.29 of [EG2], one needs to apply the dilation lemma [EG2, 2.13] and Lemma 1.6.8 above. (The number 8ε should be changed to $5 \cdot 8\varepsilon = 40\varepsilon$ which is still better than 70ε in [EG2].)
The following lemma is a direct consequence of Lemma 1.6.5.

Lemma 1.6.11. *Let X be a finite simplicial complex and $A = C(X)$. Let $F \subset A$ be a finite set and $\varepsilon > 0$. There are a finite set $G \subset A$ and a number $\delta > 0$ with the following property.*
If B is a unital C^*-algebra, $\phi_t : A \to B$, $0 \le t \le 1$ is a continuous path of G-δ multiplicative maps (i.e., $\phi_t \in Map_{G-\delta}(A, B)$), then there are a positive integer L, a homomorphism $\lambda : A \to M_L(B)$ with finite dimensional image, and a unitary $u \in M_{L+1}(B)$ such that

$$\|(\phi_0 \oplus \lambda)(f) - u(\phi_1 \oplus \lambda)(f)u^*\| < \varepsilon, \quad \forall f \in F.$$

The proof of the following corollary has some similarity to the proof of Corollary 1.6.6. Such method will be used frequently.

Corollary 1.6.12. *Let X be a finite simplicial complex and $A = C(X)$. Let $F \subset A$ be a finite set and $\varepsilon > 0$. There are a finite set $G \subset A$ and a number $\delta > 0$ with the following property.*
If B is a unital C^*-algebra, $p \in B$ is a projection, $\phi_t : A \to pBp$, $0 \le t \le 1$ is a continuous path of G-δ multiplicative maps , then there are a positive integer L

and a number $\eta > 0$ such that for any η-dense subset $\{x_1, x_2, \cdots, x_\bullet\} \subset X$, any set of mutually orthogonal projections $\{p_1, p_2, \cdots, p_\bullet\} \subset B \otimes \mathcal{K}$ with $[p_i] \geq L \cdot [p]$ and $p_i \perp p$, we have

$$\|\phi_0(f) \oplus \sum_{i=1}^{\bullet} f(x_i)p_i - u(\phi_1(f) \oplus \sum_{i=1}^{\bullet} f(x_i)p_i)u^*\| < \varepsilon, \quad \forall f \in F,$$

for a certain unitary $u \in (p \oplus p_1 \oplus \cdots p_\bullet)(B \otimes \mathcal{K})(p \oplus p_1 \oplus \cdots p_\bullet)$.

Proof: For the finite set $F \subset A$, choose η small enough such that if $\mathrm{dist}(x, x') < \eta$, then $\|f(x) - f(x')\| < \frac{\varepsilon}{3}$ for all $f \in F$. Apply Lemma 1.6.11 to F and $\frac{\varepsilon}{3}$ to obtain G and δ. For the path $\phi_t : A \to pBp$, there exist a positive integer L, a homomorphism $\lambda : A \to M_L(pBp)$, and a unitary $u_1 \in M_{L+1}(pBp)$ as in Lemma 1.6.11. That is

$$\|(\phi_0 \oplus \lambda)(f) - u_1(\phi_1 \oplus \lambda)(f)u_1^*\| < \frac{\varepsilon}{3}, \quad \forall f \in F.$$

From 1.6.4, λ is of the form

$$\lambda(f) = \sum_{i=1}^{l} f(y_i)q_i,$$

where $\{y_1, y_2, \cdots, y_l\} \subset X$, and $\{q_1, q_2, \cdots, q_l\} \subset M_L(pBp)$ is a set of mutually orthogonal projections.

Since $\{x_1, x_2, \cdots, x_\bullet\}$ is an η-dense subset of X, we can divide the set $\{y_1, y_2, \cdots, y_l\}$ into a disjoint union of subsets $X_1 \cup X_2 \cup \cdots \cup X_\bullet$ (some X_i may be empty) such that $\mathrm{dist}(y, x_i) < \eta$ for any $y \in X_i$. Set $p_i' := \sum_{y_j \in X_i} q_j$ and define $\lambda' : C(X) \to M_L(pBp)$ by $\lambda'(f) = \sum_{i=1}^{\bullet} f(x_i)p_i'$. (Note that for some i, p_i' might be 0.) Then from the way η is chosen, we have

$$\|\lambda'(f) - \lambda(f)\| < \frac{\varepsilon}{3}, \quad \forall f \in F.$$

Therefore,

$$\|(\phi_0 \oplus \lambda')(f) - u_1(\phi_1 \oplus \lambda')(f)u_1^*\| < \varepsilon, \quad \forall f \in F.$$

Our corollary follows from the fact $[p_i'] \leq L \cdot [p] \leq p_i$.

$\square$

If X does not contain any isolated point, then in the above corollary, we can change the condition $[p_i] \geq L \cdot [p]$ to a weaker condition $[p_i] \geq [p]$, by choosing η smaller. (Roughly speaking, this is true because η could be chosen so small that if $\{x_i\}$ is η-dense, then there are at least L points of x_i in the η'-neighborhood of any point in X for a pre-given small number η'. If X is a space of single point, then this is not true.) Therefore, the number L does not appear in the following corollary.

Corollary 1.6.13. *Let X be a finite simplicial complex without any single*

point components, $A = C(X)$. Let $F \subset A$ be a finite set and $\varepsilon > 0$. There are a finite set $G \subset A$ and a number $\delta > 0$ with the following property.
If B is a unital C^*-algebra, $p \in B$ is a projection, $\phi_t : A \to pBp$, $0 \le t \le 1$, is a continuous path of G-δ multiplicative maps , then there is a number $\eta > 0$ such that for any η-dense subset $\{x_1, x_2, \cdots, x_\bullet\} \subset X$, any set of mutually orthogonal projections $\{p_1, p_2, \cdots, p_\bullet\} \subset B \otimes \mathcal{K}$ with $[p_i] \ge [p]$ and $p_i \perp p$, we have

$$\|\phi_0(f) \oplus \sum_{i=1}^{\bullet} f(x_i)p_i - u(\phi_1(f) \oplus \sum_{i=1}^{\bullet} f(x_i)p_i)u^*\| < \varepsilon, \quad \forall f \in F$$

for a certain unitary $u \in (p \oplus p_1 \oplus \cdots p_\bullet)(B \otimes \mathcal{K})(p \oplus p_1 \oplus \cdots p_\bullet)$.

Proof: Let L and η_1 (in place of η) be as in Corollary 1.6.12. Let η_2 be the minimum of the diameters of path connected components of X, which is positive since X has no single point component. And let η_3 be a positive number such that if $\mathrm{dist}(x, x') < \eta_3$, then $\|f(x) - f(x')\| < \varepsilon$.
Define $\eta' = \min(\eta_1, \eta_2, \eta_3)$. Let $\eta = \frac{\eta'}{8L}$.
Suppose that $X' = \{x_1, x_2, \cdots, x_\bullet\}$ is an η-dense finite subset of X. Choose a η'-dense subset $\{x_{k_1}, x_{k_2}, \cdots, x_{k_l}\} \subset X'$ such that $\mathrm{dist}(x_{k_i}, x_{k_j}) \ge \frac{\eta'}{2}$ if $i \ne j$. (Such subset exists. It could be chosen to be a maximum subset of X' such that the distance of any two points in the set is at least $\frac{\eta'}{2}$. Then the η'-density follows from the maximality.) It is easy to see that there is a partition of X' as $X' = X_1 \cup X_2 \cup \cdots \cup X_l$ such that

$$X' \cap B_{\frac{\eta'}{4}}(x_{k_i}) \subset X_i \subset X' \cap B_{\eta'}(x_{k_i}).$$

Since X' is η-dense and $\eta = \frac{\eta'}{8L}$,

$$\#(X_i) \ge \#(X' \cap B_{\frac{\eta'}{4}}(x_{k_i})) \ge L.$$

(Here we also use the fact that the connected component of x_{k_i} in X has diameter at least η'.)
Let p_j, $j = 1, 2, \cdots \bullet$, be the projections as in the corollary. Define $q_i = \sum_{x_j \in X_i} p_j$, $i = 1, 2, \cdots l$. Then from $[p_j] \ge [p]$ and $\#(X_i) \ge L$, we have, $[q_i] \ge L \cdot [p]$.
Our corollary (with 3ε in place of ε) follows from an application of Corollary 1.6.12 to $\{x_{k_i}\}_{i=1}^l$ and $\{q_i\}_{i=1}^l$, and the estimation

$$\|\sum_{i=1}^{\bullet} f(x_i)p_i - \sum_{i=1}^{l} f(x_{k_i})q_i\| < \varepsilon, \quad \forall f \in F.$$

(The above estimation is a consequence of the way η_3 is chosen and the fact that $X_i \subset B_{\eta'}(x_{k_i})$ with $\eta' < \eta_3$.)

$\square$

The following lemma is proved by applying Lemma 1.6.8.

Lemma 1.6.14. *Let $A = \bigoplus A^k = \bigoplus_{k=1}^l M_{s(k)}(C(X_k))$, where X_k are connected simplicial complexes and $\{s(k)\}$ are positive integers. For any finite set $G' \subset A$, any number $\delta' > 0$, any finite sets $G_1^k \subset C(X_k)$ and any numbers $\delta_1^k > 0$, $k = 1, 2, \cdots l$, there are a finite set $G \subset A$ and a number $\delta > 0$ such that if $\phi \in Map(A, B)$ is G-δ multiplicative, then there is a map $\phi' \in Map(A, B)$ satisfying the following conditions.*
(1) ϕ' is G'-δ' multiplicative;
(2) $\|\phi'(g) - \phi(g)\| < \delta'$ for all $g \in G$;
(3) $\{\phi'(\mathbf{1}_{A^k})\}_{k=1}^l$ are mutually orthogonal projections in B and $\phi'(e_{11}^k) \in B$ are subprojections of $\phi'(\mathbf{1}_{A^k}) \in B$. And if each $\phi_1^k \in Map(C(X_k), B)$ is the restriction of ϕ' on $e_{11}^k M_{s(k)}(C(X_k)) e_{11}^k \cong C(X_k)$, then one can identify $\phi'(\mathbf{1}_{A^k}) B \phi'(\mathbf{1}_{A^k})$ with $\phi'(e_{11}^k) B \phi'(e_{11}^k) \otimes M_{s(k)}$ such that

$$\phi' = \bigoplus_{k=1}^{l} \phi_1^k \otimes id_{s(k)}.$$

Furthermore, ϕ_1^k is G_1^k-δ_1^k Multiplicative.

Proof: The part of G_1^k-δ_1^k multiplicativity of ϕ_1^k follows from the G'-δ' multiplicativity of ϕ' if we enlarge the set G' and reduce the number δ' so that $G' \supset \{g \cdot e_{11}^k \mid g \in G_1^k\}$ and $\delta' < \delta_1^k$. Also we can assume that G' contains $\{e_{ij}^k\}$—the set of all matrix units.
By Lemma 1.6.8, without loss of generality, we assume that the restriction $\phi|_{\bigoplus_{k=1}^l M_{s(k)}(\mathbb{C})}$ is a homomorphism.
Let $\phi_1^k = \phi|_{e_{11}^k A e_{11}^k} \in \mathrm{Map}(C(X_k), \phi(e_{11}^k) B \phi(e_{11}^k))$. Then $\phi' := \bigoplus_{k=1}^l \phi_1^k \otimes \mathrm{id}_{s(k)}$ is defined by

$$\phi'(f) = \sum_{i,j} \phi(e_{i1}^k) \phi(f_{ij} \cdot e_{11}^k) \phi(e_{1j}^k),$$

where $f = (f_{ij})_{s(k) \times s(k)} \in A^k$.
Note that for the above $f \in A^k$, one can write $f = \sum_{i,j} e_{i1}^k \cdot (f_{ij} \cdot e_{11}^k) \cdot e_{1j}^k$. Obviously, if we choose G to be the set of all the elements which can be expressed as products of at most ten elements from the set G', and if we choose δ small enough, then (1) and (2) will hold for ϕ'. (Notice that G' contains all the matrix units.)

$\square$

The following result follows from Lemma 1.6.14 and Corollary 1.6.12 (see also 1.2.19).

Corollary 1.6.15. *Let $A = \bigoplus_{k=1}^l M_{s(k)}(C(X_k))$, where X_k are connected finite simplicial complexes and $s(k)$ are positive integers. Let $F \subset A$ be a finite set and $\varepsilon > 0$. There are a finite set $G \subset A$ and a number $\delta > 0$ with the following property.*

If B is a unital C^*-algebra, $p \in B$ is a projection, $\phi_t : A \to pBp$, $0 \le t \le 1$ is a continuous path of G-δ multiplicative maps , then there are a positive integer L, and $\eta > 0$ such that for a homomorphism $\lambda : A \to B \otimes \mathcal{K}$ (with finite dimensional image), there is a unitary $u \in B$ satisfying:

$$\|\phi_0(f) \oplus \lambda(f) - u(\phi_1(f) \oplus \lambda(f))u^*\| < \varepsilon, \quad \forall f \in F,$$

provided that λ is of the following form: there are an η-dense subset $\{x_1, x_2, \cdots, x_\bullet\} \subset \coprod_{k=1}^{l} X_k \quad (= SP(A))$, and a set of mutually orthogonal projections $\{p_1, p_2, \cdots, p_\bullet\} \subset \lambda(\bigoplus_k e_{11}^k)(B \otimes \mathcal{K})\lambda(\bigoplus_k e_{11}^k)$ with $[p_i] \ge L \cdot [p]$, such that

$$\lambda(f) = \sum_{i=1}^{\bullet} p_i \otimes f(x_i), \quad \forall f \in A$$

under the identification $\lambda(\mathbf{1}_{A^k})B\lambda(\mathbf{1}_{A^k}) \cong (\lambda(e_{11}^k)B\lambda(e_{11}^k)) \otimes M_{s(k)}(\mathbb{C})$.

Proof: One can apply Lemma 1.6.14 to $\phi_t \in \mathrm{Map}(A, pBp[0,1])$ to reduce the problem to the case of $A = C(X_k)$ which is Corollary 1.6.12. (Here $pBp[0,1]$ is defined to be the C^*-algebra of continuous pBp valued functions on $[0,1]$.) $\square$

For convenience, we introduce the following definitions.

Definition 1.6.16. A homomorphism $\phi : A = \bigoplus_{i=1}^{n} M_{k_i}(C(X_i)) \to B = \bigoplus_{j=1}^{n'} M_{l_j}(C(Y_j))$ is called m-large if for each partial map $\phi^{ij} : A^i = M_{k_i}(C(X_i)) \to B^j = M_{l_j}(C(Y_j))$ of ϕ,

$$\mathrm{rank}(\phi^{ij}(\mathbf{1}_{A^i}) \ge m \cdot \mathrm{rank}(\mathbf{1}_{A^i})(= m \cdot k_i).$$

Definition 1.6.17. Let X be a connected finite simplicial complex, $A = M_k(C(X))$. A unital *-monomorphism $\phi : A \to M_l(A)$ is called a (unital) simple embedding if it is homotopic to the homomorphism $\mathrm{id} \oplus \lambda$, where $\lambda : A \to M_{l-1}(A)$ is defined by

$$\lambda(f) = \mathrm{diag}(\underbrace{f(x_0), f(x_0), \cdots, f(x_0)}_{l-1}),$$

for a fixed base point $x_0 \in X$.
Let $A = \bigoplus_{i=1}^{n} M_{k_i}(C(X_i))$, where X_i are connected finite simplicial complexes. A unital *-monomorphism $\phi : A \to M_l(A)$ is called a (unital) simple embedding, if ϕ is of the form $\phi = \oplus\phi^i$ defined by

$$\phi(f_1, f_2, \cdots, f_n) = (\phi^1(f_1), \phi^2(f_2), \cdots, \phi^n(f_n)),$$

where the homomorphisms $\phi^i : A^i(= M_{k_i}(C(X_i))) \to M_l(A^i)$ are unital simple embeddings.

1.6.18. For each connected finite simplicial complex X, there is a three dimensional connected simplicial complex $Y = Y_1 \vee Y_2 \vee \cdots \vee Y_\bullet$ such that $K^*(X) = K^*(Y)$, where Y_i are the following special spaces: $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^{\infty}$, $\{T_{III,k}\}_{k=2}^{\infty}$ and S^2.
The space $[0,1]$ could be avoided in the construction of Y. But we would like to use the space $[0,1]$ for the following special case: If $K^0(X) = \mathbb{Z}$ and $K^1(X) = 0$ (e.g., X is contractible) and X is not the space of a single point, then we choose $Y = [0,1]$. When X is the space of a single point, choose $Y = \{pt\}$.
The following result is Lemma 2.1 of [D2] (see Lemma 3.13 and Lemma 3.14 of [G4] also).

Lemma 1.6.19. ([D2, 2.1]) *Let $B_1 = \bigoplus_{j=1}^{s} M_{k(j)}(C(Y_j))$, where Y_j are the following spaces: $\{pt\}$, $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^{\infty}$, $\{T_{III,k}\}_{k=2}^{\infty}$, and S^2. Let X be a connected finite simplicial complex, let Y be the three dimensional space defined in 1.6.18 with $K^*(X) = K^*(Y)$, and let $A = M_N(C(X))$.*
Let $\alpha_1 : B_1 \to A$ be a homomorphism. For any finite sets $G \subset B_1$ and $F \subset A$, and any number $\delta > 0$, there exists a diagram

$$\begin{array}{ccc} A & \xrightarrow{\phi} & A' \\ \uparrow{\scriptstyle \alpha_1} & \searrow^{\beta} & \uparrow{\scriptstyle \alpha_2} \\ B_1 & \xrightarrow{\psi} & B_2 \end{array},$$

where
$A' = M_L(A), B_2 = M_S(C(Y))$;
ψ is a homomorphism, α_2 is a unital homomorphism, and ϕ is a unital simple embedding (see 1.6.17);
$\beta \in Map(A, B_2)$ is F-δ multiplicative.
Moreover there exist homotopies $\Psi \in Map(B_1, B_2[0,1])$ and $\Phi \in Map(A, A'[0,1])$ such that Ψ is G-δ multiplicative, Φ is F-δ multiplicative, and

$$\Psi|_1 = \psi, \quad \Psi|_0 = \beta \circ \alpha_1, \quad \Phi|_0 = \alpha_2 \circ \beta \quad \textit{and} \quad \Phi|_1 = \phi.$$

(In the application of this lemma, it is important to require that ϕ is a unital simple embedding. This requirement means that ϕ defines the same element in $kk(X,X)$ (connective KK theory) as the identity map id : $A \to A$. Roughly speaking, this lemma (and Theorem 6.26 below) means that an "identity map" could factor through matrix algebras over Y — a special space of dimension three.)

Proof: If one assumes that $\alpha_1 : B_1 \to A$ is m-large (see 1.6.16) for a number $m > 4\dim(X)$, then this lemma becomes Lemma 2.1 of [D2]. We make use of this special case to prove the general case as below.

Define a unital simple embedding $\lambda: A \to M_m(A)$ $(m > 4\dim(Y))$ by

$$\lambda(f) = \mathrm{diag}(f, \underbrace{f(x_0), f(x_0), \cdots, f(x_0)}_{m-1}).$$

Then $\alpha_1' = \lambda \circ \alpha_1$ is m-large. Apply Lemma 2.1 of [D2]— the special case of the lemma to α_1', $\lambda(F) \subset M_m(A)$ and $G \subset B_1$ to obtain the following diagram

$$\begin{array}{ccc} M_m(A) & \xrightarrow{\phi'} & A' \\ \Big\uparrow{\scriptstyle \alpha_1'} & \overset{\beta'}{\searrow} & \Big\uparrow{\scriptstyle \alpha_2'} \\ B_1 & \xrightarrow{\psi'} & B_2 \end{array}$$

with homotopy paths Ψ' and Φ' with properties described in the lemma for the homomorphism α_1', finite sets $\lambda(F) \subset M_m(A)$ and $G \subset B_1$.
Define $\beta = \beta' \circ \lambda$, $\phi = \phi' \circ \lambda$, $\alpha_2 = \alpha_2'$, $\psi = \psi'$, $\Psi = \Psi'$ and $\Phi = \Phi' \circ \lambda$. Then we have the desired diagram with the desired properties.

$\square$

Remark 1.6.20. From the construction of ϕ and α_2 in the proof of [D2, Lemma 2.1], we know that ϕ and α_2 take trivial projections to trivial projections. But ψ may not take trivial projections to trivial projections unless α_1 does.

1.6.21. Let X and Y be path connected finite simplicial complexes, and $C = M_k(C(Y)), D = M_l(C(X))$. Let $x_0 \in X$ and $y_0 \in Y$ be fixed base points. Then from Lemma 3.14 of [EG2], we have the following: any homomorphism $\phi \in \mathrm{Hom}(C, D)$ is homotopy equivalent to a homomorphism $\phi' \in \mathrm{Hom}(C, D)$ (within $\mathrm{Hom}(C, D)$) such that $\phi'(C^0) \subset D^0$, where C^0, D^0, are the ideals of C and D, respectively, which consist of matrix valued functions vanishing on the base points (see 1.1.7(h)). In other words, there is a unitary $U \in M_l(\mathbb{C})$ such that

$$\phi'(f)(x_0) = U \begin{pmatrix} f(y_0) & & & & & \\ & \ddots & & & & \\ & & f(y_0) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} U^* \in M_l(\mathbb{C}) \quad \forall f \in C.$$

Notice that if a homomorphism α_2 is as desired in Lemma 1.6.19, then any homomorphism, which is homotopic to α_2, is also as desired. Therefore, in Lemma 1.6.19, we can require that the homomorphism $\alpha_2 : B_2(= M_S(C(Y)) \to A'(= M_L(C(X)))$ is of the above form for certain base points $y_0 \in Y$ and $x_0 \in X$.

In the following, let us explain why we can also choose the homomorphism α_2 to be injective.
If X is the space of a single point, then Y is also the space of a single point by our choice. And therefore, α_2 is injective, since B_2 is simple.
If the connected simplicial complex X is not a single point (i.e., $\dim(X) \geq 1$), then it can be proved that there is a continuous surjective map $g : X \to Y$, using the standard idea of Peano curve. In fact, one can assume that the map g is homotopy trivial—one can make it factor through an interval.
On the other hand, by Theorem 6.4.4 of [DN], if $L \geq 3S(\dim(X)+1)$, then for the unital homomorphism $\alpha_2 : M_S(C(Y)) \to M_L(C(X))$, there is a homomorphism $\alpha' : M_S(C(Y)) \to M_{L-S}(C(X))$ such that α_2 is homotopic to the homomorphism defined by

$$f \mapsto \mathrm{diag}(\alpha'(f), f(y_0)),$$

Then α_2 is homotopic to $\mathrm{diag}(\alpha', g^*)$ defined by

$$f \mapsto \mathrm{diag}(\alpha'(f), f \circ g),$$

since g is homotopy trivial. So we can replace α_2 by $\mathrm{diag}(\alpha', g^*)$ which is injective, since g is surjective.
Similarly, if $\mathrm{SP}(B_2) = Y$ is not a single point space (i.e., X is not the space of a single point), then the homomorphism $\psi : B_1 \to B_2$ could be chosen to be injective with in the same homotopy class of $\mathrm{Hom}(B_1, B_2)$, provided that $\alpha_1(\mathbf{1}_{B_1^i}) \neq 0$, for each block B_1^i of B_1 (later on, we will always assume α_1 satisfies this condition, since otherwise this block can be deleted from B_1).

Lemma 1.6.22. *Let $B = M_k(C(Y))$, $A = M_l(C(X))$. Suppose that a unital homomorphism $\alpha : B \to A$ satisfies $\alpha(B^0) \subset A^0$, and takes any trivial projections of B to trivial projections of A, where $B^0 = M_k(C_0(Y)) := \{f \in M_k(C(Y)) \mid f(y_0) = 0\}$, and $A^0 = M_l(C_0(X)) := \{f \in M_l(C(X)) \mid f(x_0) = 0\}$, for some fixed base points $y_0 \in Y$ and $x_0 \in X$. Let $\beta_0 : B \to M_n(B)$ and $\beta_1 : A \to M_n(A)$ be unital homomorphisms defined by*

$$\beta_0(f)(y) = diag(f(y), f(y_0), \cdots, f(y_0)), \quad \forall f \in B,$$

and

$$\beta_1(f)(x) = diag(f(x), f(x_0), \cdots, f(x_0)), \quad \forall f \in A.$$

Then the following diagram commutes up to unitary equivalence.

$$\begin{array}{ccc} A & \xrightarrow{\beta_1} & M_n(A) \\ {\scriptstyle\alpha}\uparrow & & \uparrow{\scriptstyle\alpha\otimes id_n} \\ B & \xrightarrow{\beta_0} & M_n(B). \end{array}$$

I.e., there is a unitary $u \in M_n(A)$ such that

$$\beta_1 \circ \alpha = Adu \circ (\alpha \otimes id_n) \circ \beta_0.$$

Proof: $\beta_1 \circ \alpha$ is defined by:

$$f \mapsto \alpha(f) \mapsto \mathrm{diag}(\alpha(f), \alpha(f)(x_0), \cdots, \alpha(f)(x_0)),$$

and $(\alpha \otimes \mathrm{id}_n) \circ \beta_0$ is defined by:

$$f \mapsto \mathrm{diag}(f, f(y_0), \cdots, f(y_0)) \mapsto \mathrm{diag}(\alpha(f), \alpha(f(y_0)), \cdots, \alpha(f(y_0))),$$

where $\alpha(f(y_0))$ denotes the result of α acting on the constant function $g = f(y_0)$.
On the other hand, from $\alpha(B^0) \subset A^0$, we get

$$\alpha(f)(x_0) \ \sim_u \ diag(\underbrace{f(y_0), \cdots, f(y_0)}_{\frac{l}{k}}),$$

and from the fact that α takes trivial projections to trivial projections, we get

$$diag(\underbrace{f(y_0), \cdots, f(y_0)}_{\frac{l}{k}}) \ \sim_u \ \alpha(f(y_0)),$$

where the symbol $\sim_u$ means to be unitarily equivalent.

□

The following result is from [EG2] (see 5.10, 5.11 of [EG2]).

Lemma 1.6.23. *Let $Y = Y_1 \vee Y_2 \vee \cdots \vee Y_m$. If n is large enough, then any unital homomorphism $\beta : M_k(C(Y)) \to M_{nk}(C(Y))$ is homotopic to a homomorphism $\beta' : M_k(C(Y)) \to M_{nk}(C(Y))$, which factors through $\bigoplus_{i=1}^m M_{k_i}(C(Y_i))$, for certain integers $\{k_i\}$, as*

$$\beta' : \ M_k(C(Y)) \xrightarrow{\beta_1} \bigoplus_{i=1}^{m} M_{k_i}(C(Y_i)) \xrightarrow{\beta_2} M_{nk}(C(Y)).$$

Furthermore, β_1 and β_2 above can be chosen to be injective.

Proof: If $k = 1$, then the lemma is Lemma 5.11 of [EG2]. (Notice that, we choose both spaces X and Y in Lemma 5.11 of [EG2] to be the above space Y. In addition, the spaces X_i in Lemma 5.11 of [EG2] could be chosen to be spaces Y_i in our case, according to 5.10 of [EG2].)
For the general case, one writes β as $b \otimes \mathrm{id}_k$, where $b = \beta|_{e_{11}M_k(C(Y))e_{11}} : C(Y) \to \beta(e_{11})M_{nk}(C(Y))\beta(e_{11})$, then apply Lemma 5.11 of [EG2] to b.
Furthermore, one can make β_1 and β_2 injective in the same way as in the end of 1.6.21. (Or one observes that the maps β_1 and β_2 constructed in Lemma 5.11 of [EG2] are already injective for our case.)

□

Combining Lemmas 1.6.19, 1.6.21, 1.6.22 and 1.6.23, we have the following Lemma:

LEMMA 1.6.24. *Let* $B_1 = \bigoplus_{j=1}^{s} M_{k(j)}(C(Y_j))$, *where* Y_j *are spaces:* $\{pt\}$, $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^{\infty}$, $\{T_{III,k}\}_{k=2}^{\infty}$, *and* S^2. *Let* X *be a connected finite simplicial complex and let* $A = M_N(C(X))$.
Let $\alpha_1 : B_1 \to A$ *be a homomorphism with* $\alpha_1(\mathbf{1}_{B_1^i}) \neq 0$ *for each block* B_1^i *of* B_1. *For any finite sets* $G \subset B_1$ *and* $F \subset A$, *and any number* $\delta > 0$, *there exists a diagram*

$$\begin{array}{ccc} A & \xrightarrow{\phi} & A' \\ \uparrow{\scriptstyle \alpha_1} & \searrow^{\beta} & \uparrow{\scriptstyle \alpha_2} \\ B_1 & \xrightarrow{\psi} & B_2\ , \end{array}$$

where
$A' = M_L(A)$, *and* B_2 *is a direct sum of matrix algebras over the spaces:* $\{pt\}$, $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^{\infty}$, $\{T_{III,k}\}_{k=2}^{\infty}$, *and* S^2;
ψ *is a homomorphism,* α_2 *is a unital injective homomorphism, and* ϕ *is a unital simple embedding (see 1.6.17).*
$\beta \in Map(A, B_2)$ *is* F-δ *multiplicative.*
Moreover, there exist homotopies $\Psi \in Map(B_1, B_2[0,1])$ *and* $\Phi \in Map(A, A'[0,1])$ *such that* Ψ *is* G-δ *multiplicative,* Φ *is* F-δ *multiplicative, and*

$$\Psi|_1 = \psi, \quad \Psi|_0 = \beta \circ \alpha_1, \quad \Phi|_0 = \alpha_2 \circ \beta \quad \textit{and} \quad \Phi|_1 = \phi.$$

Furthermore, if X *is not the space of a single point, then at least one of the blocks of* B_2 *has spectrum different from the space of single point and* ψ *can be chosen to be injective.*

Proof: Let

$$\begin{array}{ccc} A & \xrightarrow{\phi} & A' \\ \uparrow{\scriptstyle \alpha_1} & \searrow^{\beta} & \uparrow{\scriptstyle \alpha_2} \\ B_1 & \xrightarrow{\psi} & B_2\ , \end{array}$$

be the diagram described in Lemma 1.6.19 with homotopies Φ and Ψ. Let n be the integer obtained by applying Lemma 1.6.23 to $B_2 = M_k(C(Y))$. Then apply 1.6.22 to $\alpha_2 : B_2 \to A'$ to obtain a diagram

$$\begin{array}{ccc} A' & \xrightarrow{\beta_1} & M_n(A') \\ {\scriptstyle \alpha_2}\uparrow & & \uparrow{\scriptstyle \alpha_2 \otimes \mathrm{id}_n} \\ B_2 & \xrightarrow{\beta_0} & M_n(B_2). \end{array}$$

which commutes up to homotopy. (Here we have the condition that α_2 takes trivial projections to trivial projections from Remark 1.6.20. Also, α_2 is homotopic to a homomorphism which takes B^0 to A'^0.)

Furthermore, from Lemma 1.6.23, β_0 is homotopic to a homomorphism β_0' factoring through a C^*-algebra $newB_2$ which is a direct sum of matrix algebras over spaces $\{pt\}$, $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^{\infty}$, $\{T_{III,k}\}_{k=2}^{\infty}$, and S^2. Now it is routine to finish the construction of the diagram. We omit the details.

$\square$

1.6.25. Our next task is to add a homomorphism $\lambda : A \to M_n(B_2)$ into the diagram in Lemma 1.6.24 to obtain diagrams,

$$\begin{tikzcd} A \arrow[dr, "\beta\oplus\lambda"] & \\ B_1 \arrow[u, "\alpha_1"] \arrow[r, "\psi\oplus(\lambda\circ\alpha_1)"] & M_{n+1}(B_2) \end{tikzcd}$$

and

$$\begin{tikzcd} A \arrow[r, "\phi\oplus((\alpha_2\otimes id_n)\circ\lambda)"] \arrow[dr, "\beta\oplus\lambda"] & M_{n+1}(A') \\ & M_{n+1}(B_2) \arrow[u, "\alpha_2\otimes id_{n+1}"'] \end{tikzcd}$$

which are almost commutative up to unitary equivalence, using Corollary 1.6.15.

To do so, we make the following assumption.

Assumption: $\alpha_1 : B_1 \to A$ is injective.

Let $G_1 \subset B_1$ and $F_1 \subset A$ be any finite sets, and $\varepsilon > 0$. We will make the above diagrams approximately commute on G_1 and F_1, respectively, to within ε, up to unitary equivalence.

Apply Corollary 1.6.15 to $G_1 \subset B_1$ (in place of $F \subset A$) and $\varepsilon > 0$, to obtain $G \subset B_1$ and δ_1 (in place of the set G and the number δ, respectively in Corollary 1.6.15). Similarly, apply Corollary 1.6.15 to $F_1 \subset A$ (in place of $F \subset A$) and $\varepsilon > 0$ to obtain $F \subset A$ and δ_2 (in place of the set G and the number δ in Corollary 1.6.15).

Let $\delta = \min(\delta_1, \delta_2)$.

Suppose that the diagram

$$\begin{tikzcd} A \arrow[r, "\phi"] \arrow[dr, "\beta"] & A' \\ B_1 \arrow[u, "\alpha_1"] \arrow[r, "\psi"] & B_2 \arrow[u, "\alpha_2"'] \end{tikzcd}$$

is the one constructed in Lemma 1.6.24 with homotopy path $\Psi \in \mathrm{Map}(B_1, B_2[0,1])$ between $\beta \circ \alpha_1$ and ψ, and homotopy path $\Phi \in \mathrm{Map}(A, A'[0,1])$ between $\alpha_2 \circ \beta$ and ϕ, corresponding to the sets $G \subset B_1$, $F \subset A$ and the number $\delta > 0$.

Regarding the homotopy path Ψ as the homotopy path ϕ_t in Corollary 1.6.15, we can obtain η_1, L_1 as the numbers η and L in Corollary 1.6,15. Similarly, replacing the above Ψ by Φ, we obtain η_2, L_2.

Notice that the injectivity of α_1 implies that, for each block B_1^j of B_1, $\mathrm{SP}(\alpha_1^j) = Y_j (= \mathrm{SP}(B_1^j))$. Therefore there is an η_2-dense subset $\{x_1, x_2, \cdots x_m\}$ of X such that $\bigcup_{i=1}^m \mathrm{SP}\alpha_1^j|_{x_i}$ is η_1-dense in Y_j for each $j \in \{1, 2, \cdots, s\}$.
Define $\lambda_1 : A(= M_N(C(X))) \to M_{mN}(B_2)$ by

$$\lambda_1(f) = \mathrm{diag}(\mathbf{1}_{B_2} \otimes f(x_1), \mathbf{1}_{B_2} \otimes f(x_2), \cdots, \mathbf{1}_{B_2} \otimes f(x_m)).$$

Then $\lambda_1 \circ \alpha_1 : B_1 \to M_{mN}(B_2)$ is a homomorphism defined by the point evaluations on the η_1-dense subset $\bigcup_{j=1}^s \bigcup_{i=1}^m \mathrm{SP}\alpha_1^j|_{x_i} \subset \mathrm{SP}B_1$. Also $(\alpha_2 \otimes \mathrm{id}_{mN}) \circ \lambda_1 : A \to M_{mN}(A')$ is defined by point evaluations on the η_2-dense subset $\{x_j\}_{j=1}^m \subset X$.
Let $L = \max(L_1, L_2)$ and $n = mNL$. Define $\lambda : A \to M_n(B_2) = M_L(M_{mN}(B_2))$ by $\lambda = \mathrm{diag}(\underbrace{\lambda_1, \lambda_1, \cdots, \lambda_1}_{L})$.
Then, obviously, $\lambda \circ \alpha_1 : B_1 \to M_n(B_2)$ satisfies the condition for λ in Corollary 1.6.15, for the homotopy Ψ, positive integer L_1, and $\eta_1 > 0$. And so does $(\alpha_2 \otimes \mathrm{id}_n) \circ \lambda : A \to M_n(A')$ for Φ, L_2 and η_2.
Therefore, there are unitaries $u_1 \in M_{n+1}(B_2)$ and $u_2 \in M_{n+1}(A')$ such that

$$\|((\beta \oplus \lambda) \circ \alpha_1)(f) - u_1(\psi \oplus (\lambda \circ \alpha_1))(f)u_1^*\| < \varepsilon, \quad \forall f \in G_1,$$

$$\|(\phi \oplus ((\alpha_2 \otimes \mathrm{id}_n) \circ \lambda))(f) - u_2((\alpha_2 \otimes id_{n+1}) \circ (\beta \oplus \lambda))(f)u_2^*\| < \varepsilon, \quad \forall f \in F_1.$$

In the diagram in Lemma 1.6.24, if we replace B_2 by $M_{n+1}(B_2)$, A' by $M_{n+1}(A')$, ψ by $\mathrm{Ad}u_1 \circ (\psi \oplus (\lambda \circ \alpha_1))$, β by $\beta \oplus \lambda$, α_2 by $\mathrm{Ad}u_2 \circ (\alpha_2 \otimes id_{n+1})$, and finally ϕ by $\phi \oplus ((\alpha_2 \otimes \mathrm{id}_n) \circ \lambda)$, then we have the diagram

$$\begin{tikzcd} A \arrow[r, "\phi"] & A' \\ B_1 \arrow[u, "\alpha_1"] \arrow[r, "\psi"] & B_2 \arrow[u, "\alpha_2"'] \arrow[from=1-1, to=2-2, "\beta"] \end{tikzcd}$$

for which, the lower left triangle is approximately commutative on G_1 to within ε and the upper right triangle is approximately commutative on F_1 to within ε. Since G_1 and F_1 are arbitrary finite subsets, we proved the following main factorization result.

THEOREM 1.6.26. *Let $B_1 = \bigoplus_{j=1}^s M_{k(j)}(C(Y_j))$, where Y_j are spaces: $\{pt\}$, $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^\infty$, $\{T_{III,k}\}_{k=2}^\infty$, and S^2. Let X be a connected finite simplicial complex and let $A = M_N(C(X))$.*
Let $\alpha_1 : B_1 \to A$ be an injective homomorphism. For any finite sets $G \subset B_1$ and $F \subset A$, and for any numbers $\varepsilon > 0$ and $\delta > 0$ there exists a diagram

$$\begin{tikzcd} A \arrow[r, "\phi"] & A' \\ B_1 \arrow[u, "\alpha_1"] \arrow[r, "\psi"] & B_2 \arrow[u, "\alpha_2"'] \arrow[from=1-1, to=2-2, "\beta"] \end{tikzcd} ,$$

where
$A' = M_L(A)$, *and* B_2 *is a direct sum of matrix algebras over the spaces:* $\{pt\}$, $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^{\infty}$, $\{T_{III,k}\}_{k=2}^{\infty}$, *and* S^2;
ψ *is an injective homomorphism,* α_2 *is a unital injective homomorphism, and* ϕ *is a unital simple embedding (see 1.6.17).*
$\beta \in Map(A, B_2)$ *is* F-δ *multiplicative.*
Moreover

$$\|\psi(f) - (\beta \circ \alpha_1)(f)\| < \varepsilon, \quad \forall f \in G;$$

$$\|\phi(f) - (\alpha_2 \circ \beta)(f)\| < \varepsilon, \quad \forall f \in F.$$

Corollary 1.6.27. *Theorem 1.6.26 still holds if one replaces the injectivity condition of* α_1 *by the following condition:*
For each block B_1^j *of* B_1, *either* α_1^j *is injective or* $\alpha_1^j(B_1^j)$ *is a finite dimensional subalgebra of* A.
(Of course, one still needs to assume that $\alpha_1(\mathbf{1}_{B_1^j}) \neq 0$ *for each block* B_1^j *of* B_1 *and that at least one block of* B_2 *has spectrum different from the space of single point (equivalently,* $X \neq \{pt\}$*) , if he wants the homomorphism* ψ *to be injective.*
If one does not assume the above condition, he could still get the following dichotomy condition for ψ: *For each block* B_1^j *of* B_1 *and* B_2^k *of* B_2, *either* $\psi^{j,k}$ *is injective or* $\psi^{j,k}$ *has a finite dimensional image.)*

Proof: Write $B_1 = B' \oplus B''$ such that α_1 is injective on B' and $\alpha_1(B'') \subset A$ is of finite dimension.
Consider the finite dimensional algebra

$$D := \bigoplus_{B_1^i \subset B'} (\alpha_1(\mathbf{1}_{B_1^i}) \cdot \mathbb{C}) \bigoplus \alpha_1(B'') \quad \subset \ A.$$

By Lemma 1.6.8, if $\beta : A \to B_2$ is sufficiently multiplicative, then β is close to such a map β' that the restriction $\beta'|_D$ is a homomorphism. β' can be connected to the original β by a linear path. If the original map β is sufficiently multiplicative, then the connecting path, regarded as a map from A to $B_2[0,1]$, is F-δ multiplicative for any pre-given finite set $F \subset A$ and number $\delta > 0$. Therefore, with out loss of generality, we assume that $\beta|_D$ is a homomorphism for the original map β in 1.6.24.
By Lemma 1.6.8 again, if $\Psi : \ B_1 \to B_2[0,1]$ is sufficiently multiplicative, then Ψ is close to a map Ψ' such that $\Psi'|_{r(B_1)}$ is a homomorphism, where $r(B_1)$ is defined in 1.1.7(h). Note that $\Psi|_1 = \psi$ is a homomorphism and $(\Psi|_0)|_{r(B_1)} = \beta|_D \circ (\alpha_1|_{r(B_1)})$ is also a homomorphism. From the proof of Lemma 1.6.8, we can see that the above Ψ' can be chosen such that $\Psi'|_1 = \Psi|_1$ and $\Psi'|_0 = \Psi|_0$. Therefore, without loss of generality, we can assume that the homotopy path Ψ in Lemma 1.6.24 satisfies that $\Psi|_{r(B_1)}$ is a homomorphism.

Up to a unitary equivalence, we can further assume that $\Psi_t(\mathbf{1}_{B_1^i}) = \Psi_{t'}(\mathbf{1}_{B_1^i})$ for any $t, t' \in [0,1]$ and any block B_1^i of B_1.

One can repeat the procedure in 1.6.25 to construct the homomorphism $\lambda : A \to M_n(A)$, defined by point evaluations on an η_2-dense subset $\{x_1, x_2, \cdots, x_m\} \subset X$, to satisfy the condition that $\lambda \circ \alpha_1^j$ is defined by point evaluations on an η_1-dense subset $\bigcup_{i=1}^m \mathrm{SP}\alpha_1^j|_{x_i} \subset \mathrm{SP}B_1^j$ of sufficiently large size, for each block B_1^j of the part B'. As in 1.6.25, we can define newβ to be $\beta \oplus \lambda$. At the same time, ϕ and α_2 can also be defined as in 1.6.25. To define ψ, we need to consider two cases. For the blocks B_1^j in B', ψ can be defined as in 1.6.25, since $\lambda \circ \alpha_1^j$ is defined by point evaluations on an η_1-dense subset (of sufficiently large size). For the blocks $B_1^j \subset B''$, we define ψ to be $(\beta \oplus \lambda) \circ \alpha_1^j = (\mathrm{new}\beta) \circ \alpha_1^j$. (Note that $\beta|_{\alpha_1(B'')}$ is a homomorphism.)

$\square$

Remark 1.6.28. Once the diagram in Theorem 1.6.26 (or Corollary 1.6.27) exists for $A' = M_L(A)$, then for any $L' > L$, one can construct a diagram with the same property as in the theorem or the corollary with $A' = M_{L'}(A)$. This is easily seen from the following.

Let $r(A) = M_N(\mathbb{C})$ and $r : A \to r(A)$ be as in 1.1.7(h). Let newB_2 = old$B_2 \oplus r(A)$, newβ = old$\beta \oplus r$, newϕ = $\mathrm{diag}(\mathrm{old}\phi, \underbrace{i \circ r, \cdots, i \circ r}_{L'-L})$, new$\psi$ = old$\psi \oplus (r \circ \alpha_1)$, and new$\alpha_2$ = old$\alpha_2 \oplus \mathrm{diag}(\underbrace{i, \cdots, i}_{L'-L})$, where $i : r(A) \to A \subset M_{L'-L}(A) \subset M_{L'}(A)$ is the inclusion (note that $r(A)$ is a subalgebra of A as in 1.1.7(h)), and oldB_2, oldβ, oldϕ and oldα_2 are B_2, β, ϕ and α_2, respectively, from Lemma 1.6.26 or Corollary 1.6.27.

Corollary 1.6.29. *Let $B_1 = \bigoplus_{j=1}^s M_{k(j)}(C(Y_j))$, where Y_j are spaces: $\{pt\}$, $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^\infty$, $\{T_{III,k}\}_{k=2}^\infty$, and S^2. Let $A = \bigoplus_{j=1}^t M_{l(j)}(C(X_j))$, where X_j are connected finite simplicial complex.*

Let $\alpha_1 : B_1 \to A$ be a homomorphism satisfying the following condition:

For each pair of blocks B_1^i of B_1 and A^j of A, either the partial map $\alpha_1^{i,j}$ is injective or $\alpha_1^{i,j}(B_1^i)$ is a finite dimensional subalgebra of A^j.

For any finite sets $G \subset B_1$ and $F \subset A$, and for any numbers $\varepsilon > 0$ and $\delta > 0$, there exists a diagram

$$\begin{array}{ccc} A & \xrightarrow{\phi} & A' \\ \uparrow{\scriptstyle\alpha_1} & \searrow^{\beta} & \uparrow{\scriptstyle\alpha_2} \\ B_1 & \xrightarrow{\psi} & B_2 \end{array}$$

where

$A' = M_L(A)$, and B_2 is a direct sum of matrix algebras over the spaces: $\{pt\}$, $[0,1]$, S^1, $\{T_{II,k}\}_{k=2}^\infty$, $\{T_{III,k}\}_{k=2}^\infty$, and S^2;

ψ is a homomorphism, α_2 is a unital injective homomorphism, and ϕ is a unital simple embedding (see 1.6.17).
$\beta \in Map(A, B_2)$ is F-δ multiplicative.
Moreover,

$$\|\psi(f) - (\beta \circ \alpha_1)(f)\| < \varepsilon, \quad \forall f \in G;$$

$$\|\phi(f) - (\alpha_2 \circ \beta)(f)\| < \varepsilon, \quad \forall f \in F.$$

If we further assume that α_1 satisfies the condition that $\alpha_1^{i,j}(\mathbf{1}_{B_1^i}) \neq 0 \in A^j$ for any partial map $\alpha_1^{i,j} : B_1^i \to A^j$ of α_1, then either the homomorphism ψ could be chosen to be injective, or the spectra of all blocks of B_2 could be chosen to be the spaces of a single point.

Proof: We can construct the diagram for each block A^j of A, then put them together in the obvious way. Using Remark 1.6.28, we can assume for each block A^j, $A'_j = M_L(A_j)$ for the same L.

$\square$

The following is Lemma 4.6 of [G4] (see Lemma 1.2 of [D2]).

Lemma 1.6.30. *Let $A = \bigoplus_{i=1}^t M_{l(i)}(C(X_i))$, where X_i are connected finite simplicial complexes. Let $A' = M_L(A)$. Let the algebra $r(A)$ and the homomorphism $r : A \to r(A)$ be as in 1.1.7(h). Let B be a direct sum of matrix algebras over finite simplicial complexes of dimension at most m. Let $\phi : A \to A'$ be a unital simple embedding (see Definition 1.6.17). For any (not necessarily unital) $(m \cdot L)$-large homomorphism $\phi' : A \to B$, there is a homomorphism $\lambda : A' \oplus r(A) \to B$ such that ϕ' is homotopic to $\lambda \circ (\phi \oplus r)$.*
Furthermore, λ could be chosen to satisfy the condition that for any block B^j with $SP(B^j) \neq \{pt\}$, the partial map $\lambda^{\cdot,j} : A' \oplus r(A) \to B^j$ of λ is injective as remarked in 1.6.21.
(This lemma will be applied in conjunction with Lemma 1.6.26 or Corollary 1.6.29. From here, one can see the importance of the requirement that ϕ is a unital simple embedding.)

Remark 1.6.31. In order to apply Lemma 1.6.30 later, we would like to do one more modification for Corollary 1.6.29. Let $r : A \to r(A)$ be as in 1.1.7(h). Then the diagram in Corollary 1.6.29 could be modified to the following diagram

$$\begin{tikzcd}
A \arrow[r, "\phi\oplus r"] \arrow[dr, "\beta\oplus r"] & A' \oplus r(A) \\
B_1 \arrow[u, "\alpha_1"] \arrow[r, "\psi\oplus(r\circ\alpha_1)"'] & B_2 \oplus r(A) \arrow[u, "\alpha_2\oplus id"']
\end{tikzcd}$$

which satisfies that $\beta \oplus r$ is F-δ multiplicative and

$$\|(\psi \oplus (r \circ \alpha_1))(f) - ((\beta \oplus r) \circ \alpha_1)(f)\| < \varepsilon, \quad \forall f \in G;$$

$$\|(\phi \oplus r)(f) - ((\alpha_2 \oplus \mathrm{id}) \circ (\beta \oplus r))(f)\| < \varepsilon, \quad \forall f \in F.$$

In the application of 1.6.29 and 1.6.30 in the proof of our main reduction theorem, we will still denote $B_2 \oplus r(A)$ by B_2, $\beta \oplus r$ by β, $\psi \oplus (r \circ \alpha_1)$ by ψ, and $\alpha_2 \oplus \mathrm{id}$ by α_2. So the diagram is

$$\begin{array}{ccc} A & \xrightarrow{\phi\oplus r} & A' \oplus r(A) \\ \uparrow{\scriptstyle \alpha_1} & \searrow^{\beta} & \uparrow{\scriptstyle \alpha_2} \\ B_1 & \xrightarrow{\psi} & B_2\ . \end{array}$$

2 Spectral Multiplicity

In this section, we will show how to perturb a homomorphism $\phi : C(X) \to PM_k(C(Y))P$ in such a way that the resulting homomorphism does not have large spectral multiplicities (see 1.2.4). Namely, the following result will be proved.

Theorem 2.1. *Let X and Y be connected simplicial complexes with $X \neq \{pt\}$. Set $\dim(X) + \dim(Y) = l$. For any given finite set $G \subset C(X)$, any $\varepsilon > 0$, and any unital homomorphism $\phi : C(X) \to PM_\bullet(C(Y))P$, where $P \in M_\bullet(C(Y))$ is a projection, there is a unital homomorphism $\phi' : C(X) \to PM_\bullet(C(Y))P$ such that*
(1) $\|\phi(g) - \phi'(g)\| < \varepsilon$ for all $g \in G$, and
(2) ϕ' has maximum spectral multiplicity at most l.

2.2. Let k be a positive integer. Let $\mathrm{Hom}(C(X), M_k(\mathbb{C}))_1 = F^kX$. The space F^kX is compact and metrizable. We can endow the space F^kX with a fixed metric d as below.
Choose a finite set $\{f_i\}_{i=1}^n \subset C(X)$ which generates $C(X)$ as a C^*-algebra (e.g. one can embed X into $\mathbb{R}^n$, then choose $\{f_i\}$ to be the coordinate functions). For any $\phi, \psi \in F^kX$ which, by definition, are unital homomorphisms from $C(X)$ to $M_k(\mathbb{C})$, define

$$d(\phi, \psi) = \sum_{i=1}^{n} \|\phi(f_i) - \psi(f_i)\|.$$

Without loss of generality, we can assume that the above finite set $\{f_i\}_{i=1}^n \subset G$. On the other hand, F^kX is a finite simplicial complex (see [DN], [Se] and [Bl1]).

2.3. Let $k = \mathrm{rank}(P)$, where P is the projection in Theorem 2.1.
For any fixed y, there is a unitary $u_y \in M_\bullet(\mathbb{C})$ such that $P(y) = u_y diag(\mathbf{1}_k, 0)u_y^*$ (as in 1.2.1). Using this unitary, one can identify $P(y)M_\bullet(\mathbb{C})P(y)$ with $M_k(\mathbb{C})$ by sending $a \in P(y)M_\bullet(\mathbb{C})P(y)$ to the element in $M_k(\mathbb{C})$ corresponding to the upper left corner of $u_y^* a u_y$. (Notice that for any $a \in P(y)M_\bullet(\mathbb{C})P(y)$, the matrix

$$u_y^* a u_y = u_y^* P(y) a P(y) u_y = u_y^* P(y) u_y u_y^* a u_y u_y^* P(y) u_y$$

$$= diag(\mathbf{1}_k, 0)u_y^* a u_y diag(\mathbf{1}_k, 0)$$

has the form

$$\begin{pmatrix} (*)_{k\times k} & 0 \\ 0 & 0 \end{pmatrix}.)$$

In this way, for any $y \in Y$, the space $\mathrm{Hom}(C(X), P(y)M_\bullet(\mathbb{C})P(y))_1$ can be identified with F^kX.

Consider the disjoint union

$$\bigcup_{y\in Y} \mathrm{Hom}(C(X), P(y)M_\bullet(\mathbb{C})P(y))_1$$

as a subspace of $\mathrm{Hom}(C(X), M_\bullet(\mathbb{C})) \times Y$ with the induced topology. Using the above identification we can define a locally trivial fibre bundle

$$\begin{array}{c} \bigcup_{y\in Y} \mathrm{Hom}(C(X), P(y)M_\bullet(\mathbb{C})P(y))_1 \\ \downarrow \pi \\ Y \end{array}$$

with fibre F^kX, as shown below, where π is the natural map sending any element in the set $\mathrm{Hom}(C(X), P(y)M_\bullet(\mathbb{C})P(y))_1$ to the point y.

For simplicity, write $E_P := \bigcup_{y\in Y} \mathrm{Hom}(C(X), P(y)M_\bullet(\mathbb{C})P(y))_1$.

For any point $y_0 \in Y$, there are an open set $U \ni y$, and a continuous unitary valued function $u : U \to M_\bullet(\mathbb{C})$ such that $P(y) = u(y)diag(\mathbf{1}_k, 0)u^*(y)$. (See 1.2.1.) Let $R : M_\bullet(\mathbb{C}) \to M_k(\mathbb{C})$ be the map taking any element in $M_\bullet(\mathbb{C})$ to the $k \times k$ upper left corner of the element. Let the trivialization $t_U : \pi^{-1}(U) \to U \times F^kX$ be defined as follows. For any $\phi \in \mathrm{Hom}(C(X), P(y)M_\bullet(\mathbb{C})P(y))_1 \subset \pi^{-1}(U)$, where $y \in U$, define $t_u(\phi) = (y, \psi)$, where $\psi \in F^kX$ is defined by

$$\psi(f) = R(u^*(y)\phi(f)u(y)) \quad \text{for any} \quad f \in C(X).$$

(Again, $u^*(y)\phi(f)u(y)$ is of the form $\begin{pmatrix} (*)_{k\times k} & 0 \\ 0 & 0 \end{pmatrix}$.)

Since the set Y is compact, there is a finite cover $\mathcal{U} = \{U\}$ of Y with the above trivialization for each U. This defines a fibre bundle $\pi : E_P \to Y$.

(See §1.4 for the definition and other materials of fibre bundle.)

2.4. In the above fibre bundle, the structure group $\Gamma \subset Homeo(F^kX)$ could be chosen to be the collection of all $\gamma \in Homeo(F^kX)$ of the form: there is a unitary $u \in M_k(\mathbb{C})$ such that

$$\gamma(\phi)(f) = u^*\phi(f)u \quad \text{for any} \quad \phi \in F^kX \text{ and } f \in C(X).$$

One can see this as follows.

Suppose that U and V are two open sets in $\mathcal{U}$, and t_U and t_V are trivializations, as in 2.3, defined by unitary valued functions $u: U \to M_\bullet(\mathbb{C})$ and $v: V \to M_\bullet(\mathbb{C})$, respectively.
For any point $y \in U \cap V$, the map $t_U \circ t_V^{-1}: F^k X \to F^k X$, can be computed as below.
For any $\phi \in F^k X$, define $\tilde{\phi}: C(X) \to M_\bullet(\mathbb{C})$ by

$$\tilde{\phi}(f) = \begin{pmatrix} \phi(f)_{k\times k} & 0 \\ 0 & 0 \end{pmatrix}, \forall f \in C(X).$$

Then

$$t_V^{-1}(\phi)(f) = v(y)\tilde{\phi}(f)v^*(y) \in P(y)M_\bullet(\mathbb{C})P(y),$$

and

$$t_U \circ t_V^{-1}(\phi)(f) = R(u^*(y)v(y)\tilde{\phi}(f)v^*(y)u(y)).$$

Notice that

$$u(y)diag(\mathbf{1}_k, 0)u^*(y) = P(y) = v(y)diag(\mathbf{1}_k, 0)v^*(y).$$

It follows that $v^*(y)u(y)$ commutes with $diag(\mathbf{1}_k, 0)$. This implies that this matrix has the form

$$\begin{pmatrix} (w_1)_{k\times k} & 0 \\ 0 & w_2 \end{pmatrix},$$

where both w_1 and w_2 are unitaries. This shows that

$$t_U \circ t_V^{-1}(\phi)(f) = w_1^*\phi(f)w_1, \quad \forall \phi \in F^k X, f \in C(X).$$

In other words, $t_U \circ t_V^{-1} \in \Gamma$.
Obviously, $\mathrm{Hom}(C(X), PM_\bullet(C(Y))P)_1$ can be regarded as a collection of continuous cross sections of the bundle $\pi: E_P \to Y$.
Since for any elements $a, b \in M_k(\mathbb{C})$ and unitary $u \in M_k(\mathbb{C})$,

$$\|uau^* - ubu^*\| = \|a - b\|,$$

it is easy to see that THE METRIC d ON $F^k X$ defined in 2.2 is INVARIANT UNDER THE ACTION OF ANY ELEMENT IN Γ in the sense of 1.4.6.

2.5. There is a natural map

$$\theta: F^k X \longrightarrow P^k X$$

defined as follows. For any $\phi \in F^k X$ given by $\phi: C(X) \to M_k(\mathbb{C})$, define

$$\theta(\phi) = \mathrm{SP}(\phi) \in P^k X,$$

counting multiplicities. (See 1.2.5 and 1.2.7.)
For each point $x = [x_1, x_2, \cdots, x_k] \in P^k X$, if the element x_i appears μ_i times in x for $i = 1, 2, \cdots, k$, then the MAXIMUM MULTIPLICITY of x is defined to be

the maximum of $\mu_1, \mu_2, \cdots, \mu_k$. The maximum multiplicity of a point $\phi \in F^kX$ is defined to be the maximum multiplicity of $\theta(\phi) \in P^kX$, which agrees with the maximum multiplicity of homomorphism $\phi : C(X) \to M_k(\mathbb{C})$ defined in 1.2.4.

The homomorphism $\phi \in \mathrm{Hom}(C(X), PM_\bullet(C(Y))P)_1$ corresponds to a continuous cross section $f : Y \to E_P$. This correspondence is one to one. For any cross section $f : Y \to E_P$, any point $y \in Y$, the maximum multiplicity of $f(y)$ is understood to be that obtained by regarding $f(y)$ as an element in F^kX by an identification of $\mathrm{Hom}(C(X), P(y)M_\bullet(\mathbb{C})P(y))_1$ with F^kX. Note that the maximum multiplicity of an element $\phi \in F^kX$ is invariant under the action of any element of Γ.

2.6. It is easy to see that for any finite set $G \subset C(X)$ and $\varepsilon > 0$, there is an $\varepsilon' > 0$ such that if $d(\phi_y, \phi'_y) < \varepsilon'$ for any $y \in Y$, then $\|\phi(g) - \phi'(g)\| < \varepsilon$ for any $g \in G$, where $\phi_y, \phi'_y \in F^kX$ are determined by an identification of $\mathrm{Hom}(C(X), P(y)M_\bullet(\mathbb{C})P(y))_1$ with F^kX, as above. (Again the choice of the identification is not important, because the metric d is invariant under the action of any element in Γ.)

Before proving Theorem 2.1, we prove the following weak version of Theorem 2.1.

Lemma 2.7. *Let X and Y be as in Theorem 2.1. Let $k > l = \dim(X)+\dim(Y)$. For any given finite set $G \subset C(X)$, any $\varepsilon > 0$, and any unital homomorphism $\phi : C(X) \to PM_\bullet(C(Y))P$, where $P \in M_\bullet(C(Y))$ is a projection with rank$(P) = k$, there is a unital homomorphism $\phi' : C(X) \to PM_\bullet(C(Y))P$ such that*

(1) $\|\phi(g) - \phi'(g)\| < \varepsilon$ for all $g \in G$, and

(2) ϕ' has maximum spectral multiplicity at most $k - 1$.

Comparing with Theorem 2.1, in the above result, we allow the maximum spectral multiplicity of the resulting homomorphism to be larger than l— only require it to be smaller than $k = \mathrm{rank}(P)$—the maximum possible multiplicity. Since we assume that all the generators f_i of $C(X)$ are inside the set G, Lemma 2.7 is equivalent to the following theorem.

Lemma 2.8. *Suppose that X, Y, and P are as in Theorem 2.1 and that rank$(P) = k > \dim(X) + \dim(Y)$. For any $\varepsilon > 0$ and any cross section $f : Y \to E_P$, there is a cross section $f' : Y \to E_P$ such that*

(1) $d(f(y), f'(y)) < \varepsilon$ for all $y \in Y$, and

(2) $f'(y)$ has multiplicity at most $k - 1$ for all $y \in Y$.

To prove our main theorem of this section—Theorem 2.1, we need the following result. The proof of this result will be given after the proof of Theorem 2.1.

Theorem 2.9. *Suppose that X is a connected simplicial complex and $X \neq \{pt\}$. For any $\varepsilon > 0$ and any $x \in F^mX$, there is a contractible open neighbor-*

hood $U_x \ni x$ with $U_x \subset B_\varepsilon(x) \subset F^m X$ such that

$$\pi_i(U_x \backslash \{x\}) = 0$$

for any $0 \le i \le m-2$. In other words, $F^m X$ has property $D(m)$ as in 1.4.3.
We will use Theorem 2.9 and Corollary 1.4.10 (see also Theorem 1.4.9) to prove the following relative version of Lemma 2.8 (which gives rise to Lemma 2.8, by taking $Y_1 = \emptyset$).

Lemma 2.10. *Let X, Y, and P be as in Theorem 2.1, and $Y_1 \subset Y$ be a sub simplicial complex. Suppose that rank$(P) = k > \dim(X) + \dim(Y)$. Suppose that a cross section $f : Y \to E_P$ satisfies the condition that $f(y)$ has multiplicity at most $k-1$, for any $y \in Y_1$. It follows that for any $\varepsilon > 0$, there is a cross section $f' : Y \to E_P$ such that*
(1) $d(f(y), f'(y)) < \varepsilon$ for all $y \in Y$, and
(2) $f'(y)$ has multiplicity at most $k-1$ for all $y \in Y$.
(3) $f'(y) = f(y)$, for any $y \in Y_1$.

Proof: Let $F_1 \subset F^k X$ denote the subset of all elements of maximum multiplicity equal to k. In other words, a homomorphism in F_1 has one dimensional range. Obviously, F_1 is the set of all homomorphisms $\phi \in Hom(C(X), M_k(\mathbb{C}))_1$ which are of the form

$$\phi(f) = \begin{pmatrix} f(x) & & & \\ & f(x) & & \\ & & \ddots & \\ & & & f(x) \end{pmatrix}$$

for a certain point $x \in X$. Hence F_1 is homeomorphic to X, and $\dim(F_1) = \dim(X)$.
As mentioned in 2.5, the maximum multiplicity of an element of $F^k X$ is invariant under the action of Γ. So F_1 is an invariant subset under the action of Γ.
The conclusion of the Lemma 2.10 follows from Corollary 1.4.10 with $E_P \to Y$, $F^k X$, F_1, Y_1 and k in place of $M \to X$, F, F_1 , X_1 and m in Corollary 1.4.10, respectively. (Note that, from Theorem 2.9, $F^k X$ has property $D(k)$.)

$\square$

The above lemma is equivalent to the following lemma (we stated it with projection Q instead of P to emphasis that we may use projections other than P—we will use subprojections of P).

Lemma 2.11. *Let $Y_2 \subset Y_1 \subset Y$ be sub-simplicial complexes of Y. Let $Q \in M_\bullet(C(Y_1))$ be a projection with rank $m > l = \dim(X) + \dim(Y)$. For any given finite set $G \subset C(X)$, any $\varepsilon > 0$, and any unital homomorphism $\psi : C(X) \to QM_\bullet(C(Y_1))Q$ with the property that for any $y \in Y_2$, the multiplicity of ψ at y*

is at most $m-1$*, there is a unital homomorphism* $\psi' : C(X) \to QM_\bullet(C(Y_1))Q$ *such that*
(1) $\|\psi(g)(y) - \psi'(g)(y)\| < \varepsilon$ *for all* $g \in G$ *and* $y \in Y_1$*,*
(2) ψ' *has spectral multiplicity at most* $m-1$*,*
(3) $\psi'|_{Y_2} = \psi|_{Y_2}$*.*
In the proof of Theorem 2.1, we will not use Theorem 2.9 or Lemma 2.10 directly. We will use Lemma 2.11 instead. (So we do not need anything from fibre bundles in the rest of the proof of Theorem 2.1.)

2.12 Sketch of the idea of the proof of Theorem 2.1. Note that the proof of Lemma 2.10 can not be used to prove Theorem 2.1 (or the fibre bundle version of Theorem 2.1) in a straightforward way. For example, if we let $F_1 \subset F^k X$ be the subset of all elements with maximum multiplicity at least $l+1$ (instead of k), then $\dim(F_1)$ may be very large— much larger than $\dim(X)$. In fact, $\dim(F_1)$ also depends on k and l.
In Lemma 2.11 (or Lemma 2.10), we have already perturbed the homomorphism to avoid the largest possibility of maximum multiplicity—k. Next, we will perturb it again to avoid the next largest possibility of maximum multiplicity—$k-1$. We will continue the procedure in this way.
In general, suppose that the homomorphism $\phi : C(X) \to PM_\bullet(C(Y))P$ has maximum multiplicity m, with $l < m \leq k$, we will prove that ϕ can be approximated arbitrarily well by another homomorphism ϕ' with maximum multiplicity at most $m-1$. Once this is done, Theorem 2.1 follows from a reverse induction argument beginning with $m = k > l$. (Note that for the case $k \leq l$, we have nothing to prove.)
To do the above, we need to work simplex by simplex. In fact, on each small simplex, the homomorphism ϕ can be decomposed into a direct sum of several homomorphisms $\bigoplus_i \phi_i$, such that the projections $\phi_i(\mathbf{1})$ has rank at most m. Then we can apply Lemma 2.11 to each ϕ_i to avoid maximum multiplicity m. With these ideas in mind, it will not be difficult for the reader to construct the proof of Theorem 2.1. The complete detail will be contained in the next few lemmas, in particular, see the proof of Lemma 2.16.

Lemma 2.13. *Suppose that* P, X *and* Y *are as in Theorem 2.1. For any* $\varepsilon > 0$ *and any positive integer* d*, there is a* $\delta > 0$ *such that if* $\Delta \subset Y$ *is a simplex of dimension* d*, if* P *is regarded as a projection in* $M_\bullet(C(\Delta))$*, and if* $\psi : C(X) \to PM_\bullet(C(\partial\Delta))P$ *is a homomorphism such that*

$$\|\psi(g)(y) - \psi(g)(y')\| < \delta, \quad \forall g \in G, y, y' \in \partial\Delta,$$

then there is a homomorphism $\psi' : C(X) \to PM_\bullet(C(\Delta))P$ *such that*
(1) $\|\psi'(g)(y) - \psi'(g)(y')\| < \varepsilon, \quad \forall g \in G, y, y' \in \Delta$*, and*
(2) $\psi'|_{\partial\Delta} = \psi$*.*

Proof: Note that $P|_\Delta$ is a trivial projection, since a simplex Δ is contractible.

So $PM_\bullet(C(\Delta))P \cong M_k(C(\Delta))$, where $k = \mathrm{rank}(P)$. The lemma follows from the fact that

$$F^k X = \mathrm{Hom}(C(X), M_k(\mathbb{C}))_1$$

is a simplicial complex (see [DN] and [Bl]), which is locally contractible (see 1.4.2 and 1.4.3).

$\Box$

We need the following lemma, which is obviously true.

Lemma 2.14. *Suppose that $\phi : C(X) \to PM_\bullet(C(Y))P$ has maximum spectral multiplicity at most m. Then there exist $\eta > 0$ and $\delta > 0$ such that the following statement holds.*
For any subset $Z \subset Y$ with diameter$(Z) < \eta$, and homomorphism $\psi : C(X) \to PM_\bullet(C(Z))P$ with the property that

$$\|\psi(g)(z) - \phi(g)(z)\| < \delta, \quad \forall z \in Z, g \in G,$$

there is a decomposition of ψ such as described below.
There are open sets $O_1, O_2, \cdots, O_t \subset X$, with mutually disjoint closures (i.e., $\bar{O}_i \cap \bar{O}_j = \emptyset$, $\forall i \neq j$), and there are mutually orthogonal projections $Q_1, Q_2, \cdots, Q_t \in M_\bullet(C(Z))$ and homomorphisms $\psi_i : C(X) \to Q_i M_\bullet(C(Z))Q_i$ such that
1. $\psi = \sum_{i=1}^t \psi_i$
2. $P(z) = \sum_{i=1}^t Q_i(z), \forall z \in Z$,
3. rank$(Q_i) \leq m$, and
4. $SP\psi_i \subset O_i$ for all i.

Proof: One can prove it using the following fact. Suppose that $\mathrm{SP}\phi \subset \cup O_i$. If ψ is close enough to ϕ, then $\mathrm{SP}\psi \subset \cup O_i$ and $\#(\mathrm{SP}\phi \cap O_i) = \#(\mathrm{SP}\psi \cap O_i)$ (see 1.2.12).
Q_i in the lemma should be chosen to be the spectral projections of ψ corresponding to the open sets O_i (see 1.2.9).
(Notice that Y is compact and that $G \subset C(X)$ contains a set of the generators.)

$\Box$

Lemma 2.15. *Suppose that $\phi : C(X) \to PM_\bullet(C(Y))P$ has maximum spectral multiplicity at most m. Then there exist $\eta > 0$ and $\delta > 0$ such that the following statement holds.*
For any $\varepsilon > 0$, any simplex $\Delta \subset Y$ (of any simplicial decomposition of Y) with diameter$(\Delta) < \eta$, and any homomorphism $\psi : C(X) \to PM_\bullet(C(\Delta))P$ with the following properties:
(i) $\|\psi(g)(z) - \phi(g)(z)\| < \delta$, $\forall z \in \Delta, g \in G$, and
(ii) $\psi|_{\partial\Delta}$ has maximum multiplicity at most $m-1$,
there exists a homomorphism $\psi' : C(X) \to PM_\bullet(C(\Delta))P$ such that
(1) $\|\psi(g)(y) - \psi'(g)(y)\| < \varepsilon$ for all $g \in G$ and $y \in \Delta$;
(2) ψ' has spectral multiplicity at most $m-1$;

(3) $\psi'|_{\partial\Delta} = \psi|_{\partial\Delta}$.

Proof: . Suppose that η and δ are as in Lemma 2.14. If ψ is as described in this lemma, then one can obtain the decomposition $\psi = \sum_{i=1}^{t} \psi_i$ of ψ as in Lemma 2.14.
Then we only need to apply Lemma 2.11 to each map ψ_i to obtain $\psi_i' : C(X) \to Q_i M_\bullet(C(Z))Q_i$ to satisfy the conclusion of Lemma 2.11 with ψ_i, Δ, $\partial\Delta$, and Q_i in place of ψ, Y_1, Y_2, and Q, respectively.
If ε is small enough, then $\mathrm{SP}\psi_i' \subset O_i$, where the open sets O_i are from Lemma 2.14. Hence the sum $\psi' = \sum_{i=1}^{t} \psi_i'$ is as desired.

□

LEMMA 2.16. *Suppose that $\phi : C(X) \to PM_\bullet(C(Y))P$ has maximum spectral multiplicity at most $m > l = \dim(X) + \dim(Y)$. For any simplicial subcomplex $Y_1 \subset Y$, with respect to any simplicial decomposition of Y, and any $\varepsilon > 0$, there is a homomorphism $\phi' : C(X) \to PM_\bullet(C(Y_1))P$ with multiplicity at most $m-1$ such that*

$$\|\phi'(g)(y) - \phi(g)(y)\| < \varepsilon, \quad \forall g \in G, y \in Y_1.$$

(In particular, the above is true for $Y_1 = Y$.)

Proof: We will prove the lemma by induction on $\dim(Y_1)$.
If $\dim(Y_1) = 0$, the lemma follows from the fact that, for a connected simplicial complex X with $X \neq \{pt\}$, the subset of homomorphisms with distinct spectrum (maximum spectral multiplicity one) is dense in
$\mathrm{Hom}(C(X), M_k(\mathbb{C}))_1$.
Suppose that the lemma is true for any simplicial subcomplex of dimension n, with respect to any simplicial decomposition.
Let $Y_1 \subset Y$ be a simplicial complex of dimension $n + 1 \leq \dim Y$, with respect to some simplicial decomposition of Y.
Let $\varepsilon > 0$.
Let δ_1, η_1 be the δ and η of Lemma 2.15.
Apply Lemma 2.13 with $n+1$ in place of d, and $\frac{1}{4}\min(\varepsilon, \delta_1)$ in place of ε, to find δ_2 as δ in the lemma.
Choose $\eta_2 > 0$ such that if $\mathrm{dist}(y, y') < \eta_2$, then

$$(*) \qquad \|\phi(g)(y) - \phi(g)(y')\| < \frac{1}{4}\min(\varepsilon, \delta_1, \delta_2), \forall g \in G.$$

Endow Y_1 with a simplicial complex structure such that diameter$(\Delta) < \min(\eta_1, \eta_2)$ for any simplex Δ of Y_1. Let $Y' \subset Y_1$ be the n-skeleton of Y_1 with respect to the simplicial structure.
From the inductive assumption, there is a homomorphism $\phi_1 : C(X) \to PM_\bullet(C(Y'))P$, with multiplicity at most $m-1$, such that

$$(**) \qquad \|\phi_1(g)(y) - \phi(g)(y)\| < \frac{1}{4}\min(\varepsilon, \delta_1, \delta_2), \quad \forall g \in G \text{ and } y \in Y'.(**)$$

Consider a fixed simplex $\Delta \subset Y_1$ of top dimension (i.e., $\dim(\Delta) = n+1$). Let us extend $\phi_1|_{\partial\Delta}$ to Δ (notice that $\partial\Delta \subset Y'$).
For any two points $y, y' \in \partial\Delta$, applying (*) to the pair of points y and y', applying (**) to the points y and y' separately, and combining all these three inequalities together, we get

$$\|\phi_1(g)(y) - \phi_1(g)(y')\| < \frac{3}{4}\min(\varepsilon, \delta_1, \delta_2) < \delta_2, \quad \forall g \in G.$$

By Lemma 2.13 and the way δ_2 is chosen, there is a homomorphism (let us still denote it by ϕ_1) $\phi_1 : C(X) \to PM_\bullet(C(\Delta))P$, which extends the original $\phi_1|_{\partial\Delta}$, such that

$$(***) \qquad \|\phi_1(g)(y) - \phi_1(g)(y')\| < \frac{1}{4}\min(\varepsilon, \delta_1), \quad \forall g \in G \quad \text{and} \quad y, y' \in \Delta.$$

For any point $y \in \Delta$, choose a point $y' \in \partial\Delta$. Applying both (*) and (***) to the pair (y, y'), applying (**) to the point $y' \in \partial\Delta \subset Y'$ and combining all these three inequalities together, we get

$$(****) \qquad \|\phi_1(g)(y) - \phi(g)(y)\| < \frac{3}{4}\min(\varepsilon, \delta_1), \forall g \in G \text{ and } y \in \Delta.$$

Since δ_1 and η_1 are chosen as in Lemma 2.15, and diameter$(\Delta) < \eta_1$, it follows from $(* * **)$ and Lemma 2.15 that there is a homomorphism $\phi' : C(X) \to PM_\bullet(C(\Delta))P$ such that
(1) $\|\phi'(g)(y) - \phi_1(g)(y)\| < \frac{1}{4}\varepsilon, \quad \forall g \in G$ and $y \in \Delta$.
(2) ϕ' has spectral multiplicity at most $m-1$.
(3) $\phi'|_{\partial\Delta} = \phi_1|_{\partial\Delta}$.
Combining (1) above with (****), yields

$$\|\phi'(g)(y) - \phi(g)(y)\| < \frac{3}{4}\min(\varepsilon, \delta_1) + \frac{1}{4}\varepsilon \le \varepsilon, \forall g \in G \text{ and } y \in \Delta.$$

Carry out the above construction independently for each simplex Δ. Since the definition of ϕ' on $\partial\Delta$ is as same as ϕ_1, the definitions of ϕ' on different simplices are agree on their intersection. By Lemma 1.2.14, this yields a homomorphism over the whole set Y_1. The lemma follows.

□

Obviously, Theorem 2.1 follows from Lemma 2.16 by reverse induction argument beginning with $m = k$. (Note that we only need Lemma 2.16 for the case $Y_1 = Y$.)
Now we are going to prove Theorem 2.9, which is the only missing part in the proof of Theorem 2.1. The proof is somewhat similar to the proof of Theorem 6.4.2 of [DN]. It will therefore be convenient to recall some of the terminology and notation of [DN]. (It will be important to consider a certain method of decomposing the space F^kX.)

2.17. Recall from 6.17 of [DN] (cf. 1.2.4 above) that there is a map λ :

$X^k \times U(k) \to F^k X$, defined as follows. If $u \in U(k)$ and $(x_1, x_2, \cdots, x_k) \in X^k$, then

$$(\lambda(x_1, x_2, \cdots, x_k, u))(f) = u \begin{pmatrix} f(x_1) & & & \\ & f(x_2) & & \\ & & \ddots & \\ & & & f(x_k) \end{pmatrix} u^*$$

for any $f \in C(X)$. Since λ is surjective, $F^k X$ can be regarded as a quotient space $X^k \times U(k)$. Therefore, for convenience, a point in $F^k X$ will be written as

$$[x_1, x_2, \cdots, x_k, u]$$

which means $\lambda(x_1, x_2, \cdots, x_k, u)$.

With the above notation, it is easy to see that, if X is path connected and is not a single point, then any element in $F^k X$ can be approximated arbitrarily well by elements in $F^k X$ with distinct spectra.

2.18. If $X_1 \subset X$ is a subset, then define $F^k X_1$ to be the subset of $F^k X$ consisting of those homomorphisms $\phi \in \mathrm{Hom}(C(X), M_k(\mathbb{C}))$ with $\mathrm{SP}(\phi) \subset X_1$ as a set. Obviously, if X_1 is open (closed resp.), then $F^k X_1$ is open (closed resp.).

If $X_1, X_2, \cdots, X_i$ are disjoint subspaces of X, and $k_1, k_2, \cdots, k_i$ are nonnegative integers with

$$k_1 + k_2 + \cdots + k_i = k,$$

then define $F^{(k_1, k_2, \cdots, k_i)}(X_1, X_2, \cdots, X_i)$ to be the subset of $F^k X$ consisting of all ϕ with

$$\#(\mathrm{SP}(\phi) \cap X_i) = k_i$$

counting multiplicity.

Usually when we use the above notation, we suppose that $\bar{X}_{i_1} \cap \bar{X}_{i_2} = \emptyset$ if $i_1 \neq i_2$, where $\bar{X}_i$ is the closure of X_i. In this case,

$$F^k(X_1 \cup X_2 \cup \cdots \cup X_i) = \coprod_{k_1 + k_2 + \cdots + k_i = k} F^{(k_1, k_2, \cdots, k_i)}(X_1, X_2, \cdots, X_i)$$

is a disjoint union of separate components.

2.19. For each i-tuple $(k_1, k_2, \cdots, k_i)$ with

$$k_1 + k_2 + \cdots + k_i = k,$$

one can define $G_{(k_1, k_2, \cdots, k_i)}(\mathbb{C}^k)$ to be the collection of i-tuples $(p_1, p_2, \cdots, p_i)$ of orthogonal projections $p_j \in M_k(\mathbb{C})$ with $\mathrm{rank}(p_j) = k_j$ $(1 \leq j \leq i)$ and $\sum_{j=1}^{i} p_j = \mathbf{1} \in M_k(\mathbb{C})$. Note that if $i = 2$, $G_{(k_1, k_2)}(\mathbb{C}^k)$ is the ordinary complex Grassmannian manifold $G_{k_1}(\mathbb{C}^k) = G_{k_2}(\mathbb{C}^k)$.

For each fixed i-tuple $(k_1, k_2, \cdots, k_i)$, there is a locally trivial fibre bundle

$$\begin{array}{ccc} F^{k_1}(X_1) \times F^{k_2}(X_2) \times \cdots \times F^{k_i}(X_i) & \longrightarrow & F^{(k_1,k_2,\cdots,k_i)}(X_1, X_2, \cdots, X_i) \\ & & \downarrow \\ & & G_{(k_1,k_2,\cdots,k_i)}(\mathbb{C}^k). \end{array}$$

2.20. For certain purposes, it is more convenient to use CW complexes (instead of simplicial complexes).

For the terminology used below, see [Wh].

Suppose that (X, A) is a relative CW complex pair. If X is path connected, then (X, A) is zero connected CW complex pair, no matter A is connected or not. In particular, (X, A) is homotopy equivalent to (X_1, A), where X_1 is obtained from A by attaching finitely many cells of dimension ≥ 1 (see Theorem 2.6 of Chapter five of [Wh]). This can not be done if one only uses simplicial complex pair. (Note, we always assume our CW complexes to be finite CW complexes without saying so.)

For a relative CW complex pair (X, A), define $F^m_A X \subset F^m X$ to be the subspace consisting of those elements $x \in F^m X$, with

$$\mathrm{SP}(x) \cap A \neq \emptyset.$$

(This is different from the set $F^m A$ (defined in 2.18) which consists of elements $x \in F^m X$ such that $\mathrm{SP}(x) \subset A$.)

Lemma 2.21. *Suppose that (X, A) is a relative CW complex pair. Suppose that X is obtained from A by attaching cells of dimension at least* 1. *It follows that the inclusion*

$$F^m_A X \hookrightarrow F^m X$$

is $m-1$ equivalent, i.e., $i_* : \pi_j(F^m_A X) \to \pi_j(F^m X)$ is an isomorphism for any $0 \leq j \leq m-2$ and a surjection for $j = m-1$, where i_* is induced by the inclusion map.

The proof of this lemma is divided into two steps.

Lemma 2.22. *Lemma 2.21 is true if X is obtained from A by attaching several cells of dimension 1.*

Proof: Let $X = A \cup e_1 \cup e_2 \cup \cdots \cup e_t$, where $e_1, e_2, \cdots, e_t$ are 1-cells with $\partial e_i \subset A$. Then $F^m X \backslash F^m_A X$ consists of those points whose spectra are contained in

$$\overset{\circ}{e}_1 \cup \overset{\circ}{e}_2 \cup \cdots \cup \overset{\circ}{e}_t,$$

where each $\overset{\circ}{e}_j = e_j \backslash \partial e_j$ is homeomorphic to $(0,1)$.

In other words,

$$F^m X \backslash F^m_A X = \coprod_{k_1+k_2+\cdots+k_t=m} F^{(k_1,k_2,\cdots,k_t)}(\overset{\circ}{e}_1, \overset{\circ}{e}_2, \cdots, \overset{\circ}{e}_t).$$

For each fixed t-tuple $(k_1, k_2, \cdots, k_t)$ with $\sum k_i = m$, the space $F^{(k_1,k_2,\cdots,k_t)}(\overset{\circ}{e}_1, \overset{\circ}{e}_2, \cdots, \overset{\circ}{e}_t)$ is a smooth manifold. To see this, we can consider the fibre bundle

$$\begin{array}{ccc} F^{k_1}(\overset{\circ}{e}_1) \times F^{k_2}(\overset{\circ}{e}_2) \times \cdots \times F^{k_t}(\overset{\circ}{e}_t) & \longrightarrow & F^{(k_1,k_2,\cdots,k_t)}(\overset{\circ}{e}_1, \overset{\circ}{e}_2, \cdots, \overset{\circ}{e}_t) \\ & & \downarrow \\ & & G_{(k_1,k_2,\cdots,k_t)}(\mathbb{C}) \end{array}$$

introduced in 2.19. Evidently, the fibre of the bundle is

$$F^{k_1}(\overset{\circ}{e}_1) \times F^{k_2}(\overset{\circ}{e}_2) \times \cdots \times F^{k_t}(\overset{\circ}{e}_t) \cong \mathbb{R}^{k_1^2} \times \mathbb{R}^{k_2^2} \times \cdots \times \mathbb{R}^{k_t^2}.$$

Note that the above fibre bundle has an obvious cross section (see [DN]). Therefore, the fibre bundle can be regarded as a smooth vector bundle with the vector space $\mathbb{R}^{k_1^2+k_2^2+\cdots+k_t^2}$ as the fibre. The zero section of the bundle has codimension

$$k_1^2 + k_2^2 + \cdots + k_t^2 \geq k_1 + k_2 + \cdots + k_t = m.$$

By a standard argument from topology, using the transversality theorem, the lemma can be proved. (See [DN, 6.3.4] for details.)

$\square$

2.23. The next step is to prove Lemma 2.21 by induction, starting with 2.22. Since the proof is a complete repetition of the proof of Theorem 6.4.2 of [DN], we omit the detail—only point out how to define the collection $W(n, r)$ in our setting, and several small modifications.
Let $W(1, 0) = \{A\}$—the set with single element: the space A. For $r > 0$, $W(n, r)$ is the class of all finite CW complexes, each of which is obtained by attaching at most one cell of dimension n to a space in $W(n, r - 1)$. Let

$$W(n + 1, 0) = \bigcup_{r=0}^{+\infty} W(n, r).$$

Lemma 2.22 says that $F_A^m X \to F^m X$ is $m - 1$ equivalent if $X \in W(2, 0)$. In applying the argument in [DN, 6.4.2], $F_A^m X$ is in place of $F^k(X)$ and $F^m X$ is in place of $F^{k+1}(X)$. All the other parts of the proof follow from [DN]. The only thing needs mentioning is that the inclusion

$$F_A^m(X \backslash \alpha_I) \hookrightarrow F_A^m X$$

is 1-equivalent, where α_I is a set of finitely many points inside one of the n-cells of X, and $n \geq 2$. To prove this statement, one needs to prove that any continuous map from S^1 to $F_A^m X$ can be perturbed to a map from S^1 to $F_A^m(X \backslash \alpha_I)$. To do this, he can first perturb a map to a piecewise linear map for which the image will be one dimensional. And the resulting map can be easily perturbed again to a map whose spectrum avoids any given finite set of points in any cell of dimension at least 2.

$\square$

2.24. Let X be a simplicial complex, and $x_0 \in X$ be a vertex.
Define X' to be the sub-simplicial complex consisting of all the simplices Δ (and their faces) with $\Delta \ni x_0$. Then X' is contractible.
We also use x_0 to denote the point in $F^m X$ defined by

$$\phi(f) = f(x_0) \cdot \mathbf{1}_m \in M_m(\mathbb{C})$$

for each $f \in C(X)$.
We can easily prove the following claim: The map $F^m X' \backslash \{x_0\} \hookrightarrow F^m X'$ is $(m-1)$ equivalent.
To see this, let

$$A = \cup\{\Delta \mid \Delta \text{ is a simplex of } X' \text{ and } x_0 \notin \Delta\}.$$

Then A is a sub-simplicial complex of X'. A may not be connected, but (X', A) is 0-connected. In the notation of 1.4.2,

$$X' = \overline{\mathrm{Star}(\{x_0\})} \quad \text{and} \quad A = \overline{\mathrm{Star}(\{x_0\})} \backslash \mathrm{Star}(\{x_0\}).$$

It is obvious that $A \hookrightarrow X' \backslash \{x_0\}$ is a homotopy equivalence. Therefore, $F_A^m X' \hookrightarrow (F^m X') \backslash \{x_0\}$ is a homotopy equivalence. By Lemma 2.21, the claim holds. In particular,

$$\pi_i(F^m X' \backslash \{x_0\}) = 0$$

for any $0 \leq i \leq m-2$, since $F^m X'$ is contractible. Equivalently,

$$\pi_i(F^m (X' \backslash A) \backslash \{x_0\}) = 0$$

for any $0 \leq i \leq m-2$. Note that $X' \backslash A$ is an open neighborhood of $x_0 \in X$, which is the interior of X'.

2.25. PROOF OF THEOREM 2.9. Suppose that

$$\mathrm{SP}(x) = \{ \underbrace{\lambda_1, \lambda_1, \cdots, \lambda_1}_{k_1}, \underbrace{\lambda_2, \lambda_2, \cdots, \lambda_2}_{k_2}, \cdots, \underbrace{\lambda_i, \lambda_i, \cdots, \lambda_i}_{k_i} \},$$

where $\lambda_1, \lambda_2, \cdots, \lambda_i \in X$ are distinct points and $k_1 + k_2 + \cdots + k_i = m$. Choose mutually disjoint open sets $U_1 \ni \lambda_1, U_2 \ni \lambda_2, \cdots, U_i \ni \lambda_i$, in X. Then there is a locally trivial fibre bundle

$$\begin{array}{ccc} F^{k_1}(U_1) \times F^{k_2}(U_2) \times \cdots \times F^{k_i}(U_i) & \longrightarrow & F^{(k_1, k_2, \cdots, k_i)}(U_1, U_2, \cdots, U_i) \\ & & \downarrow \\ & & G_{(k_1, k_2, \cdots, k_i)}(\mathbb{C}). \end{array}$$

Note that $G_{(k_1,k_2,\cdots,k_i)}(\mathbb{C}) = U(m)/(U(k_1)\times U(k_2)\times\cdots\times U(k_i))$ is a smooth manifold of dimension $t := m^2 - \sum_{j=1}^{i} k_j^2$. There is a small contractible neighborhood $U_x \subset B_\varepsilon(x)$ which is homeomorphic to the space

$$F^{k_1}(X_1)\times F^{k_2}(X_2)\times\cdots F^{k_i}(X_i)\times\mathbb{R}^t,$$

where $X_1, X_2, \cdots, X_i$ are mutually disjoint open subsets of X. The space X_i can be chosen so that $\bar{X}_i$ is the simplicial complex X' as in 2.24 corresponding to vertex λ_i, with respect to some simplicial decomposition of X.
The following fact is well known in topology. Suppose that X and Y are connected CW complexes with base points x_0, y_0 respectively. If $X\backslash\{x_0\}$ is l_1-connected and $X\backslash\{y_0\}$ is l_2-connected, then $(X\times Y)\backslash\{x_0,y_0\}$ is (l_1+l_2+2)-connected.
Combining this fact with 2.24, we conclude that $U_x\backslash\{x\}$ is

$$\sum_{j=1}^{i} k_j - 2 + t = m - 2 + t$$

connected. This ends the proof.

$\square$

3 Combinatorial Results

In this section, we will prove certain results of a combinatorial nature, for the preparation of the proof of the Decomposition Theorem—Theorem 4.1 of the next section.
We will need the results in the case that certain multiplicities are general—not just equal to one. For the sake of clarity, we will first state and prove the results in the special case of multiplicity one. We will then consider the general case.

3.1. Suppose that X is a simplicial complex. Let σ denote the simplicial complex structure of X—which tells what the simplices of X are, and what the faces of each simplex of X are. In this section, we will use (X,σ) to denote the simplicial complex X with simplicial structure σ, to emphasize that we may endow the same space X with different simplicial complex structures. In this section, we will reserve the notation, $\sigma, \tau, \sigma_1, \tau_1, \cdots$, etc., for simplicial complex structures.
Recall that, if Δ is a simplex, its boundary is denoted by $\partial\Delta$. For example, if $\dim(\Delta) = 0$, i.e., $\Delta = \{pt\}$, the set consisting of a single point, then $\partial\Delta = \emptyset$; if $\dim(\Delta) = 1$, i.e., Δ is an interval, then $\partial\Delta$ is the set consisting of the two extreme points of the interval. Let us also consider the set $\Delta\backslash\partial\Delta$, and denote it by interior(Δ).
If (X,σ) is a simplicial complex, then for any point $x\in X$, there is a unique simplex Δ such that $x\in$ interior(Δ).

As usual, if each simplex of (X, σ_1) is a union of certain simplices of (X, σ_2), then we shall call σ_2 a subdivision of σ_1. This is equivalent to the property that any simplex of (X, σ_2) is contained in a simplex of (X, σ_1).
The notation $\mathrm{Vertex}(X, \sigma)$ (respectively, $\mathrm{Vertex}(\Delta)$) will be used to denote the set of vertices of (X, σ) (or of the simplex Δ).

Definition 3.2. Let $E = \{1, 2, \cdots, K\}$ be an index set. (The index set E can be any set with exactly K elements.) Let $K_1, K_2, \cdots, K_m$ be non negative integers with

$$K_1 + K_2 + \cdots + K_m = K.$$

A grouping of E of type $(K_1, K_2, \cdots, K_m)$ is a collection of m mutually disjoint index sets $E_1, E_2, \cdots, E_m$ with

$$E = E_1 \cup E_2 \cup \cdots \cup E_m,$$

and $\#(E_j) = K_j$ for each $1 \le j \le m$. (Cf. 1.5.18.)
Usually, we will keep the tuple $(K_1, K_2, \cdots, K_m)$ fixed and just call the collection $E_1, E_2, \cdots, E_m$ a grouping of E (without mentioning the type).
(Most of the time, $K_1, K_2, \cdots K_m$ will be positive integers, i.e., nonzero. But for convenience, we allow some numbers $K_i = 0$, and then the corresponding sets E_i should be the empty set.)

3.3. Let (X, σ) be a simplicial complex. Suppose that, associated to each vertex $x \in X$, there is a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ of E of type $(K_1, K_2, \cdots, K_m)$. (In our application in the proof of Theorem 4.1, the index set E will be the spectrum of a homomorphism at the given vertex, see 1.5.13, 1.5.17–1.5.22.)
Suppose that these groupings for all the vertices are chosen arbitrarily. Then, in general, for a simplex Δ with vertices $x_0, x_1, \cdots, x_n$, the intersections

$$\bigcap_{x \in \mathrm{Vertex}(\Delta)} E_j(x) = E_j(x_0) \cap E_j(x_1) \cap \cdots \cap E_j(x_n),\ \ j = 1, 2, \cdots, m,$$

may have very few elements—the sets $E_j(x_0), E_j(x_1), \cdots, E_j(x_n)$ may be very different.
The purpose of this section is to introduce a subdivision (X, τ) of (X, σ), and to associate to each new vertex of (X, τ) a grouping to make the following true: For any simplex Δ of (X, τ) (after the subdivision), for each j, the number of elements in the intersection

$$\bigcap_{x \in \mathrm{Vertex}(\Delta)} E_j(x)$$

is not much less than the number of elements in each individual set $E_j(x)$ (note that $\#(E_j(x)) = K_j$ for each x); in other words, the groupings of adjacent vertices (after subdivision) should be almost as the same as each other.

First we will state the following lemma (the proof will be given in 3.15). Later on, we will need a relative version of the lemma.
(See 1.5.17 to 1.5.23 for the explanations of the role of this lemma in §4. To visualize the following lemma, see 1.5.21 for the explanation of the one dimensional case.)

Lemma 3.4. *Let (X,σ) be a simplicial complex consisting of a single simplex X and its faces. Suppose that associated to each $x \in Vertex(X,\sigma)$, there is a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ of E (of type $(K_1, K_2, \cdots, K_m)$).*
It follows that there is a subdivision (X,τ) of (X,σ), and associated to each new vertex $x \in Vertex(X,\tau)$, there is a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ of E (of type $(K_1, K_2, \cdots, K_m)$), (for any old vertex of (X,σ), the grouping should not be changed), such that the following hold.
For each newly introduced vertex $x \in Vertex(X,\tau)$,

$$(1) \qquad \bigcap_{y \in Vertex(X,\sigma)} E_j(y) \subset E_j(x), \qquad j = 1, 2, \cdots, m,$$

and

$$(2) \qquad E_j(x) \subset \bigcup_{y \in Vertex(X,\sigma)} E_j(y), \qquad j = 1, 2, \cdots, m.$$

For any simplex Δ of (X,τ) (after subdivision),

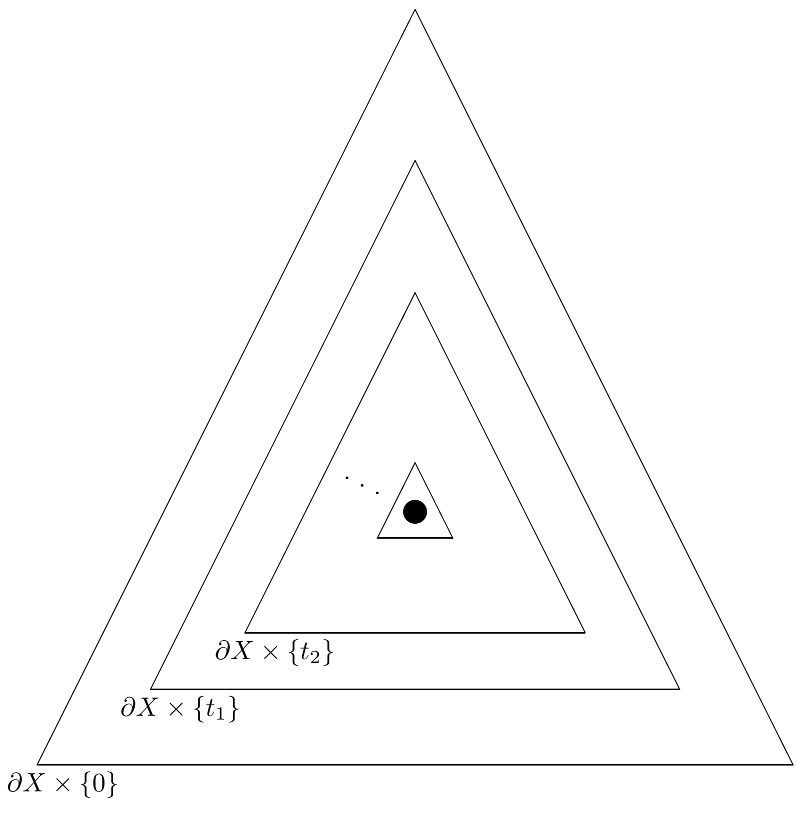
$$(3) \qquad \#\Big(\bigcap_{x \in Vertex(\Delta)} E_j(x)\Big) \geq K_j - \frac{n(n+1)}{2}, \qquad j = 1, 2, \cdots, m,$$

where $n = \dim X$.
(When we apply this lemma in §4, the simplex X will be a simplex of a simplicial complex Y, and $K_j >> (\dim Y)^3$; from this it follows that

$$\#\Big(\bigcap_{x \in \mathrm{Vertex}(\Delta)} E_j(x)\Big) \geq K_j - \frac{n(n+1)}{2} >> (\dim Y)^3, \qquad j = 1, 2, \cdots, m.)$$

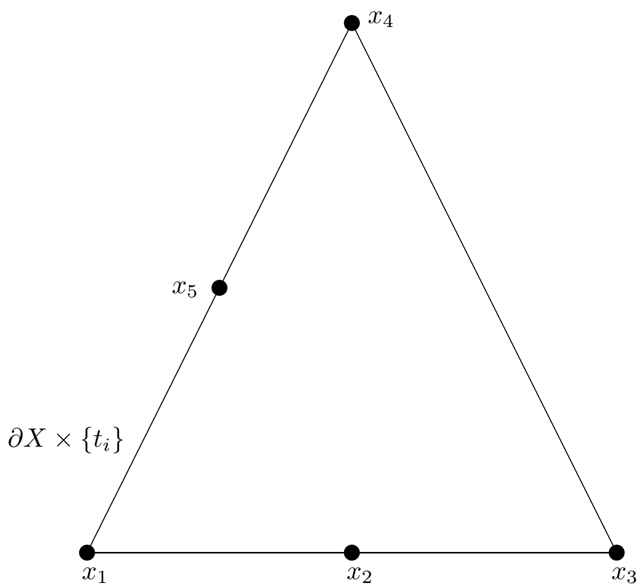
Remark 3.5. The inclusion $E_j(x) \subset \bigcup_{y \in \mathrm{Vertex}(X,\sigma)} E_j(y)$ in the condition (2) of Lemma 3.4 is important for our application in §4 (see 1.5.19 for the explanation). We will put this inclusion into a more general context, 3.7. So we will only discuss the condition (1) in this remark.
The inclusion $\bigcap_{y \in \mathrm{Vertex}(X,\sigma)} E_j(y) \subset E_j(x)$ in the condition (1) above, will not be used in our application in §4. But taking this inclusion as a part of the conclusion will make the induction argument easier in the proof.
We would like to point out that the weak version of the above lemma without requiring the inclusion in (1) will automatically imply the above stronger version. (This can be seen from the proof of Corollary 3.14 below.)

3.6. Strategy and logistics of the proof of Lemma 3.4
We shall prove it by induction. So we assume that the lemma is true for simplex of dimension at most $n-1$, and prove the case that $\dim(X)=n$.
First, we introduce a new vertex, which is the barycenter of X, and introduce a model grouping $E_1^{model}, E_2^{model}, \cdots, E_m^{model}$ of E for the new vertex.
We shall view X as many layers similar to the boundary ∂X: $\partial X\times\{0\}, \partial X\times\{t_1\}, \cdots,$ and the top layer $\partial X\times\{1\}$ is identified into a single point which is the barycenter, where $t_1, t_2, \cdots,$ is a finite sequence of increasing numbers between 0 and 1 (the number of terms in this sequence depending in a certain sense on the distance between the giving groupings at the vertices and the model grouping at the barycenter). See the picture below.



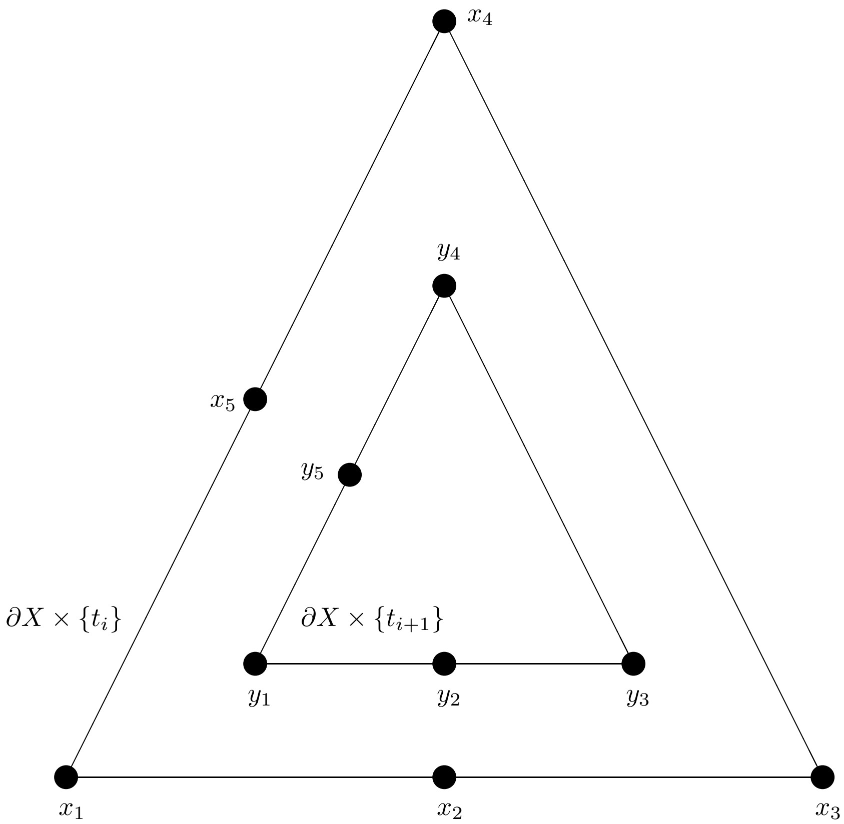
We will introduce a subdivision of each layer $\partial X\times\{t_i\}$ (identifying the layer as a set with ∂X and thereby endow it with a simplicial complex structure), and a grouping for each vertex on this layer. The general principle we shall follow

is: the higher the layer is, the closer the groupings are to the model grouping. We should, gradually, change the groupings from each layer to the next higher layer.
Let us explain it for the case $\dim(X) = 2$ and $\dim(\partial X) = 1$. Fix a t_i, and suppose that we have the simplicial structure and groupings for all vertices on $\partial X \times \{t_i\}$. Let us using the following picture to show the vertices of $\partial X \times \{t_i\}$.



(I.e., there are five 1-dimensional simplices $[x_1, x_2], [x_2, x_3], [x_3, x_4], [x_4, x_5]$ and $[x_5, x_1]$.)
Let us assume that the condition (3) holds for any simplex of $\partial X \times \{t_i\}$ with $\dim(X)$ replaced by $\dim(\partial X)$. (We will also discuss the condition (1) below, but not the condition (2).)
We shall construct simplicial structure and groupings on $\partial X \times \{t_{i+1}\}$. To begin with, let us provisionally define the simplicial structure on $\partial X \times \{t_{i+1}\}$ to be as the same as that on $\partial X \times \{t_i\}$, as in the picture on the next page.
Fix an element $\lambda \in E_1^{model}$ such that $\lambda \notin \bigcap_{k=1}^{5} E_1(x_k)$. (If such an element does not exist, then the groupings are already good for E_1. In other words, $E_1(x_k)$ contains and therefore equals E_1^{model} for every k. Then we should go on to E_2 or other parts.)
The grouping on the vertex $y_j, j = 1, \cdots, 5$ will be taken to be either the grouping on the corresponding vertex x_j, if $E_1(x_j) \ni \lambda$, or the grouping on the corresponding vertex x_j, with a certain element of $E_1(x_j) \backslash E_1^{model}$ replaced by $\lambda \in E_1^{model}$ if $E_1(x_j) \not\ni \lambda$. Lemma 3.9 below tells which element should be chosen to be replaced. Of course the other part E_t, $t > 1$ of the grouping must also

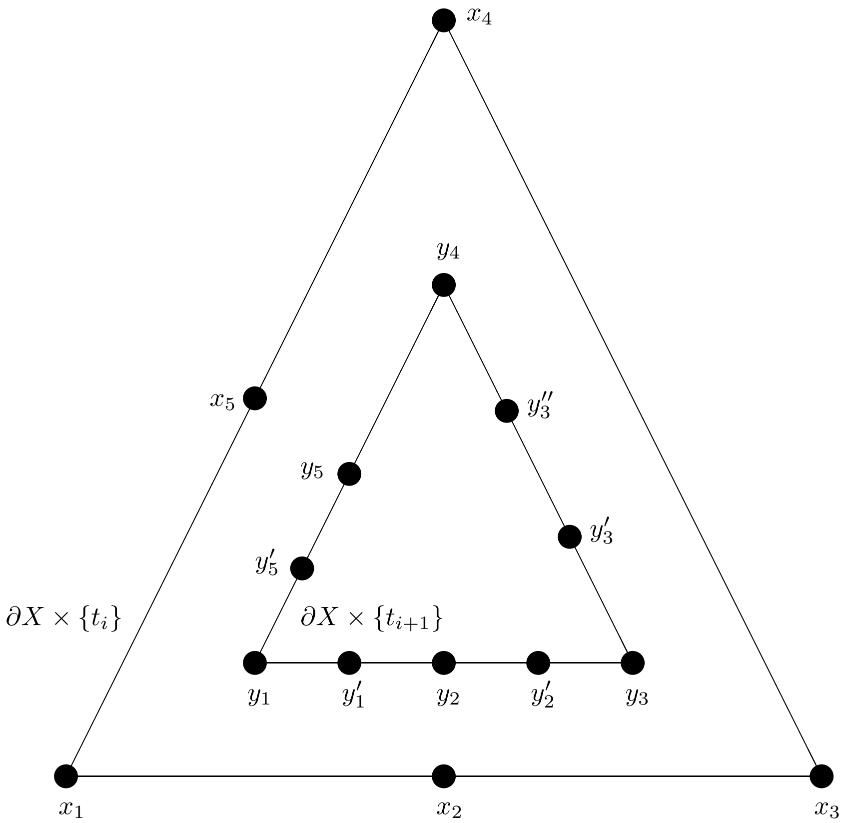
be slightly modified. Lemma 3.9 also guarantees that such modification exists. Subsections 3.7 and 3.8 give the definition used in 3.9. (This consideration are all in order to ensure the condition (2).)



Now, $\bigcap_{k=1}^{5} E_1(y_k)$ contains one more element of E_1^{model} than $\bigcap_{k=1}^{5} E_1(x_k)$, namely, λ. So for E_1, the groupings on $\partial X \times \{t_{i+1}\}$ are (globally) closer to the model grouping than that on $\partial X \times \{t_i\}$.

But the groupings on $\partial X \times \{t_{i+1}\}$ may not satisfy the condition (3) with $\dim(X)$ replaced by $\dim(\partial X)$, as the groupings on $\partial X \times \{t_i\}$ do .

By the induction assumption, applied to each individual simplex of $\partial X \times \{t_{i+1}\}$ (with the provisional simplicial structure), we can introduce a subdivision for $\partial X \times \{t_{i+1}\}$ and groupings for the new vertices to make the condition (3), with $\dim(X)$ replaced by $\dim(\partial X)$, hold for $\partial X \times \{t_{i+1}\}$. The picture now looks like



In this picture, y_1', y_2', y_3', y_3'', and y_5' are the new vertices introduced in the subdivision.

(Of course this picture only shows a special case.)

It goes without saying that we wish to ensure the condition (1) (and also the condition (2)) for the groupings associated to the new vertices inside each provisional simplex of $\partial X \times \{t_{i+1}\}$. In the other words, when we introduce the groupings for a new vertex inside a fixed provisional simplex of $\partial X \times \{t_{i+1}\}$ (e.g., y_2' inside $[y_2, y_3]$), for each k we should keep the intersection of the sets E_k over vertices of this simplex (e.g., $E_k(y_2) \cap E_k(y_3)$) inside the set E_k for the new vertex (e.g., inside $E_k(y_2')$). This is the condition (1) for this provisional simplex. The condition (1) for all the individual simplices implies that after the subdivision, the intersection over the whole layer $\bigcap_{y \in \mathrm{Vertex}(X \times \{t_{i+1}\})} E_1(y)$ is equal to the intersection over the vertices of provisional simplicial structure $\bigcap_{k=1}^{5} E_1(y_k)$, and therefore still contains one more element of E_1^{model} than $\bigcap_{k=1}^{5} E_1(x_k)$ (namely , λ).

One may notice that the subset $\partial X \times [t_i, t_{i+1}]$ is not automatically a simplicial complex. We shall use Lemma 3.10 below to decompose it into a simplicial complex.

Because we do not change much from the grouping of x_j to the grouping of y_j and because we make (1) true when introduce groupings for new vertices y_j', y_j'', etc., the groupings for any simplex inside $\partial X \times [t_i, t_{i+1}]$ will satisfy the condition (3) (of course with $\dim(X)$ NOT replaced by $\dim(\partial X)$).

Finally, let us mention that, we carry out the above construction separately for E_1, E_2, etc. Once this has been done for E_1, the same method can be used for E_2. The condition (1) will guarantee that when we work on E_2, we will not affect the condition (3) for E_1, which was supposed to be already satisfied. (The details will be contained in the proof of Lemma 3.11)

As we mentioned above, when we construct $E_1(y_j)$ from $E_1(x_j)$, we need to replace one element of $E_1(x_j)\backslash E_1^{model}$ by the element $\lambda \in E_1^{model}\backslash E_1(x_j)$. If we choose an arbitrary element $\mu \in E_1(x_j)\backslash E_1^{model}$ to be replaced by λ to define $E_1(y_j)$, then in general, $E_1(y_j)$ may not be extended to a grouping satisfying the condition (2), in other words, there may not exist a grouping $E_1, E_2, \cdots, E_m$ of E of type $(K_1, K_2, \cdots, K_m)$ such that $E_1 = E_1(y_j)$ and

$$E_k \subset \bigcup_{x \in \mathrm{Vertex}(X)} E_k(x), \quad k = 1, 2, \cdots, m.$$

So we need to give a condition to ensure that a subset $E_1 \subset E$ can be extended to a grouping satisfying condition (2). This will be discussed in 3.7 and 3.8. (See condition $(**)$ in 3.8.)

The proof of Lemma 3.4 will be given in 3.7 to 3.16.

3.7. We will put the inclusion $E_j(x) \subset \bigcup_{y \in \mathrm{Vertex}(X,\sigma)} E_j(y)$ in the condition (2) of Lemma 3.4, into a more general form, as follows. (In fact, we will use this more general form in our application.)

Suppose that $H_1, H_2, \cdots, H_m$ are (not necessarily disjoint) subsets of E, satisfying the following condition (called Condition $(*)$). For each subset $I \subset \{1, 2, \cdots, m\}$,

$$\#(\bigcup_{i \in I} H_i) \geq \sum_{i \in I} K_i \ . \qquad (*)$$

It follows obviously that $H_1 \cup H_2 \cup \cdots \cup H_m = E$, since $\#(H_1 \cup H_2 \cup \cdots \cup H_m) \geq \sum_{i=1}^m K_i = \#(E)$.

From the Marriage Lemma of [HV] (or the Pairing Lemma in [Su]), the condition $(*)$ is a necessary and sufficient condition for the existence of a grouping $E_1, E_2, \cdots, E_m$ of E of type $(K_1, K_2, \cdots, K_m)$ with the condition $E_i \subset H_i$.

(Recall that, the Marriage Lemma of [HV] is stated as follows.

Suppose that there are two groups of K boys and K girls. Suppose that the following condition holds:

For any subset of K_1 girls ($K_1 = 1, 2, \cdots, K$), there are at least K_1 boys, each of them knows at least one girl from this subset.
Then there is a way to arrange marriage between them such that each boy marries one of the girls he knows.
Our claim above is a special case of this Marriage Lemma. One can see this as follows. Suppose that the K girls are from m different clubs, and the i-th club has exactly K_i girls. Number the boys by $1, 2, \cdots, K$. Let us define the relation consisting of a boy knowing a girl as follows. If $j \in H_i$, then the boy j knows all the girls in the i-th club. Otherwise, he does not know any girl in the i-th club. (Notice that the boy j could be in different H_i, so he could know girls from different clubs.) Obviously the condition (*) becomes the above condition in the Marriage Lemma. So if the condition $(*)$ holds, then there is a way to arrange the marriage as in the lemma. One can define E_i to be the set of boys each of whom marries a girl from the i-th club. Obviously $E_i \subset H_i$. This proves the sufficiency part of the condition. The necessary part is trivial.)
If we let $H_i = \bigcup_{y \in \mathrm{Vertex}(X,\sigma)} E_i(y)$, then the inclusion in (2) of Lemma 3.4 becomes $E_i(x) \subset H_i$ for each $x \in \mathrm{Vertex}(X, \sigma)$.
For any subset $I \subset \{1, 2, \cdots, m\}$, let

$$H_I = \bigcup_{i \in I} H_i \ .$$

3.8. We say that a subset $E_1 \subset H_1$, of K_1 elements, satisfies Condition $(**)$ if for any $I \subset \{2, 3, \cdots, m\}$,

$$(**) \qquad \#(H_I \backslash E_1) \geq \sum_{i \in I} K_i \ .$$

(Caution: $1 \notin I$.) Again, from the Marriage Lemma, $E_1 \subset H_1$ satisfies $(**)$ if and only if E_1 can be extended to a grouping $E_1,\ E_2, \cdots, E_m$ of E of type $(K_1, K_2, \cdots, K_m)$ such that $E_i \subset H_i$.

Lemma 3.9. *Suppose that E_1, F_1 $(\subset H_1)$ are two subsets satisfying $(**)$. If $\lambda \in F_1 \backslash E_1$, then there is a $\mu \in E_1 \backslash F_1$ such that*

$$E_1' = (E_1 \backslash \{\mu\}) \cup \{\lambda\},$$

satisfies $(**)$.

Proof: Let $G = E_1 \cup \{\lambda\}$. Since E_1 satisfies $(**)$, necessarily,

$$\#(H_I \backslash G) \geq \sum_{i \in I} K_i - 1 \ ,$$

for all subsets $I \subset \{2, 3, \cdots, m\}$.
Let $\tilde{H}_i = H_i \backslash G,\ i \in \{2, 3, \cdots, m\}$. And let $\tilde{H}_I = \cup_{i \in I} \tilde{H}_i$ for any $I \subset \{2, 3, \cdots, m\}$. The above inequality becomes $\#(\tilde{H}_I) \geq \sum_{i \in I} K_i - 1$.

Let I_0 be a minimum subset of $\{2,3,\cdots,m\}$ such that

$$\#(\tilde{H}_{I_0}) = \sum_{i\in I_0} K_i - 1 \ .$$

Note that such set I_0 exists, since if $I = \{2,3,\cdots,m\}$,

$$\#(\tilde{H}_I) = \sum_{i\in I} K_i - 1 \ .$$

Using the fact that $\#(H_{I_0}\backslash F_1) \geq \sum_{i\in I_0} K_i$, we can prove that

$$E_1 \cap H_{I_0} \not\subset F_1.$$

If it is not true, then $G\cap H_{I_0} \subset F_1$, since $\lambda \in F_1$ and $G = E_1 \cup \{\lambda\}$. And therefore,

$$\#(\tilde{H}_{I_0}) = \#(H_{I_0}\backslash G) \geq \#(H_{I_0}\backslash F_1) \geq \sum_{i\in I_0} K_i \ ,$$

which contradicts with the above equation.
Choose any element $\mu \in (E_1\cap H_{I_0})\backslash F_1$; we will prove that μ is as desired in the lemma. I.e., the set

$$E_1' = (E_1\backslash\{\mu\}) \cup \{\lambda\} = G\backslash\{\mu\}$$

satisfies $(**)$. That is, for any $J \subset \{2,3,\cdots,m\}$, $\#(H_J\backslash E_1') \geq \sum_{i\in J} K_i$.
The proof is divided into three cases.
(i) *The case that* $J\cap I_0 = \emptyset$. By the relations

$$\tilde{H}_{I_0\cup J} = (\tilde{H}_J\backslash\tilde{H}_{I_0}) \cup \tilde{H}_{I_0} \qquad \text{(disjoint union)}$$

and

$$\#(\tilde{H}_{I_0\cup J}) \ \geq \sum_{i\in I_0\cup J} K_i - 1 \ ,$$

combined with the definition of I_0, one knows that

$$(a) \qquad \#(\tilde{H}_J\backslash\tilde{H}_{I_0}) \ \geq \sum_{i\in J} K_i \ ,$$

which is stronger than the condition

$$\#(H_J\backslash E_1') \ \geq \sum_{i\in J} K_i \ .$$

(ii) *The case that* $J\subset I_0$. Obviously, for $J = I_0$, we have

$$\#(H_{I_0}\backslash E_1') \ = \#(H_{I_0}\backslash G) + 1 = \sum_{i\in I_0} K_i \ ,$$

since $E_1' = G\backslash\{\mu\}$ and $\mu \in H_{I_0} \cap G$.
So we can suppose that $J \subsetneqq I_0$.

By the minimality of I_0, we know that

$$\#(H_J\backslash G) \geq \sum_{i\in J} K_i \ .$$

Therefore,

$$\#(H_J\backslash E_1') \ \geq \sum_{i\in J} K_i \ .$$

(iii) *The general case.* Let $J_0 = J \cap I_0$, and $J_1 = J\backslash J_0$. Then

$$(H_J\backslash E_1') \supset (\tilde{H}_{J_1}\backslash\tilde{H}_{I_0}) \cup (H_{J_0}\backslash E_1'),$$

where the right hand side is a disjoint union since $J_0 \subset I_0$.
Evidently, this case follows from (a) above and case (ii).

$\square$

The following lemma is perhaps well known.

Lemma 3.10. *Let (X,σ_0) be a simplicial complex and (X,σ_1) be a subdivision of (X,σ_0). It follows that there is a simplicial structure σ of $X\times[0,1]$ such that*
(1) all vertices of $(X\times[0,1],\sigma)$ are on three subsets $X\times\{0\}$, $X\times\{\frac{1}{2}\}$, and $X\times\{1\}$;
(2) $(X\times[0,1],\sigma)|_{X\times\{0\}} = (X,\sigma_0)$, and $(X\times[0,1],\sigma)|_{X\times\{1\}} = (X,\sigma_1)$;
(3) For a simplex Δ of $(X\times[0,1],\sigma)$, there is a simplex Δ_0 of (X,σ_0) (caution: we do not use (X,σ_1)) such that

$$\Delta \subset \Delta_0 \times [0,1],$$

as a subset.

Proof: We prove it by induction on $\dim(X)$.
If X is 0-dimensional simplicial complex which consists of finitely many points, the conclusion is obvious, since $X\times[0,1]$ is finitely many disjoint intervals. (Note that, at this case, necessarily, $(X,\sigma_0) = (X,\sigma_1)$.) For us to visualize the general case later on, we introduce a new vertex $(x,\frac{1}{2}) \in X\times\{\frac{1}{2}\}$ for each $x\in X$. That is, we divide the interval $\{x\}\times[0,1]$ into two simplices $\{x\}\times[0,\frac{1}{2}]$ and $\{x\}\times[\frac{1}{2},1]$.
As the induction assumption, let us assume that the lemma is true for any n-dimensional complex. Let $\dim(X) = n+1$.
Let $X^{(n)}$ be the n-skeleton of (X,σ_0) (we use σ_0 not σ_1 here). By the induction assumption, there is a simplicial structure σ' of $X^{(n)}\times[0,1]$ such that
(1) all vertices of $(X^{(n)}\times[0,1],\sigma')$ are on three subsets $X^{(n)}\times\{0\}$, $X^{(n)}\times\{\frac{1}{2}\}$, and $X^{(n)}\times\{1\}$;

(2) $(X^{(n)} \times [0,1], \sigma')|_{X^{(n)}\times\{0\}} = (X^{(n)}, \sigma_0|_{X^{(n)}})$, and
$(X^{(n)} \times [0,1], \sigma')|_{X^{(n)}\times\{1\}} = (X^{(n)}, \sigma_1|_{X^{(n)}})$;

(3) For an simplex Δ of $(X^{(n)} \times [0,1], \sigma')$, there is a simplex Δ_0 of $(X^{(n)}, \sigma_0)$ such that

$$\Delta \subset \Delta_0 \times [0,1],$$

as a subset.

Let us introduce the simplicial structure σ on $X \times [0,1]$ such that $(X \times [0,1], \sigma)|_{X^{(n)}\times[0,1]} = (X^{(n)} \times [0,1], \sigma')$, $(X \times [0,1], \sigma)|_{X\times\{0\}} = (X, \sigma_0)$, and $(X \times [0,1], \sigma)|_{X\times\{1\}} = (X, \sigma_1)$.

Consider each $\Delta \times [0,1]$ for any $(n+1)$-simplex Δ of (X, σ_0) (again, we use σ_0 not σ_1). From the above, we already have the simplicial structure on the boundary

$$\partial(\Delta \times [0,1]) = (\Delta \times \{0\}) \cup (\partial\Delta \times [0,1]) \cup (\Delta \times \{1\}).$$

Namely, on $\Delta \times \{0\}$, we use σ_0; on $\partial\Delta \times [0,1]$, we use σ'; and on $\Delta \times \{1\}$, we use σ_1.

Let c be the barycenter of Δ, introduce a new vertex $C = (c, \frac{1}{2}) \in X \times \{\frac{1}{2}\}$. The simplices of σ on $\Delta \times [0,1]$ are of the following forms.

(i) C itself is a zero dimensional simplex;

(ii) Any simplex of the boundary $\partial(\Delta \times [0,1])$ is a simplex for σ on $\Delta \times [0,1]$; and

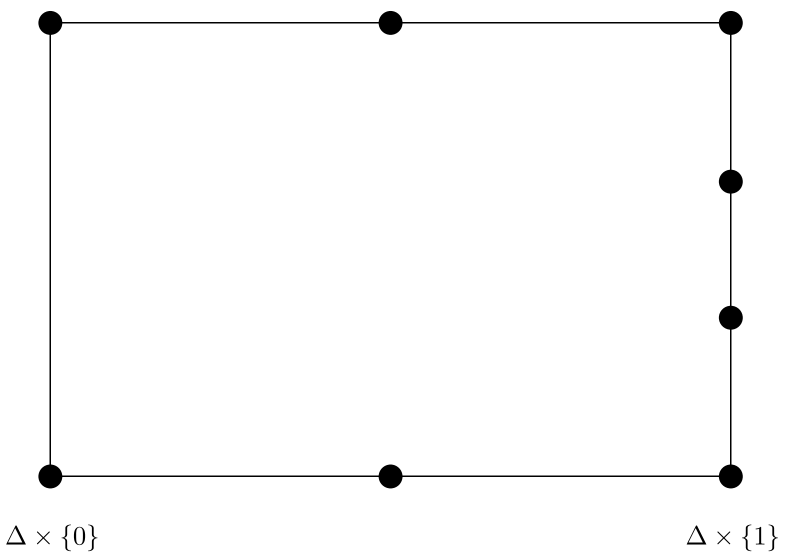
(iii) For any simplex Δ' of the boundary $\partial(\Delta\times[0,1])$, the convex hull of $\Delta'\cup\{C\}$ is a simplex of dimension $(\dim(\Delta')+1)$ for σ on $\Delta \times [0,1]$.

Define such simplicial structure for each $(n+1)$-simplex separately, and put them together give rise to a simplicial structure of $X \times [0,1]$, which obviously satisfies the conditions (1), (2), and (3).

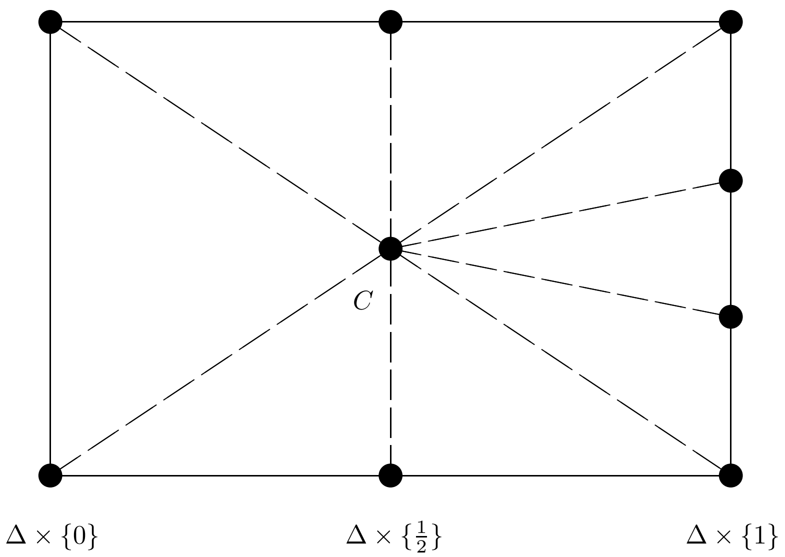
(Note that the simplicial structure on $\partial\Delta \times [0,1]$ is as the same as σ', therefore the simplicial structure on $\Delta \times [0,1]$ and on $\Delta_1 \times [0,1]$ for different $(n+1)$-dimensional simplices Δ and Δ_1 are compatible on the intersection $(\Delta \cap \Delta_1) \times [0,1]$.)

The following pictures may help the reader to visualize the construction. They are pictures only for the case $n = 0$, $\dim(\Delta) = 1$, and $\dim(\Delta \times [0,1]) = 2$.

Suppose the simplicial structure for the boundary $\partial(\Delta \times [0,1])$ is as follows. (The dots represent vertices.)



Then the simplicial structure on $\Delta \times [0,1]$ will be described by the following picture.



$\square$

The following lemma presents the main technical step of this section.

LEMMA 3.11. *Suppose that $\{H_1, H_2, \cdots, H_m\}$ satisfies the condition $(*)$. Suppose that (X, σ) is a simplicial complex consisting of a single simplex Δ_0 and its faces. Let $(Y, \sigma) = (\partial\Delta_0, \sigma)$, and (Y, τ) be a subdivision of (Y, σ). Suppose that it is assigned, for each vertex $x \in (Y, \tau)$, a set $E_1(x) \subset H_1$ which satisfies the condition $(**)$. Furthermore, suppose that for any simplex Δ of (Y, τ),*

$$\#\left(\bigcap_{y \in Vertex(\Delta)} E_1(y)\right) \geq K_1 - \frac{\dim Y(\dim Y + 1)}{2} .$$

It follows that there are a subdivision $(X, \tilde{\tau})$ of (X, σ) and an assignment, for each vertex $x \in Vertex(X, \tilde{\tau})$, a set $E_1(x) \subset H_1$, satisfying condition $(**)$, with the following conditions.
(1) $(X, \tilde{\tau})|_Y = (Y, \tau)$, and for each vertex $y \in Vertex(Y, \tau)$, the assignment $E_1(y)$ is as same as the original one.
(2) For any $x \in (X, \tilde{\tau})$,

$$E_1(x) \supset \bigcap_{y \in Vertex(Y,\tau)} E_1(y) .$$

(3) For any simplex Δ of $(X, \tilde{\tau})$,

$$\#\left(\bigcap_{x \in Vertex(\Delta)} E_1(x)\right) \geq K_1 - \frac{\dim X(\dim X + 1)}{2} .$$

Proof: The Lemma is proved by induction on the dimension of the simplex. If $dim(\Delta_0) = 0$, then $\Delta_0 = \{pt\}$, a set of single point, and $\partial\Delta_0 = \emptyset$. Obviously, the lemma holds by choosing any $E_1(pt) \subseteq H_1$ of K_1 element to satisfy $(**)$.
Let us prove the 1-dimensional case. Logically, this part could be skipped. But the proof of this case will be easier to visualize which can be used to understand the general case.
Suppose that $dim(\Delta_0) = 1$. Δ_0 is a line segment $[0, 1]$. Divide $[0, 1]$ into several subintervals by

$$0 = t_0^0 < t_1^0 < t_2^0 < \cdots < t_{a-1}^0 < t_a^0 = \frac{1}{2} = t_a^1 < t_{a-1}^1 < \cdots < t_2^1 < t_1^1 < t_0^1 = 1.$$

(The natural number a is to be determined later.) The points $\{t_j^i\}_{i=0,1; j=1,2\cdots,a}$ will be the new vertices of $(\Delta_0, \tilde{\tau})$. (Note that t_a^0 is the same vertex as t_a^1.) Choose a model $E_1^{model} \subset H_1$ to satisfy $(**)$ and

$$E_1^{model} \supset E_1(t_0^0) \cap E_1(t_0^1).$$

In fact, one can choose E_1^{model} to be either $E_1(t_0^0)$ or $E_0(t_0^1)$. (Note that $t_0^0 = 0$ and $t_0^1 = 1$ are vertices of (Δ_0, σ).)
Let $G = E_1(t_0^0) \cap E_1(t_0^1) \cap E_1^{model}$. Without loss of generality, we can assume that there is $\lambda \in E_1^{model} \backslash G$. Otherwise, $E_1^{model} = G = E_1(t_0^0) = E_1(t_0^1)$, and the conclusion already holds before introducing any subdivision.
By Lemma 3.9, if $\lambda \notin E_1(t_0^i)$ $(i = 0, 1)$, then there is a $\mu \in E_1(t_0^i) \backslash E_1^{model}$ such that $E_1(t_0^i) \cup \{\lambda\} \backslash \{\mu\}$ satisfies $(**)$. Define

$$E_1(t_1^i) = \begin{cases} E_1(t_0^i) & \text{if } \lambda \in E_1(t_0^i) \\ (E_1(t_0^i) \cup \{\lambda\}) \backslash \{\mu\} & \text{if } \lambda \notin E_1(t_0^i) \end{cases}.$$

Then $E_1(t_1^i) \supset G \cup \{\lambda\}$. Therefore,

$$E_1(t_1^0) \cap E_1(t_1^1) \cap E_1^{model} \underset{\neq}{\supset} E_1(t_0^0) \cap E_1(t_0^1) \cap E_1^{model}.$$

Suppose that we already have the definitions of $E_1(t_i^0)$ and $E_1(t_i^1)$, we can define $E_1(t_{i+1}^0)$ and $E_1(t_{i+1}^1)$ exactly the same as above (i in place of 0, and $i+1$ in place of 1), and obtain

$$E_1(t_{i+1}^0) \cap E_1(t_{i+1}^1) \cap E_1^{model} \underset{\neq}{\supset} E_1(t_i^0) \cap E_1(t_i^1) \cap E_1^{model}.$$

Carrying out this procedure for at most finitely many times, we will reach $E_1(t_{a-1}^0) \cap E_1(t_{a-1}^1) \cap E_1^{model} = E_1^{model}$. Then define $E_1(t_a^i) = E_1^{model}$. (Note that $t_a^1 = t_a^0 = \frac{1}{2}$.)
For $i = 0, 1; j = 0, 1, 2, \cdots, a-1$,

$$\#(E_1(t_j^i) \cap E_1(t_{j+1}^i)) \geq K_1 - 1 = K_1 - \frac{dim(\Delta_0)(dim(\Delta_0)+1)}{2},$$

since we take out at most one point from $E_1(t_j^i)$ to define $E_1(t_{j+1}^i)$. This proves that the lemma holds for $n = 1$.
(Let us point out that for one dimensional case, the proof could be simpler. We choose the above proof to present some idea for the general case below.)
Suppose that the lemma is true for any simplex of dimension $\leq n-1$. We will prove it for $\dim(X) = n$. (One should compare to the explanation in 3.6.)
Step 1. Identify Δ_0 with $\partial\Delta_0 \times [0,1]/\partial\Delta_0 \times \{1\}$. Regard $\partial\Delta_0$ as $\partial\Delta_0 \times \{0\} \subset \partial\Delta_0 \times [0,1]$. Note that $\partial\Delta_0 \times \{1\}$ is identified as a single point which is the center of Δ_0, and is NOT a vertex of (Δ_0, σ).
Choose $0 = t_0 < t_1 < \cdots < t_a = 1$. (The natural number a is to be determined later.)
We will first introduce some new vertices (for the subdivision $(X, \tilde{\tau})$) on $\partial\Delta_0 \times \{t_1\}, \partial\Delta_0 \times \{t_2\}, \cdots, \partial\Delta_0 \times \{t_a\}$, and define E_1 for those vertices.
Later on (in Step 4), we will consider each $\partial\Delta \times [t_i, t_{i+1}]$ to be $X \times [0,1]$ in Lemma 3.10, and introduce new vertices on $\partial\Delta \times \{\frac{t_i+t_{i+1}}{2}\}$ (in place of $X \times \{\frac{1}{2}\}$). (We need to do this, because $\partial\Delta \times [t_i, t_{i+1}]$ is not automatically a simplicial complex.)

Choose a model $E_1^{model} \subset H_1$ to satisfy $(**)$. We also require that

$$E_1^{model} \supset \bigcap_{y \in \text{Vertex}(Y,\tau)} E_1(y) \ .$$

(One can choose E_1^{model} to be $E_1(y)$ for any vertex $y \in \text{Vertex}(Y,\tau)$.)
Define $E_1\{(x,1)\} = E_1^{model}$. (Note that $\{(x,1)\} \subset \partial\Delta_0 \times [0,1]$ is identified to a single point, the center of Δ_0.)
The construction will be carried out in Step 2, 3 and 4. The procedure can be outlined as follows. If we already have the construction of simplicial structure $\tilde{\tau}$ for $\partial\Delta_0 \times \{t_{i-1}\}$ and the definition of E_1 on all vertices in $\text{Vertex}(\partial\Delta_0 \times \{t_{i-1}\})$, then, to define the simplicial structure on $\partial\Delta_0 \times [t_{i-1}, t_i]$ (in particular, to introduce vertices on $\partial\Delta_0 \times \{t_i\}$), and to define E_1 on the newly introduced vertices (on $\partial\Delta_0 \times \{t_i\}$), we will only use the simplicial structure and the definition of E_1 on $\partial\Delta_0 \times \{t_{i-1}\}$.
In this procedure, if there is a vertex $x \in \text{Vertex}(\partial\Delta_0 \times \{t_{i-1}\}, \tilde{\tau})$ such that

$$E_1(x) \neq E_1^{model} \ ,$$

then we will require that

$$(a) \qquad \bigcap_{x \in \text{Vertex}(\partial\Delta_0 \times \{t_i\}, \tilde{\tau})} E_1(x) \cap E_1^{model} \underset{\neq}{\supset} \bigcap_{x \in \text{Vertex}(\partial\Delta_0 \times \{t_{i-1}\}, \tilde{\tau})} E_1(x) \cap E_1^{model} \ .$$

(That is, the sets E_1's on $\partial\Delta_0 \times \{t_i\}$ are globally closer to E_1^{model} than those on $\partial\Delta_0 \times \{t_{i-1}\}$.) Finally, within finitely many steps, we will reach that, for certain $i-1$, and for all vertices $x \in \text{Vertex}(\partial\Delta_0 \times \{t_{i-1}\}, \tilde{\tau})$,

$$E_1(x) = E_1^{model} \ .$$

Then we choose $t_i = t_a = 1$, and choose any simplicial structure on $\partial\Delta_0 \times [t_{a-1}, 1]/\partial\Delta_0 \times \{1\}$ with vertex set to be $\text{Vertex}(\partial\Delta_0 \times \{t_{a-1}\}) \cup \partial\Delta_0 \times \{1\}$. Recall that the set $\partial\Delta_0 \times \{1\}$ is identified as a single point with $E_1(\partial\Delta_0 \times \{1\}) = E_1^{model}$.)
Furthermore, in this procedure, we not only make (3) true for any simplex in $\partial\Delta_0 \times [t_{i-1}, t_i]$, but also make the following stronger statement true for any simplex Δ lies on $\partial\Delta_0 \times \{t_i\}$:

$$(b) \qquad \#\left(\bigcap_{y \in \text{Vertex}(\Delta)} E_1(y) \right) \ \geq \ K_1 - \frac{(n-1)n}{2} \ .$$

(Note that $n-1 = \dim(\partial\Delta_0 \times \{t_i\}) = \dim(\partial\Delta_0)$.) This condition has to be satisfied for the construction of the next step by induction.
Step 2. We will do all the above construction only for $\partial\Delta_0 \times [t_0, t_1]$. For the other part of the construction, one uses induction argument with aid of (b) (i.e., let t_{i-1} play the role of t_0, and t_i play the role of t_1.)

Let $\{(y_1,t_0),(y_2,t_0),\cdots,(y_p,t_0)\}$ be the vertices of $\partial\Delta_0\times\{t_0\}=Y$. There is a simplicial complex structure on $\partial\Delta_0\times\{t_1\}$, which is exactly the same as that of $(\partial\Delta_0\times\{t_0\},\tilde{\tau})$, since both $\partial\Delta_0\times\{t_1\}$ and $\partial\Delta_0\times\{t_0\}$ can be regarded as $\partial\Delta_0$. We call such simplicial complex $\tilde{\tau}_{pre}$. Therefore, each point (y_i,t_1) $(1\le i\le p)$ is a vertex of $(\partial\Delta_0\times\{t_1\},\tilde{\tau}_{pre})$. We will introduce more vertices later.

Let $G=E_1(y_1,t_0)\cap E_1(y_2,t_0)\cap\cdots\cap E_1(y_p,t_0)\cap E_1^{model}$. If $G=E_1^{model}$, then $E_1(y_i,t_0)=E_1^{model}$ for each $1\le i\le p$ and the construction is done. So we assume

$$G\neq E_1^{model}.$$

Choose $\lambda\in E_1^{model}\backslash G$. When we define $E_1(x)$ for any vertex $x\in\partial\Delta_0\times\{t_1\}$, it is always required that

$$E_1(x)\supset G\cup\{\lambda\}.$$

Therefore, (a) holds for the pair $\{t_0,t_1\}$.

For each point (y_i,t_0), if $\lambda\notin E_1(y_i,t_0)$, by Lemma 3.9, there is $\mu\in E_1(y_i,t_0)\backslash E_1^{model}$ such that $(E_1(y_i,t_0)\cup\{\lambda\})\backslash\{\mu\}$ satisfies $(**)$. Define

$$E_1(y_i,t_1)=\begin{cases} E_1(y_i,t_0) & \text{if } \lambda\in E_1(y_i,t_0)\\ (E_1(y_i,t_0)\cup\{\lambda\})\backslash\{\mu\} & \text{if } \lambda\notin E_1(y_i,t_0)\ .\end{cases}$$

In this way, obviously, $E_1(x)\supset G\cup\{\lambda\}$ for each vertex $x=(y_i,t_1)\in \mathrm{Vertex}(\partial\Delta_0\times\{t_1\},\tilde{\tau}_{pre})$.

Step 3. Note that the definition of E_1 on $\mathrm{Vertex}(\partial\Delta_0\times\{t_1\},\tilde{\tau}_{pre})$ may not satisfies (b). Therefore we can not use the simplicial structure $\tilde{\tau}_{pre}$ and the definition of E_1 on $\mathrm{Vertex}(\partial\Delta_0\times\{t_1\},\tilde{\tau}_{pre})$ to construct simplicial structure and the definition of E_1 for $\partial\Delta_0\times\{t_2\}$. We need to introduce a subdivision for $(\partial\Delta_0\times\{t_1\},\tilde{\tau}_{pre})$ and the definitions of E_1 for new vertices to make (b) true. (This step is not needed in the one dimensional case, since for any zero dimensional simplex (which is a point), (b) automatically holds.)

Apply the induction assumption to each simplex of $(\partial\Delta_0\times\{t_1\},\tilde{\tau}_{pre})$ with the above definition of E_1 on $\mathrm{Vertex}(\partial\Delta_0\times\{t_1\},\tilde{\tau}_{pre})$, from the simplices of the lowest dimension (dimension 1) to the simplices of the highest dimension (dimension $n-1$). (Note that each such simplex has dimension at most $n-1$.) One should begin with each 1-simplex (with boundary being two points — two 0-simplices), then each 2-simplex, and so on.

First, let e be any 1-simplex of $(\partial\Delta_0\times\{t_1\},\tilde{\tau}_{pre})$ with boundary $\partial e=\{v_0,v_1\}$. Obviously, the condition of Lemma 3.11 automatically holds for simplex e in place of Δ_0 and ∂e in place of $\partial\Delta_0$, since ∂e is zero-dimensional. By the induction assumption, there is a subdivision $(e,\tilde{\tau})$ of $(e,\tilde{\tau}_{pre})$ and the definition of E_1 for each vertex of $(e,\tilde{\tau})$ such that

(1) The definition of E_1 on the original vertices $\{v_0,v_1\}$ are the same as before.

(2) For any $x\in\mathrm{Vertex}(e,\tilde{\tau})$,

$$E_1(x)\supset E_1(v_0)\cap E_1(v_1).$$

(3) For any simplex e' of $(e, \tilde{\tau})$ (a line segment of e)

$$\bigcap_{x \in \mathrm{Vertex}(e')} E_1(x) \geq K_1 - \frac{dim(e)(dim(e)+1)}{2}.$$

After we have done the above procedure for each 1-simplex, we can do it for each 2-simplex, since we already have simplicial structure and the definition of E_1 for the boundary of any 2-simplex as required in the condition of Lemma 3.11.
Going through this way, finally, one obtains a subdivision $(\partial\Delta_0 \times \{t_1\}, \tilde{\tau})$ of $(\partial\Delta_0 \times \{t_1\}, \tilde{\tau}_{pre})$ and the definition of E_1 for each newly introduced vertex, such that the following two statements hold.

1. For each old simplex Δ of $(\partial\Delta_0 \times \{t_1\}, \tilde{\tau}_{pre})$ and any new vertex $x \in \Delta$,

$$(c) \qquad E_1(x) \supset \bigcap_{y \in \mathrm{Vertex}(\Delta, \tilde{\tau}_{pre})} E_1(y) \ .$$

2. If Δ is a simplex of $(\partial\Delta_0 \times \{t_1\}, \tilde{\tau})$, then

$$\#\left(\bigcap_{y \in \mathrm{Vertex}(\Delta)} E_1(y)\right) \quad \geq \quad K_1 - \frac{\dim Y(\dim Y + 1)}{2} = K_1 - \frac{(n-1)n}{2} \ .$$

(This is the requirement (b) in Step 1.)
The first statement is the induction assumption of validity of (2) and the second statement is the induction assumption of validity of (3).
Step 4. In this step, we will apply Lemma 3.10 to define the simplicial structure $\tilde{\tau}$ on $\partial\Delta_0 \times [t_0, t_1]$ and the definitions of E_1 on all vertices. Note that we already have simplicial structure $\tilde{\tau}$ on $\partial\Delta_0 \times \{t_0\}$ and on $\partial\Delta_0 \times \{t_1\}$. Furthermore, $\tilde{\tau}|_{\partial\Delta_0 \times \{t_1\}}$ is a subdivision of $\tilde{\tau}|_{\partial\Delta_0 \times \{t_0\}}$ if we regard both $\partial\Delta_0 \times \{t_0\}$ and $\partial\Delta_0 \times \{t_1\}$ as $\partial\Delta_0$. Apply Lemma 3.10 (with $\partial\Delta_0$ in place of X) to obtain the simplicial structure on $\partial\Delta_0 \times [t_0, t_1]$ (we only need to introduce new vertices on $\partial\Delta_0 \times \{\frac{t_0+t_1}{2}\}$).
For each new vertex $(u, \frac{t_0+t_1}{2}) \in \partial\Delta_0 \times \{\frac{t_0+t_1}{2}\}$, consider $(u, t_0) \in \partial\Delta_0 \times \{t_0\}$. From 3.1, there is a unique simplex Δ of $(\partial\Delta_0 \times \{t_0\}, \tilde{\tau})$ such that $(u, t_0) \in$ interior(Δ). Choose any vertex x of Δ and define $E_1(u, \frac{t_0+t_1}{2}) = E_1(x)$.
So we have the simplicial structure $\tilde{\tau}$ on $\partial\Delta_0 \times [t_0, t_1]$ and the definition of $E_1(x)$ for each $x \in \mathrm{Vertex}(\partial\Delta_0 \times [t_0, t_1])$. We need to verify the condition (3). Let Δ be any simplex of $(\partial\Delta_0 \times \{t_0\}, \tilde{\tau})$ with vertices $\{(u_0, t_0), (u_1, t_0), \cdots, (u_i, t_0)\}$. Then

$$\#(E_1(u_0, t_0) \cap E_1(u_1, t_0) \cap \cdots \cap E_1(u_i, t_0)) \geq K_1 - \frac{\dim Y(\dim Y + 1)}{2} \ .$$

Let $G_1 = E_1(u_0, t_0) \cap E_1(u_1, t_0) \cap \cdots \cap E_1(u_i, t_0)$. From the above definition of E_1 for vertices of $\partial\Delta_0 \times \{\frac{t_0+t_1}{2}\}$, we know that if $(u, t_0) \in \Delta$ and $(u, \frac{t_0+t_1}{2}) \in \mathrm{Vertex}(\partial\Delta_0 \times [0,1], \tilde{\tau})$, then

$$E_1(u, \frac{t_0+t_1}{2}) \supset G_1. \tag{d}$$

Since each $E_1(u_j, t_1)$ is either $E_1(u_j, t_0)$ or is obtained by replacing one element of $E_1(u_j, t_0)$ by λ, we have

$$\begin{aligned} &\#(G_1 \cap E_1(u_0, t_1) \cap E_1(u_1, t_1) \cap \cdots \cap E_1(u_i, t_1)) \\ &\quad\geq K_1 - \tfrac{\dim Y(\dim Y+1)}{2} - (i+1) \\ &\quad\geq K_1 - \tfrac{(n-1)n}{2} - n \\ &\quad= K_1 - \tfrac{n(n+1)}{2}\ . \end{aligned} \tag{e}$$

(Note that there are $i+1$ $(\leq n)$ sets of $\{E_1(u_j, t_1)\}_{j=0}^{i}$, and, therefore, at most $i+1$ points were taken out from G_1.)
Recall that $\tilde{\tau}$ on $\Delta_0 \times \{t_1\}$ is the subdivision of $\tilde{\tau}_{pre}$. By (c) of Step 3, we have

$$\begin{aligned} \bigcap_{x \in \mathrm{Vertex}(\Delta \times \{t_1\}, \tilde{\tau})} E_1(x) &\supset \bigcap_{y \in \mathrm{Vertex}(\Delta \times \{t_1\}, \tilde{\tau}_{pre})} E_1(y) \\ &= E_1(u_0, t_1) \cap E_1(u_1, t_1) \cap \cdots \cap E_1(u_i, t_1). \end{aligned}$$

(Note that (c) implies that the above "$\supset$" holds if the left hand side of "$\supset$" is replaced by $E_1(x)$ for any $x \in \mathrm{Vertex}(\Delta \times \{t_1\}, \tilde{\tau})$, so it also holds for the intersection of these $E_1(x)$. In fact, the above "$\supset$" can be replaced by "$=$".)
Then combining it with (d), we have

$$\bigcap_{x \in \mathrm{Vertex}(\Delta \times [t_0, t_1], \tilde{\tau})} E_1(x) = G_1 \cap E_1(u_0, t_1) \cap E_1(u_1, t_1) \cap \cdots \cap E_1(u_i, t_1)$$

which has at least $K_1 - \frac{n(n+1)}{2}$ elements by (e). Combining this fact with (3) of Lemma 3.10, we know that the desired condition (3) holds for any simplex of $(\partial\Delta_0 \times [t_0, t_1], \tilde{\tau})$.
Evidently, (2) holds from the construction.
Since (b) holds for $\partial\Delta_0 \times \{t_1\}$, one can continue this procedure. This ends the proof.

$\square$

Corollary 3.12. *Suppose that $\{H_1, H_2, \cdots, H_m\}$ satisfies the condition $(*)$. Suppose that (X, σ) is a simplicial complex consisting of a single simplex and its faces. Suppose that there is assigned, for each vertex $x \in (X, \sigma)$, a set $E_1(x) \subset H_1$ which satisfies the condition $(**)$.*
It follows that there are a subdivision (X, τ) of (X, σ) and an assignment, for each new vertex $x \in$ Vertex(X, τ), a set $E_1(x) \subset H_1$, satisfying condition $(**)$,

with the following conditions. (The definition of E_1 for the old vertex should not be changed.)
(1) For any $x \in Vertex(X,\tau)$,

$$E_1(x) \supset \bigcap_{y \in Vertex(X,\sigma)} E_1(y) \ .$$

(2) For any simplex Δ of (X,τ),

$$\#\left(\bigcap_{x \in Vertex(\Delta)} E_1(x)\right) \geq K_1 - \frac{\dim X(\dim X+1)}{2} \ .$$

Proof: To prove this corollary, one needs to apply Lemma 3.11 to simplices from the lowest dimension (e.g. dimension one simplex whose boundary consists two vertices of (X,σ)) to the highest dimension (e.g the simplex X itself with boundary ∂X). Each time, we only work on a single simplex Δ of (X,σ). And when we work on Δ, we should assume that we already have the subdivision and the definition of E_1 on the boundary $\partial\Delta$ to satisfy the condition in Lemma 3.11 with $\dim(\partial\Delta)$ in place of $\dim(Y)$.

$\square$

Corollary 3.13. *Let (X,σ) be a simplicial complex consisting of a single simplex X and all its faces. Suppose that associated to each $x \in Vertex(X,\sigma)$, there is a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ of E.*
It follows that there is a subdivision (X,τ) of (X,σ), and associated to each new vertex $x \in Vertex(\Delta,\tau)$, there is a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ of E (for any old vertex of (Δ,σ), the grouping should not be changed), such that the following hold.
For each newly introduced vertex $x \in Vertex(X,\tau)$,

$$E_j(x) \subset \bigcup_{y \in Vertex(X,\sigma)} E_j(y), \qquad j = 1,2,\cdots,m. \tag{2}$$

For any simplex Δ of (X,τ) (after subdivision),

$$\#(\bigcap_{x \in Vertex(\Delta)} E_1(x)) \geq K_1 - \frac{n(n+1)}{2} \tag{3}$$

where $n = \dim X$.
(In this corollary, we do not require the condition (1) in Lemma 3.4. This will be done in the next corollary.)

Proof: Set $\bigcup_{y \in \mathrm{Vertex}(X,\sigma)} E_j(y) := H_j, \quad j = 1,2,\cdots,m.$ Then

$H_1, H_2, \cdots, H_m$ satisfy condition $(*)$, and for each $x \in \mathrm{Vertex}(X,\sigma)$, $E_1(x) \subset H_1$ satisfies $(**)$.
Applying Corollary 3.12, we obtain the subdivision (X,τ) and the definition of $E_1(x)$, for each new vertex, to satisfy condition $(**)$, and (1) and (2) in the Corollary 3.12.
For each new vertex x, since $E_1(x)$ satisfies $(**)$, we can extend it to a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ such that $E_i(x) \subset H_i$. Therefore this grouping satisfy the condition (2) of our corollary.
The condition (3) follows from the condition (2) of Corollary 3.12. Thus the corollary is proved.

□

Corollary 3.14. *Let (X,σ) be a simplicial complex consisting of a single simplex X and all its faces. Suppose that associated to each $x \in \mathrm{Vertex}(X,\sigma)$, there is a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ of E.*
It follows that there is a subdivision (X,τ) of (X,σ), and associated to each new vertex $x \in \mathrm{Vertex}(\Delta,\tau)$, there is a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ of E (for any old vertex of (Δ,σ), the grouping should not be changed), such that the following hold.
For each newly introduced vertex $x \in \mathrm{Vertex}(X,\tau)$,

$$(1) \qquad \bigcap_{y \in \mathit{Vertex}(X,\sigma)} E_j(y) \subset E_j(x), \qquad j = 1, 2, \cdots, m,$$

and

$$(2) \qquad E_j(x) \subset \bigcup_{y \in \mathit{Vertex}(X,\sigma)} E_j(y), \qquad j = 1, 2, \cdots, m.$$

For any simplex Δ of (X,τ) (after subdivision),

$$(3) \qquad \#\Big(\bigcap_{x \in \mathit{Vertex}(\Delta)} E_1(x)\Big) \geq K_1 - \frac{n(n+1)}{2}.$$

where $n = \dim X$.
(Comparing this Corollary to Lemma 3.4, the only difference is that we require (3) holds only for E_1 in the corollary.)

Proof: The only difference between this corollary and Corollary 3.13 is that we require condition (1) holds. To make (1) hold, we need to do the following. Reserve all the subsets $\bigcap_{y\in \mathrm{Vertex}(X,\sigma)} E_j(y)$, $j = 1, 2, \cdots, m$, which are supposed to be in $E_j(x)$ (if we want the condition (1) to hold), for any newly introduced vertex x; group the rest of the elements of E (using Corollary 3.13); and finally put $\bigcap_{y\in \mathrm{Vertex}(X,\sigma)} E_j(y)$ into each $E_j(x)$. The details are as follows.
Set $\bigcap_{y\in \mathrm{Vertex}(X,\sigma)} E_j(y) := D_j$, $j = 1, 2, \cdots, m$. Then D_j, $j = 1, 2, \cdots, m$ are mutually disjoint. To see this, we fix a $y \in \mathrm{Vertex}(X,\sigma)$, and notice

that $D_j \subset E_j(y)$, and $E_j(y)$, $j = 1, 2, \cdots, m$ are mutually disjoint, from the definition of grouping. Similarly, if $j_1 \neq j_2$, then $D_{j_1} \cap E_{j_2}(y) = \emptyset$, for any $y \in \mathrm{Vertex}(X, \sigma)$.
Consider $E' = E \backslash (\cup_j D_j)$ and the m-tuple

$$(K'_1, K'_2, \cdots, K'_m) = (K_1 - \#(D_1), K_2 - \#(D_2), \cdots, K_m - \#(D_m)).$$

For any $y \in \mathrm{Vertex}(X, \sigma)$, the grouping $E_1(y), E_2(y), \cdots, E_m(y)$ of E of type $(K_1, K_2, \cdots, K_m)$ induces a grouping $E'_1(y), E'_2(y), \cdots, E'_m(y)$ of E' of type $(K'_1, K'_2, \cdots, K'_m)$, by setting $E'_j(y) = E_j(y) \backslash D_j$.
Apply Corollary 3.13 to the simplex (X, σ) and those groupings of E', to obtain a subdivision (X, τ) and groupings $E'_1(x), E'_2(x), \cdots, E'_m(x)$ of E' for all newly introduced vertices $x \in \mathrm{Vertex}(X, \tau)$ such that the following hold.
For each newly introduced vertex $x \in \mathrm{Vertex}(X, \tau)$,

$$(2') \qquad E'_j(x) \subset \bigcup_{y \in \mathrm{Vertex}(X,\sigma)} E'_j(y), \qquad j = 1, 2, \cdots, m.$$

For any simplex Δ of (X, τ) (after subdivision),

$$(3') \qquad \#(\bigcap_{x \in \mathrm{Vertex}(\Delta)} E'_1(x)) \geq K'_1 - \frac{n(n+1)}{2},$$

where $n = \dim X$.
Finally, let $E_j(x) = E'_j \cup D_j$ for any $x \in \mathrm{Vertex}(X, \tau)$. Then the desired condition (1) of the corollary means $D_j \subset E_j(x)$, which is true from the definition. Also the conditions (2) and (3) of the corollary follows from (2′) and (3′).

□

Corollary 3.15. *Suppose that (X, σ) is a simplicial complex. Suppose that for each vertex $x \in$ Vertex(X, σ), there is a grouping $E_1(x), E_2(x), \cdots, E_m(x)$ of E.*
It follows that there is a subdivision (X, τ) of (X, σ), and there is an extension of the definition of the groupings of E for Vertex(X, σ) to the groupings of E for Vertex$(X, \tau) \supset$ Vertex(X, σ) such that the following properties hold.
For each newly introduced vertex $x \in$ Vertex(X, τ), if $x \in \Delta$, where Δ is a simplex of (X, σ)(before subdivision), then

$$(1) \qquad \bigcap_{y \in \mathit{Vertex}(\Delta,\sigma)} E_j(y) \subset E_j(x), \qquad j = 1, 2, \cdots, m,$$

and

$$(2) \qquad E_j(x) \subset \bigcup_{y \in \mathit{Vertex}(\Delta,\sigma)} E_j(y), \qquad j = 1, 2, \cdots, m.$$

For any simplex Δ_1 *of* (X,τ) *(after subdivision),*

$$\#(\bigcap_{x\in \mathit{Vertex}(\Delta_1)} E_1(x)) \geq K_1 - \frac{n(n+1)}{2}, \tag{3}$$

where $n = \dim X$.
(The above (1) and (2) imply that for any $x \in \mathrm{Vertex}(X,\tau)$,

$$\bigcap_{y\in\mathrm{Vertex}(X,\sigma)} E_j(y) \subset E_j(x) \subset \bigcup_{y\in\mathrm{Vertex}(X,\sigma)} E_j(y), \qquad j = 1,2,\cdots,m.)$$

Proof: The proof is exactly the same as that of Corollaries 3.13 and 3.14. In fact, in the proof of Corollary 3.13, we were working simplex by simplex from the lowest dimension to the highest dimension. As same as Corollary 3.13, when we work on simplex Δ, we should suppose that, we have already done with $\partial\Delta$. The only difference is the following. We should choose the sets H_i, D_i differently according to the simplex we are working on. For simplex Δ, choose $H_i = \bigcup_{y\in\mathrm{Vertex}(\Delta)} E_i(y)$, $D_i = \bigcap_{y\in\mathrm{Vertex}(\Delta)} E_i(y)$, $i = 1,2\cdots,m$. $\square$

Lemma 3.4 is a special case of the following theorem.

THEOREM 3.16. *Suppose that* (X,σ) *is a simplicial complex. Suppose that for each vertex* $x \in$ *Vertex*(X,σ)*, there is a grouping* $E_1(x), E_2(x),\cdots,E_m(x)$ *of* E.
It follows that there is a subdivision (X,τ) *of* (X,σ)*, and there is an extension of the definition of the groupings of* E *for Vertex*(X,σ) *to the groupings of* E *for Vertex*$(X,\tau) \supset$ *Vertex*(X,σ) *such that the following properties hold.*
For each newly introduced vertex $x \in$ *Vertex*(X,τ)*, if* $x \in \Delta$*, where* Δ *is a simplex of* (X,σ) *(before subdivision), then*

$$\bigcap_{y\in \mathit{Vertex}(\Delta,\sigma)} E_j(y) \subset E_j(x), \qquad j = 1,2,\cdots,m, \tag{1}$$

and

$$E_j(x) \subset \bigcup_{y\in \mathit{Vertex}(\Delta,\sigma)} E_j(y), \qquad j = 1,2,\cdots,m. \tag{2}$$

For any simplex Δ_1 *of* (X,τ) *(after subdivision),*

$$\#(\bigcap_{x\in \mathit{Vertex}(\Delta_1)} E_j(x)) \geq K_j - \frac{n(n+1)}{2}, \quad j = 1,2,\cdots,m, \tag{3}$$

where $n = \dim X$.

Proof: We will apply Corollary 3.15 to prove our theorem. First we can apply

Corollary 3.15 to E_1 and (X,σ) to make the condition (3) of the theorem hold for E_1 and any simplex of the subdivision, and also the conditions (1) and (2) of the theorem hold. We call the simplicial structure after this step, τ_1.
Then we apply Corollary 3.15 to E_2 (in place of E_1) and (X,τ_1) (in place of (X,σ)). We call this new subdivision τ_2. Now (3) for E_2 holds for any simplex of new subdivision τ_2. Furthermore (1) and (2) of Corollary 3.15 hold for (X,τ_1) as the simplicial structure before subdivision (i.e., in place of (X,σ)) and (X,τ_2) as the subdivision (i.e., in place of (X,τ)).
The important point is, (3) for E_1 holds for any simplex Δ_2 of (X,τ_2), because (3) for E_1 holds for the simplex Δ_1 of (X,τ_1) which supports Δ_2 (i.e., $\Delta_1 \supset \Delta_2$), and because (1) holds for τ_1 (in place of σ) and τ_2 (in place of τ). So, now (3) holds for both E_1 and E_2.
(It is also obvious that (1) for σ and τ_2 (in place of τ) follows from (1) for σ and τ_1 (in place of τ), together with (1) for τ_1 (in place of σ) and τ_2 (in place of τ). The same thing also holds for (2).)
Repeating this procedure, we can define τ_3, τ_4, and so on, until τ_m. Then (1), (2), (3) hold for σ and τ_m and any $E_j,\ j=1,2,\cdots,m$. Let $\tau=\tau_m$.

$\square$

Remark 3.17. Let us remark that, in the proof of Lemma 3.11 when we construct the sets E_1, simplex by simplex for (X,σ), it is impossible to obtain

$$\#\left(\bigcap_{x\in \mathrm{Vertex}(\Delta')} E_1(x)\right) \quad \geq \quad K_1-\frac{\dim(\Delta')(\dim(\Delta')+1)}{2}\ ,$$

for each simplex Δ' of subdivision (X,τ) of (X,σ). (Explained below.)
But from the proof of Corollary 3.12, we can make the following hold,

$$\#\left(\bigcap_{x\in \mathrm{Vertex}(\Delta')} E_1(x)\right) \quad \geq \quad K_1-\frac{\dim(\Delta)(\dim(\Delta)+1)}{2}\ ,$$

where Δ is any simplex of (X,σ) which support Δ' (i.e., $\Delta'\subset\Delta$ as spaces). In other words,

$$\#\left(\bigcap_{x\in \mathrm{Vertex}(\Delta')} E_1(x)\right) \quad \geq \quad K_1-\frac{l(l+1)}{2}\ ,$$

if Δ' is a subspace of l-skeleton $X^{(l)}$ of (X,σ).
In the induction construction from dimension not larger than $n-1$ to dimension n (see the proof of Lemma 3.11), in particular, from $(\partial\Delta_0\times\{t_0\},\tilde{\tau})$ to $(\partial\Delta_0\times[t_0,t_1],\tilde{\tau})$, for any simplex Δ inside one of $(\partial\Delta_0\times\{t_0\},\tilde{\tau})$ and $(\partial\Delta_0\times\{t_1\},\tilde{\tau})$, we do have

$$\bigcap_{x\in \mathrm{Vertex}(\Delta)} E_1(x)\geq K_1-\frac{(n-1)n}{2},$$

from our construction (see condition (b) in the proof of Lemma 3.11). But for simplices Δ which are not completely sitting inside one of $(\partial\Delta_0 \times \{t_0\}, \tilde{\tau})$ and $(\partial\Delta_0 \times \{t_1\}, \tilde{\tau})$, we do NOT have

$$\bigcap_{x\in \mathrm{Vertex}(\Delta)} E_1(x) \geq K_1 - \frac{(n-1)n}{2},$$

even if we assume $dim(\Delta) \leq n-1$.
For the application we have in mind, we need the following strengthened form of Theorem 3.16 (in fact, we will need the version of the following result which allow multiplicities; see Theorem 3.32).

Theorem 3.18. *Let (X,σ) be a simplicial complex and $Y = X^{(l)}$ be the l-skeleton of X. Suppose that there is a subdivision (Y,τ) of (Y,σ) and a grouping for each vertex of (Y,τ) (and (X,σ)), such that*
(a) if Δ is a simplex of (Y,τ), then

$$\#\left(\bigcap_{y\in \mathit{Vertex}(\Delta,\tau)} E_j(y)\right) \geq K_j - \frac{l(l+1)}{2}\ . \qquad j = 1, 2, \cdots, m;$$

(b) if Δ is a simplex of $(Y,\sigma) \subset (X,\sigma)$, and $y \in \Delta$ is a vertex of (Y,τ), then

$$\bigcap_{x\in \mathit{Vertex}(\Delta,\sigma)} E_j(x) \subset E_j(y) \subset \bigcup_{x\in \mathit{Vertex}(\Delta,\sigma)} E_j(x), \qquad j = 1, 2, \cdots, m.$$

It follows that there is a subdivision $(X,\tilde{\tau})$ of (X,σ) and groupings for all the vertices, such that
(1) $(X,\tilde{\tau})|_Y = (Y,\tau)$, and groupings on Vertex(Y,τ) are the same as the old ones.
(2) if Δ is a simplex of (X,σ), and $x_1 \in \Delta$ is a newly introduced vertex of $(X,\tilde{\tau})$, then

$$\bigcap_{x\in \mathit{Vertex}(\Delta,\sigma)} E_j(x) \subset E_j(x_1) \subset \bigcup_{x\in \mathit{Vertex}(\Delta,\sigma)} E_j(x), \qquad j = 1, 2, \cdots, m;$$

(3) for each simplex Δ of $(X,\tilde{\tau})$, if Δ is inside the l'-skeleton $(X,\sigma)^{(l')}$ $(l' \geq l)$ of (X,σ), then

$$\#\left(\bigcap_{x\in \mathit{Vertex}(\Delta,\tilde{\tau})} E_j(x)\right) \geq K_j - \frac{l'(l'+1)}{2}\ . \qquad j = 1, 2, \cdots, m.$$

Proof: If one does not require (1) (i.e., if it is allowed to introduce more vertices into (Y,τ)), then the theorem is Theorem 3.16 (see 3.17 also).

Recall, in the proof of Theorem 3.16, we first constructed a subdivision (X, τ_1) and the groupings to make the above (3) hold for E_1. Then based on (X, τ_1), we constructed a new subdivision (X, τ_2) and groupings to make the above (3) hold also for E_2, and so on. If we use the same procedure to prove Theorem 3.18, we will encounter a difficulty in the second step. We have no problem for the first step, since we can begin with what we already have on $(X^{(l)}, \tau)$ and work on each of the simplexes of dimension larger than l (see Lemma 3.11, the proof of Corollary 3.12 and Remark 3.17). But for the second step, the condition (a) may not hold for l-skeleton $(X, \tau_1)^{(l)}$ of (X, τ_1). So we need to start with the simplex of the lowest dimension, which forced us to introduce vertices on $(Y, \tau) = (X^{(l)}, \tau)$.

The following small trick can be used to avoid the difficulty mentioned above. Consider simplex Δ. Suppose that the subdivision $(\partial\Delta, \tilde{\tau})$ and the groupings for those vertices are chosen. Identify Δ with $\partial\Delta \times [0,1]/\partial\Delta \times \{1\}$ as in the proof of Lemma 3.11. Choose a point $t_0 \in (0,1)$, and write

$$\Delta = \partial\Delta \times [0, t_0] \cup (\partial\Delta \times [t_0, 1]/\partial\Delta \times \{1\}).$$

Substitute Δ by $\Delta^{sub} = \partial\Delta \times [t_0, 1]/\partial\Delta \times \{1\}$. The simplicial structure $\tilde{\tau}_{pre}$ and the groupings on $\partial\Delta^{sub} = \partial\Delta \times \{t_0\}$ should be endowed the same as $\tilde{\tau}$ and the groupings on $\partial\Delta = \partial\Delta \times \{0\}$. Then apply Theorem 3.16 to Δ^{sub}. One may introduce new vertices on $(\partial\Delta \times \{t_0\}, \tilde{\tau}_{pre})$, but no new vertices are introduced on $\partial\Delta = \partial\Delta \times \{0\}$. Finally, for the part $\partial\Delta \times [0, t_0]$, same as in the Step 4 of the proof of Lemma 3.11, we apply Lemma 3.10 to make this part a simplicial complex, in which we do not introduce any new vertices on $\partial\Delta \times \{0\}$.

□

3.19. For convenience, define

$$E_j(\Delta) = \bigcap_{x \in \mathrm{Vertex}(\Delta)} E_j(x), \qquad j = 1, 2, \cdots, m$$

for each simplex Δ of $(X, \tilde{\tau})$. Then (3) of 3.18 becomes

$$\#(E_j(\Delta)) \geq K_j - \frac{l'(l'+1)}{2},$$

if Δ is in the l'-skeleton of (X, σ) $(l' \geq l)$.

3.20. We need a different version of Theorem 3.18 which allows multiplicity. Let $w_1, w_2, \cdots, w_k$ be a k-tuple of positive integers. Let

$$E = \{\lambda_1^{\sim w_1}, \lambda_2^{\sim w_2}, \cdots, \lambda_k^{\sim w_k}\}$$

be an index set with multiplicity and $\lambda_i \neq \lambda_j$ if $i \neq j$. (See 1.1.7 (b) for the notation $\lambda^{\sim w}$.) Let $w_1 + w_2 + \cdots + w_k = K$. (3.2 is a special case with each $w_i =$

1.) Let $K_1, K_2, \cdots, K_m$ be non negative integers with $K_1+K_2+\cdots+K_m = K$. Suppose that

$$E_j = \{\lambda_1^{\sim p_1^j}, \lambda_2^{\sim p_2^j}, \cdots, \lambda_k^{\sim p_k^j}\}, \qquad j = 1, 2, \cdots, m,$$

where p_i^j are nonnegative integers. If $\{E_1, E_2, \cdots, E_m\}$ satisfies

$$\sum_{i=1}^{k} p_i^j = K_j \qquad \text{for each } j = 1, 2, \cdots, m, \quad \text{and}$$

$$\sum_{j=1}^{m} p_i^j = w_i \qquad \text{for each } i = 1, 2, \cdots, k,$$

then we call $\{E_1, E_2, \cdots, E_m\}$ a grouping of E of type $(K_1, K_2, \cdots, K_m)$, or just a grouping of E.

3.21. It is convenient to use the notations of union, intersection etc. for the sets with multiplicity. A is called a subset of E if A is of the form

$$\{\lambda_1^{\sim t_1}, \lambda_2^{\sim t_2}, \cdots, \lambda_k^{\sim t_k}\}$$

with $0 \le t_i \le w_i$, for each $1 \le i \le k$. Note that if all $t_i = 0$, then A is called the empty set. If $t_i = w_i$, then $A = E$. Let B be another subset of E of form $\{\lambda_1^{\sim s_1}, \lambda_2^{\sim s_2}, \cdots, \lambda_k^{\sim s_k}\}$. A is a subset of B (denoted by $A \subset B$) if $t_i \le s_i$ for all i. In general, define the union, intersection, and difference of two subsets A and B of E as follows.

$$A \cup B = \{\lambda_1^{\sim \max(t_1,s_1)}, \lambda_2^{\sim \max(t_2,s_2)}, \cdots, \lambda_k^{\sim \max(t_k,s_k)}\}$$

$$A \cap B = \{\lambda_1^{\sim \min(t_1,s_1)}, \lambda_2^{\sim \min(t_2,s_2)}, \cdots, \lambda_k^{\sim \min(t_k,s_k)}\}$$

$$A \backslash B = \{\lambda_1^{\sim \max(0,t_1-s_1)}, \lambda_2^{\sim \max(0,t_2-s_2)}, \cdots, \lambda_k^{\sim \max(0,t_k-s_k)}\} .$$

(The definitions of union and intersection can be easily generalized to finitely many subsets of E.)
Warning 1: $B \cap (A\backslash B)$ may be a nonempty set.
Warning 2: The assumption that $\{E_1, E_2, \cdots E_m\}$ is a grouping of E does NOT imply that $\cup_{i=1}^m E_i = E$ or that $E_i \cap E_{i'} = \emptyset$ for $i \ne i'$. (See 3.20.)
But $A\backslash B = A\backslash(A \cap B)$ still holds.

3.22. Let (X, σ) be a simplicial complex. Suppose that each vertex $x \in X$ is associated with a grouping $\{E_1, E_2, \cdots, E_m\}$, satisfying

$$\#(E_j) = \sum_{i=1}^{k} p_i^j = K_j.$$

(Recall that, for the notation of #, we count multiplicity.)
In Theorem 3.18, we introduced subdivisions (X,τ) of (X,σ), and groupings for newly introduced vertices to make $\#(E_j(\Delta))$ large, for all simplices Δ of (X,τ). In order to prove the decomposition theorem in the next section, we need a stronger result, since the multiplicity of the spectrum of the homomorphism is involved. (See §2. We can not always perturb the map to have distinct spectrum, like the one dimensional case.) Fortunately, this stronger result can be proved in the same way as that for Theorem 3.18, with a few modifications. For any subset $F = \{\lambda_1^{\sim u_1}, \lambda_2^{\sim u_2}, \cdots, \lambda_k^{\sim u_k}\} \subset E$, define

$$\mathring{F} = \{\lambda_1^{\sim v_1}, \lambda_2^{\sim v_2}, \cdots, \lambda_k^{\sim v_k}\},$$

where

$$v_i = \begin{cases} u_i & \text{if } u_i = w_i \\ 0 & \text{if } u_i < w_i \ . \end{cases}$$

That is, $\mathring{F}$ is the set of all those elements λ_i, which are entirely inside F. Evidently,

$$\mathring{E}_j(\Delta) = \bigcap_{x \in \text{Vertex}(\Delta)} \mathring{E}_j(x).$$

Instead of the condition that $E_j(\Delta)$ is large (see 3.18 and 3.19), we need to make $\mathring{E}_j(\Delta)$ large for any simplex Δ of $(X,\tilde{\tau})$. For this purpose, $\mathring{E}_j(x)$ should be large for each vertex of (X,σ) at the beginning.

3.23. For each set $F = \{\lambda_1^{\sim u_1}, \lambda_2^{\sim u_2}, \cdots, \lambda_k^{\sim u_k}\} \subset E$, define

$$\bar{F} = \{\lambda_1^{\sim v_1}, \lambda_2^{\sim v_2}, \cdots, \lambda_k^{\sim v_k}\}$$

where

$$v_i = \begin{cases} w_i & \text{if } \ u_i > 0 \\ 0 & \text{if } \ u_i = 0 \ . \end{cases}$$

Obviously,

$$\mathring{F} \subset F \subset \bar{F}.$$

3.24. Let $H_1, H_2, \cdots, H_m$ (not necessarily disjoint) be finite subsets of E satisfying condition $(*)$ in 3.7, and $E = H_1 \cup H_2 \cup \cdots \cup H_m$. Suppose that

$$\mathring{H}_i = H_i = \bar{H}_i \qquad \text{for each } i = 1, 2, \cdots, m.$$

In what follows, we will require that

$$E_j \subset H_j, \qquad j = 1, 2, \cdots, m$$

(comparing with the condition (2) in Theorem 3.18).

3.25. For each subset $I \subset \{1, 2, \cdots, m\}$, define

$$H_I = \bigcup_{j \in I} H_j.$$

Let $G'_I = \cap_{j \in I} H_j$. Then define

$$G_I = G'_I \setminus \bigcup_{J \supsetneq I} G'_J \ .$$

Another way to define G_I is by

$$G_I = \{\lambda \in E \mid \lambda \in H_i \text{ if and only if } i \in I\}.$$

(Note that G_I may be an empty set for some I.) Obviously,

$$\mathring{G}_I = G_I = \bar{G}_I \subset \mathring{H}_I = H_I = \bar{H}_I.$$

If $I \cap J = \emptyset$, then $H_I \cap G_J = \emptyset$.
Furthermore, for any $\lambda \in E$, there is a unique set I (defined by $I = \{i \mid \lambda \in H_i\}$) such that $\lambda \in G_I$. Hence E is a disjoint union of

$$\{G_I, \emptyset \neq I \subset \{1, 2, \cdots, m\}\}.$$

Similarly, we have

$$H_I = \bigcup_{J \cap I \neq \emptyset} G_J.$$

3.26. Under the above partition G_I of E, two elements $\lambda, \mu \in E$ are in the same part, if and only if the following is true. For any $i = 1, 2, \cdots, m$, either H_i contains both λ and μ, or H_i contains none of λ and μ.
For any $E_1, E'_1 \subset H_1$, if $\#(E_1 \cap G_I) = \#(E'_1 \cap G_I)$ for any $I \subset \{1, 2, \cdots, m\}$, then from the end of 3.25,

$$\#(E_1 \cap H_I) = \sum_{J \cap I \neq \emptyset} \#(E_1 \cap G_J) = \sum_{J \cap I \neq \emptyset} \#(E'_1 \cap G_J) = \#(E'_1 \cap H_I)$$

for any $I \subset \{1, 2, \cdots, m\}$. Hence $\#(H_I \setminus E_1) = \#(H_I \setminus E'_1)$. At this circumstance, either both of E_1 and E'_1 satisfy $(**)$ in 3.8, or both of them do not satisfy $(**)$ in 3.8.
Note that $H_i = \cup_{I \ni i} G_I$. A grouping $\{E_1, E_2, \cdots E_m\}$ satisfies $E_i \subset H_i$ $(i = 1, 2, \cdots m)$ if and only if for any $i \notin I$, $E_i \cap G_I = \emptyset$.
FOR THE REST OF THE SECTION, LET $\Omega \geq \max(w_1, w_2, \cdots, w_m)$ BE A FIXED NUMBER, WHERE $w_1, w_2, \cdots, w_m$ ARE THE MULTIPLICITIES IN E. Note that for our application, sometimes, we have to allow Ω to be larger than the maximum multiplicity.

Assumption 3.27. For each grouping $\{E_1(x), E_2(x), \cdots, E_m(x)\}$, we always assume that

$$\#(\mathring{E}_j(x)) \geq \#(E_j(x)) - M\Omega = K_j - M\Omega, \qquad j = 1, 2, \cdots, m,$$

where $M = 2^m - 1$. We not only require each initial grouping for (X, σ) to satisfy the above assumption, but also require any new groupings for vertices of (X, τ) to satisfy the assumption.
Since $M = 2^m - 1$, there are totally M non-empty subsets $I \subset \{1, 2, \cdots, m\}$. If the grouping $\{E_1, E_2, \cdots, E_m\}$ satisfies

$$\#(\mathring{E}_j \cap G_I) \geq \#(E_j \cap G_I) - \Omega$$

for all $j = 1, 2, \cdots, m$ and for all $I \subset \{1, 2, \cdots, m\}$, then it also satisfies Assumption 3.27.

Lemma 3.28. *If $\{E_1, E_2, \cdots, E_m\}$ is a grouping of E with $E_i \subset H_i$, then there is a grouping $\{E'_1, E'_2, \cdots, E'_m\}$ of E satisfying Assumption 3.27, and*

$$E'_i \subset H_i, \quad E'_i \supset \mathring{E}_i \qquad \textit{for all } \ i = 1, 2, \cdots, m.$$

Proof: The proof is straight forward. Consider $E'_1, E'_2, \cdots, E'_m$ to be m boxes with no element at the beginning, and put each element of E into one of the boxes, following the procedures described below.
Step 1. Put all the elements of $\mathring{E}_i$ into box E'_i for each $i = 1, 2, \cdots, m$. (Thus $E'_i \supset \mathring{E}_i$.)
Step 2. Fix $I \subset \{1, 2, \cdots, m\}$. For the set E'_1, if there is a $\lambda_i \in G_I \backslash (E'_1 \cup E'_2 \cup \cdots \cup E'_m)$ such that

$$\#(E'_1 \cap G_I) + w_i \leq \#(E_1 \cap G_I),$$

where w_i is the multiplicity of λ_i in E, then put the entire set $\{\lambda_i^{\sim w_i}\}$ into E'_1. (Note that if $1 \notin I$, then $E_1 \cap G_I = \emptyset$. Hence for I, we need not do anything for E_1.) Repeat this procedure until no such i exists. Thus, so far,

$$\#(\mathring{E}'_1 \cap G_I) = \#(E'_1 \cap G_I) \geq \#(E_1 \cap G_I) - (\Omega - 1).$$

For the same I above, repeat the above construction for the set E'_2, then E'_3, etc.
After this step has been completed for each I, (it is done for each set I separately) we have the following:

$$\#(\mathring{E}'_j \cap G_I) = \#(E'_j \cap G_I) \geq \#(E_j \cap G_I) - (\Omega - 1).$$

Step 3. Put what left for each G_I from the previous steps, arbitrarily into the boxes

$$E'_1, E'_2, \cdots, E'_m$$

to make the following condition hold:

$$\#(E'_j \cap G_I) = \#(E_j \cap G_I).$$

From the end of 3.26, $E'_i \subset H_i$ is a consequence of the above equation. (Note that $E_i \subset H_i$.) Evidently, $\{E'_1, E'_2, \cdots, E'_m\}$ is as desired.

□

3.29. A set $E_1 (\subset H_1)$ of K_1 elements is said to satisfy the condition $(***)$, if there is a grouping $(E_1, E_2, \cdots, E_m)$ of E (of type $(K_1, K_2, \cdots K_m)$), $E_i \subset H_i$, satisfying Assumption 3.27.
Obviously, $(***)$ implies $(**)$.
The following corollary is a direct consequence of Lemma 3.28.

COROLLARY 3.30. *For any set E_1 $(\subset H_1)$ satisfying $(**)$, there is a set E'_1 $(\subset H_1)$ satisfying $(***)$ such that*

$$E'_1 \supset \mathring{E}_1.$$

Proof: Since E_1 satisfies $(**)$, we can extend E_1 to a grouping $\{E_1, E_2, \cdots, E_m\}$ of E such that $E_i \subset H_i$ for each i. By Lemma 3.28, there is a grouping $\{E'_1, E'_2, \cdots, E'_m\}$ satisfying Assumption 3.27, and $E'_i \subset H_i$ for each i. This is condition $(***)$ for E'_1.

□

LEMMA 3.31. *Let E_1 and F_1 be two sets satisfying condition $(***)$. Suppose that there is a $\lambda \in E$ such that*

$$\{\lambda^{\sim w}\} \subset \mathring{F}_1 \backslash \mathring{E}_1,$$

where w is the multiplicity of λ in E. Then there are (perhaps repeating) elements $\mu_1, \mu_2, \cdots, \mu_t \in E_1 \backslash \mathring{F}_1$, where $t = w - \#(\{\lambda^{\sim w}\} \cap E_1)$, such that

$$E'_1 = (E_1 \cup \{\lambda^{\sim w}\}) \backslash \{\mu_1, \mu_2, \cdots, \mu_t\}$$

satisfies $(**)$ and

$$\begin{aligned} \#((\mathring{E}'_1 \cap \mathring{E}_1) \cap G_I) &\geq \#(\mathring{E}_1 \cap G_I) - (w + \Omega) \\ &\geq \#(\mathring{E}_1 \cap G_I) - 2\Omega \end{aligned}$$

for each $I \subset \{1, 2, \cdots, m\}$. As a consequence,

$$\#(\mathring{E}'_1 \cap \mathring{E}_1) \geq \#(\mathring{E}_1) - 2M\Omega.$$

Proof: Let $t_1 = \#(\{\lambda^{\sim w}\} \cap E_1)$. Then $t_1 < w$. Applying Lemma 3.9 $t := w - t_1$ times, one can obtain a (possibly repeating) set $T' = \{\nu_1, \nu_2, \cdots, \nu_t\} \subset E_1 \backslash F_1 \subset E_1 \backslash \overset{\circ}{F}_1$, such that

$$\tilde{E}_1 = (E_1 \cup \{\lambda^{\sim w}\}) \backslash T'$$

satisfies $(**)$.
From 3.26, if another set $T = \{\mu_1, \mu_2, \cdots, \mu_t\} \subset E_1 \backslash \overset{\circ}{F}_1$, satisfies that

$$\#(T \cap G_I) = \#(T' \cap G_I)$$

for each $I \subset \{1, 2, \cdots, m\}$, then $E_1' = (E_1 \cup \{\lambda^{\sim w}\}) \backslash T$ also satisfies $(**)$.
$T \subset E_1 \backslash \overset{\circ}{F}_1$ will be constructed to satisfy the following condition. For each $I \subset \{1, 2, \cdots, m\}$,

$$\#(T \cap G_I) = \#(T' \cap G_I),$$

and $(T \cap (E_1 \backslash T)) \cap G_I$ is either empty or $\{\mu_i^{\sim s}\}$ for a certain $\mu_i \in \{\mu_1, \mu_2, \cdots, \mu_t\}$. (Note that $T \cap (E_1 \backslash T)$ may not be empty, since we are dealing with sets with multiplicities.)
To do the above, write

$$(E_1 \backslash \overset{\circ}{F}_1) \cap G_I = \{\lambda_{i_1}^{\sim s_1}, \lambda_{i_2}^{\sim s_2}, \cdots, \}.$$

Then put each of the sets $\{\lambda_{i_1}^{\sim s_1}\}, \{\lambda_{i_2}^{\sim s_2}\}, \cdots$, entirely into T one by one until we can not do it without violating the restriction

$$\#(T \cap G_I) \leq \#(T' \cap G_I).$$

Then make T to satisfy $\#(T \cap G_I) = \#(T' \cap G_I)$ by putting part of $\{\lambda_{i_j}^{\sim s_j}\}$ into T if necessary.
Since $\#(T) \leq w$, combining with the above condition for $(T \cap (E_1 \backslash T)) \cap G_I$, after a moment thinking, one can obtain,

$$\#((E_1 \backslash T)^{\circ} \cap G_I) \geq \#(\overset{\circ}{E}_1 \cap G_I) - (w + \Omega).$$

(In fact, $\overset{\circ}{E}_1 \backslash (E_1 \backslash T)^{\circ} \subset \overset{\circ}{T} \cup (T \cap (E_1 \backslash T))$, and $(\overset{\circ}{T} \cup (T \cap (E_1 \backslash T))) \cap G_I$ has at most $w + \Omega$ elements.)
Hence

$$\#((\overset{\circ}{E}{}_1' \cap \overset{\circ}{E}_1) \cap G_I) \geq \#(\overset{\circ}{E}_1 \cap G_I) - (w + \Omega).$$

$\square$

The following is the main result of this section. Together with Lemma 3.28, it will be used in §4.

Theorem 3.32. *Let (X,σ) be a simplicial complex, and $Y = X^{(l)}$, the l-skeleton of X. Suppose that (Y,τ) is a subdivision of (Y,σ) and, for each vertex $y \in$ Vertex(Y,τ), there is a grouping $E_1(y), E_2(y), \cdots, E_m(y)$ of E (of type $(K_1, K_2, \cdots, K_m)$). Suppose that the groupings satisfy the following three conditions:*

(a) For each simplex Δ of (Y,τ), and $i = 1, 2, \cdots, m$,

$$\#(\mathring{E}_i(\Delta)) \geq K_i - (M\Omega + M\Omega \dim Y \cdot (\dim Y + 1)),$$

where $M = 2^m - 1$.
(b) $E_i(x) \subset H_i,\ i = 1, 2, \cdots, m$, for each $x \in$ Vertex(Y,τ).

(c) Each grouping, for a vertex of (Y,τ), satisfies Assumption 3.27.

It follows that there exist a subdivision $(X,\tilde{\tau})$ of (X,σ) and a grouping for each vertex of $(X,\tilde{\tau})$, satisfying the following conditions.
(1) $(X,\tilde{\tau})|_Y = (Y,\tau)$, and each grouping on Vertex(Y,τ) is as same as the old one.
(2) $E_i(x) \subset H_i,\ i = 1, 2, \cdots, m$, for each $x \in$ Vertex$(X,\tilde{\tau})$, and if Δ is a simplex of (X,σ) (before the subdivision), and $x \in \Delta$ is a newly introduced vertex of $(X,\tilde{\tau})$, then

$$\mathring{E}_j(x_1) \ \supset \bigcap_{y \in \mathit{Vertex}(\Delta \cap Y, \tau)} \mathring{E}_j(y).$$

(3) For each simplex Δ of $(X,\tilde{\tau})$, if Δ is inside the l'-skeleton $(X,\sigma)^{(l')}$, $(l' > l)$, of (X,σ), then

$$\#(\mathring{E}_j(\Delta)) \geq K_j - (M\Omega + M\Omega l'(l'+1)).$$

(4) Each grouping on Vertex$(X,\tilde{\tau})$ satisfies Assumption 3.27.

Proof: (Sketch) The proof is the same as the one of 3.18 (see 3.11 to 3.18), using Lemma 3.31 to replace Lemma 3.9. The arguments in 3.12 – 3.18 are easily adopted in this new setting. We only give the proof for the part corresponding to 3.11 and sketch the differences for other parts.
As in 3.11, consider only one simplex $X = \Delta_0$ with $Y = \partial\Delta_0$, and only one set $E_1(x)$.
Similar to Step 1 of 3.11, choose E_1^{model} to satisfy condition $(***)$ and

$$\mathring{E}_1^{\,model} \ \supset \bigcap_{x \in \mathrm{Vertex}(\partial\Delta_0, \tau)} \mathring{E}_1(x).$$

Replace (a) of 3.11 by

$$\bigcap_{x \in \mathrm{Vertex}(\partial\Delta \times \{t_i\}, \tilde{\tau})} (\mathring{E}_1(x) \cap \mathring{E}_1^{\,model}) \ \underset{\neq}{\supset} \bigcap_{x \in \mathrm{Vertex}(\partial\Delta \times \{t_{i-1}\}, \tilde{\tau})} (\mathring{E}_1(x) \cap \mathring{E}_1^{\,model}) \ .$$

Keeping the notations in 3.11, in Step 2, replace G by

$$G=\mathring{E}_1(y_1,t_0)\cap\mathring{E}_1(y_2,t_0)\cap\cdots\cap\mathring{E}_1(y_p,t_0)\cap\mathring{E}_1{}^{model}.$$

If $G=\mathring{E}_1{}^{model}$, then define

$$E_1(y_i,t_1)=E_1{}^{model}, \qquad i=1,2,\cdots,p.$$

Suppose that $G\neq\mathring{E}_1{}^{model}$. Choose $\lambda\in\mathring{E}_1{}^{model}\backslash G$. Let w be the multiplicity of λ. Then

$$\{\lambda^{\sim w}\}\subset\mathring{E}_1{}^{model}\backslash G.$$

For each (y_i,t_0), if $\lambda\in\mathring{E}_1(y_i,t_0)$, then define $E_1(y_i,t_1)=E_1(y_i,t_0)$ (as in Step 2 of 3.11). If $\lambda\notin\mathring{E}_1(y_i,t_0)$, apply Lemma 3.31 to obtain E_1' satisfying$(**)$, $\mathring{E}_1'\supset G\cup\{\lambda^{\sim w}\}$ and

$$(A) \qquad \#(\mathring{E}_1'\cap\mathring{E}_1)\geq\#(\mathring{E}_1)-2M\Omega.$$

Then we can apply Corollary 3.29 to find E_1'' satisfying $(***)$ and $\mathring{E}_1''\supset\mathring{E}_1'$. Define

$$E_1(y_i,t_1)=E_1''.$$

Then

$$\mathring{E}_1(y_i,t_1)\supset\mathring{E}_1'\supset G\cup\{\lambda^{\sim w}\}.$$

The arguments in Step 3 and Step 4 of 3.11 can also be employed here. (Of course, at many places (not all places), one needs to replace E_i by $\mathring{E}_i$.) The estimation (e) in Step 4 of 3.11 will be changed to

$$\begin{aligned}\#\Big(\mathring{E}_1(u_0,t_0)\cap\mathring{E}_1(u_1,t_0)\cap\cdots\cap\mathring{E}_1(u_i,t_0)\cap\mathring{E}_1(u_0,t_1)\cap\mathring{E}_1(u_1,t_1)\cap\cdots\cap\mathring{E}_1(u_i,t_1)\Big)&\\ \geq K_1-[M\Omega+M\Omega\dim Y\cdot(\dim Y+1)]-2M\Omega(i+1)&\\ =K_1-[M\Omega+M\Omega\cdot(n-1)\cdot n]-2M\Omega\cdot n&\\ =K_1-[M\Omega+M\Omega\cdot n\cdot(n+1)].&\end{aligned}$$

(Here we used the above estimation (A) which is from Lemma 3.31.)
Since $E_1(y_i,t_1)$ satisfies $(***)$, all the other parts (e.g., induction arguments) in 3.11—3.18 can go through easily. In the part corresponding to the proof of Corollary 3.14, the definition of D_i should be changed to

$$D_i=\bigcap_{x\in\mathrm{Vertex}(\Delta)}\mathring{E}_i(x).$$

$\square$

We remark that in §4, we will only use the theorem of the case that $X = \Delta$, a single simplex with $Y = \partial\Delta$.

Remark 3.33. The condition $E_i(x) \subset H_i$ in (2) can be strengthened as

$$E_i(x) \subset \bigcup_{y \in \mathrm{Vertex}(\Delta \cap Y, \tau)} \bar{E}_i(y),$$

where Δ is a simplex of (X, σ) (before the subdivision) such that $x \in \Delta$ is a newly introduced vertex. (The notation is from 3.23.)

4 Decomposition Theorems

In this section, we will prove the decomposition theorems which are needed for the proof of our main Reduction Theorem and the main results in [EGL]. The following Theorem 4.1 is one version of the Decomposition Theorem. After Theorem 4.1 has been proved, we will use [Li 2] to verify that the condition of Theorem 4.1 holds for connecting homomorphisms $\phi_{n,m}$(for each fixed n, m should be large enough), with the maps $a_1, a_2, \cdots a_L$(see below) factoring through interval $[0,1]$ or the single point space $\{pt\}$. In such a way, we can prove our main decomposition theorems (Theorem 4.35 and Theorem 4.37).

Theorem 4.1. *Let X be a connected finite simplicial complex, and $F \subset C(X)$ be a finite set which generates $C(X)$. For any $\varepsilon > 0$, there is an $\eta > 0$ such that the following statement is true.*
Suppose that a unital homomorphism $\phi : C(X) \to PM_{K'}(C(Y))P$ (rank$(P) = K$) (where Y is a finite simplicial complex) satisfies the following condition: There are L continuous maps

$$a_1, a_2, \cdots, a_L : \; Y \longrightarrow X$$

such that for each $y \in Y$, $SP\phi_y$ and $\Theta(y)$ can be paired within η, where

$$\Theta(y) = \{a_1(y)^{\sim T_1}, a_2(y)^{\sim T_2}, \cdots, a_L(y)^{\sim T_L}\}$$

and $T_1, T_2, \cdots, T_L$ are positive integers with

$$T_1 + T_2 + \cdots + T_L = K = \mathit{rank}(P).$$

(See 1.1.7(b) for notation $x^{\sim T_i}$.) Let $T = 2^L(\dim X + \dim Y)^3$. It follows that there are L mutually orthogonal projections $p_1, p_2, \cdots, p_L \in PM_{K'}(C(Y))P$ such that
(i) $\|\phi(f)(y) - p_0(y)\phi(f)(y)p_0(y) \oplus \sum_{i=1}^{L} f(a_i(y))p_i(y)\| < \varepsilon$, for any $f \in F$ and $y \in Y$, where $p_0 = P - \sum_{i=1}^{L} p_i$;
(ii) $\|p_0(y)\phi(f)(y) - \phi(f)(y)p_0(y)\| < \varepsilon$ for any $f \in F$ and $y \in Y$;
(iii) rank$(p_i) \geq T_i - T$ for $1 \leq i \leq L$, and hence rank$(p_0) \leq LT$.

4.2. In the above theorem, some of the p_i may be zero projections if $T_i \leq T$. But when the theorem is applied later in this article, the positive integers T_i are always very large compared with $T = 2^L(\dim X + \dim Y)^3$.
The proof of this theorem will be divided into several small steps. The results in §2 and §3 will be used. In fact, §2 and §3 will only be used in the proof of Theorem 4.1, no other place in this paper or [EGL]. (The results in §2 have some other applications.)

4.3. The theorem is trivial if $X = \{pt\}$, applying Theorem 1.2 of [Hu,Chapter 8]. Without loss of generality, we assume that $X \neq \{pt\}$. By the results in §2, we can assume that ϕ has maximum spectral multiplicity at most $\Omega := \dim X + \dim Y$.

4.4. For any $\varepsilon > 0$, there is an $\eta > 0$ such that for any $x_1, x_2 \in X$, if $\mathrm{dist}(x_1, x_2) < 2\eta$, then

$$|f(x_1) - f(x_2)| < \frac{\varepsilon}{3} \qquad\qquad \text{for all } f \in F.$$

We will prove that this η is as desired.

4.5. Recall, from 1.2.5, for any positive integer n, $P^n X$ is the symmetric product of n-copies of X. Also, any element $\Lambda \in P^K X$ can be considered as a set with multiplicity. So $\mathrm{SP}\phi_y \in P^K X$.
Suppose that $\Lambda_1 \in P^{k_1}X, \Lambda_2 \in P^{k_2}X, \cdots, \Lambda_t \in P^{k_t}X$. Write

$$\Lambda_1 = \{\lambda_1, \lambda_2, \cdots, \lambda_{k_1}\}$$

$$\Lambda_2 = \{\lambda_{k_1+1}, \lambda_{k_1+2}, \cdots, \lambda_{k_1+k_2}\}$$

$$\vdots$$

$$\Lambda_t = \{\lambda_{k_1+\cdots+k_{t-1}+1}, \cdots, \lambda_{k_1+\cdots+k_t}\}$$

as sets with multiplicity. By abusing the notation, we use $\{\Lambda_1, \Lambda_2, \cdots, \Lambda_t\}$ to denote

$$\{\lambda_1, \lambda_2, \cdots, \lambda_{k_1},\ \lambda_{k_1+1},\ \cdots,\ \lambda_{k_1+k_2},\ \cdots\cdots,\ \lambda_{k_1+k_2+\cdots+k_t}\},$$

which defines an element in $P^{k_1+k_2+\cdots+k_t}X$.
(Note that $\{\Lambda_1, \Lambda_2, \cdots, \Lambda_t\} = \Lambda_1 \cup \Lambda_2 \cup \cdots \cup \Lambda_t$, if $\{\Lambda_j\}$ are mutually disjoint. See 3.21 for the definition of unions of sets with multiplicity.)

4.6. For any fixed point $y \in Y$, write

$$\mathrm{SP}\phi_y = \{\lambda_1^{\sim w_1}, \lambda_2^{\sim w_2}, \cdots, \lambda_k^{\sim w_k}\}$$

with $\lambda_i \neq \lambda_j$ if $i \neq j$. Note that $w_1 + w_2 + \cdots + w_k = K$. We denote the above k by t_y to indicate that this integer depends on y. Define

$$(a) \qquad \theta(y) = \frac{1}{4} \min_{1 \le i < j \le t_y} \operatorname{dist}(\lambda_i, \lambda_j).$$

Then $\theta(y) > 0$ for any $y \in Y$. (Of course, $\theta(y)$, in general, does not depend on y continuously.)
For each $i = 1, 2, \cdots, t_y$, there is an open set $U(y,i) \ni \lambda_i$ such that

$$(b) \qquad \textit{diameter}(U(y,i)) \le \min\left(\frac{\eta}{2(\dim Y + 1)}, \theta(y)\right).$$

Then, obviously,

$$(c) \qquad \operatorname{dist}(U(y,i), U(y,j)) \ge 2\theta(y) \qquad \text{if} \qquad i \neq j.$$

Applying Lemma 1.2.10, there is an (connected) open neighborhood $O(y)$ of y such that

$$\mathrm{SP}\phi_{y'} \subset U(y,1) \bigcup U(y,2) \bigcup \cdots \bigcup U(y,t_y)$$

for all $y' \in O(y)$. Define the continuous maps

$$\Lambda_1 : O(y) \longrightarrow P^{w_1} U(y,1) \ (\subset P^{w_1} X)$$

$$\Lambda_2 : O(y) \longrightarrow P^{w_2} U(y,2) \ (\subset P^{w_2} X)$$

$$\vdots$$

$$\Lambda_{t_y} : O(y) \longrightarrow P^{w_{t_y}} U(y,1) \ (\subset P^{w_{t_y}} X)$$

by $\Lambda_i(y') = \mathrm{SP}\phi_{y'} \cap U(y,i)$. Then

$$\mathrm{SP}\phi_{y'} = \{\Lambda_1(y'),\ \Lambda_2(y'),\ \cdots,\ \Lambda_{t_y}(y')\}$$

for each $y' \in O(y)$. Later on, we will use the disjoint open cover

$$U(y,1) \bigcup U(y,2) \bigcup \cdots \bigcup U(y,t_y) \supset \mathrm{SP}\phi_{y'}$$

of $\mathrm{SP}\phi_{y'}$ to decompose $\mathrm{SP}\phi_{y'}$ into a disjoint union of $\mathrm{SP}\phi_{y'} \cap U(y,t)$ and to identify the elements in each set $\mathrm{SP}\phi_{y'} \cap U(y,t)$ as a single element with multiplicity w_t.
We further require that $O(y)$ is so small that

$$(c) \qquad \operatorname{diameter}(a_i(O(y))) \le \frac{\eta}{2(\dim Y + 1)},$$

where $a_i : Y \to X$ is any one of the continuous maps $a_1, a_2, \cdots, a_L$, appeared in Theorem 4.1.

4.7. Considering the open cover $\{O(y)\}_{y\in Y}$ of Y, where the open sets $O(y)$ are from 4.6, there exists a finite sub-cover

$$\mathcal{O} = \{O_1, O_2, \cdots, O_\bullet\} \subset \{O(y)\}_{y\in Y},$$

of Y.
Without loss of generality, we assume that the simplicial complex structure (Y,σ) of Y satisfies the following condition, because we can always refine it if necessary.
For each simplex Δ of (Y,σ), the closure of

$$\mathrm{Star}(\Delta) := \bigcup_{\Delta'\cap\Delta\neq\emptyset} \mathrm{interior}(\Delta')$$

can be covered by an open set $O_i \in \mathcal{O}$, where

$$\mathrm{interior}(\Delta') = \Delta'\backslash\partial\Delta'.$$

(Note that $\mathrm{Star}(\Delta)$ is an open set, see 1.4.2)

4.8. For each $y \in Y$, in order to construct $p_1(y), p_2(y), \cdots, p_L(y)$, as in Theorem 4.1, we need to split $\mathrm{SP}\phi_y$ into L sets $E_1(y), E_2(y), \cdots, E_L(y)$ such that each set $E_i(y)$ is contained in an open ball of $a_i(y)$ with small radius (smaller than 2η) and that $\#(E_i(y)) = T_i$, $1 \le i \le L$, where a_i and T_i are maps and positive integers appeared in Theorem 4.1. Since $E_1(y), E_2(y), \cdots, E_L(y)$ may have non-empty intersection (because of the multiplicity of the spectrum), we need to introduce certain subsets of them, which are $\mathring{E}_1 \subset E_1(y)$, $\mathring{E}_2 \subset E_2(y)$, $\cdots$, $\mathring{E}_L \subset E_L(y)$, in the notations of 3.22. This will become precise when the index set with multiplicity is introduced later. The projections $p_1(y), p_2(y), \cdots, p_L(y)$, to be constructed, will be certain sub-projections of the spectral projections corresponding to $\mathring{E}_1, \mathring{E}_2, \cdots, \mathring{E}_L$, respectively. (See Definition 1.2.8 for the spectral projection.)
Following §3, a split of $\mathrm{SP}\phi_y$ into L sets $E_1(y), E_2(y), \cdots, E_L(y)$ will be called a GROUPING of $\mathrm{SP}\phi_y$. THE WORD "GROUPING" IS RESERVED ONLY FOR THIS PURPOSE.
Recall, from 4.6, $\mathrm{SP}\phi_{y'}$ can be written as a disjoint union

$$\mathrm{SP}\phi_{y'} = (\mathrm{SP}\phi_{y'} \cap U(y,1)) \bigcup (\mathrm{SP}\phi_{y'} \cap U(y,2)) \bigcup \cdots \bigcup (\mathrm{SP}\phi_{y'} \cap U(y,t_y)).$$

And the elements in each set $\mathrm{SP}\phi_{y'} \cap U(y,t)$ can be identified as a single element with multiplicity. This will serve as the index set for the groupings. To avoid confusion, the above decomposition is NOT called a "grouping" of $\mathrm{SP}\phi_{y'}$. It is called a DECOMPOSITION instead.
In the next few paragraphs, we apply §3 to construct a subdivision (Y,τ) of (Y,σ) and useful groupings for all vertices $y \in \mathrm{Vertex}(Y,\tau)$.

4.9. Let Δ be a simplex of (Y,σ) and

$$O(y_1), O(y_2), \cdots, O(y_i),$$

the list of all open sets in $\mathcal{O}$, each of which covers Δ. Suppose that

$$\theta(y_1) \le \theta(y_2) \le \cdots \le \theta(y_i).$$

From 4.6, for any $y \in \Delta \subset \cap_{k=1}^{i} O(y_k)$, and $j \in \{1,2,\cdots,i\}$,

$$\mathrm{SP}\phi_y \subset U(y_j,1) \bigcup U(y_j,2) \bigcup \cdots \bigcup U(y_j,t_{y_j}).$$

Claim: If $j < j' \in \{1,2,\cdots,i\}$, then each open set $U(y_j,t)$ $(t = 1,2\cdots,t_{y_j})$ intersects with at most one of $\{U(y_{j'},s)\}_{s=1}^{t_{y_{j'}}}$.
Proof of the Claim: Suppose that the claim is not true, that is, for some $t \in \{1,2,\cdots,t_{y_j}\}$, there are two different $s_1, s_2 \in \{1,2\cdots,t_{y_{j'}}\}$ such that

$$U(y_j,t)\cap U(y_{j'},s_1) \ne \emptyset \quad \text{and} \quad U(y_j,t)\cap U(y_{j'},s_2) \ne \emptyset.$$

Together with the fact that diameter$(U(y_j,t)) \le \theta(y_j)$ (see (b) in 4.6), it yields

$$dist(U(y_{j'},s_1), U(y_{j'},s_2)) \le \theta(y_j).$$

This contradicts with (c) in 4.6 which gives

$$dist(U(y_{j'},s_1), U(y_{j'},s_2)) \ge 2\theta(y_{j'}) > \theta(y_j).$$

(Recall $\theta(y_j) \le \theta(y_{j'})$.) This proves the claim.
Still suppose that $j < j'$. From the claim, we have the following. For each $y \in \Delta$, if two different elements of $\mathrm{SP}\phi_y$ are identified as a single element in the decomposition

$$U(y_j,1) \bigcup U(y_j,2) \bigcup \cdots \bigcup U(y_j,t_{y_j})$$

(i.e., if these two elements are in the same open set $U(y_j,t)$ for some $t \in \{1,2,\cdots,t_{y_j}\}$), then these two elements are also identified as a single element in the decomposition

$$U(y_{j'},1) \bigcup U(y_{j'},2) \bigcup \cdots \bigcup U(y_{j'},t_{y_{j'}})$$

(i.e., these two elements are also in the same open set $U(y_{j'},s)$ for some $s \in \{1,2,\cdots,t_{y_{j'}}\}$).
Therefore, the decompositions of $\mathrm{SP}\phi_y$ corresponding to y_1 and y_i are the finest and coarsest decompositions, respectively, among all the above decompositions (corresponding to $y_1, y_2, \cdots, y_i$). The coarsest decomposition will be used to decompose $\mathrm{SP}\phi_y$ into several sets. The elements in each of the sets will be

identified as a single element with multiplicity. Denote $\theta(y_i)$ by $\theta(\Delta)$. (Recall that $O(y_1), O(y_2), \cdots, O(y_i)$ is the list of all open sets in $\mathcal{O}$, each of which covers Δ. Therefore, $\theta(y_i)$ — the maximum of all $\{\theta(y_j)\}_{j=1}^{i}$ — depends only on Δ.)
Introduce the following notations.

$$\Lambda(\Delta,1)(y) = U(y_i,1) \cap \mathrm{SP}\phi_y,$$
$$\Lambda(\Delta,2)(y) = U(y_i,2) \cap \mathrm{SP}\phi_y,$$
$$\vdots$$
$$\Lambda(\Delta,t_\Delta)(y) = U(y_i,t_{y_i}) \cap \mathrm{SP}\phi_y,$$

where $t_\Delta = t_{y_i}$. Recall (see 4.6) that $\mathrm{SP}\phi_{y_i}$ is written as

$$\mathrm{SP}\phi_{y_i} = \{\lambda_1^{\sim w_1}, \lambda_2^{\sim w_2}, \cdots, \lambda_k^{\sim w_k}\},$$

where $k = t_\Delta = t_{y_i}$. Since $y \in \Delta \subset O(y_i)$,

$$\#(\Lambda(\Delta,t)(y)) = w_t, \qquad 1 \le t \le t_\Delta,$$

counting multiplicity. Define set

$$\Lambda(\Delta) = \{\Lambda(\Delta,1)^{\sim w_1}, \Lambda(\Delta,2)^{\sim w_2}, \cdots, \Lambda(\Delta,k)^{\sim w_k}\},$$

where $k = t_\Delta$. That is, identify all the elements of $\mathrm{SP}\phi_y$ in $\Lambda(\Delta,t)(y)$ as a single element (denoted by $\Lambda(\Delta,t)$) with the multiplicity.
As above, we will use $\Lambda(\Delta,t)(y)$ for two purposes. It is a subset of $\mathrm{SP}\phi_y$, or it is a single element in $\Lambda(\Delta)$ which repeats w_t times.
Strictly speaking, w_t $(t = 1,2,\cdots,t_\Delta)$ should be written as $w_t(\Delta)$, and the set $\Lambda(\Delta)$ should be written as
$\{\Lambda(\Delta,1)^{\sim w_1(\Delta)}, \Lambda(\Delta,2)^{\sim w_2(\Delta)}, \cdots, \Lambda(\Delta,k)^{\sim w_k(\Delta)}\}$. When there is a danger of confusion, we will use $w_t(\Delta)$ instead of w_t.

4.10. Let $Y' \subset Y$ be a path connected subspace. Usually we will let Y' be either an open or a closed subset. Suppose that there are positive integers $u_1, u_2, \cdots, u_t$ and continuous maps

$$A(Y',i) : Y' \to P^{u_i}X, \quad i = 1,2,\cdots,t,$$

such that $\{\mathrm{SP}\phi_y\}_{y\in Y'}$ can be decomposed as

$$\mathrm{SP}\phi_y = \{A(Y',1)(y), A(Y',2)(y), \cdots, A(Y',t)(y)\}$$

for all $y \in Y'$. We say that THE ABOVE DECOMPOSITION OF $\{\mathrm{SP}\phi_y\}_{y\in Y'}$ SATISFIES THE CONDITION (S) (S stands for separation) if
(S): there are mutually disjoint open sets $U_1, U_2, \cdots, U_t \subset X$ satisfying

$$A(Y',i)(y) \subset U_i, \qquad \forall y \in Y',\ i = 1,2,\cdots,t.$$

Define
$$A(Y') = \{A(Y',1)^{\sim u_1}, A(Y',2)^{\sim u_2}, \cdots, A(Y',t)^{\sim u_t}\},$$
where $u_i = \#(A(Y',i))$, counting multiplicity. Again, $A(Y',s)(y)$ is used for two purposes. It is regarded as a subset of $\mathrm{SP}\phi_y$ or as a single element of $A(Y')$ with multiplicity u_s.
In fact, if $U_1, U_2, \cdots U_t$ are open sets, with mutually disjoint closure, such that $\mathrm{SP}\phi_y \subset U_1 \cup U_2 \cup \cdots \cup U_t, \forall y \in Y'$, then $\#(\mathrm{SP}\phi_y \cap U_i),\ y \in Y'$ are constants, denoted by u_i, (note that Y' is path connected). Furthermore, the maps
$$A(Y',i) : Y' \to P^{u_i}X, \quad i = 1,2,\cdots t,$$
defined by $A(Y',i)(y) = \mathrm{SP}\phi_y \cap U_i$, are continuous, and they determine a decomposition of $\{\mathrm{SP}\phi_y\}_{y\in Y'}$ as
$$\mathrm{SP}\phi_y = \{A(Y',1)(y), A(Y',2)(y), \cdots, A(Y',t)(y)\}$$
satisfying the condition (S). (See Lemmas 1.2.9 and 1.2.10.)
For any $y \in Y'$, a grouping $E_1, E_2, \cdots, E_L$ of $\mathrm{SP}\phi_y$ induces a UNIQUE grouping $E_1^{A(Y')}, E_2^{A(Y')}, \cdots E_L^{A(Y')}$ of $A(Y')$, defined by
$$E_i^{A(Y')} = \{A(Y',1)^{\sim v_1}, A(Y',2)^{\sim v_2}, \cdots, A(Y',t)^{\sim v_t}\},\ i = 1,2,\cdots,L,$$
where $v_j = \#(E_i \cap A(Y',j)(y))$, counting multiplicity. (Here the intersection of sets is defined as for the sets with multiplicity, as in 3.21.)
On the other hand, let $E_1, E_2, \cdots, E_L$ be a grouping of $A(Y')$. Define a grouping of the set $\mathrm{SP}\phi_y$, for any $y \in Y'$, in the following way. For any $j = 1,2,\cdots,L$, if the part E_j (for the grouping of $A(Y')$) contains exactly w elements of $\{A(Y',s)^{\sim u_s}\}$ ($w \le u_s$), then the part E_j (for the grouping of the set of $\mathrm{SP}\phi_y$) contains exactly w elements (counting multiplicity) which are contained in $A(Y',s)(y)$. Since these w elements are to be chosen, the induced grouping is not unique. But we will always fix one of them for use.
Let $E_1, E_2, \cdots, E_L$ be a grouping of
$$A(Y') = \{A(Y',1)^{\sim u_1}, A(Y',2)^{\sim u_2}, \cdots, A(Y',t)^{\sim u_t}\}.$$
Define $\mathring{E}_1, \mathring{E}_2, \cdots, \mathring{E}_L$ as in 3.22.
Although the subsets of $\mathrm{SP}\phi_y$ corresponding to E_i are not unique, the subsets of $\mathrm{SP}\phi_y$ corresponding to $\mathring{E}_i$ are unique. We denote them by $\mathring{E}_i|_y$. Also, $\#(\mathring{E}_i) = \#(\mathring{E}_i|_y)$ counting multiplicity. Note that we use $\mathring{E}_i|_y$ instead of $\mathring{E}_i(y)$ for the following reason (also see the next paragraph). We reserve the notation $\{E_i(y)\}_{i=1}^L$ for the grouping of $\mathrm{SP}\phi_y$ which is associated to a vertex y in a certain simplicial complex (Y,τ). (τ is a subdivision of σ.)
Suppose that $y \in Y'$. Let $E_1(y), E_2(y), \cdots E_L(y)$ be a grouping of $\mathrm{SP}\phi_y$. Then it induces a grouping $E_1^{A(Y')}(y), E_2^{A(Y')}(y), \cdots E_L^{A(Y')}(y)$ of $A(Y')$ as above. The sets $\mathring{E}_i^{A(Y')}(y)$ are well defined as subsets of $A(Y')$. (Warning: $\mathring{E}_i^{A(Y')}(y)$

are not subsets of $\mathrm{SP}\phi_y$.) Also, from the last paragraph, the sets $\mathring{E}_i^{A(Y)}(y)|_z$ are well defined as subsets of $\mathrm{SP}\phi_z$ for any $z \in Y'$ (may be different from y). Furthermore, $\#(\mathring{E}_i^{A(Y')}(y)) = \#(\mathring{E}_i^{A(Y')}(y)|_z)$.
In the next few paragraphs, each grouping of $\mathrm{SP}\phi_y$ can be referred as a grouping of $A(Y')$ for different space Y' and different decomposition $A(Y')$ provided that $y \in Y'$, or vise versa.

4.11. Note that the collection of sets $\{\Lambda(\Delta, i)\}_{i=1}^{t_\Delta}$ (in 4.9) could be regarded as a decomposition of $\mathrm{SP}\phi_y$, $y \in \Delta$ (see 4.6 also). And this decomposition satisfies condition (S) for Δ in place of Y'; therefore 4.10 can be applied to $\Lambda(\Delta)$ as $A(Y')$.
As mentioned in 4.8, we will introduce the groupings of $\mathrm{SP}\phi_y$ for all vertices of a certain subdivision of (Y, σ). As in section 3, for a simplex Δ of (Y, σ), once we have the subdivision $(\partial\Delta, \tau)$ of $(\partial\Delta, \sigma)$ and groupings for all vertices in $\mathrm{Vertex}(\partial\Delta, \tau)$, then we can define the subdivision (Δ, τ) of (Δ, σ) and introduce the groupings for all newly introduced vertices. One may notice that, in section 3, for different vertices, the index sets involved are the SAME. But in the setting here, the index sets $\mathrm{SP}\phi_y$ are DIFFERENT for different vertices y. So some special care should be taken.
Suppose that (Δ, σ) is a simplicial complex consisting of a single simplex Δ and all its faces. Suppose that there is a subdivision $(\partial\Delta, \tau)$ of $(\partial\Delta, \sigma)$ and the groupings $E_1(y), E_2(y), \cdots, E_L(y)$ of $\mathrm{SP}\phi_y$ for all vertices $y \in \mathrm{Vertex}(\partial\Delta, \tau)$ (see notation in 3.1–3.3). ATTENTION: When we introduce a grouping $E_1(z), E_2(z), \cdots, E_L(z)$ of $\mathrm{SP}\phi_z$ for any newly introduced vertex $z \in \mathrm{interior}(\Delta) = \Delta\backslash\partial\Delta$, THE FOLLOWING PROCEDURE WILL ALWAYS BE USED.
First, as in 4.10, we can regard the groupings $E_1(y), E_2(y), \cdots, E_L(y)$ of $\mathrm{SP}\phi_y$ as groupings $E_1^{\Lambda(\Delta)}(y), E_2^{\Lambda(\Delta)}(y), \cdots, E_L^{\Lambda(\Delta)}(y)$ of $\Lambda(\Delta)$ for all vertices $y \in \mathrm{Vertex}(\partial\Delta, \tau)$. (Then the set $E_i^{\Lambda(\Delta)}(y) \cap E_i^{\Lambda(\Delta)}(y') \cap \cdots$, as a subset of $\Lambda(\Delta)$, makes sense, for vertices $y, y', \cdots \in \mathrm{Vertex}(\partial\Delta, \tau)$. Also $\{\mathring{E}_i^{\Lambda(\Delta)}(y)\}_{i=1}^{L}$ are subsets of $\Lambda(\Delta)$.) Then we use these groupings of the SAME index set, $\Lambda(\Delta)$, applying the results from section 3 (see 3.32), to introduce subdivision (Δ, τ) of (Δ, σ) and groupings of $\Lambda(\Delta)$ for all newly introduced vertices $z \in \Delta\backslash\partial\Delta$. Finally, theses groupings of $\Lambda(\Delta)$ will induce the groupings $E_1(z), E_2(z), \cdots, E_L(z)$ of $\mathrm{SP}\phi_z$ as in 4.10 (not unique, but we fix one of them for our use). Furthermore, as in 4.10, $\mathring{E}_1^{\Lambda(\Delta)}(z), \mathring{E}_2^{\Lambda(\Delta)}(z), \cdots, \mathring{E}_L^{\Lambda(\Delta)}(z)$ are well defined subsets of $\Lambda(\Delta)$ and $\mathring{E}_1^{\Lambda(\Delta)}(z)|_{z'}, \mathring{E}_2^{\Lambda(\Delta)}(z)|_{z'}, \cdots, \mathring{E}_L^{\Lambda(\Delta)}(z)|_{z'}$ are well defined subsets of $\mathrm{SP}\phi_{z'}$ for any $z' \in \Delta$ (not necessarily a vertex).

4.12. Let Δ' be a face of Δ. Then for $y \in \Delta' \subset \Delta$, both $\Lambda(\Delta')$ and $\Lambda(\Delta)$ can be viewed as decompositions of $\mathrm{SP}\phi_y$. Recall, in 4.9, the decomposition corresponding to $\Lambda(\Delta)$ is the coarsest decomposition among those corresponding to $O(y_j)$ such that $O(y_j) \supset \Delta$ and that $O(y_j) \in \mathcal{O}$. Since $\Delta' \subset \Delta$, any open set in $\mathcal{O}$ which covers Δ will also cover Δ'. Therefore, the decomposition of $\mathrm{SP}\phi_y$

corresponding to $\Lambda(\Delta')$ is coarser than that corresponding to $\Lambda(\Delta)$. That is, each set $\Lambda(\Delta', s)(y)$ is a finite union of certain sets,

$$\Lambda(\Delta, t_1)(y) \bigcup \Lambda(\Delta, t_2)(y) \bigcup \cdots ,$$

as subsets of $\mathrm{SP}\phi_y$. (Notice that if $w'_s = w_s(\Delta')$ is the multiplicity appeared in $\Lambda(\Delta')$ for $\Lambda(\Delta', s)$ and $w_t = w_t(\Delta)$ is the multiplicity appeared in $\Lambda(\Delta)$ for $\Lambda(\Delta, t)$, then $w'_s = w_{t_1} + w_{t_2} + \cdots$, a finite sum.)
It follows that if $y \in \Delta' \subset \Delta$ and $E_1(y), E_2(y), \cdots, E_L(y)$ is a grouping of $\mathrm{SP}\phi_y$, then

$$\overset{\circ}{E}{}_j^{\Lambda(\Delta')}(y)|_{y'} \subset \overset{\circ}{E}{}_j^{\Lambda(\Delta)}(y)|_{y'},$$

regarded as subsets of $\mathrm{SP}\phi_{y'}$ for any $y' \in \Delta'$. Now we are ready to construct the subdivision (Y, τ) of (Y, σ), and the grouping for each vertex of (Y, σ) and each vertex of the complex (Y, τ) (after subdivision).
Since the notations $y_1, y_2, \cdots, y_\bullet$ have been already used for the open cover $\mathcal{O} = \{O(y_1), O(y_1), \cdots, O(y_\bullet)\}$, we use $z_1, z_2, \cdots$ to denote the points in Y, especially the vertices of certain simplicial structure.

4.13. Let $\Omega = \dim X + \dim Y$, $M = 2^L - 1$, where L is the number of the continuous maps $\{a_i\}$ appeared in the statement of 4.1. Note that all the multiplicities w_t appearing in any of $\Lambda(\Delta)$ do not exceed Ω, by 4.3 and the construction of $\Lambda(\Delta)$ (see 4.6 and 4.9).
For each vertex $z \in \mathrm{Vertex}(Y, \sigma)$, by the condition of Theorem 4.1, $\mathrm{SP}\phi_z$ and

$$\Theta(z) = \{a_1(z)^{\sim T_1}, a_2(z)^{\sim T_2}, \cdots, a_L(z)^{\sim T_L}\}$$

can be paired within η. Therefore, we can define a grouping
$E_{pre,1}(z), E_{pre,2}(z), \cdots, E_{pre,L}(z)$ of $\mathrm{SP}\phi_z$, with $T_1, T_2, \cdots, T_L$ elements, respectively, counting multiplicity, such that

$$\mathrm{dist}(\lambda, a_i(z)) < \eta \tag{1}$$

if $\lambda \in E_{pre,i}(z)$, where η is as in 4.4. (We denote them by $E_{pre,i}$ because this grouping will be modified later.)
We can regard such a grouping of $\mathrm{SP}\phi_z$ as a grouping of $\Lambda(\Delta)$, where $\Delta \ni z$ is a simplex.
First we regard it as the grouping of $\Lambda(\{z\})$, where $\{z\}$ is the 0-dimensional simplex of (Y, σ) corresponding to vertex z. By Lemma 3.28, we can modify the grouping to satisfy the Assumption 3.27. Then this modified grouping of $\Lambda(\{z\})$ could induce a grouping on $\mathrm{SP}\phi_z$, for which the condition (1) above may not hold. But if we carefully choose the sets H_i in Lemma 3.28, we could still guarantee that any elements $\lambda \in E_i$ are close to $a_i(z)$ (see (2) below). In this subsection, we will also introduce the sets $H_i(\Delta)$ to serve as the sets H_i of Lemma 3.28 and Theorem 3.32, when we construct groupings on Δ from the groupings on $\partial\Delta$, by applying Theorem 3.32.

For each vertex $z_0 \in \mathrm{Vertex}(Y,\sigma)$, the notation $\{z_0\}$ is used to denote the corresponding zero dimensional simplex of (Y,σ). The above grouping induces a grouping $\{E_{pre,i}^{\Lambda(\{z_0\})}\}$ of

$$\Lambda(\{z_0\}) = \{\Lambda(\{z_0\},1)^{\sim w_1}, \Lambda(\{z_0\},2)^{\sim w_2}, \cdots, \Lambda(\{z_0\},t_{\{z_0\}})^{\sim w_{t_{\{z_0\}}}}\}.$$

Define subsets $H_1(\{z_0\}), H_2(\{z_0\}), \cdots, H_L(\{z_0\})$ of $\Lambda(\{z_0\})$ as follows. For any $i=1,2,\cdots,L,\ H_i(\{z_0\})$ is the collection of all
$\Lambda(\{z_0\},t)^{\sim w_t}$ $(\subset \Lambda(\{z_0\}))$ satisfying

$$\Lambda(\{z_0\},t)(z_0) \subset \left\{x:\ \mathrm{dist}(x,a_i(z_0)) < \eta + \frac{1}{(\dim Y+1)}\cdot\eta\right\}.$$

Note that each set $\Lambda(\{z_0\},t)(z_0)$ (as a subset of certain $U(y_j,t)$, from 4.9) has diameter at most $\frac{\eta}{2(\dim Y+1)} < \frac{\eta}{(\dim Y+1)}$ (see (b) in 4.6). Combining this fact with (1) above, we know that if $\lambda_1 \in E_{pre,i}(z_0)$ and $\lambda_1 \in \Lambda(\{z_0\},t)(z_0)$, then

$$\mathrm{dist}(\lambda, a_i(z_0)) < \eta + \frac{\eta}{(\dim Y+1)} \tag{2}$$

for any $\lambda \in \Lambda(\{z_0\},t)(z_0)$. That means $E_{pre,i}^{\Lambda(\{z_0\})} \subset H_i(\{z_0\})$.
By Lemma 3.28, the above grouping can be modified to another grouping $E_i^{\Lambda(\{z_0\})}$ of $\Lambda(\{z_0\})$ satisfying

$$\#(\overset{\circ}{E}_i{}^{\Lambda(\{z_0\})}) \geq T_i - M\Omega.$$

(This is Assumption 3.27.) And $E_i^{\Lambda(\{z_0\})} \subset H_i(\{z_0\})$ still holds, regarded as a grouping of $\Lambda(\{z_0\})$.
The above grouping of $\Lambda(\{z_0\})$ could induce a grouping $E_1(z_0), E_2(z_0), \cdots, E_L(z_0)$ of $\mathrm{SP}\phi_{z_0}$ (see 4.10 and 4.11). This grouping will be used as the grouping for vertex z_0. Even though (1) may not hold for λ in the new $E_i(z_0)$, (2) holds for any λ in the new $E_i(z_0)$, from the definition of $H_i(\{z_0\})$, and $E_i^{\Lambda(\{z_0\})} \subset H_i(\{z_0\})$.
For each simplex Δ of (Y,σ), let us also define the subsets $H_1(\Delta)$, $H_2(\Delta), \cdots$, $H_L(\Delta)$ of $\Lambda(\Delta)$ as follows. For each $j=1,2,\cdots,L$, $H_j(\Delta)$ is the collection of all such $\Lambda(\Delta,t)^{\sim w_t}$ $(\subset \Lambda(\Delta))$ that $\Lambda(\Delta,t)(z)$, as a subset of $\mathrm{SP}\phi_z$, satisfies

$$\Lambda(\Delta,t)(z) \subset \{x:\ \mathrm{dist}(x,a_i(z)) < \eta + \frac{\dim(\Delta)+1}{(\dim Y+1)}\cdot\eta\}$$

for any $z \in \Delta$. These sets will serve as the sets $H_1, H_2, \cdots, H_L$ when we apply Theorem 3.32.
The following fact follows directly from the definition of $\Lambda(\Delta,t)$ and $H_i(\Delta)$, which will be used in 4.14:
Suppose that $z \in \Delta$. A grouping $E_1, E_2, \cdots E_L$ of $\mathrm{SP}\phi_z$, regarded as a grouping of $\Lambda(\Delta)$, satisfies $E_i \subset H_i(\Delta)$ if and only if for any $\lambda \in E_i$ (as a subset of

SPϕ_z), for the index t satisfying $\lambda \in \Lambda(\Delta, t)(z)$ (such t exists; see 4.9), we have $\{\Lambda(\Delta, t)^{\sim w_t(\Delta)}\} \subset H_i(\Delta)$.

4.14. Beginning with the simplicial structure (Y, σ) and the above groupings for all 0-dimensional simplex (i.e., vertex) of (Y, σ), we will construct a subdivision (Y, τ) of (Y, σ) and the groupings for newly introduced vertices. We will refine (Y, σ), simplex by simplex, from the lowest dimension to the highest dimension by use of Theorem 3.32.
To avoid confusion, use $\Gamma, \Gamma_1, \Gamma_2, \Gamma', etc.$ to denote the simplices of (Y, τ), after subdivision, and reserve the notations $\Delta, \Delta', \Delta_1, etc.$ for the simplices of (Y, σ)—with original simplicial complex structure σ introduced in 4.7.
As the induction assumption, we suppose that there are a subdivision $(\partial\Delta, \tau)$ of $(\partial\Delta, \sigma)$ and the groupings of SPϕ_z for all vertices $z \in \mathrm{Vertex}(\partial\Delta, \tau)$ with the following properties.
(1) If Δ' is a proper face of (Δ, σ) (by a proper face of Δ, we mean a face Δ' with $\Delta' \subset \partial\Delta$) and $z \in \Delta'$, then the grouping of $\Lambda(\Delta')$, induced by the grouping of SPϕ_z satisfies

$$E_i \subset H_i(\Delta').$$

In other words, $E_i^{\Lambda(\Delta')}(z) \subset H_i(\Delta')$.
(2) Let Γ be a simplex of $(\partial\Delta, \tau)$ with vertices $z_0, z_1, \cdots, z_j$. If $\Gamma \subset \Delta'$, where Δ' is a proper face of Δ, then

$$\begin{aligned} \# \left(\overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_0) \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_1) \cap \cdots \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_j) \right) \\ \geq T_i - [M\Omega + M\Omega \dim \Delta'(\dim \Delta' + 1)] \\ (\geq T_i - [M\Omega + M\Omega \dim \partial\Delta(\dim \partial\Delta + 1)]) \,. \end{aligned}$$

(3) For each vertex z of $(\partial\Delta, \tau)$, Assumption 3.27 holds. I.e.,

$$\#(\overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z)) \geq T_i - M\Omega$$

for any proper face Δ' of Δ with $z \in \Delta'$.
(In the above conditions (1), (2) and (3), $\{E_i^{\Lambda(\Delta')}(z)\}_{i=1}^{L}$ are regarded as groupings of the set $\Lambda(\Delta')$ (with multiplicity); see 4.11.)
Now we define the subdivision (Δ, τ) of (Δ, σ) and the groupings for all newly introduced vertices. The restriction of the simplicial structure (Δ, τ) on $\partial\Delta$ will be the same as $(\partial\Delta, \tau)$, that is, we will only introduce new vertices inside interior$(\Delta) = \Delta \backslash \partial\Delta$. We need to define the groupings as groupings of $\Lambda(\Delta)$. Then they will induce groupings of SPϕ_z.
Claim: For any vertex z of $(\partial\Delta, \tau)$, if the grouping of SPϕ_z is regarded as the grouping of $\Lambda(\Delta)$, then
(1') $E_i(z) \subset H_i(\Delta)$ for $i = 1, 2, \cdots, L$. In other words, $E_i^{\Lambda(\Delta)}(z) \subset H_i(\Delta)$.
Proof of the Claim: Let y be the point y_i in the definition of $\Lambda(\Delta)$ in 4.9. Then $\Delta \subset O(y) \in \mathcal{O}$. (We avoid the notation y_i, since i is used for E_i above. So we use y instead.)

Let $\lambda \in E_i(z)$. Since $z \in \partial\Delta$, there is a proper face Δ' of Δ such that $z \in \Delta'$. By (1) above, $E_i(z) \subset H_i(\Delta')$ (regarded as a grouping of $\Lambda(\Delta')$). From the end of 4.13, there is an index s such that $\lambda \in \Lambda(\Delta', s)(z)$ and that $\{\Lambda(\Delta', s)^{\sim w_s(\Delta')}\} \subset H_i(\Delta')$.
Recall, from 4.12, $\Lambda(\Delta', s)(z')$ is a finite union $\Lambda(\Delta, t_1)(z') \cup \Lambda(\Delta, t_2)(z') \cup \cdots$, for any $z' \in \Delta' \subset \Delta$. Hence, there is an index t such that

$$\lambda \in \Lambda(\Delta, t)(z) \subset \Lambda(\Delta', s)(z),$$

where both $\Lambda(\Delta, t)(z)$ and $\Lambda(\Delta', s)(z)$ are regarded as subsets of $\mathrm{SP}\phi_z$. To prove the claim, by the end of 4.13, we only need to prove $\{\Lambda(\Delta, t)^{\sim w_t(\Delta)}\} \subset H_i(\Delta)$. From the definition of $H_i(\Delta)$, this is equivalent to

$$(A) \qquad \Lambda(\Delta, t)(z_1) \subset \{x : \ \mathrm{dist}(x, a_i(z_1)) < \eta + \frac{\dim(\Delta) + 1}{(\dim Y + 1)} \cdot \eta\}$$

for any $z_1 \in \Delta$. From $\{\Lambda(\Delta', s)^{\sim w_s(\Delta')}\} \subset H_i(\Delta')$ and the definition of $H_i(\Delta')$, we have

$$\Lambda(\Delta', s)(z') \subset \{x : \ \mathrm{dist}(x, a_i(z')) < \eta + \frac{\dim(\Delta') + 1}{(\dim Y + 1)} \cdot \eta\}$$

for any $z' \in \Delta'$. In the above, if we choose $z' = z$—the vertex in the claim—(and note that $\Lambda(\Delta, t)(z) \subset \Lambda(\Delta', s)(z)$), then

$$(a) \qquad \Lambda(\Delta, t)(z) \subset \{x : \ \mathrm{dist}(x, a_i(z)) < \eta + \frac{\dim(\Delta') + 1}{(\dim Y + 1)} \cdot \eta\}.$$

On the other hand, from (d) in 4.6, we have

$$(b) \qquad \mathrm{diameter}(a_i(\Delta)) \leq \mathrm{diameter}(a_i(O(y))) < \frac{\eta}{2(\dim Y + 1)}.$$

And from (b) in 4.6, we have

$$(c) \qquad \mathrm{diameter}(U(y, t)) < \frac{\eta}{2(\dim Y + 1)}.$$

From 4.6 and 4.9, $\Lambda(\Delta, t)(z_1) \subset U(y, t)$ for any $z_1 \in \Delta \subset O(y)$. Combining this with (c) above, for any $\mu \in \Lambda(\Delta, t)(z_1)$ ($z_1 \in \Delta$), we have

$$dist(\mu, \Lambda(\Delta, t)(z)) < \frac{\eta}{2(\dim Y + 1)}.$$

Then combining it with (a) above, we have

$$dist(\mu, a_i(z)) < \eta + \frac{(\dim(\Delta') + 1)\eta}{(\dim Y + 1)} + \frac{\eta}{2(\dim Y + 1)}.$$

Finally, combining it with (b),

$$dist(\mu, a_i(z_1)) < \eta + \frac{(\dim(\Delta')+1)\eta}{(\dim Y+1)} + \frac{\eta}{2(\dim Y+1)} + \frac{\eta}{2(\dim Y+1)}$$
$$\leq \eta + \frac{\dim(\Delta)+1}{(\dim Y+1)},$$

since $dim(\Delta') \leq dim(\Delta) - 1$. Note that $z_1 \in \Delta$ and $\mu \in \Lambda(\Delta,t)(z_1)$ are arbitrary, this proves (A) and the Claim.

Suppose that Γ is a simplex of $(\partial\Delta,\tau)$ with vertices $z_0, z_1, \cdots z_j$. Suppose that $\Gamma \subset \Delta'$, where Δ' is a face of Δ. As mentioned in 4.12,

$$\overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z)|_{z'} \subset \overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z)|_{z'}$$

as a subset of $\mathrm{SP}\phi_{z'}$ for all $z' \in \Delta'$ and all $z = z_0, z_1, \cdots, z_j$. Therefore, from (2) and (3) above, we have the following (2') and (3').

(2') $\#\left(\overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z_0) \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z_1) \cap \cdots \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z_j)\right)$

$$= \#\left(\overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z_0)|_{z'} \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z_1)|_{z'} \cap \cdots \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z_j)|_{z'}\right)$$
$$\geq \#\left(\overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_0)|_{z'} \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_1)|_{z'} \cap \cdots \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_j)|_{z'}\right)$$
$$= \#\left(\overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_0) \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_1) \cap \cdots \cap \overset{\circ}{E}{}_i^{\Lambda(\Delta')}(z_j)\right)$$
$$\geq T_i - [M\Omega + M\Omega \dim \partial\Delta(\dim \partial\Delta + 1)],$$

for every simplex $\Gamma \subset (\partial\Delta,\tau)$ with vertices $z_0, z_1, \cdots, z_j$.

(3') The Assumption 3.27 holds for the grouping $\{E_i(z)\}_{i=1}^L$ regarded as a grouping of $\Lambda(\Delta)$, i.e.,

$$\#(\overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z)) \geq T_i - M\Omega,$$

where z is a vertex of $(\partial\Delta,\tau)$.

Apply Theorem 3.32 to obtain a subdivision (Δ,τ) of (Δ,σ), and, for each newly introduced vertex $z \in \Delta$, a grouping $E_1(z), E_2(z), \cdots E_L(z)$ of $\mathrm{SP}\phi_z$ such that (1), (2) and (3) hold with the version obtained by replacing Δ' by Δ, and $\dim(\partial\Delta)$ by $\dim\Delta$. (As mentioned in 4.11, for each vertex z, we should first get the groupings of $\Lambda(\Delta)$, then this grouping induces a grouping of $\mathrm{SP}\phi_z$.) Using Mathematical Induction, combined with 4.13, we obtain our subdivision (Y,τ) of (Y,σ) and the groupings.

We summarize what we obtained in 4.13 and 4.14 as in the following proposition.

Proposition: *There is a subdivision (Y,τ) of (Y,σ), and for all vertices $z \in$ Vertex(Y,τ), there are groupings $E_1(z), E_2(z), \cdots, E_L(z)$ of $SP\phi_z$ of type*

$(T_1, T_2, \cdots, T_L)$ (i.e., $\#(E_i(z)) = T_i \ \ \forall i$) such that the following are true.
(1) If Δ is a simplex of (Y, σ) (before subdivision) and $z \in \Delta$, then the grouping $(E_1^{\Lambda(\Delta)}(z), E_2^{\Lambda(\Delta)}(z), \cdots E_L^{\Lambda(\Delta)}(z))$ of $\Lambda(\Delta)$, induced by the grouping $(E_1(z), E_2(z), \cdots, E_L(z))$ of $SP\phi_z$, satisfies

$$E_i^{\Lambda(\Delta)}(z) \subset H_i(\Delta).$$

(2) Let Γ be a simplex of (Y, τ) with vertices $z_0, z_1, \cdots, z_j$. If $\Gamma \subset \Delta$, where Δ is a simplex of (Y, σ) (before subdivision), then

$$\begin{aligned}\# \left(\mathring{E}_i^{\Lambda(\Delta)}(z_0) \cap \mathring{E}_i^{\Lambda(\Delta)}(z_1) \cap \cdots \cap \mathring{E}_i^{\Lambda(\Delta)}(z_j)\right) \qquad & \\ \geq T_i - [M\Omega + M\Omega \dim \Delta(\dim \Delta + 1)] & \\ (\geq T_i - [M\Omega + M\Omega \dim Y(\dim Y + 1)]) \, . &\end{aligned}$$

(We do not need the condition (3) any more.)

4.15. For the simplicial complex (Y, τ), there is a finite open cover

$$\{W(\Gamma) : \ \Gamma \text{ is a simplex of } \ (Y, \tau)\}$$

of Y, with the following properties.
(a) $W(\Gamma) \supset \text{interior}(\Gamma) = \Gamma \backslash \partial\Gamma$.
(b) If $W(\Gamma_1) \cap W(\Gamma_2) \neq \emptyset$, then either Γ_1 is a face of Γ_2 or Γ_2 is a face of Γ_1.
(Such open cover has been constructed in 1.4.2 (b).)
For any simplex Γ, we will construct an open set $O(\Gamma) \supset \Gamma$ and introduce a decomposition $\Xi(\Gamma)$ of $\{\mathrm{SP}\phi_y\}_{y \in O(\Gamma)}$, which is the finest possible decomposition satisfying the condition (S) for Γ in place Y' in 4.10.
Recall that $K = \text{rank}(P)$, and $y \mapsto \mathrm{SP}\phi_y$ defines a map $\mathrm{SP}\phi : Y \to P^K X$. We will prove the following easy fact.
Claim 1: $\mathrm{SP}\phi|_\Gamma := \overline{\cup_{z \in \Gamma} \mathrm{SP}\phi_z}$ $(\subset X)$ has at most K connected components.
(For $K = 1$, the claim says that the image of a connected space Γ under a continuous map $\mathrm{SP}\phi : \Gamma \to P^1 X = X$ is connected. This is a trivial fact.)
Proof of Claim 1: Suppose that by the contrary, $\mathrm{SP}\phi|_\Gamma$ has more than K connected components. Write $\mathrm{SP}\phi|_\Gamma = X_1 \cup X_2 \cup \cdots \cup X_{K+1}$, where $X_1, X_2, \cdots, X_{K+1}$ are mutually disjoint non empty closed subsets (which are not necessary connected).
There are open sets $U_1, U_2, \cdots, U_{K+1}$ with mutually disjoint closures such that $U_i \supset X_i$. Then for any $z \in \Gamma$, $\mathrm{SP}\phi_z \subset \cup_{i=1}^{K+1} U_i$. By Lemma 1.2.9, for each i, $\#(\mathrm{SP}\phi_z \cap U_i)$ is a nonzero constant. Hence $\#(\mathrm{SP}\phi_z) \geq K + 1$, contradicting with $\#(\mathrm{SP}\phi_z) = K = \text{rank}(P)$, counting multiplicity. This proves the claim.
We are back to our construction of open set $O(\Gamma)$ and decomposition $\Xi(\Gamma)$. Write $\mathrm{SP}\phi|_\Gamma = X_1 \cup X_2 \cup \cdots \cup X_t$, where $X_1, X_2, \cdots, X_t$ (with $t \leq K$) are mutually disjoint connected components of $\mathrm{SP}\phi|_\Gamma$. Choose open sets $U_1, U_2, \cdots, U_t$

with mutually disjoint closures such that $X_i \subset U_i$. By Lemma 1.2.9, there is an open set $O(\Gamma) \supset \Gamma$ such that $\mathrm{SP}\phi|_{O(\Gamma)} \subset \cup_{i=1}^t U_i$.
As in 4.10, define

$$\Xi(\Gamma, t)(z) = \mathrm{SP}\phi_z \cap U_i, \ \forall z \in O(\Gamma), \ i = 1, 2, \cdots, t.$$

This gives a decomposition

$$\mathrm{SP}\phi_z = \ \{\Xi(\Gamma, 1)(z), \Xi(\Gamma, 2)(z), \cdots, \Xi(\Gamma, t)(z)\}, \ \forall z \in O(\Gamma).$$

Let $c_i = \#(\Xi(\Gamma, i)$, counting multiplicity. And write

$$\Xi(\Gamma) := \ \{\Xi(\Gamma, 1)^{\sim c_1}, \Xi(\Gamma, 2)^{\sim c_2}, \cdots, \Xi(\Gamma, t)^{\sim c_t}\}.$$

Note that the above decomposition satisfies condition (S) in 4.10 as the decomposition of spectrum on $O(\Gamma)$ (not only on Γ). In 4.16 below, when we apply 4.10, we will use $U(\Gamma)$ (a subset of $O(\Gamma)$) in place of Y' of 4.10. Obviously, $\Xi(\Gamma)$ is the finest decomposition among all the decompositions of $(\mathrm{SP}\phi_z)_{z\in\Gamma}$ satisfying condition (S) on Γ, since each X_i is connected. In particular, if $z \in \Gamma \subset \Delta$, where Δ is a simplex of (Y, σ) (before subdivision), then the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Xi(\Gamma)$ is finer than the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Lambda(\Delta)$.
We will use the following fact later.
Claim 2: If $\Gamma' \subset \Gamma$ is a face, then for any $z \in O(\Gamma') \cap O(\Gamma)$, the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Xi(\Gamma')$ is finer than the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Xi(\Gamma)$.
Proof of Claim 2: The Claim follows from the definition of $\Xi(\Gamma)$ and the fact that any connected component of $\mathrm{SP}\phi|_{\Gamma'}$ is completely contained in a connected component of $\mathrm{SP}\phi|_{\Gamma}$.

4.16. For each simplex Γ, define $U(\Gamma) = W(\Gamma) \cap O(\Gamma)$.
$\{U(\Gamma); \ \Gamma \ \text{ is a simplex of } \ (Y, \tau)\}$ is an open covering of Y since $U(\Gamma) \supset$ interior(Γ).
For each $U = U(\Gamma)$, we will define mutually orthogonal projection valued functions

$$P_1^U, P_2^U, \cdots, P_L^U : \ U(\Gamma) \ni y \mapsto \{ \text{ sub-projections of } P(y)\}.$$

Then apply Proposition 3.2 of [DNNP] to construct the globally defined projections $p_1, p_2, \cdots, p_L$ for our Theorem 4.1.
(Attention: For each vertex z of (Y, τ), we have a grouping
$E_1(z), E_2(z), \cdots, E_L(z)$ of $\mathrm{SP}\phi_z$. It will induce a grouping of $\Xi(\Gamma)$, as in 4.10, if $\Gamma \ni z$. In the following construction of $P_i^U(y)$, this grouping will be used. That is, we will use the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Xi(\Gamma)$. The decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Lambda(\Delta)$ will NOT be used in the definition of $P_i^U(y)$ at all— it is only used in the estimation of rank(P_i^U).
In the definition of the grouping $E_1(z), E_2(z), \cdots, E_L(z)$ of $\mathrm{SP}\phi_z$, it involves the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Lambda(\Delta)$. But once it has been defined, it

makes sense by itself without the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Lambda(\Delta)$ as a reference (though $\overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z)|_z$ only makes sense with the decomposition as the reference).)

Back to our construction. For each $y \in U$ $(= U(\Gamma))$ and each $i = 1, 2, \cdots, L$, define $P_i^U(y)$ to be the spectral projection of ϕ_y corresponding to

$$\left(\overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_0) \cap \overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_1) \cap \cdots \cap \overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_j)\right)|_y ,$$

where $z_0, z_1, \cdots, z_j$ are all vertices of Γ, and the notations $\overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z)$ and $\overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z)|_y$ are as in 4.10. (That is, $\overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z)$ is a subset of $\Xi(\Gamma)$ and $\overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z)|_y$ is a subset of $\mathrm{SP}\phi_y$.)

By Lemma 1.2.9, the above functions $P_i^U(y)$ depend on y continuously. In fact, for each i and any $y \in U(\Gamma) \subset Y$, $P_i^U(y)$ is the spectral projection of ϕ_y corresponding to an open subset (of X)— in the notation of 4.15 (see the paragraph after the proof of Claim 1 in 4.15), the open subset is the union of all open subsets $U_j \subset X$ such that

$$\Xi(\Gamma, j) \in \overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_0) \cap \overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_1) \cap \cdots \cap \overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_j)(\subset \Xi(\Gamma)).$$

(Note that when we apply Lemma 1.2.9, we use the fact that $\{U_i\}$ have mutually disjoint closures and $\mathrm{SP}\phi_y \subset \bigcup U_i$ from 4.15.) Recall, the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Xi(\Gamma)$ is finer than any decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Lambda(\Delta)$, if $\Gamma \subset \Delta$. Therefore, $\overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_0)|_z \supset \overset{\circ}{E}{}_i^{\Lambda(\Delta)}(z_0)|_z$, regarded as a subset of $\mathrm{SP}\phi_z$, for any vertex $z_0 \in \Gamma$ and any point $z \in \Gamma$. By Condition (2) of the grouping (see 4.14),

$$\mathrm{rank}(P_j^U) \geq T_j - [M\Omega + M\Omega \dim Y(\dim Y + 1)]$$

for each U.

The projections $P_i^U, i = 1, 2, \cdots, L$ are mutually orthogonal, since they are spectral projections corresponding to mutually disjoint subsets of X.

Let Γ' be a face of Γ and $z \in U(\Gamma) \cap U(\Gamma')$. By Claim 2 in 4.15, opposite to the case of decompositions corresponding to $\Lambda(\Delta)$ and $\Lambda(\Delta')$, the decomposition of $\mathrm{SP}\phi_z$ corresponding to $\Xi(\Gamma')$ is finer than that corresponding to $\Xi(\Gamma)$. Therefore,

$$\overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_0)|_y \subset \overset{\circ}{E}{}_i^{\Xi(\Gamma')}(z_0)|_y$$

for all $z_0 \in \mathrm{Vertex}(\Gamma', \tau) \subset \mathrm{Vertex}(\Gamma, \tau)$. Combining it with the fact that $\mathrm{Vertex}(\Gamma', \tau) \subset \mathrm{Vertex}(\Gamma, \tau)$, we get

$$\left(\bigcap_{z_j \in \mathrm{Vertex}(\Gamma,\tau)} \overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z_j)\right)|_y \subset \left(\bigcap_{z_j \in \mathrm{Vertex}(\Gamma',\tau)} \overset{\circ}{E}{}_i^{\Xi(\Gamma')}(z_j)\right)|_y.$$

Consequently,

$$P_i^{U(\Gamma)}(y) \leq P_i^{U(\Gamma')}(y) \qquad\qquad \text{if } y \in U(\Gamma) \cap U(\Gamma').$$

Finally, from the condition (1) of the groupings (see the proposition in the end of 4.14) and the definition of $H_i(\Delta)$, we have

$$\left(\overset{\circ}{E}{}_i^{\Xi(\Gamma)}(z)\right)|_y \subset \left\{\lambda;\ \mathrm{dist}(\lambda, a_i(y)) < \eta + \ \frac{(\dim Y+1)}{(\dim Y+1)} \ \cdot \eta = 2\eta\right\},$$

where Δ is any simplex of (Y, σ) satisfying $\Gamma \subset \Delta$. Therefore, $P_i^U(y)$ is the spectral projection of ϕ_y corresponding to a subset of

$$\{\lambda;\ \mathrm{dist}(\lambda, a_i(y)) \leq 2\eta\} \ \subset \ X.$$

We have proved the following lemma.

Lemma 4.17. *There is a collection $\mathcal{U}$ of finitely many open sets which covers Y. For each open set $U \in \mathcal{U}$, there are mutually orthogonal projection valued continuous functions*

$$P_1^U, P_2^U, \cdots, P_L^U : \ U \ni y \mapsto \{ \textit{ sub-projections of } P(y)\}$$

with the following properties.
(1) If $U_1, U_2 \in \mathcal{U}$, and $U_1 \cap U_2 \neq \emptyset$, then either

$$P_i^{U_1}(z) \leq P_i^{U_2}(z)$$

is true for all $i = 1, 2, \cdots, L$ and all $z \in U_1 \cap U_2$, or

$$P_i^{U_2}(z) \leq P_i^{U_1}(z)$$

is true for all $i = 1, 2, \cdots, L$ and all $z \in U_1 \cap U_2$.
(2) rank$(P_i^U(z)) \geq T_i - [M\Omega + M\Omega \dim Y(\dim Y + 1)]$.
(3) Each $P_i^U(z)$ is a spectral projection of ϕ_z corresponding to a subset of

$$\{\lambda;\ \textit{dist}(\lambda, a_i(z)) < 2\eta\}.$$

4.18. For $i = 1, 2, \cdots, L$, applying Proposition 3.2 of [DNNP] to $\{P_i^U\}_{U\in\mathcal{U}}$, there exist continuous projection valued functions

$$p_1^U, p_2^U, \cdots, p_L^U : \ Y \ni y \mapsto \{ \text{ sub-projections of } P(y)\}$$

such that

$$p_i(y) \leq \bigvee\{P_i^U(y);\ y \in U \in \mathcal{U}\}$$

and that

$$\mathrm{rank}(p_i) \quad \geq T_i - [M\Omega + M\Omega \dim Y(\dim Y + 1)] - \dim Y > T_i - T\ .$$

(Note that $T = 2^L(\dim X + \dim Y)^3$.)

By Condition (1) of 4.17, for each y,

$$\mathrm{span}\{P_i^U;\ y \in U \in \mathcal{U}\} = P_i^{U_0}$$

for a certain $U_0 \ni y$ which does not depend on i. Therefore, $\{p_i(y)\}_{i=1}^L$ are mutually orthogonal since $\{P_i^{U_0}\}_{i=1}^L$ are mutually orthogonal.

4.19. We will prove that the above projections $\{p_i\}_{i=1}^L$ and $p_0 = P - \sum_{i=1}^L p_i$ are as desired in Theorem 4.1. This is a routine calculation, as in the proof of Theorem 2.7 of [GL1] or the last part of the proof of Theorem 2.21 of [EG2]. (See 1.5.4 and 1.5.7 also.) Since we need an extra property of $p_0\phi p_0$ (described in 4.20 below), we write down the complete proof.
For each $y \in Y$, as mentioned in 4.18, there exists an open set $U_0 \in \mathcal{U}$ with $U_0 \ni y$ such that

$$\mathrm{span}\{P_i^U;\ y \in U \in \mathcal{U}\} = P_i^{U_0}, \qquad i = 1, 2, \cdots, L.$$

Let $P_i(y) = P_i^{U_0}(y)$. Then

$$p_i(y) \le P_i(y), \qquad i = 1, 2, \cdots, L,$$

and each $P_i(y)$ is the spectral projection corresponding to a certain subset of

$$\{\lambda;\ \lambda \in \mathrm{SP}\phi_y,\ \mathrm{dist}(\lambda, a_i(y)) < 2\eta\}\ .$$

Let mutually different elements $\mu_1, \mu_2, \cdots, \mu_s \in \mathrm{SP}\phi_y$ be the list of spectra which are not in the set of those spectra belonging to the projections $\{P_i(y)\}_{i=1}^L$. Let $q_1, q_2, \cdots, q_s$ be spectral projections corresponding to $\{\mu_1\}, \{\mu_2\}, \cdots, \{\mu_s\}$, respectively. (The rank of each q_i is the multiplicity of μ_i in $\mathrm{SP}\phi_y$.) Then

$$P(y) = \sum_{i=1}^L P_i(y) + \sum_{i=1}^s q_i.$$

Therefore,

$$p_0(y) = P(y) - \sum_{i=1}^L p_i(y) = \sum_{i=1}^L (P_i(y) - p_i(y)) + \sum_{i=1}^s q_i.$$

Since the spectra belonging to $P_i(y)$ are within distance 2η of $a_i(y)$, by the way η is chosen in 4.4, for each $f \in F$,

$$\|\phi(f)(y) - [\sum_{i=1}^L f(a_i(y))P_i + \sum_{i=1}^s f(\mu_i)q_i]\| < \frac{\varepsilon}{3}.$$

Therefore, for each $f \in F$, $\|p_0(y)\phi(f)(y) - \phi(f)(y)p_0(y)\| < \frac{2\varepsilon}{3}$, and

$$(*) \qquad \|p_0(y)\phi(f)(y)p_0(y) - [\sum_{i=1}^L f(a_i(y))(P_i(y) - p_i(y)) + \sum_{i=1}^s f(\mu_i)q_i]\| < \frac{\varepsilon}{3}.$$

Also, for all $f \in F$,

$$\|p_i(y)\phi(f)(y) - f(a_i(y))p_i(y)\| < \frac{\varepsilon}{3} \qquad \text{and}$$

$$\|\phi(f)(y)p_i(y) - f(a_i(y))p_i(y)\| < \frac{\varepsilon}{3}.$$

Let $P' = \sum_{i=1}^{L} p_i$. Then

$$\begin{aligned}
&\|P'(y)\phi(f)(y)p_0(y)\| = \|\textstyle\sum_{i=1}^{L} p_i(y)\phi(f)(y)p_0(y)\| \\
&\quad \leq \|\textstyle\sum_{i=1}^{L}[p_i(y)\phi(f)(y) - p_i(y)f(a_i(y))]p_0(y)\| + \|\textstyle\sum_{i=1}^{L} p_i(y)f(a_i(y))p_0(y)\| \\
&\quad \leq \tfrac{\varepsilon}{3} + 0 = \tfrac{\varepsilon}{3}\ .
\end{aligned}$$

Similarly, for all $f \in F$,

$$\|p_0(y)\phi(f)(y)P'(y)\| < \frac{\varepsilon}{3}.$$

Also,

$$\|P'(y)\phi(f)(y)P'(y) - \bigoplus_{i=1}^{L} f(a_i(y))p_i(y)\| < \frac{\varepsilon}{3}.$$

Combining all the above estimations, we have, for $f \in F$,

$$\|\phi(f)(y) - p_0(y)\phi(f)(y)p_0(y) \oplus \bigoplus_{i=1}^{L} f(a_i(y))p_i(y)\| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

This ends the proof of Theorem 4.1.

$\square$

Attention: In fact, we proved that the conclusion of Theorem 4.1 holds not only for f in the finite set F, but also for any f satisfying the condition that if $\mathrm{dist}(x, x') < 2\eta$, then $\|f(x) - f(x')\| < \frac{\varepsilon}{3}$.

Remark 4.20. The following is the $(*)$ from 4.19:

$$(*) \qquad \|p_0(y)\phi(f)(y)p_0(y) - [\sum_{i=1}^{L} f(a_i(y))(P_i(y) - p_i(y)) + \sum_{i=1}^{s} f(\mu_i)q_i]\| < \frac{\varepsilon}{3}.$$

Recall that for any $x, x' \in X$, if $\mathrm{dist}(x, x') < 2\eta$, then

$$\|f(x) - f(x')\| < \frac{\varepsilon}{3}$$

for all $f \in F$.

Note that $\xi_y : C(X) \to p_0(y)M_\bullet(\mathbb{C})p_0(y)$, defined by
$\xi_y(f) = \sum_{i=1}^{L} f(a_i(y))(P_i(y) - p_i(y)) + \sum_{i=1}^{s} f(\mu_i)q_i$, is a homomorphism. By 1.2.18, we have the following claim.

Claim: Let $\{x_1, x_2, \cdots, x_r\}$ be an η-dense subset of X. Suppose that mutually orthogonal projections $p^1, p^2, \cdots, p^r \in (P-p_0)M_{K'}(C(Y))(P-p_0)$ satisfy

$$\operatorname{rank}(p^i) \geq \operatorname{rank}(p_0).$$

Let $\psi : C(X) \to (p_0 \oplus p^1 \oplus p^2 \oplus \cdots \oplus p^r)M_{K'}(C(Y))(p_0 \oplus p^1 \oplus p^2 \oplus \cdots \oplus p^r)$ be the positive linear map defined by

$$\psi(g) = p_0\phi(g)p_0 \oplus \sum_{i=1}^{r} g(x_i)p^i \ ,$$

for all $g \in C(X)$. Then $\psi(F)$ is weakly approximately constant to within ε. This fact will be used later.

Remark 4.21. The proof of Theorem 4.1 is very long and complicated. We point out that the following direct approaches will encounter difficulties. (These discussions have appeared in §1.5.)

1. One may let $P_i^U(y)$ be the spectral projections corresponding to the open sets

$$\{\lambda;\ \operatorname{dist}(\lambda, a_i(y)) < \eta\}$$

and make use of Proposition 3.2 of [DNNP] to construct the projection p_i. The trouble is that such $\{p_i\}_{i=1}^{L}$ are not mutually orthogonal since P_i^U are not mutually orthogonal.

2. For each sufficiently small neighborhood U, applying the theorem about spectral multiplicity from §2, one can construct mutually orthogonal projections $\{P_i^U(y)\}_{i=1}^{L}$ with relatively large rank such that each $P_i^U(y)$ is the spectral projection corresponding to a subset of

$$\{\lambda;\ \operatorname{dist}(\lambda, a_i(y)) < \eta\}.$$

But one cannot guarantee that the projection associated to $\bigvee\{P_i^U;\ U \ni y\}$ is orthogonal to the projection associated to $\bigvee\{P_j^U;\ U \ni y\}$, for $i \neq j$. So one still can not obtain orthogonal projections $\{p_i\}_{i=1}^{L}$.

3. One may try to define $p_1, p_2, \cdots, p_L$, one by one. For example, after $p_1(y)$ is defined, try to choose $P_2^U(y)$ to be orthogonal to $p_1(y)$ and to be the spectral projection of a certain subset of X. Then this subset can not be chosen to be a subset of $\{\lambda;\ \operatorname{dist}(\lambda, a_2(y)) < \eta\}$ since some spectra may have been taken out when $p_1(y)$ is defined. In fact, this subset can be chosen to be a subset of $\{\lambda;\ \operatorname{dist}(\lambda, a_2(y)) < 2\eta\}$. In this way, when we define $P_i^U(y)$, it will be a spectral projection corresponding to a subset of

$$\{\lambda;\ \operatorname{dist}(\lambda, a_i(y)) < i\cdot\eta\}.$$

In order for the theorem to hold, $L\cdot\eta$ needs to be small, which makes η depend on L. This is not useful at all for the application.

Remark 4.22. Note that in 4.19, when we prove that the projections $\{p_i\}_{i=1}^L$ satisfy the desired conditions (i) and (ii) of Theorem 4.1, we only use the property that for any $y \in Y$, $p_i(y)$, $i = 1, 2, \cdots, L$, are subprojections of $P_i(y), i = 1, 2, \cdots, L$, respectively. This means, (i) and (ii) of Theorem 4.1 hold for any set of projections $\{p_i'\}_{i=1}^L$ with $p_i' \leq p_i$, $i = 1, 2, \cdots, L$. So we have the freedom to replace any p_i by its subprojection (with suitable rank). This fact is important for the discussion below and in 4.41 and 4.44.

In what follows, we will use the fact that, for the projections in $M_{K'}(C(Y))$ of rank at least $\dim(Y)$, cancellation always holds. That is , if three projections p, q and r in $M_\bullet(C(Y))$ satisfy that $\mathrm{rank}(p) > \dim(Y)$, $\mathrm{rank}(q) > \dim(Y)$ and $p \oplus r$ is Murry von Neumann equivalent to $q \oplus r$, then p is Murry von Neumann equivalent to q.

(a) In fact, in 4.19, $\mathrm{rank}(p_i)$ for our projections p_i satisfy the stronger condition (see 4.18):

$$\mathrm{rank}(p_i) \geq T_i - [M\Omega + M\Omega \dim Y(\dim Y + 1)] - \dim Y.$$

From Theorem 1.2 of [Hu, Chapter 8], there is a trivial projection $p_i' < p_i$ such that

$$\begin{aligned} \mathrm{rank}(p_i') &\geq \mathrm{rank}(p_i) - \dim Y \\ &\geq T_i - [M\Omega + M\Omega \dim Y(\dim Y + 1)] - 2\dim Y. \end{aligned}$$

That is, $\mathrm{rank}(p_i')$ is still larger than $T_i - T$, where $T = 2^L(\dim X + \dim Y)^3$. (In fact it is larger than $T_i - T + 2\dim Y$.) In Theorem 4.1, replacing p_i by p_i', one makes all the projections $\{p_i\}_{i=1}^L$ trivial.

(b) Suppose that there is an $i_0 \in \{1, 2, \cdots, L\}$ such that $T_{i_0} > T + \dim Y$. Suppose that the projections $p_1, p_2, \cdots, p_L$ are trivial as in (a). In particular, suppose that $\mathrm{rank}(p_{i_0}) \geq T_{i_0} - T + 2\dim Y$ as mentioned in (a). By [Hu], $P \in M_{K'}(C(Y))$ (the total projection of the target algebra $PM_{K'}(C(Y))P$ in Theorem 4.1) can be written in the form

$$q \oplus (\text{trivial projection}),$$

where q is of rank $T_{i_0} - T + \dim Y$. It follows from [Hu], that there is a subprojection p_{i_0}' of p_{i_0} which is unitarily equivalent to q. Replacing p_{i_0} by p_{i_0}', and keeping all the other projections p_i, then P will be unitarily equivalent to a projection of the form

$$\bigoplus_{i=1}^L p_i \oplus \ (\text{trivial projection}).$$

Therefore, $p_0 = P - \bigoplus_{i=1}^L p_i$ is a trivial projection. (Note that $\mathrm{rank}(p_0) \geq \dim Y$.)

In other words, in Theorem 4.1, we can choose all the projections $p_0, p_1, \cdots, p_L$ to be trivial except one of them, p_{i_0}, where $i_0 \neq 0$. In particular, p_0 is a trivial projection, as comparing with (a) above.

The following theorem is proved in [EGL].

Theorem 4.23. ([EGL]) *Let $A = \lim_{n\to\infty}(A_n, \phi_{n,m})$ be an inductive limit C^*-algebra (not necessarily unital) with*

$$A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})),$$

where $X_{n,i}$ are simplicial complexs. Then one can write $A = \lim_{n\to\infty}(B_n, \psi_{n,m})$ with

$$B_n = \bigoplus_{i=1}^{t_n} M_{\{n,i\}}(C(Y_{n,i})),$$

where $Y_{n,i}$ are (not necessarily connected) simplicial complexs, with $\dim(Y_{n,i}) \leq \dim(X_{n,i})$, such that all the connecting maps $\psi_{n,m}$ are injective.
Furthermore, if $(A_n, \phi_{n,m})$ satisfies the very slow dimension growth condition, then so does $(B_n, \psi_{n,m})$.

4.24. Without loss of generality, in the rest of this article, we will assume that the connecting maps $\phi_{n,m}$ in the inductive limit system are injective. Without this assumption one can still prove all the theorems in this paper by modifying our arguments, and by passing to some good subsets of $X_{n,i}$. But this assumption makes the discussions much simpler. As mentioned in 1.1.5, we will suppose that the inductive limit algebra $A = \lim_{n\to\infty}(A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$ satisfies the very slow dimension growth condition.

4.25. As a consequence of Theorem 4.1 and the lemma inside 1.5.11— a result due to Li—, one can obtain a decomposition for each (partial map of a) connecting map $\phi^{i,j}_{n,m}$ (m large enough), with the major part factoring though an interval algebra. But for our application, we need a certain part of the decomposition to be defined by point evaluations and (even if it is not large absolutely) to be relatively large compared to the "bad" part $p_0\phi p_0$, where p_0 is the projection in Theorem 4.1, and ϕ is the map corresponding to $\phi^{i,j}_{n,m}$ (see 1.2.18 and 1.2.19), i.e., $\phi = \phi^{i,j}_{n,m}|_{e_{11}A^i_n e_{11}}$.
Following Section 2 of [Li3] (see the proof of Theorem 2.28 in [Li3]), we can prove our main Decomposition Theorem (see Theorem 4.37 below). [Li3] only proves the special case that $X_{n,i}$ = graphs (one dimensional spaces). Although the idea behind Li's proof is reasonably simple and clear (see the explanation in 2.29 of [Li3]), the proof itself is complicated and long. It combines several difficulties together. For convenience in the higher dimensional case, we will give a slightly different approach. (See 1.5.25 for the explanation of the difference between our approach and Li's approach.) Our proof will be a little shorter, and perhaps easier to follow (hopefully). More importantly, using this approach, we

will be able to prove the Decomposition Theorem for any homomorphism provided that the homomorphism satisfies a certain quantitative condition (see Theorem 4.35 below). (Li's theorem is for the homomorphism $\phi_{n,m}$ with m sufficiently large.) This slightly stronger version of the theorem is needed in [EGL] to prove the Uniqueness Theorem. It should be emphasized that our proof is essentially the same as Li's proof in spirit.

The idea behind our proof is roughly as follows.

In [Li2, 2.18–2.19] (see 1.5.11), Li proves that for fixed $\eta > 0$, for m large enough, and for any (partial) connecting map $\phi_{n,m}^{i,j}$—denoted by ϕ—, there are L continuous maps $\beta_1, \beta_2, \cdots, \beta_L : Y\ (= X_{m,j}) \longrightarrow X\ (= X_{n,i})$, factoring through the interval $[0,1]$, such that for each $y \in Y$, the set $\mathrm{SP}\phi_y$ and the set

$$\Theta(y) = \left\{\beta_1(y)^{\sim L_2}, \beta_2(y)^{\sim L_2}, \cdots, \beta_{L-1}(y)^{\sim L_2}, \beta_L(y)^{\sim L_2+L_1}\right\}$$

can be paired within η, where L_2 could be very large compared with $L \cdot 2^L \cdot (\dim(X) + \dim(Y))^3$, if the inductive limit system satisfies the very slow dimension growth condition.

What we are going to prove is that, if $\mathrm{SP}\phi_y$ and $\Theta(y)$ can be paired within η, then they can still be paired within some small number (e.g., 2η), if one changes a number—a small number compared with L—of maps β_i to ARBITRARY maps (in particular, to constant maps), provided that X is path connected and ϕ has a certain spectral distribution property related to the number η and another number δ (see 4.26 below). (Note that, how many maps are allowed to be changed, also depends on η and δ.) (Those constant maps form the part of the homomorphism defined by point evaluations.) At first sight, it might seem impossible for this to be true. But, with the spectral distribution property of the homomorphism ϕ, Lemma 2.15 of [Li2] (see Lemma 4.29 below) says that if ϕ and another homomorphism ψ (in the application, ψ should be chosen to be a homomorphism with the family of spectral functions $\Theta(y)$, i.e., $\mathrm{SP}\psi_y = \Theta(y)$ for all $y \in Y$) are close on the level of Aff$T$, then their spectra $\mathrm{SP}\phi_y$ and $\mathrm{SP}\psi_y$ can be paired within a small number. On the other hand, changing a very few spectral functions (no matter how large a change in each function), will NOT create a big change on the level of Aff$T$ (see 4.28 and the claims in 4.31 below). Since the results of [Li 2] are not of a quantitative nature—they are for connecting homomorphisms $\phi_{n,m}$ with m large—, we can not apply them (2.18 and 2.19 of [Li 2]) directly. So we repeat part of the arguments in [Li 2].

The above method will lead us to Lemma 4.33 (see 4.26–4.33 for details). Then our main decomposition theorems—Theorem 4.35 and Theorem 4.37— will be more or less consequences.

Finally we remark that, in our decomposition, we cannot require that both parts of the decomposition be homomorphisms as in 2.28 of [Li3], since in general, $C(X)$ is not stably generated (see [Lo]).

4.26. For the reader's convenience, we will quote some notations, terminologies and results from [Li1] and [Li2].

The following notation is inspired by a similar notation in [Li1].
For any $\eta > 0$, $\delta > 0$, a homomorphism $\phi\colon PM_k(C(X))P \to QM_{k'}(C(Y))Q$ is said to have the property $\mathrm{sdp}(\eta,\delta)$ (SPECTRAL DISTRIBUTION PROPERTY WITH RESPECT TO η AND δ) if for any $\eta-$ball

$$B_\eta(x) := \{x' \in X;\ \mathrm{dist}(x',x) < \eta\} \subset X$$

and any point $y \in Y$,

$$\#(\mathrm{SP}\phi_y \cap B_\eta(x)) \ \geq\ \delta\#(\mathrm{SP}\phi_y),$$

counting multiplicity.
(Attension: The property $\mathrm{sdp}(r,\delta)$ in [Li1] corresponds to $\mathrm{sdp}(\frac{1}{2r},\delta)$ above.)
Any homomorphism $\phi: \ \oplus M_k(C(X)) \to \oplus M_l(C(Y))$ is said to have the property $\mathrm{sdp}(\eta,\delta)$ if each partial map has $\mathrm{sdp}(\eta,\delta)$.

4.27. The following notations can be found in Section 2 of [Li2]. Let X be a connected simplicial complex. For any closed set $X_1 \subset X$, $M > 0$, let

$$\chi_{X_1,M}(x) = \begin{cases} 1 & \text{if } \ x \in X_1 \\ 1 - M\cdot \mathrm{dist}(x,X_1) & \text{if } \ \mathrm{dist}(x,X_1) \leq \frac{1}{M} \\ 0 & \text{if } \ \mathrm{dist}(x,X_1) \geq \frac{1}{M} \end{cases}.$$

For $\eta > 0$ and $\delta > 0$, let

$$H_1(\eta) = \{\chi_{X_1,\frac{32}{\eta}}:\ X_1 \subset X \ \text{ closed } \}.$$

Then there is a finite set $H \subset H_1(\eta)$ such that for all $h \in H_1(\eta)$, $\mathrm{dist}(h,H) < \frac{\delta}{8}$ (the distance is the distance defined by uniform norm). Denote such set by $H(\eta,\delta,X)$ $(\subset C(X))$. Although such a set is not unique, we fix one for each triple (η,δ,X) for our purpose. (As pointed out in [Li1], the existence of such finite set $H(\eta,\delta,X)$ follows from equi-continuity of the functions in $H_1(\eta)$.)

4.28. For a unital C^*-algebra A, let TA denote the space of all tracial states of A, i.e., $\tau \in TA$ if and only if τ is a positive linear map from A to $\mathbb{C}$, with $\tau(xy) = \tau(yx)$ and $\tau(\mathbf{1}) = 1$. $\mathrm{Aff}TA$ is the collection of all the affine maps from TA to $\mathbb{R}$.
Any unital homomorphism $\phi:\ A \to B$ induces an affine map

$$\mathrm{Aff}T\phi:\ \mathrm{Aff}TA \longrightarrow \mathrm{Aff}TB.$$

It is well known, for any connected metrizable space X and any projection $P \in M_k(C(X))$,

$$\mathrm{Aff}T(PM_k(C(X))P) = \mathrm{Aff}T(C(X)) = C_{\mathbb{R}}(X).$$

We would like to quote some easy facts about the $\mathrm{Aff}T$ map from [Li1] and [Li2].

If $\phi : C(X) \to PM_l(C(Y))P$ is a unital homomorphism and $\mathrm{rank}(P) = k$, then $\mathrm{Aff}T\phi : C(X) \to C(Y)$ is given by

$$\mathrm{Aff}T\phi(f) = \frac{1}{k}\sum_{i=1}^{l} \phi(f)_{ii},$$

where each $\phi(f)_{ii}$ is the diagonal entry of $\phi(f) \in PM_l(C(Y))P \subset M_l(C(Y))$ at the place (i, i)

For a continuous map $\beta : Y \to X$, let $\beta^* : C(X) \to C(Y)$ be defined by

$$\beta^*(f) = f \circ \beta \quad (\in C(Y)) \text{ for any } f \in C(X).$$

Suppose that $\beta_1, \beta_2, \cdots \beta_l : Y \to X$ are continuous maps. If $\psi : C(X) \to M_l(C(Y))$ is a homomorphism with $\{\beta_i\}_{i=1}^l$ as the set of spectral functions, (e.g., ψ is defined by $\psi(f) = diag(\beta_1^*(f), \beta_2^*(f), \cdots, \beta_l^*(f)),$) then

$$\mathrm{Aff}T\psi(f) = \frac{1}{l}\sum_{i=1}^{l} \beta_i^*(f).$$

(Let $H \subset C_{\mathbb{R}}(X)$ be a finite subset satisfying $\|f\| \le 1$ for any $f \in H$. If one modifies the above homomorphism ψ to a new homomorphism ψ', by replacing k functions from the set of spectral functions $\{\beta_i\}_{i=1}^l$ by other functions (from Y to X), then

$$\|\mathrm{Aff}T\psi(f) - \mathrm{Aff}T\psi'(f)\| \le \frac{k}{l}, \ \forall f \in H.$$

In particular, this modification (from ψ to ψ') does not create a big change on the level of $\mathrm{Aff}T$, provided that k is very small compared with l, as mentioned in 4.25.)

For a unital homomorphism $\phi : C(X) \to PM_l(C(Y))P$ with $rank(P) = k$, quoting from 1.9 of [Li1], we have

$$\mathrm{Aff}T\phi(f)(y) = \frac{1}{k}\sum_{x_i(y)\in \mathrm{SP}\phi_y} f(x_i(y)).$$

Consider $e_y : PM_l(C(Y))P \to P(y)M_l(\mathbb{C})P(y) \quad (\cong M_{rank(P)}(\mathbb{C}))$, which is the homomorphism defined by evaluation at the point y. Then from the above paragraph, we know that $\mathrm{Aff}T(e_y \circ \phi)$ depends only on $\mathrm{SP}\phi_y$. We can denote $e_y \circ \phi$ by $\phi|_y$ (this is the homomorphism $\phi|_{\{y\}}$ in 1.2.13 for the single point set $\{y\}$).

Lemma 4.29. ([Li2,2.15]) *Suppose that two unital homomorphisms $\phi : C(X) \to PM_k(C(Y))P$ and $\psi : C(X) \to QM_k(C(Y))Q$ with $rank(P) = rank(Q)$, satisfy the following two conditions:*
(1) ϕ has the property $sdp(\frac{\eta}{32}, \delta)$;
(2) $\|\mathit{Aff}T\phi(h) - \mathit{Aff}T\psi(h)\| < \frac{\delta}{4}$, for all $h \in H(\eta, \delta, X)$.

Then $SP\phi_y$ and $SP\psi_y$ can be paired within $\frac{\eta}{4}$ for any $y \in Y$.
(Notice that, no matter how small the δ is , the above conditions (1) and (2) do not imply the other assumption $\mathrm{rank}(P) = \mathrm{rank}(Q)$, which is necessary for our conclusion.)

Proof: If $P = Q$, this is exactly 2.15 of [Li2]. (Notice that, we use $\frac{\eta}{32}$ in place of $\frac{\eta}{8}$ of [Li2 2.15], so our conclusion is that, $\mathrm{SP}\phi_y$ and $\mathrm{SP}\psi_y$ can be paired within $\frac{\eta}{4}$ (instead of η). Also notice that the set H in [Li2 2.15] is chosen to be the same as the above set $H(\eta, \delta, X)$.)
To see the general case, fix $y \in Y$. We can consider two maps $\phi|_y$ and $\psi|_y$ which are unital homomorphisms from $C(X)$ to C^*-algebras which are isomorphic to the same C^*-algebra $M_{rank(P)}(\mathbb{C})$. (Note that $rank(P) = rank(Q)$.) The conditions (1) and (2) above imply the same conditions for $\phi|_y$ and $\psi|_y$, since $\mathrm{Aff}T\phi(h)(y) = \mathrm{Aff}T(\phi|_y)(h)$. Therefore, by 2.15 of [Li 2], $\mathrm{SP}\phi_y = SP(\phi|_y)$ and $\mathrm{SP}\psi_y = SP(\psi|_y)$ can be paired within $\frac{\eta}{4}$.
(If one checks the proof of 2.15 of [Li2] carefully, then he will easily recognize that the above Lemma is already proved there.)

$\square$

4.30. In the following paragraphs (4.30—4.32), we will apply the materials from 2.8 – 2.10 of [Li2].
For any $\eta > 0$ and $\delta > 0$, from 2.9 of [Li2], there exist a continuous map $\alpha : [0,1] \to X$, and a unital positive linear map $\xi : C[0,1] \to C(X)$ such that

$$\|\xi \circ \alpha^*(f) - f\| < \frac{\delta}{16},$$

for each $f \in H(\eta, \delta, X)$, where $\alpha^* : C(X) \to C[0,1]$ is induced by α. Furthermore, we can choose α such that image(α) is $\frac{\eta}{32}$-dense in X.
For $\alpha : [0,1] \to X$, there is a $\sigma > 0$ such that $|t - t'| < 2\sigma$ implies that

$$\mathrm{dist}(\alpha(t), \alpha(t')) < \frac{\eta}{32}.$$

For a fixed space X, the number σ depends only on η and δ, since so does the continuous map α. We denote the number σ by $\sigma(\eta, \delta)$.

4.31. Let $\tilde{H} = \alpha^*(H(\eta, \delta, X)) \subset C[0,1]$. For the finite set $\tilde{H}$ and $\frac{\delta}{16} > 0$, there is an integer N (as in Theorem 2.1 of [Li2]) such that for any positive linear map $\zeta : C[0,1] \to C(Y)$, and for any $r \geq N$, there are r continuous maps

$$\beta_1, \beta_2, \cdots, \beta_r : Y \longrightarrow [0,1]$$

such that

$$\left\| \zeta(f) - \frac{1}{r} \sum_{i=1}^{r} \beta_i^*(f) \right\| < \frac{\delta}{16}$$

for all $f \in \tilde{H}$, where $\beta_i^* : C[0,1] \to C(Y)$ is induced by β_i.

We will also assume $\frac{1}{N} < \frac{\delta}{64}$. Then we can prove the following claim.
Claim 1: For any $r \geq N$, if

$$\left\| \zeta(f) - \frac{1}{r}\sum_{i=1}^{r} \beta_i^*(f) \right\| < \frac{\delta}{16} \qquad \text{for all} \quad f \in \tilde{H},$$

then for any other two continuous maps $\tau_1, \tau_2 : Y \longrightarrow [0,1]$,

$$\left\| \zeta(f) - \frac{1}{r+2}\left(\sum_{i=1}^{r} \beta_i^*(f) + \tau_1^*(f) + \tau_2^*(f) \right) \right\| < \frac{\delta}{8}, \quad \forall f \in \tilde{H}$$

Proof of the claim: The claim follows from

$$\begin{aligned} &\left\| \zeta(f) - \frac{1}{r+2}\left(\sum_{i=1}^{r} \beta_i^*(f) + \tau_1^*(f) + \tau_2^*(f) \right) \right\| \\ &\quad \leq \left\| \zeta(f) - \frac{1}{r}\sum_{i=1}^{r} \beta_i^*(f) \right\| + \left\| \frac{1}{r}\sum_{i=1}^{r} \beta_i^*(f) - \frac{1}{r+2}\sum_{i=1}^{r} \beta_i^*(f) \right\| \\ &\qquad\qquad + \left\| \frac{1}{r+2}(\tau_1^*(f) + \tau_2^*(f)) \right\| \\ &\quad < \frac{\delta}{16} + 2\cdot\frac{\delta}{64} + 2\cdot\frac{\delta}{64} = \frac{\delta}{8} \end{aligned}$$

for any $f \in \tilde{H}$. In the above estimation, we use the facts that $\|f\| \leq 1$, $\|\beta_i^*(f)\| \leq 1$ and $\|\tau_i^*(f)\| \leq 1$ for any $f \in \tilde{H}$.
In the above claim, if we replace the condition $r \geq N$ by the condition $r \geq mN$, then in the conclusion, we can allow $2m$ continuous maps $\tau_1, \tau_2, \cdots, \tau_{2m} : Y \longrightarrow [0,1]$, instead of two maps. Namely, the following claim can be proved in exactly the same way.
Claim 2: For any $r \geq mN$, if

$$\left\| \zeta(f) - \frac{1}{r}\sum_{i=1}^{r} \beta_i^*(f) \right\| < \frac{\delta}{16}, \qquad \forall\ f \in \tilde{H},$$

then for any $2m$ continuous maps $\tau_1, \tau_2, \cdots, \tau_{2m} : Y \longrightarrow [0,1]$,

$$\left\| \zeta(f) - \frac{1}{r+2m}\left(\sum_{i=1}^{r} \beta_i^*(f) + \sum_{i=1}^{2m} \tau_i^*(f) \right) \right\| < \frac{\delta}{8}, \quad \forall f \in \tilde{H}.$$

4.32. Let $n = \mathrm{int}\left(\frac{1}{\sigma(\eta,\delta)}\right) + 1$, where $\mathrm{int}(\cdot)$ denote the integer part of the number (see 1.1.7 (c)).
Divide $[0,1]$ into n intervals such that each of them has length at most $\sigma(\eta,\delta)$. Choose n points

$$t_1, t_2, \cdots, t_n,$$

one from each of the intervals. Let $x_i = \alpha(t_i) \in X,\ i = 1, 2, \cdots, m$. Then the set

$$\{x_1, x_2, \cdots, x_n\}$$

is $\frac{\eta}{16}$-dense in X by the way σ is chosen in 4.30.

From the above discussion, for fixed $\eta > 0, \delta > 0$, and the space X, we can find $\alpha, \xi, \sigma, N, n, H(\eta, \delta, X), \tilde{H}$, the set $\{t_1, t_2, \cdots, t_n\} \subset [0,1]$, and the $\frac{\eta}{16}$-dense set $\{x_1, x_2, \cdots, x_n\} \subset X$. All of them depend only on η, δ, and the space X.

Lemma 4.33. *For any connected simplicial complex X, any numbers $\eta > 0$ and $\delta > 0$, there are integers n, N, a continuous map $\alpha :\ [0,1] \to X$, and finitely many points $\{t_1, t_2, \cdots, t_n\} \subset [0,1]$ with $\{\alpha(t_1), \alpha(t_2), \cdots, \alpha(t_n)\}$ $\frac{\eta}{16}$-dense in X, such that the following is true. (Denote $L := n(N+2)$.)*

If a unital homomorphism $\phi : C(X) \to PM_k(C(Y))P$ satisfies the following two conditions:

(i) ϕ has the property sdp$(\frac{\eta}{32}, \delta)$;

(ii) rank$\phi(\mathbf{1}) := K \geq L^2 = (n(N+2))^2$,

and write $K = LL_2 + L_1$ with $L_2 = int(\frac{K}{L})$ and $0 \leq L_1 < L$, (note that $L \leq L_2$, since $K \geq L^2$,)

then there are L continuous functions

$$\beta_1, \beta_2, \cdots, \beta_n, \beta_{n+1}, \cdots, \beta_L :\ Y \longrightarrow [0,1]$$

such that

(1) $\beta_i(y) = t_i$ for $1 \leq i \leq n$;

(2) For each $y \in Y$, $SP\phi_y$ and the set

$$\Theta(y) = \left\{\alpha{\circ}\beta_1(y)^{\sim L_2}, \alpha{\circ}\beta_2(y)^{\sim L_2}, \cdots, \alpha{\circ}\beta_{L-1}(y)^{\sim L_2}, \alpha{\circ}\beta_L(y)^{\sim L_2+L_1}\right\}$$

can be paired within $\frac{\eta}{2}$.

(3) If Y is a connected finite simplicial complex and $Y \neq \{pt\}$, then the map $\beta_{n+1} :\ Y \to [0,1]$—the first nonconstant map above—, is a surjection.

(This lemma is similar to Lemma 2.18 of [Li2], but we require some of the functions β_i $(1 \leq i \leq n)$ to be constant functions.)

(Attention: To apply Theorem 4.1, one only needs SPϕ_y and $\Theta(y)$ to be paired within η. The advantage of using $\frac{\eta}{2}$ is the following. If ψ is another homomorphism such that SPψ_y and SPϕ_y can be paired within $\frac{\eta}{2}$ for any y, then we can apply Theorem 4.1 to both ϕ and ψ without requiring ψ to have the property sdp$(\frac{\eta}{32}, \delta)$. This observation will not be used in the proof of the main theorem of this paper. But it will be used in the proof of the Uniqueness Theorem in [EGL] (part II of the series), see 4.41–4.48 below.)

Proof: Follow the notations in 4.26 – 4.32. Let

$$\zeta :\ C[0,1] \longrightarrow C(Y)$$

be defined by $\zeta = \mathrm{Aff}T\phi\circ\xi$. Since $K - 2nL_2 \geq nL_2N$, there are $K - 2nL_2$ continuous maps

$$\gamma_1, \gamma_2, \cdots, \gamma_{K-2nL_2} :\ Y \longrightarrow [0,1]$$

such that

$$\left\| \zeta(f) - \frac{1}{K-2nL_2} \sum_{i=1}^{K-2nL_2} \gamma_i^*(f) \right\| < \frac{\delta}{16}$$

for all $f \in \tilde{H}$. Let

$$\beta_1, \beta_2, \cdots, \beta_n :\ Y \longrightarrow [0,1]$$

be defined by $\beta_i(y) = t_i$. Then by Claim 2 of 4.31 (taking $m = nL_2$),

$$\left\| \zeta(f) - \frac{1}{K} \left(\sum_{i=1}^{K-2nL_2} \gamma_i^*(f) + 2L_2 \sum_{i=1}^{n} \beta_i^*(f) \right) \right\| < \frac{\delta}{8}$$

for all $f \in \tilde{H} \subset C[0,1]$. Therefore,

$$\left\| (\zeta\circ\alpha^*)(f) - \frac{1}{K} \left(\sum_{i=1}^{K-2nL_2} (\alpha\circ\gamma_i)^*(f) + 2L_2 \sum_{i=1}^{n} (\alpha\circ\beta_i)^*(f) \right) \right\| < \frac{\delta}{8}$$

for all $f \in H(\eta, \delta, X)$. On the other hand, by 4.30 and $\zeta = \mathrm{Aff}T\phi\circ\xi$,

$$\|\mathrm{Aff}T\phi(f) - (\zeta\circ\alpha^*)(f)\| < \frac{\delta}{16} \qquad \text{for } \ f \in H(\eta, \delta, X).$$

One can define a unital homomorphism $\psi :\ C(X) \to M_K(C(Y))$ with $\{\alpha\circ\gamma_i\}_{i=1}^{K-2nL_2} \cup \{(\alpha\circ\beta_i)^{\sim 2L_2}\}_{i=1}^{n}$ as the family of the spectral functions. Then from 4.28,

$$\mathrm{Aff}T\psi(f) = \frac{1}{K} \left(\sum_{i=1}^{K-2nL_2} (\alpha\circ\gamma_i)^*(f) + 2L_2 \sum_{i=1}^{n} (\alpha\circ\beta_i)^*(f) \right).$$

Hence,

$$\|\mathrm{Aff}T\phi(f) - \mathrm{Aff}T\psi(f)\| \leq \frac{\delta}{8} + \frac{\delta}{16} < \frac{\delta}{4}$$

for all $f \in H(\eta, \delta, X)$. Note that $\mathrm{rank}(P) = K$. By Lemma 4.29, $\mathrm{SP}\phi_y$ and

$$\mathrm{SP}\psi_y = \left\{ \alpha\circ\beta_1(y)^{\sim 2L_2}, \alpha\circ\beta_2(y)^{\sim 2L_2}, \cdots, \alpha\circ\beta_n(y)^{\sim 2L_2}, \alpha\circ\gamma_1(y), \cdots, \alpha\circ\gamma_{L-2nL_2}(y) \right\}$$

can be paired within $\frac{\eta}{4}$.

Note that in our lemma, we only need L_2 copies of each constant maps β_i ($i = 1, 2, \cdots, n$). One may wonder why we put $2L_2$ copies of each of maps β_i in the above set. The reason is that, after taking out L_2 copies of β_i, we still want the set $\Theta(y)$ to have enough elements in each small interval of length σ, and the other L_2 copies of β_i can serve for this purpose.

Consider the following set (of $K - nL_2$ elements)

$$\{\beta_1^{\sim L_2}(y), \beta_2^{\sim L_2}(y), \cdots, \beta_n^{\sim L_2}(y),\ \gamma_1(y), \gamma_2(y), \cdots, \gamma_{K-2nL_2}(y)\}.$$

In each interval of $[0,1]$ of length σ, there are at least L_2 points (counting multiplicities) in the above set.
The following argument appeared in 2.18 of [Li2].
For each fixed y, we can rearrange all the elements in the above set in the increasing order. I.e., write them as $\gamma_1'(y), \gamma_2'(y), \cdots, \gamma_{K-nL_2}'(y)$ such that for each fixed y

$$\begin{aligned}\{\gamma_1'(y), \gamma_2'(y), \cdots, \gamma_{K-nL_2}'(y)\} = \\ = \{\beta_1(y)^{\sim L_2}, \beta_2(y)^{\sim L_2}, \cdots, \beta_n(y)^{\sim L_2},\ \gamma_1(y), \gamma_2(y), \cdots, \gamma_{K-2nL_2}(y)\}\end{aligned}$$

(as a set with multiplicity), and such that

$$0 \le \gamma_1'(y) \le \gamma_2'(y) \le \cdots \le \gamma_{K-nL_2}'(y) \le 1.$$

It is easy to prove that $\gamma_i'(y),\ 1 \le i \le K - nL_2$ are continuous (real-valued) functions, using the following well known fact repeatedly: For any two real-valued continuous functions f and g, the functions $\max(f,g)$ and $\min(f,g)$ are also continuous.
We can put each group of L_2 consecutive functions of $\{\gamma_i'\}$ (beginning with smallest one) together except the last $L_2 + L_1$ functions which will be put into a single group—the last group. Then we replace all the functions in a same group by the smallest function in the group. Namely, let

$$\beta_{n+1} = \gamma_1', \beta_{n+2} = \gamma_{L_2+1}', \cdots, \beta_L = \gamma_{(L-n-1)L_2+1}' \ .$$

Then from the fact that in each interval of $[0,1]$ of length σ, there are at least L_2 points (counting multiplicity) in the set $\{\gamma_1'(y), \gamma_2'(y), \cdots, \gamma_{K-nL_2}'(y)\}$, we know that $\{\gamma_1'(y), \gamma_2'(y), \cdots, \gamma_{K-nL_2}'(y)\}$ and

$$\{\beta_{n+1}(y)^{\sim L_2}, \beta_{n+2}(y)^{\sim L_2}, \cdots, \beta_{L-1}(y)^{\sim L_2}, \beta_L(y)^{\sim L_2+L_1}\}$$

can be paired within 2σ. Recall that $|t-t'| < 2\sigma$ implies that $\mathrm{dist}(\alpha(t), \alpha(t')) < \frac{\eta}{16}$. Hence

$$\begin{aligned}\big\{\alpha\circ\beta_1(y)^{\sim 2L_2}, \alpha\circ\beta_2(y)^{\sim 2L_2}, \cdots, \alpha\circ\beta_n(y)^{\sim 2L_2}, \\ \alpha\circ\gamma_1(y), \alpha\circ\gamma_2(y), \cdots, \alpha\circ\gamma_{L-2nL_2}(y)\}\end{aligned}$$

and

$$\begin{aligned}\Theta(y) = \big\{\alpha\circ\beta_1(y)^{\sim L_2}, \alpha\circ\beta_2(y)^{\sim L_2}, \cdots, \alpha\circ\beta_n(y)^{\sim L_2}, \\ \alpha\circ\beta_{n+1}(y)^{\sim L_2}, \cdots, \alpha\circ\beta_{L-1}(y)^{\sim L_2}, \alpha\circ\beta_L(y)^{\sim L_2+L_1}\big\}\end{aligned}$$

can be paired within $\frac{\eta}{16}$. Therefore, $\mathrm{SP}\phi_y$ and $\Theta(y)$ can be paired within $\frac{\eta}{4}+\frac{\eta}{16}<\frac{5\eta}{16}$.
Note that $\{\alpha\circ\beta_1(y),\alpha\circ\beta_2(y),\cdots,\alpha\circ\beta_n(y)\}$ is $\frac{\eta}{16}$ dense in X. From the proof of Lemma 1.2.17, if we replace only one map (say β_{n+1}) by an arbitrary map from Y to $[0,1]$, then the new $\Theta(y)$ can be paired with the old $\Theta(y)$ to within $\frac{\eta}{8}$. As a consequence, we still have that $\mathrm{SP}\phi_y$ and the new $\Theta(y)$ can be paired within $\frac{5\eta}{16}+\frac{\eta}{8}<\frac{\eta}{2}$. In particular, if Y is a connected finite simplicial complex which is not a single point, then β_{n+1} could be chosen to be a surjection, again using the Peano curve.

$\square$

4.34. Fix a large positive integer J. We require that the decomposition in 4.1 to satisfy the condition

$$J\cdot(\mathrm{rank}(p_0)+2\dim(Y))\le rank(p_i),\ \forall i\ge 1.$$

To do so, we need $\mathrm{rank}(\phi(1))$ to be large enough. We describe it as follows. For a connected simplicial complex X, and for numbers $\eta>0$ and $\delta>0$, let N,n and $\alpha:[0,1]\to X$ be as in Lemma 4.33. Let $L=n(N+2)$. Suppose that $\phi:C(X)\to M_k(C(Y))$ is a homomorphism. If ϕ has the property $\mathrm{sdp}(\frac{\eta}{32},\delta)$ and

$$\mathrm{rank}(\phi(1))\ge 2JL^2\cdot 2^L(\dim X+\dim Y+1)^3,$$

then there are continuous functions

$$\beta_1,\beta_2,\cdots,\beta_L:\ Y\longrightarrow[0,1]$$

(as in Lemma 4.33) such that $\mathrm{SP}\phi_y$ and the set

$$\left\{\alpha\circ\beta_1(y)^{\sim L_2},\alpha\circ\beta_2(y)^{\sim L_2},\cdots,\alpha\circ\beta_{L-1}(y)^{\sim L_2},\alpha\circ\beta_L(y)^{\sim L_2+L_1}\right\}$$

can be paired within $\frac{\eta}{2}$, where

$$L_2=\mathrm{int}\left(\frac{\mathrm{rank}\phi(\mathbf{1})}{L}\right)\ge 2JL\cdot 2^L(\dim X+\dim Y+1)^3,$$

and $0\le L_1<L$. For any given set $F\subset C(X)$, if η is chosen as in Theorem 4.1 (see 4.4), then by Theorem 4.1, there are mutually orthogonal projections $p_1,p_2,\cdots,p_L$ and $p_0=\phi(\mathbf{1})-\sum_{i=1}^{L}p_i$ such that
(1) For all $f\in F$ and $y\in Y$,

$$\|\phi(f)(y)-p_0(y)\phi(f)(y)p_0(y)\oplus\bigoplus_{i=1}^{L}f(a_i(y))p_i(y)\|<\varepsilon;$$

(2) For each $i=1,2,\cdots,L$, $\mathrm{rank}(p_i)\ge L_2-2^L(\dim X+\dim Y+1)^3$, and

$$J(\mathrm{rank}(p_0)+2\dim(Y))\le J(L\cdot 2^L(\dim X+\dim Y+1)^3+2\dim Y)\le\mathrm{rank}(p_i).$$

By [Hu], $\underbrace{p_0 \oplus p_0 \oplus \cdots \oplus p_0}_{J}$ is (unitarily) equivalent to a subprojection of p_i, since every complex vector bundle (over Y) of dimension $J \cdot \mathrm{rank}(p_0)$ is a subbundle of any vector bundle (over Y) of dimension at least $J \cdot \mathrm{rank}(p_0) + dim(Y)$. We denote this fact by $J[p_0] < [p_i]$.

Let $Q_0 = p_0$, $Q_1 = p_1 + p_2 + \cdots + p_n$ and $Q_2 = p_{n+1} + p_{n+2} + \cdots + p_L$. Then

$$\phi(\mathbf{1}) = Q_0 + Q_1 + Q_2.$$

Let $\phi_0 : C(X) \to Q_0 M_k(C(Y))Q_0$, $\phi_1 : C(X) \to Q_1 M_k(C(Y))Q_1$ and $\phi_2 : C(X) \to Q_2 M_k(C(Y))Q_2$ be defined by

$$\phi_0(f)(y) = p_0 \phi(f)(y) p_0,$$

$$\phi_1(f) = \sum_{i=1}^{n} f(\alpha_\circ \beta_i(y)) p_i, \quad \text{and}$$

$$\phi_2(f)(y) = \sum_{i=n+1}^{L} f(\alpha_\circ \beta_i(y)) p_i.$$

Then we have the following facts.

(a) ϕ_2 is a homomorphism factoring through $C[0,1]$ as

$$\phi_2 : C(X) \xrightarrow{\xi_1} C[0,1] \xrightarrow{\xi_2} Q_2 M_k(C(Y))Q_2.$$

Furthermore, if $Y \neq \{pt\}$, then ξ_2 is injective. (This follows from the surjection of β_{n+1}.)

(b) Note that

$$\alpha_\circ \beta_1(y) = x_1, \alpha_\circ \beta_2(y) = x_2, \cdots, \alpha_\circ \beta_n(y) = x_n$$

are n constant maps with $\{x_1, x_2, \cdots, x_n\}$ η-dense in X. By the claim in 4.20, $(\phi_0 \oplus \phi_1)(F)$ is approximately constant to within ε. (Note that ϕ_0 is not a homomorphism, it is a completely positive linear $*$-contraction.)

Furthermore, if $\eta < \varepsilon$, then the set $\{x_1, x_2, \cdots, x_n\}$ is ε-dense in X.

Therefore, we have proved the following theorem.

Theorem 4.35. *Let X be a connected finite simplicial complex, and $\varepsilon > \eta > 0$. For any $\delta > 0$, there is an integer $L > 0$ such that the following holds.*

Suppose that $F \subset C(X)$ is a finite set such that $dist(x, x') < 2\eta$ implies $|f(x) - f(x')| < \frac{\varepsilon}{3}$ for all $f \in F$.

If $\phi : C(X) \to M_k(C(Y))$ is a homomorphism with the property $sdp(\frac{\eta}{32}, \delta)$, and $rank(\phi(1)) \geq 2J \cdot L^2 \cdot 2^L(\dim X + \dim Y + 1)^3$, where Y is a connected finite simplicial complex and J is any fixed positive integer, then there are three mutually orthogonal projections $Q_0, Q_1, Q_2 \in M_k(C(Y))$, a map $\phi_0 \in Map(C(X), Q_0 M_k(C(Y))Q_0)_1$ and two homomorphisms

$\phi_1 \in Hom(C(X), Q_1M_k(C(Y))Q_1)_1$ and $\phi_2 \in Hom(C(X), Q_2M_k(C(Y))Q_2)_1$ such that
(1) $\phi(\mathbf{1}) = Q_0 + Q_1 + Q_2$;
(2) $\|\phi(f) - \phi_0(f) \oplus \phi_1(f) \oplus \phi_2(f)\| < \varepsilon$ for all $f \in F$;
(3) The homomorphism ϕ_2 factors through $C[0,1]$ as

$$\phi_2 : C(X) \xrightarrow{\xi_1} C[0,1] \xrightarrow{\xi_2} Q_2M_k(C(Y))Q_2.$$

Furthermore, if $Y \neq \{pt\}$, then ξ_2 is injective;
(4) The set $(\phi_0 \oplus \phi_1)(F)$ is approximately constant to within ε;
(5) $Q_1 = p_1 + \cdots + p_n$, with $J[Q_0] \leq [p_i]$ $(i = 1, 2 \cdots n)$, ϕ_0 is defined by $\phi_0(f) = Q_0\phi(f)Q_0$, and ϕ_1 is defined by

$$\phi_1(f) = \sum_{i=1}^{n} f(x_i)p_i, \ \forall f \in C(X),$$

where $p_0, p_1, \cdots p_n$ are mutually orthogonal projections and $\{x_1, x_2, \cdots x_n\} \subset X$ is an ε-dense subset of X. (Again by $J[p] \leq [q]$, we mean that $\underbrace{p \oplus p \oplus \cdots \oplus p}_{J}$ is (unitarily) equivalent to a subprojection of q.)
Furthermore, we can choose any two of projections Q_0, Q_1, Q_2 to be trivial, if we wish. If $\phi(\mathbf{1})$ is trivial, then all of them can be chosen to be trivial projections. (This is remark 4.22.)

4.36. Let a simple C^*-algebra A be an inductive limit of matrix algebras over simplicial complexes $(A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$ with injective homomorphisms. Suppose that this inductive limit system possesses the very slow dimension growth condition.
In what follows, we will use the material from 1.2.19.
Fix A_n, finite set $F_n = \bigoplus_{i=1}^{t_n} F_n^i \subset A_n$, and $\varepsilon > 0$. Let $\varepsilon' = \frac{\varepsilon}{max_{1\leq i\leq t_n}\{[n,i]\}}$.
Let ${F'}^i \subset C(X_{n,i})$ be the finite set consisting of all the entries of elements in F_n^i $(\subset M_{[n,i]}(C(X_{n,i})))$. Let $\eta > 0$ $(\eta \leq \varepsilon)$ be such that if $x, x' \in X_{n,i}$ $(i = 1, 2 \cdots t_n)$ and $dist(x, x') < 2\eta$, then $|f(x) - f(x')| < \frac{\varepsilon'}{3}$ for any $f \in {F'}^i$.
For the above $\eta > 0$, there is a $\delta > 0$ such that for sufficiently large m, each partial map $\phi_{n,m}^{i,j} : A_n^i \to A_m^j$ has the property sdp$(\frac{\eta}{32}, \delta)$. (This is a consequence of simplicity of the algebra A and injectivity of $\phi_{n,m}$. See [DNNP], [Ell], [Li1-2] for details.)
For these numbers η and δ, and the simplicial complexes $X_{n,i}$, there are $L(i)$, $i = 1, 2, \cdots, t_n$, as in Theorem 4.35. (Note that the numbers L_i only depend on η, δ and the spaces.) Let $L = \max_i L(i)$. Fix a positive integer J. By the very slow dimension growth condition, there is an integer M such that for any $m \geq M$,

$$\frac{\text{rank}\phi_{n,m}^{i,j}(\mathbf{1}_{A_n^i})}{\text{rank}(\mathbf{1}_{A_n^i})} > 2J \cdot L^2 \cdot 2^L(\dim X_{n,i} + \dim X_{m,j} + 1)^3.$$

As in 1.2.16, (also see 1.2.19) each partial map
$\phi^{i,j}_{n,m} : A^i_n \to \phi^{i,j}_{n,m}(\mathbf{1}_{A^i_n})A^j_m\phi^{i,j}_{n,m}(\mathbf{1}_{A^i_n})$ can be written as $\phi' \otimes \mathbf{1}_{[n,k]}$ for some homomorphism $\phi' : C(X_{n,i}) \to EA^j_mE$, where $E = \phi^{i,j}_{n,m}(e_{11})$, and e_{11} is the canonical matrix unit corresponding to the upper left corner. The map ϕ' also has the property sdp$(\frac{\eta}{32}, \delta)$.
Applying Theorem 4.35 to ${F'}^i \subset C(X_{n,i})$, η, δ, and ϕ' (as the above) and using 1.2.19, one can obtain the following Theorem.

Theorem 4.37. *For any A_n, finite set $F = \bigoplus_{i=1}^{t_n} F^i \subset A_n$, positive integer J, and number $\varepsilon > 0$, there are an A_m, mutually orthogonal projections $Q_0, Q_1, Q_2 \in A_m$ with $Q_0 + Q_1 + Q_2 = \phi_{n,m}(\mathbf{1}_{A_n})$, a unital map $\psi_0 \in$ Map$(A_n, Q_0A_mQ_0)_1$, and unital homomorphisms $\psi_1 \in$ Hom$(A_n, Q_1A_mQ_1)_1$, $\psi_2 \in$ Hom$(A_n, Q_2A_mQ_2)_1$, such that*
(1) $\|\phi_{n,m}(f) - \psi_0(f) \oplus \psi_1(f) \oplus \psi_2(f)\| < \varepsilon$ for all $f \in F$;
(2) The set $(\psi_0 \oplus \psi_1)(F)$ is weakly approximately constant to within ε;
(3) The homomorphism ψ_2 factors through $\bigoplus_{i=1}^{t_n} M_{[n,i]}(C[0,1])$ as

$$\psi_2 : A_n \xrightarrow{\xi_1} \bigoplus_{i=1}^{t_n} M_{[n,i]}(C[0,1]) \xrightarrow{\xi_2} Q_2A_mQ_2,$$

and ξ_2 satisfies the following condition: if $X_{m,j} \neq \{pt\}$, then $\xi_2^{i,j} : M_{[n,i]}(C[0,1]) \to A^j_m$ is injective;
(4) Each partial map $\psi_0^{i,j} : A^i_n \to Q_0^{i,j}A^j_mQ_0^{i,j}$ (where $Q_0^{i,j} = \psi_0^{i,j}(\mathbf{1}_{A^i_n})$) is of the form $\psi'_0 \otimes id_{[n,k]}$ with $\psi'_0 : C(X_{n,i}) \to q_0A^j_mq_0$ (where $q_0 = \psi_0^{i,j}(e_{11})$ is a projection). Each partial map $\psi_1^{i,j} : A^i_n \to Q_1^{i,j}A^j_mQ_1^{i,j}$ (where $Q_1^{i,j} = \psi_1^{i,j}(\mathbf{1}_{A^i_n})$) is of the form $\psi'_1 \otimes id_{[n,k]}$ and $\psi'_1 : A^i_n \to p^{i,j}A^j_mp^{i,j}$ (where $p^{i,j} = \psi_1^{i,j}(e_{11})$), satisfies the following

$$\psi'_1(f) = \sum_{i=1}^{n} f(x_i)p_i$$

for any $f \in C(X_{n,i})$, where $p_1, \cdots, p_n$ are mutually orthogonal projections with $p^{i,j} = p_1 + \cdots + p_n$, and with $J \cdot [q_0] \le [p_s]$ $(s = 1, 2 \cdots n)$ and $\{x_1, x_2, \cdots x_n\} \subset X_{n,i}$ is an ε-dense subset in $X_{n,i}$.
(When we apply this theorem in Section 6, $Q_0 + Q_1$ will be chosen to be a trivial projection.)

Definition 4.38. Let $A = PM_l(C(X))P$, and L be a positive integer and $\eta > 0$. A homomorphism $\lambda :\ A \to B = QM_{l_1}(C(Y))Q$ is said to be defined by point evaluations of size at least L at an η-dense subset if there are mutually orthogonal projections $Q_1, Q_2, \cdots, Q_n$ with rank$(Q_i) \ge L$, an η-dense subset $\{x_1, x_2, \cdots, x_n\} \subset X$, and unital homomorphisms $\lambda_i :\ A \to Q_iBQ_i$, $i = 1, 2, \cdots, n$ such that
(1) $\lambda(\mathbf{1}) = \sum_{i=1}^n Q_i$, and $\lambda = \bigoplus_{i=1}^n \lambda_i$;

(2) The homomorphisms λ_i factor through $P(x_i)M_l(\mathbb{C})P(x_i)$ ($\cong M_{\mathrm{rank}(P)}(\mathbb{C})$) as

$$\lambda_i = \lambda_i' \circ e_{x_i} : PM_l(C(X))P \xrightarrow{e_{x_i}} P(x_i)M_l(\mathbb{C})P(x_i) \xrightarrow{\lambda_i'} Q_iBQ_i,$$

where e_{x_i} are evaluation maps defined by $e_{x_i}(f) = f(x_i)$ and $\lambda_i' \in \mathrm{Hom}(M_{\mathrm{rank}(P)}(\mathbb{C}), Q_iBQ_i)_1$.

We will also call the above homomorphism λ to have the PROPERTY PE$(L,\ \eta)$. (PE stands for point evaluation.)

A homomorphism $\lambda : A \to B = QM_{l_1}(C(Y))Q$ is said TO CONTAIN A PART OF POINT EVALUATION AT POINT x OF SIZE AT LEAST L, if $\lambda = \lambda_1 \oplus \lambda'$, where λ_1 factor through $P(x)M_l(\mathbb{C})P(x)$ as

$$\lambda_1 = \lambda_1' \circ e_x : PM_l(C(X))P \xrightarrow{e_x} P(x)M_l(\mathbb{C})P(x) \xrightarrow{\lambda_1'} Q_1BQ_1,$$

and λ_1' is a unital homomorphism with $\mathrm{rank}(Q_1) \geq L$.

The following result is a corollary of Theorem 4.37, and will also be used in the proof of our main reduction theorem.

Corollary 4.39. *For any A_n, finite set $F = \bigoplus_{i=1}^{t_n} F^i \subset A_n$, positive integer J, any numbers $\varepsilon > 0$ and $\eta > 0$, and any projection $P = \oplus P^i \in \oplus A_n^i$, there are an A_m, mutually orthogonal projections $Q_0, Q_1, Q_2 \in A_m$ with $Q_0 + Q_1 + Q_2 = \phi_{n,m}(\mathbf{1}_{A_n})$, a unital map $\psi_0 \in Map(A_n, Q_0A_mQ_0)_1$, and unital homomorphisms $\psi_1 \in Hom(A_n, Q_1A_mQ_1)_1$, $\psi_2 \in Hom(A_n, Q_2A_mQ_2)_1$, such that*

Part I:

(1) $\|\phi_{n,m}(f) - \psi_0(f) \oplus \psi_1(f) \oplus \psi_2(f)\| < \varepsilon$ for all $f \in F$;

(2) The homomorphism ψ_2 factors through a direct sum of matrix algebras over $C[0,1]$ as

$$\psi_2 : A_n \xrightarrow{\xi_1} \bigoplus_{i=1}^{t_n} M_{[n,i]}(C[0,1]) \xrightarrow{\xi_2} Q_2A_mQ_2,$$

and ξ_2 satisfies the condition that, if $X_{m,j} \neq \{pt\}$, then $\xi_2^{i,j} : M_{[n,i]}(C[0,1]) \to A_m^j$ is injective.

(3) For any blocks $A_n^i \subset A_n$, $A_m^j \subset A_m$, and for the partial maps $\psi_0^{i,j}$ and $\psi_1^{i,j}$, we have that $\psi_0^{i,j}(\mathbf{1}_{A_n^i}) := Q_0^{i,j}$ is a projection and $\psi_1^{i,j}$ has the property PE$(J \cdot rank(Q_0^{i,j}),\ \eta)$.

(4) The set $(\psi_0 \oplus \psi_1)(F)$ is weakly approximately constant to within ε.

Part II:

$\psi_0^{i,j}(P^i)$ and $\psi_0^{i,j}(\mathbf{1}_{A_n^i} - P^i)$ are mutually orthogonal projections, and the decomposition of $\phi_{n,m}' := \phi_{n,m}|_{PA_nP}$ as the direct sum of $\psi_0' := \psi_0|_{PA_nP}$, $\psi_1' := \psi_1|_{PA_nP}$, and $\psi_2' := \psi_2|_{PA_nP}$ satisfies the following conditions:

(1) $\|\phi_{n,m}'(f) - \psi_0'(f) \oplus \psi_1'(f) \oplus \psi_2'(f)\| < \varepsilon$ for all $f \in PFP = \oplus P^iF^iP^i$;

(2) The homomorphism ψ_2' factors through a C^*-algebra C which is a direct sum of matrix algebras over $C[0,1]$ as

$$\psi_2' : PA_nP \xrightarrow{\xi_1'} C \xrightarrow{\xi_2'} Q_2'A_mQ_2',$$

and ξ_2' satisfies the following condition, if $X_{m,j} \neq \{pt\}$, then $\xi_2^{i,j} : C^i \to A_m^j$ is injective, where $Q_2' = \psi_2(P)$.
(3) For any blocks $A_n^i \subset A_n$, $A_m^j \subset A_m$, and for the partial maps $\psi_0^{'i,j}$ and $\psi_1^{'i,j}$, we have that $\psi_0^{'i,j}(P^i) := Q_0^{'i,j}$ is a projection and $\psi_1^{'i,j}$ has property PE($J \cdot rank(Q_0^{'i,j})$, η).

Proof: Obviously, the first part of the corollary follows from Theorem 4.37. To prove the second part, we only need to perturb $\psi_0 \in \mathrm{Map}(A, Q_0BQ_0)_1$ to $new\psi_0$ such that the restriction $new\psi_0|_D$ is a homomorphism, where

$$D := \bigoplus_i \mathbb{C} \cdot P^i \oplus \bigoplus_i \mathbb{C} \cdot (\mathbf{1}_{A_n^i} - P^i)$$

is a finite dimensional subalgebra of A_n.
By Lemma 1.6.8, such perturbation exists if ψ_0 is sufficiently multiplicative, which is automatically true if the set F is large enough and the number ε is small enough, using the next lemma.
(Note that $C := \xi_1(P)(\bigoplus_{i=1}^{t_n} M_{[n,i]}(C[0,1]))\xi_1(P)$ is still a direct sum of matrix algebras over $C[0,1]$, since all the projections in $M_\bullet(C[0,1])$ are trivial.)

$\square$

Lemma 4.40. *Let A be a unital C^*-algebra. Suppose that $G \subset A$ is a finite set containing $\mathbf{1}_A$, and $G_1 = G \times G := \{gh \mid g \in G, h \in G\}$. Suppose that $\delta > 0$, and $\delta' = \frac{1}{3}\frac{1}{\|G\|}\delta$, where $\|G\| = \max_{g\in G}\{\|g\|\}$.*
Suppose that B is a unital C^*-algebra and $p \in B$ is a projection. If a homomorphism $\phi \in Hom(A,B)$ and two maps $\phi_1 \in Map(A, pBp), \phi_2 \in Map(A, (1-p)B(1-p))$ satisfy

$$\|\phi(g) - \phi_1(g) \oplus \phi_2(g)\| < \delta', \quad \forall g \in G_1,$$

then both ϕ_1 and ϕ_2 are G-δ multiplicative.

Proof: The proof is straight forward, we omit it.
Theorem 4.37 and Corollary 4.39 will be used in the proof of our Main Reduction Theorem in this article. Theorem 4.35 will be used in the proof of the Uniqueness Theorem in [EGL]. The rest of this section will not be used in this paper. They are important to [EGL].

4.41. In the rest of this section, we will compare the decompositions of two different homomorphisms. Such comparison will be used in the proof of the Uniqueness Theorem in [EGL].
Let $X, \eta, \delta, \phi, \{\beta_i\}_{i=1}^{L}$, and $\Theta(y)$ be as in 4.34. (Take $J = 1$.) Suppose that $\phi : C(X) \to M_k(C(Y))$ is as in Theorem 4.35, and $\psi : C(X) \to M_k(C(Y))$ is another homomorphism with $\phi(\mathbf{1}) = \psi(\mathbf{1})$. If

$$\|\mathrm{Aff}T\phi(f) - \mathrm{Aff}T\psi(f)\| < \frac{\delta}{4}$$

for all $f \in H(\eta, \delta, X)$, then by Lemma 4.29, $\mathrm{SP}\phi_y$ and $\mathrm{SP}\psi_y$ can be paired within $\frac{\eta}{2}$. Since $\mathrm{SP}\phi_y$ and $\Theta(y)$ can be paired within $\frac{\eta}{2}$, $\mathrm{SP}\psi_y$ and $\Theta(y)$ can be paired within η. Similar to 4.34, by Theorem 4.1, there are mutually orthogonal projections $q_1, q_2, \cdots, q_n, q_{n+1}, \cdots, q_L$ and $q_0 = \psi(\mathbf{1}) - \sum_{i=1}^{L} q_i$ such that
(1) For all $y \in Y$ and $f \in F$,

$$\|\psi(f)(y) - q_0\psi(f)(y)q_0 \oplus \sum_{i=1}^{L} f(\alpha{\circ}\beta_i(y))q_i\| < \varepsilon.$$

(2) $\mathrm{rank}(q_0) + 2\dim(Y) \le \mathrm{rank}(q_i)$.
As Remark 4.22, we can choose projections p_i for ϕ and q_i for ψ to be trivial projections with $\mathrm{rank}(p_i) = \mathrm{rank}(q_i)$. (Note that, in 4.34, the number L_2 and $L_2 + L_1$, which serve as T_i, $i = 1, 2, \cdots, L$ (i.e., $T_i = L_2$, for $1 \le i \le L-1$, and $T_L = L_2 + L_1$) in Theorem 4.1, are very larger.) Therefore, there is a unitary $u \in M_k(C(Y))$ such that

$$uq_iu^* = p_i, \qquad\qquad i = 1, 2, \cdots, L.$$

Let $\tilde{\psi} = \mathrm{Ad}u{\circ}\psi$. Then

$$\|\tilde{\psi}(f)(y) - p_0\tilde{\psi}(f)(y)p_0 \oplus \sum_{i=1}^{L} f(\alpha{\circ}\beta_i(y))p_i\| < \varepsilon$$

for all $y \in Y$ and $f \in F$.
Note that the above decomposition has the same form as that of ϕ, even with the same projections p_i and the part $\sum_{i=1}^{L} f(\alpha{\circ}\beta_i(y))p_i$. Also, in the part $\sum_{i=1}^{L} f(\alpha{\circ}\beta_i(y))p_i$, there is a map defined by point evaluations:

$$\phi'(f) = \sum_{i=1}^{n} f(x_i)p_i,$$

with $\{x_1, x_2, \cdots, x_n\}$ η-dense in X, and $\mathrm{rank}(p_i) \ge \mathrm{rank}(p_0) + 2\dim(Y)$. This means that two different homomorphisms which are close at the level of AffT can be decomposed in the same way. This result will be useful in the proof of the Uniqueness Theorem for certain spaces X with $K_1(C(X))$ a torsion group. We summarize what we obtained as the following proposition which will be used in the proof of the Uniqueness Theorem for certain spaces X with $K_1(C(X))$ a torsion group.

Proposition 4.42. *Let X be a connected simplicial complex, $\varepsilon > 0$, and $F \subset C(X)$ be a finite set.*
Suppose that $\eta \in (0, \varepsilon)$ satisfies that if dist$(x, x') < 2\eta$, then $|f(x) - f(x')| < \frac{\varepsilon}{3}$ for all $f \in F$.
For any $\delta > 0$, there is an integer $L > 0$ and a finite set $H \subset AffT(C(X))(= C(X))$ such that the following holds.

If $\phi, \psi:\ C(X) \to M_k(C(Y))$ *are homomorphisms with properties*
(a) ϕ *has* $sdp(\frac{\eta}{32}, \delta)$*;*
(b) $rank(\phi(\mathbf{1})) \geq 2L^2 \cdot 2^L(\dim X + \dim Y + 1)^3$*;*
(c) $\phi(\mathbf{1}) = \psi(\mathbf{1})$ *and*

$$\|AffT\phi(h) - AffT\psi(h)\| < \frac{\delta}{4},\ \forall h \in H,$$

then there are two orthogonal projections $Q_0, Q_1 \in M_k(C(Y))$*, two maps* $\phi_0, \psi_0 \in Map(C(X), Q_0M_k(C(Y))Q_0)_1$*, a homomorphism* $\phi_1 \in Hom(C(X), Q_1M_k(C(Y))Q_1)_1$*, and a unitary* $u \in M_k(C(Y))$ *such that*
(1) $\phi(\mathbf{1}) = \psi(\mathbf{1}) = Q_0 + Q_1$*;*
(2) $\|\phi(f) - \phi_0(f) \oplus \phi_1(f)\| < \varepsilon$*, and* $\|(Adu \circ \psi)(f) - \psi_0(f) \oplus \phi_1(f)\| < \varepsilon$ *for all* $f \in F$*;*
(3) ϕ_1 *factors through* $C[0,1]$*.*
(4) $Q_0 = p_0 + p_1 + \cdots + p_n$ *with* $rank(p_0) + 2\dim(Y) \leq rank(p_i)\ \ (i = 1, 2 \cdots n)$*, and* ϕ_0 *and* ψ_0 *are defined by*

$$\phi_0(f) = p_0\phi(f)p_0 + \sum_{i=1}^{n} f(x_i)p_i,\ \forall f \in C(X),$$

$$\psi_0(f) = p_0(Adu \circ \psi)(f)p_0 + \sum_{i=1}^{n} f(x_i)p_i,\ \forall f \in C(X),$$

where $p_0, p_1, \cdots p_n$ *are mutually orthogonal projections and* $\{x_1, x_2, \cdots x_n\} \subset X$ *is an* ε*-dense subset in* X*.*
(Comparing with Theorem 4.35, the maps ϕ_0 and ϕ_1 in 4.35 have been put together to form the map ϕ_0 in the above proposition.)
(In [EGL], we will prove that the above ϕ_0 and ψ_0 are approximately unitarily equivalent to each other to within some small number (under the condition $KK(\phi) = KK(\psi)$), then so also are ϕ and ψ.)

4.43. The above proposition is not strong enough to prove the Uniqueness Theorem for homomorphisms from $C(S^1)$ to $M_k(C(Y))$, since $K_1(C(S^1))$ is infinite. Before we conclude this section, we introduce a result which can be used to deal with this case (i.e, the case S^1).
We will discuss briefly what the problem is, and how to solve the problem.
Suppose that ϕ and ψ are two homomorphisms from $C(S^1)$ to another C^*-algebra, For ϕ and ψ to be approximately unitarily equivalent to each other, they should agree not only on Aff$T(C(S^1))$ and $K_*(C(S^1))$, but also on the determinant functions. That is, $\phi(z)\psi(z)^*$ should have only a small variation in the determinant, where $z \in C(S^1)$ is the standard generator. (All these things will be made precise in [EGL].) This idea has appeared in [Ell2] and [NT].
Roughly speaking, if ϕ and ψ agree (approximately) to within ε at the level of the determinant (this will also be made precise in [EGL]), then the maps $p_0\phi p_0$

and $p_0(\mathrm{Ad}u \circ \psi)p_0$ from Proposition 4.42 agree only to within $\frac{\mathrm{rank}(\phi(\mathbf{1}))}{\mathrm{rank}(p_0)}\varepsilon$ at the level of the determinant. So for the decomposition to be useful for the proof of the uniqueness theorem, $\mathrm{rank}(p_0)$ should not be too small compared with $\mathrm{rank}(\phi(\mathbf{1}))$. (This will be the property (2) of Theorem 4.45 below.) On the other hand, in the decompositions of ϕ and $\mathrm{Ad}u \circ \psi$, we also need the homomorphism defined by point evaluations, which by Proposition 4.42 is the same for both of these decompositions, to be large in order to absorb the parts $p_0\phi p_0$ and $p_0(\mathrm{Ad}u \circ \psi)p_0$. This will be the property (3) of Theorem 4.45. Therefore, $\mathrm{rank}(p_0)$ should not be too large either.

To do that, besides the property $\mathrm{sdp}(\frac{\eta}{32}, \delta)$, we also need sdp property for an extra pair $(\frac{\tilde{\eta}}{32}, \tilde{\delta})$, where $\tilde{\eta}$ depends on δ. For those readers who are familiar with [Ell2] and [NT], we encourage them to compare the sdp property for the two pairs $(\frac{\eta}{32}, \delta)$ and $(\frac{\tilde{\eta}}{32}, \tilde{\delta})$, with the conditions of Theorem 4 of [Ell2], and Lemma 2.3 and Theorem 2.4 of [NT], in the following way. In Theorem 4 of [Ell2] (see page 100 of [Ell2]), roughly speaking, the sentence on lines 16–20 corresponds to our property $\mathrm{sdp}(\frac{\tilde{\eta}}{32}, \tilde{\delta})$, and the sentence on lines 21–22 corresponds to our property $\mathrm{sdp}(\frac{\eta}{32}, \delta)$. That is, $\frac{1}{m}$ corresponds to our $\frac{\eta}{32}$ (or $\frac{\eta}{16}$ in some sense), $\frac{3}{n}$ corresponds to our δ, $\frac{1}{n}$ corresponds to our $\frac{\tilde{\eta}}{32}$, and δ corresponds to our $\tilde{\delta}$. Similarly, in Lemma 2.3 of [NT], condition (1) corresponds to our $\mathrm{sdp}(\frac{\eta}{32}, \delta)$ and condition (2) corresponds to our $\mathrm{sdp}(\frac{\tilde{\eta}}{32}, \tilde{\delta})$. Also in Theorem 2.4 of [NT], condition (2) corresponds to our $\mathrm{sdp}(\frac{\eta}{32}, \delta)$ and condition (3) corresponds to our $\mathrm{sdp}(\frac{\tilde{\eta}}{32}, \tilde{\delta})$.

Such a construction will be given in 4.44 below.

In 4.44, we will first describe the condition that ϕ should satisfy. Then we will carry out the construction in three steps.

In Step 1, we will follow the procedure in 4.34, to decompose ϕ into $p_0\phi p_0 \oplus \phi_1$, corresponding to the property $\mathrm{sdp}(\frac{\tilde{\eta}}{32}, \tilde{\delta})$ (not $\mathrm{sdp}(\frac{\eta}{32}, \delta)$). (Here, the map ϕ_1 is $\phi_1 \oplus \phi_2$ in the notation of 4.34 or 4.35.)

In Step 2, we will take a part $p'\phi_1 p'$ out of ϕ_1 and add it to $p_0\phi p_0$ to obtain $P_0\phi P_0$, where $P_0 = p_0 + p'$. The rest of ϕ_1 will be defined to be newϕ_1. In this way, we can get the projection P_0 with suitable size (neither too small nor too large). The size depends on δ, which explains why $\tilde{\eta}$ depends on δ.

In Step 3. we will prove that newϕ_1 can be decomposed again in such a way that the point evaluation part of its decomposition is sufficiently large that it can be used to control $P_0\phi P_0$, in the proof of the uniqueness theorem in [EGL]. (See the property (3) of Theorem 4.45.) The property $\mathrm{sdp}(\frac{\eta}{32}, \delta)$ is used in this step.

4.44. Let $F \subset C(X)$ be a finite set, $\varepsilon > 0$ and $\varepsilon_1 > 0$. Suppose that the positive number $\eta < \frac{\varepsilon_1}{4}$ satisfies the condition that, if $\mathrm{dist}(x, x') < 2\eta$, then

$$\|f(x) - f(x')\| < \frac{\varepsilon}{3}.$$

For any $\delta > 0$, consider the pair (η, δ) as in 4.33. Let N, n be as in 4.33. Instead of choosing $L = n(N+2)$, we choose

$$L \geq \max\{n(N+2), \tfrac{8}{\delta}, \tfrac{4}{\varepsilon}, \tfrac{4}{\varepsilon_1}\}.$$

Consider $\tilde{\varepsilon} = \frac{1}{8L} < \min(\varepsilon, \varepsilon_1)$. Let positive number $\tilde{\eta} < \frac{\eta}{4}$ satisfy that, if $\mathrm{dist}(x, x') < 2\tilde{\eta}$, then

$$\|f(x) - f(x')\| < \frac{\tilde{\varepsilon}}{3}.$$

Let $\tilde{\delta} > 0$ be any number. Then for the pair $(\tilde{\eta}, \tilde{\delta})$, there exists an integer $\tilde{L}$ playing the role of L as in Lemma 4.33. We can assume $\tilde{L} > L$. Let

$$\Lambda = 6\tilde{L}^2 \cdot 2^{\tilde{L}}(\dim X + M + 1)^3,$$

where M is a positive integer.
Now let Y be a simplicial complex with $\dim Y \leq M$, and $\phi : C(X) \to PM_k(C(Y))P$ be a unital homomorphism satisfying the following two conditions:
(a) ϕ has both $\mathrm{sdp}(\frac{\eta}{32}, \delta)$ and $\mathrm{sdp}(\frac{\tilde{\eta}}{32}, \tilde{\delta})$;
(b) $\mathrm{rank}(P) \geq \Lambda$.
We will construct a decomposition for ϕ.
Step 1. By the discussion in 4.34 corresponding to $\mathrm{sdp}(\frac{\tilde{\eta}}{32}, \tilde{\delta})$, there is a set

$$\Theta(y) = \left\{\alpha_\circ\beta_1(y)^{\sim L_2}, \alpha_\circ\beta_2(y)^{\sim L_2}, \cdots, \alpha_\circ\beta_{\tilde{L}-1}(y)^{\sim L_2}, \alpha_\circ\beta_{\tilde{L}}(y)^{\sim L_2+L_1}\right\},$$

where

$$L_2 = \mathrm{int}\left(\frac{\mathrm{rank}(P)}{\tilde{L}}\right) \geq \mathrm{int}\left(\frac{\Lambda}{\tilde{L}}\right),$$

such that $\mathrm{SP}\phi_y$ and $\Theta(y)$ can be paired within $\frac{\tilde{\eta}}{2}$.
As in 4.34, there are mutually orthogonal projections p_0 and $P_1 = \sum_{i=1}^{\tilde{L}} p_i$ and a homomorphism $\phi_1 : C(X) \to P_1M_k(C(Y))P_1$, such that
(1) $\|\phi(f) - p_0\phi(f)p_0 \oplus \phi_1(f)\| \leq \tilde{\varepsilon} \leq \frac{1}{8L}$,
(2) $\mathrm{rank}(p_0) \leq \tilde{L} \cdot 2^{\tilde{L}}(\dim X + M + 1)^3 \leq \mathrm{int}\left(\frac{\Lambda}{6\tilde{L}}\right)$,
where ϕ_1 is defined by

$$\phi_1(f)(y) = \sum_{i=1}^{\tilde{L}} f(\alpha_\circ\beta_i(y))p_i$$

with $\mathrm{rank}(p_i) \geq L_2 - 2^{\tilde{L}}(\dim X + M + 1)^3$.
Step 2. We will take a part $p'\phi_1p'$ out from ϕ_1 and add it into $p_0\phi p_0$, such that the projection $P_0 = p_0 + p'$ has rank about $\frac{\mathrm{rank}(P)}{L}$, which is neither too large nor too small. (Here we use L not $\tilde{L}$.)
There exists a projection p' satisfying the following two conditions.
(c) $p' = \sum_{i=1}^{\tilde{L}} p'_i$, with $p'_i < p_i$, $i = 1, 2, \cdots, \tilde{L}$.

(d) $\operatorname{rank}(p') = \operatorname{int}\left(\frac{\operatorname{rank}(P)}{L}\right)$ (here we use L, not $\tilde{L}$), where L was chosen in the beginning of this subsection.
We can make the above (d) hold for the following reason. First,

$$\operatorname{rank}(\sum_{i=1}^{\tilde{L}} p_i) \geq \operatorname{rank}(P) - \operatorname{int}\left(\frac{\Lambda}{6\tilde{L}}\right) > \operatorname{int}\left(\frac{\operatorname{rank}(P)}{L}\right) + \tilde{L}\dim(Y).$$

So one can choose non negative integers $k_1, k_2, \cdots, k_{\tilde{L}}$ such that $\sum_{i=1}^{\tilde{L}} k_i = \operatorname{int}\left(\frac{\operatorname{rank}(P)}{L}\right)$ and that $k_i \leq rank(p_i) - \dim(Y)$. Therefore, by [Hu], we can choose trivial projections $p_i' < p_i$ with $\operatorname{rank}(p_i') = k_i$.
Define

$$P_0 = p_0 \oplus p' \qquad \text{and} \qquad \text{new}P_1 = P_1 \ominus p'.$$

Note that p' is a sub-projection of $P_1 = \sum_{i=1}^{\tilde{L}} p_i$. Define $\text{new}\phi_1 : C(X) \to \text{new}P_1 M_k(C(Y))\text{new}P_1$ by

$$(\text{new}\phi_1(f))(y) = \sum_{i=1}^{\tilde{L}} f(\alpha \circ \beta_i(y))(p_i \ominus p_i').$$

$\text{new}P_1$ and $\text{new}\phi_1$ are still denoted by P_1 and ϕ_1, respectively. Evidently, the following are true.

$$(1') \qquad \|\phi(f) - P_0\phi(f)P_0 \oplus \phi_1(f)\| < \frac{1}{4L}.$$

$$(2') \qquad \frac{\operatorname{rank}(P)}{L} \leq \operatorname{rank}(P_0) \leq 2 \cdot \operatorname{int}\left(\frac{\operatorname{rank}(P)}{L}\right).$$

(Notice that, to get the above decomposition, one only needs the condition that $\mathrm{SP}\phi_y$ and $\Theta(y)$ can be paired within $\tilde{\eta}$ (see the way η is chosen in 4.4 for Theorem 4.1 and the way $\tilde{\eta}$ is chosen above). On the other hand, $\mathrm{SP}\phi_y$ and $\Theta(y)$ can be paired within $\frac{\tilde{\eta}}{2}$ in our case. So if ψ satisfies the condition that $\mathrm{SP}\psi_y$ and $\mathrm{SP}\phi_y$ can be paired within $\frac{\tilde{\eta}}{2}$, then the above decomposition also holds for ψ, as discussed in 4.41. In particular, for a certain unitary u, $\mathrm{Ad}u \circ \psi$ can have same form of decomposition as ϕ does— same projection P_0 and even exactly the same part of the above ϕ_1. This will be used in 4.46 and Proposition 4.47.)
Step 3. Now, we can decompose ϕ_1 again to obtain a large part of the homomorphism defined by point evaluations, which will be used to absorb the part of $P_0\phi P_0$, in the proof of the uniqueness theorem in [EGL].
For the compact metric space X, and $\eta > 0$ (now we use η not $\tilde{\eta}$), there exists a finite η-dense subset $\{x_1, x_2, \cdots, x_m\}$ such that $\operatorname{dist}(x_i, x_j) \geq \eta$, if $i \neq j$.

(Such set could be chosen to be a maximum set of finite many points which have mutual distance at least η. Then the η-density of the set follows from the maximality.)
We will prove the following claim.
Claim: There are mutually orthogonal projections $q_1, q_2, \cdots, q_m < P_1$ with $\mathrm{rank}(q_i) > \mathrm{rank}(P_0) + \dim(Y)$, such that

$$\left\| \phi_1(f) - (P_1 - \sum_{i=1}^{m} q_i)\phi_1(f)(P_1 - \sum_{i=1}^{m} q_i) \oplus \sum_{i=1}^{m} f(x_i)q_i \right\| < \varepsilon$$

for all $f \in F$.
Proof of the claim:
First, we know that the set $(\mathrm{SP}\phi_1)_y$ is obtained by deleting $\mathrm{rank}(P_0)$ points (counting multiplicity) from the set $\Theta(y)$. Also $\mathrm{SP}\phi_y$ and $\Theta(y)$ can be paired within $\frac{\tilde{\eta}}{2}$. Recall that,

$$\Theta(y) = \left\{\alpha_\circ\beta_1(y)^{\sim L_2}, \alpha_\circ\beta_2(y)^{\sim L_2}, \cdots, \alpha_\circ\beta_{\tilde{L}-1}(y)^{\sim L_2}, \alpha_\circ\beta_{\tilde{L}}(y)^{\sim L_2+L_1}\right\}$$

is the set corresponding to ϕ and the pair $(\tilde{\eta}, \tilde{\delta})$ in 4.33. And recall that $L_2 = \mathrm{int}\left(\frac{\mathrm{rank}(P)}{\tilde{L}}\right)$. From (a), ϕ has the property $\mathrm{sdp}(\frac{\eta}{32}, \delta)$. So $\Theta(y)$ has the property $\mathrm{sdp}(\frac{\eta}{32} + \frac{\tilde{\eta}}{2}, \delta)$. But $(\mathrm{SP}\phi_1)_y$ is obtained by deleting

$$\mathrm{rank}(P_0) \left(\leq 2 \cdot \mathrm{int}\left(\frac{\mathrm{rank}(P)}{L}\right) \leq \frac{\delta}{4}\mathrm{rank}(P) \right)$$

points from $\Theta(y)$. (Note that $\frac{1}{L} < \frac{\delta}{8}$.) Therefore, in the $(\frac{\eta}{32} + \frac{\tilde{\eta}}{2})$-ball of any point in X, $(\mathrm{SP}\phi_1)_y$ contains at least

$$\delta \cdot \mathrm{rank}(P) - \frac{\delta}{4}\mathrm{rank}(P) = \frac{3\delta}{4}\mathrm{rank}(P)$$

points (counting multiplicity). That is, ϕ_1 has the property $\mathrm{sdp}(\frac{\eta}{32} + \frac{\tilde{\eta}}{2}, \frac{3\delta}{4})$. Therefore ϕ_1 has the property $\mathrm{sdp}(\frac{\eta}{4}, \frac{3\delta}{4})$, since $\tilde{\eta} < \frac{\eta}{4}$.
Set $U_i = B_{\frac{\eta}{2}}(x_i)$, $i = 1, 2, \cdots, m$. Then U_i, $i = 1, 2, \cdots, m$ are mutually disjoint open sets, since $\mathrm{dist}(x_i, x_j) \geq \eta$, if $i \neq j$. By the property $\mathrm{sdp}(\frac{\eta}{4}, \frac{3\delta}{4})$ of ϕ_1, for any $y \in Y$,

$$\#(\mathrm{SP}(\phi_1)_y \cap U_i) \geq \frac{3\delta}{4}\mathrm{rank}(P) > \frac{2}{L}\mathrm{rank}(P) + 3\dim Y > \mathrm{rank}(P_0) + 3\dim(Y).$$

The claim follows from the following proposition:
Proposition. *Let X be a simplicial complex, and $F \subset C(X)$ a finite subset. Let $\varepsilon > 0$ and $\eta > 0$ be such that if $\mathrm{dist}(x, x') < 2\eta$, then $|f(x) - f(x')| < \frac{\varepsilon}{3}$ for any $f \in F$.*
Suppose that $U_1, U_2, \cdots, U_m$ are disjoint open neighborhoods of points $x_1, x_2, \cdots, x_m \in X$, respectively, such that $U_i \subset B_\eta(x_i)$ for all $1 \leq i \leq m$.

Suppose that $\phi : C(X) \to PM_\bullet(C(Y))P$ *is a unital homomorphism, where* Y *is a simplicial complex, such that*

$$\#(SP\phi_y \cap U_i) \geq k_i, \qquad \text{for } 1 \leq i \leq m \text{ and for all } y \in Y.$$

Then there are mutually orthogonal projections $q_1, q_2, \cdots, q_m \in PM_\bullet(C(Y))P$ *with rank*$(q_i) \geq k_i -$ *dim*(Y) *such that*

$$\|\phi(f) - p_0\phi(f)p_0 \oplus \sum_{i=1}^{m} f(x_i)q_i\| < \varepsilon, \qquad \text{for all } f \in F,$$

where $p_0 = P - \sum q_i$.
This is Proposition 1.5.7 of this paper (see 1.5.4—1.5.6 for the proof). Since the expert reader may skip §1.5, we point out that the above result was essentially proved in [EG2, Theorem 2.21].
So we obtain the projections q_i with

$$\mathrm{rank}(q_i) \geq \min_y(\#(\mathrm{SP}(\phi_1)_y \cap U_i)) - \dim(Y) \geq \mathrm{rank}(p_0) + 2\dim(Y).$$

Summarizing the above, we obtain the following theorem.

Theorem 4.45. *Let* $F \subset C(X)$ *be a finite set,* $\varepsilon > 0$, $\varepsilon_1 > 0$, *and let* M *be a positive integer (in the application in [EGL], we will let* $M = 3$*). Let the positive number* $\eta < \frac{\varepsilon_1}{4}$ *satisfy that, if dist*$(x, x') < 2\eta$, *then*

$$\|f(x) - f(x')\| < \frac{\varepsilon}{3} \qquad \text{for all } f \in F.$$

Let $\delta > 0$ *be any positive number. There is an integer* $L > \max\{\frac{8}{\delta}, \frac{4}{\varepsilon}, \frac{4}{\varepsilon_1}\}$ *satisfying the following condition. The rest of the theorem describes this condition. Suppose that* $\tilde{\eta} > 0$ *satisfies that, if dist*$(x, x') < 2\tilde{\eta}$, *then*

$$\|f(x) - f(x')\| < \frac{1}{24L} \qquad \text{for all } f \in F.$$

For any $\tilde{\delta} > 0$, *there is a positive integer* Λ *such that if a unital homomorphism* $\phi : C(X) \to PM_k(C(Y))P$ *(with* $\dim Y \leq M$*) satisfies the following conditions*
(a) ϕ *has the properties sdp*$(\frac{\eta}{32}, \delta)$ *and sdp*$(\frac{\tilde{\eta}}{32}, \tilde{\delta})$;
(b) rank$(P) \geq \Lambda$,
then there are projections $P_0, P_1 \in PM_k(C(Y))P$ *(with* $P_0 + P_1 = P$*) and a homomorphism* $\phi_1 : C(X) \to P_1M_k(C(Y))P_1$ *such that*
(1) $\|\phi(f) - P_0\phi(f)P_0 \oplus \phi_1(f)\| < \frac{1}{4L}$ *for all* $f \in F$;
(2) rank$(P_0) \geq \frac{rank(P)}{L}$;
(3) There are mutually orthogonal projections $q_1, q_2, \cdots, q_m \in P_1M_k(C(Y))P_1$ *and an* η*-dense finite subset* $\{x_1, x_2, \cdots, x_m\} \subset X$ *with the following properties.*

(i) $rank(q_i) > rank(P_0) + 2\dim(Y)$, $i = 1, 2, \cdots, m$;
(ii) $\|\phi_1(f) - (P_1 - \sum_{i=1}^m q_i)\phi_1(f)(P_1 - \sum_{i=1}^m q_i) \oplus \sum_{i=1}^m f(x_i)q_i\| < \varepsilon$ *for all* $f \in F$.

4.46. Let $\tilde{\eta}$ and $\tilde{\delta}$ be as in 4.44 (or 4.45), and $H(\tilde{\eta}, \tilde{\delta}, X) \subset C(X)$ the subset defined in 4.27. Suppose that $\phi : C(X) \to PM_k(C(Y))P$ satisfies the conditions (a) and (b) in Theorem 4.45. And suppose that $\psi : C(X) \to PM_k(C(Y))P$ is another homomorphism satisfying

$$\|\mathrm{Aff}T\phi(h) - \mathrm{Aff}T\psi(h)\| < \frac{\tilde{\delta}}{4},$$

for all $h \in H(\tilde{\eta}, \tilde{\delta}, X)$. Similar to 4.41, there is a unitary $u \in PM_k(C(Y))P$ such that

$$\|\mathrm{Ad}u{\circ}\psi(f) - P_0\mathrm{Ad}u{\circ}\psi(f)P_0 \oplus \phi_1\| < \frac{1}{4L}, \quad \forall f \in F$$

where P_0 and ϕ_1 are exactly the same as those for ϕ in Theorem 4.45. (See the end of step 2 of 4.44.)
So we have the following proposition.

Proposition 4.47. *Let* $F \subset C(X)$ *be a finite set,* $\varepsilon > 0$, $\varepsilon_1 > 0$, *and let* M *be a positive integer (in the application in [EGL], we will let* $M = 3$*). Let the positive number* $\eta < \frac{\varepsilon_1}{4}$ *satisfy that, if* $dist(x, x') < 2\eta$, *then*

$$\|f(x) - f(x')\| < \frac{\varepsilon}{3} \qquad \text{for all } f \in F.$$

Let $\delta > 0$ *be any positive number. There is an integer* $L > \max\{\frac{8}{\delta}, \frac{4}{\varepsilon}, \frac{4}{\varepsilon_1}\}$ *satisfying the following condition. The rest of the proposition describes this condition.*
Suppose that $\tilde{\eta} > 0$ *satisfies that, if* $dist(x, x') < 2\tilde{\eta}$, *then*

$$\|f(x) - f(x')\| < \frac{1}{24L} \qquad \text{for all } f \in F.$$

For any $\tilde{\delta} > 0$, *there is a positive integer* Λ *and a finite set* $H \subset AffT(C(X))(= C(X))$ *such that if unital homomorphisms* $\phi, \psi : C(X) \to PM_k(C(Y))P$ *(with* $\dim Y \leq M$*) satisfy the following conditions:*
(a) ϕ *has the properties* $sdp(\frac{\eta}{32}, \delta)$ *and* $sdp(\frac{\tilde{\eta}}{32}, \tilde{\delta})$;
(b) $rank(P) \geq \Lambda$;
(c) $\|AffT\phi(h) - AffT\psi(h)\| < \frac{\tilde{\delta}}{4}$, $\forall h \in H$,
then there are projections $P_0, P_1 \in PM_k(C(Y))P$ *(with* $P_0 + P_1 = P$*), a homomorphism* $\phi_1 : C(X) \to P_1M_k(C(Y))P_1$ *factoring through* $C[0,1]$, *and a unitary* $u \in PM_k(C(Y))P$ *such that*
(1) $\|\phi(f) - P_0\phi(f)P_0 \oplus \phi_1(f)\| < \frac{1}{4L}$ *and*
$\|(Adu \circ \psi)(f) - P_0(Adu \circ \psi)(f)P_0 \oplus \phi_1(f)\| < \frac{1}{4L}$ *for all* $f \in F$;

(2) rank$(P_0) \geq \frac{rank(P)}{L}$;
(3) There are mutually orthogonal projections $q_1, q_2, \cdots, q_m \in P_1M_k(C(Y))P_1$ *and an* η*-dense finite subset* $\{x_1, x_2, \cdots, x_m\} \subset X$ *with the following properties.*
(i) rank$(q_i) >$ *rank*$(P_0) + 2\dim(Y)$;
(ii) $\|\phi_1(f) - (P_1 - \sum_{i=1}^m q_i)\phi_1(f)(P_1 - \sum_{i=1}^m q_i) \oplus \sum_{i=1}^m f(x_i)q_i\| < \varepsilon$ *for all* $f \in F$.
In order to be consistent in notation with the application in [EGL], let us rewrite the above proposition in the following form.

PROPOSITION 4.47'. *For any finite set* $F \subset C(X)$, $\varepsilon > 0$, $\varepsilon_1 > 0$, *there is a number* $\eta > 0$ *with the property described below.*
For any $\delta > 0$, *there are an integer* $K > \frac{4}{\varepsilon}$ *and a a number* $\tilde{\eta} > 0$ *satisfying the following condition.*
For any $\tilde{\delta} > 0$, *there is a positive integer* L *and a finite set* $H \subset AffT(C(X))$*(*H *can be chosen to be* $H(\tilde{\eta}, \tilde{\delta}, X)$ *in 4.27) such that if two unital homomorphisms* $\phi, \psi : C(X) \to PM_k(C(Y))P$ *(with* $\dim Y \leq 3$*) satisfy the following conditions:*
(a) ϕ *has the properties sdp*$(\frac{\eta}{32}, \delta)$ *and sdp*$(\frac{\tilde{\eta}}{32}, \tilde{\delta})$;
(b) rank$(P) \geq L$;
(c) $\|AffT\phi(h) - AffT\psi(h)\| < \frac{\tilde{\delta}}{4}$, $\forall h \in H$,
then there are projections $P_0, P_1 \in PM_k(C(Y))P$ *(with* $P_0 + P_1 = P$*), a homomorphism* $\phi_1 : C(X) \to P_1M_k(C(Y))P_1$ *factoring through* $C[0,1]$, *and a unitary* $u \in PM_k(C(Y))P$ *such that*
(1) $\|\phi(f) - P_0\phi(f)P_0 \oplus \phi_1(f)\| < \frac{1}{4K}$ *and*
$\|(Adu \circ \psi)(f) - P_0(Adu \circ \psi)(f)P_0 \oplus \phi_1(f)\| < \frac{1}{4K}$ *for all* $f \in F$;
(2) rank$(P_0) \geq \frac{rank(P)}{K}$;
(3) There are mutually orthogonal projections $q_1, q_2, \cdots, q_m \in P_1M_k(C(Y))P_1$ *and an* $\frac{\varepsilon_1}{4}$*-dense finite subset* $\{x_1, x_2, \cdots, x_m\} \subset X$ *with the following properties.*
(i) rank$(q_i) >$ *rank*$(P_0) + 2\dim(Y)$;
(ii) $\|\phi_1(f) - (P_1 - \sum_{i=1}^m q_i)\phi_1(f)(P_1 - \sum_{i=1}^m q_i) \oplus \sum_{i=1}^m f(x_i)q_i\| < \varepsilon$ *for all* $f \in F$.
(Notice that in the above statement, we change the notation L and Λ to K and L respectively. Also, in condition (3), we change η-density to $\frac{\varepsilon_1}{4}$-density.)

4.48. Proposition 4.47' will be used in the proof of the Uniqueness Theorem in [EGL]. Namely, we will prove that, under certain conditions about KK(ϕ) and KK(ψ) and the determinants of $\phi(z)$ and $\psi(z)$ (see (4) of Theorem 2.4 of [NT]), where $z \in C(S^1)$ is the standard generator,

$$P_0\phi(f)P_0 \oplus \sum_{i=1}^m f(x_i)q_i, \qquad f \in F$$

is approximately unitarily equivalent to

$$P_0 \mathrm{Ad} u_\circ \psi(f) P_0 \oplus \sum_{i=1}^{m} f(x_i) q_i, \qquad f \in F.$$

Therefore, $\{\phi(f),\ f \in F\}$ is approximately unitarily equivalent to $\{\psi(f),\ f \in F\}$. In [EGL], we need both of the following conditions:

$$\mathrm{rank}(P_0) \geq \frac{\mathrm{rank}(P)}{L} \qquad \text{and} \qquad [q_i] > [P_0] \ \text{ in } \ K_0(C(Y)).$$

In comparison with Theorem 2.4 of [NT], in the Uniqueness Theorem in [EGL], we also have a condition similar to (4) of Theorem 2.4 of [NT]. But this condition will be useful only when it is combined with the condition (2) above (see [EGL] for details).

5 Almost Multiplicative Maps

In this section, we study almost multiplicative maps

$$\phi \in \mathrm{Map}\left(M_l(C(X)), M_{l_1}(C(Y))\right),$$

where $X = T_{II,k}, T_{III,k}$, or S^2, and Y is a simplicial complex of dimension at most M with M a fixed number. In this section, all the simplicial complexes are assumed to have dimension at most M.

5.1. Suppose that $B_1, B_2, \cdots, B_n, \cdots$ are unital C^*-algebras. Let $B = \bigoplus_{n=1}^{+\infty} B_n$. Then the multiplier algebra $M(B)$ of B is $\prod_{n=1}^{+\infty} B_n$. The Six Term Exact Sequence associated to

$$0 \longrightarrow B \longrightarrow M(B) \longrightarrow M(B)/B \longrightarrow 0$$

breaks into two exact sequences

$$0 \longrightarrow K_0(B) \longrightarrow K_0(M(B)) \longrightarrow K_0(M(B)/B) \longrightarrow 0 \qquad \text{and}$$

$$0 \longrightarrow K_1(B) \longrightarrow K_1(M(B)) \longrightarrow K_1(M(B))/B \longrightarrow 0.$$

since each projection (or unitary) in $M_n(M(B)/B)$ can be lifted to a projection (or a unitary) in $M_n(M(B))$.
Furthermore,

$$K_0(B) = \bigoplus_{n=1}^{+\infty} K_0(B_n) \qquad \text{and} \qquad K_1(B) = \bigoplus_{n=1}^{+\infty} K_1(B_n).$$

But in general, it is NOT true that

$$K_0(M(B)) = \prod_{n=1}^{+\infty} K_0(B_n) \qquad \text{or} \qquad K_1(M(B)) = \prod_{n=1}^{+\infty} K_1(B_n).$$

In fact, $K_0(M(B))$ is a subgroup of $\prod_{n=1}^{+\infty} K_0(B_n)$. But $K_1(M(B))$ is more complicated. In the first part of this section, we will calculate the K-theory of $M(B)$ (and of $M(B)/B$) for the case

$$B_n = M_{k_n}(C(X_n)),$$

where X_n are simplicial complexes of dimension at most M. For convenience, we always suppose that the spaces X_n are connected.

5.2. Consider $S^1 = \{z;\ |z| = 1\} \subset \mathbb{C}$. Let

$$F:\ (S^1\backslash\{-1\}) \times [0,1] \longrightarrow S^1\backslash\{-1\}$$

be defined by

$$F(e^{i\theta}, t) = e^{it\theta}, \qquad\qquad -\pi < \theta < \pi.$$

Then $|t-t'| < \varepsilon$ implies

$$|F(x,t) - F(x,t')| < \pi\varepsilon.$$

This fact implies the following. If u and v are unitaries such that $\|u - v\| < 1$, then there is a path of unitaries u_t with $u_0 = u,\ u_1 = v$ such that $|t - t'| < \varepsilon$ implies $\|u_t - u_{t'}\| < \pi\varepsilon$.
Let $SU(n)(\subset U(n))$ denote the collection of $n \times n$ unitaries with determinant 1. Let $SU_n(X)$ denote the collection of continuous functions from X to $SU(n)$. Note $SU_n(X) \subset U_n(X) \subset M_n(C(X))$.
From the proof of Theorem 3.3 (and Lemma 3.1) of [Phi2] (in particular $(***)$ in Step 4 of 3.3 of [Phi2]), one can prove the following useful fact.

Lemma 5.3. ([Phi2]) *For each positive integer M, there is an $M' > 0$ satisfying the following condition. For any connected finite CW-complex X of dimension at most M, and $u, v \in SU_n(X)$, if u and v can be connected to each other in $U_n(X)$, then there is a path $u_t \in SU_n(X)$ such that*
1. $u_0 = u, u_1 = v$ and
2. $|t-t'| < \varepsilon$ implies $\|u_t - u_{t'}\| < M' \cdot \varepsilon$.
(Note that M' does not depend on n, the size of the unitaries.)

5.4. Let $B_n = M_{k_n}(C(X_n)),\ \dim(X_n) \le M$. Let $B = \bigoplus_{n=1}^{+\infty} B_n$. Then we can describe $K_0(M(B))$ as below. Let $(K_0(B_n), K_0(B_n)^+, \mathbf{1}_{B_n})$ be the scaled ordered K-group of B_n (see 1.2 of [EG2]). Let $\Pi_b K_0(B_n)$ be the subgroup of $\prod_{n=1}^{+\infty} K_0(B_n)$ consisting of elements

$$(x_1, x_2, \cdots, x_n, \cdots) \in \prod_{n=1}^{+\infty} K_0(B_n)$$

with the property that there is a positive integer L such that

$$-L[\mathbf{1}_{B_n}] < x_n < L[\mathbf{1}_{B_n}] \ \in\ K_0(B_n)$$

for all n.

Lemma 5.5. $K_0\left(\prod_{n=1}^{+\infty} B_n\right) = \Pi_b K_0(B_n)$.

Proof: Any element in $K_0\left(\prod_{n=1}^{+\infty} B_n\right)$ is of the form $[p]-[q]$, where $p, q \in M_L\left(\prod_{n=1}^{+\infty} B_n\right)$ are projections. Let

$$p = (p_1, p_2, \cdots, p_n, \cdots),\ q = (q_1, q_2, \cdots, q_n, \cdots) \in M_L\left(\prod_{n=1}^{+\infty} B_n\right).$$

Then $[p]-[q] \in K_0\left(\prod_{n=1}^{+\infty} B_n\right)$ corresponds to the element

$$([p_1]-[q_1], [p_2]-[q_2], \cdots, [p_n]-[q_n], \cdots) \in \Pi_b K_0(B_n).$$

We will prove that this correspondence is bijective.
Surjectivity: Let

$$([p_1]-[q_1], [p_2]-[q_2], \cdots, [p_n]-[q_n], \cdots) \in \Pi_b K_0(B_n).$$

Then there is an $L > M$ such that

$$-L[\mathbf{1}_{B_n}] < [p_n]-[q_n] < L[\mathbf{1}_{B_n}], \quad \forall n.$$

Therefore,

$$-L\cdot k_n \leq \mathrm{rank}(p_n) - \mathrm{rank}(q_n) < L \cdot k_n, \quad \forall n.$$

It is well known that (see [Hu]) any vector bundle of dimension $M+T$ over an M dimensional space has a T dimensional trivial sub-bundle. Thus one can replace p_n by p'_n, q_n by q'_n, with properties

$$[p'_n] < 2L[\mathbf{1}_{B_n}], \quad [q'_n] < 2L[\mathbf{1}_{B_n}] \qquad \text{and}$$
$$[p'_n]-[q'_n] = [p_n]-[q_n] \qquad \text{in } K_0(B_n).$$

$([p'_1]-[q'_1], [p'_2]-[q'_2], \cdots)$ is in the image of the correspondence, since every element $[p'_n] < 2L[\mathbf{1}_{B_n}]$ can be realized by a projection in $M_{4L}(B_n)$ (recall that $L > M$).
Injectivity. Let $p = (p_1, p_2, \cdots, p_n, \cdots)$ and $q = (q_1, q_2, \cdots, q_n, \cdots)$ be projections in $M_L\left(\prod_{n=1}^{+\infty} B_n\right)$. Suppose that for each n, $[p_n] = [q_n] \in K_0(B_n)$. We have to prove that $[(p_1, p_2, \cdots, p_n, \cdots)] = [(q_1, q_2, \cdots, q_n, \cdots)] \in K_0\left(\prod_{n=1}^{+\infty} B_n\right)$.
Without loss of generality, assume that $L > M$. Let $\mathbf{1}_n \in M_L(B_n)$ be the unit. By [Hu], for each n, the projection $p_n \oplus \mathbf{1}_n$ is unitary equivalent to $q_n \oplus \mathbf{1}_n$. That is, there is a unitary $u_n \in M_{2L}(B_n)$ such that $q_n \oplus \mathbf{1}_n = u_n(p_n \oplus \mathbf{1}_n)u_n^*$. Hence the unitary $u = (u_1, u_2, \cdots, u_n, \cdots) \in M_{2L}(\prod_{n=1}^{+\infty} B_n)$ satisfies $q \oplus \mathbf{1} = u(p \oplus \mathbf{1})u^*$. It follows that $[q] = [p]$.

□

5.6. Let X be a finite CW complex. Then $K_1(C(X)) = K^1(X)$ is defined to be the collection of homotopy equivalence classes of continuous maps from X to $U(\infty)$, denoted by $[X, U(\infty)]$, (or from X to $U(n)$, denoted by $[X, U(n)]$, for n large enough). Consider the fibration

$$SU(n) \longrightarrow U(n) \xrightarrow{b} S^1,$$

where $S^1 \subset \mathbb{C}$ is the unit circle, b is defined by sending a unitary to its determinant, and $SU(n)$ is the special unitary group consisting the unitaries of determinant 1. The fibration has a splitting $S^1 \xrightarrow{b^{-1}} U(n)$, defined by

$$S^1 \ni z \xmapsto{b^{-1}} \begin{pmatrix} z & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \in U(n).$$

One can identify $U(n)$ with $SU(n) \times S^1$ by $U(n) \ni u \mapsto ((b^{-1} \circ b(u))^* u, b(u)) \in SU(n) \times S^1$.
Therefore, $[X, U(n)] = [X, SU(n)] \oplus [X, S^1]$ as a group. We use notation $SK_1(C(X))$ or $SK^1(X)$ to denote $[X, SU(n)]$, n large enough, and $\pi^1(X)$ to denote $[X, S^1]$. Then

$$K_1(C(X)) = SK_1(C(X)) \oplus \pi^1(X).$$

(The splitting is not a natural splitting.)

5.7. Let $\{X_n\}$ be a sequence of connected finite CW complexes of dimension at most M. Let $B_n = M_{k_n}(C(X_n))$. Define a map

$$\tau: \ K_1\left(\prod_{n=1}^{+\infty} B_n\right) \longrightarrow \prod_{n=1}^{+\infty} K_1 B_n \qquad \text{by}$$

$$\tau[(u_1, u_2, \cdots, u_n, \cdots)] = ([u_1], [u_2], \cdots, [u_n], \cdots),$$

where $(u_1, u_2, \cdots, u_n, \cdots)$ is a unitary in $M_L\left(\prod_{n=1}^{+\infty} B_n\right)$. If $L \geq M$, then any element in $K_1(B_n)$ can be realized by a unitary in $M_L(B_n)$. Based on this fact, we know that τ is surjective. We will prove that

$$0 \longrightarrow Ker\tau \longrightarrow K_1\left(\prod_{n=1}^{+\infty} B_n\right) \longrightarrow \prod_{n=1}^{+\infty} K_1 B_n \longrightarrow 0$$

is a splitting exact sequence. A splitting

$$\tilde{\tau}: \prod_{n=1}^{+\infty} K_1 B_n \longrightarrow K_1\left(\prod_{n=1}^{+\infty} B_n\right)$$

will be defined such that $\tau \circ \tilde{\tau} = \mathrm{id}$ on $\prod_{n=1}^{+\infty} K_1 B_n$. By 5.6,

$$K_1 B_n = SK_1 B_n \oplus \pi^1(X_n).$$

Hence we define $\tilde{\tau}$ on $\prod_{n=1}^{+\infty} SK_1 B_n$ and $\prod_{n=1}^{+\infty} \pi^1(X_n)$ separately.
If $x \in \prod_{n=1}^{+\infty} SK_1 B_n$ is represented by a sequence of unitaries

$$u_1 \in M_L(B_1),\ u_2 \in M_L(B_2),\ \cdots,\ u_n \in M_L(B_n),\ \cdots,$$

each with determinant 1, then define

$$\tilde{\tau}(x) = [(u_1, u_2, \cdots, u_n, \cdots)] \in K_1\left(\prod_{n=1}^{+\infty} B_n\right).$$

To see that $\tilde{\tau}$ is well defined, let $v_1, v_2, \cdots, v_n, \cdots$ be another sequence with determinant 1 and

$$[u_n] = [v_n] \qquad\qquad \text{in } \ K_1(B_n).$$

Without loss of generality, we assume that $L > M$. By Lemma 5.3, for each n, there is a unitary path u_n such that $u_n(0) = u_n$, $u_n(1) = v_n$, and $\|u_n(t) - u_n(t')\| < M' \cdot |t - t'|$, $\forall t \in [0,1]$, where M' is a constant which does not depend on n. Obviously,

$$(u_1(t), u_2(t), \cdots, u_n(t), \cdots) \in (M_L(\prod_{n=1}^{+\infty} B_n)) \otimes C([0,1]).$$

Hence

$$[(u_1, u_2, \cdots, u_n, \cdots)] = [(v_1, v_2, \cdots, v_n, \cdots)].$$

That is, the above map is well defined.
(Warning: It is not enough to prove that each u_n can be connected to v_n, since a sequence of paths, each connecting u_n and v_n $(n = 1, 2, \cdots)$, only defines an element in $M_L(\prod_{n=1}^{+\infty}(B_n \otimes C[0,1]))$, but

$$\left(\prod_{n=1}^{+\infty} B_n\right) \otimes C[0,1] \ \subsetneq\ \prod_{n=1}^{+\infty}(B_n \otimes C[0,1]).\ \ )$$

The following claim is a well known folklore result in topology. Since we can not find a precise reference, we present a proof here.
Claim: For any connected simplicial complex X, the cohomotopy group $\pi^1(X)$ is a finitely generated free abelian group.
Proof of the claim. Let $X^{(1)}$ be the 1-skeleton of X. Then $X^{(1)}$ is homotopy equivalent to a finite wedge of S^1. Evidently, $\pi^1(X^{(1)})$ is a finitely generated free abelian group. (In comparison with the above cohomotopy group, we point out that the fundamental group $\pi_1(X^{(1)})$ of a finite wedge $X^{(1)}$ of S^1 is a free group (not a free abelian group).)

On the other hand, we can prove that

$$i^* : \pi^1(X) \to \pi^1(X^{(1)}),$$

induced by the inclusion $i : X^{(1)} \to X$, is an injective map as below. Once this is done, the claim follows from the result in group theory that any subgroup of a free abelian group is still a free abelian group.

Let us prove the injectivity of i^*. Suppose that $f,\ g :\ X \to S^1$ are two maps satisfying that

$$i^*([f]) = i^*([g]),$$

where $[f],[g] \in \pi^1(X)$ are elements represented by f and g, respectively. Then $f|_{X^{(1)}}$ is homotopic to $g|_{X^{(1)}}$. Let $F :\ X^{(1)} \times [0,1] \to S^1$ be a homotopy path connecting $f|_{X^{(1)}}$ and $g|_{X^{(1)}}$. That is

$$F|_{X^{(1)}\times\{0\}} = f|_{X^{(1)}} \quad \text{and} \quad F|_{X^{(1)}\times\{1\}} = g|_{X^{(1)}}.$$

We are going to extend the homotopy F to a homotopy on the entire space $X \times [0,1]$. The construction is done by induction. Suppose that F has been extended to a homotopy (let us still denote it by F) $F :\ X^{(n)} \times [0,1] \to S^1$ between $f|_{X^{(n)}}$ and $g|_{X^{(n)}}$ on the n-skeleton (where $n \geq 1$) of X. I.e., $F|_{X^{(n)}\times\{0\}} = f|_{X^{(n)}}$ and $F|_{X^{(n)}\times\{1\}} = g|_{X^{(n)}}$. We need to prove that it can be extended to a homotopy on the $(n+1)$-skeleton. Let Δ be any $(n+1)$-simplex. Then $\partial\Delta \subset X^{(n)}$. Let $G : \partial\Delta \times [0,1] \cup \Delta \times \{0\} \cup \Delta \times \{1\} \to S^1$ be defined by

$$G(x) = \begin{cases} F(x) & \text{if } x \in \partial\Delta \times [0,1] \\ f(x) & \text{if } x \in \Delta \times \{0\} \\ g(x) & \text{if } x \in \Delta \times \{1\}\ . \end{cases}$$

Then $G(x)$ is a continuous map from $\partial(\Delta \times [0,1])$ to S^1. Since $\pi_{n+1}(S^1) = 0$ and $\partial(\Delta \times [0,1]) = S^{n+1}$, G can be extended to a map $G : \Delta \times [0,1] \to S^1$. Define F on each simplex Δ to be this G. Then F is the desired extension. This ends the proof of the claim.

Let us go back to the construction of $\tilde{\tau}$ on $\prod_{n=1}^{+\infty} \pi^1(X_n)$. Let $x_1^0 \in X_1,\ x_2^0 \in X_2,\ \cdots, x_n^0 \in X_n,\ \cdots$, be chosen as the base points of the spaces. Let

$$\theta_{n,1}, \theta_{n,2}, \cdots,\ \theta_{n,t_n} :\ X_n \longrightarrow S^1$$

be the functions representing the generators

$$[\theta_{n,1}],\ [\theta_{n,2}],\ \cdots,\ [\theta_{n,t_n}]\ \in\ \pi^1(X_n).$$

Suppose that

$$\theta_{n,j}(x_n^0) = 1 \in S^1 \subset \mathbb{C}, \qquad j = 1, 2, \cdots, t_n.$$

For any element $(x_1, x_2, \cdots, x_n, \cdots) \in \prod_{n=1}^{+\infty} \pi^1(X_n)$, define $\tilde{\tau}(x)$ as below. Let $u_n \in B_n$ be defined by

$$u_n(y) = \begin{pmatrix} \theta_{n,1}(y)^{m_1} \cdot \theta_{n,2}(y)^{m_2} \cdots \theta_{n,t_n}(y)^{m_{t_n}} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}_{k_n \times k_n} \in M_{k_n}(\mathbb{C})$$

for each $y \in X_n$, where $m_1, m_2, \cdots, m_{t_n}$ are integers with

$$(*) \qquad x_n = m_1[\theta_{n,1}] + m_2[\theta_{n,2}] + \cdots + m_{t_n}[\theta_{n,t_n}] \in \pi^1(X_n).$$

Define

$$\tilde{\tau}(x) = [(u_1, u_2, \cdots, u_n, \cdots)] \in K_1\left(\prod_{n=1}^{+\infty} B_n\right).$$

Since each $\pi^1(X_n)$ is a free abelian group, the expression $(*)$ for x_n is unique. It is easy to check that $\tilde{\tau}$ on $\prod_{n=1}^{+\infty} \pi^1(X_n)$ is a well defined group homomorphism. It is straight forward to check that

$$\tau \circ \tilde{\tau} = \mathrm{id} : \prod_{n=1}^{+\infty} \pi^1(X_n) \longrightarrow \prod_{n=1}^{+\infty} \pi^1(X_n)$$

And that

$$\tau \circ \tilde{\tau} = \mathrm{id} : \prod_{n=1}^{+\infty} SK_1(B_n) \longrightarrow \prod_{n=1}^{+\infty} SK_1(B_n).$$

That is,

$$\tau \circ \tilde{\tau} = \mathrm{id} : \prod_{n=1}^{+\infty} K_1(B_n) \longrightarrow \prod_{n=1}^{+\infty} K_1(B_n).$$

The splitting $\tilde{\tau} : \prod_{n=1}^{+\infty} K_1(B_n) \longrightarrow K_1\left(\prod_{n=1}^{+\infty} B_n\right)$ of the exact sequence

$$0 \longrightarrow Ker(\tau) \longrightarrow K_1\left(\prod_{n=1}^{+\infty} B_n\right) \xrightarrow{\tau} \prod_{n=1}^{+\infty} K_1 B_n \longrightarrow 0$$

gives an isomorphism

$$K_1\left(\prod_{n=1}^{+\infty} B_n\right) = \prod_{n=1}^{+\infty} K_1 B_n \oplus Ker(\tau).$$

5.8. In order to identify $Ker(\tau)$, suppose that

$$u = [(u_1, u_2, \cdots, u_n, \cdots)] \in K_1\left(\prod_{n=1}^{+\infty} B_n\right)$$

satisfies that

$$\tau(u) = 0 \in \prod_{n=1}^{+\infty} K_1 B_n.$$

Note that any unitary matrix $v \in M_\bullet(\mathbb{C})$ can be connected to $\mathbf{1} \in M_\bullet(\mathbb{C})$, by a path $v(t)$ satisfying that, if $|t - t'| < \varepsilon$, then

$$\|v(t) - v(t')\| < 2\pi\varepsilon.$$

Based on this fact, we have

$$[(u_1, u_2, \cdots)] = [(u_1^*(x_1^0)u_1, u_2^*(x_2^0)u_2, \cdots)] \in K_1\left(\prod_{n=1}^{+\infty} B_n\right).$$

Therefore, without loss of generality, we assume that

$$u_n(x_n^0) = \mathbf{1} \in M_L(B_n),$$

where $x_n^0 \in X_n$ are the base points.
Since $\tau(u) = 0$, if we assume $L \geq M$, then each u_n can be connected to $\mathbf{1} \in M_L(B_n)$. This implies that the map

$$\text{determinant}(u_n) : \ X_n \longrightarrow S^1$$

is homotopy trivial. Therefore, this map can be lifted to a unique map

$$\det(u_n) : \ X_n \longrightarrow \mathbb{R}$$

such that $\det(u_n)(x_n^0) = 0 \in \mathbb{R}$ and

$$exp(2\pi i \det(u_n)) = \text{determinant}(u_n).$$

Let $\coprod X_n$ be the disjoint union of X_n and $\text{Map}\left(\coprod X_n, \mathbb{R}\right)_0$ the set of all continuous maps $f : \coprod X_n \to \mathbb{R}$ with $f(x_n^0) = 0$ for all x_n^0. Let $\text{Map}_b\left(\coprod X_n, \mathbb{R}\right)_0$ be the set of those maps with bounded images.
Define a map $d : Ker(\tau) \to \dfrac{\text{Map}\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}{\text{Map}_b\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}$ by

$$d(u) = [(\frac{\det(u_1)}{k_1}, \frac{\det(u_2)}{k_2}, \cdots, \frac{\det(u_n)}{k_n}, \cdots)].$$

We will prove that d is a well defined isomorphism.
Suppose that u can be represented by another unitary

$$(v_1, v_2, \cdots, v_n, \cdots) \in M_L\left(\prod_{n=1}^{+\infty} B_n\right)$$

with $v(x_n^0) = \mathbf{1} \in M_L(B_n)$. Then for the unit of a certain matrix algebra over $\prod_{n=1}^{+\infty} B_n$

$$\mathbf{1}_{L_1} \in M_{L_1}\left(\prod_{n=1}^{+\infty} B_n\right),$$

we have that, the element

$$(u_1 \oplus \mathbf{1}, u_2 \oplus \mathbf{1}, \cdots, u_n \oplus \mathbf{1}, \cdots) \in M_{L+L_1}\left(\prod_{n=1}^{+\infty} B_n\right)$$

can be connected to the element

$$(v_1 \oplus \mathbf{1}, v_2 \oplus \mathbf{1}, \cdots, v_n \oplus \mathbf{1}, \cdots) \in M_{L+L_1}\left(\prod_{n=1}^{+\infty} B_n\right)$$

by a unitary path

$$(u_1(t), u_2(t), \cdots, u_n(t), \cdots) \in \left(M_{L+L_1}\left(\prod_{n=1}^{+\infty} B_n\right)\right) \otimes C[0,1].$$

We need to prove that

$$\left(\frac{\det(u_1) - \det(v_1)}{k_1}, \frac{\det(u_2) - \det(v_2)}{k_2}, \cdots, \frac{\det(u_n) - \det(v_n)}{k_n}, \cdots\right)$$

has a uniformly bounded image in $\mathbb{R}$. This follows from the following fact. If two unitaries $w_1, w_2 \in M_{(L+L_1)k_n}(\mathbb{C})$ satisfying $|w_1 - w_2| < \varepsilon < \frac{1}{4}$, then

$$|\text{determinant}(w_1^* w_2) - 1| < \pi(L + L_1)k_n\varepsilon.$$

Now, we have to prove that d is an isomorphism.
Obviously, d is surjective. In fact, for any function

$$(f_1, f_2, \cdots, f_n, \cdots) \in \text{Map}\left(\coprod_{n=1}^{+\infty} \mathbb{R}\right)_0,$$

let $u \in K_1\left(\prod_{n=1}^{+\infty} B_n\right)$ be the element represented by

$$(exp(2\pi i f_1), exp(2\pi i f_2), \cdots, exp(2\pi i f_n), \cdots) \in \prod_{n=1}^{+\infty} B_n = \prod_{n=1}^{+\infty} M_{k_n}(C(X_n)).$$

Then $d(u) = [(f_1, f_2, \cdots, f_n, \cdots)]$.
Finally, we have to prove that d is injective. Suppose that $u \in Ker(\tau)$ is represented by

$$(u_1, u_2, \cdots, u_n, \cdots) \in M_L\left(\prod_{n=1}^{+\infty} B_n\right)$$

satisfying

$$\left(\frac{\det(u_1)}{k_1}, \frac{\det(u_2)}{k_2}, \cdots, \frac{\det(u_n)}{k_n}, \cdots\right) \in \mathrm{Map}_b\left(\coprod X_n, \mathbb{R}\right)_0.$$

Let $f_n = \frac{\det(u_n)}{k_n} : X_n \to \mathbb{R}$ and

$$v_n = exp\frac{2\pi i f_n}{L} \in M_{Lk_n}(C(X_n)).$$

Then $v_n^* u_n \in SU_{Lk_n}(X_n)$, i.e., it has determinant 1 every where. Since $(f_1, f_2, \cdots, f_n, \cdots)$ is of uniformly bounded image, we know that

$$(u_1, u_2, \cdots, u_n, \cdots) \qquad \text{and} \qquad (v_1^* u_1, v_2^* u_2, \cdots v_n^* u_n, \cdots)$$

can be connected by a continuous path in

$$\left(M_L\left(\prod_{n=1}^{+\infty} B_n\right)\right) \otimes C[0,1].$$

Therefore, $u = [(v_1^* u_1, v_2^* u_2, \cdots v_n^* u_n, \cdots)]$. The latter is zero by Lemma 5.3 and the fact that $u \in Ker(\tau)$.
Summarizing the above, we obtain

Lemma 5.9. $K_1\left(\prod_{n=1}^{+\infty} B_n\right) = \prod_{n=1}^{+\infty} K_1 B_n \oplus \frac{Map\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}{Map_b\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}$

$$= \prod_{n=1}^{+\infty} SK_1 B_n \oplus \prod_{n=1}^{+\infty} \pi^1(X_n) \oplus \frac{Map\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}{Map_b\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}.$$

Corollary 5.10.

$$K_1\left(\prod_{n=1}^{+\infty} B_n \Big/ \bigoplus_{n=1}^{+\infty} B_n\right) = \left(\prod_{n=1}^{+\infty} K_1 B_n \Big/ \bigoplus_{n=1}^{+\infty} K_1 B_n\right) \oplus \frac{\mathrm{Map}\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}{\mathrm{Map}_b\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}.$$

5.11. From [Sch], for any C^*-algebra A in the bootstrap class and any C^*-algebra B (not necessarily separable), there is a splitting short exact sequence

$$0 \longrightarrow K_*(A) \otimes K_*(B) \longrightarrow K_*(A \otimes B) \longrightarrow Tor(K_*(A), K_*(B)) \longrightarrow 0.$$

Let $A = C_0(W_k)$, where $W_k = T_{II,k}$ as in the introduction. W_k is used for $T_{II,k}$ only when involving mod k K-theory $K_*(B, \mathbb{Z}/k)$. From the definition

$$K_*(B, \mathbb{Z}/k) := K_*(A \otimes B),$$

one has

$$0 \longrightarrow K_0(B)\otimes \mathbb{Z}/k \longrightarrow K_0(B,\mathbb{Z}/k) \longrightarrow Tor(\mathbb{Z}/k, K_1(B)) \longrightarrow 0$$

and

$$0 \longrightarrow K_1(B)\otimes \mathbb{Z}/k \longrightarrow K_1(B,\mathbb{Z}/k) \longrightarrow Tor(\mathbb{Z}/k, K_0(B)) \longrightarrow 0.$$

Since $G\otimes \mathbb{Z}/k$ can be identified with the cokernel of

$$G \xrightarrow{\times k} G,$$

and $Tor(\mathbb{Z}/k, G)$ can be identified with the kernel of

$$G \xrightarrow{\times k} G,$$

one has the following well known exact sequences

$$K_0(B) \xrightarrow{\times k} K_0(B) \longrightarrow K_0(B,\mathbb{Z}/k) \longrightarrow K_1(B) \xrightarrow{\times k} K_1(B)$$

and

$$K_1(B) \xrightarrow{\times k} K_1(B) \longrightarrow K_1(B,\mathbb{Z}/k) \longrightarrow K_0(B) \xrightarrow{\times k} K_0(B).$$

5.12. Let $\{X_n\}_{n=1}^{+\infty}$, $B_n = M_{k_n}(C(X_n))$ be as above, and $B = \bigoplus_{n=1}^{+\infty} B_n$, $M(B) = \prod_{n=1}^{+\infty} B_n$, $Q(B) = M(B)/B$.
From Lemma 5.5 and Corollary 5.10, we have

$$K_0(Q(B)) = \Pi_b K_0(B_n)/\oplus K_0(B_n) \qquad \text{and}$$

$$K_1(Q(B)) = \left\{\frac{\prod_{n=1}^{+\infty} K_1(B_n)}{\bigoplus_{n=1}^{+\infty} K_1(B_n)}\right\} \bigoplus \left\{\frac{\mathrm{Map}\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}{\mathrm{Map}_b\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}\right\}.$$

It is easy to see that the map

$$\left\{\frac{\mathrm{Map}\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}{\mathrm{Map}_b\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}\right\} \xrightarrow{\times k} \left\{\frac{\mathrm{Map}\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}{\mathrm{Map}_b\left(\coprod_{n=1}^{+\infty} X_n, \mathbb{R}\right)_0}\right\}$$

is an isomorphism.
Any torsion element $x_n \in K_0(B_n)$ can be realized as a formal difference of two projections $p, q \in M_\infty(B_n)$ of the same rank. (The rank of a projection makes sense since X_n are connected. Also for any element $x \in K_0(B_n)$ represented by $[p]-[q]$, we define $\mathrm{rank}(x) = \mathrm{rank}(p) - \mathrm{rank}(q)$, which is always a (possibly negative) integer.) By [Hu], if a projection $p \in M_\bullet(C(X_n))$ has rank larger than $\dim(X_n) + r$, then p has a trivial sub projection of rank r. Therefore, any torsion element $x_n \in K_0(B_n)$ can be realized as a formal difference of two

projections $p, q \in M_L(B_n)$ if $L > \dim(X_n)$. Based on this fact, one can directly compute that

$$\text{Kernel}\left(\Pi_b K_0(B_n) \xrightarrow{\times k} \Pi_b K_0(B_n)\right) = \text{Kernel}\left(\prod_{n=1}^{+\infty} K_0(B_n) \xrightarrow{\times k} \prod_{n=1}^{+\infty} K_0(B_n)\right).$$

Fixed a positive integer k. let $x = (x_1, x_2, \cdots, x_n, \cdots) \in \prod_{n=1}^{+\infty} K_0(B_n)$. For each element $x_n \in K_0(B_n)$, one can write $\text{rank}(x_n) = k \cdot M \cdot l_n + r_n$, where l_n is a (possibly negative) integer, M is the maximum of $\{\dim(X_n)\}_n$, and $0 < k \cdot M \leq r_n < 2k \cdot M$. Let p_n be the trivial rank one projection in B_n. Then x_n can be written as $k \cdot M \cdot l_n[p_n] + [q_n]$, where q_n is a projection of rank r_n. Therefore, x can can be written as $x = x' + x''$, where $x' \in k(\prod_{n=1}^{+\infty} K_0(B_n))$ and $x'' \in \Pi_b K_0(B_n)$. As a consequence, one can compute that

$$\begin{aligned}&\text{Cokernel}\left(\Pi_b K_0(B_n) \xrightarrow{\times k} \Pi_b K_0(B_n)\right)\\ &\qquad = \text{Cokernel}\left(\prod_{n=1}^{+\infty} K_0(B_n) \xrightarrow{\times k} \prod_{n=1}^{+\infty} K_0(B_n)\right).\end{aligned}$$

Combined with 5.11, yields

$$K_0(Q(B), \mathbb{Z}/k) = \prod_{n=1}^{+\infty} K_0(B_n, \mathbb{Z}/k) \Big/ \bigoplus_{n=1}^{+\infty} K_0(B_n, \mathbb{Z}/k),$$

and

$$K_1(Q(B), \mathbb{Z}/k) = \prod_{n=1}^{+\infty} K_1(B_n, \mathbb{Z}/k) \Big/ \bigoplus_{n=1}^{+\infty} K_1(B_n, \mathbb{Z}/k).$$

5.13. Following [DG], denote

$$\underline{K}(A) = K_*(A) \;\oplus\; \bigoplus_{n=2}^{+\infty} K_*(A, \mathbb{Z}/n).$$

For any finite CW complex X and two KK-elements $\alpha, \beta \in KK(C(X), A)$, from [DL] (also see [DG]), we know that $\alpha = \beta$ if and only if

$$\alpha_* = \beta_* : \; \underline{K}(C(X)) \longrightarrow \underline{K}(A).$$

We will discuss the special cases of $X = \{pt\}, [0,1], T_{II,k}, T_{III,k}$ and S^2, where $T_{II,k}, T_{III,k}$ are defined in the Introduction. (See §4 of [EG2] for details.) (The case $X = \{pt\}$ or $[0,1]$ is similar to the case $X = S^2$, so we will not discuss the spaces $\{pt\}$ and $[0,1]$ separately.)
From [DL], there is an isomorphism

$$KK(C(X), B) \longrightarrow \text{Hom}_\Lambda(\underline{K}(C(X)), \underline{K}(B)),$$

where $\mathrm{Hom}_\Lambda(\underline{K}(C(X)), \underline{K}(B))$ is the set of systems of group homomorphisms which is compatible with all the Bockstein Operations (see [DL] for details). For any fixed finite CW complex X, an element $\alpha \in KK(C(X), B)$ is determined by the system of maps

$$\alpha_n^* : \; K_*(C(X), \mathbb{Z}/n) \longrightarrow K_*(B, \mathbb{Z}/n), \; n = 0, 2, 3, \cdots$$

which are induced by α. In fact α would be determined by a few maps from the above list—all the other maps in the system $\{\alpha_n^*\}_{n=0}^{+\infty} : \underline{K}(C(X)) \to \underline{K}(B)$ would be completely determined by these few maps via the Bockstein Operations. We will choose those few maps for the cases $X = \{pt\}$, $[0,1]$, S^2, $T_{II,k}$, or $T_{III,k}$.
1. $X = S^2$. Then

$$\underline{K}(C(S^2)) \longrightarrow \underline{K}(B)$$

is completely determined by

$$K_0(C(S^2)) \longrightarrow K_0(B)$$

via the Bockstein Operation

$$\begin{array}{ccc} K_0(C(S^2)) & \longrightarrow & K_0(C(S^2), \mathbb{Z}/k) \\ \downarrow & & \downarrow \\ K_0(B) & \longrightarrow & K_0(B, \mathbb{Z}/k) \end{array}$$

since the top horizontal map is surjective. (Note that $K_1(C(S^2))$ and $K_1(C(S^2), \mathbb{Z}/k)$ are trivial groups.) Therefore,

$$KK(C(S^2), B) \cong \mathrm{Hom}(K_0(C(S^2)), K_0(B)).$$

(This is also a well known consequence of the Universal Coefficient Theorem.) (The case $X = \{pt\}$ or $[0,1]$ is similar to the above case.)
2. $X = T_{II,k}$. Let $rC(T_{II,k}) \cong \mathbb{C}$ and let $C_0(T_{II,k})$ be the ideal of $C(T_{II,k})$ consisting of the continuous functions vanishing at the base point. (See 1.6 of [EG2] and 1.1.7 for the notations.) Consider the splitting exact sequence

$$0 \longrightarrow K_0(C_0(T_{II,k})) \longrightarrow K_0(C(T_{II,k})) \longrightarrow K_0(rC(T_{II,k})) \longrightarrow 0.$$

Each KK-element $\alpha \in KK(C(T_{II,k}), B)$ induces two group homomorphisms

$$\alpha_0^0 : \; K_0(rC(T_{II,k})) \; (= \mathbb{Z}) \longrightarrow K_0(B) \qquad \text{and}$$

$$\alpha_k^1 : \; K_1(C(T_{II,k}), \mathbb{Z}/k) \; (= \mathbb{Z}/k) \longrightarrow K_1(B, \mathbb{Z}/k).$$

This induces a map

$$\begin{aligned} KK(C(T_{II,k}), B) &\longrightarrow \mathrm{Hom}(K_0(rC(T_{II,k})), K_0(B)) \bigoplus \mathrm{Hom}(K_1(C(T_{II,k}), \mathbb{Z}/k), K_1(B, \mathbb{Z}/k)) \\ &= \mathrm{Hom}(\mathbb{Z}, K_0(B)) \bigoplus \mathrm{Hom}(\mathbb{Z}/k, K_1(B, \mathbb{Z}/k)). \end{aligned}$$

It can be verified that any two homomorphisms

$$\alpha_0^0: \ K_0(rC(T_{II,k}))\ (=\mathbb{Z}) \longrightarrow K_0(B) \qquad \text{and}$$

$$\alpha_k^1: \ K_1(C(T_{II,k}), \mathbb{Z}/k)\ (=\mathbb{Z}/k) \longrightarrow K_1(B, \mathbb{Z}/k)$$

induces a unique system of homomorphisms in $\mathrm{Hom}_\Lambda(\underline{K}(C(T_{II,k})), \underline{K}(B))$. Therefore, the above map is an isomorphism.
Another way to see it, is as follows. Note that

$$K_1(C(T_{II,k}), \mathbb{Z}/k) = K_0(C_0(T_{II,k})) \subset K_0(C(T_{II,k})).$$

Considering

$$K_1(B) \xrightarrow{\times k} K_1(B) \longrightarrow K_1(B, \mathbb{Z}/k) \longrightarrow K_0(B) \xrightarrow{\times k} K_0(B),$$

we obtain

$$\begin{aligned}&\mathrm{Hom}(K_1(C(T_{II,k}), \mathbb{Z}/k), K_1(B, \mathbb{Z}/k))\\ &\qquad \cong \mathrm{Hom}(K_0(C_0(T_{II,k})), K_0(B)) \oplus \mathrm{Ext}(K_0(C_0(T_{II,k})), K_1(B)).\end{aligned}$$

Then from the Universal Coefficient Theorem,

$$\begin{aligned}&KK(C(T_{II,k}), B)\\ &\cong \mathrm{Hom}(K_0(C(T_{II,k})), K_0(B)) \oplus \mathrm{Ext}(K_0(C(T_{II,k})), K_1(B))\\ &\cong \ \mathrm{Hom}(K_0(rC(T_{II,k})), K_0(B)) \oplus \mathrm{Hom}(K_0(C_0(T_{II,k})), K_0(B))\\ &\qquad\qquad \oplus \mathrm{Ext}(K_0(C_0(T_{II,k})), K_1(B)).\end{aligned}$$

(Note that $K_1(C(T_{II,k})) = 0$.) Hence one can see again, the map mentioned above is an isomorphism.
3. $X = T_{III,k}$. Also, let $rC(T_{III,k}) = \mathbb{C}$ and let $C_0(T_{III,k})$ be the ideal consisting of functions vanishing at the base point. Notice that

$$K_0(C(T_{III,k})) = \mathbb{Z} \qquad \text{and} \qquad K_0(C_0(T_{III,k}), \mathbb{Z}/k) = \mathbb{Z}/k.$$

By the splitting exact sequence

$$0 \to K_0(C_0(T_{III,k}), \mathbb{Z}/k) \to K_0(C(T_{III,k}), \mathbb{Z}/k) \to K_0(rC(T_{III,k}), \mathbb{Z}/k) \to 0$$

we know that each $\alpha \in KK(C(T_{III,k}), B)$ induces an element

$$\alpha_k^0: \ K_0(C_0(T_{III,k}), \mathbb{Z}/k) \longrightarrow K_0(B, \mathbb{Z}/k).$$

It can be proved that

$$\begin{aligned}&KK(C(T_{III,k}), B)\\ &\ \cong \mathrm{Hom}(K_0(C(T_{III,k})), K_0(B)) \bigoplus \mathrm{Hom}(K_0(C_0(T_{III,k}), \mathbb{Z}/k), K_0(B, \mathbb{Z}/k))\\ &\ = \mathrm{Hom}(\mathbb{Z}, K_0(B)) \bigoplus \mathrm{Hom}(\mathbb{Z}/k, K_0(B, \mathbb{Z}/k)),\end{aligned}$$

as what we did for $T_{II,k}$.
(Notice that the map $K_1(C(T_{III,k})) \to K_1(B)$ is completely determined by the map $K_0(C_0(T_{III,k}), \mathbb{Z}/k) \to K_0(B, \mathbb{Z}/k)$.)
Summarizing the above, we have the following.
For any two elements $\alpha, \beta \in KK(C(X), B)$, $\alpha = \beta$ if and only if
(1) $\alpha_0^0 = \beta_0^0 : K_0(C(X)) \longrightarrow K_0(B)$, when $X = S^2$;
(2) $\alpha_0^0 = \beta_0^0 : K_0(rC(X)) \longrightarrow K_0(B)$ and $\alpha_k^1 = \beta_k^1 : K_1(C(X), \mathbb{Z}/k) \longrightarrow K_1(B, \mathbb{Z}/k)$, when $X = T_{II,k}$;
(3) $\alpha_0^0 = \beta_0^0 : K_0(C(X)) \longrightarrow K_0(B)$ and $\alpha_k^0 = \beta_k^0 : K_0(C_0(X), \mathbb{Z}/k) \longrightarrow K_0(B, \mathbb{Z}/k)$, when $X = T_{III,k}$.
Therefore, we have the following lemma.

Lemma 5.14. *Let $A = PM_l(C(X))P$, and X one of $\{pt\}, [0,1], T_{II,k}, T_{III,k}$, or S^2. Let $\alpha, \beta \in KK(A, B)$, where B is a C^*-algebra. Then $\alpha = \beta$ if and only if the following hold:*
1. When $X = \{pt\}$, $[0,1]$ or S^2,

$$\alpha_* = \beta_* : K_0(A) \longrightarrow K_0(B);$$

2. When $X = T_{II,k}$,

$$\alpha_* = \beta_* : K_0(A) \longrightarrow K_0(B) \qquad \textit{and}$$

$$\alpha_* = \beta_* : K_1(A, \mathbb{Z}/k) \longrightarrow K_1(B, \mathbb{Z}/k);$$

3. When $X = T_{III,k}$,

$$\alpha_* = \beta_* : K_0(A) \longrightarrow K_0(B) \qquad \textit{and}$$

$$\alpha_* = \beta_* : K_0(A, \mathbb{Z}/k) \longrightarrow K_0(B, \mathbb{Z}/k).$$

Combined with Theorem 6.1 of [DG], yields the following lemma.

Lemma 5.15. *Let $A = PM_l(C(X))P$, and X, one of $\{pt\}, [0,1], T_{II,k}, T_{III,k}$ or S^2, and let B be any C^*-algebra. Let $\phi, \psi \in Hom(A, B)$. Suppose that the following statements hold.*
1. When $X = \{pt\}, [0,1]$ or S^2,

$$[\phi]_* = [\psi]_* : K_0(A) \longrightarrow K_0(B);$$

2. When $X = T_{II,k}$,

$$[\phi]_* = [\psi]_* : K_0(A) \longrightarrow K_0(B) \qquad \textit{and}$$

$$[\phi]_* = [\psi]_* : K_1(A, \mathbb{Z}/k) \longrightarrow K_1(B, \mathbb{Z}/k);$$

3. When $X = T_{III,k}$,

$$[\phi]_* = [\psi]_* : K_0(A) \longrightarrow K_0(B) \qquad \textit{and}$$

$$[\phi]_* = [\psi]_* : \ K_0(A, \mathbb{Z}/k) \longrightarrow K_0(B, \mathbb{Z}/k).$$

It follows that, for any finite set $F \subset A$ and any number $\varepsilon > 0$, there exist $n \in \mathbb{N}$, $\mu \in Hom(A, M_n(B))$ with finite dimensional image and a unitary $u \in M_{n+1}(B)$ such that

$$\|u(\phi(a) \oplus \mu(a))u^* - \psi(a) \oplus \mu(a)\| < \varepsilon$$

for all $a \in F$.

5.16. Fix $A = PM_l(C(X))P$, $X = \{pt\}, [0,1], T_{II,k}, T_{III,k}$ or S^2. Then A is stably isomorphic to $C(X)$. By 5.14, an element $\alpha \in KK(A,B)$ is completely determined by

$$\alpha_0^0 : \ K_0(A) \to K_0(B),$$
$$\alpha_k^0 : \ K_0(A, \mathbb{Z}/k) \to K_0(B, \mathbb{Z}/k), \qquad \text{and}$$
$$\alpha_k^1 : \ K_1(A, \mathbb{Z}/k) \to K_1(B, \mathbb{Z}/k).$$

Note that, for any C^*-algebra A,

$$K_0(A \otimes C(W_k \times S^1)) \cong K_0(A) \oplus K_1(A) \oplus K_0(A, \mathbb{Z}/k) \oplus K_1(A, \mathbb{Z}/k).$$

Each projection $p \in M_\infty(A \otimes C(W_k \times S^1))$ defines an element

$$[p] \in K_0(A) \oplus K_1(A) \oplus K_0(A, \mathbb{Z}/k) \oplus K_1(A, \mathbb{Z}/k) \subset \underline{K}(A).$$

This defines a map from the set of projections in $\bigcup_{k=2}^\infty M_\infty(A \otimes C(W_k \times S^1))$ to $\underline{K}(A)$.
For any finite set $\mathcal{P} \subset \bigcup_{k=2}^\infty M_\infty(A \otimes C(W_k \times S^1))$ of projections, denoted by $\mathcal{P}\underline{K}(A)$ the finite subset of $\underline{K}(A)$ consisting of elements coming from the projections $p \in \mathcal{P}$, that is

$$\mathcal{P}\underline{K}(A) = \{[p] \in \underline{K}(A) | \ p \in \mathcal{P}\}.$$

In particular, if $A = PM_l(C(X))P, X = T_{II,k}$, or $T_{III,k}$, then we can choose a finite set of projections $\mathcal{P}_A \subset M_\bullet(A \otimes C(W_k \times S^1))$ such that the set $\{[p] \in K_0(A) \oplus K_1(A) \oplus K_0(A, \mathbb{Z}/k) \oplus K_1(A, \mathbb{Z}/k) \mid p \in \mathcal{P}_A\} = \mathcal{P}_A\underline{K}(A)$ generates $K_0(A) \oplus K_1(A) \oplus K_0(A, \mathbb{Z}/k) \oplus K_1(A, \mathbb{Z}/k) \subset \underline{K}(A)$. For $X = \{pt\}, [0,1]$ or S^2, choose $\mathcal{P}_A \subset M_\bullet(A)$ such that $\{[p] \in K_0(A) \mid p \in \mathcal{P}_A\}$ generates $K_0(A) \subset \underline{K}(A)$. We will use $\mathcal{P}$ to denote $\mathcal{P}_A$ if there is no danger of confusion.

5.17. Let $A = PM_l(C(X))P, X = \{pt\}, [0,1], T_{II,k}, T_{III,k}$ or S^2, and $\mathcal{P} \subset M_\bullet(A \otimes C(W_k \times S^1))$ or $\mathcal{P} \subset M_\bullet(A)$ be as in 5.16. There are a finite subset $G(\mathcal{P}) \subset A$ and a number $\delta(\mathcal{P}) > 0$ such that if B is any C^*-algebra and $\phi \in \mathrm{Map}(A,B)$ is $G(\mathcal{P}) - \delta(\mathcal{P})$ multiplicative, then

$$\|((\phi \otimes \mathrm{id})(p))^2 - (\phi \otimes \mathrm{id})(p)\| < \frac{1}{4}, \ \ \forall p \in \mathcal{P},$$

where id is the identity map on $M_\bullet(C(W_k \times S^1))$ or on $M_\bullet(\mathbb{C})$. Hence for any $p \in \mathcal{P}$, there is a projection $q \in M_\bullet(B \otimes C(W_k \times S^1))$ (or $q \in M_\bullet(B)$) such that

$$\|(\phi \otimes \mathrm{id})(p) - q\| < \frac{1}{2}.$$

So q defines an element in $\underline{K}(B)$. (If q' is another projection satisfying the same condition, then $\|q - q'\| < 1$, hence q' is unitarily equivalent to q.)
Therefore, if ϕ is $G(\mathcal{P}) - \delta(\mathcal{P})$ multiplicative, then it induces a map

$$\phi_* : \mathcal{P}\underline{K}(A) \to \underline{K}(B).$$

Note that such $G(\mathcal{P})$ and $\delta(\mathcal{P})$ could be defined for any finite set $\mathcal{P} \subset M_\infty(A) \cup M_\infty(A \otimes C(S^1)) \cup \bigcup_{k=2}^{\infty} M_\infty(A \otimes C(W_k \times S^1))$ of projections.

Theorem 5.18. *Let X be one of the spaces $\{pt\}, [0,1], T_{II,k}, T_{III,k}$ or S^2. Let $A = PM_l(C(X))P$ and $\mathcal{P}$ be as in 5.16. For any finite set $F \subset A$, any positive number $\varepsilon > 0$, and any positive integer M, there are a finite set $G \subset A$ ($G \supset G(\mathcal{P})$ large enough), a positive number $\delta > 0$ ($\delta < \delta(\mathcal{P})$ small enough), and a positive integer L (large enough) such that the following statement is true.*
If $\phi, \psi \in Map(A, B)$ are G-δ multiplicative and

$$\phi_* = \psi_* : \ \mathcal{P}\underline{K}(A) \longrightarrow \underline{K}(B),$$

where $B = QM_\bullet(C(Y))Q$ with $\dim(Y) \le M$, then there is a homomorphism $\nu \in Hom(A, M_L(B))$, with finite dimensional image, and there is a unitary $u \in M_{L+1}(B)$ such that

$$\|u(\phi \oplus \nu)(a)u^* - (\psi \oplus \nu)(a)\| < \varepsilon, \quad \forall f \in F.$$

Proof: We first prove the theorem for the case $B = M_\bullet(C(Y))$. Then we apply Lemma 1.3.6 to reduce the general case to this special case.
We prove the theorem by contradiction.
Let $G(\mathcal{P}) \subset G_1 \subset G_2 \subset \cdots \subset G_n \subset \cdots$ be a sequence of finite subsets with

$$\overline{\cup G_n} = \text{unit ball of } A.$$

Let $\delta(\mathcal{P}) > \delta_1 > \delta_2 > \cdots > \delta_n > \cdots$ be a sequence of positive numbers with $\delta_n \to 0$. Let $L_1 < L_2 < \cdots < L_n < \cdots$ be a sequence of positive integers with $L_n \to +\infty$.
Suppose that the theorem does not hold for (G_n, δ_n, L_n). That is, there exist a C^*-algebra $B_n = M_{k_n}(C(Y_n))$ and two G_n-δ_n multiplicative maps

$$\phi_n, \psi_n : \ A \longrightarrow B_n$$

with $(\phi_n)_* = (\psi_n)_* : \mathcal{P}\underline{K}(A) \longrightarrow \underline{K}(B_n)$ and

$$\inf_{\nu,u} \sup_{a\in F} \|u(\phi_n \oplus \nu)(a)u^* - (\psi_n \oplus \nu)(a)\| \geq \varepsilon, \tag{$*$}$$

where ν runs over all subsets of $\mathrm{Hom}(A, M_{L_n}(B_n))$ consisting of those homomorphisms with finite dimensional images, and u runs over $U(M_{L_n+1}(B_n))$. The above $\{\phi_n\}_{n=1}^{+\infty}, \{\psi_n\}_{n=1}^{+\infty}$ induce two homomorphisms

$$\tilde{\phi}, \tilde{\psi} : \ A \longrightarrow \prod_{n=1}^{+\infty} B_n \Big/ \bigoplus_{n=1}^{+\infty} B_n = Q(B).$$

We will prove that $KK(\phi) = KK(\psi)$.
1. $X = \{pt\}, [0,1]$ or S^2. By Lemma 5.14, $KK(\tilde{\phi})$ is completely determined by

$$([\tilde{\phi}]_*)_0^0 : \ K_0(A) \longrightarrow K_0(Q(B)).$$

From 5.12,

$$K_0(Q(B)) = \Pi_b K_0(B_n) \Big/ \bigoplus_{n=1}^{+\infty} K_0(B_n).$$

That is, the above $([\tilde{\phi}]_*)_0^0$ is completely determined by the component $([\phi_n]_*)_0^0$. From the condition that

$$[\phi_n]_* = [\psi_n]_* : \ \mathcal{P}\underline{K}(A) \longrightarrow \underline{K}(B_n)$$

and the condition that the group generated by $\mathcal{P}\underline{K}(A)$ is $K_0(A)$, we know that $KK(\phi) = KK(\psi)$.
2. $X = T_{II,k}$. By Lemma 5.14, $KK(\tilde{\phi})$ is completely determined by

$$([\tilde{\phi}]_*)_0^0 : \ K_0(A) \longrightarrow K_0(Q(B)) \qquad \text{and}$$

$$([\tilde{\phi}]_*)_k^1 : \ K_1(A, \mathbb{Z}/k) \longrightarrow K_1(Q(B), \mathbb{Z}/k).$$

Furthermore, by 5.12,

$$K_1(Q(B), \mathbb{Z}/k) = \prod_{n=1}^{+\infty} K_1(B_n, \mathbb{Z}/k) \Big/ \bigoplus_{n=1}^{+\infty} K_1(B_n, \mathbb{Z}/k).$$

Again, $([\tilde{\phi}]_*)_0^0$ and $([\tilde{\phi}]_*)_k^1$ are completely determined by the components corresponding to $[\phi_n]_*$. And from $[\phi_n]_* = [\psi_n]_*$ on $\mathcal{P}\underline{K}(A)$, we obtain $KK(\tilde{\phi}) = KK(\tilde{\psi})$. (Note that, we also use the fact that $\mathcal{P}\underline{K}(A)$ generates a subgroup of $\underline{K}(A)$ containing $K_0(A)$ and $K_1(A, \mathbb{Z}/k)$.) (The subgroup of $\underline{K}(A)$ generated by $\mathcal{P}\underline{K}(A)$ also contains $K_1(A)$, though we do not use this fact.)
3. $X = T_{III,k}$. It can be proved that $KK(\tilde{\phi}) = KK(\tilde{\psi})$ as above. Note that $K_0(B_n, \mathbb{Z}/k) = \prod_{n=1}^{+\infty} K_0(B_n, \mathbb{Z}/k) / \bigoplus_{n=1}^{+\infty} K_0(B_n, \mathbb{Z}/k)$, by 5.12.

By Lemma 5.15, there are a positive integer L and a homomorphism $\tilde{\nu}: A \to M_L(\prod_{n=1}^{+\infty} B_n / \bigoplus_{n=1}^{+\infty} B_n)$ with finite dimensional image, and a unitary $\tilde{u} \in M_{L+1}(\prod_{n=1}^{+\infty} B_n / \bigoplus_{n=1}^{+\infty} B_n)$, such that

$$\|\tilde{u}(\tilde{\phi} \oplus \tilde{\nu})(a)\tilde{u}^* - (\tilde{\psi} \oplus \tilde{\nu})(a)\| < \frac{\varepsilon}{2}$$

for all $a \in F$. Since $\tilde{\nu}$ has finite dimensional image, one can find a sequence of homomorphisms

$$\nu_n: \ A \longrightarrow M_L(B_n)$$

of finite dimensional images such that $\{\nu_n\}_{n=1}^{+\infty}$ induces $\tilde{\nu}$. One can also lift $\tilde{u}$ to a sequence of unitaries $u_n \in M_{L+1}(B_n)$. Then if n is large enough, we have

$$\|u_n(\phi_n \oplus \nu_n)(a)u_n^* - (\psi_n \oplus \nu_n)(a)\| < \varepsilon$$

for all $a \in F$. This contradicts with $(*)$ if one choose n to satisfy $L_n \geq L$.

Now we apply Lemma 1.3.6 to prove the general case. Let G, δ and L_1 (in place of L) be as above for the case of full matrix algebras over $C(Y)$ with $\dim(Y) \leq M$. Choose $L = (2M+2)L_1 - 1$. We will verify that G, δ and L satisfies the condition of the theorem even for $B = QM_\bullet(C(Y))Q$—cutting down of full matrix algebras by projections, as follows.

Let $n = \mathrm{rank}(Q) + \dim(Y)$ and $m = 2M+1$. Then by Lemma 1.3.6, $QM_\bullet(C(Y))Q$ can be identified as a corner subalgebra of $M_n(C(Y))$, and $M_n(C(Y))$ can be identified as corner subalgebra of $M_m(QM_\bullet(C(Y))Q)$.

If $\phi, \psi \in \mathrm{Map}_{G-\delta}(A, B)$ satisfy the condition in the theorem, then regarding B as a corner subalgebra of $M_n(C(Y))$, we can regard ϕ, ψ as elements in $\mathrm{Map}_{G-\delta}(A, M_n(C(Y)))$ which still satisfy the condition. Hence from the above special case of the theorem, there are $\nu: A \to M_{L_1}(M_n(C(Y)))$ and a unitary $u_1 \in M_{L_1+1}(M_n(C(Y)))$ such that

$$\|u_1(\phi \oplus \nu)(a)u_1^* - (\psi \oplus \nu)(a)\| < \varepsilon, \quad \forall a \in F.$$

Also, $M_n(C(Y))$ can be regarded as a corner subalgebra of $M_m(QM_\bullet(C(Y))Q)$, so $\phi \oplus \nu$ and $\psi \oplus \nu$ can be regarded as maps from A to $M_{L_1+1}(M_m(QM_\bullet(C(Y))Q)) = M_{L+1}(QM_\bullet(C(Y))Q)$. Therefore, there is a unitary $u \in M_{L+1}(B)$

$$\|u(\phi \oplus \nu)(a)u^* - (\psi \oplus \nu)(a)\| < \varepsilon, \quad \forall a \in F.$$

$\square$

Remark 5.19. The theorem is not true for $X = S^1$, even if we assume that both ϕ and ψ are homomorphisms. A counterexample is given below .

Let $\phi_n, \ \psi_n: \ C(S^1) \to C[0,1]$ be defined by

$$\phi_n(f)(t) = f(e^{2\pi int}) \qquad \text{and} \qquad \psi_n(f)(t) = f(\mathbf{1}).$$

Then $KK(\phi_n) = KK(\psi_n)$. Let $F = \{z\}$ and $\varepsilon = \frac{1}{4}$, where $z \in C(S^1)$ is a canonical generator. One can prove that there is no integer L which is good for all (ϕ_n, ψ_n) as in Theorem 5.18, by using the variation of determinant.

5.20. If X is any finite CW complex such that $K_1(C(X))$ is a torsion group, then Theorem 5.18 holds for X — one needs to choose $\mathcal{P}_A$ accordingly, which is described below.

Suppose that $m_1, m_2, \cdots, m_i$ are the degrees of all the torsion elements in $K_0(A)$ and $K_1(A)$. Let m be the least common multiple of $m_1, m_2, \cdots, m_i$. Since $K_1(A)$ is a torsion group, similar to the discussion in 5.13, an element $\alpha \in KK(A, B)$ is completely determined by the K-theory maps

$$\alpha^0 : K_0(A) \longrightarrow K_0(B),$$

$$\alpha_p^0 : K_0(A, \mathbb{Z}/p) \longrightarrow K_0(B, \mathbb{Z}/p),$$

$$\alpha_p^1 : K_1(A, \mathbb{Z}/p) \longrightarrow K_1(B, \mathbb{Z}/p),$$

where p are all the numbers with $p|m$. (In particular, the map $\alpha^1 : K_1(A) \longrightarrow K_1(B)$ is determined by the above maps.)

One can choose $\mathcal{P}$ to be a finite set of projections in $M_\bullet(A) \bigcup \bigcup_{p|m} M_\bullet(A \otimes C(W_p \otimes S^1))$ such that the set $\mathcal{P}\underline{K}(A)$ (defined in 5.16) generates a sub group containing the group

$$K_0(A) \oplus \bigoplus_{k|m} K_*(A, \mathbb{Z}/k).$$

Similar to the proof of Theorem 5.18, we can prove the following theorem, since, to determine a KK-element $\alpha \in KK(A, B)$, one does not need the map from $K_1(A)$. ($G(\mathcal{P})$ and $\delta(\mathcal{P})$ can be chosen accordingly as in 5.17.)

Theorem 5.21. *Suppose that X is a finite CW complex with $K_1(X)$ a torsion group. Suppose that $A = PM_l(C(X))P$ and $\mathcal{P}$ are as in 5.20. For any finite set $F \subset A$, positive number $\varepsilon > 0$, and positive integer M, there are a finite set $G \subset A$ ($G \supset G(\mathcal{P})$ large enough), a positive number $\delta > 0$ ($\delta < \delta(\mathcal{P})$ small enough), and a positive integer L (large enough) such that the following statement is true.*

If $\phi, \psi \in Map(A, B)$ are $G(\mathcal{P})$-$\delta(\mathcal{P})$ multiplicative and

$$\phi_* = \psi_* : \mathcal{P}\underline{K}(A) \longrightarrow \underline{K}(B),$$

where $B = QM_\bullet(C(Y))Q$ with $\dim(Y) \leq M$, then there is a homomorphism $\nu \in Hom(A, M_L(B))$ with finite dimensional image, and there is a unitary $u \in M_{L+1}(B)$ such that

$$\|u(\phi \oplus \nu)(a)u^* - (\psi \oplus \nu)(a)\| < \varepsilon$$

for all $a \in F$.

The following is a direct consequence of Theorem 5.18.

Corollary 5.22. *Let $A = C(X)$, where X is one of the spaces: $[0,1], S^2, T_{II,k}$ or $T_{III,k}$, and let $\mathcal{P}$ be as in 5.16. For any finite set $F \subset C(X)$, any positive number $\varepsilon > 0$ and any positive integer M, there are a finite set $G \subset C(X)$ ($G \supset G(\mathcal{P})$ large enough), positive numbers $\delta > 0$ ($\delta \le \delta(\mathcal{P})$ small enough) and $\eta > 0$ (small enough) such that the following statement is true.*
Let $B = M_\bullet(C(Y))$ with $\dim(Y) \le M$, and $p \in B$, a projection.
If $\phi, \psi \in Map(C(X), pBp)$ are G-δ multiplicative maps inducing the same maps $\phi_* = \psi_* : \ \mathcal{P}\underline{K}(C(X)) \longrightarrow \underline{K}(B)$, and $\{x_1, x_2, \cdots, x_n\}$ is an η-dense subset of X, and $q_1, q_2, \cdots, q_n$ are mutually orthogonal projections in $(\mathbf{1}-p)B(\mathbf{1}-p)$ with rank$(q_i) \ge$ rank(p), then there is a unitary

$$u \in (p \oplus q_1 \oplus q_2 \oplus \cdots \oplus q_n)B(p \oplus q_1 \oplus q_2 \oplus \cdots \oplus q_n)$$

such that

$$\left\| \phi(f) \oplus \sum_{i=1}^{n} f(x_i)q_i - u\left(\psi(f) \oplus \sum_{i=1}^{n} f(x_i)q_i\right)u^* \right\| < \varepsilon, \qquad \forall f \in F.$$

In particular, if ψ is a homomorphism, then there is a homomorphism $\tilde{\phi} \in Hom(C(X), (p \oplus q_1 \oplus q_2 \oplus \cdots \oplus q_n)B(p \oplus q_1 \oplus q_2 \oplus \cdots \oplus q_n))$ (defined by $\tilde{\phi}(f) = u\,(\psi(f) \oplus \sum_{i=1}^{n} f(x_i)q_i)\,u^*$) such that

$$\left\| \tilde{\phi}(f) - \left(\phi(f) \oplus \sum_{i=1}^{n} f(x_i)q_i\right) \right\| < \varepsilon, \qquad \forall\ f \in F.$$

Proof: Since X is not the space of a single point, we can assume that X, as a metric space, satisfies that diameter$(X) = 1$. Apply Theorem 5.18 to the finite set $F \subset A$, the positive number $\frac{\varepsilon}{3}$ and the integer M to obtain G, δ, L as in Theorem 5.18. Choose a positive number $\eta < \frac{1}{8ML^2}$ such that if dist$(x, x') < 8ML^2 \cdot \eta$, then $\|f(x) - f(x')\| < \frac{\varepsilon}{3}$ for all $f \in F$.
Let $\{x_1, x_2, \cdots, x_n\}$ be an η-dense subset of X and let $q_1, q_2, \cdots, q_n \in (1-p)B(1-p)$ be mutually orthogonal projections with rank$(q_i) \ge$ rank(p). Similar to the proof of Corollary 1.6.13, one can find a $8ML \cdot \eta$-dense subset $\{x_{k_1}, x_{k_2}, \cdots, x_{k_l}\} \subset \{x_1, x_2, \cdots, x_n\}$ and mutually orthogonal projections $Q_1, Q_2, \cdots, Q_l$ with rank$(Q_j) \ge ML \cdot$ rank(p), such that

$$\| \sum_{i=1}^{n} f(x_i)q_i - \sum_{j=1}^{l} f(x_{k_j})Q_j \| < \frac{\varepsilon}{3}, \quad \forall f \in F.$$

Since rank$(Q_i) \ge ML \cdot$ rank(p), it follows that $[Q_i] \ge L \cdot [p]$.
Again, similar to the proofs of Corollaries 1.6.12 and 1.6.13, it can be proved that a homomorphism $\nu \in$ Hom$(A, M_L(pBp))$ with finite dimensional image

(from 5.18) can be perturbed, at the expense of at most $\frac{\varepsilon}{3}$ on the finite set F, to a homomorphism ν' which is of the form

$$\nu'(f) = \sum_{j=1}^{l} f(x_{k_j})q'_j$$

with $[q'_j] \le [Q_j]$ (SOME OF THE PROJECTIONS q'_j COULD BE ZERO). Hence the corollary follows (see the proof of Corollary 1.6.12).

□

By the discussion in 1.2.19, we have the following corollary.

COROLLARY 5.23. *Let $A = M_l(C(X))$, where X is one of the spaces: $[0,1], S^2, T_{II,k}$ or $T_{III,k}$, and let $\mathcal{P}$ be as in 5.16. For any finite set $F \subset A$, any positive number $\varepsilon > 0$ and any positive integer M, there are a finite set $G \subset A$ ($G \supset G(\mathcal{P})$ large enough), numbers $\delta > 0$ ($\delta \le \delta(\mathcal{P})$ small enough) and $\eta > 0$ (small enough) such that the following statement is true.*
Let $B = M_\bullet(C(Y))$ with $\dim(Y) \le M$, and $p \in B$ a projection.
If $\phi, \psi \in Map(A, pBp)$ are G-δ multiplicative maps inducing the same map $\phi_* = \psi_* : \mathcal{P}\underline{K}(A) \longrightarrow \underline{K}(B)$, and $\{x_1, x_2, \cdots, x_n\}$ is an η-dense subset of X, and

$$q_1 = \underbrace{q'_1 \oplus q'_1 \oplus \cdots \oplus q'_1}_{l}, q_2 = \underbrace{q'_2 \oplus q'_2 \oplus \cdots \oplus q'_2}_{l}, \cdots, q_n = \underbrace{q'_n \oplus q'_n \oplus \cdots \oplus q'_n}_{l}$$

are mutually orthogonal projections in $(\mathbf{1}-p)B(\mathbf{1}-p)$ with rank$(q_i) \ge$ rank(p), then there is a unitary

$$u \in (p \oplus q_1 \oplus q_2 \oplus \cdots \oplus q_n)B(p \oplus q_1 \oplus q_2 \oplus \cdots \oplus q_n)$$

such that

$$\left\| \phi(f) \oplus \sum_{i=1}^{n} q'_i \otimes f(x_i) - u\left(\psi(f) \oplus \sum_{i=1}^{n} q'_i \otimes f(x_i)\right)u^* \right\| < \varepsilon, \qquad \forall f \in F.$$

In particular, if ψ is a homomorphism, then there is a homomorphism $\tilde{\phi} \in Hom(C(X), (p \oplus q_1 \oplus q_2 \oplus \cdots \oplus q_n)B(p \oplus q_1 \oplus q_2 \oplus \cdots \oplus q_n))$ such that

$$\left\| \tilde{\phi}(f) - \left(\phi(f) \oplus \sum_{i=1}^{n} q'_i \otimes f(x_i)\right) \right\| < \varepsilon, \qquad \forall\ f \in F.$$

Proof: Thanks to Lemma 1.6.8, we can always assume that the two maps ϕ and ψ satisfy the condition that $\phi|_{M_l(\mathbb{C})}$ and $\psi|_{M_l(\mathbb{C})}$ are homomorphisms. Using the condition $\phi_* = \psi_* : \mathcal{P}\underline{K}(A) \longrightarrow \underline{K}(B)$, we can assume

$$\phi|_{M_l(\mathbb{C})} = \psi|_{M_l(\mathbb{C})}$$

after conjugating with a unit.
Now the corollary follows from the following claim.
Claim: For any finite set $F \subset A = M_l(C(X))$, any $\varepsilon > 0$, there are a finite set $G \subset A$ and a positive number $\delta > 0$ such that if a map $\phi : A \to B$ is G-δ multiplicative and $\phi|_{M_l(\mathbb{C})}$ is a homomorphism, then there are a map $\phi_1 : C(X) \to \phi(e_{11})B\phi(e_{11})$ and an identification of $\phi(\mathbf{1})B\phi(\mathbf{1}) \cong M_l(\phi(e_{11})B\phi(e_{11}))$ such that

$$\|\phi(f) - (\phi_1 \otimes \mathbf{1}_l)(f)\| < \varepsilon \qquad\qquad \forall f \in F.$$

Furthermore, if $G_1 \subset C(X)$ and $\delta_1 > 0$ are a pregiven finite set and a pregiven positive number, then one can modify the set G and the number δ so that the map ϕ_1 above can be chosen to be G_1-δ_1 multiplicative.
Proof of Claim: Suppose $\mathbf{1} \in F$. Let $F_1 = \{a_{ij} | (a_{ij})_{l\times l} \in F\} (\subset C(X))$ be the set of all entries of the elements in F. Let $G = \{(b_{ij})_{l\times l} = \sum b_{ij}e_{ij} \mid b_{ij} \in F_1 \cup G_1 \subset C(X)\} (\subset A)$ and $\delta = \min\left(\frac{\varepsilon}{2l^2}, \delta_1\right)$. Suppose that $\phi : \ M_l(C(X)) \to B$ is G-δ multiplicative. Let

$$\phi_1 = \phi|_{e_{11}M_l(C(X))e_{11}} : \ C(X) \to \phi(e_{11})B\phi(e_{11}).$$

Obviously the G-δ multiplicativity of ϕ implies the G_1-δ_1 multiplicativity of ϕ_1. Identify $\phi(\mathbf{1})B\phi(\mathbf{1}) \cong (\phi(e_{11})B\phi(e_{11})) \otimes M_l$ by sending $\phi(e_{ij})$ to $e_{ij} \in M_l \subset (\phi(e_{11})B\phi(e_{11})) \otimes M_l$. Under this identification, we have

$$(\phi_1 \otimes \mathbf{1}_l)(a) = \sum_{i,j} \phi(e_{i1})\phi_1(a_{ij})\phi(e_{1j}),$$

where $a = (a_{ij})_{l\times l} \in M_l(C(X))$. On the other hand, writing $a = \sum e_{1i}(a_{ij}e_{11})e_{1j}$ and using the G-δ multiplicativity of ϕ, we have

$$\begin{aligned}\|\phi(a) - (\phi_1 \otimes \mathbf{1}_l)(a)\| &\le \sum_{i,j} \|\phi(e_{1i}(a_{ij}e_{11})e_{1j}) - \phi(e_{i1})\phi_1(a_{ij})\phi(e_{1j})\| \\ &= \sum_{i,j} \|\phi(e_{1i}(a_{ij}e_{11})e_{1j}) - \phi(e_{i1})\phi(a_{ij}e_{11})\phi(e_{1j})\| \\ &\le \sum_{i,j} 2\delta = 2l^2\delta \le \varepsilon.\end{aligned}$$

This proves the Claim.
Applying the Claim, one can reduce the proof to the case $A = C(X)$ which is Corollary 5.22.

$\square$

Definition 5.24. Let A be a unital C^*-algebra, let

$$\mathcal{P} \subset M_\bullet(A) \cup M_\bullet(A \otimes C(S^1)) \cup \bigcup_{k=2}^{\infty} M_\bullet(A \otimes C(W_k \times S^1))$$

be a finite set of projections, and let $G(\mathcal{P})$, $\delta(\mathcal{P})$ be as in 5.17. A $G(\mathcal{P})-\delta(\mathcal{P})$ multiplicative map $\phi: A \to B$ is called QUASI-$\mathcal{P}\underline{K}$-HOMOMORPHISM if there is a homomorphism $\psi: A \to B$ with $\phi(\mathbf{1}_A) = \psi(\mathbf{1}_A)$ such that

$$[\phi]_* = [\psi]_* : \mathcal{P}\underline{K}(A) \to \underline{K}(B).$$

Using the above definition and Definition 4.38, we can restate the second part of Corollary 5.23 as below.

LEMMA 5.25. *Let $A = M_l(C(X))$, where X is one of the spaces: $[0,1], S^2, T_{II,k}$ or $T_{III,k}$, and let $\mathcal{P}$ be as in 5.16. For any finite set $F \subset A$, any positive number $\varepsilon > 0$ and any positive integer M, there are a finite set $G \subset A$ ($G \supset G(\mathcal{P})$ large enough), positive numbers $\delta > 0$ ($\delta \le \delta(\mathcal{P})$ small enough) and $\eta > 0$ (small enough) such that the following statement is true. Let $B = M_\bullet(C(Y))$ with $\dim(Y) \le M$, and let $p \in B$ be a projection. If $\phi \in Map(A, pBp)$ is a G-δ multiplicative quasi-$\mathcal{P}\underline{K}$-homomorphism, and $\lambda \in Hom(A, (\mathbf{1}-p)B(\mathbf{1}-B))$ has the property PE(rank(p), η), then there is a homomorphism $\tilde{\phi} \in Hom(A,B)$ such that*

$$\|\tilde{\phi}(f) - (\phi \oplus \lambda)(f)\| < \varepsilon, \quad \forall f \in F.$$

Furthermore if Y is a connected simplicial complex different from the single point space, then $\tilde{\phi}$ can be chosen to be injective.

Proof: The main body of the lemma is a restatement of Corollary 5.23. So we only need to prove the last sentence of the lemma. We need the following fact: Let $X = [0,1], S^2, T_{II,k}$ or $T_{III,k}$, and let Y be a connected finite simplicial complex different from $\{pt\}$. If $\lambda_1 : M_l(C(X)) \to p_1 M_\bullet(C(Y))p_1$ is a homomorphism defined by the point evaluation at a point $x_1 \in X$ as

$$M_l(C(X)) \xrightarrow{e_{x_1}} M_l(\mathbb{C}) \longrightarrow p_1 M_\bullet(C(Y))p_1,$$

then λ_1 is homotopic to an injective homomorphism $\lambda_1' : M_l(C(X)) \to p_1 M_\bullet(C(Y))p_1$. (Again, this fact can be proved by using the Peano Curve.)
Let η' be as the η desired in the main body of the lemma for $\frac{\varepsilon}{2}$ (in place of ε). We can also assume that η' satisfies the condition that if $\mathrm{dist}(x,x') < \eta'$, then $\|f(x) - f(x')\| < \frac{\varepsilon}{2}$ for all $f \in F$. Choose $\eta = \frac{\eta'}{4}$. Suppose that $\lambda \in \mathrm{Hom}(A, (1-p)B(1-p))$ has the property PE(rank(p), η). Write $\lambda = \bigoplus_{i=1}^n \lambda_i$, where

$$\lambda_i : M_l(C(X)) \xrightarrow{e_{x_i}} M_l(\mathbb{C}) \xrightarrow{\phi_i} p_i B p_i,$$

are point evaluations at an η-dense set $\{x_1, x_2, \cdots, x_n\}$ and ϕ_i are unital homomorphisms.
Let p_1 be a projection with minimum rank among all the projections $p_1, p_2, \cdots, p_n$. Let $\phi \in \mathrm{Map}(A, pBp)$ be a G-δ multiplicative quasi-$\mathcal{P}\underline{K}$-homomorphism. Then $\phi \oplus \lambda_1 \in \mathrm{Map}(A, (p \oplus p_1)B(p \oplus p_1))$ is also a G-δ

multiplicative quasi-$\mathcal{P}\underline{K}$-homomorphism. Furthermore, from the above fact, it defines the same map on the level of $\mathcal{P}\underline{K}(A)$ as an injective homomorphism $\psi \in \mathrm{Hom}(A, (p \oplus p_1)B(p \oplus p_1))$. On the other hand, $\lambda' = \bigoplus_{i=2}^{n} \lambda_i$ has the properties $\mathrm{PE}(\mathrm{rank}(p), 2\eta)$ and $\mathrm{PE}(\mathrm{rank}(p_1), 2\eta)$, since $\lambda = \bigoplus_{i=1}^{n} \lambda_i$ has $\mathrm{PE}(\mathrm{rank}(p), \eta)$, and $\mathrm{rank}(p_1) \le \mathrm{rank}(p_i)$, $i = 2, \cdots, n$. Similar to the proof of Corollary 5.22, λ' can be perturbed to a homomorphism λ'' which has the property $\mathrm{PE}(\mathrm{rank}(p) + \mathrm{rank}(p_1), 4\eta)$ at the expense of at most $\frac{\varepsilon}{2}$ on the finite set F. Note that $4\eta = \eta'$, and ψ is injective. Hence the homomorphism $\mathrm{Ad}u \circ (\psi \oplus \lambda'')$ (for a certain unitary u), as desired in the main body of the lemma, is also injective.

$\square$

Lemma 5.26. *Let X and Y be connected finite simplicial complexes. Suppose that $\phi_1 : PM_k(C(X))P \to Q_1M_l(C(Y))Q_1$ and $\phi_2 : PM_k(C(X))P \to Q_2M_l(C(Y))Q_2$ are unital homomorphisms, where P, Q_1, and Q_2 are projections with*

$$rank(Q_2) - rank(Q_1) \ge 2\dim(Y) \cdot rank(P).$$

Then there exists a homomorphism $\psi : PM_k(C(X))P \to M_\bullet(C(Y))$ such that

$$[\psi] = [\phi_2] - [\phi_1] \in KK(C(X), C(Y)).$$

Proof: First, we suppose that $A = C(X)$. As in Lemma 3.14 of [EG 2] (see Remark 1.6.21 above), we can assume that $\phi_1(C_0(X)) \subset M_l(C_0(Y))$ and $\phi_2(C_0(X)) \subset M_l(C_0(Y))$, where $C_0(X)$ and $C_0(Y)$ are sets of functions vanishing on fixed base points of X and Y, respectively. Hence ϕ_i defines an element $kk(\phi_i) \in kk(Y, X)$ (see [DN]). Furthermore, $[\phi_i] \in KK(C(X), C(Y))$ is completely determined by $kk(\phi_i)$ and $\phi_{i*}([\mathbf{1}_A]) \in K_0(B)$. Let $\alpha = kk(\phi_2) - kk(\phi_1) \in kk(Y, X)$ (note that $kk(Y, X)$ is an abelian group, see [DN]). Since $\mathrm{rank}(Q_2) - \mathrm{rank}(Q_1) \ge 2\dim(Y)$, by [Hu], there is a projection $Q_3 \in M_\bullet(C(Y))$ such that $[Q_3] = [Q_2] - [Q_1] \in K_0(C(Y))$. By Theorem 4.11 of [DN] or Lemma 3.16 of [EG2], there is a unital homomorphism $\psi : C(X) \to Q_3M_\bullet(C(Y))Q_3$ to realize $\alpha \in kk(Y, X)$. Obviously ψ is as desired.
For the general case, using the Dilation Lemma (Lemma 1.3.1), one can prove that $[\phi_i] \in KK(C(X), C(Y))$ can be realized by homomorphism $\phi_i' : C(X) \to M_\bullet(C(Y))$. This reduces the proof to the above case.

$\square$

Remark 5.27. In the above lemma, if $Q_1 < Q_2$, then one can choose ψ to satisfy $\psi(\mathbf{1}_A) = Q_2 - Q_1$.

Lemma 5.28. *Let X be a finite simplicial complex, and $A = PM_l(C(X))P$.*

For any finite set

$$\mathcal{P} \subset M_\bullet(A) \cup M_\bullet(A \otimes C(S^1)) \cup \bigcup_{k=2}^{\infty} M_\bullet(A \otimes C(W_k \times S^1)),$$

there are a finite set $G \subset A$ and a number $\delta > 0$, such that the following is true.
If Y is a simplicial complex, $Q > Q_1$ are two projections in $M_\bullet(C(Y))$ with $rank(Q) - rank(Q_1) \geq 2\dim(Y) rank(P)$, and two unital homomorphisms $\phi \in Hom(A, QM_\bullet(C(Y))Q)_1$, $\phi_1 \in Hom(A, Q_1 M_\bullet(C(Y))Q_1)_1$ and a unital map $\phi_2 \in Map(A, (Q - Q_1)M_\bullet(C(Y))(Q - Q_1))_1$ satisfy that

$$(*) \qquad \|\phi(f) - \phi_1(f) \oplus \phi_2(f)\| < \delta, \quad \forall g \in G,$$

then there is a homomorphism $\psi : A \to (Q - Q_1)M_\bullet(C(Y))(Q - Q_1)$ such that

$$[\psi]_* = [\phi_2]_* : \mathcal{P}\underline{K}(A) \to \underline{K}(C(Y)).$$

In other words, ϕ_2 is a quasi-$\mathcal{P}\underline{K}$-homomorphism.
(Notice that, from Lemma 4.40, if G is large enough and δ is small enough, then (*) above implies that
$\phi_2 \in \mathrm{Map}(A, (Q - Q_1)M_\bullet(C(Y))(Q - Q_1))_1$ is $G(\mathcal{P}) - \delta(\mathcal{P})$ multiplicative, and hence $[\phi_2]_* : \mathcal{P}\underline{K}(A) \to \underline{K}(C(Y))$ makes sense.)

Proof: If G is large enough and δ is small enough, then (*) implies

$$[\phi_2]_* = [\phi]_* - [\phi_1]_* : \mathcal{P}\underline{K}(A) \to \underline{K}(C(Y)).$$

Then the lemma follows from Lemma 5.26 and Remark 5.27.

□

REMARK 5.29. In Corollary 4.39, we can choose ψ_0 (or ψ_0') such that $\psi_0^{i,j}$ (or $\psi_0'^{i,j}$) is a quasi-$\mathcal{P}\underline{K}$-homomorphism for any pre-given set of projections

$$\mathcal{P} \subset M_\bullet(A) \cup M_\bullet(A \otimes C(S^1)) \cup \bigcup_{k=2}^{\infty} M_\bullet(A \otimes C(W_k \times S^1)).$$

To do so, by Lemma 5.28, one only needs to choose the projection $Q_0^{i,j}$ to have rank at least $2\dim(X_{m,j}) \cdot \mathrm{rank}(1_{A_n^i})$. But from the construction in 4.34, we have freedom to do so.

LEMMA 5.30. *Fix a positive integer M. Suppose that $B = \bigoplus_{i=1}^{s} M_{l_i}(C(Y_i))$, where Y_i are the spaces: $\{pt\}, [0,1], S^1, T_{II,k}, T_{III,k}, S^2$. For any finite set $G \subset B$ and positive number $\varepsilon > 0$, there exist a finite set $G_1 \subset B$, numbers $\delta_1 > 0$ and $\eta > 0$ such that the following is true.*
If a map $\alpha = \alpha_0 \oplus \alpha_1 : B \to A = \bigoplus_{j=1}^{t} M_{k_j}(C(X_j))$, with $\dim(X_j) \leq M$, satisfies the following conditions:

(1) α_0 is G_1-δ_1 multiplicative, $\{\alpha_0(\mathbf{1}_{B^i})\}_{i=1}^s$ are mutually orthogonal projections, and α_1 is a homomorphism with finite dimensional image (i.e., defined by point evaluations);
(2) For any block B^i with $Y_i = T_{II,k}, T_{III,k}$ or S^2 and any block A^j, the partial map $\alpha_0^{i,j}$ is quasi-$\mathcal{P}\underline{K}$-homomorphism, where $\mathcal{P}$ is the set of projections associated to B^i as in 5.16, and the homomorphism $\alpha_1^{i,j}$ has the property PE(rank$\alpha_0^{i,j}(\mathbf{1}_{B^i})$, η);

then there is a unital homomorphism $\alpha' : B \to \alpha(\mathbf{1}_B)A\alpha(\mathbf{1}_B)$ such that

$$\|\alpha'(g) - \alpha(g)\| < \varepsilon, \quad \forall g \in G.$$

Proof: We only need to perturb all the individual maps $\alpha^{i,j}$ to homomorphisms ${\alpha'}^{i,j}$ within $\alpha^{i,j}(\mathbf{1}_{B^i})A^j\alpha^{i,j}(\mathbf{1}_{B^i})$.
For a block of B^i with spectrum $\{pt\}$, $[0,1]$ or S^1, such perturbation exists by Lemma 1.6.1. For a block of B^i with spectrum $T_{II,k}, T_{III,k}$ or S^2, such perturbation exists by Lemma 5.25.

$\square$

Lemma 5.31. *Let M be a fixed positive integer. Let $B = M_l(C(Y)), Y = T_{II,k}, T_{III,k}$ or S^2. Let the set of projections $\mathcal{P} \subset M_\bullet(B) \cup M_\bullet(B \otimes C(W_k \times S^1))$ be as in 5.16.*
Let $A = RM_{l_1}(C(X))R$ with $\dim(X) \le M$, where $R \in M_{l_1}(C(X))$ is a projection. Let $\alpha : B \to A$ be an injective homomorphism. Let a finite set of projections $\mathcal{P}'$ be given by $\mathcal{P}' := (\alpha \otimes id)(\mathcal{P}) \subset M_\bullet(A) \cup M_\bullet(A \otimes C(W_k \times S^1))$. Let $\eta > 0$. Choose $\eta_1 > 0$ such that if a finite set $\{x_1, x_2, \cdots, x_n\} \subset X$ is η_1-dense in X, then $\bigcup_{i=1}^n SP\alpha_{x_i}$ is η-dense in Y. (Such η_1 exists because of injectivity of α.)
For any finite subset $G_1 \subset B$ and any number $\delta_1 > 0$, there are a finite subset $G_2 \subset A$ and a number $\delta_2 > 0$ such that the following are true.
Let $C = M_\bullet(Z)$ with $\dim(Z) \le M$.
(1) If $\psi_0 : A \to Q_0CQ_0$ is a G_2-δ_2 multiplicative quasi-$\mathcal{P}'\underline{K}$-homomorphism and $\psi_0(\alpha(\mathbf{1}_B))$ is a projection, then $\psi_0 \circ \alpha$ is a G_1-δ_1 multiplicative quasi-$\mathcal{P}\underline{K}$-homomorphism.
(2) If $\psi_1 : A \to Q_1CQ_1$ has the property PE($J \cdot L$, η_1), where $J =$ rank(R), then $\psi_1 \circ \alpha : B \to \psi_1(\alpha(\mathbf{1}_B))C\psi_1(\alpha(\mathbf{1}_B))$ has property PE(L, η).
In particular, if ψ_1 has the property PE($J \cdot$ rank(Q_0), η_1), (this is the condition (3) of Corollary 4.39), then $\psi_1 \circ \alpha$ has the property PE(rank$((\psi_0 \circ \alpha)(\mathbf{1}_B))$, η). (Note that rank$((\psi_0 \circ \alpha)(\mathbf{1}_B)) \le$ rank(Q_0).) Consequently, if we further assume that G_1, δ_1 and η are as those chosen in 5.30 for a finite set $G \subset B$ and $\varepsilon > 0$, and Q_1 is orthogonal to Q_0, then there is a homomorphism $\psi : B \to (Q_0 \oplus Q_1)C(Q_0 \oplus Q_1)$ such that

$$\|\psi(g) - (\psi_0 \oplus \psi_1)(\alpha(g))\| < \varepsilon, \quad \forall g \in G.$$

Proof: (1) holds if we choose $G_2 \supset \alpha(G)$ and $\delta_2 < \delta_1$.
(2) follows from the following fact: if a homomorphism $\phi : RM_{l_1}(C(X))R \to M_t(C(Z))$ contains a part of point evaluation at a point $x \in X$ of size at least L (see Definition 4.38), then for any $y \in \mathrm{SP}\alpha_x \subset Y$, $\phi \circ \alpha$ contains a part of point evaluation at point y of size at least $\frac{L}{\mathrm{rank}(R)}$.

$\square$

The following two theorems are important for the proof of our main theorem.

Theorem 5.32a. *Let M be a positive integer. Let $\lim\limits_{n\to\infty}(A_n = \bigoplus_{i=1}^{k_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$ be a simple inductive limit with injective connecting homomorphisms $\phi_{n,m}$ and with $\dim(X_{n,i}) \leq M$, for any n, i. Let $B = \bigoplus_{i=1}^{s} M_{l_i}(C(Y_i))$, where Y_i are the spaces: $\{pt\}, [0,1], S^1, T_{II,k}$, $T_{III,k}$, and S^2.*
Suppose that a homomorphism $\alpha : B \to A_n$ satisfies the following dichotomy condition*:*
For any block B^i of B and any block A_n^j of A_n, either the partial map $\alpha^{i,j} : B^i \to A_n^j$ is injective or it has a finite dimensional image.
Denote $\alpha(\mathbf{1}_B) := R(=\bigoplus R^i) \in A_n(=\bigoplus A_n^i)$. For any finite sets $G \subset B$ and $F \subset RA_nR$, any positive number $\varepsilon > 0$, and any positive integer L, there are A_m and mutually orthogonal projections $Q_0, Q_1, Q_2 \in A_m$, with $\phi_{n,m}(R) = Q_0 + Q_1 + Q_2$, a unital map $\theta_0 \in Map(RA_nR, Q_0A_mQ_0)_1$, two unital homomorphisms $\theta_1 \in Hom(RA_nR, Q_1A_mQ_1)_1$ and
$\xi \in Hom(RA_nR, Q_2A_mQ_2)_1$ such that
(1) $\|\phi_{n,m}(f) - (\theta_0(f) + \theta_1(f) + \xi(f))\| < \varepsilon, \quad \forall f \in F$;
(2) there is a homomorphism $\alpha_1 : B \to (Q_0 \oplus Q_1)A_m(Q_0 \oplus Q_1)$ such that

$$\|\alpha_1(g) - (\theta_0 + \theta_1) \circ \alpha(g)\| < \varepsilon, \quad \forall g \in G;$$

(3) θ_0 is $F - \varepsilon$ multiplicative and θ_1 satisfies that for any nonzero projection (including any rank 1 projection) $e \in R^iA_n^iR^i$

$$\theta_1^{i,j}([e]) \geq L \cdot [\theta_0^{i,j}(R^i)],$$

(the condition (3) will be used when we apply Theorem 1.6.9 in the proof of the Main Theorem);
(4) ξ factors through a C^*-algebra C—a direct sum of matrix algebras over $C[0,1]$ or $\mathbb{C}$— as

$$\xi : \ RA_nR \xrightarrow{\xi_1} C \xrightarrow{\xi_2} Q_2A_mQ_2,$$

and the partial maps of ξ_2 satisfy the dichotomy condition;
(5) the partial maps of α_1 satisfies the dichotomy condition.

Proof: Let $E^{i,j} = \alpha^{i,j}(\mathbf{1}_{B^i}) \in A_n^j$. Let

$$I = \{(i,j) \mid \alpha^{i,j} : B^i \to A_n^j \text{ has finite dimensional image}\}.$$

Let the subalgebra $D \subset A_n = \bigoplus A_n^i$ be defined by

$$D = \bigoplus_j \left(\bigoplus_{(i,j)\in I} \alpha^{i,j}(B^i) \oplus \bigoplus_{(i,j)\notin I} \alpha^{i,j}(\mathbb{C} \cdot \mathbf{1}_{B^i}) \right) \subset \bigoplus_j A_n^j.$$

Notice that D is a finite dimensional subalgebra of A_n containing the mutually orthogonal projections $\{E^{i,j} = \alpha^{i,j}(\mathbf{1}_{B_i})\}_{i,j}$.
Apply part 2 of Corollary 4.39 for sufficiently large set $F' \subset RA_nR$, sufficiently small number $\varepsilon' > 0$ and $\eta' > 0$, and positive integer $J = L \cdot \max_i \operatorname{rank}(R^i)$, to obtain A_m and the decomposition $\theta_0 \oplus \theta_1 \oplus \xi$ of $\phi_{n,m}|_{RA_nR}$ as $\psi_0' \oplus \psi_1' \oplus \psi_2'$ in 4.39. By Lemma 1.6.8, we can assume that the restriction $\theta_0|_D$ is a homomorphism. The condition (1) follows if we choose $F' \supset F$, and $\varepsilon' < \varepsilon$.
The $F - \varepsilon$ multiplicativity of θ_0 in (3) follows from Lemma 4.40, if F' is large enough and ε' is small enough, and the desired property of θ_1 in (3) follows from the choice of J and Lemma 5.31.
To construct α_1 as desired in the condition (2), we need to construct

$$\alpha_1^{i,j,k} : \; B^i \to \theta^{j,k}(E^{i,j}) A_m^k \theta^{j,k}(E^{i,j}),$$

where $\theta = \theta_0 \oplus \theta_1$, to satisfy

$$\|\alpha_1^{i,j,k}(g) - \theta^{j,k} \circ \alpha^{i,j}(g)\| < \varepsilon, \quad \forall g \in G.$$

The construction are divided into three cases.
1. If $(i,j) \in I$, then $\theta^{j,k} \circ \alpha^{i,j}$ is already a homomorphism and can be chosen to be $\alpha_1^{i,j,k}$.
2. If $(i,j) \notin I$, and $Y_i = [0,1]$ or S^1, then the existence of $\alpha_1^{i,j,k}$ follows from Lemma 1.6.1 and Lemma 4.40, if F' is large enough and ε' is small enough. (See Lemma 5.30 also.) In fact, in this case, the map $\theta_0^{j,k} \circ \alpha^{i,j}$ itself can be perturbed to a homomorphism. On the other hand, the homomorphism $\theta_1^{j,k} \circ \alpha^{i,j}$ is defined by the point evaluations on an η-dense set for a certain small number η. Evidently, such a homomorphism $\theta_1^{j,k} \circ \alpha^{i,j}$ from $M_{l_i}(C(S^1))$ or $M_{l_i}(C([0,1]))$ (to A_m^k) can be perturbed to an injective homomorphism, provided that η is sufficiently small and that the path connected simplicial complex $X_{m,k}$ is not the space of a single point. Therefore, in this case, the homomorphism $\alpha_1^{i,j,k}$ can be chosen to be injective.
3. If $(i,j) \notin I$, and $Y_i = T_{II,k}, T_{III,k}$ or S^2, then $\alpha^{i,j}$ is injective, and the existence of $\alpha_1^{i,j,k}$ follows from Lemma 5.30 and the choice of J, if F' is large enough and ε' is small enough, and if we choose η' to be the number η_1 in Lemma 5.31 corresponding to the η in Lemma 5.30. The homomorphism $\alpha_1^{i,j,k}$ can also be chosen to be injective, if $X_{m,k}$ is not the space of a single point, according to the last part of Lemma 5.25.
Finally, define the partial map $\alpha_1^{i,k}$ of α_1 to be $\bigoplus_j \alpha_1^{i,j,k}$ to complete the construction. Obviously, it follows, from the discussion of the injectivity in case 2 and case 3, that α_1 satisfies the dichotomy condition.

$\square$

Theorem 5.32b. *Let M be a positive integer. Let $\lim\limits_{n\to\infty}(A_n = \bigoplus_{i=1}^{k_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$ be a simple inductive limit with injective connecting homomorphisms $\phi_{n,m}$ and with $\dim(X_{n,i}) \le M$, for any n, i. Let $B = \bigoplus_{i=1}^{s} M_{l_i}(C(Y_i))$, where Y_i are the spaces:$\{pt\}, [0,1], S^1, T_{II,k}$, $T_{III,k}$, and S^2.*

Suppose that a homomorphism $\alpha : B \to A_n$ satisfies the following dichotomy condition*:*

For any block B^i of B and any block A_n^j of A_n, either the partial map $\alpha^{i,j}$: $B^i \to A_n^j$ is injective or it has a finite dimensional image.

For any finite sets $G \subset B$ and $F \subset A_n$, and any number $\varepsilon > 0$, there are A_m and mutually orthogonal projections $P, Q \in A_m$, with $\phi_{n,m}(\mathbf{1}_{A_n}) = P + Q$, a unital map $\theta \in Map(A_n, PA_mP)_1$, and a unital homomorphism $\xi \in Hom(A_n, QA_mQ)_1$ such that

(1) $\|\phi_{n,m}(f) - (\theta(f) \oplus \xi(f))\| < \varepsilon, \quad \forall f \in F$;

(2) there is a homomorphism $\alpha_1 : B \to PA_mP$ such that

$$\|\alpha_1(g) - (\theta \circ \alpha)(g)\| < \varepsilon, \quad \forall g \in G;$$

(3) $\theta(F)$ is weakly approximately constant to within ε;

(4) ξ factors through a C^*-algebra C—a direct sum of matrix algebras over $C[0,1]$ or $\mathbb{C}$— as

$$\xi : \ A_n \xrightarrow{\xi_1} C \xrightarrow{\xi_2} QA_mQ,$$

and the partial maps of ξ_2 satisfy the dichotomy condition;

(5) the partial maps of α_1 satisfy the dichotomy condition.

The proof is similar to the proof of Theorem 5.32a, we omit it.

6 The proof of the main theorem

In this section, we will combine §4, §5 and §1.6 to prove our Main Theorem — the Reduction Theorem.

The following is Proposition 3.1 of [D2].

Proposition 6.1. ([D2, 3.1]) *Consider the diagram*

$$\begin{array}{ccccccccccc}
A_1 & \xrightarrow{\phi_{1,2}} & A_2 & \xrightarrow{\phi_{2,3}} & \cdots & \longrightarrow & A_n & \xrightarrow{\phi_{n,n+1}} & A_{n+1} & \longrightarrow & \cdots \\
\uparrow{\scriptstyle \alpha_1} & \searrow^{\beta_1} & \uparrow{\scriptstyle \alpha_2} & \searrow^{\beta_2} & \cdots & & \uparrow{\scriptstyle \alpha_n} & \searrow^{\beta_n} & \uparrow{\scriptstyle \alpha_{n+1}} & \searrow & \\
B_1 & \xrightarrow{\psi_{1,2}} & B_2 & \xrightarrow{\psi_{2,3}} & \cdots & \longrightarrow & B_n & \xrightarrow{\psi_{n,n+1}} & B_{n+1} & \longrightarrow & \cdots
\end{array} ,$$

where A_n, B_n are C^*-algebras, $\phi_{n,n+1}, \psi_{n,n+1}$ are homomorphisms and α_n, β_n are linear $*$-contractions.

Suppose that $F_n \subset A_n,\ E_n \subset B_n$ are finite sets satisfying the following conditions.

$$\phi_{n,n+1}(F_n) \cup \alpha_{n+1}(E_{n+1}) \subset F_{n+1}, \quad \psi_{n,n+1}(E_n) \cup \beta_n(F_n) \subset E_{n+1},$$

and $\overline{\bigcup_{n=1}^{\infty}(\phi_{n,\infty}(F_n))}$ and $\overline{\bigcup_{n=1}^{\infty}(\psi_{n,\infty}(E_n))}$ are the unit balls of $A = \lim(A_n, \phi_{n,m})$ and $B = \lim(B_n, \psi_{n,m})$, respectively. Suppose that there is a sequence $\varepsilon_1, \varepsilon_2, \cdots$ of positive numbers with $\sum \varepsilon_n < +\infty$ such that α_n and β_n are F_n-ε_n multiplicative and E_n-ε_n multiplicative, respectively, and

$$\|\phi_{n,n+1}(f) - \alpha_{n+1} \circ \beta_n(f)\| < \varepsilon_n, \quad \text{and} \quad \|\psi_{n,n+1}(g) - \beta_n \circ \alpha_n(g)\| < \varepsilon_n$$

for all $f \in F_n$ and $g \in E_n$.
Then A is isomorphic to B.

Lemma 6.2. *Let $\lim(A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$ be a simple inductive limit C^*-algebra with $\phi_{n,m}$ injective, where $X_{n,i}$ are path connected finite simplicial complexes with uniformly bounded dimensions. Let C^*-algebra C be a direct sum of matrix algebras over the spaces: $\{pt\}, [0,1], S^1, T_{II,k}, T_{III,k}$ and S^2, and $\phi:\ C \to A_n$ be an injective homomorphism. Then for any finite set $F \subset C$ and $\varepsilon > 0$, there is a positive integer $N > n$ such that for any $m > N$, there is a homomorphism $\psi : C \to A_m$ satisfying the following conditions.*
(1) $\psi(\mathbf{1}_{C^i}) = (\phi_{n,m} \circ \phi)(\mathbf{1}_{C^i})$, for any block C^i of C.
(2) $\|\psi(f) - (\phi_{n,m} \circ \phi)(f)\| < \varepsilon, \quad \forall f \in F$.
(3) ψ satisfies the following dichotomy condition:
For any block C^i of C and A_m^j of A_m, either $\psi^{i,j}$ is injective or $\psi^{i,j}$ has finite dimensional image.
(For the proof of the main theorem of this article—Theorem 6.3 below, we only need this lemma for the case that C is a direct sum of matrix algebras over spaces $\{pt\}$ and $[0,1]$. The full generality of the lemma will be used in the proof of Corollary 6.11 below.)

Proof: We only need to prove for the case that C has only one block $C = M_k(C(X))$. And, by the discussion in 1.2.19, this case can further be reduced to the case $C = C(X)$.
For the finite set $F \subset C$, there is an $\eta > 0$ such that if $\mathrm{dist}(t, t') < 4\eta$, then

$$\|f(t) - f(t')\| < \varepsilon, \quad \forall f \in F.$$

Let (X, σ) be a simplicial decomposition of X such that for any simplex $\Delta \subset (X, \sigma), \mathrm{diameter}(\Delta) < \eta$. We call a simplex Δ a top simplex if Δ is not a proper face of any simplex. Obviously, Δ is a top simplex if and only if the interior $\mathring{\Delta}$ is an open subset of X.

From [DNNP, Proposition 2.1], using the injectivity of ϕ and $\phi_{n,m}$, it follows that there is an integer $N > n$ such that for any open set $\mathring{\Delta}$ — the interior of a top simplex $\Delta \subset (X, \sigma)$, one has

$$SP(\phi_{n,m} \circ \phi)_y \cap \overset{\circ}{\Delta} \neq \emptyset.$$

We can define the homomorphism $\psi : C \to A_m$ for each block A_m^j of A_m separately. That is, we need to define $\psi^j : C \to A_m^j$, then let $\psi := \oplus \psi^j$.
If $\mathrm{SP}(A_m^j) = X_{m,j} = \{pt\}$, then the partial map $(\phi_{n,m} \circ \phi)^j$ has finite dimensional image, and we can define it to be ψ^j. Hence we assume that the connected finite simplicial complex $X_{m,j}$ is not the space of single point $\{pt\}$.
Let $\alpha = (\phi_{n,m} \circ \phi)^j : C \to A_m^j$.
Let Y be the union of all such top simplices Δ that $\overset{\circ}{\Delta} \cap \mathrm{SP}\alpha$ is uncountable.
Let Z be the union of all simplices Δ which are not top simplices. Both Y and Z are closed subset of X. Let $\Delta_1, \Delta_2, \cdots, \Delta_l$ be the list of all top simplices such that $\Delta_i \not\subset Y, i = 1, 2, \cdots, l$. Then

$$X = Y \cup \Delta_1 \cup \Delta_2 \cdots \cup \Delta_l.$$

(This fact will be used later.) (Here we use the fact that X is equal to the union of all top simplices, since each simplex is a face of a top simplex.)
For each Δ_i, $\overset{\circ}{\Delta}_i \cap \mathrm{SP}\alpha$ is a countable nonempty set. There is a point x_i and an open disk $U_i = B_{\varepsilon_i}(x_i) \ni x_i$ such that

$$\mathrm{SP}\alpha \cap U_i = \{x_i\}.$$

We can assume that $U_i \subset \overset{\circ}{\Delta}_i$. Obviously, $\partial\Delta_i$ is a deformation retract of $\Delta_i \backslash U_i$.
Set $\left(X \backslash \left(\cup_{i=1}^l U_i\right)\right) \cap \mathrm{SP}\alpha = T$. Then $\mathrm{SP}\alpha = T \cup \{x_1, x_2, \cdots, x_l\}$.
Define a function $g : T \to Y \cup Z (\subset X)$ as below.
Let $g' : Z \to Y \cup Z$ be the identity map, that is,

$$g'(z) = z, \qquad \forall z \in Z.$$

We will extend the map g' to a map (let us still denote it by g').

$$g' : X \backslash \left(\cup_{i=1}^l U_i\right) \longrightarrow Y \cup Z.$$

For each top simplex $\Delta \subset Y$, extend $g'|_{\partial\Delta}$ to a map $g' : \Delta \to \Delta$ satisfying

$$g'(T \cap \Delta) = \Delta.$$

(Such extension exists since $T \cap \overset{\circ}{\Delta}$ is uncountable, see Lemma 2.6 of [EGL].)
For any simplex Δ_i, $i = 1, 2, \cdots, l$, one can extend $g'|_{\partial\Delta}$ to a map $g' : \Delta_i \backslash U_i \to \partial\Delta_i$, since $\partial\Delta_i$ is a deformation retract of $\Delta_i \backslash U_i$.

Thus we obtain the extension $g' : X\backslash \left(\cup_{i=1}^{l} U_i\right) \longrightarrow Y \cup Z$. Let $g = g'|_T$. Then

$$g(T) \supset Y, \quad \text{ and } \quad \mathrm{dist}(g(x), x) < \eta, \ \ \forall x \in T.$$

Since $\mathrm{SP}\alpha = T \cup \{x_1, x_2, \cdots, x_l\}$, there are homomorphisms $\alpha_0 : C(T) \to A_m^j$ and $\alpha_i : \mathbb{C} = C(\{x_i\}) \to A_m^j$, $j = 1, 2, \cdots, l$, with mutually orthogonal images, such that

$$\alpha(f) = \alpha_0(f|_T) + \sum_{i=1}^{l} (f|_{\{x_i\}}), \qquad \forall f \in C(X).$$

Define $\beta_0 : C(Y \cup Z) \to A_m^j$ by

$$\beta_0(f) = \alpha_0(f \circ g), \qquad \forall f \in C(Y \cup Z),$$

where $g : T \to Y \cup Z$ is defined as above. For each Δ_i, there is a surjective map $g_i : X_{m,j} \to \Delta_i$, since $X_{m,j} \neq \{pt\}$. Define $\beta_i : C(\Delta_i) \to A_m^j$ by

$$\beta_i(f)(x) = f(g_i(x)) \cdot \alpha_i(\mathbf{1}_{\mathbb{C}}), \qquad \forall f \in C(\Delta_i), x \in X_{m,j}.$$

Then, obviously, we have
$(1')$ $\beta_0(\mathbf{1}_{C(Y\cup Z)}) = \alpha_0(\mathbf{1}_{C(T)})$, and $\beta_i(\mathbf{1}_{C(\Delta_i)}) = \alpha_i(\mathbf{1}_{\mathbb{C}})$, for $i = 1, 2, \cdots, l$.
From the way η is chosen and the properties that $\mathrm{dist}(g(x), x) < \eta$ for any $x \in T$, and that diameter$(\Delta_i) < \eta$ for any $i = 1, 2, \cdots, l$, we have
$(2')$ $\|\beta_0(f|_{Y\cup Z}) - \alpha_0(f|_T)\| < \varepsilon$, and $\|\beta_i(f|_{\Delta_i}) - \alpha_i(f|_{\{x_i\}})\| < \varepsilon$ for $i = 1, 2, \cdots, l$, and $f \in F$.
Finally, let the partial homomorphism $\psi^j : C(X) \to A_m^j$ be defined by

$$\psi^j(f) = \beta_0(f|_{Y\cup Z}) + \sum_{i=1}^{l} \beta_i(f|_{\Delta_i}).$$

Since $T \subset \mathrm{SP}\alpha_0$ and the map $g : T \to Y \cup Z$ satisfies $g(T) \supset Y$, we have $\mathrm{SP}(\beta_0) \supset Y$. Hence $\mathrm{SP}\psi^j = \mathrm{SP}\beta_0 \cup \cup_{i=1}^{l} \mathrm{SP}\beta_i \supset Y \cup \cup_{i=1}^{l} \Delta_i = X$. That is, ψ^j is injective.
The property (1) follows from $(1')$ and (2) follows from $(2')$.

$\square$

We will use 5.32a, 5.32b, 1.6.9, 1.6.29, 1.6.30 to prove the following main theorem of this article.

Theorem 6.3. *Suppose that* $\lim(A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$ *is a simple inductive limit* C^**-algebra with* $\dim(X_{n,i}) \leq M$ *for a fixed positive integer* M*. Then there is another inductive system* $(B_n = \bigoplus_{i=1}^{t_n} M_{\{n,i\}}(C(Y_{n,i})),\ \phi_{n,m})$ *with the same limit algebra as the above system, where all* $Y_{n,i}$ *are spaces of forms* $\{pt\}, [0,1], S^1, S^2, T_{II,k}$*, or* $T_{III,k}$*.*

Proof: Without loss of generality, assume that the spaces $X_{n,i}$ are connected finite simplicial complexes and the connecting maps $\phi_{n,m}$ are injective (see Theorem 4.23).

Let $\varepsilon_1 > \varepsilon_2 > \varepsilon_3 > \cdots > 0$ be a sequence of positive numbers satisfying $\sum \varepsilon_n < +\infty$.

We need to construct the intertwining diagram

$$
\begin{array}{cccccccccc}
F_1 & & F_2 & & & F_n & & F_{n+1} & \\
\bigcap & & \bigcap & & & \bigcap & & \bigcap & \\
A_{s(1)} & \xrightarrow{\phi_{s(1),s(2)}} & A_{s(2)} & \xrightarrow{\phi_{s(2),s(3)}} & \cdots \longrightarrow & A_{s(n)} & \xrightarrow{\phi_{s(n),s(n+1)}} & A_{s(n+1)} & \longrightarrow \cdots \\
\uparrow{\scriptstyle \alpha_1} & \searrow^{\beta_1} & \uparrow{\scriptstyle \alpha_2} & \searrow^{\beta_2} & & \uparrow{\scriptstyle \alpha_n} & \searrow^{\beta_n} & \uparrow{\scriptstyle \alpha_{n+1}} & \searrow \\
B_1 & \xrightarrow{\psi_{1,2}} & B_2 & \xrightarrow{\psi_{2,3}} & \cdots \longrightarrow & B_n & \xrightarrow{\psi_{n,n+1}} & B_{n+1} & \longrightarrow \cdots \\
\bigcup & & \bigcup & & & \bigcup & & \bigcup & \\
E_1 & & E_2 & & & E_n & & E_{n+1} &
\end{array}
$$

satisfying the following conditions.

(0.1) $(A_{s(n)}, \phi_{s(n),s(m)})$ is a sub-inductive system of $(A_n, \phi_{n,m})$. $(B_n, \psi_{n,m})$ is an inductive system of matrix algebras over the spaces: $\{pt\}, [0,1], S^1$, $\{T_{II,k}\}_{i=2}^{\infty}, \{T_{III,k}\}_{i=2}^{\infty}, S^2$.

(0.2) Choose $\{a_{ij}\}_{j=1}^{\infty} \subset A_{s(i)}$ and $\{b_{ij}\}_{j=1}^{\infty} \subset B_i$ to be countable dense subsets of the unit balls of $A_{s(i)}$ and B_i, respectively. F_n are subsets of the unit balls of $A_{s(n)}$, and E_n are subsets of the unit balls of B_n satisfying

$$\phi_{s(n),s(n+1)}(F_n) \cup \alpha_{n+1}(E_{n+1}) \cup \bigcup_{i=1}^{n+1} \phi_{s(i),s(n+1)}(\{a_{i1}, a_{i2}, \cdots, a_{i\ n+1}\}) \subset F_{n+1}$$

and

$$\psi_{n,n+1}(E_n) \cup \beta_n(F_n) \cup \bigcup_{i=1}^{n+1} \psi_{i,n+1}(\{b_{i1}, b_{i2}, \cdots, b_{i\ n+1}\}) \subset E_{n+1}.$$

(Here we use the convention that $\phi_{n,n} = id : A_n \to A_n$.)

(0.3) β_n are $F_n - 2\varepsilon_n$ multiplicative and α_n are homomorphisms.

(0.4) $\|\psi_{n,n+1}(g) - \beta_n \circ \alpha_n(g)\| < 2\varepsilon_n$ for all $g \in E_n$, and $\|\phi_{s(n),s(n+1)}(f) - \alpha_{n+1} \circ \beta_n(f)\| < 12\varepsilon_n$ for all $f \in F_n$.

(0.5) For any block B_n^i of B_n and any block $A_{s(n)}^j$ of $A_{s(n)}$, the map $\alpha_n^{i,j}$ satisfies the following dichotomy condition:

either $\alpha_n^{i,j}$ is injective or $\alpha_n^{i,j}$ has a finite dimensional image.

The diagram will be constructed inductively.

First, let $B_1 = \{0\}$, $A_{s(1)} = A_1$, $\alpha_1 = 0$. Let $b_{1j} = 0 \in B_1$ for $j = 1, 2, \cdots$, and let $\{a_{1j}\}_{j=1}^{\infty}$ be a countable dense subset of the unit ball of $A_{s(1)}$. And let $E_1 = \{b_{11}\} = B_1$ and $F_1 = \{a_{11}\} \subset A_{s(1)}$.

As an inductive assumption, assume that we already have the diagram

$$\begin{array}{ccccccc}
F_1 & & F_2 & & & & F_n \\
\bigcap & & \bigcap & & & & \bigcap \\
A_{s(1)} & \xrightarrow{\phi_{s(1),s(2)}} & A_{s(2)} & \xrightarrow{\phi_{s(2),s(3)}} & \cdots & \longrightarrow & A_{s(n)} \\
\Big\uparrow{\scriptstyle \alpha_1} & \searrow^{\beta_1} & \Big\uparrow{\scriptstyle \alpha_2} & \searrow^{\beta_2} & \cdots & \searrow^{\beta_{n-1}} & \Big\uparrow{\scriptstyle \alpha_n} \\
B_1 & \xrightarrow{\psi_{1,2}} & B_2 & \xrightarrow{\psi_{2,3}} & \cdots & \longrightarrow & B_n \\
\bigcup & & \bigcup & & & & \bigcup \\
E_1 & & E_2 & & & & E_n
\end{array}$$

and, for each $i = 1, 2\cdots, n$, we have countable dense subsets $\{a_{ij}\}_{j=1}^{\infty} \subset$ unit ball of $A_{s(i)}$ and $\{b_{ij}\}_{j=1}^{\infty} \subset$ unit ball of B_i to satisfy the conditions (0.1)-(0.5) above. We have to construct the next piece of the diagram,

$$\begin{array}{cccccc}
F_n \subset & A_{s(n)} & \xrightarrow{\phi_{s(n),s(n+1)}} & A_{s(n+1)} & \supset F_{n+1} & \\
 & \Big\uparrow{\scriptstyle \alpha_n} & \searrow^{\beta_n} & \Big\uparrow{\scriptstyle \alpha_{n+1}} & & \\
E_n \subset & B_n & \xrightarrow{\psi_{n,n+1}} & B_{n+1} & \supset E_{n+1} & ,
\end{array}$$

to satisfy the conditions (0.1)-(0.5).

Our construction are divided into several steps. In order to provide the reader with a whole picture of the construction, we first give an outline of it. Then the detailed construction will follow.

Outline of the construction. We will construct the following diagram.

$$\begin{tikzcd}
C \arrow[r, "\xi_2"] & P_1A_{m_1}P_1 \arrow[rrr, "\phi_{m_1,s(n+1)}"] & & & \phi_{m_1,s(n+1)}(P_1)A_{s(n+1)}\phi_{m_1,s(n+1)}(P_1) \\
 & \approx\phi_{s(n),m_1} \quad \bigoplus & & & \bigoplus \\
 & & RA_{m_2}R \arrow[r, "\xi_3"] \arrow[ddrr, dashed, "\theta_0+\theta_1"'] & D \arrow[r, "\xi_4"] & (Q_2)A_{s(n+1)}(Q_2) \\
 & & \mathrm{u} & & \approx\phi_{m_2,s(n+1)} \quad \bigoplus \\
A_{s(n)} \arrow[uuuu, "\xi_1"] \arrow[r, dashed, "\theta"'] & P_0A_{m_1}P_0 \arrow[uur, "\phi_{m_1,m_2}"] \arrow[d, dashed, "\beta"'] & RA_{m_2}R \arrow[rr, dashed, "\theta_0+\theta_1"] & & (Q_0+Q_1)A_{s(n+1)}(Q_0+Q_1) \arrow[loop below, "Adu"] \\
B_n \arrow[u, "\alpha_n"] \arrow[ur, "\alpha"] \arrow[r, "\psi"'] \arrow[u, phantom, shift left=6, "1"] \arrow[ur, phantom, shift right=3, "2"'] & B \arrow[ur, "\lambda\circ\alpha'"] \arrow[urrr, "\alpha''"'] \arrow[ur, phantom, shift right=6, "3"'] & & &
\end{tikzcd}$$

This large picture consists of several smaller diagrams, each of which is called a sub-diagram. There are two kinds of sub-diagrams. The sub-diagrams of the first kind are labeled by the numbers 1, 2, 3 and the letter u (in the centers of the sub-diagrams). These sub-diagrams are almost commutative in some sense. For example, the one in the center of the large picture, labeled by the letter u consists of two composite maps $(\theta_0 + \theta_1) \circ (\lambda \circ \alpha') \circ \beta$ and $(\theta_0 + \theta_1) \circ (\phi_{m_1,m_2}|_{P_0A_{m_1}P_0})$. They are almost equal to each other on a given finite set up to unitary equivalence.

The sub-diagrams of the second kind are those two labeled by "$\approx \phi_{s(n),m_1}$" and "$\approx \phi_{m_2,s(n+1)}$". They describe the approximate decompositions of the given maps "$\phi_{s(n),m_1}$" and "$\phi_{m_2,s(n+1)}|_{RA_{m_2}R}$".

All the maps in the above picture are homomorphisms except β, θ, and $\theta_0 + \theta_1$ (which are represented by broken line arrows). These maps are linear $*$-contractions which are almost multiplicative on some given finite sets (i.e., on the sets $F_n \subset A_{s(n)}$, $F := \theta(F_n) \subset P_0 A_{m_1} P_0$, or a certain (large enough) finite subset $F' \subset RA_{m_2}R_2$) to within given small numbers (i.e., ε_n or some related small numbers).

The sub-diagrams labeled by the numbers $1, 2$, and 3 are approximately commutative on certain given finite sets (i.e., $E_n \subset B_n$, $G := \psi(E_n) \cup \beta(F) \subset B$ (F is from the above paragraph)) to within a small number (i.e., ε_n). The sub-diagram labeled by the letter u is approximately commutative on a finite set ($F := \theta(F_n)$) to within a given small number ($9\varepsilon_n$) up to unitary equivalence.

The sub-diagrams labeled by "$\approx \phi_{s(n),m_1}$" and "$\approx \phi_{m_2,s(n+1)}$" are approximate decompositions of $\phi_{s(n),m_1}$ and $\phi_{m_2,s(n+1)}|_{RA_{m_2}R}$, respectively. (E.g., the direct sum $\theta \oplus (\xi_2 \circ \xi_1)$ of the two maps θ and $\xi_2 \circ \xi_1$ is close to $\phi_{s(n),m_1}$ to within a small number ε_n on a given finite set F_n.)

The above decomposition of $\phi_{s(n),m_1}$ and the almost commutative sub-diagram labeled by the number 1 are obtained in Step 1 in the detailed proof, applying Theorem 5.32b to $A_{s(n)}$ and $\alpha_n : B_n \to A_{s(n)}$ (and to the finite sets E_n and F_n). The main purpose of this step is to make the set $\theta(F_n) := F$ weakly approximately constant to within ε_n (the other part $\xi_2 \circ \xi_1$ of the decomposition factors through an interval algebra C), which will be useful later when we apply Theorem 1.6.9. (If one assumes in the beginning that the set F_n is weakly approximately constant to within ε_n, then he does not need this step.)

The sub-diagrams labeled by 2 or u will be explained by another picture later. The almost commutative sub-diagram labeled by the number 3 and the decomposition of $\phi_{m_2,s(n+1)}|_{RA_{m_2}R}$ (i.e., "$\approx \phi_{m_2,s(n+1)}$" in the picture), are obtained in Step 4, applying Theorem 5.32a to $RA_{m_2}R$ and $\lambda \circ \alpha' : B \to RA_{m_2}R$ (and certain finite subsets of B and $RA_{m_2}R$). The purpose of applying Theorem 5.32a is to construct the map $\theta_0 + \theta_1$ to satisfy the condition in Theorem 1.6.9 for the two homotopic homomorphisms $\lambda \circ (\phi \oplus r)$ and $\phi_{m_1,m_2}|_{P_0 A_{m_1} P_0}$ in the next picture, and therefore to obtain the almost commutative sub-diagram up to unitary equivalence—the sub-diagram labeled by u—, (the other part $\xi_4 \circ \xi_3$ of the decomposition factors through an interval algebra D).

In order to get the parts of the sub-diagram labeled by 2 and u, we need to start with $\alpha : B_n \to P_0 A_{m_1} P_0$. We describe it in the next picture.

$$\begin{tikzcd}
 & RA_{m_2}R \arrow[r, dashed, "\theta_0+\theta_1"] & (Q_0+Q_1)A_{s(n+1)}(Q_0+Q_1) \arrow[d, phantom, "Adu"] \\
A(=P_0A_{m_1}P_0) \arrow[ur, "\phi_{m_1,m_2}"] \arrow[r, "\phi\oplus r"] \arrow[dr, dashed, "\beta"] \arrow[r, phantom, "h", shift left=4ex] & M_L(A)\oplus r(A) \arrow[u, "\lambda"'] & {} \\
B_n \arrow[u, "\alpha"] \arrow[r, "\psi"'] \arrow[r, phantom, "2", shift left=5ex, near start] & B \arrow[u, "\alpha'"'] \arrow[u, phantom, "4", shift left=5ex] &
\end{tikzcd}$$

By Corollary 1.6.29 (see 1.6.31 also), applied to the homomorphism α from the first picture, we obtain the almost commutative sub-diagrams labeled by the numbers 2 and 4. Then we apply Lemma 1.6.30 to obtain the sub-diagram labeled by the letter h which commutes up to homotopy equivalence. By Theorem 1.6.9 and the property of the map $\theta_0+\theta_1$ (from Theorem 5.32a), this sub-diagram leads to the sub-diagram labeled by u in the first picture.
With the first picture in mind, we define

$$B_{n+1} = C\oplus B\oplus D,$$

$$\psi_{n,n+1} = (\xi_1\circ\alpha_n)\oplus\psi\oplus(\xi_3\circ\phi_{m_1,m_2}|_{P_0A_{m_1}P_0}\circ\alpha),$$
$$\beta_n = \xi_1\oplus(\beta\circ\theta)\oplus(\xi_3\circ\phi_{m_1,m_2}|_{P_0A_{m_1}P_0}\circ\theta),$$

and

$$\alpha_{n+1} = (\phi_{m_1,s(n+1)}|_{P_1A_{m_1}P_1}\circ\xi_2)\oplus(\mathrm{Ad}u\circ\alpha'')\oplus\xi_4.$$

In the definitions of $\psi_{n,n+1}$ and α_{n+1}, we use solid line arrows only since these maps are supposed to be homomorphisms (but in the definition of β_n, we can use broken line arrows).
One can easily verify the conditions (0.1)–(0.5) except that the map $\phi_{m_1,s(n+1)}|_{P_1A_{m_1}P_1}\circ\xi_2$ may not automatically satisfy the dichotomy condition (0.5), for which we have to apply Lemma 6.2 to make some modification.
Details of the construction. The above outline can be used as a guide to understand the following construction. But the proof below is complete by itself. (We encourage readers to compare the following detailed proof with the two diagrams in the outline.)
Among the conditions in the induction assumption, only the dichotomy condition (0.5) of α_n is used in the following construction.
Step 1. By Theorem 5.32b, applied to $\alpha_n: B_n\to A_{s(n)}, E_n\subset B_n$, $F_n\subset A_{s(n)}$, and $\varepsilon>0$, there are A_{m_1} ($m_1>s(n)$), two orthogonal projections $P_0,P_1\in A_{m_1}$ with $\phi_{s(n),m_1}(\mathbf{1}_{A_{s(n)}}) = P_0+P_1$ and P_0 trivial, a C^*-algebra C — a direct sum of matrix algebras over $C[0,1]$ or $\mathbb{C}$—, a unital map $\theta\in\mathrm{Map}(A_{s(n)},P_0A_{m_1}P_0)_1$,

a unital homomorphism $\xi_1 \in \mathrm{Hom}(A_{s(n)}, C)_1$, an injective unital homomorphism $\xi_2 \in \mathrm{Hom}(C, P_1 A_{m_1} P_1)_1$ and a (not necessarily unital) homomorphism $\alpha \in \mathrm{Hom}(B_n, P_0 A_{m_1} P_0)$ such that
(1.1) $\|\phi_{s(n),m_1}(f) - \theta(f) \oplus (\xi_2 \circ \xi_1)(f)\| < \varepsilon_n$ for all $f \in F_n$.
(1.2) θ is F_n-ε_n multiplicative and $F := \theta(F_n)$ is weakly approximately constant to within ε_n.
(1.3) $\|\alpha(g) - \theta \circ \alpha_n(g)\| < \varepsilon_n$ for all $g \in E_n$.
(1.4) Both $\alpha : B_n \to P_0 A_{m_1} P_0$ and $\xi_2 : C \to P_1 A_{m_1} P_1$ satisfy the dichotomy condition in (0.5).
(Thus we finished the construction of the sub-diagrams labeled by the number "1" and "$\approx \phi_{s(n),m_1}$" of the large diagram in the outline.)
Let all the blocks of C be parts of C^*-algebra B_{n+1}. That is,

$$B_{n+1} = C \oplus (\text{some other blocks}).$$

The map $\beta_n : A_{s(n)} \to B_{n+1}$ and the homomorphism $\psi_{n,n+1} : B_n \to B_{n+1}$ are defined by

$$\beta_n = \xi_1 : A_{s(n)} \to C(\subset B_{n+1}) \quad \text{and} \quad \psi_{n,n+1} = \xi_1 \circ \alpha_n : B_n \to C\ (\ \subset B_{n+1})$$

for the blocks of C ($\subset B_{n+1}$). For this part, β_n is also a homomorphism.
Step 2. Let $A = P_0 A_{m_1} P_0,\ F = \theta(F_n)$. Since P_0 is a trivial projection,

$$A \cong \bigoplus M_{l_i}(C(X_{m_1,i})).$$

Let $rA := \bigoplus M_{l_i}(\mathbb{C}) \subset A$, and $r : A \to rA$ be the homomorphism defined by evaluation at certain base points $x_i^0 \in X_{m_1,i}$ (see 1.1.7(h)).
Applying Corollary 1.6.29 (see Remark 1.6.31 also) to $\alpha : B_n \to A$ (notice that α satisfies the dichotomy condition), $E_n \subset B_n$ and $F \subset A$, we obtain the following diagram:

$$\begin{array}{ccc} A & \xrightarrow{\phi\oplus r} & M_L(A) \oplus r(A) \\ \uparrow{\scriptstyle\alpha} & \searrow^{\beta} & \uparrow{\scriptstyle\alpha'} \\ B_n & \xrightarrow{\psi} & B \end{array}$$

such that
(2.1) B is a direct sum of matrix algebras over $\{pt\}, [0,1], S^1, T_{II,k}, T_{III,k}$, or S^2.
(2.2) α' is an injective homomorphism, and β is an F-ε_n multiplicative map.
(2.3) $\phi : A \to M_L(A)$ is a unital simple embedding. $r : A \to r(A)$ is the homomorphism defined by evaluations as in 1.1.7(h).
(2.4) $\|\beta \circ \alpha(g) - \psi(g)\| < \varepsilon_n$ for all $g \in E_n$, $\|(\phi \oplus r)(f) - \alpha' \circ \beta(f)\| < \varepsilon_n$ for all $f \in F(:= \theta(F_n))$.
(Thus we finished the construction of the sub-diagrams labeled by the number "2" and "4" of the second diagram in the outline.)
Let all the blocks B be also parts of B_{n+1}, that is,

$$B_{n+1} = C \oplus B \oplus (\text{some other blocks}).$$

The maps $\beta_n : A_{s(n)} \to B_{n+1}, \psi_{n,n+1} : B_n \to B_{n+1}$ are defined by

$$\beta_n := \beta \circ \theta : \ A_{s(n)} \xrightarrow{\theta} A \xrightarrow{\beta} B \ \ (\subset B_{n+1})$$

and

$$\psi_{n,n+1} := \psi : \ B_n \longrightarrow B \ \ (\subset B_{n+1})$$

for the blocks of $B(\subset B_{n+1})$. This part of β_n is F_n-$2\varepsilon_n$ multiplicative, since θ is F_n-ε_n multiplicative, β is F-ε_n multiplicative, and $F = \theta(F_n)$.

Step 3. By the simplicity of $\lim(A_n, \phi_{n,m})$, for m large enough, the homomorphism $\phi_{m_1,m}|_{P_0 A_{m_1} P_0}$ is $4M$-large in the sense of 1.6.16. By Lemma 1.6.30, applied to $\phi \oplus r : A \to M_L(A) \oplus r(A)$, there is an A_{m_2} and unital homomorphism $\lambda : M_L(A) \oplus r(A) \to RA_{m_2}R$, where $R = \phi_{m_1,m_2}(P_0)$ (write R as $\bigoplus_j R^j \in \bigoplus_j A^j_{m_i}$) such that the diagram

$$\begin{tikzcd} & & RA_{m_2}R \\ A(= P_0 A_{m_1} P_0) \arrow[urr, "\phi_{m_1,m_2}"] \arrow[rr, "\phi \oplus r"] & & M_L(A) \oplus r(A) \arrow[u, "\lambda"'] \end{tikzcd}$$

satisfies the following conditions:

(3.1) For each block $A^j_{m_2}$, the partial map

$$\lambda^{\cdot,j} : \ M_L(A) \oplus r(A) \longrightarrow R^j A^j_{m_2} R^j$$

is non zero. Furthermore, either it is injective or it has finite dimensional image — depending on whether $\mathrm{SP}(A^j_{m_2})$ is a single point space.

(3.2) $\lambda \circ (\phi \oplus r)$ is homotopy equivalent to

$$\phi' := \phi_{m_1,m_2}|_A.$$

(Thus we finished the construction of the sub-diagram labeled by the letter "h" of the second diagram in the outline.)

Step 4. Applying Theorem 1.6.9 to the finite set $F \subset A$ (which is weakly approximately constant to within ε_n), and to two homotopic homomorphisms

$$\phi' \ \text{ and } \ \lambda \circ (\phi \oplus r) : \ A \longrightarrow RA_{m_2}R$$

(with $RA_{m_2}R$ in place of C), we obtain a finite set $F' \subset RA_{m_2}R$, $\delta > 0$ and $L > 0$ as in the Theorem 1.6.9.

Let $G := \psi(E_n) \cup \beta(F) \subset B$. By Theorem 5.32a, applied to $RA_{m_2}R$,

$$\lambda \circ \alpha' : B \longrightarrow RA_{m_2}R$$

(which satisfies the dichotomy condition by (2.2) and (3.1)), finite sets $G \subset B$, $F' \subset RA_{m_2}R$, $\min(\varepsilon_n, \delta) > 0$ (in place of ε), and $L > 0$, there are $A_{s(n+1)}$, mutually orthogonal projections $Q_0, Q_1, Q_2 \in A_{s(n+1)}$ with $\phi_{m_2,s(n+1)}(R) = Q_0 + Q_1 + Q_2$, a C^*-algebra D — a direct sum of matrix algebras over $C[0,1]$—, a unital map $\theta_0 \in \mathrm{Map}(RA_{m_2}R, Q_0A_{s(n+1)}Q_0)$ and four unital homomorphisms $\theta_1 \in \mathrm{Hom}(RA_{m_2}R, Q_1A_{s(n+1)}Q_1)_1$, $\xi_3 \in \mathrm{Hom}(RA_{m_2}R, D)_1, \xi_4 \in \mathrm{Hom}(D, Q_2A_{s(n+1)}Q_2)_1$, and $\alpha'' \in Hom(B, (Q_0+Q_1)A_{s(n+1)}(Q_0+Q_1))_1$ such that the following are true:
(4.1) $\|\phi_{m_2,s(n+1)}(f) - ((\theta_0 + \theta_1) \oplus (\xi_4 \circ \xi_3))(f)\| < \varepsilon_n$ for all $f \in F' \subset RA_{m_2}R$.
(4.2) $\|\alpha''(g) - (\theta_0 + \theta_1) \circ \lambda \circ \alpha'(g)\| < \varepsilon_n$ for all $g \in G$.
(4.3) θ_0 is F'-$\min(\varepsilon_n, \delta)$ multiplicative and θ_1 satisfies that

$$\theta_1^{i,j}([q]) > L \cdot [\theta_0^{i,j}(R^i)]$$

for any non zero projection $q \in R^iA_{m_2}R^i$.
(4.4) Both $\alpha'' : B \longrightarrow (Q_0 + Q_1)A_{s(n+1)}(Q_0 + Q_1)$ and $\xi_4 : D \to Q_2A_{s(n+1)}Q_2$ satisfy the dichotomy condition (0.5).
(Thus we finished the construction of the sub-diagrams labeled by the number "3" and "$\approx \phi_{m_2,s(n+1)}$" of the large diagram in the outline. Combined with Step 2 and Step 3, these two sub-diagrams will lead to the sub-diagram labeled by the letter "u" of the large diagram as below.)
By the end of 1.1.4, for any blocks A^i, $A^k_{m_2}$ and any non zero projection $e \in A^i$, $\phi^{i,k}_{m_1,m_2}(e) \in A^k_{m_2}$ is a non zero projection. As a consequence of (4.3), we have

$$[(\theta_1 \circ \phi')(e)] \geq L \cdot [\theta_0(R)] \ (= L \cdot [Q_0]),$$

(Recall that $\phi' = \phi_{m_1,m_2}|_A$). Therefore, θ_0 and θ_1 (in place of λ_0 and λ_1) satisfy the condition in Theorem 1.6.9. By Theorem 1.6.9, there is a unitary $u \in (Q_0 + Q_1)A_{s(n+1)}(Q_0 + Q_1)$ such that

$$\|(\theta_0 + \theta_1) \circ \phi'(f) - \mathrm{Ad}u \circ (\theta_0 + \theta_1) \circ \lambda \circ (\phi \oplus r)(f)\| < 8\varepsilon_n, \qquad \forall f \in F.$$

Combining it with the second inequality of (2.4), we have
(4.5) $\|(\theta_0 + \theta_1) \circ \phi'(f) - \mathrm{Ad}u \circ (\theta_0 + \theta_1) \circ \lambda \circ \alpha' \circ \beta(f)\| < 9\varepsilon_n, \qquad \forall f \in F.$
Step 5. Finally, let all the blocks of D be the rest of B_{n+1}. Namely, let

$$B_{n+1} = C \oplus B \oplus D,$$

where C is from *Step 1*, B is from *Step 2*, and D is from *Step 4*.
We already have the definitions of $\beta_n : A_{s(n)} \to B_{n+1}$ and $\psi_{n,n+1} : B_n \to B_{n+1}$ for those blocks of $C \oplus B \subset B_{n+1}$ (from *Step 1* and *Step 2*). The definitions of β_n and $\psi_{n,n+1}$ for blocks of D, and the homomorphism $\alpha_{n+1} : C \oplus B \oplus D \to A_{s(n+1)}$ will be given below.
The part of $\beta_n : A_{s(n)} \to D$ ($\subset B_{n+1}$) is defined by

$$\beta_n = \xi_3 \circ \phi' \circ \theta : A_{s(n)} \xrightarrow{\theta} A \xrightarrow{\phi'} RA_{m_2}R \xrightarrow{\xi_3} D.$$

(Recall that $A = P_0 A_{m_1} P_0$ and $\phi' = \phi_{m_1,m_2}|_A$.) Since θ is F_n-ε_n multiplicative, and ϕ' and ξ_3 are homomorphisms, we know that this part of β_n is F_n-ε_n multiplicative.
The part of $\psi_{n,n+1} : B_n \to D$ $(\subset B_{n+1})$ is defined by

$$\psi_{n,n+1} = \xi_3 \circ \phi' \circ \alpha : B_n \xrightarrow{\alpha} A \xrightarrow{\phi'} RA_{m_2}R \xrightarrow{\xi_3} D$$

which is a homomorphism. The homomorphism $\alpha_{n+1} : C \oplus B \oplus D \to A_{s(n+1)}$ is defined as follows.
Consider the composition

$$\phi'' \circ \xi_2 : C \xrightarrow{\xi_2} P_1 A_{m_1} P_1 \xrightarrow{\phi''} \phi_{m_1,s(n+1)}(P_1) A_{s(n+1)} \phi_{m_1,s(n+1)}(P_1),$$

where P_1 and ξ_2 are from *Step 1*, $\phi'' = \phi_{m_1,s(n+1)}|_{P_1 A_{m_1} P_1}$. Using the dichotomy condition of ξ_2, by Lemma 6.2, there is a homomorphism $\tau : C \to \phi_{m_1,s(n+1)}(P_1) A_{s(n+1)} \phi_{m_1,s(n+1)}(P_1)$ such that
(5.1) $\quad \|\tau(f) - (\phi'' \circ \xi_2)(f)\| < \varepsilon_n, \quad \forall f \in \xi_1(F_n) \subset C$, and
(5.2) τ satisfies the dichotomy condition (0.5).
Define

$$\alpha_{n+1}|_C = \tau : C \to \phi_{m_1,s(n+1)}(P_1) A_{s(n+1)} \phi_{m_1,s(n+1)}(P_1),$$
$$\alpha_{n+1}|_B = \mathrm{Ad}u \circ \alpha'' : B \xrightarrow{\alpha''} (Q_0 + Q_1) A_{s(n+1)} (Q_0 + Q_1) \qquad \circlearrowleft_{Adu} \quad ,$$

where α'' is from *Step 4*, and define

$$\alpha_{n+1}|_D = \xi_4 : D \longrightarrow Q_2 A_{s(n+1)} Q_2.$$

Finally, choose $\{a_{n+1\ j}\}_{j=1}^{\infty} \subset A_{s(n+1)}$ and $\{b_{n+1\ j}\}_{j=1}^{\infty} \subset B_{n+1}$ to be countable dense subsets of unit balls of $A_{s(n+1)}$ and B_{n+1}, respectively. And choose

$$F_{n+1} = \phi_{s(n),s(n+1)}(F_n) \cup \alpha_{n+1}(E_{n+1}) \cup \bigcup_{i=1}^{n+1} \phi_{s(i),s(n+1)}(\{a_{i1}, a_{i2}, \cdots, a_{i\ n+1}\})$$

and

$$E_{n+1} = \psi_{n,n+1}(E_n) \cup \beta_n(F_n) \cup \bigcup_{i=1}^{n+1} \psi_{i,n+1}(\{b_{i1}, b_{i2}, \cdots, b_{i\ n+1}\}).$$

Thus we obtain the following diagram:

$$\begin{tikzcd}
F_n \subset & A_{s(n)} \arrow[rr, "\phi_{s(n),s(n+1)}"] & & A_{s(n+1)} & \supset F_{n+1} \\
& & {} & & \\
E_n \subset & B_n \arrow[uu, "\alpha_n"'] \arrow[rr, "\psi_{n,n+1}"] & & B_{n+1} \arrow[uu, "\alpha_{n+1}"'] & \supset E_{n+1} .
\end{tikzcd}$$

with $A_{s(n)} \xrightarrow{\beta_n} B_{n+1}$ drawn diagonally ($\searrow^{\beta_n}$).

Step 6. Now we need to verify all the conditions (0.1)–(0.5) for the above diagram.

(0.1)–(0.2) hold from the construction (see the constructions of B, C, D in *Step 1*, *Step 2* and *Step 4*, and E_{n+1}, F_{n+1} in the end of *Step 5*.)
(0.3) follows from the end of *Step 1*, the end of *Step 2*, and the part of the definition of β_n for D from *Step 5*.
(0.5) follows from (4.4) and (5.2).
So we only need to verify (0.4).
Combining (1.1) with (4.1), we have

$$\|\phi_{s(n),s(n+1)}(f) - [(\phi'' \circ \xi_2 \circ \xi_1) \oplus ((\theta_0 + \theta_1) \circ \phi' \circ \theta) \oplus (\xi_4 \circ \xi_3 \circ \phi' \circ \theta)](f)\| \\ < \varepsilon_n + \varepsilon_n = 2\varepsilon_n$$

for all $f \in F_n$ (recall that $\phi'' = \phi_{m_1,s(n+1)}|_{P_1 A_{m_1} P_1}, \phi' := \phi_{m_1,m_2}|_{P_0 A_{m_1} P_0}$).
Combined with (4.2), (4.5), (5.1) and the definitions of β_n and α_{n+1}, the proceeding inequality yields

$$\|\phi_{s(n),s(n+1)}(f) - (\alpha_{n+1} \circ \beta_n)(f)\| < 9\varepsilon_n + \varepsilon_n + 2\varepsilon_n = 12\varepsilon_n \qquad \forall f \in F_n.$$

Combining (1.3), the first inequality of (2.4), and the definitions of β_n and $\psi_{n,n+1}$, we have

$$\|\psi_{n,n+1}(g) - \beta_n \circ \alpha_n(g)\| < \varepsilon_n + \varepsilon_n = 2\varepsilon_n \qquad \forall g \in E_n.$$

So we obtain (0.4).
The theorem follows from Proposition 6.1.

$\square$

Remark 6.4. In the proof of the above theorem, if there is at least one block of B_{n+1} having spectrum of forms $S^1, T_{II,k}, T_{III,k}$, or S^2, then we can chose the map $\psi_{n,n+1}$ to be injective (e.g., the map ψ in *Step 2* can be chosen to be injective). Hence, in general, we can make the maps $\psi_{n,m}$, in the inductive system $(B_n, \psi_{n,m})$, injective. (Note that if no space of $S^1, T_{II,k}, T_{III,k}$ or S^2 appears, then it is easy to make the maps injective; see Theorem 2.2.1 of [Li2]).

Remark 6.5. By Lemma 1.3.3, our main result Theorem 6.3 also holds for general simple AH inductive limit C^*-algebras
$\lim(A_n = \bigoplus_{i=1}^{t_n} P_{n,i} M_{[n,i]}(C(X_{n,i}))P_{n,i}, \phi_{n,m})$ with uniformly bounded dimensions of $X_{n,i}$, where $P_{n,i} \in M_{[n,i]}(C(X_{n,i}))$ are projections. That is, such an AH algebra can be written as an inductive limit of a system $(B_n = \bigoplus_{i=1}^{s_n} Q_{n,i} M_{\{n,i\}}(C(Y_{n,i}))Q_{n,i}, \psi_{n,m})$, where $Y_{n,i}$ are the spaces:
$\{pt\}, [0,1], S^1, T_{II,k}, T_{III,k}$ and S^2, and $Q_{n,i} \in M_{\{n,i\}}(C(Y_{n,i}))$ are projections.

6.6. Suppose that a simple C^*-algebra A is an inductive limit of matrix algebras over $X_{n,i}$, where $X_{n,i}$ are the spaces of forms $\{pt\}, [0,1], S^1, S^2, T_{II,k}$ or $T_{III,k}$. Suppose that $K_*(A)$ is torsion free. Then it can be proved that for each fixed algebra A_n, integer $N > 0$, there is an A_m such that

$$\frac{\operatorname{rank}\phi_{n,m}^{i,j}(\mathbf{1}_{A_n^i})}{\operatorname{rank}(\mathbf{1}_{A_n^i})} \geq N,$$

and that $(\phi_{n,m})_*(tor K_*(A_n)) = 0$. Based on this, using the argument from §4 of [G2], we know that for any $F \subset A_n, \varepsilon > 0$, if N is large enough, then the above $\phi_{n,m}$ is homotopic to a homomorphism $\psi : A_n \to \phi_{n,m}(\mathbf{1}_{A_n})A_m\phi_{n,m}(\mathbf{1}_{A_n})$ satisfying $\psi(F) \subset_\varepsilon C$, where C is a direct sum of matrix algebras over spaces $\{pt\}, [0,1]$ and S^1. (See [G1] and the proof of Lemma 5.6 of [EGL] also.) Using the above fact, the following Corollary is a direct consequence of our Main Theorem and its proof. (In fact, since the algebras $M_k(\mathbb{C}), M_k(C[0,1])$ and $M_k(C(S^1))$ are stably generated, the proof is much simpler (see §3 of [Li3]).

Corollary 6.7. *Suppose that A is a simple C^*-algebra which is an inductive limit of an AH system with uniformly bounded dimensions of local spectra. If $K_*(A)$ is torsion free, then it is an inductive limit of matrix algebras over $C(S^1)$.*

Combining the above corollary with [El2] (see [NT] also), we have the following theorem.

Theorem 6.8. *Suppose that $A = \lim\limits_{n\to\infty}(A_n = \bigoplus_{i=1}^{t_n} P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}, \phi_{n,m})$ and $B = \lim\limits_{n\to\infty}(B_n = \bigoplus_{i=1}^{s_n} Q_{n,i}M_{\{n,i\}}(C(Y_{n,i}))Q_{n,i}, \psi_{n,m})$ are unital simple inductive limit algebras with uniformly bounded dimensions of local spectra $X_{n,i}$ and $Y_{n,i}$, respectively. Suppose that $K_*(A) = K_*(B)$ are torsion free. Suppose that there is an isomorphism of ordered groups*

$$\phi_0 : \ K_0A \longrightarrow K_0B$$

taking $[\mathbf{1}] \in K_0A$ into $[\mathbf{1}] \in K_0B$, that there is a group isomorphism

$$\phi_1 : \ K_1A \longrightarrow K_1B$$

and that there is an isomorphism between compact convex sets

$$\phi_T : TB \longrightarrow TA,$$

where TA and TB denote the simplices of tracial states of A and B, respectively. Suppose that ϕ_0 and ϕ_T are compatible, in the sense that

$$\tau(\phi_0 g) = \phi_T(\tau)(g), \qquad g \in K_0A, \ \tau \in TB.$$

It follows that there exists an isomorphism

$$\phi : \ A \longrightarrow B$$

giving rise to ϕ_0, ϕ_1, ϕ_T.

Remark 6.9. Since the C^*-algebras $C(T_{II,k}), C(T_{III,k})$ and $C(S^2)$ are not stably generated, our proof heavily depends on the results that, certain G-δ

multiplicative maps (with parts of point evaluations of sufficiently large sizes) are approximated by true homomorphisms in §5. We believe that such results should play important role in the future study of general simple C^*-algebras (with or without real rank zero property).

Remark 6.10. From a result of J. Villadsen, [V1], one knows that the restriction on the dimensions of the spaces $X_{n,i}$ can not be removed.
In [G5]—an appendix to this article, we will show that the condition of uniformly bounded dimensions of local spectra can be replaced by the condition of very slow dimension growth. The main difficulty for this case is that we can not obtain the homomorphism from B_n to $A_{s(n)}$ as the homomorphism α_n in the above proof. (The α_n in this case will be only a sufficiently multiplicative map.) But we can still construct homomorphisms $\psi_n : B_n \to B_{n+1}$, if we carefully choose α_n and β_n. This case does not create essential difficulty, but makes the proof much longer. We refer it to [G5] , a separate appendix to this paper.
It could be an improvement if one can replace the very slow dimension growth condition by the slow dimension growth condition. The author believes that the theorem is also true for this case. In fact, if one can prove the corresponding decomposition results (see Section 4) for the AH-algebras with slow dimension growth, then the Main Theorem in this article would also hold, by the same proof as in [G5].

Corollary 6.11. *Suppose that* $A = \lim\limits_{n\to\infty}(A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i})), \phi_{n,m})$ *is a simple inductive limit C^*-algebras. Suppose that each of the spaces $X_{n,i}$ is of the forms: $\{pt\}, [0,1], S^1, S^2, T_{II,k}$ or $T_{III,k}$. And suppose that all the connecting maps $\phi_{n,m}$ are injective. For any $F \subset A_n, \varepsilon > 0$, if m is large enough, then there are two mutually orthogonal projections $P, Q \in A_m$ and two homomorphisms $\phi : A_n \to PA_mP$ and $\psi : A_n \to QA_mQ$ such that*
(1) $\|\phi_{n,m}(f) - (\phi \oplus \psi)(f)\| < \varepsilon$ for all $f \in F$;
(2) $\phi(F)$ is weakly approximately constant to within ε and $SPV(\phi) < \varepsilon$;
(3) ψ factors through matrix algebras over $C[0,1]$.
Furthermore, if for some i, j, the partial map $\phi_{n,m}^{i,j} : A_n^i \to A_m^j$ is homotopic to a homomorphism with finite dimensional image, then the part ϕ of the decomposition $\phi \oplus \psi$ corresponding to this partial map can be chosen to be zero (or, equivalently, $\phi_{n,m}^{i,j}$ itself is close to a homomorphism factoring through a matrix algebra over $C[0,1]$).

Proof: It follows from the corollary of 2.3 of [Su] that for any $M_l(C(X))$, $\varepsilon > 0$, there are $\varepsilon_1 > 0$ and a finite subset F of self adjoint elements of $M_l(C(X))$ (i.e., $F \subset (M_l(C(X)))_{s.a}$) such that for any homomorphism $\phi : M_l(C(X)) \to M_{l_1}(C(Y))$, if $\phi(F)$ is weakly approximately constant to within ε_1, then $\mathrm{SPV}(\phi) < \varepsilon$. Therefore, for the desired condition (2) above, we only need to make $\phi(F)$ weakly approximately constant to within $\min(\varepsilon, \varepsilon_1)$.

To simplify the notation, we still denote $\min(\varepsilon, \varepsilon_1)$ by ε.
Now, the main body of the corollary follows from Lemma 6.2 and Theorem 5.32b. Namely, first apply Lemma 6.2 to id : $A_n \to A_n$ (in place of ϕ) and A_n in place of both B and A_n to find A_{n_1} (in place of A_m) and homomorphism $\alpha : A_n \to A_{n_1}$ such that α satisfies the dichotomy condition and such that α is sufficiently close to ϕ_{n,n_1} on the finite set F. Then apply Lemma 5.32b to A_n and $F \subset A_n$ (in place of B and $G \subset B$), A_{n_1} and $\phi_{n,n_1}(F) \subset A_{n_1}$ (in place of A_n and $F \subset A$), and $\alpha : A_n \to A_{n_1}$ (in place of $\alpha : B \to A_n$) to construct the desired decomposition. (Note that we use the following trivial fact: If two maps $\phi_1, \phi_2 : A_n \to A_m$ are approximately equal to each other to within ε_1 on the finite set F and the set $\phi_1(F)$ is weakly approximately constant to within ε_2, then the set $\phi_2(F)$ is weakly approximately constant to within $2\varepsilon_1 + \varepsilon_2$.)
For the last part of the Corollary, one needs to notice the following facts.
(i) In the additional parts of Corollaries 5.22 and 5.23, if the homomorphisms ψ are homomorphisms with finite dimensional images, then the homomorphisms $\tilde{\phi}$ in the corollaries 5.22 and 5.23 are also homomorphisms with finite dimensional images.
(ii) In Lemma 5.28, if both ϕ and ϕ_1 are homomorphisms factoring through interval algebras (this condition implies that they are homotopic to homomorphisms with finite dimensional images), then the homomorphism ψ in Lemma 5.28 (with $[\psi]_* = [\phi_2]_*$) can be chosen to be a homomorphism with finite dimensional image.
With the above facts, if $\phi^{i,j}_{n,m}$ is homotopic to a homomorphism with finite dimensional image and if $X_{n,i} \neq S^1$, then the corresponding part of ϕ in our corollary could be chosen to be a homomorphism with finite dimensional image, and therefore it can also factor through matrix algebras over $C[0,1]$. So, we can put it together with the part ψ and hence the part ϕ disappears from the decomposition of this partial map. This proves the additional part for the case $X_{n,i} \neq S^1$.
For the case that $X_{n,i} = S^1$, the additional part of the corollary follows from the following claim.
Claim: For any unitary $u \in A^i_n$ and any $\varepsilon > 0$, there is an integer $N > n$ such that if $m > N$, and if $\phi^{i,j}_{n,m}(u)$ is in the path connected component of the unit in the unitary group of $\phi^{i,j}_{n,m}(\mathbf{1}_{A^i_n})A^j_m\phi^{i,j}_{n,m}(\mathbf{1}_{A^i_n})$, then there is a self adjoint element $a \in \phi^{i,j}_{n,m}(\mathbf{1}_{A^i_n})A^j_m\phi^{i,j}_{n,m}(\mathbf{1}_{A^i_n})$ such that

$$\|\phi^{i,j}_{n,m}(u) - e^{2\pi i a}\| < \varepsilon.$$

(Obviously, if $\phi^{i,j}_{n,m}$ is homotopic to a homomorphism with finite dimensional image, then $\phi^{i,j}_{n,m}(u)$ is in the path connected component of the unit element in the unitary group of $\phi^{i,j}_{n,m}(\mathbf{1}_{A^i_n})A^j_m\phi^{i,j}_{n,m}(\mathbf{1}_{A^i_n})$.)
The proof of the above claim is exactly the same as the proof of the main theorem of [Phi3]: the simple inductive limit C^*-algebra in our corollary has exponential rank at most $1 + \varepsilon$. We omit the details.

We point out that, in [EGL], we will only need this result for the case $X_{m,j} = S^2$. Since $\dim(S^2) \leq 2$, $PM_\bullet(C(S^2))P$ has exponential rank at most $1+\varepsilon$. Therefore, the claim for the case $X_{m,j} = S^2$ $(X_{n,i} = S^1)$ is trivial.

$\square$

By Lemma 1.3.3, the above corollary also holds for the case of $A_n = \bigoplus_{i=1}^{t_n} P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}$, instead of $A_n = \bigoplus_{i=1}^{t_n} M_{[n,i]}(C(X_{n,i}))$.

COROLLARY 6.12. *Suppose that* $A = \lim_{n\to\infty}(A_n = \bigoplus_{i=1}^{t_n} P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}$, $\phi_{n,m})$ *is a simple inductive limit C^*-algebra. Suppose that each of the spaces $X_{n,i}$ is of the forms: $\{pt\}, [0,1], S^1, S^2, T_{II,k}$ or $T_{III,k}$. And suppose that all the connecting maps $\phi_{n,m}$ are injective. For any $F \subset A_n, \varepsilon > 0$, if m is large enough, then there are two mutually orthogonal projections $P, Q \in A_m$ and two homomorphisms $\phi : A_n \to PA_mP$ and $\psi : A_n \to QA_mQ$ such that*
(1) $\|\phi_{n,m}(f) - (\phi \oplus \psi)(f)\| < \varepsilon$ for all $f \in F$;
(2) $\phi(F)$ is weakly approximately constant to within ε and SPV$(\phi) < \varepsilon$;
(3) ψ factors through matrix algebras over $C[0,1]$.
Furthermore, if for some i, j, the partial map $\phi_{n,m}^{i,j} : A_n^i \to A_m^j$ is homotopic to a homomorphism with finite dimensional image, then the part ϕ of the decomposition $\phi \oplus \psi$ corresponding to this partial map can be chosen to be zero (or, equivalently, $\phi_{n,m}^{i,j}$ itself is close to a homomorphism factoring through a matrix algebra over $C[0,1]$).

Proof: By Lemma 1.3.3, there is an inductive system

$$\tilde{A} = \lim_{n\to\infty}(\tilde{A}_n = \bigoplus_{i=1}^{t_n} M_{\{n,i\}}(C(X_{n,i})), \tilde{\phi}_{n,m})$$

such that each $P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}$ is a corner of $M_{\{n,i\}}(C(X_{n,i}))$ and $\phi_{n,m} = \tilde{\phi}_{n,m}|_{P_{n,i}M_{[n,i]}(C(X_{n,i}))P_{n,i}}$. $\tilde{A}$ is simple since it is stably isomorphic to a simple C^*-algebra A. $\tilde{\phi}_{n,m}$ are injective since $\phi_{n,m}$ are injective. Apply Corollary 6.11 to $F \cup \{\mathbf{1}_{A_n^i}\}_{i=1}^{t_n} \subset A_n \subset \tilde{A}_n$ and $\frac{\varepsilon}{4} > 0$ to obtain $\tilde{\phi}$ and $\tilde{\psi}$ as the homomorphisms ϕ and ψ in Corollary 6.11. Since

$$\|(\tilde{\phi} + \tilde{\psi})(\mathbf{1}_{A_n^i}) - \tilde{\phi}_{n,m}(\mathbf{1}_{A_n^i})\| < \frac{\varepsilon}{4}, \quad \forall i,$$

there is a unitary $u \in \tilde{A}_m$ such that $\|u - \mathbf{1}\| < \frac{\varepsilon}{2}$ and

$$u((\tilde{\phi} + \tilde{\psi})(\mathbf{1}_{A_n^i}))u^* = \tilde{\phi}_{n,m}(\mathbf{1}_{A_n^i}) = \phi_{n,m}(\mathbf{1}_{A_n^i}), \quad \forall i$$

Finally, let

$$\phi = (\mathrm{Ad}u \circ \tilde{\phi})|_{A_n} \quad \text{and} \quad \psi = (\mathrm{Ad}u \circ \tilde{\psi})|_{A_n}$$

to obtain our corollary.

$\square$

References

[Bl1] B. Blackadar, Matricial and ultra-matricial topology, Operator Algebras, Mathematical Physics, and Low Dimensional Topology (R. H. Herman and B. Tanbay, eds.), A K Peters, Massachusetts, 1993, pp. 11–38.

[Bl2] B. Blackadar, *K-Theory for Operator Algebras*, Springer-Verlag, New York/Berlin/Heidelberg, 1986.

[Bl3] B. Blackadar, Symmetries of the CAR Algebras, *Ann. of Math.,* 131(1990), 589–623.

[BDR] B. Blackadar, M. Dadarlat, and M. Rørdam, M. The real rank of inductive limit C*-algebras, *Math. Scand.* 69 (1991), 211–216.

[Br] O. Bratteli, Inductive limits of finite dimensional C^*-algebras, *Trans. A.M.S.*, 171(1972), 195–234.

[Ch] E. Christensen, Near inclusions of C^*-algebras, *Acta Math.* 144(1980) 249-265.

[Con] A. Connes, *Noncommutative Geometry,* Academic press, New York, Tokyo, 1995.

[Cu] J. Cuntz, K-theory for certain C^*-algebras, *Ann. Math.* 113 (1981), 181-197.

[D1] M. Dadarlat, Approximately unitarily equivalent morphisms and inductive limit C^*-algebras, *K-theory* 9 (1995), 117–137.

[D2] M. Dadarlat, Reduction to dimension three of local spectra of real rank zero C^*-algebras, *J. Reine Angew. Math.* 460 (1995), 189–212.

[DG] M. Dadarlat and G. Gong, A classification result for approximately homogeneous C^*-algebras of real rank zero , *Geometric and Functional Analysis,* 7(1997) 646–711.

[DNNP] M. Dadarlat, G. Nagy, A. Nemethi, and C. Pasnicu, Reduction of topological stable rank in inductive limits of C*-algebras, *Pacific J. Math.* 153 (1992), 267–276.

[DN] M. Dadarlat and Nemethi, Shape theory and (connective) K-theory, *J. Operator Theory* 23 (1990), 207–291.

[Da] K. Davidson, C^*-algebras by examples, Fields Institute Monographs, 6, A.M.S. Providence, R.I.

[Ell1] G. A. Elliott, A classification of certain simple C^*-algebras, *Quantum and Non-Commutative Analysis* (editors, H. Araki et al.), Kluwer, Dordrecht, 1993, pp. 373–385.

[Ell2] G. A. Elliott, A classification of certain simple C^*-algebras, II, *J. Ramanujan Math. Soc.* 12 (1997), 97–134.

[Ell3] G. A. Elliott, The classification problem for amenable C^*-algebras, *Proceedings of the International Congress of Mathematicians, Zürich, Switzerland, 1994 (editors, S.D. Chattrji)*, Birkhäuser, Basel, 1995, pp. 922–932.

[Ell4] Elliott, G.A., On the classification of inductive limits of sequences of semisimple finite dimensional algebras, *J. Algebra* 38 (1976), 29–44.

[Ell5] Elliott, G.A., On the classification of C*-algebras of real rank zero, *J. Reine Angew. Math.* 443 (1993), 179–219.

[EE] G. A. Elliott and D. E. Evans, The structure of irrational rotation C^*-algebras, *Ann. of Math.* 138 (1993), 477–501.

[EG1] G. A. Elliott and G. Gong, On inductive limits of matrix algebras over two-tori, *Amer. J. Math.* 118 (1996), 263–290.

[EG2] G. A. Elliott and G. Gong, On the classification of C^*-algebras of real rank zero, II, *Ann. of Math.* 144 (1996), 497–610.

[EGJS] G. A. Elliott, G. Gong, X. Jiang, and H. Su, A classification of simple limits of dimension drop C^*-algebras, *Fields Institute Communications* 13 (1997), 125–143.

[EGL] G. A. Elliott, G. Gong, and L. Li, On the classification of simple inductive limit C^*-algebras, II : The isomorphism theorem, preprint.

[EGLP] G. A. Elliott, G. Gong, H. Lin, and C. Pasnicu, Abelian C^*-subalgebras of C^*-algebras of real rank zero and inductive limit C^*-algebras, *Duke Math. J.* 83 (1996), 511–554.

[G1] G. Gong, Approximation by dimension drop C^*-algebras and classification, *C. R. Math. Rep. Acad. Sci. Canada* 16 (1994), 40–44.

[G2] G. Gong, Classification of C^*-algebras of real rank zero and unsuspended E-equivalence types, *J. Funct. Anal.* 152(1998), 281–329.

[G3] G. Gong, On inductive limits of matrix algebras over higher dimensional spaces, Part I, *Math. Scand.* 80 (1997), 45–60.

[G4] G. Gong, On inductive limits of matrix algebras over higher dimensional spaces, Part II, *Math. Scand.* 80(1997), 61–100.

[G5] G. Gong, Simple inductive limit C^*-algebras with very slow dimension growth: An appendix for "On the classification of simple inductive limit C^*-algebras, I: The reduction theorem", Preprint.

[GL1] G. Gong and H. Lin, The exponential rank of inductive limit C^*-algebras,, *Math. Scand.* 71 (1992), 301–319.

[GL2] G. Gong and H. Lin, Almost multiplicative morphisms and K-theory, *International J. of Math.* 11(2000) 983-1000.

[HV] P. Halmos and H. Vaughan, Marriage problems, *Amer. J. of Math.* 72 (1950), 214–215.

[Hu] D. Husemoller, *Fibre Bundles*, McGraw–Hill, New York, 1966; reprinted in Springer-Verlag Graduate Texts in Mathematics.

[JS1] X. Jiang and H. Su, On a simple unital projectionless C^*-algebra, *American J. of Math.* 121(1999), 359–413.

[JS2] X. Jiang and H. Su, A classification of simple inductive limits of splitting interval algebras, *J. Funct. Anal.*, 151(1997), 50–76.

[Kir] E. Kirchberg, The classification of purely infinite C^*-algebras using Kasparov's theory, preprint.

[Li1] L. Li, *On the classification of simple C^*-algebras: Inductive limits of matrix algebras over trees*, Mem. Amer. Math. Soc., no. 605, vol. 127, 1997.

[Li2] L. Li, *Simple inductive limit C^*-algebras: Spectra and approximation by interval algebras,* J. Reine. Angew Math. 507 (1999), 57–79.

[Li3] L. Li, Classification of simple C^*-algebras: Inductive limits of matrix algebras over 1-dimensional spaces, *J. Funct. Anal.*192 (2002) 1-51.

[Lo] T. Loring, Lifting solutions to perturbing problems in C^*-algebras, 8, Fields Institute monograph, Providence, R.I.

[NT] K. Nielsen, and K. Thomsen, *Limit of circle algebras,* Exposition Math 14(1996) 17–56.

[Phi1] N. C. Phillips, A classification theorem for nuclear purely infinite simple C^*-algebras, *Doc. Math.* 5 (2000), 49-114.

[Phi2] N. C. Phillips, How many exponentials?, American J. of Math. 116 (1994) 1513–1543.

[Phi3] N. C. Phillips, Reduction of exponential rank in direct limits of C^*-algebras, *Canad. J. Math.* 46 (1994), 818–853.

[R] M. Rørdam, Classification of inductive limits of Cuntz algebras, *J. Reine Angew. Math.* 440 (1993), 175–200.

[Sch] C. Schochet, Topological methods for C^*-algebras IV: mod p homology, *Pacific J. Math.* 114 (1984), 447–468.

[Se] G. Segal, K-homology theory in K-theory and operator algebras, *Lecture Notes in Mathematics* 575, Springer-Verlag, 1977, pp. 113–127.

[St] J. R. Stallings, *Lectures on Polyhedral Topology* (notes by G. A. Swarup) 1967, Tata Institute of Fundamental Research, Bombay.

[Su] H. Su, *On the classification of C^*-algebras of real rank zero: Inductive limits of matrix algebras over non-Hausdorff graphs*, Mem. Amer. Math. Soc. no. 547, vol. 114 (1995).

[Th1] K. Thomsen, Limits of certain subhomogeneous C^*-algebras, Mem. Soc. Math. Fr. 71 (1999).

[V1] J. Villadsen, Simple C^*-algebras with perforation, *J. Funct. Anal.*, 154(1998), 110–116.

[V2] J. Villadsen, On stable rank of Simple C^*-algebras, *J. of Amer. Math. Soc.* 12(1999), 1091–1102.

[Wh] G. Whitehead, Elements of Homotopy theory, Springer-Verlag, 1978.

[Zh] S. Zhang, A property of purely infinite simple C^*-algebras, *Proc. Amer. Math. Soc.* 109 (1990), 717–720.

Guihua Gong
Department of Mathematics
University of Puerto Rico, Rio Piedras
San Juan, PR 00931-3355, USA
ggong@goliath.cnnet.clu.edu

Unimodular Covers of Multiples of Polytopes

Winfried Bruns and Joseph Gubeladze[1]

Received: November 11, 2001
Revised: December 4, 2002

Communicated by Günter M. Ziegler

Abstract. Let P be a d-dimensional lattice polytope. We show that there exists a natural number c_d, only depending on d, such that the multiples cP have a unimodular cover for every natural number $c \geq c_d$. Actually, an explicit upper bound for c_d is provided, together with an analogous result for unimodular covers of rational cones.

2000 Mathematics Subject Classification: Primary 52B20, 52C07, Secondary 11H06
Keywords and Phrases: lattice polytope, rational cone, unimodular covering

1. Statement of results

All polytopes and cones considered in this paper are assumed to be convex. A polytope $P \subset \mathbb{R}^d$ is called a *lattice polytope*, or *integral polytope*, if its vertices belong to the standard lattice $\mathbb{Z}^d$. For a (not necessarily integral) polytope $P \subset \mathbb{R}^d$ and a real number $c \geq 0$ we let cP denote the image of P under the dilatation with factor c and center at the origin $O \in \mathbb{R}^d$. A polytope of dimension e is called an *e-polytope*.

A *simplex* Δ is a polytope whose vertices $v_0, \dots, v_e$ are affinely independent (so that $e = \dim \Delta$). The *multiplicity* $\mu(\Delta)$ of a lattice simplex is the index of the subgroup U generated by the vectors $v_1 - v_0, \dots, v_e - v_0$ in the smallest direct summand of $\mathbb{Z}^d$ containing U, or, in other words, the order of the torsion subgroup of $\mathbb{Z}^d/U$. A simplex of multiplicity 1 is called *unimodular*. If $\Delta \subset \mathbb{R}^d$ has the full dimension d, then $\mu(\Delta) = d! \operatorname{vol}(\Delta)$, where vol is the Euclidean volume. The union of all unimodular d-simplices inside a d-polytope P is denoted by $\mathrm{UC}(P)$.

In this paper we investigate for which multiples cP of a lattice d-polytope one can guarantee that $cP = \mathrm{UC}(cP)$. To this end we let $\mathfrak{c}_d^{\mathrm{pol}}$ denote the infimum of the natural numbers c such that $c'P = \mathrm{UC}(c'P)$ for all lattice d-polytopes

[1]The second author was supported by the Deutsche Forschungsgemeinschaft.

P and all natural numbers $c' \geq c$. A priori, it is not excluded that $\mathfrak{c}_d^{\mathrm{pol}} = \infty$ and, to the best of our knowledge, it has not been known up till now whether $\mathfrak{c}_d^{\mathrm{pol}}$ is finite except for the cases $d = 1, 2, 3$: $\mathfrak{c}_1^{\mathrm{pol}} = \mathfrak{c}_2^{\mathrm{pol}} = 1$ and $\mathfrak{c}_3^{\mathrm{pol}} = 2$, where the first equation is trivial, the second is a crucial step in the derivation of Pick's theorem, and a proof of the third can be found in Kantor and Sarkaria [KS]. Previous results in this direction were obtained by Lagarias and Ziegler (Berkeley 1997, unpublished).
The main result of this paper is the following upper bound, positively answering Problem 4 in [BGT2]:

Theorem 1.1. *For all natural numbers* $d > 1$ *one has*

$$\mathfrak{c}_d^{\mathrm{pol}} \leq O\big(d^5\big)\left(\frac{3}{2}\right)^{\left\lceil\sqrt{d-1}\right\rceil(d-1)}.$$

Theorem 1.1 is proved by passage to cones, for which we establish a similar result on covers by unimodular subcones (Theorem 1.3 below). This result, while interesting of its own, implies Theorem 1.1 and has the advantage of being amenable to a proof by induction on d.
We now explain some notation and terminology. The convex hull of a set $X \subset \mathbb{R}^d$ is denoted by $\operatorname{conv}(X)$, and $\operatorname{Aff}(X)$ is its affine hull. Moreover, $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$ and $\mathbb{Z}_+ = \mathbb{Z} \cap \mathbb{R}_+$.
A lattice simplex is called *empty* if its vertices are the only lattice points in it. Every unimodular simplex is empty, but the opposite implication is false in dimensions ≥ 3. (In dimension 2 empty simplices are unimodular.)
A *cone* (without further predicates) is a subset of $\mathbb{R}^d$ that is closed under linear combinations with coefficients in $\mathbb{R}_+$. All cones considered in this paper are assumed to be *polyhedral*, *rational* and *pointed* (i. e. not to contain an affine line); in particular they are generated by finitely many rational vectors. For such a cone C the semigroup $C \cap \mathbb{Z}^d$ has a unique finite minimal set of generators, called the *Hilbert basis* and denoted by $\operatorname{Hilb}(C)$. The *extreme (integral) generators* of a rational cone $C \subset \mathbb{R}^d$ are, by definition, the generators of the semigroups $l \cap \mathbb{Z}^d \approx \mathbb{Z}_+$ where l runs through the edges of C. The extreme integral generators of C are members of $\operatorname{Hilb}(C)$. We define Δ_C to be the convex hull of O and the extreme integral generators of C.
A cone C is *simplicial* if it has a linearly independent system of generators. Thus C is simplicial if and only if Δ_C is a simplex. We say that C is *empty simplicial* if Δ_C is an empty simplex. The *multiplicity* of a simplicial cone is $\mu(\Delta_C)$. If Δ is a lattice simplex with vertex O, then the multiplicity of the cone $\mathbb{R}_+\Delta$ divides $\mu(\Delta)$. This follows easily from the fact that each non-zero vertex of Δ is an integral multiple of an extreme integral generator of $\mathbb{R}_+\Delta$.
A *unimodular* cone $C \subset \mathbb{R}^d$ is a rational simplicial cone for which Δ_C is a unimodular simplex. Equivalently we could require that C is simplicial and its extreme integral generators generate a direct summand of $\mathbb{Z}^d$. A *unimodular cover* of an arbitrary rational cone C is a finite system of unimodular cones

whose union is C. A *unimodular triangulation* of a cone is defined in the usual way – it is a unimodular cover whose member cones coincide along faces.
In addition to the cones C with apex in the origin O, as just introduced, we will sometimes have to deal with sets of the form $v + C$ where $v \in \mathbb{R}^d$. We call $v + C$ a *cone with apex* v.
We define $\mathfrak{c}_d^{\text{cone}}$ to be the infimum of all natural numbers c such that every rational d-dimensional cone $C \subset \mathbb{R}^d$ admits a unimodular cover $C = \bigcup_{j=1}^k C_j$ for which

$$\operatorname{Hilb}(C_j) \subset c\Delta_C \quad j \in [1, k].$$

Remark 1.2. We will often use that a cone C can be triangulated into empty simplicial cones C' such that $\Delta_{C'} \subset \Delta_C$. In fact, one first triangulates C into simplicial cones generated by extreme generators of C. After this step one can assume that C is simplicial with extreme generators $v_1, \dots, v_d$. If Δ_C is not empty, then we use stellar subdivision along a ray through some $v \in \Delta_C \cap \mathbb{Z}^d$, $v \neq 0, v_1, \dots, v_d$, and for each of the resulting cones C' the simplex $\Delta_{C'}$ has a smaller number of integral vectors than Δ_C. In proving a bound on $\mathfrak{c}_d^{\text{cone}}$ it is therefore enough to consider empty simplicial cones.
Similarly one triangulates every lattice polytope into empty simplices.

Results on $\mathfrak{c}_d^{\text{cone}}$ seem to be known only in dimensions ≤ 3. Since the empty simplicial cones in dimension 2 are exactly the unimodular 2-cones (by a well known description of Hilbert bases in dimension 2, see Remark 4.2) we have $\mathfrak{c}_2^{\text{cone}} = 1$. Moreover, it follows from a theorem of Sebő [S1] that $\mathfrak{c}_3^{\text{cone}} = 2$. In fact Sebő has shown that a 3-dimensional cone C can be triangulated into unimodular cones generated by elements of $\operatorname{Hilb}(C)$ and that $\operatorname{Hilb}(C) \subset (d-1)\Delta_C$ in all dimensions d (see Remark 1.4(f)).
We can now formulate the main result for unimodular covers of rational cones:

Theorem 1.3. *For all $d \geq 2$ one has*

$$\mathfrak{c}_d^{\text{cone}} \leq \left\lceil \sqrt{d-1}\, \right\rceil (d-1) \frac{d(d+1)}{2} \left(\frac{3}{2}\right)^{\left\lceil \sqrt{d-1} \right\rceil (d-1) - 2}.$$

Remark 1.4. (a) We have proved in [BGT1, Theorem 1.3.1] that there is a natural number c_P for a lattice polytope $P \subset \mathbb{R}^d$ such that $cP = \operatorname{UC}(cP)$ whenever $c \geq c_P$, $c \in \mathbb{N}$. However, neither did the proof in [BGT1] provide an explicit bound for c_P, nor was it clear that the numbers c_P can be uniformly bounded with respect to all d-dimensional polytopes. The proof we present below is an essential extension of that of [BGT1, Theorem 1.3.1].
(b) It has been proved in [KKMS, Theorem 4, Ch. III] that for every lattice polytope P there exists a natural number c such that cP admits even a regular triangulation into unimodular simplices. This implies that $c'cP$ also admits such a triangulation for $c' \in \mathbb{N}$. However, the question whether there exists a natural number c_P^{triang} such that the multiples $c'P$ admit unimodular triangulations for all $c' \geq c_P^{\text{triang}}$ remains open. In particular, the existence of a uniform bound $\mathfrak{c}_d^{\text{triang}}$ (independent of P) remains open.

(c) The main difficulty in deriving better estimates for $\mathfrak{c}_d^{\mathrm{pol}}$ lies in the fundamental open problem of an effective description of the empty lattice d-simplices; see Haase and Ziegler [HZ] and Sebő [S2] and the references therein.
(d) A chance for improving the upper bound in Theorem 1.1 to, say, a polynomial function in d would be provided by an algorithm for resolving toric singularities which is faster then the standard one used in the proof of Theorem 4.1 below. Only there exponential terms enter our arguments.
(e) A lattice polytope $P \subset \mathbb{R}^d$ which is covered by unimodular simplices is *normal*, i. e. the additive subsemigroup

$$S_P = \sum_{x\in P\cap\mathbb{Z}^d} \mathbb{Z}_+(x,1) \subset \mathbb{Z}^{d+1}$$

is normal and, moreover, $\mathrm{gp}(S_P) = \mathbb{Z}^{d+1}$. (The normality of S_P is equivalent to the normality of the K-algebra $K[S_P]$ for a field K.) However, there are normal lattice polytopes in dimension ≥ 5 which are not unimodularly covered [BG]. On the other hand, if $\dim P = d$ then cP is normal for arbitrary $c \geq d-1$ [BGT1, Theorem 1.3.3(a)] (and $\mathrm{gp}(S_{cP}) = \mathbb{Z}^{d+1}$, as is easily seen). The example found in [BG] is far from being of type cP with $c > 1$ and, correspondingly, we raise the following question: is $\mathfrak{c}_d^{\mathrm{pol}} = d-1$ for all natural numbers $d > 1$? As mentioned above, the answer is 'yes' for $d = 2, 3$, but we cannot provide further evidence for a positive answer.
(f) Suppose $C_1, \dots, C_k$ form a unimodular cover of C. Then $\mathrm{Hilb}(C_1) \cup \dots \cup \mathrm{Hilb}(C_k)$ generates $C \cap \mathbb{Z}^d$. Therefore $\mathrm{Hilb}(C) \subset \mathrm{Hilb}(C_1) \cup \dots \cup \mathrm{Hilb}(C_k)$, and so $\mathrm{Hilb}(C)$ sets a lower bound to the size of $\mathrm{Hilb}(C_1) \cup \dots \cup \mathrm{Hilb}(C_k)$ relative to Δ_C. For $d \geq 3$ there exist cones C such that $\mathrm{Hilb}(C)$ is not contained in $(d-2)\Delta_C$ (see Ewald and Wessels [EW]), and so one must have $\mathfrak{c}_d^{\mathrm{cone}} \geq d-1$. On the other hand, $d-1$ is the best lower bound for $\mathfrak{c}_d^{\mathrm{cone}}$ that can be obtained by this argument since $\mathrm{Hilb}(C) \subset (d-1)\Delta_C$ for all cones C. We may assume that C is empty simplicial by Remark 1.2, and for an empty simplicial cone C we have

$$\mathrm{Hilb}(C) \subset \square_C \setminus (v_1 + \dots + v_d - \Delta_C) \subset (d-1)\Delta_C$$

where

(i) $v_1, \dots, v_d$ are the extreme integral generators of C,
(ii) $\square_C$ is the semi-open parallelotope spanned by $v_1, \dots, v_d$, that is,

$$\square_C = \{\xi_1 v_1 + \dots + \xi_d v_d : \xi_1, \dots, \xi_d \in [0,1)\}.$$

Acknowledgement. We thank the referees for their careful reading of the paper. It led to a number of improvements in the exposition, and helped us to correct an error in the first version of Lemma 4.1.

2. Slope independence

By $[0,1]^d = \{(z_1, \dots, z_d) \mid 0 \leq z_1, \dots, z_d \leq 1\}$ we denote the standard unit d-cube. Consider the system of simplices

$$\Delta_\sigma \subset [0,1]^d, \quad \sigma \in S_d,$$

where S_d is the permutation group of $\{1,\dots,d\}$, and Δ_σ is defined as follows:

(i) $\Delta_\sigma = \operatorname{conv}(x_0, x_1, \dots, x_d)$,

(ii) $x_0 = O$ and $x_d = (1,\dots,1)$,

(iii) x_{i+1} differs from x_i only in the $\sigma(i+1)$st coordinate and $x_{i+1,\sigma(i+1)} = 1$ for $i \in [0, d-1]$.

Then $\{\Delta_\sigma\}_{\sigma\in S_d}$ is a unimodular triangulation of $[0,1]^d$ with additional good properties [BGT1, Section 2.3]. The simplices Δ_σ and their integral parallel translates triangulate the entire space $\mathbb{R}^d$ into affine Weyl chambers of type A_d. The induced triangulations of the integral multiples of the simplex

$$\operatorname{conv}(O, e_1, e_1+e_2, \dots, e_1+\dots+e_d\} \subset \mathbb{R}^d$$

are studied in great detail in [KKMS, Ch. III]. All we need here is the very existence of these triangulations. In particular, the integral parallel translates of the simplices Δ_σ cover (actually, triangulate) the cone

$$\mathbb{R}_+e_1 + \mathbb{R}_+(e_1+e_2) + \dots + \mathbb{R}_+(e_1+\dots+e_d) \approx \mathbb{R}^d_+$$

into unimodular simplices.

Suppose we are given a real linear form

$$\alpha(X_1,\dots,X_d) = a_1X_1 + \dots + a_dX_d \neq 0.$$

The *width* of a polytope $P \subset \mathbb{R}^d$ in direction $(a_1,\dots,a_d)$, denoted by $\operatorname{width}_\alpha(P)$, is defined to be the Euclidean distance between the two extreme hyperplanes that are parallel to the hyperplane $a_1X_1+\dots+a_dX_d = 0$ and intersect P. Since $[0,1]^d$ is inscribed in a sphere of radius $\sqrt{d}/2$, we have $\operatorname{width}_\alpha(\Delta_\sigma) \le \sqrt{d}$ whatever the linear form α and the permutation σ are. We arrive at

Proposition 2.1. *All integral parallel translates of Δ_σ, $\sigma \in S_d$, that intersect a hyperplane H are contained in the $\sqrt{d}$-neighborhood of H.*

In the following we will have to consider simplices that are unimodular with respect to an affine sublattice of $\mathbb{R}^d$ different from $\mathbb{Z}^d$. Such lattices are sets

$$\mathcal{L} = v_0 + \sum_{i=1}^{e} \mathbb{Z}(v_i - v_0)$$

where $v_0,\dots,v_e$, $e \le d$, are affinely independent vectors. (Note that $\mathcal{L}$ is independent of the enumeration of the vectors $v_0,\dots,v_d$.) An e-simplex $\Delta = \operatorname{conv}(w_0,\dots,w_e)$ defines the lattice

$$\mathcal{L}_\Delta = w_0 + \sum_{i=0}^{e} \mathbb{Z}(w_i - w_0).$$

Let $\mathcal{L}$ be an affine lattice. A simplex Δ is called $\mathcal{L}$-*unimodular* if $\mathcal{L} = \mathcal{L}_\Delta$, and the union of all $\mathcal{L}$-unimodular simplices inside a polytope $P \subset \mathbb{R}^d$ is denoted by $\operatorname{UC}_\mathcal{L}(P)$. For simplicity we set $\operatorname{UC}_\Delta(P) = \operatorname{UC}_{\mathcal{L}_\Delta}(P)$.

Let $\Delta \subset \Delta'$ be (not necessarily integral) d-simplices in $\mathbb{R}^d$ such that the origin $O \in \mathbb{R}^d$ is a common vertex and the two simplicial cones spanned by Δ and Δ'

at O are the same. The following lemma says that the $\mathcal{L}_\Delta$-unimodularly covered area in a multiple $c\Delta'$, $c \in \mathbb{N}$, approximates $c\Delta'$ with a precision independent of Δ'. The precision is therefore independent of the "slope" of the facets of Δ and Δ' opposite to O. The lemma will be critical both in the passage to cones (Section 3) and in the treatment of the cones themselves (Section 6).

Lemma 2.2. *For all d-simplices $\Delta \subset \Delta'$ having O as a common vertex at which they span the same cone, all real numbers ε, $0 < \varepsilon < 1$, and $c \geq \sqrt{d}/\varepsilon$ one has*

$$(c - \varepsilon c)\Delta' \subset \mathrm{UC}_\Delta(c\Delta').$$

Proof. Let $v_1, \dots, v_d$ be the vertices of Δ different from O, and let w_i, $i \in [1, d]$ be the vertex of Δ' on the ray $\mathbb{R}_+ v_i$. By a rearrangement of the indices we can achieve that

$$\frac{|w_1|}{|v_1|} \geq \frac{|w_2|}{|v_2|} \geq \cdots \geq \frac{|w_d|}{|v_d|} \geq 1.$$

where $|\ |$ denotes Euclidean norm. Moreover, the assertion of the lemma is invariant under linear transformations of $\mathbb{R}^d$. Therefore we can assume that

$$\Delta = \mathrm{conv}(O, e_1, e_1 + e_2, \dots, e_1 + \cdots + e_d).$$

Then $\mathcal{L}_\Delta = \mathbb{Z}^d$. The ratios above are also invariant under linear transformations. Thus

$$\frac{|w_1|}{|e_1|} \geq \frac{|w_2|}{|e_1 + e_2|} \geq \cdots \geq \frac{|w_d|}{|e_1 + \cdots + e_d|} \geq 1.$$

Now Lemma 2.4 below shows that the distance h from O to the affine hyperplane $\mathcal{H}$ through $w_1, \dots, w_d$ is at least 1.
By Proposition 2.1, the subset

$$(c\Delta') \setminus U_{\sqrt{d}}(c\mathcal{H}) \subset c\Delta'$$

is covered by integral parallel translates of the simplices Δ_σ, $\sigma \in S_d$ that are contained in $c\Delta$. ($U_\delta(M)$ is the δ-neighborhood of M.) In particular,

$$(1) \qquad (c\Delta') \setminus U_{\sqrt{d}}(c\mathcal{H}) \subset \mathrm{UC}_\Delta(c\Delta').$$

Therefore we have

$$(1-\varepsilon)c\Delta' \subset \left(1-\frac{\sqrt{d}}{c}\right)c\Delta' \subset \left(1-\frac{\sqrt{d}}{ch}\right)c\Delta' = \frac{ch-\sqrt{d}}{ch}c\Delta' = (c\Delta')\setminus U_{\sqrt{d}}(c\mathcal{H}),$$

and the lemma follows from (1). □

Remark 2.3. One can derive an analogous result using the trivial tiling of $\mathbb{R}^d_+$ by the integral parallel translates of $[0,1]^d$ and the fact that $[0,1]^d$ itself is unimodularly covered. The argument would then get simplified, but the estimate obtained is $c \geq d/\varepsilon$, and thus worse than $c \geq \sqrt{d}/\varepsilon$.

We have formulated the Lemma 2.2 only for full dimensional simplices, but it holds for simplices of smaller dimension as well: one simply chooses all data relative to the affine subspace generated by Δ'.
Above we have used the following

Lemma 2.4. *Let $e_1, \dots, e_d$ be the canonical basis of $\mathbb{R}^d$ and set $w_i = \lambda_i(e_1 + \cdots + e_i)$ where $\lambda_1 \geq \cdots \geq \lambda_d > 0$. Then the affine hyperplane $\mathcal{H}$ through $w_1, \dots, w_d$ intersects the set $Q = \lambda_d(e_1 + \cdots + e_d) - \mathbb{R}^d_+$ only in the boundary ∂Q. In particular the Euclidean distance from O to $\mathcal{H}$ is $\geq \lambda_d$.*

Proof. The hyperplane $\mathcal{H}$ is given by the equation

$$\frac{1}{\lambda_1} X_1 + \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right) X_2 + \cdots + \left(\frac{1}{\lambda_d} - \frac{1}{\lambda_{d-1}}\right) X_d = 1.$$

The linear form α on the left hand side has non-negative coefficients and $w_d \in \mathcal{H}$. Thus a point whose coordinates are strictly smaller than λ_d cannot be contained in $\mathcal{H}$. □

3. Passage to cones

In this section we want to relate the bounds for $\mathfrak{c}_d^{\mathrm{pol}}$ and $\mathfrak{c}_d^{\mathrm{cone}}$. This allows us to derive Theorem 1.1 from Theorem 1.3.

Proposition 3.1. *Let d be a natural number. Then $\mathfrak{c}_d^{\mathrm{pol}}$ is finite if and only if $\mathfrak{c}_d^{\mathrm{cone}}$ is finite, and, moreover,*

$$\mathfrak{c}_d^{\mathrm{cone}} \leq \mathfrak{c}_d^{\mathrm{pol}} \leq \sqrt{d}(d+1)\mathfrak{c}_d^{\mathrm{cone}}. \tag{2}$$

Proof. Suppose that $\mathfrak{c}_d^{\mathrm{pol}}$ is finite. Then the left inequality is easily obtained by considering the multiples of the polytope Δ_C for a cone C: the cones spanned by those unimodular simplices in a multiple of Δ_C that contain O as a vertex constitute a unimodular cover of C.

Now suppose that $\mathfrak{c}_d^{\mathrm{cone}}$ is finite. For the right inequality we first triangulate a polytope P into lattice simplices. Then it is enough to consider a lattice d-simplex $\Delta \subset \mathbb{R}^d$ with vertices $v_0, \dots, v_d$.

Set $c' = \mathfrak{c}_d^{\mathrm{cone}}$. For each i there exists a unimodular cover (D_{ij}) of the corner cone C_i of Δ with respect to the vertex v_i such that $c'\Delta - c'v_i$ contains $\Delta_{D_{ij}}$ for all j. Thus the simplices $\Delta_{D_{ij}} + c'v_i$ cover the corner of $c'\Delta$ at $c'v_i$, that is, their union contains a neighborhood of $c'v_i$ in $c'\Delta$.

We replace Δ by $c'\Delta$ and can assume that each corner of Δ has a cover by unimodular simplices. It remains to show that the multiples $c''\Delta$ are unimodularly covered for every number $c'' \geq \sqrt{d}(d+1)$ for which $c''P$ is an integral polytope.

Let

$$\omega = \frac{1}{d+1}(v_0 + \cdots + v_d)$$

be the barycenter of Δ. We define the subsimplex $\Delta_i \subset \Delta$ as follows: Δ_i is the homothetic image of Δ with respect to the center v_i so that ω lies on the facet of Δ_i opposite to v_i. In dimension 2 this is illustrated by Figure 1. The factor of the homothety that transforms Δ into Δ_i is $d/(d+1)$. In particular, the simplices Δ_i are pairwise congruent. It is also clear that

$$\bigcup_{i=0}^{d} \Delta_i = \Delta. \tag{3}$$

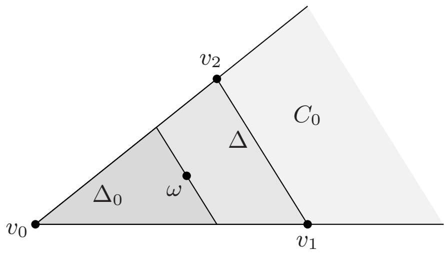


Figure 1.

The construction of ω and the subsimplices Δ_i commutes with taking multiples of Δ. It is therefore enough to show that $c''\Delta_i \subset \mathrm{UC}(c''\Delta)$ for all i. In order to simplify the use of dilatations we move v_i to O by a parallel translation. In the case in which $v_i = O$ the simplices $c''\Delta$ and $c''\Delta_i$ are the unions of their intersections with the cones D_{ij}. This observation reduces the critical inclusion $c''\Delta_i \subset c''\Delta$ to

$$c''(\Delta_i \cap D_{ij}) \subset c''(\Delta \cap D_{ij})$$

for all j. But now we are in the situation of Lemma 2.2, with the unimodular simplex $\Delta_{D_{ij}}$ in the role of the Δ of 2.2 and $\Delta \cap D_{ij}$ in that of Δ'. For $\varepsilon = 1/(d+1)$ we have $c'' \geq \sqrt{d}/\varepsilon$ and so

$$c''(\Delta_i \cap D_{ij}) = c''\frac{d}{d+1}(\Delta \cap D_{ij}) = c''(1-\varepsilon)(\Delta \cap D_{ij}) \subset \mathrm{UC}(\Delta \cap D_{ij}),$$

as desired. $\square$

At this point we can deduce Theorem 1.1 from Theorem 1.3. In fact, using the bound for $\mathfrak{c}_d^{\mathrm{cone}}$ given in Theorem 1.3 we obtain

$$\begin{aligned}
\mathfrak{c}_d^{\mathrm{pol}} &\leq \sqrt{d}(d+1)\mathfrak{c}_d^{\mathrm{cone}} \\
&\leq \sqrt{d}(d+1)\left\lceil \sqrt{d-1}\right\rceil (d-1)\frac{d(d+1)}{2}\left(\frac{3}{2}\right)^{\left\lceil \sqrt{d-1}\right\rceil (d-1)-2} \\
&\leq O\big(d^5\big)\left(\frac{3}{2}\right)^{\left\lceil \sqrt{d-1}\right\rceil (d-1)},
\end{aligned}$$

as desired. (The left inequality in (2) has only been stated for completeness; it will not be used later on.)

4. Bounding toric resolutions

Let C be a simplicial rational d-cone. The following lemma gives an upper bound for the number of steps in the standard procedure to equivariantly resolve the toric singularity $\mathrm{Spec}(k[\mathbb{Z}^d \cap C])$ (see [F, Section 2.6] and [O, Section 1.5] for the background). It depends on d and the multiplicity of Δ_C. Exponential factors enter our estimates only at this place. Therefore any improvement of the toric resolution bound would critically affect the order of magnitude of the estimates of $\mathfrak{c}_d^{\mathrm{pol}}$ and $\mathfrak{c}_d^{\mathrm{cone}}$.

Theorem 4.1. *Every rational simplicial d-cone $C \subset \mathbb{R}^d$, $d \geq 3$, admits a unimodular triangulation $C = D_1 \cup \cdots \cup D_T$ such that*

$$\mathrm{Hilb}(D_t) \subset \left(\frac{d}{2}\left(\frac{3}{2}\right)^{\mu(\Delta_C)-2}\right)\Delta_C, \quad t \in [1, T].$$

Proof. We use the sequence h_k, $k \geq -(d-2)$, of real numbers defined recursively as follows:

$$h_k = 1, \quad k \leq 1, \qquad h_2 = \frac{d}{2}, \qquad h_k = \frac{1}{2}(h_{k-1} + \cdots + h_{k-d}), \quad k \geq 3.$$

One sees easily that this sequence is increasing, and that

$$\begin{aligned} h_k &= \frac{1}{2}h_{k-1} + \frac{1}{2}(h_{k-2} + \cdots + h_{k-d-1}) - \frac{1}{2}h_{k-d-1} = \frac{3}{2}h_{k-1} - \frac{1}{2}h_{k-d-1} \\ &\leq \frac{d}{2}\left(\frac{3}{2}\right)^{k-2} \end{aligned}$$

for all $k \geq 2$.

Let $v_1, \dots, v_d$ be the extreme integral generators of C and denote by $\square_C$ the semi-open parallelotope

$$\left\{z \mid z = \xi_1 v_1 + \cdots + \xi_d v_d, \ \ 0 \leq \xi_1, \dots, \xi_d < 1\right\} \subset \mathbb{R}^d.$$

The cone C is unimodular if and only if

$$\square_C \cap \mathbb{Z}^d = \{O\}.$$

If C is unimodular then the bound given in the theorem is satisfied (note that $d \geq 3$). Otherwise we choose a non-zero lattice point, say w, from $\square_C$,

$$w = \xi_{i_1} v_{i_1} + \cdots + \xi_{i_k} v_{i_k}, \quad 0 < \xi_{i_j} < 1.$$

We can assume that w is in $(d/2)\Delta_C$. If not, then we replace w by

$$v_{i_1} + \cdots + v_{i_k} - w. \tag{4}$$

The cone C is triangulated into the simplicial d-cones

$$C_j = \mathbb{R}_+ v_1 + \cdots + \mathbb{R}_+ v_{i_j - 1} + \mathbb{R}_+ w + \mathbb{R}_+ v_{i_j+1} + \cdots + \mathbb{R}_+ v_d, \quad j = 1, \dots, k.$$

Call these cones the *second generation cones*, C itself being of *first generation.* (The construction of the cones C_j is called *stellar subdivision* with respect to w.)

For the second generation cones we have $\mu(\Delta_{C_i}) < \mu(\Delta_C)$ because the volumes of the corresponding parallelotopes are in the same relation. Therefore we are done if $\mu(\Delta_C) = 2$.

If $\mu(\Delta_C) \geq 3$, we generate the $(k+1)$st generation cones by successively subdividing the kth generation *non-unimodular* cones. It is clear that we obtain a triangulation of C if we use each vector produced to subdivide all kth generation cones to which it belongs. Figure 2 shows a typical situation after 2 generations of subdivision in the cross-section of a 3-cone.

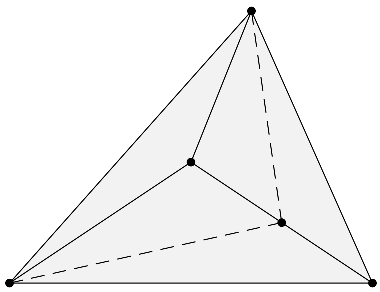


FIGURE 2.

If C'' is a next generation cone produced from a cone C', then $\mu(\Delta_{C''}) < \mu(\Delta_{C'})$, and it is clear that there exists $g \le \mu(\Delta_C)$ for which all cones of generation g are unimodular.
We claim that each vector $w^{(k)}$ subdividing a $(k-1)$st generation cone $C^{(k-1)}$ is in

$$h_k\ \Delta_C.$$

For $k = 2$ this has been shown already. So assume that $k \ge 3$. Note that all the extreme generators $u_1, \dots, u_d$ of $C^{(k-1)}$ either belong to the original vectors $v_1, \dots, v_d$ or were created in *different* generations. By induction we therefore have

$$u_i \in h_{t_i}\ \Delta_C, \qquad t_1, \dots, t_d \text{ pairwise different.}$$

Using the trick (4) if necessary, one can achieve that

$$w^{(k)} \in c\Delta_C, \qquad c \le \frac{1}{2}(h_{t_1} + \dots + h_{t_d}).$$

Since the sequence (h_i) is increasing,

$$c \le \frac{1}{2}(h_{k-1} + \dots + h_{k-d}) = h_k. \qquad \square$$

REMARK 4.2. (a) In dimension $d = 2$ the algorithm constructs a triangulation into unimodular cones D_t with $\operatorname{Hilb}(D_t) \subset \Delta_C$.
(b) For $d = 3$ one has Sebő's [S1] result $\operatorname{Hilb}(D_t) \subset 2\Delta_C$. It needs a rather tricky argument for the choice of w.

5. CORNER COVERS

Let C be a rational cone and v one of its extreme generators. We say that a system $\{C_j\}_{j=1}^k$ of subcones $C_j \subset C$ *covers the corner of* C *at* v if $v \in \operatorname{Hilb}(C_j)$ for all j and the union $\bigcup_{j=1}^k C_j$ contains a neighborhood of v in C.

LEMMA 5.1. *Suppose that* $\mathfrak{c}_{d-1}^{\text{cone}} < \infty$*, and let* C *be a simplicial rational* d*-cone with extreme generators* $v_1, \dots, v_d$*.*

(a) *Then there is a system of unimodular subcones* $C_1, \dots, C_k \subset C$ *covering the corner of* C *at* v_1 *such that* $\operatorname{Hilb}(C_1), \dots, \operatorname{Hilb}(C_k) \subset (\mathfrak{c}_{d-1}^{\text{cone}}+1)\Delta_C$.

(b) *Moreover, each element $w \neq v_1$ of a Hilbert basis of C_j, $j \in [1,k]$, has a representation $w = \xi_1 v_1 + \cdots + \xi_d v_d$ with $\xi_1 < 1$.*

Proof. For simplicity of notation we set $\mathfrak{c} = \mathfrak{c}_{d-1}^{\text{cone}}$. Let C' be the cone generated by $w_i = v_i - v_1$, $i \in [2,d]$, and let V be the vector subspace of $\mathbb{R}^d$ generated by the w_i. We consider the linear map $\pi : \mathbb{R}^d \to V$ given by $\pi(v_1) = 0$, $\pi(v_i) = w_i$ for $i > 0$, and endow V with a lattice structure by setting $\mathcal{L} = \pi(\mathbb{Z}^d)$. (One has $\mathcal{L} = \mathbb{Z}^d \cap V$ if and only if $\mathbb{Z}^d = \mathbb{Z}v_1 + (\mathbb{Z}^d \cap V)$.) Note that $v_1, z_2, \dots, z_d$ with $z_j \in \mathbb{Z}^d$ form a $\mathbb{Z}$-basis of $\mathbb{Z}^d$ if and only if $\pi(z_2), \dots, \pi(z_d)$ are a $\mathbb{Z}$-basis of $\mathcal{L}$. This holds since $\mathbb{Z}v_1 = \mathbb{Z}^d \cap \mathbb{R}v_1$, and explains the unimodularity of the cones C_j constructed below.
Note that $w_i \in \mathcal{L}$ for all i. Therefore $\Delta_{C'} \subset \operatorname{conv}(O, w_2, \dots, w_d)$. The cone C' has a unimodular covering (with respect to $\mathcal{L}$) by cones C'_j, $j \in [1,k]$, with $\operatorname{Hilb}(C'_j) \subset \mathfrak{c}\Delta_{C'}$. We "lift" the vectors $x \in \operatorname{Hilb}(C'_j)$ to elements $\tilde{x} \in C$ as follows. Let $x = \alpha_2 w_2 + \cdots + \alpha_d w_d$ (with $\alpha_i \in \mathbb{Q}_+$). Then there exists a unique integer $n \geq 0$ such that

$$\begin{aligned}\tilde{x} := nv_1 + x &= nv_1 + \alpha_2(v_2 - v_1) + \cdots + \alpha_d(v_d - v_1)\\ &= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_d v_d\end{aligned}$$

with $0 \leq \alpha_1 < 1$. (See Figure 3.) If $x \in \mathfrak{c}\Delta_{C'} \subset \mathfrak{c} \cdot \operatorname{conv}(O, w_2, \dots, w_d)$, then $\tilde{x} \in (\mathfrak{c}+1)\Delta_C$.

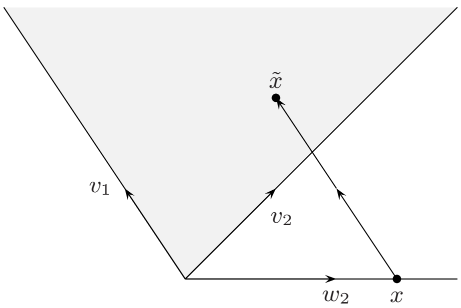


Figure 3.

We now define C_j as the cone generated by v_1 and the vectors $\tilde{x}$ where $x \in \operatorname{Hilb}(C'_j)$. It only remains to show that the C_j cover a neighborhood of v_1 in C. To this end we intersect C with the affine hyperplane $\mathcal{H}$ through $v_1, \dots, v_d$. It is enough that a neighborhood of v_1 in $C \cap \mathcal{H}$ is contained in $C_1 \cup \cdots \cup C_k$. For each $j \in [1,k]$ the coordinate transformation from the basis $w_2, \dots, w_d$ of V to the basis $x_2, \dots, x_d$ with $\{x_2, \dots, x_d\} = \operatorname{Hilb}(C'_j)$ defines a linear operator on $\mathbb{R}^{d-1}$. Let M_j be its $\| \ \|_\infty$ norm.
Moreover, let N_j be the maximum of the numbers n_i, $i \in [2,d]$ defined by the equation $\tilde{x}_i = n_i v_1 + x_i$ as above. Choose ε with

$$0 < \varepsilon \leq \frac{1}{(d-1)M_j N_j}, \qquad j \in [1,k].$$

and consider

$$y = v_1 + \beta_2 w_2 + \cdots + \beta_d w_d, \qquad 0 \le \beta_i < \varepsilon.$$

Since the C'_j cover C', one has $\beta_2 w_2 + \cdots + \beta_d w_d \in C'_j$ for some j, and therefore

$$y = v_1 + \gamma_2 x_2 + \cdots + \gamma_d x_d,$$

where $\{x_2, \dots, x_d\} = \operatorname{Hilb}(C'_j)$ and $0 \le \gamma_i \le M_j \varepsilon$ for $i \in [2, d]$. Then

$$y = \left(1 - \sum_{i=2}^{d} n_i \gamma_i\right) v_1 + \gamma_2 \tilde{x}_2 + \cdots + \gamma_d \tilde{x}_d$$

and

$$\sum_{i=2}^{d} n_i \gamma_i \le (d-1) N_j M_j \varepsilon \le 1,$$

whence $\left(1 - \sum_{i=2}^{d} n_i \gamma_i\right) \ge 0$ and $y \in C_j$, as desired. $\square$

6. The bound for cones

Before we embark on the proof of Theorem 1.3, we single out a technical step. Let $\{v_1, \dots, v_d\} \subset \mathbb{R}^d$ be a linearly independent subset. Consider the hyperplane

$$\mathcal{H} = \operatorname{Aff}(O, v_1 + (d-1)v_2, v_1 + (d-1)v_3, \dots, v_1 + (d-1)v_d) \subset \mathbb{R}^d$$

It cuts a simplex δ off the simplex $\operatorname{conv}(v_1, \dots, v_d)$ so that $v_1 \in \delta$. Let Φ denote the closure of

$$\mathbb{R}_+ \delta \setminus \big(\big((1 + \mathbb{R}_+) v_1 + \mathbb{R}_+ e_2 + \cdots + \mathbb{R}_+ v_d \big) \cup \Delta \big) \subset \mathbb{R}^d.$$

where $\Delta = \operatorname{conv}(O, v_1, \dots, v_d)$. See Figure 4 for the case $d = 2$. The polytope

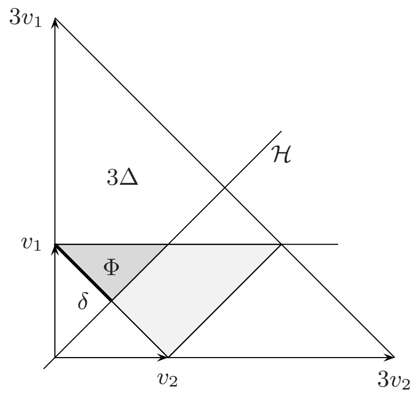


FIGURE 4.

$$\Phi' = -\frac{1}{d-1} v_1 + \frac{d}{d-1} \Phi$$

is the homothetic image of the polytope Φ under the dilatation with factor $d/(d-1)$ and center v_1. We will need that

$$\Phi' \subset (d+1)\Delta. \tag{5}$$

The easy proof is left to the reader.

Proof of Theorem 1.3. We want to prove the inequality

$$\mathfrak{c}_d^{\text{cone}} \le \left\lceil \sqrt{d-1} \right\rceil (d-1) \frac{d(d+1)}{2} \left(\frac{3}{2}\right)^{\left\lceil \sqrt{d-1} \right\rceil (d-1)-2} \tag{6}$$

for all $d \ge 2$ by induction on d.
The inequality holds for $d = 2$ since $\mathfrak{c}_2^{\text{cone}} = 1$ (see the remarks preceding Theorem 1.3 in Section 1), and the right hand side above is 2 for $d = 2$. By induction we can assume that (6) has been shown for all dimensions $< d$. We set

$$\gamma = \left\lceil \sqrt{d-1}\ \right\rceil (d-1) \quad \text{and} \quad \kappa = \gamma \frac{d(d+1)}{2} \left(\frac{3}{2}\right)^{\gamma-2}.$$

As pointed out in Remark 1.2, we can right away assume that C is empty simplicial with extreme generators $v_1, \dots, v_d$.

Outline. The following arguments are subdivided into four major steps. The first three of them are very similar to their analogues in the proof of Proposition 3.1. In Step 1 we cover the d-cone C by $d+1$ smaller cones each of which is bounded by the hyperplane that passes through the barycenter of $\operatorname{conv}(v_1, \dots, v_d)$ and is parallel to the facet of $\operatorname{conv}(v_1, \dots, v_d)$ opposite of v_i, $i = 1, \dots, d$. We summarize this step in Claim A below.
In Step 2 Lemma 5.1 is applied for the construction of unimodular corner covers. Claim B states that it is enough to cover the subcones of C 'in direction' of the cones forming the corner cover.
In Step 3 we extend the corner cover far enough into C. Lemma 2.2 allows us to do this within a suitable multiple of Δ_C. The most difficult part of the proof is to control the size of all vectors involved.
However, Lemma 2.2 is applied to simplices $\Gamma = \operatorname{conv}(w_1, \dots, w_e)$ where $w_1, \dots, w_e$ span a unimodular cone of dimension $e \le d$. The cones over the unimodular simplices covering $c\Gamma$ have multiplicity dividing c, and possibly equal to c. Nevertheless we obtain a cover of C by cones with *bounded* multiplicities. So we can apply Theorem 4.1 in Step 4 to obtain a unimodular cover.

Step 1. The facet $\operatorname{conv}(v_1, \dots, v_d)$ of Δ_C is denoted by Γ_0. (We use the letter Γ for $(d-1)$-dimensional simplices, and Δ for d-dimensional ones.) For $i \in [1, d]$ we put

$$\mathcal{H}_i = \operatorname{Aff}(O,\ v_i+(d-1)v_1, \dots, v_i+(d-1)v_{i-1},\ v_i+(d-1)v_{i+1}, \dots, v_i+(d-1)v_d)$$

and

$$\Gamma_i = \operatorname{conv}\big(v_i, \Gamma_0 \cap \mathcal{H}_i\big).$$

Observe that $v_1 + \cdots + v_d \in \mathcal{H}_i$. In particular, the hyperplanes $\mathcal{H}_i$, $i \in [1, d]$ contain the barycenter of Γ_0, i. e. $(1/d)(v_1 + \cdots + v_d)$. In fact, $\mathcal{H}_i$ is the vector subspace of dimension $d-1$ through the barycenter of Γ_0 that is parallel to the facet of Γ_0 opposite to v_i. Clearly, we have the representation $\bigcup_{i=1}^{d} \Gamma_i = \Gamma_0$, similar to (3) in Section 3. In particular, each of the Γ_i is homothetic to Γ_0 with factor $(d-1)/d$.
To prove (6) it is enough to show the following

Claim A. For each index $i \in [1, d]$ there exists a system of unimodular cones

$$C_{i1}, \ldots, C_{ik_i} \subset C$$

such that $\operatorname{Hilb}(C_{ij}) \subset \kappa\Delta_C$, $j \in [1, k_i]$, and $\Gamma_i \subset \bigcup_{j=1}^{k_i} C_{ij}$.

The step from the original claim to the reduction expressed by Claim A seems rather small – we have only covered the cross-section Γ_0 by the Γ_i, and state that it is enough to cover each Γ_i by unimodular subcones. The essential point is that these subcones need not be contained in the cone spanned by Γ_i, but just in C. This gives us the freedom to start with a corner cover at v_i and to extend it far enough into C, namely beyond $\mathcal{H}_i$. This is made more precise in the next step.

Step 2. To prove Claim A it is enough to treat the case $i = 1$. The induction hypothesis implies $\mathfrak{c}_{d-1}^{\text{cone}} \leq \kappa - 1$ because the right hand side of the inequality (6) is a strictly increasing function of d. Thus Lemma 5.1 provides a system of unimodular cones $C_1, \ldots, C_k \subset C$ covering the corner of C at v_1 such that

$$\operatorname{Hilb}(C_j) \setminus \{v_1, \ldots, v_d\} \subset \big(\kappa\Delta_C\big) \setminus \Delta_C, \quad j \in [1, k]. \tag{7}$$

Here we use the emptiness of Δ_C – it guarantees that $\operatorname{Hilb}(C_j) \cap (\Delta_C \setminus \Gamma_0) = \emptyset$ which is crucial for the inclusion (9) in Step 3.
With a suitable enumeration $\{v_{j1}, \ldots, v_{jd}\} = \operatorname{Hilb}(C_j)$, $j \in [1, k]$ we have $v_{11} = v_{21} = \cdots = v_{k1} = v_1$ and

$$0 \leq (v_{jl})_{v_1} < 1, \quad j \in [1, k], \quad l \in [2, d], \tag{8}$$

where $(-)_{v_1}$ is the first coordinate of an element of $\mathbb{R}^d$ with respect to the basis $v_1, \ldots, v_d$ of $\mathbb{R}^d$ (see Lemma 5.1(b)).
Now we formulate precisely what it means to extend the corner cover beyond the hyperplane $\mathcal{H}_1$. Fix an index $j \in [1, k]$ and let $D \subset \mathbb{R}^d$ denote the simplicial d-cone determined by the following conditions:

(i) $C_j \subset D$,
(ii) the facets of D contain those facets of C_j that pass through O and v_1,
(iii) the remaining facet of D is in $\mathcal{H}_1$.

Figure 5 describes the situation in the cross-section Γ_0 of C.
By considering all possible values $j = 1, \ldots, k$, it becomes clear that to prove Claim A it is enough to prove

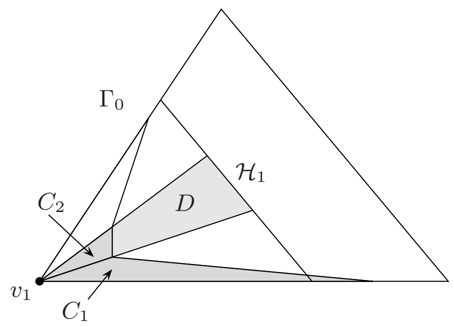


Figure 5.

Claim B. There exists a system of unimodular cones $D_1, \ldots, D_T \subset C$ such that

$$\operatorname{Hilb}(D_t) \subset \kappa\Delta_C, \quad t \in [1, T] \qquad \text{and} \qquad D \subset \bigcup_{t=1}^{T} D_t.$$

Step 3. For simplicity of notation we put $\Delta = \Delta_{C_j}$, $\mathcal{H} = \mathcal{H}_1$. (Recall that Δ is of dimension d, spanned by O and the extreme integral generators of C_j.) The vertices of Δ, different from O and v_1 are denoted by $w_2, \ldots, w_d$ in such a way that there exists i_0, $1 \leq i_0 \leq d$, for which

(i) $w_2, \ldots, w_{i_0} \in D \setminus \mathcal{H}$ ('bad' vertices, on the same side of $\mathcal{H}$ as v_1),
(ii) $w_{i_0+1}, \ldots w_d \in \overline{C_j \setminus D}$ ('good' vertices, beyond or on $\mathcal{H}$),

neither $i_0 = 1$ nor $i_0 = d$ being excluded. ($\overline{X}$ is the closure of $X \subset \mathbb{R}^d$ with respect to the Euclidean topology.) In the situation of Figure 5 the cone C_2 has two bad vertices, whereas C_1 has one good and one bad vertex. (Of course, we see only the intersection points of the cross-section Γ_0 with the rays from O through the vertices.)
If all vertices are good, there is nothing to prove since $D \subset C_j$ in this case. So assume that there are bad vertices, i. e. $i_0 \geq 2$. We now show that the bad vertices are caught in a compact set whose size with respect to Δ_C depends only on d, and this fact makes the whole proof work.
Consider the $(d-1)$-dimensional cone

$$E = v_1 + \mathbb{R}_+(w_2 - v_1) + \cdots + \mathbb{R}_+(w_d - v_1).$$

In other words, E is the $(d-1)$-dimensional cone with apex v_1 spanned by the facet $\operatorname{conv}(v_1, w_2, \ldots, w_d)$ of Δ opposite to O. It is crucial in the following that the simplex $\operatorname{conv}(v_1, w_2, \ldots, w_d)$ is unimodular (with respect to $\mathbb{Z}^d \cap \operatorname{Aff}(v_1, w_2, \ldots, w_d)$), as follows from the unimodularity of C_j.
Due to the inequality (8) the hyperplane $\mathcal{H}$ cuts a $(d-1)$-dimensional (possibly non-lattice) simplex off the cone E. We denote this simplex by Γ. Figure 6 illustrates the situation by a vertical cross-section of the cone C.

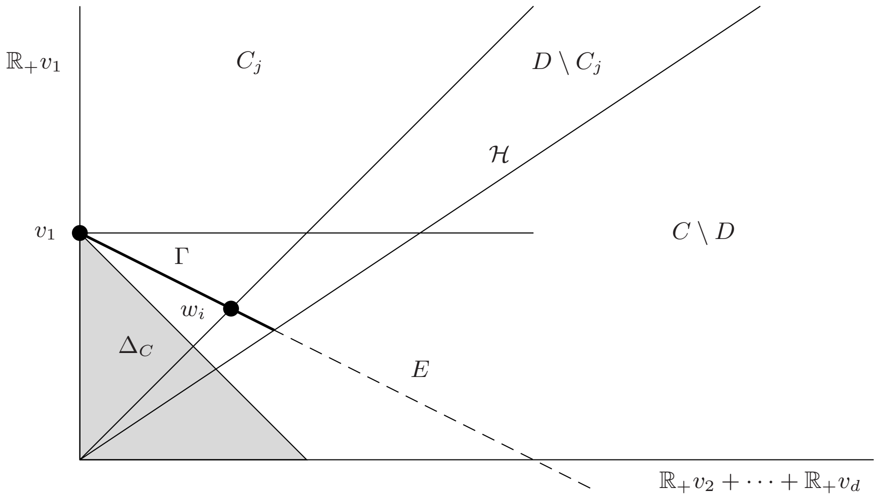


Figure 6.

By (7) and (8) we have

$$\Gamma \subset \Phi = \overline{\mathbb{R}_+\Gamma_1 \setminus \big((v_1 + C) \cup \Delta_C\big)}.$$

Let ϑ be the dilatation with center v_1 and factor $d/(d-1)$. Then by (5) we have the inclusion

$$\vartheta(\Gamma) \subset (d+1)\Delta_C. \tag{9}$$

One should note that this inclusion has two aspects: first it shows that Γ is not too big with respect to Δ_C. Second, it guarantees that there is some $\zeta > 0$ *only depending on* d, namely $\zeta = 1/(d-1)$, such that the dilatation with factor $1+\zeta$ and center v_1 keeps Γ inside C. If ζ depended on C, there would be no control on the factor c introduced below.

Let $\Sigma_1 = \operatorname{conv}(v_1, w_2, \dots, w_{i_0})$ and Σ_2 be the smallest face of Γ that contains Σ_1. These are d'-dimensional simplices, $d' = i_0 - 1$. Note that $\Sigma_2 \subset \vartheta(\Sigma_2)$.

We want to apply Lemma 2.2 to the pair

$$\gamma v_1 + (\Sigma_1 - v_1) \subset \gamma v_1 + (\Sigma_2 - v_1).$$

of simplices with the common vertex γv_1. The lattice of reference for the unimodular covering is

$$\mathcal{L} = \mathcal{L}_{\gamma v_1 + (\Sigma_1 - v_1)} = \gamma v_1 + \sum_{j=2}^{i_0} \mathbb{Z}(w_j - v_1).$$

Set

$$\varepsilon = \frac{1}{d} \quad \text{and} \quad c = \frac{d}{d-1}\gamma = \left\lceil \sqrt{d-1}\, \right\rceil d.$$

Since $d' \leq d-1$, Lemma 2.2 (after the parallel translation of the common vertex to O and then back to γv_1) and (9) imply

$$\gamma\Sigma_2 \subset \mathrm{UC}_{\mathcal{L}}\big(\gamma\vartheta(\Sigma_2)\big) \subset \gamma(d+1)\Delta_C. \tag{10}$$

Step 4. Consider the i_0-dimensional simplices spanned by O and the unimodular (i_0-1)-simplices appearing in (10). Their multiplicities with respect to the i_0-rank lattice $\mathbb{Z}\mathcal{L}_{\Sigma_1}$ are all equal to γ, since Σ_1, a face of $\mathrm{conv}(v_1, w_2, \dots, w_d)$ is unimodular and, thus, we have unimodular simplices σ on height γ. The cones $\mathbb{R}_+\sigma$ have multiplicity dividing γ. Therefore, by Lemma 4.1 we conclude that the i_0-cone $\mathbb{R}_+\Sigma_2$ is in the union $\delta_1 \cup \cdots \cup \delta_T$ of unimodular (with respect to the lattice $\mathbb{Z}\mathcal{L}_{\Sigma_1}$) cones such that

$$\begin{aligned}\mathrm{Hilb}(\delta_1), \dots, \mathrm{Hilb}(\delta_T) &\subset \left(\frac{d}{2}\left(\frac{3}{2}\right)^{\gamma-2}\right)\Delta_{\mathbb{R}_+\Sigma_2}\\ &\subset \left(\frac{d}{2}\left(\frac{3}{2}\right)^{\gamma-2}\right)\gamma(d+1)\Delta_C = \kappa\Delta_C.\end{aligned}$$

In view of the unimodularity of $\mathrm{conv}(v_1, w_2, \dots, w_d)$, the subgroup $\mathbb{Z}\mathcal{L}_{\Sigma_1}$ is a direct summand of $\mathbb{Z}^d$. It follows that

$$D_t = \delta_t + \mathbb{R}_+w_{i_0+1} + \cdots + \mathbb{R}_+w_d, \quad t \in [1, T],$$

is the desired system of unimodular cones. $\square$

References

[BG] W. Bruns and J. Gubeladze, *Normality and covering properties of affine semigroups*, J. Reine Angew. Math. 510, 151–178, 1999.

[BGT1] W. Bruns, J. Gubeladze, and N. V. Trung, *Normal polytopes, triangulations, and Koszul algebras*, J. Reine Angew. Math. 485 (1997), 123–160.

[BGT2] W. Bruns, J. Gubeladze, and N. V. Trung, *Problems and algorithms for affine semigroups*, Semigroup Forum 64 (2002), 180–212.

[EW] G. Ewald and U. Wessels, *On the ampleness of invertible sheaves in complete projective toric varieties*, Result. Math. 19 (1991), 275-278.

[F] W. Fulton, *Introduction to toric varieties*, Princeton University Press, 1993.

[HZ] C. Haase and G. M. Ziegler, *On the maximal width of empty lattice simplices*, Eur. J. Comb. 21 (2000), 111-119.

[KS] J.-M. Kantor and K. S. Sarkaria, *On primitive subdivisions of an elementary tetrahedron*, preprint.

[KKMS] G. Kempf, F. Knudsen, D. Mumford, and B. Saint-Donat, *Toroidal embeddings I*, Lecture Notes in Math. 339, Springer, 1973.

[O] T. Oda, *Convex bodies and algebraic geometry (An introduction to the theory of toric varieties)*, Springer, 1988.

[S1] A. Sebő, *Hilbert bases, Carathéodory's theorem, and combinatorial optimization*, in 'Integer Programming and Combinatorial Optimization'

(R. Kannan, W. Pulleyblank, eds.), University of Waterloo Press, Waterloo 1990, 431–456.

[S2] A. Sebő, *An introduction to empty lattice simplices*, Lect. Notes Comput. Sci. 1610 (1999), 400–414.

Winfried Bruns
Universität Osnabrück
FB Mathematik/Informatik
49069 Osnabrück
Germany
winfried@mathematik.
uni-osnabrueck.de

Joseph Gubeladze
A. Razmadze Mathematical
Institute
Alexidze St. 1
380093 Tbilisi
Georgia
gubel@rmi.acnet.ge

Rost Projectors and Steenrod Operations

Nikita Karpenko and Alexander Merkurjev[1]

Received: September 29, 2002

Communicated by Ulf Rehmann

Abstract. Let X be an anisotropic projective quadric possessing a Rost projector ρ. We compute the 0-dimensional component of the total Steenrod operation on the modulo 2 Chow group of the Rost motive given by the projector ρ. The computation allows to determine the whole Chow group of the Rost motive and the Chow group of every excellent quadric (the results announced by Rost). On the other hand, the computation is being applied to give a simpler proof of Vishik's theorem stating that the integer $\dim X + 1$ is a power of 2.

2000 Mathematics Subject Classification: 11E04; 14C25
Keywords and Phrases: quadratic forms, Chow groups and motives, Steenrod operations

M. Rost noticed that certain smooth projective anisotropic quadric hypersurfaces are decomposable in the category of Chow motives into a direct sum of some motives. The (in some sense) smallest direct summands are called the *Rost motives*. For example, the motive of a Pfister quadric is a direct sum of Rost motives and their Tate twists. The *Rost projectors* split off the Rost motives as direct summands of quadrics. In the present paper we study Rost projectors by means of modulo 2 Steenrod operations on the Chow groups of quadrics. The Steenrod operations in motivic cohomology were defined by V. Voevodsky. We use results of P. Brosnan who found in [1] an elementary construction of the Steenrod operations on the Chow groups.

As a consequence of our computations we give a description of the Chow groups of a Rost motive (Corollary 8.2). This result (which has been announced by M. Rost in [11]) allows to compute all the Chow groups of every excellent quadric (see Remark 8.4).

We also give a simpler proof of a theorem of A. Vishik [3, th. 6.1] stating that if an anisotropic quadric X possesses a Rost projector, then $\dim X + 1$ is a power of 2 (Theorem 5.1).

[1]The second author was supported in part by NSF Grant #0098111.

Contents

1. Parity of binomial coefficients

Lemma 1.1. *Let i, n be any non-negative integers. The binomial coefficient $\binom{n+i}{i}$ is odd if and only if we don't carry over units while adding n and i in base 2.*

Proof. For any integer $a \geq 0$, let $s_2(a)$ be the sum of the digits in the base 2 expansion of a. By [9, Lemma 5.4(a)], $\binom{n+i}{i}$ is odd if and only if $s_2(n+i) = s_2(n) + s_2(i)$. □

The following statement is obvious:

Lemma 1.2. *For any non-negative integer m, we don't carry over units while adding m and $m+1$ in base 2 if and only if $m+1$ is a power of 2.* □

The following statement will be applied in the proof of Theorem 4.8:

Corollary 1.3. *For any non-negative integer m, the binomial coefficient $\binom{-m-2}{m}$ is odd if and only if $m+1$ is a power of 2.*

Proof. By Lemma 1.1, the binomial coefficient

$$\binom{-m-2}{m} = (-1)^m \binom{2m+1}{m}$$

is odd if and only if we don't carry over units while adding m and $m+1$ in base 2. It remains to apply Lemma 1.2. □

2. Integral and modulo 2 Rost projectors

Let F be a field, X a quasi-projective smooth equidimensional variety over F. We write $\mathrm{CH}(X)$ for the *modulo* 2 Chow group of X. The usual (integral) Chow group is denoted by $\mathbb{CH}(X)$. We are working mostly with $\mathrm{CH}(X)$, but several times we have to use $\mathbb{CH}(X)$ (for example, already the definition of a modulo 2 Rost correspondence cannot be given on the level of the modulo 2 Chow group).

Both groups are graded. We use the upper indices for the gradation by codimension of cycles and we use the lower indices for the gradation by the dimension of cycles.

For projective X_1 and X_2, an element $\rho \in \mathbb{CH}(X_1 \times X_2)$ (we do not consider the gradation on $\mathbb{CH}$ for the moment) can be viewed as a correspondence from X_1 to X_2 ([2, §16.1]). In particular, it gives a homomorphism [2, def. 16.1.2]

$$\rho_* \colon \mathbb{CH}(X_1) \to \mathbb{CH}(X_2), \quad \rho_*(\alpha) = pr_{2*}\left(pr_1^*(\alpha) \cdot \rho\right),$$

where pr_1 and pr_2 are the two projections of $X_1 \times X_2$ onto X_1 and X_2, and can be composed with another correspondence $\rho' \in \mathbb{CH}(X_2 \times X_3)$ [2, def. 16.1.1]. The same can be said and defined with $\mathbb{CH}$ replaced by CH.

Starting from here, we always assume that $\operatorname{char} F \neq 2$. Let φ be a non-degenerate quadratic form over F, and let X be the projective quadric $\varphi = 0$. We set $n = \dim X = \dim \varphi - 2$ and we assume that $n \geq 1$.

An element $\varrho \in \mathbb{CH}^n(X \times X)$ is called an (integral) *Rost correspondence*, if over an algebraic closure $\bar{F}$ of F one has:

$$\varrho_{\bar{F}} = [\bar{X} \times x] + [x \times \bar{X}] \in \mathbb{CH}^n(\bar{X} \times \bar{X})$$

with $\bar{X} = X_{\bar{F}}$ and a rational point $x \in \bar{X}$. A *Rost projector* is a Rost correspondence which is an idempotent with respect to the composition of correspondences.

Remark 2.1. Assume that the quadric X is isotropic, i.e., contains a rational closed point $x \in X$. Then $[X \times x] + [x \times X]$ is a Rost projector. Moreover, this is the unique Rost projector on X ([7, lemma 4.1]).

Remark 2.2. Let ϱ be a Rost correspondence on X. It follows from the Rost nilpotence theorem ([12, prop. 1]) that a certain power of ϱ is a Rost projector (see [7, cor. 3.2]). In particular, a quadric X possesses a Rost projector if and only if it possesses a Rost correspondence.

A *modulo* 2 *Rost correspondence* $\rho \in \mathrm{CH}^n(X \times X)$ is a correspondence which can be represented by an integral Rost correspondence. A *modulo* 2 *Rost projector* is an idempotent modulo 2 Rost correspondence. Clearly, a modulo 2 Rost correspondence represented by an integral Rost projector is a modulo 2 Rost projector. Conversely,

Lemma 2.3. *A modulo* 2 *Rost projector is represented by an integral Rost projector.*

Proof. Let ρ be a modulo 2 Rost projector and let ϱ be an integral Rost correspondence representing ρ. The correspondence $\varrho_{\bar{F}}$ is idempotent, therefore, by the Rost nilpotence theorem (see [7, th. 3.1]), ϱ^r is idempotent for some r; so, ϱ^r is an integral Rost projector. Since ρ is idempotent as well, ϱ^r still represents ρ. □

Lemma 2.4. *Let ϱ be an integral Rost correspondence and let ρ be a modulo* 2 *Rost correspondence. Then ϱ_* is the identity on $\mathbb{CH}^0(X)$ and on $\mathbb{CH}_0(X)$;*

also ρ_* *is the identity on* $\mathrm{CH}^0(X)$ *and on* $\mathrm{CH}_0(X)$*. Moreover, for every* i *with* $0 < i < n$*, the group* $\varrho_* \,\mathbb{CH}^i(X)$ *vanishes over* $\bar{F}$*.*

Proof. It suffices to prove the statements on ϱ. Since $\mathbb{CH}^0(X)$ and $\mathbb{CH}_0(X)$ inject into $\mathbb{CH}^0(X_{\bar{F}})$ and $\mathbb{CH}_0(X_{\bar{F}})$ (see [5, prop. 2.6] or [13] for the statement on $\mathbb{CH}_0(X)$), it suffices to consider the case where the quadric X has a rational closed point x and $\varrho = [X \times x] + [x \times X]$. Since $[X \times x]_*([X]) = 0$, $[X \times x]_*([x]) = [x]$, $[x \times X]_*([X]) = [X]$, $[x \times X]_*([x]) = 0$ and since $[X]$ generates $\mathbb{CH}^0(X)$ while $[x]$ generates $\mathbb{CH}_0(X)$, we are done with the statements on $\mathbb{CH}^0(X)$ and on $\mathbb{CH}_0(X)$. Since $[X \times x]_*([Z]) = 0 = [x \times X]_*([Z])$ for any closed subvariety $Z \subset X$ of codimension $\neq 0, n$, we are done with the rest. $\square$

3. Steenrod operations

In this section we briefly recall the basic properties of the Steenrod operations on the modulo 2 Chow groups constructed in [1].

Let X be a smooth quasi-projective equidimensional variety over a field F. For every $i \geq 0$ there are certain homomorphisms $S^i : \mathrm{CH}^*(X) \to \mathrm{CH}^{*+i}(X)$ called *Steenrod operations*; their sum (which is in fact finite because $S^i = 0$ for $i > \dim X$)

$$S = S_X = S^0 + S^1 + \cdots : \mathrm{CH}(X) \to \mathrm{CH}(X)$$

is the *total Steenrod operation* (we omit the $*$ in the notation of the Chow group to indicate that S is not homogeneous). They have the following basic properties (see [1] for the proofs): for any smooth quasi-projective F-scheme X, the total operation $S\colon \mathrm{CH}(X) \to \mathrm{CH}(X)$ is a ring homomorphism such that for every morphism $f\colon Y \to X$ of smooth quasi-projective F-schemes and for every field extension E/F, the squares

$$\begin{array}{ccc} \mathrm{CH}(Y) & \xrightarrow{S_Y} & \mathrm{CH}(Y) \\ \uparrow{\scriptstyle f^*} & & {\scriptstyle f^*}\uparrow \\ \mathrm{CH}(X) & \xrightarrow{S_X} & \mathrm{CH}(X) \end{array} \quad \text{and} \quad \begin{array}{ccc} \mathrm{CH}(X_E) & \xrightarrow{S_{X_E}} & \mathrm{CH}(X_E) \\ \uparrow{\scriptstyle \mathrm{res}_{E/F}} & & {\scriptstyle \mathrm{res}_{E/F}}\uparrow \\ \mathrm{CH}(X) & \xrightarrow{S_X} & \mathrm{CH}(X) \end{array}$$

are commutative. Moreover, the restriction $S^i|_{\mathrm{CH}^n(X)}$ is 0 for $n < i$ and the map $\alpha \mapsto \alpha^2$ for $n = i$; finally S^0 is the identity.

Also, the total Steenrod operation satisfies the following Riemann-Roch type formula:

$$f_*\big(S_Y(\alpha) \cdot c(-T_Y)\big) = S_X\big(f_*(\alpha)\big) \cdot c(-T_X)$$

(in other words, S modified by $c(-T)$ this way, commutes with the push-forwards) for any proper $f\colon Y \to X$ and any $\alpha \in \mathrm{CH}(Y)$, where $f_*\colon \mathrm{CH}(Y) \to \mathrm{CH}(X)$ is the push-forward, c is the total Chern class, T_X is the tangent bundle of X, and $c(-T_X) = c^{-1}(T_X)$ (the expression $-T_X$ makes sense if one considers T_X as an element of $K_0(X)$). This formula is proved in [1]. It also follows from the previously formulated properties of S by the general Riemann-Roch theorem of Panin [10].

Lemma 3.1. *Assume that X is projective. For any $\alpha \in \mathrm{CH}(X)$ and for any $\rho \in \mathrm{CH}(X \times X)$, one has*

$$S_X(\rho_*(\alpha)) = S_{X\times X}(\rho)_*\big(S_X(\alpha)\cdot c(-T_X)\big) ,$$

where T_X is the class in $K_0(X)$ of the tangent bundle of X.

Proof. Let $pr_1, pr_2 \colon X \times X \to X$ be the first and the second projections. By the Riemann-Roch formula applied to the morphism pr_2, one has

$$S_X(\rho_*(\alpha)) = pr_{2*}\Big(S_{X\times X}\big(\, pr_1^*(\alpha)\cdot\rho\big)\cdot c(-T_{X\times X})\Big)\cdot c(T_X) .$$

By the projection formula for pr_2, this gives

$$S_X(\rho_*(\alpha)) = pr_{2*}\Big(S_{X\times X}(pr_1^*(\alpha)\cdot\rho)\cdot c\big(-T_{X\times X}+pr_2^*(T_X)\big)\Big) .$$

Since

$$T_{X\times X} = pr_1^*(T_X) + pr_2^*(T_X) \in K_0(X\times X)$$

and since S (as well as c) commutes with the products and the pull-backs, we get

$$pr_{2*}\Big(\, pr_1^*\left(S_X(\alpha)\cdot c(-T_X)\right)\cdot S_{X\times X}(\rho)\Big) = S_{X\times X}(\rho)_*\Big(S_X(\alpha)\cdot c(-T_X)\Big) .$$

□

4. Main theorem

In this section, let φ be an anisotropic quadratic form over F, and let X be the projective quadric $\varphi = 0$ with $n = \dim X = \dim\varphi - 2 \geq 1$. We are assuming that an integral Rost projector (see §2 for the definition) $\varrho \in \mathbb{CH}^n(X\times X)$ exists for our X and we write $\rho \in \mathrm{CH}^n(X\times X)$ for the modulo 2 Rost projector. We write h for the class in $\mathrm{CH}^1(X)$ (as well as in $\mathbb{CH}^1(X)$) of a hyperplane section of X.

Proposition 4.1. *One has for every $i \geq 0$:*

$$S\big(\rho_*(h^i)\big) = S(\rho)_*\big(h^i\cdot(1+h)^{i-n-2}\big) .$$

Proof. Since $h \in \mathrm{CH}^1(X)$, we have $S(h) = S^0(h) + S^1(h)$. Since $S^0 = \mathrm{id}$ while S^1 on $\mathrm{CH}^1(X)$ is the squaring, $S(h) = h + h^2 = h(1+h)$, hence $S(h^i) = S(h)^i = h^i(1+h)^i$. To finish the proof it suffices to check that $c(T_X) = (1+h)^{n+2}$ for the tangent bundle T_X of the quadric X: then the formula of Lemma 3.1 will give the formula of Proposition 4.1.

Let $i\colon X \hookrightarrow P$ be the embedding of X into the $(n+1)$-dimensional projective space P. Let us write H for the class in $\mathrm{CH}^1(P)$ of a hyperplane. Note that $h = i^*(H)$.

The exact sequence of vector X-bundles

$$0 \to T_X \to i^*(T_P) \to i^*(\mathcal{O}_P(2)) \to 0$$

gives the equality $c(T_X)\cdot i^*\big(c(\mathcal{O}_P(2))\big) = i^*\big(c(T_P)\big)$. Since $c(\mathcal{O}_P(2)) = 1+2H = 1$ (we are working with the *modulo* 2 Chow groups) and $c(T_P) = (1+H)^{n+2}$, we get $c(T_X) = (1+h)^{n+2}$. □

Lemma 4.2. *Let L/F be a field extension such that the quadric X_L is isotropic. Then $S^i(\rho_L) = 0$ for every $i > 0$.*

Proof. By the uniqueness of a modulo 2 Rost projector on an isotropic quadric (Remark 2.1) $\rho_L = [X] \times [x] + [x] \times [X]$, where $x \in X_L$ is a rational point. Since $S = S^0 = \mathrm{id}$ on $\mathrm{CH}^0(X_L) \ni [X]$ as well as on $\mathrm{CH}^n(X_L) \ni [x]$, we have

$$S(\rho_L) = S([X] \times [x] + [x] \times [X]) = S([X]) \times S([x]) + \\ S([x]) \times S([X]) = [X] \times [x] + [x] \times [X] = \rho_L = S^0(\rho_L)\ .$$

□

Lemma 4.3. *The Witt index of the quadratic form $\varphi_{F(X)}$ is 1.*

Proof. Let $Y \subset X$ be a subquadric of codimension 1. If $i_W(\varphi_{F(X)}) > 1$, Y has a rational point over $F(X)$. Therefore there exists a rational morphism $X \to Y$. Let $\alpha \in \mathbb{CH}^n(X \times X)$ be the correspondence given by the closure of the graph of this morphism. Let us show that $\Delta^*(\varrho \circ \alpha) \in \mathbb{CH}_0(X)$, where $\Delta \colon X \to X \times X$ is the diagonal morphism, is an element of $\mathbb{CH}_0(X)$ of degree 1 (giving a contradiction with the fact that the quadric X is anisotropic, this will finish the proof).
Clearly, verifying the assertion on the degree, we may replace F by a field extension of F. Therefore, we may assume that X has a rational point x. In this case $\varrho = [X \times x] + [x \times X]$. Since $[x \times X] \circ \alpha = 0$ (because $\dim Y < \dim X$) while $[X \times x] \circ \alpha = [X \times x]$, we get $\Delta^*(\varrho \circ \alpha) = [x]$. □

Lemma 4.4. *Let X be an anisotropic F-quadric such that the Witt index of the quadratic form $\varphi_{F(X)}$ is 1. Then for every $\alpha \in \mathbb{CH}_i(X_{F(X)})$, $i > 0$, the degree of the 0-cycle class $h^i \cdot \alpha$ is even.*

Proof. It is sufficient to consider the case $i = 1$. We have $\varphi_{F(X)} \simeq \psi \perp \mathbb{H}$ for an anisotropic quadratic form ψ over $F(X)$ (where $\mathbb{H}$ is a hyperbolic plane). Let X' be the quadric $\psi = 0$ over $F(X)$. There is an isomorphism [5, §2.2]

$$f : \mathbb{CH}_1(X_{F(X)}) \to \mathbb{CH}_0(X')$$

taking h^{n-1} to the class of a closed point of degree 2. Since the quadric X' is anisotropic, the group $\mathbb{CH}_0(X')$ is generated by the class of a degree 2 closed point (see [5, prop. 2.6] or [13]); therefore the group $\mathbb{CH}_1(X_{F(X)})$ is generated by h^{n-1}. Since $\deg(h \cdot h^{n-1}) = 2$, it follows that the integer $\deg(h \cdot \alpha)$ is even for every $\alpha \in \mathbb{CH}_1(X_{F(X)})$. □

For the modulo 2 Chow groups we get

Corollary 4.5. *Let $\mu \in \mathrm{CH}_i(X \times X)$ for some $0 < i \le n$ be a correspondence such that $\mu_{F(X)} = 0$. Then $\mu_*(h^i) = 0$.*

Proof. We replace μ by its representative in $\mathbb{CH}_i(X \times X)$ and we mean by h the integral class of a hyperplane section of X in the proof (while in the statement h is the class of a hyperplane section in the modulo 2 Chow group). Since the degree homomorphism $\deg : \mathbb{CH}^n(X) \to \mathbb{Z}$ is injective ([5, prop. 2.6] or [13]) with the image $2\mathbb{Z}$, it suffices to show that $\deg(\mu_*(h^i))$ is divisible by 4. Let us compute this degree. By definition of μ_*, we have $\mu_*(h^i) = pr_{2*}\big(\mu \cdot pr_1^*(h^i)\big)$. Note that the product $\mu \cdot pr_1^*(h^i)$ is in $\mathbb{CH}_0(X \times X)$ and the square

$$\begin{array}{ccc} \mathbb{CH}_0(X \times X) & \xrightarrow{pr_{2*}} & \mathbb{CH}_0(X) \\ {\scriptstyle pr_{1*}}\downarrow & & {\scriptstyle \deg}\downarrow \\ \mathbb{CH}_0(X) & \xrightarrow{\deg} & \mathbb{Z} \end{array}$$

commutes (the two compositions being the degree homomorphism of the group $\mathbb{CH}_0(X \times X)$). Therefore the degree of $\mu_*(h^i)$ coincides with the degree of $pr_{1*}\big(\mu \cdot pr_1^*(h^i)\big)$. By the projection formula for pr_{1*} the latter element coincides with the product $h^i \cdot pr_{1*}(\mu)$.
We are going to check that the degree of this element is divisible by 4. Since the degree does not change under extensions of the base field, it suffices to verify the divisibility relation over $F(X)$. The class $pr_{1*}(\mu)_{F(X)}$ is divisible by 2 by assumption, therefore the statement follows from Lemmas 4.3 and 4.4. □

Corollary 4.6. *$S^{n-i}(\rho)_*(h^i) = 0$ for every i with $0 < i < n$.*

Proof. We take $\mu = S^{n-i}(\rho)$. Since $i < n$, we have $\mu_{F(X)} = 0$ by Lemma 4.2. Since $i > 0$, we may apply Corollary 4.5 obtaining $\mu_*(h^i) = 0$. □

Putting together Corollary 4.6 and Proposition 4.1, we get

Corollary 4.7. *For every $i > 0$, one has:*

$$S^{n-i}\big(\rho_*(h^i)\big) = \binom{i-n-2}{n-i} \cdot \rho_*(h^n)\ .$$

Proof. By Proposition 4.1, $S^{n-i}\big(\rho_*(h^i)\big)$ is the n-codimensional component of $S(\rho)_*\big(h^i \cdot (1+h)^{i-n-2}\big)$; moreover, according to Corollary 4.6, $S(\rho)$ can be replaced by $S^0(\rho) = \rho$. □

Finally, by Corollary 1.3, computing the binomial coefficient modulo 2, together with Lemma 2.4, computing $\rho_*(h^n)$, we get

Theorem 4.8. *Suppose that the anisotropic quadric X of dimension n possesses a Rost projector. Let ρ be a modulo 2 Rost projector on X and let i be an integer with $0 < i < n$. Then*

$$S^{n-i}\big(\rho_*(h^i)\big) = h^n$$

in $\mathrm{CH}_0(X)$ if (and only if) the integer $n - i + 1$ is a power of 2. □

As the quadric X is anisotropic, $\mathbb{CH}_0(X)$ is an infinite cyclic group generated by h^n (see [5, prop. 2.6] or [13]); in particular, h^n in $\mathrm{CH}_0(X)$ is not 0. Therefore we get

Corollary 4.9. *For every i such that $0 < i < n$ and $n - i + 1$ is a power of 2, the element $S^{n-i}(\rho_*(h^i))$ (and consequently $\rho_*(h^i)$) is non-zero.* □

5. Dimensions of quadrics with Rost projectors

The following Theorem is proved in [3]. The proof given there makes use of the Steenrod operations in the motivic cohomology constructed by Voevodsky (since Voevodsky has announced that the operations were constructed in any characteristic $\neq 2$ only quite recently, the assumption $\operatorname{char} F = 0$ was made in [3]). Here we give an elementary proof.

Theorem 5.1 ([3, th. 6.1]). *If X is an anisotropic smooth projective quadric possessing a Rost projector, then $\dim X + 1$ is a power of 2.*

Proof. Let us assume that this is not the case. Let r be the largest integer such that $n > 2^r - 1$ where $n = \dim X$. Then Theorem 4.8 applies to $i = n - (2^r - 1)$, stating that $S^{n-i}\big(\rho_*(h^i)\big) \neq 0$. Note that $n - i \geq i$. Since the Steenrod operation S^i is trivial on $\mathrm{CH}^j(X)$ with $i > j$, it follows that $n - i = i$ and therefore $S^{n-i}\big(\rho_*(h^i)\big) = \rho_*(h^i)^2$. Since the element $\varrho_*(h^i)$ (where ϱ is the integral Rost projector) vanishes over $\bar{F}$ (Lemma 2.4), its square vanishes over $\bar{F}$ as well. The group $\mathbb{CH}_0(X)$ injects however into $\mathbb{CH}_0(X_{\bar{F}})$, hence $\varrho_*(h^i)^2 = 0$ and therefore $S^{n-i}\big(\rho_*(h^i)\big) = 0$, giving a contradiction with Corollary 4.9. $\square$

Remark 5.2. It turns out that Theorem 5.1 is extremely useful in the theory of quadratic forms. For example, it is the main ingredient of Vishik's proof of the theorem that there is no anisotropic quadratic forms satisfying $2^r < \dim\varphi < 2^r + 2^{r-1}$ and $[\varphi] \in I^r(F)$ (see [14], [15]).

6. Rost motives

Let Λ be an associative commutative ring with 1. We set $\Lambda\mathbb{CH} = \Lambda \otimes_{\mathbb{Z}} \mathbb{CH}$ (we will only need $\Lambda = \mathbb{Z}$ or $\Lambda = \mathbb{Z}/2$).

We briefly recall the construction of the category of Grothendieck $\Lambda\mathbb{CH}$-motives similar to that of [4]. A motive is a triple (X, p, n), where X is a smooth projective equidimensional F-variety, $p \in \Lambda\mathbb{CH}^{\dim X}(X \times X)$ an idempotent correspondence, and n an integer. Sometimes the reduced notations are used: (X, n) for (X, p, n) with p the diagonal class; (X, p) for (X, p, n) with $n = 0$; and (X) for $(X, 0)$, the motive of the variety X.

For a motive $M = (X, p, n)$ and an integer m, the m-th twist $M(m)$ of M is defined as $(X, p, n + m)$.

The set of morphisms is defined as

$$\operatorname{Hom}\big((X, p, n), (X', p', n')\big) = p' \circ \Lambda\mathbb{CH}^{\dim X' - n + n'}(X \times X') \circ p\ .$$

In particular, *every* homogeneous correspondence $\alpha \in \Lambda\mathbb{CH}(X \times X')$ determines a morphism of *every* twist of (X, p) to a certain twist of (X', p').

The Chow group $\Lambda\mathbb{CH}_*(X, p, n)$ of a motive (X, p, n) is defined as

$$\Lambda\mathbb{CH}_*(X, p, n) = p_*\, \Lambda\mathbb{CH}_{*-n}(X)\ .$$

It gives an additive functor of the category of $\Lambda\mathbb{CH}$-motives to the category of graded abelian groups (namely, the functor $\operatorname{Hom}(M(*), -)$, where M is the motive of a point).

For any Λ, there is an evident additive functor of the category of $\mathbb{CH}$-motives to the category of $\Lambda\mathbb{CH}$-motives (identical on the motives of varieties). In particular, every isomorphism of $\mathbb{CH}$-motives automatically produces an isomorphism of the corresponding $\Lambda\mathbb{CH}$-motives. This is why bellow we mostly formulate the results only on the integral motives.

We are coming back to the quadratic forms.

Definition 6.1. Let ϱ be an integral Rost projector on a projective quadric X. We refer to the motive (X, ϱ) as to an (integral) *Rost motive*. (While the CH-motive given by a modulo 2 Rost projector can be called a *modulo* 2 *Rost motive*.) A Rost motive is *anisotropic*, if the quadric X is so.

Let now π be a Pfister form and let φ be a neighbor of π which is *minimal*, that is, has dimension $\dim \pi/2+1$. As noticed by M. Rost (see [7, 5.2] for a proof), the projective quadric X given by φ possesses an integral Rost projector ϱ.

Proposition 6.2. *Let φ be as above. Let ϱ' be the Rost projector on the quadric X' given by a minimal neighbor φ' of another Pfister form π'. The Rost motives (X, ϱ) and (X', ϱ') are isomorphic if and only if the Pfister forms π and π' are isomorphic.*

Proof. First we assume that $(X, \varrho) \simeq (X', \varrho')$. Looking at the degrees of 0-cycles on X and on X', we see that φ is isotropic if and only if φ' is isotropic, thus π is isotropic if and only if π' is isotropic. Therefore, the forms $\pi_{F(\pi')}$ and $\pi'_{F(\pi)}$ are isotropic. Since π and π' are Pfister forms, it follows that $\pi \simeq \pi'$.
Conversely, assume that $\pi \simeq \pi'$. By [12, §3], in order to show that $(X, \varrho) \simeq (X', \varrho')$, it suffices to construct a morphism of motives $(X, \varrho) \to (X', \varrho')$ which becomes mutually inverse isomorphism over an algebraic closure $\bar{F}$ of F. We will do a little bit more: we construct two morphisms $(X, \varrho) \rightleftarrows (X', \varrho')$ which become mutually inverse isomorphisms over an algebraic closure $\bar{F}$ of F (in this case, the initial F-morphisms are isomorphisms, although possibly not mutually inverse ones, [7, cor. 3.3]).
Since $\pi \simeq \pi'$, the quadratic forms $\varphi'_{F(\varphi)}$ and $\varphi_{F(\varphi')}$ are isotropic. Therefore there exist rational morphisms $X \to X'$ and $X' \to X$. The closures of their graphs give two correspondences $\alpha \in \mathbb{CH}(X \times X')$ and $\beta \in \mathbb{CH}(X' \times X)$.
Over $\bar{F}$ we have: $\varrho' \circ \alpha \circ \varrho = [X \times x'] + a[x \times X']$, where $x \in X_{\bar{F}}$ and $x' \in X'_{\bar{F}}$ are closed rational points, while a is an integer (which coincides, in fact, with the degree of the rational morphism $X \to X'$). Similarly, $\varrho \circ \beta \circ \varrho' = [X' \times x] + b[x' \times X]$ with some $b \in \mathbb{Z}$ over $\bar{F}$.
We are going to check that the integers a and b are odd. For this we consider the composition

$$(\varrho \circ \beta \circ \varrho') \circ (\varrho' \circ \alpha \circ \varrho) \in \mathbb{CH}(X \times X) .$$

Over $\bar{F}$ this composition gives $[X \times x] + ab[x \times X]$. Consequently, by [6, th. 6.4] and Lemma 4.3, the integer ab is odd.

Let us take now

$$\alpha' = \alpha - \frac{a-1}{2}\cdot[y\times X'] \quad\text{and}\quad \beta' = \beta - \frac{b-1}{2}\cdot[y'\times X]$$

with some degree 2 closed points $y \in X$ and $y' \in X'$. Then over $\bar{F}$

$$\varrho'\circ\alpha'\circ\varrho = [X\times x'] + [x\times X'] \quad\text{while}\quad \varrho\circ\beta'\circ\varrho' = [X'\times x] + [x'\times X]\ ,$$

therefore the two F-morphisms $(X,\varrho) \rightleftarrows (X',\varrho')$ given by these α' and β' become mutually inverse isomorphisms over $\bar{F}$. □

Definition 6.3. The motive (X,ϱ) for X and ϱ as in Proposition 6.2 (more precisely, the isomorphism class of motives) is called the *Rost motive of the Pfister form* π and denoted $R(\pi)$.

Remark 6.4. It is conjectured in [7, conj. 1.6] (with a proof given for 3 and 7-dimensional quadrics) that every anisotropic Rost motive is the Rost motive of some Pfister form.

7. Motivic decompositions of excellent quadrics

Theorem 7.1 (announced in [11]). *Let φ be a neighbor of a Pfister form π and let φ' be the complementary form (that is, φ' is such that the form $\varphi\perp\varphi'$ is similar to π). Then*

$$(X) \simeq \Big(\bigoplus_{i=0}^{m-1} R(\pi)(i)\Big)\bigoplus (X')(m)\ ,$$

where $m = (\dim\varphi - \dim\varphi')/2$, X is the quadric defined by φ, and X' is the quadric defined by φ'.

Proof. Similar to [12, th. 17] (see also [7, prop. 5.3]). □

We recall that a quadratic form φ over F is called *excellent*, if for every field extension E/F the anisotropic part of the form φ_E is defined over F. An anisotropic quadratic form is excellent if and only if it is a Pfister neighbor whose complementary form is excellent as well [8, §7].

Let $\pi_0 \supset \pi_1 \supset \cdots \supset \pi_r$ be a strictly decreasing sequence of embedded Pfister forms. Let φ be the quadratic form such that the class $[\varphi]$ of φ in the Witt ring of F is the alternating sum $[\pi_0]-[\pi_1]+\cdots+(-1)^r[\pi_r]$, while the dimension of φ is the alternating sum of the dimensions of the Pfister forms. Clearly, φ is excellent. Moreover, every anisotropic excellent quadratic form is similar to a form obtained this way. Let us require additionally that $2\dim\pi_r < \dim\pi_{r-1}$. Then every anisotropic excellent quadratic form is still similar to a form obtained this way and, moreover, the Pfister forms $\pi_0,\dots,\pi_r$ are uniquely determined by the initial excellent quadratic form.

Let X be an *excellent* quadric, that is, the quadratic form φ giving X is excellent. As Theorem 7.1 shows, the motive of X is a direct sum of twisted Rost motives. More precisely, Corollary 7.2 (announced in [11]). *Let X be the excellent quadric determined by Pfister forms $\pi_0 \supset \cdots \supset \pi_r$. Then*

$$(X) \simeq \Big(\bigoplus_{i=0}^{m_0-1} R(\pi_0)(i) \Big) \bigoplus \Big(\bigoplus_{i=m_0}^{m_0+m_1-1} R(\pi_1)(i) \Big) \bigoplus \dots$$
$$\dots \bigoplus \Big(\bigoplus_{i=m_0+\cdots+m_{r-1}}^{m_0+\cdots+m_r} R(\pi_r)(i) \Big)$$

with $m_j = \dim \pi_j/2 - \dim \pi_{j+1} + \dim \pi_{j+2} - \dots$. □

Here are three examples of excellent forms which are most important for us:

Example 7.3 (Pfister forms, [12, prop. 19]). Let $\varphi = \pi$ be a Pfister form. Then

$$(X) \simeq \bigoplus_{i=0}^{\dim \pi/2-1} R(\pi)(i) .$$

Example 7.4 (Maximal neighbors, [12, th. 17]). Let φ be a *maximal* neighbor of a Pfister form π (that is, $\dim \varphi = \dim \pi - 1$). Then

$$(X) \simeq \bigoplus_{i=0}^{\dim \pi/2-2} R(\pi)(i)$$

Example 7.5 (Norm forms, [12, th. 17]). Let φ be a norm quadratic form, that is, φ is a minimal neighbor of a Pfister form π containing a 1-codimensional subform which is similar to a Pfister form π'. Then

$$(X) \simeq R(\pi) \bigoplus \Big(\bigoplus_{i=1}^{\dim \pi'/2-1} R(\pi')(i) \Big) .$$

8. Chow groups of Rost motives

The following theorem computes the Chow groups of the modulo 2 Rost motive of a Pfister form.

Theorem 8.1 (announced in [11]). *Let ρ be the modulo 2 Rost projector on the projective n-dimensional quadric X given by an anisotropic minimal Pfister neighbor. Let i be an integer with $0 \le i \le n$. If $i+1$ is a power of 2, then the Chow group $\mathrm{CH}_i(X, \rho) = \rho_* \mathrm{CH}_i(X)$ is cyclic of order 2 generated by $\rho_*(h^{n-i})$. Otherwise this group is 0.*

Proof. According to Proposition 6.2, we may assume that X is a *norm quadric*, that is, X contains a 1-codimensional subquadric Y being a Pfister quadric. Let r be the integer such that $n = \dim X = 2^r - 1$.

We proceed by induction on r. Let $Y' \subset Y$ be a subquadric of dimension $2^{r-1} - 2$ which is a Pfister quadric. Let X' be a norm quadric of dimension $2^{r-1} - 1$ such that $Y' \subset X' \subset Y$. Let ρ' be a modulo 2 Rost projector on X'. By Example 7.5, passing from $\mathbb{CH}$-motives to the category of CH-motives, we see that the motive of X is the direct sum of the motive (X, ρ) and the motives

(X', ρ', i) with $i = 1, \dots, 2^{r-1} - 1$. Therefore

$$\mathrm{CH}(X) \simeq \rho_* \mathrm{CH}(X) \bigoplus \Big(\bigoplus_{i=1}^{2^{r-1}-1} \rho'_* \mathrm{CH}(X') \Big)$$

(we do not care about the gradations on the Chow groups).
Also the motive of Y decomposes in the direct sum of the motives (X', ρ', i) with $i = 0, \dots, 2^{r-1} - 1$ (Example 7.3). Therefore

$$\mathrm{CH}(Y) \simeq \bigoplus_{i=0}^{2^{r-1}-1} \rho'_* \mathrm{CH}(X') \,.$$

It follows that the order of the group $\mathrm{CH}(Y)$ is $|\rho'_* \mathrm{CH}(X')|^{2^{r-1}}$, while the order of $\mathrm{CH}(X)$ is $|\rho'_* \mathrm{CH}(X')|^{2^{r-1}-1} \cdot |\rho_* \mathrm{CH}(X)|$.
In the exact sequence

$$\mathrm{CH}(Y) \to \mathrm{CH}(X) \to \mathrm{CH}(U) \to 0$$

with $U = X \setminus Y$, the Chow group $\mathrm{CH}(U)$ of the affine norm quadric U is computed by M. Rost ([7, th. A.4]): $\mathrm{CH}(U) = \mathrm{CH}^0(U) \simeq \mathbb{Z}/2$. Therefore, the orders of these groups satisfy

$$|\,\mathrm{CH}(X)| \leq |\,\mathrm{CH}(Y)| \cdot |\,\mathrm{CH}(U)| = 2|\,\mathrm{CH}(Y)| \,,$$

thus $|\rho_* \mathrm{CH}(X)| \leq 2|\rho'_* \mathrm{CH}(X')|$.
The group $\rho'_* \mathrm{CH}(X')$ is known by induction. In particular, the order of this group is 2^r. It follows that the order of $\rho_* \mathrm{CH}(X)$ is at most 2^{r+1}. Corollary 4.9 gives already $r+1$ non-zero elements of $\rho_* \mathrm{CH}_*(X)$ living in different dimensions (more precisely, $\rho_*(h^{n-2^s+1}) \neq 0$ for $s = 1, \dots, r-1$ by Corollary 4.9 and for $s = 0, r$ by Lemma 2.4) and therefore generating a subgroup of order 2^{r+1}. It follows that the order of $\rho_* \mathrm{CH}(X)$ is precisely 2^{r+1} and the non-zero elements we have found generate the group $\rho_* \mathrm{CH}(X)$. □

The integral version of Theorem 8.1 is given by

Corollary 8.2 (announced in [11]). *For X as in Theorem 8.1, let ϱ be the integral Rost projector on X. Then for every i with $0 \leq i \leq n$, the Chow group $\mathbb{CH}_i(X, \varrho) = \varrho_* \mathbb{CH}_i(X)$ is a cyclic group generated by $\varrho_*(h^{n-i})$. Moreover, the element $\varrho_*(h^{n-i})$ is*

- *0, if $i+1$ is not a power of 2;*
- *of order 2, if $i+1$ is a power of 2 and $i \notin \{0, n\}$;*
- *of infinite order, if $i \in \{0, n\}$.*

Proof. The statements on $\mathbb{CH}_n(X)$ and on $\mathbb{CH}_0(X)$ are clear. The rest follows from Theorem 8.1, if we show that $2 \cdot \varrho_* \mathbb{CH}_i(X) = 0$ for every i with $0 < i < n$. Let L/F be a quadratic extension such that X_L is isotropic. Then $(\varrho_L)_* \mathbb{CH}_i(X_L) = 0$ for such i by [7, cor. 4.2] (cf. Lemma 2.4). Since the composition of the restriction $\mathbb{CH}_i(X) \to \mathbb{CH}_i(X_L)$ with the transfer $\mathbb{CH}_i(X_L) \to \mathbb{CH}_i(X)$ coincides with the multiplication by 2, it follows that $2 \cdot \varrho_* \mathbb{CH}_i(X) = 0$. □

Remark 8.3. The result of Corollary 8.2 was announced in [11]. A proof has never appeared.

Remark 8.4. Clearly, Corollary 8.2 describes the Chow group of the Rost motive of an anisotropic Pfister form. Since the motive of any anisotropic excellent quadric is a direct sum of twists of such Rost motives (Corollary 7.2), we have computed the Chow group of an arbitrary anisotropic excellent projective quadric. Note that the answer depends only on the dimension of the quadric.

References

[1] P. Brosnan, *Steenrod operations in Chow theory*, K-theory Preprint Archives 0370 (1999), 1–19 (see `www.math.uiuc.edu/K-theory`). To appear in Trans. Amer. Math. Soc.

[2] W. Fulton, *Intersection Theory*, Springer-Verlag, Berlin, 1984.

[3] O. Izhboldin and A. Vishik, *Quadratic forms with absolutely maximal splitting*, Contemp. Math. 272 (2000), 103–125.

[4] U. Jannsen, *Motives, numerical equivalence, and semi-simplicity*, Invent. Math. 107 (1992), 447–452.

[5] N. A. Karpenko, *Algebro-geometric invariants of quadratic forms*, Algebra i Analiz 2 (1990), no. 1, 141–162 (in Russian). Engl. transl.: Leningrad (St. Petersburg) Math. J. 2 (1991), no. 1, 119–138.

[6] N. A. Karpenko, *On anisotropy of orthogonal involutions*, J. Ramanujan Math. Soc. 15 (2000), no. 1, 1–22.

[7] N. A. Karpenko, *Characterization of minimal Pfister neighbors via Rost projectors*, J. Pure Appl. Algebra 160 (2001), 195–227.

[8] M. Knebusch, *Generic splitting of quadratic forms, II*, Proc. London Math. Soc. 34 (1977), 1–31.

[9] A. S. Merkurjev, I. A. Panin, and A. R. Wadsworth, *Index reduction formulas for twisted flag varieties. II*, K-Theory 14 (1998), no. 2, 101–196.

[10] I. Panin, *Riemann-Roch theorem for oriented cohomology*, K-theory Preprint Archives 0552 (2002), (see `www.math.uiuc.edu/K-theory/0552`).

[11] M. Rost, *Some new results on the Chowgroups of quadrics*, Preprint (1990) (see `www.math.ohio-state.edu/~rost`).

[12] M. Rost, *The motive of a Pfister form*, Preprint (1998) (see `www.math.ohio-state.edu/~rost`).

[13] R. G. Swan, *Zero cycles on quadric hypersurfaces*, Proc. Amer. Math. Soc. 107 (1989), no. 1, 43–46.

[14] A. Vishik, *On the dimension of anisotropic forms in* I^n, Max-Planck-Institut für Mathematik in Bonn, preprint MPI 2000-11 (2000), 1–41 (see `www.mpim-bonn.mpg.de`).

[15] A. Vishik, *On the dimensions of quadratic forms*, Rossiĭskaya Akademiya Nauk. Doklady Akademii Nauk (Russian) 373 (2000), no. 4, 445–447.

Nikita Karpenko
Laboratoire des Mathématiques
Faculté des Sciences
Université d'Artois
rue Jean Souvraz SP 18
62307 Lens Cedex
France
karpenko@euler.univ-artois.fr

Alexander Merkurjev
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
USA
merkurev@math.ucla.edu

Toric Hyperkähler Varieties

Tamás Hausel and Bernd Sturmfels

Received: August 15, 2002
Revised: December 8, 2002

Communicated by Günter M. Ziegler

Abstract. Extending work of Bielawski-Dancer [3] and Konno [14], we develop a theory of toric hyperkähler varieties, which involves toric geometry, matroid theory and convex polyhedra. The framework is a detailed study of semi-projective toric varieties, meaning GIT quotients of affine spaces by torus actions, and specifically, of Lawrence toric varieties, meaning GIT quotients of even-dimensional affine spaces by symplectic torus actions. A toric hyperkähler variety is a complete intersection in a Lawrence toric variety. Both varieties are non-compact, and they share the same cohomology ring, namely, the Stanley-Reisner ring of a matroid modulo a linear system of parameters. Familiar applications of toric geometry to combinatorics, including the Hard Lefschetz Theorem and the volume polynomials of Khovanskii-Pukhlikov [11], are extended to the hyperkähler setting. When the matroid is graphic, our construction gives the toric quiver varieties, in the sense of Nakajima [17].

1 Introduction

Hyperkähler geometry has emerged as an important new direction in differential and algebraic geometry, with numerous applications to mathematical physics and representation theory. Roughly speaking, a *hyperkähler manifold* is a Riemannian manifold of dimension $4n$, whose holonomy is in the unitary symplectic group $Sp(n) \subset SO(4n)$. The key example is the quaternionic space $\mathbb{H}^n \simeq \mathbb{C}^{2n} \simeq \mathbb{R}^{4n}$. Our aim is to relate hyperkähler geometry to the combinatorics of convex polyhedra. We believe that this connection is fruitful for both subjects. Our objects of study are the *toric hyperkähler manifolds* of Bielawski and Dancer [3]. They are obtained from $\mathbb{H}^n$ by taking the hyperkähler quotient [10] by an abelian subgroup of $Sp(n)$. Bialewski and Dancer found that the geometry and topology of toric hyperkähler manifolds is governed by hyperplane arrangements, and Konno [14] gave an explicit presentation of their cohomology rings. The present paper is self-contained and contains new proofs for the relevant results of [3] and [14].

We start out in Section 2 with a discussion of semi-projective toric varieties. A toric variety X is called *semi-projective* if X has a torus-fixed point and X is projective over its affinization $\mathrm{Spec}(H^0(X,\mathcal{O}_X))$. We show that semi-projective toric varieties are exactly the ones which arise as GIT quotients of a complex vector space by an abelian group. Then we calculate the cohomology ring of a semi-projective toric orbifold X. It coincides with the cohomology of the *core* of X, which is defined as the union of all compact torus orbit closures. This result and further properties of the core are derived in Section 3.

The lead characters in the present paper are the *Lawrence toric varieties*, to be introduced in Section 4 as the GIT quotients of symplectic torus actions on even-dimensional affine spaces. They can be regarded as the "most non-compact" among all semi-projective toric varieties. The combinatorics of Lawrence toric varieties is governed by the Lawrence construction of convex polytopes [22, §6.6] and its intriguing interplay with matroids and hyperplane arrangements.

In Section 6 we define *toric hyperkähler varieties* as subvarieties of Lawrence toric varieties cut out by certain natural bilinear equations. In the smooth case, they are shown to be biholomorphic with the toric hyperkähler manifolds of Bielawski and Dancer, whose differential-geometric construction is reviewed in Section 5 for the reader's convenience. Under this identification the core of the toric hyperkähler variety coincides with the core of the ambient Lawrence toric variety. We shall prove that these spaces have the same cohomology ring which has the following description. All terms and symbols appearing in Theorem 1.1 are defined in Sections 4 and 6.

Theorem 1.1 *Let $A : \mathbb{Z}^n \to \mathbb{Z}^d$ be an epimorphism, defining an inclusion $\mathbb{T}^d_{\mathbb{R}} \subset \mathbb{T}^n_{\mathbb{R}}$ of compact tori, and let $\theta \in \mathbb{Z}^d$ be generic. Then the following graded $\mathbb{Q}$-algebras are isomorphic:*

1. *the cohomology ring of the toric hyperkähler variety $Y(A,\theta) = \mathbb{H}^n////_{(\theta,0)}\mathbb{T}^d_{\mathbb{R}}$,*
2. *the cohomology ring of the Lawrence toric variety $X(A^{\pm},\theta) = \mathbb{C}^{2n}//_{\theta}\,\mathbb{T}^d_{\mathbb{R}}$,*
3. *the cohomology ring of the core $C(A^{\pm},\theta)$, which is the preimage of the origin under the affinization map of either the Lawrence toric variety or the toric hyperkähler variety,*
4. *the quotient ring $\mathbb{Q}[x_1,\ldots,x_n]/(M^*(\mathcal{A})+\mathrm{Circ}(\mathcal{A}))$, where $M^*(\mathcal{A})$ is the matroid ideal which is generated by squarefree monomials representing cocircuits of A, and $\mathrm{Circ}(\mathcal{A})$ is the ideal generated by the linear forms that correspond to elements in the kernel of A.*

If the matrix A is unimodular then $X(A^{\pm},\theta)$ and $Y(A,\theta)$ are smooth and $\mathbb{Q}$ can be replaced by $\mathbb{Z}$.

Here is a simple example where all three spaces are manifolds: take $A:\mathbb{Z}^3\to\mathbb{Z}$, $(u_1,u_2,u_3)\mapsto u_1+u_2+u_3$ with $\theta\neq 0$. Then $C(A^{\pm},\theta)$ is the complex projective plane $\mathbb{P}^2$. The Lawrence toric variety $X(A^{\pm},\theta)$ is the quotient of $\mathbb{C}^6=\mathbb{C}^3\oplus\mathbb{C}^3$ modulo the symplectic torus action $(x,y)\mapsto(t\cdot x,t^{-1}\cdot y)$. Geometrically, X is a rank 3 bundle over $\mathbb{P}^2$, visualized as an unbounded 5-dimensional polyhedron with a bounded 2-face, which is a triangle. The toric hyperkähler variety $Y(A,\theta)$ is embedded into $X(A^{\pm},\theta)$ as the hypersurface $x_1y_1+x_2y_2+x_3y_3=0$. It is isomorphic to the cotangent bundle of $\mathbb{P}^2$. Note that $Y(A,\theta)$ itself is not a toric variety.

For general matrices A, the varieties $X(A^{\pm},\theta)$ and $Y(A,\theta)$ are orbifolds, by the genericity hypothesis on θ, and they are always non-compact. The core $C(A^{\pm},\theta)$ is projective but almost always reducible. Each of its irreducible components is a projective toric orbifold.

In Section 7 we give a dual presentation, in terms of *cogenerators*, for the cohomology ring. These cogenerators are the volume polynomials of Khovanskii-Pukhlikov [11] of the bounded faces of our unbounded polyhedra. As an application we prove the injectivity part of the Hard Lefschetz Theorem for toric hyperkähler varieties (Theorem 7.4). In light of the following corollary to Theorem 1.1, this provides new inequalities for the h-numbers of rationally representable matroids.

Corollary 1.2 *The Betti numbers of the toric hyperkähler variety $Y(A,\theta)$ are the h-numbers (defined in Stanley's book [18, §III.3]) of the rank $n-d$ matroid given by the integer matrix A.*

The *quiver varieties* of Nakajima [17] are hyperkähler quotients of $\mathbb{H}^n$ by some subgroup $G\subset Sp(n)$ which is a product of unitary groups indexed by a quiver (i.e. a directed graph). In Section 8 we examine *toric quiver varieties* which arise when G is a compact torus. They are the toric hyperkähler manifolds obtained when A is the differential $\mathbb{Z}^{\mathrm{edges}}\to\mathbb{Z}^{\mathrm{vertices}}$ of a quiver. Note that our notion of toric quiver variety is not the same as that of Altmann and Hille [1]. Theirs are toric and projective: in fact, they are the irreducible components of our core $C(A^{\pm},\theta)$.

We close the paper by studying two examples in detail. First in Section 9 we illustrate the main results of this paper for a particular example of a toric quiver variety, corresponding to the complete bipartite graph $K_{2,3}$. In the final Section 10 we examine the ALE spaces of type A_n. Curiously, these manifolds are both toric and hyperkähler, and we show that they and their products are the only toric hyperkähler manifolds which are toric varieties in the usual sense.

Acknowledgment. This paper grew out of a lecture on toric aspects of Nakajima's quiver varieties [17] given by the second author in the Fall 2000

Quiver Varieties seminar at UC Berkeley, organized by the first author. We are grateful to the participants of this seminar for their contributions. In particular we thank Mark Haiman, Allen Knutson and Valerio Toledano. We thank Roger Bielawski for drawing our attention to Konno's work [14], and we thank Manoj Chari for explaining the importance of [15] for Betti numbers of toric quiver varieties. Both authors were supported by the Miller Institute for Basic Research in Science, in the form of a Miller Research Fellowship (1999-2002) for the first author and a Miller Professorship (2000-2001) for the second author. The second author was also supported by the National Science Foundation (DMS-9970254).

2 Semi-projective toric varieties

Projective toric varieties are associated with rational polytopes, that is, bounded convex polyhedra with rational vertices. This section describes toric varieties associated with (typically unbounded) rational polyhedra. The resulting class of semi-projective toric varieties will be seen to equal the GIT-quotients of affine space $\mathbb{C}^n$ modulo a subtorus of $\mathbb{T}^n_{\mathbb{C}}$.

Let $A = [a_1, \dots, a_n]$ be a $d \times n$-integer matrix whose $d \times d$-minors are relatively prime. We choose an $n \times (n-d)$-matrix $B = [b_1, \dots, b_n]^T$ which makes the following sequence exact:

$$0 \longrightarrow \mathbb{Z}^{n-d} \xrightarrow{B} \mathbb{Z}^n \xrightarrow{A} \mathbb{Z}^d \longrightarrow 0. \tag{1}$$

The choice of B is equivalent to choosing a basis in $\ker(A)$. The configuration $\mathcal{B} := \{b_1, \dots, b_n\}$ in $\mathbb{Z}^{n-d}$ is said to be a *Gale dual* of the given vector configuration $\mathcal{A} := \{a_1, \dots, a_n\}$ in $\mathbb{Z}^d$.

We denote by $\mathbb{T}_{\mathbb{C}}$ the complex group $\mathbb{C}^*$ and by $\mathbb{T}_{\mathbb{R}}$ the circle $U(1)$. Their Lie algebras are denoted by $\mathfrak{t}_{\mathbb{C}}$ and $\mathfrak{t}_{\mathbb{R}}$ respectively. We apply the contravariant functor $\mathrm{Hom}(\,\cdot\,, \mathbb{T}_{\mathbb{C}})$ to the short exact sequence (1). This gives a short exact sequence of abelian groups:

$$1 \longleftarrow \mathbb{T}^{n-d}_{\mathbb{C}} \xleftarrow{B^T} \mathbb{T}^n_{\mathbb{C}} \xleftarrow{A^T} \mathbb{T}^d_{\mathbb{C}} \longleftarrow 1. \tag{2}$$

Thus $\mathbb{T}^d_{\mathbb{C}}$ is embedded as a d-dimensional subtorus of $\mathbb{T}^n_{\mathbb{C}}$. It acts on the affine space $\mathbb{C}^n$. We shall construct the quotients of this action in the sense of *geometric invariant theory* (= GIT). The ring of polynomial functions on $\mathbb{C}^n$ is graded by the semigroup $\mathbb{N}\mathcal{A} \subseteq \mathbb{Z}^d$:

$$S = \mathbb{C}[x_1, \dots, x_n], \quad \deg(x_i) = a_i \in \mathbb{N}\mathcal{A}. \tag{3}$$

A polynomial in S is homogeneous if and only if it is a $\mathbb{T}^d_{\mathbb{C}}$-eigenvector. For $\theta \in \mathbb{N}\mathcal{A}$, let S_θ denote the (typically infinite-dimensional) $\mathbb{C}$-vector space of homogeneous polynomials of degree θ. Note that S_θ is a module over the subalgebra S_0 of degree zero polynomials in $S = \bigoplus_{\theta \in \mathbb{N}\mathcal{A}} S_\theta$. The following lemma is a standard fact in combinatorial commutative algebra.

Lemma 2.1 *The $\mathbb{C}$-algebra S_0 is generated by a finite set of monomials, corresponding to the minimal generators of the semigroup $\mathbb{N}^n \cap \mathrm{im}(B)$. For any $\theta \in \mathbb{N}\mathcal{A}$, the graded component S_θ is a finitely generated S_0-module, and the ring $S_{(\theta)} = \bigoplus_{r=0}^{\infty} S_{r\theta}$ is a finitely generated S_0-algebra.*

The $\mathbb{C}$-algebra S_0 coincides with the ring of invariants $S^{\mathbb{T}^d_{\mathbb{C}}}$. The S_0-algebra $S_{(\theta)}$ is isomorphic to $\bigoplus_{r=0}^{\infty} t^r S_{r\theta}$. We regard it as $\mathbb{N}$-graded by the degree of t.

Definition 2.2 *The* affine GIT quotient *of $\mathbb{C}^n$ by the d-torus $\mathbb{T}^d_{\mathbb{C}}$ is the affine toric variety*

$$X(A,0) \;:=\; \mathbb{C}^n//_0\, \mathbb{T}^d_{\mathbb{C}} \;:=\; \mathrm{Spec}\,(S^{\mathbb{T}^d_{\mathbb{C}}}) = \mathrm{Spec}\,(S_0) = \mathrm{Spec}\,\big(\mathbb{C}[\mathbb{N}^n \cap \mathrm{im}(B)]\big). \quad (4)$$

For any $\theta \in \mathbb{N}\mathcal{A}$, the projective GIT quotient *of $\mathbb{C}^n$ by the d-torus $\mathbb{T}^d_{\mathbb{C}}$ is the toric variety*

$$X(A,\theta) \quad := \quad \mathbb{C}^n//_\theta\, \mathbb{T}^d_{\mathbb{C}} \;:=\; \mathrm{Proj}\,(S_{(\theta)}) \quad = \quad \mathrm{Proj} \bigoplus_{r=0}^{\infty} t^r \cdot S_{r\theta}. \quad (5)$$

Recall that the isomorphism class of any toric variety is given by a fan in a lattice. A toric variety is a *toric orbifold* if its fan is simplicial. We shall describe the fans of the toric varieties $X(A,0)$ and $X(A,\theta)$ using the notation in Fulton's book [8]. We write M for the lattice $\mathbb{Z}^{n-d}$ in (1) and $N = \mathrm{Hom}(M,\mathbb{Z})$ for its dual. The torus $\mathbb{T}^{n-d}_{\mathbb{C}}$ in (2) is identified with $N \otimes \mathbb{T}_{\mathbb{C}}$. The column vectors $\mathcal{B} = \{b_1,\ldots,b_n\}$ of the matrix B^T form a configuration in $N \simeq \mathbb{Z}^{n-d}$. We write $\mathrm{pos}(\mathcal{B})$ for the convex polyhedral cone spanned by $\mathcal{B}$ in the vector space $N_{\mathbb{R}} = N \otimes \mathbb{R} \simeq \mathbb{R}^{n-d}$. Note that the affine toric variety associated with the cone $\mathrm{pos}(\mathcal{B})$ equals $X(A,0)$.

A *triangulation* of the configuration $\mathcal{B}$ is a simplicial fan Σ whose rays lie in $\mathcal{B}$ and whose support equals $\mathrm{pos}(\mathcal{B})$. A *T-Cartier divisor* on Σ is a continuous function $\Psi : \mathrm{pos}(\mathcal{B}) \to \mathbb{R}$ which is linear on each cone of Σ and takes integer values on $N \cap \mathrm{pos}(\mathcal{B})$. The triangulation Σ is called *regular* if there exists a T-Cartier divisor Ψ which is *ample*, i.e. the function $\Psi : \mathrm{pos}(\mathcal{B}) \to \mathbb{R}$ is convex and restricts to a different linear function on each maximal cone of Σ. Two T-Cartier divisors Ψ_1 and Ψ_2 are *equivalent* if $\Psi_1 - \Psi_2$ is a linear map on $\mathrm{pos}(\mathcal{B})$, i.e. it is an element of M. A *divisor* on Σ is an equivalence class of T-Cartier divisors on Σ. Since Ψ_1 is ample if and only if Ψ_2 is ample, ampleness is well-defined for divisors $[\Psi]$. Finally, we define a *polarized triangulation* of $\mathcal{B}$ to be a pair consisting of a triangulation Σ of $\mathcal{B}$ and an ample divisor $[\Psi]$.

The cokernel of $M \xrightarrow{B} \mathbb{Z}^n$ is identified with $\mathbb{Z}^d$ in (1) and we call it the *Picard group*. Hence $\mathcal{A} = \{a_1,\ldots,a_n\}$ is a vector configuration in the Picard group. The *chamber complex* $\Gamma(\mathcal{A})$ of $\mathcal{A}$ is defined to be the coarsest fan with support $\mathrm{pos}(\mathcal{A})$ that refines all triangulations of $\mathcal{A}$. Experts in toric geometry will note that $\Gamma(\mathcal{A})$ equals the *secondary fan* of $\mathcal{B}$ as in [7]. We say that $\theta \in \mathbb{N}\mathcal{A}$ is *generic* if it lies in an open chamber of $\Gamma(\mathcal{A})$. Thus $\theta \in \mathbb{N}\mathcal{A}$ is generic if it is not in any lower-dimensional cone $\operatorname{pos}\{a_{i_1}, \ldots, a_{i_{d-1}}\}$ spanned by columns of A. The chamber complex $\Gamma(\mathcal{A})$ parameterizes the different combinatorial types of the convex polyhedra

$$P_\theta \quad = \quad \bigl\{ u \in \mathbb{R}^n \,:\, Au = \theta,\, u \geq 0 \bigr\}$$

as θ ranges over $\mathbb{N}\mathcal{A}$. In particular, θ is generic if and only if P_θ is $(n-d)$-dimensional and each of its vertices has exactly d non-zero coordinates (i.e. P_θ is simple). A vector θ in $\mathbb{N}\mathcal{A}$ is called an *integral degree* if every vertex of the polyhedron P_θ is a lattice point in $\mathbb{Z}^n$.

Proposition 2.3 *There is a one-to-one correspondence between generic integral degrees θ in $\mathbb{N}\mathcal{A}$ and polarized triangulations $\bigl(\Sigma, [\Psi]\bigr)$ of $\mathcal{B}$. When forgetting the polarization this correspondence gives a bijection between open chambers of $\Gamma(\mathcal{A})$ and regular triangulations Σ of $\mathcal{B}$.*

Proof: Given a generic integral degree θ, we construct the corresponding polarized triangulation $\bigl(\Sigma, [\Psi]\bigr)$. First choose any $\psi \in \mathbb{Z}^n$ such that $A\psi = -\theta$. Then consider the polyhedron

$$Q_\psi \quad := \quad \bigl\{ v \in M_{\mathbb{R}} \,:\, Bv \,\geq\, \psi \bigr\}.$$

The map $v \mapsto Bv - \psi$ is an affine-linear isomorphism from Q_ψ onto P_θ which identifies the set of lattice points $Q_\psi \cap M$ with the set of lattice points $P_\theta \cap \mathbb{Z}^n$. The set of linear functionals which are bounded below on Q_ψ is precisely the cone $\operatorname{pos}(\mathcal{B}) \subset N$. Finally, define the function

$$\Psi \,:\, \operatorname{pos}(\mathcal{B}) \to \mathbb{R}\,, \quad w \mapsto \min\{\, w \cdot v \,:\, v \in Q_\psi \,\}.$$

This is the *support function* of Q_ψ, which is piecewise-linear, convex and continuous. It takes integer values on $N \cap \operatorname{pos}(\mathcal{B})$ because each vertex of Q_ψ lies in M. Since Q_ψ is a simple polyhedron, its *normal fan* is a regular triangulation Σ_θ of $\mathcal{B}$, and Ψ restricts to a different linear function on each maximal face of Σ_θ. Hence $\bigl(\Sigma_\theta, [\Psi]\bigr)$ is a polarized triangulation of $\mathcal{B}$.
Conversely, if we are given a polarized triangulation $\bigl(\Sigma, [\Psi]\bigr)$ of $\mathcal{B}$, then we define $\psi := (\Psi(b_1), \ldots, \Psi(b_n)) \in \mathbb{Z}^n$, and $\theta = -A\psi$ is the corresponding generic integral degree in $\mathbb{N}\mathcal{A}$. □

Theorem 2.4 *Let $\theta \in \mathbb{N}\mathcal{A}$ be a generic integral degree. Then $X(A, \theta)$ is an orbifold and equals the toric variety $X(\Sigma_\theta)$, where Σ_θ is the regular triangulation of $\mathcal{B}$ given by θ as in Proposition 2.3.*

Proof: First note that the multigraded polynomial ring S is the *homogeneous coordinate ring* in the sense of Cox [6] of the toric variety $X(\Sigma_\theta)$. Specifically, our sequence (1) is precisely the second row in (1) on page 19 of [6]. The irrelevant ideal B_{Σ_θ} of $X(\Sigma_\theta)$ equals the radical of the ideal generated by

$\bigoplus_{r=1}^{\infty} S_{r\theta}$. Since Σ_θ is a simplicial fan, by [6, Theorem 2.1], $X(\Sigma_\theta)$ is the geometric quotient of $\mathbb{C}^n \backslash \mathcal{V}(B_{\Sigma_\theta})$ modulo $\mathbb{T}_{\mathbb{C}}^d$. The variety $\mathcal{V}(B_{\Sigma_\theta})$ consists of the points in $\mathbb{C}^n$ which are not semi-stable with respect to the $\mathbb{T}_{\mathbb{C}}^d$-action. By standard results in Geometric Invariant Theory, the geometric quotient of the semi-stable locus in $\mathbb{C}^n$ modulo $\mathbb{T}_{\mathbb{C}}^d$ coincides with $X(A,\theta) = \mathrm{Proj}\,(S_{(\theta)}) = \mathbb{C}^n /\!/_\theta\, \mathbb{T}_{\mathbb{C}}^d$. Therefore $X(A,\theta)$ is isomorphic to $X(\Sigma_\theta)$. □

Corollary 2.5 *The distinct GIT quotients $X(A,\theta) = \mathbb{C}^n /\!/_\theta\, \mathbb{T}_{\mathbb{C}}^d$ which are toric orbifolds are in bijection with the open chambers in $\Gamma(\mathcal{A})$, and hence with the regular triangulations of $\mathcal{B}$.*

Recall that for every scheme X there is a canonical morphism

$$\pi_X : X \mapsto X_0 \tag{6}$$

to the affine scheme $X_0 = \mathrm{Spec}(H^0(X, \mathcal{O}_X))$ of regular functions on X. We call a toric variety X *semi-projective* if X has at least one torus-fixed point and the morphism π_X is projective.

Theorem 2.6 *The following three classes of toric varieties coincide:*

1. *semi-projective toric orbifolds,*
2. *the GIT-quotients $X(A,\theta)$ constructed in (5) where $\theta \in \mathbb{N}\mathcal{A}$ is a generic integral degree,*
3. *toric varieties $X(\Sigma)$ where Σ is a regular triangulation of a set $\mathcal{B}$ which spans the lattice N.*

Proof: The equivalence of the classes 2 and 3 follows from Theorem 2.4. Let $X(\Sigma)$ be a toric variety in class 3. Since $\mathcal{B}$ spans the lattice, the fan Σ has a full-dimensional cone, and hence $X(\Sigma)$ has a torus-fixed point. Since Σ is simplicial, $X(\Sigma)$ is an orbifold. The morphism π_X can be described as follows. The ring of global sections $H^0(X(\Sigma), \mathcal{O}_{X(\Sigma)})$ is the semigroup algebra of the semigroup in M consisting of all linear functionals on N which are non-negative on the support $|\Sigma|$ of Σ. Its spectrum is the affine toric variety whose cone is $|\Sigma|$. The triangulation Σ supports an ample T-Cartier divisor Ψ. The morphism π_X is projective since it is induced by Ψ. Hence $X(\Sigma)$ is in class 1. Finally, let X be any semi-projective toric orbifold. It is represented by a fan Σ in a lattice N. The fan Σ is simplicial since X is an orbifold, and $|\Sigma|$ spans $N_{\mathbb{R}}$ since X has at least one fixed point. Since the morphism π_X is projective, the fan Σ is a regular triangulation of a subset $\mathcal{B}'$ of $|\Sigma|$ which includes the rays of Σ. The set $\mathcal{B}'$ need not span the lattice N. We choose any superset $\mathcal{B}$ of $\mathcal{B}'$ which is contained in $\mathrm{pos}(\mathcal{B}') = |\Sigma|$ and which spans the lattice N. Then Σ can also be regarded as a regular triangulation of $\mathcal{B}$, and we conclude that X is in class 3. □

Remark. 1. The passage from $\mathcal{B}'$ to $\mathcal{B}$ in the last step means that any GIT quotient of $\mathbb{C}^{n'}$ modulo any abelian subgroup of $\mathbb{T}_{\mathbb{C}}^{n'}$ can be rewritten as a GIT quotient of some bigger affine space $\mathbb{C}^n$ modulo a subtorus of $\mathbb{T}_{\mathbb{C}}^n$. This construction applies in particular when the given abelian group is finite, in which case the initial subset $\mathcal{B}'$ of N is linearly independent.
2. Our proof can be extended to show the following: if X is any toric variety where the morphism π_X is projective then X is the product of a semi-projective toric variety and a torus.
3. The affinization map (6) for $X(A,\theta)$ is the canonical map to $X(A,0)$.

A triangulation Σ of a subset $\mathcal{B}$ of $N \simeq \mathbb{Z}^{n-d}$ is called *unimodular* if every maximal cone of Σ is spanned by a basis of N. This property holds if and only if $X(\Sigma)$ is a toric manifold (= smooth toric variety). We say that a vector θ in $\mathbb{N}\mathcal{A}$ is a *smooth degree* if $C^{-1} \cdot \theta \geq 0$ implies $\det(C) = \pm 1$ for every non-singular $d \times d$-submatrix C of A. Equivalently, the edges at any vertex of the polyhedron P_θ generate $\ker_{\mathbb{Z}} A \cong \mathbb{Z}^{n-d}$. From Theorem 2.6 we conclude:

Corollary 2.7 *The following three classes of smooth toric varieties coincide:*

1. *semi-projective toric manifolds,*
2. *the GIT-quotients $X(A,\theta)$ constructed in* (5) *where $\theta \in \mathbb{N}\mathcal{A}$ is a generic smooth degree,*
3. *toric varieties $X(\Sigma)$ where Σ is a regular unimodular triangulation of a spanning set $\mathcal{B} \subset N$.*

Definition 2.8 *The matrix A is called* unimodular *if the following equivalent conditions hold:*

- *all non-zero $d \times d$-minors of A have the same absolute value,*
- *all $(n-d) \times (n-d)$-minors of the matrix B in (1) are -1, 0 or $+1$,*
- *every triangulation of $\mathcal{B}$ is unimodular,*
- *every vector θ in $\mathbb{N}\mathcal{A}$ is an integral degree,*
- *every vector θ in $\mathbb{N}\mathcal{A}$ is a smooth degree.*

Corollary 2.9 *For A unimodular, every GIT quotient $X(A,\theta)$ is a semi-projective toric manifold, and the distinct smooth quotients $X(A,\theta)$ are in bijection with the open chambers in $\Gamma(\mathcal{A})$.*

Every affine toric variety has a natural moment map onto a polyhedral cone, and every projective toric variety has a moment map onto a polytope. These are described in Section 4.2 of [8]. It is straightforward to extend this description to semi-projective toric varieties. Suppose that the S_0-algebra $S_{(\theta)}$ in Lemma

2.1 is generated by a set of $m+1$ monomials in S_θ, possibly after replacing θ by a multiple in the non-unimodular case. Let $\mathbb{P}^m_{\mathbb{C}}$ be the projective space whose coordinates are these monomials. Then, by definition of "Proj", the toric variety $X(A,\theta)$ is embedded as a closed subscheme in the product $\mathbb{P}^m_{\mathbb{C}} \times \mathrm{Spec}(S_0)$. We have an action of the $(n-d)$-torus $\mathbb{T}^n_{\mathbb{C}}/\mathbb{T}^d_{\mathbb{C}}$ on $\mathbb{P}^m_{\mathbb{C}}$, since S_θ is an eigenspace of $\mathbb{T}^d_{\mathbb{C}}$. This gives rise to a moment map $\mu_1 : \mathbb{P}^m_{\mathbb{C}} \to \mathbb{R}^{n-d}$, whose image is a convex polytope. Likewise, we have the affine moment map $\mu_2 : \mathrm{Spec}(S_0) \to \mathbb{R}^{n-d}$ whose image is the cone polar to $\mathrm{pos}(\mathcal{B})$. This defines the moment map

$$\mu \,:\, X(A,\theta) \,\subset\, \mathbb{P}^m_{\mathbb{C}} \times \mathrm{Spec}(S_0) \,\to\, \mathbb{R}^{n-d}, \;\; (u,v) \mapsto \mu_1(u) + \mu_2(v). \tag{7}$$

The image of $X(A,\theta)$ under the moment map μ is the polyhedron $P_\theta \simeq Q_\psi$, since the convex hull of its vertices equals the image of μ_1 and the cone $P_0 \simeq Q_0$ equals the image of μ_2.

Given an arbitrary fan Σ in N, Section 2.3 in [8] describes how a one-parameter subgroup λ_v, given by $v \in N$, acts on the toric variety $X(\Sigma)$. Consider any point x in $X(\Sigma)$ and let $\gamma \in \Sigma$ be the unique cone such that x lies in the orbit O_γ. The orbit O_γ is fixed pointwise by the one-parameter subgroup λ_v if and only if v lies in the $\mathbb{R}$-linear span $\mathbb{R}\gamma$ of γ. Thus the irreducible components F_i of the fixed point locus of the λ_v-action on $X(\Sigma)$ are the orbit closures $\overline{O}_{\sigma_i}$ where σ_i runs over all cones in Σ which are minimal with respect to the property $v \in \mathbb{R}\sigma_i$.

The closure of O_γ in $X(\Sigma)$ is the toric variety $X(\mathrm{Star}(\gamma))$ given by the quotient fan $\mathrm{Star}(\gamma)$ in $N(\gamma) = N/(N \cap \mathbb{R}\gamma)$; see [8, page 52]. From this we can derive the following lemma.

Lemma 2.10 *For $v \in N$ and $x \in O_\gamma$ the limit $\lim_{z\to 0} \lambda_v(z)\, x$ exists and lies in $F_i = \overline{O}_{\sigma_i}$ if and only if $\gamma \subseteq \sigma_i$ is a face and the image of v in $N_{\mathbb{R}}/\mathbb{R}\gamma$ is in the relative interior of $\sigma_i/\mathbb{R}\gamma$.*

The set of all faces γ of σ_i with this property is closed under taking intersections and hence this set has a unique minimal element. We denote this minimal element by τ_i. Thus if we denote

$$U_i^v \quad = \quad \bigl\{\, x \in X(\Sigma) \,:\, \lim_{z\to 0} \lambda_v(z)\, x \;\text{ exists and lies in } F_i \,\bigr\},$$

or just U_i for short, then this set decomposes as a union of orbits as follows:

$$U_i \quad = \quad \cup_{\tau_i \subseteq \gamma \subseteq \sigma_i} O_\gamma. \tag{8}$$

In what follows we further suppose $v \in |\Sigma|$. Then Lemma 2.10 implies $X(\Sigma) = \cup_i U_i$, which is the *Bialynicki-Birula decomposition* [2] of the toric variety with respect to the one-parameter subgroup λ_v.

We now apply this to our semi-projective toric variety $X(A,\theta)$ with fan $\Sigma = \Sigma_\theta$. The moment map μ_v for the circle action induced by λ_v is given by the inner product $\mu_v(x) = \langle v, \mu(x)\rangle$ with μ as in (7). We relabel the fixed components F_i according to the values of this moment map, so that

$$\mu_v(F_i) < \mu_v(F_j) \text{ implies } i < j. \tag{9}$$

Given this labeling, the distinguished faces $\tau_i \subseteq \sigma_i$ have the following important property:

$$\tau_i \subseteq \sigma_j \quad \text{implies} \quad i \leq j. \tag{10}$$

This generalizes the property $(*)$ in [8, Chapter 5.2], and it is equivalent to

$$U_j \quad \text{is closed in} \quad U_{\leq j} = \cup_{i \leq j} U_i. \tag{11}$$

This means that the Bialynicki-Birula decomposition of $X(A,\theta)$ is *filtrable* in the sense of [2]. The following is well-known in the projective case.

Proposition 2.11 *The integral cohomology of a smooth semi-projective toric variety $X(A,\theta)$ equals*

$$H^*(X(A,\theta);\mathbb{Z}) \quad \cong \quad \mathbb{Z}[x_1, x_2, \ldots, x_n]/(\mathrm{Circ}(\mathcal{A}) + I_\theta),$$

where I_θ is the Stanley-Reisner ideal of the simplicial fan Σ_θ, i.e. I_θ is generated by square-free monomials $x_{i_1}x_{i_2}\cdots x_{i_k}$ corresponding to non-faces $\{b_{i_1}, b_{i_2}, \ldots, b_{i_k}\}$ of Σ_θ, and $\mathrm{Circ}(\mathcal{A})$ *is the* circuit ideal

$$\mathrm{Circ}(\mathcal{A}) \quad := \quad \langle \sum_{i=1}^{n} \lambda_i x_i \mid \lambda \in \mathbb{Z}^n, A \cdot \lambda = 0 \rangle.$$

Proof: Let $D_1, D_2, \ldots, D_n$ denote the divisors corresponding to the rays $b_1, b_2, \ldots, b_n$ in Σ_θ. The cohomology class of any torus orbit closure $\overline{O}_\sigma$ can be expressed in terms of the D_i's, namely if the rays in σ are $b_{i_1}, b_{i_2}, \ldots, b_{i_k}$, then $[\overline{O}_\sigma] = [D_{i_1}][D_{i_2}]\cdots[D_{i_k}]$. Following the reasoning in [8, Section 5.2], we first prove that certain torus orbit closures linearly span $H^*(X(A,\theta);\mathbb{Z})$ and hence the cohomology classes $[D_1], [D_2], \ldots, [D_n]$ generate $H^*(X(A,\theta);\mathbb{Z})$ as a $\mathbb{Z}$-algebra.

We choose $v \in |\Sigma|$ to be generic, so that each σ_i is $(n-d)$-dimensional and each F_i is just a point. Then (8) shows that U_i is isomorphic with the affine space $\mathbb{C}^{n-k_i}$, where $k_i = \dim(\tau_i)$.

We set $U_{\leq j} = \cup_{i \leq j} U_i$ and $U_{<j} = \cup_{i<j} U_i$. Note that U_j is closed in $U_{\leq j}$. Thus writing down the cohomology long exact sequence of the pair $(U_{\leq j}, U_{<j})$, we can show by induction on j that the cohomology classes of the closures of the cells U_i generate $H^*(X(A,\theta);\mathbb{Z})$ additively. Because the closure of a cell U_i is the closure of a torus orbit, it follows that the cohomology classes $[D_1], [D_2], \ldots, [D_n]$ generate $H^*(X(A,\theta);\mathbb{Z})$. Thus sending $x_i \mapsto [D_i]$ defines a surjective ring map $\mathbb{Z}[x_1, \ldots, x_n] \to H^*(X(A,\theta);\mathbb{Z})$, whose kernel is seen

to contain $\mathrm{Circ}(\mathcal{A}) + I_\theta$. That this is precisely the kernel follows from the "algebraic moving lemma" of [8, page 107]. $\square$

A similar proof works with $\mathbb{Q}$-coefficients when $X(A, \theta)$ is not smooth but just an orbifold.

Corollary 2.12 *The rational cohomology ring of a semi-projective toric orbifold $X(A, \theta)$ equals*

$$H^*(X(A, \theta); \mathbb{Q}) \quad \cong \quad \mathbb{Q}[u_1, u_2, \ldots, u_n]/(\mathrm{Circ}(\mathcal{A}) + I_\theta).$$

In light of Corollary 2.12, the Betti numbers of $X(A, \theta)$ satisfy $b_{2i} = h_i(\Sigma_\theta)$, where $h_i(\Sigma_\theta)$ are the h-numbers of the Stanley-Reisner ideal I_θ, cf. [18, Section III.3]. This observation leads to the following result.

Corollary 2.13 *If $f_i(P_\theta^{bd})$ denotes the number of i-dimensional bounded faces of P_θ then the Betti numbers of the semi-projective toric orbifold $X(A, \theta)$ are given by the following formula:*

$$b_{2k} \quad = \quad \dim_{\mathbb{Q}} H^{2k}(X(A, \theta); \mathbb{Q}) \quad = \quad \sum_{i=k}^{n-d} (-1)^{i-k} \binom{i}{k} f_i(P_\theta^{bd}). \tag{12}$$

Proof: Lemma 2.3 of [19] implies that

$$\sum_{i=0}^{n-d} h_i(\Sigma_\theta) \cdot x^i \;=\; \sum_{\sigma \in \Sigma_\theta \backslash \partial \Sigma_\theta} (x-1)^{n-d-\dim(\sigma)}, \tag{13}$$

where $\partial \Sigma_\theta$ denotes the boundary of Σ_θ. Hence the right hand sum is over all interior cones σ of the fan Σ_θ. These cones are in order-reversing bijection with the bounded faces of P_θ. Hence (13) is the sum of $(x-1)^{\dim(F)}$ where F runs over all bounded faces of P_θ. This proves (12). $\square$

3 The core of a toric variety

The proof of Corollary 2.13 shows the importance of interior cones of Σ_θ. They are the ones for which the closure of the corresponding torus orbit in $X(A, \theta)$ is compact. This suggests the following

Definition 3.1 *The* core *of a semi-projective toric variety $X(A, \theta)$ is $C(A, \theta) = \cup_{\sigma \in \Sigma_\theta \backslash \partial \Sigma_\theta} O_\sigma$. Thus the core $C(A, \theta)$ is the union of all compact torus orbit closures in $X(A, \theta)$.*

Theorem 3.2 *The core of a semi-projective toric orbifold $X(A, \theta)$ is the inverse image of the origin under the canonical projective morphism $X(A, \theta) \to X(A, 0)$ as in (6). It also equals the inverse image of the bounded faces of the polyhedron P_θ under the moment map (7) from $X(A, \theta)$ onto P_θ. In particular, the core of $X(A, \theta)$ is a union of projective toric orbifolds.*

Proof: On the level of fans, the toric morphism $X(A,\theta) \to X(A,0)$ corresponds to forgetting the triangulation of the cone $|\Sigma| = \mathrm{pos}(\mathcal{B})$. It follows from the description of toric morphisms in Section 1.4 of [8] that the inverse image of the origin is the union of the orbit closures corresponding to interior faces of Σ. This was our first assertion. Each face of a simple polyhedron is a simple polyhedron, and each bounded face is a simple polytope. If σ is the interior cone of Σ dual to a bounded face of P_θ then the corresponding orbit closure is the projective toric orbifold $X(\mathrm{Star}(\sigma))$. The core $C(A,\theta)$ is the union of these orbifolds. □

We fix a generic vector $v \in \mathrm{int}|\Sigma|$. Then the F_i above are points and lie in $C(A,\theta)$. In what follows we shall study the action of the one-parameter subgroup λ_v on the core $C(A,\theta)$. We define

$$D_i = U_i^{-v} \quad = \quad \bigl\{ x \in X(A,\theta) : \lim_{z\to\infty} \lambda_v(z)\, x \text{ exists and equals } F_i \bigr\}.$$

Lemma 2.10 implies that this gives a decomposition of the core: $C(A,\theta) = \cup_i D_i$. The closure $\overline{D}_i$ is a projective toric orbifold, and it is the preimage of a bounded face of P_θ via the moment map (7). If we now introduce an ordering as in (9) then the counterpart of (11) is the following:

$$D_{\leq j} \ = \ \cup_{i\leq j} D_i \ \text{ is compact.} \tag{14}$$

This property of the decomposition $C(A,\theta) = \cup_i D_i$ translates into a non-trivial statement about the convex polyhedron P_θ. Let P_θ^{bd} denote the *bounded complex*, that is, the polyhedral complex consisting of all bounded faces of P_θ. Let P_j denote the bounded face of P_θ corresponding to $\overline{D}_j$, and let p_j denote the vertex of P_θ corresponding to F_j. Then $P_{\leq j} = \cup_{i\leq j} P_i$ is a subcomplex of the bounded complex P_θ^{bd}, and $P_{\leq j} \backslash P_{<j}$ consists precisely of those faces of P_j which contain p_j. This property is called *star-collapsibility*. It implies that $P_{<j}$ is a deformation retract of $P_{\leq j}$ and in turn that P_θ^{bd} is contractible. The contractibility also follows from [4, Exercise 4.27 (a)]. In summary we have proven the following result.

Theorem 3.3 *The bounded complex P_θ^{bd} of P_θ is star-collapsible; in particular, it is contractible.*

This theorem implies that the core of any semi-projective toric variety is connected, since $C(A,\theta)$ is the preimage of the bounded complex P_θ^{bd} under the continuous moment map. Moreover, since the cohomology of P_θ^{bd} vanishes, the bounded complex does not contribute to the cohomology of $C(A,\theta)$. This fact is expressed in the following proposition, which will be crucial in Section 7.

Proposition 3.4 *Let $C(A,\theta)$ be the core of a semi-projective toric orbifold and consider a class α in $H^*(C(A,\theta);\mathbb{Q})$. If α vanishes on every irreducible component of $C(A,\theta)$ then $\alpha = 0$.*

Proof: Let $v \in \mathrm{int}|\Sigma|$, F_i and D_i as above. We prove by induction on j that

$$\text{if } \alpha \in H^*(D_{\leq j};\mathbb{Q}) \text{ and } \alpha\,|_{\overline{D}_i} = 0 \text{ for } i \leq j, \text{ then } \alpha = 0. \tag{15}$$

This implies the proposition, because if α vanishes on every irreducible component of the core then it vanishes on every irreducible projective subvariety $\overline{D}_i$ of the core. The statement (15) then implies by induction that α vanishes on the core.
To prove (15) consider the Mayer-Vietoris sequence of the covering $D_{\leq j} = D_{<j} \cup \overline{D}_j$.

$$\ldots \to H^k(D_{\leq j};\mathbb{Q}) \xrightarrow{\alpha} H^k(D_{<j};\mathbb{Q}) \oplus H^k(\overline{D}_j;\mathbb{Q}) \xrightarrow{\beta} H^k(D_{<j} \cap \overline{D}_j;\mathbb{Q}) \to \ldots$$

We show that the map α is injective, which will prove our claim. For this we show that β is surjective. This follows from the surjectivity of $H^k(\overline{D}_j;\mathbb{Q}) \to H^k(\overline{D}_j \backslash D_j;\mathbb{Q})$, because clearly $D_{<j} \cap \overline{D}_j = \overline{D}_j \backslash D_j$.
To prove this we do Morse theory on the projective toric orbifold $\overline{D}_j$. First it follows from Morse theory that $H^*(\overline{D}_j;\mathbb{Q}) \to H^*(\overline{D}_j \backslash F_j;\mathbb{Q})$ surjects. Moreover we have that $\overline{D}_j \backslash D_j$ is the core of the quasi-projective variety $\overline{D}_j \backslash F_j$. This means that $\overline{D}_j \backslash D_j$ is the set of points x in $\overline{D}_j$ such that $\lim_{z\to\infty} \lambda_v(z)x$ is not in F_j. Then the proof of Theorem 3.5 shows that $H^*(\overline{D}_j \backslash F_j;\mathbb{Q})$ is isomorphic with $H^*(\overline{D}_j \backslash D_j;\mathbb{Q})$. This proves (15) and in turn our Proposition 3.4. □

We finish this section with an explicit description of the cohomology ring of $C(A,\theta)$, namely, we identify it with the cohomology of the ambient semi-projective toric orbifold $X(A,\theta)$:

Theorem 3.5 *The embedding of the core $C(A,\theta)$ in $X(A,\theta)$ induces an isomorphism on cohomology with integer coefficients.*

Proof: Let $v \in \mathrm{int}|\Sigma|$, F_i, U_i and D_i as above. We clearly have an inclusion $D_{\leq j} \subset U_{\leq j}$. We show by induction on j that this inclusion induces an isomorphism on cohomology. Consider the following commutative diagram:

$$\begin{array}{ccccccc} \ldots \to & H^k(U_{\leq j}, U_{<j};\mathbb{Z}) & \to & H^k(U_{\leq j};\mathbb{Z}) & \to & H^k(U_{<j};\mathbb{Z}) & \to \ldots \\ & \downarrow & & \downarrow & & \downarrow & \\ \ldots \to & H^k(D_{\leq j}, D_{<j};\mathbb{Z}) & \to & H^k(D_{\leq j};\mathbb{Z}) & \to & H^k(D_{<j};\mathbb{Z}) & \to \ldots \end{array}.$$

The rows are the long exact sequence of the pairs $(U_{\leq j}, U_{<j})$ and $(D_{\leq j}, D_{<j})$ respectively. The vertical arrows are induced by inclusion. The last vertical arrow is an isomorphism by induction.
By excision $H^k(U_{\leq j}, U_{<j};\mathbb{Z}) \cong H^k(T(N_j), t_0; Z)$, where N_j is the normal (orbi-)bundle to U_j and $T(N_j)$ is the Thom space $N_j \cup t_0$, where t_0 is the point at infinity. Similarly $H^k(D_{\leq j}, D_{<j};\mathbb{Z}) \cong H^k(T(D_j), t_0;\mathbb{Z})$, where $T(D_j) = D_{\leq j}/D_{<j}$ is the one point compactification of D_j, which is homeomorphic to the Thom space of $N_j|_{F_j}$, the negative bundle at F_j. Because F_j is a deformation retract of U_j and because the normal bundle N_j to U_j in $U_{\leq j}$

restricts to the normal bundle of F_j in D_j, we find that $T(D_j)$ is a deformation retract of $T(N_j)$. Consequently the first vertical arrow is also an isomorphism. The Five Lemma now delivers our assertion. □

Remark. One can prove more, namely, that $C(A,\theta)$ is a deformation retract of $X(A,\theta)$. This follows from Theorem 3.5 and the analogous statement about the fundamental group, which vanishes for both spaces. Alternatively, one can use Bott-Morse theory in the spirit of the proof of [16, Theorem 3.2] to get the homotopy equivalence.

4 Lawrence toric varieties

In this section we examine an important class of toric varieties which are semi-projective but not projective. We fix an integer $d\times n$-matrix A as in (1), and we write $A^{\pm} = [A, -A]$ for the $d \times 2n$-matrix obtained by appending the negative of A to A. The corresponding vector configuration $\mathcal{A}^{\pm} = \mathcal{A} \cup -\mathcal{A}$ spans $\mathbb{Z}^d$ as a semigroup; in symbols, $\mathbb{N}\mathcal{A}^{\pm} = \mathbb{Z}\mathcal{A} = \mathbb{Z}^d$. A vector θ is *generic* with respect to $\mathcal{A}^{\pm}$ if it does not lie on any hyperplane spanned by a subset of $\mathcal{A}$.

Definition 4.1 *We call $X(A^{\pm},\theta)$ a* Lawrence toric variety*, for any generic vector $\theta \in \mathbb{Z}^d$.*

Our choice of name comes from the Lawrence construction in polytope theory; see e.g. Chapter 6 in [22]. The Gale dual of the centrally symmetric configuration $\mathcal{A}^{\pm}$ is denoted $\Lambda(\mathcal{B})$ and is called the *Lawrence lifting* of $\mathcal{B}$. It consists of $2n$ vectors which span $\mathbb{Z}^{2n-d}$. The cone $\mathrm{pos}(\Lambda(\mathcal{B}))$ is the cone over the $(2n-d-1)$-dimensional *Lawrence polytope* with Gale transform $\mathcal{A}^{\pm}$.
Consider the even-dimensional affine space $\mathbb{C}^{2n}$ with coordinates $z_1,\dots,z_n,w_1,\dots,w_n$. We call a torus action on $\mathbb{C}^{2n}$ *symplectic* if the products $z_1w_1,\dots,z_nw_n$ are fixed under this action.

Proposition 4.2 *The following three classes of toric varieties coincide:*

1. *Lawrence toric varieties,*
2. *toric orbifolds which are GIT-quotients of a symplectic torus action on $\mathbb{C}^{2n}$ for some $n \in \mathbb{N}$,*
3. *toric varieties $X(\Sigma)$ where Σ is the cone over a regular triangulation of a Lawrence polytope.*

Proof: This follows from Theorem 2.6 using the observation that a torus action on $\mathbb{C}^{2n}$ is symplectic if and only if it arises from a matrix of the form $A^{\pm}$. This means the action looks like

$$z_i \mapsto t^{a_i}\cdot z_i\ , \qquad w_i \mapsto t^{-a_i}\cdot w_i \qquad\qquad (i=1,2,\dots,n)$$

Note that a polytope is Lawrence if and only if its Gale transform is centrally symmetric. $\square$

The matrix $A^{\pm}$ is unimodular if and only if the smaller matrix A is unimodular. Therefore unimodularity of A implies the smoothness of the Lawrence toric variety, by Corollary 2.9. An interesting feature of Lawrence toric varieties is that the converse to this statement also holds:

Proposition 4.3 *The Lawrence toric variety $X(A^{\pm}, \theta)$ is smooth if and only if A is unimodular.*

Proof: The chamber complex $\Gamma(\mathcal{A}^{\pm})$ is the arrangement of hyperplanes spanned by subsets of $\mathcal{A}$. The vector θ is assumed to lie in an open cell of that arrangement. For any column basis $C = \{a_{i_1}, \ldots, a_{i_d}\}$ of the $d \times n$-matrix A there exists a unique linear combination

$$\lambda_1 a_{i_1} + \lambda_2 a_{i_2} + \cdots + \lambda_d a_{i_d} \quad = \quad \theta.$$

Here all the coefficients λ_j are non-zero rational numbers. We consider the polynomial ring

$$\mathbb{Z}[x, y] \quad = \quad \mathbb{Z}[x_1, \ldots, x_n, y_1, \ldots, y_n].$$

The $2n$ variables are used to index the elements of $\mathcal{A}^{\pm}$ and the elements of $\Lambda(\mathcal{B})$. We set

$$\sigma(C, \theta) \quad = \quad \{\, x_{i_j} \,:\, \lambda_j > 0 \,\} \,\cup\, \{\, y_{i_j} \,:\, \lambda_j < 0 \,\}.$$

Its complement $\overline{\sigma}(C, \theta) = \big\{x_1, \ldots, x_n, y_1, \ldots, y_n\big\} \backslash \sigma(C, \theta)$ corresponds to a subset of $\Lambda(\mathcal{B})$ which forms a basis of $\mathbb{R}^{2n-d}$. The triangulation Σ_θ of the Lawrence polytope defined by θ is identified with its set of maximal faces. This set equals

$$\Sigma_\theta \quad = \quad \big\{\, \overline{\sigma}(C, \theta) \;:\; C \text{ is any column basis of } A \,\big\}. \tag{16}$$

Hence the Lawrence toric variety $X(A^{\pm}, \theta) = X(\Sigma_\theta)$ is smooth if and only if every basis in $\Lambda(\mathcal{B})$ spans the lattice $\mathbb{Z}^{2n-d}$ if and only if every column basis C of A spans $\mathbb{Z}^d$. The latter condition is equivalent to saying that A is a unimodular matrix. $\square$

Corollary 4.4 *The Stanley-Reisner ideal of the fan Σ_θ equals*

$$I_\theta \quad = \quad \bigcap_C \langle \sigma(C, \theta) \rangle \quad \subset \quad \mathbb{Z}[x, y], \tag{17}$$

i.e. I_θ is the intersection of the monomial prime ideals generated by the sets $\sigma(C, \theta)$ where C runs over all column bases of A. The irrelevant ideal of the Lawrence toric variety $X(\Sigma_\theta)$ equals

$$B_\theta \quad = \quad \langle \prod \sigma(C, \theta) \,:\, C \textit{ is any column basis of } A \,\rangle \quad \subset \quad \mathbb{Z}[x, y]. \tag{18}$$

We now compute the cohomology of a Lawrence toric variety. For simplicity of exposition we assume A is unimodular so that $X(A^{\pm},\theta)$ is smooth. The orbifold case is analogous. First note

$$\mathrm{Circ}(\mathcal{A}^{\pm}) \quad = \quad \langle x_1+y_1, x_2+y_2,\ldots,x_n+x_n\rangle \;+\; \mathrm{Circ}(\mathcal{A}),$$

where $\mathrm{Circ}(\mathcal{A})$ is generated by all linear forms $\sum_{i=1}^n \lambda_i x_i$ such that $\lambda = (\lambda_1,\ldots,\lambda_n)$ lies in $\ker(A) = \mathrm{im}(B)$. From Proposition 2.11, we have

$$H^*\big(X(A^{\pm},\theta) = \mathbb{Z}[x,y]\,/\Big(\langle x_1+y_1, x_2+y_2,\ldots,x_n+y_n\rangle \;+\; \mathrm{Circ}(\mathcal{A}) \;+\; I_\theta\Big).$$

Let ϕ denote the $\mathbb{Z}$-algebra epimorphism which collapses the variables pairwise:

$$\phi\,:\,\mathbb{Z}[x_1,\ldots,x_n,y_1,\ldots,y_n] \,\rightarrow\, \mathbb{Z}[x_1,\ldots,x_n], \quad x_i\mapsto x_i,\; y_i\mapsto -x_i.$$

Then we can rewrite the presentation of the cohomology ring as follows:

$$H^*\big(X(A^{\pm},\theta);\mathbb{Z}\big) \quad = \quad \mathbb{Z}[x_1,\ldots,x_n]\,/\Big(\mathrm{Circ}(\mathcal{A}) \;+\; \phi(I_\theta)\Big).$$

Clearly, the image of the ideal (17) under ϕ is the intersection of the ideals

$$\phi\,\big(\,\langle\sigma(C,\theta)\rangle\,\big) \quad = \quad \langle\, x_i\,:\, i\in C\,\rangle$$

where C runs over the column bases of A. Note that this ideal is independent of the choice of θ. It depends only on A. This ideal is called the *matroid ideal* of $\mathcal{B}$ and it is abbreviated by

$$\begin{aligned} M^*(\mathcal{A}) \quad &= \quad \bigcap\big\{\langle x_{i_1},\ldots,x_{i_d}\rangle \,:\, \{a_{i_1},\ldots,a_{i_d}\}\subseteq\mathcal{A} \quad \text{is linearly independent}\,\big\} \\ &= \quad \langle x_{i_1}\cdots x_{i_k} : \{b_{i_1},\ldots,b_{i_k}\}\subseteq\mathcal{B} \;\text{ is linearly dependent }\rangle \\ &= \quad M(\mathcal{B}). \end{aligned}$$

We summarize what we have proven concerning the cohomology of a Lawrence toric variety.

Theorem 4.5 *The integral cohomology ring of a smooth Lawrence toric variety $X(A^{\pm},\theta)$ is independent of the choice of the generic vector θ in $\mathbb{Z}^d$. It equals*

$$H^*\big(X(A^{\pm},\theta);\mathbb{Z}\big) \quad = \quad \mathbb{Z}[x_1,\ldots,x_n]/\big(\mathrm{Circ}(\mathcal{A}) \,+\, M^*(\mathcal{A})\,\big). \tag{19}$$

The same holds for Lawrence toric orbifolds with $\mathbb{Z}$ replaced by $\mathbb{Q}$.

Remark. The independence of the cohomology ring on θ is an unusual phenomenon in the GIT-construction. Usually, the topology of the quotient changes when one crosses a wall. Theorem 4.5 says that this is not the case for symplectic torus actions. An explanation of this fact is offered through our Theorem 1.1, as there are no walls in the hyperkähler quotient construction.

The ring $\mathbb{Q}[x_1, \ldots, x_n]/M^*(\mathcal{A})$ is the Stanley-Reisner ring of the matroid complex (of linearly independent subsets) of the $(n-d)$-dimensional configuration $\mathcal{B}$. This ring is Cohen-Macaulay, and $\mathrm{Circ}(\mathcal{A})$ provides a linear system of parameters. We write $h(\mathcal{B}) = (h_0, h_1, \ldots, h_{n-d})$ for its h-vector. This is a well-studied quantity in combinatorics; see e.g. [5] and [18, Section III.3].

Corollary 4.6 *The Betti numbers of the Lawrence toric variety $X(A^{\pm}, \theta)$ are independent of θ, and they coincide with the entries in the h-vector of the rank $n-d$ matroid given by $\mathcal{B}$:*

$$\dim_{\mathbb{Q}} H^{2i}(X(A^{\pm}, \theta); \mathbb{Q}) \quad = \quad h_i(\mathcal{B}). \qquad \text{for} \quad i = 0, 1, \ldots, n-d.$$

Our second result in this section concerns the core of a Lawrence toric variety of dimension $2n-d$. We fix a generic vector θ in $\mathbb{Z}^d$. The fan Σ_θ is the normal fan of the unbounded polyhedron

$$P_\theta \quad = \quad \big\{ (u, v) \in \mathbb{R}^n \oplus \mathbb{R}^n \ : \ Au - Av \, = \, \theta, \ u, v \geq 0 \big\}.$$

As in the proof of Proposition 2.3, we chose any vector $\psi \in \mathbb{Z}^n$ such that $A\psi = -\theta$, and we consider the following full-dimensional unbounded polyhedron in $\mathbb{R}^{2n-d}$:

$$Q_\psi \quad = \quad \big\{ (w, t) \in \mathbb{R}^{n-d} \oplus \mathbb{R}^n \ : t \, \geq \, 0 \, , \ \ Bw \, + \, t \, \geq \, \psi \ \big\}.$$

The map $(w, t) \mapsto (Bw + t - \psi, t)$ is an affine-linear isomorphism from Q_ψ onto P_θ. We define $\mathcal{H}(B, \psi)$ to be the arrangement of the following n hyperplanes in $\mathbb{R}^{n-d}$:

$$\{ \, w \in \mathbb{R}^{n-d} \ : \ b_i \cdot w = \psi_i \, \} \qquad \qquad (i = 1, 2, \ldots, n).$$

The arrangement $\mathcal{H}(B, \psi)$ is regarded as a polyhedral subdivision of $\mathbb{R}^{n-d}$ into relatively open polyhedra of various dimensions. The collection of all such polyhedra which are bounded form a subcomplex, called the *bounded complex* of $\mathcal{H}(B, \psi)$ and denoted by $\mathcal{H}^{bd}(B, \psi)$.

Theorem 4.7 *The bounded complex $\mathcal{H}^{bd}(B, \psi)$ of the hyperplane arrangement $\mathcal{H}(B, \psi)$ in $\mathbb{R}^{n-d}$ is isomorphic to the complex of bounded faces of the $(2n-d)$-dimensional polyhedron $Q_\psi \simeq P_\theta$.*

Proof: We define an injective map from $\mathbb{R}^{n-d}$ into the polyhedron Q_ψ as follows

$$w \, \mapsto \, \big(w, t\big), \qquad \text{where} \ \ t_i \, = \, \max\{0, \psi_i - b_i \cdot w\}. \tag{20}$$

This map is linear on each cell of the hyperplane arrangement $\mathcal{H}(B, \psi)$, and the image of each cell is a face of Q_ψ. In particular, every bounded cell of $\mathcal{H}(B, \psi)$ is mapped to a bounded face of Q_ψ and each unbounded cell of $\mathcal{H}(B, \psi)$ is mapped to an unbounded face of Q_ψ. It remains to be shown that every bounded face of Q_ψ lies in the image of the map (20).

Now, the image of (20) is the following subcomplex in the boundary of our polyhedron:

$$\begin{aligned} & \{\, (w,t) \in Q_\psi \ : \ t_i \cdot (b_i \cdot w + t_i - \psi_i) \,=\, 0 \ \text{ for } i = 1,2,\ldots,n \,\} \\ \simeq \quad & \{\, (u,v) \in P_\theta \ : \ u_i \cdot v_i \,=\, 0 \ \text{ for } i = 1,2,\ldots,n \,\} \end{aligned}$$

Consider any face F of P_θ which is not in this subcomplex, and let (u,v) be a point in the relative interior of F. There exists an index i with $u_i > 0$ and $v_i > 0$. Let e_i denote the i-th unit vector in $\mathbb{R}^n$. For every positive real λ, the vector $(u + \lambda e_i, v + \lambda e_i)$ lies in P_θ and has the support as (u,v). Hence $(u + \lambda e_i, v + \lambda e_i)$ lies in F for all $\lambda \geq 0$. This shows that F is unbounded. $\square$

Theorem 4.7 and Corollary 2.13 imply the following enumerative result:

Corollary 4.8 *The Betti numbers of the Lawrence toric variety $X(A^{\pm},\theta)$ satisfy*

$$\dim_{\mathbb{Q}} H^{2i}(X(A^{\pm},\theta);\mathbb{Q}) \quad = \quad \sum_{i=k}^{n-d} (-1)^{i-k} \binom{i}{k} f_i(\mathcal{H}^{bd}(B,\psi)),$$

where $f_i(\mathcal{H}^{bd}(B,\psi))$ denotes the number of i-dimensional bounded regions in $\mathcal{H}(B,\psi)$.

There are two natural geometric structures on any Lawrence toric variety. First the canonical bundle of $X(A^{\pm},\theta)$ is trivial, because the vectors in $\mathcal{A}^{\pm}$ add to 0. This means that $X(A^{\pm},\theta)$ is a *Calabi-Yau variety.* Moreover, since the symplectic $\mathbb{T}^d_{\mathbb{C}}$-action preserves the natural Poisson structure on $\mathbb{C}^{2n} \cong \mathbb{C}^n \oplus (\mathbb{C}^n)^*$, the GIT quotient $X(A^{\pm},\theta)$ inherits a natural *holomorphic Poisson structure.* The holomorphic symplectic leaves of this Poisson structure are what we call *toric hyperkähler manifolds.* The special leaf which contains the core of $X(A^{\pm},\theta)$ will be called the *toric hyperkähler variety.* We present these definitions in complete detail in the following two sections.

5 Hyperkähler quotients

Our aim is to describe an algebraic approach to the toric hyperkähler manifolds of Bielawski and Dancer [3]. In this section we sketch the original differential geometric construction in [3]. This construction is the hyperkähler analogue to the construction of toric varieties using Kähler quotients. We first briefly review the latter. Fix the standard Euclidean bilinear form on $\mathbb{C}^n$,

$$g(z,w) \quad = \quad \sum_{i=1}^{n} \left(\mathrm{re}(z_i)\mathrm{re}(w_i) + \mathrm{im}(z_i)\mathrm{im}(w_i) \right).$$

The corresponding *Kähler form* is

$$\omega(z,w) \quad = \quad g(iz,w) \quad = \quad \sum_{i=1}^{n} \left(\mathrm{re}(z_i)\mathrm{im}(w_i) - \mathrm{im}(z_i)\mathrm{re}(w_i) \right).$$

Let A be as in (1) and consider the real torus $\mathbb{T}^d_{\mathbb{R}}$ which is the maximal compact subgroup of $\mathbb{T}^d_{\mathbb{C}}$. The group $\mathbb{T}^d_{\mathbb{R}}$ acts on $\mathbb{C}^n$ preserving the Kähler structure. This action has the moment map

$$\mu_{\mathbb{R}} : \mathbb{C}^n \to (\mathfrak{t}^d_{\mathbb{R}})^* \cong \mathbb{R}^d, \quad (z_1, \dots, z_n) \mapsto \frac{1}{2}\sum_{i=1}^n |z_i|^2 a_i. \tag{21}$$

Fix $\xi_{\mathbb{R}} \in \mathbb{R}^d$. The *Kähler quotient* $X(A, \xi_{\mathbb{R}}) = \mathbb{C}^n//_{\xi_{\mathbb{R}}}\mathbb{T}^d_{\mathbb{R}} = \mu_{\mathbb{R}}^{-1}(\xi_{\mathbb{R}})/\mathbb{T}^d_{\mathbb{R}}$ inherits a Kähler structure from $\mathbb{C}^n$ at its smooth points. If $\xi_{\mathbb{R}} = \theta$ lies in the lattice $\mathbb{Z}^d$ then there is a biholomorphism between the smooth loci in the GIT quotient $X(A, \theta)$ and the Kähler quotient $X(A, \xi_{\mathbb{R}})$. Hence if A is unimodular and θ generic then the complex manifolds $X(A, \theta)$ and $X(A, \xi_{\mathbb{R}})$ are biholomorphic.
Now we turn to toric hyperkähler manifolds. Let $\mathbb{H}$ be the skew field of *quaternions*, the 4-dimensional real vector space with basis $1, i, j, k$ and associative algebra structure given by $i^2 = j^2 = k^2 = ijk = -1$. Left multiplication by i (resp. j and k) defines complex structures $I : \mathbb{H} \to \mathbb{H}$, with $I^2 = -\mathrm{Id}_{\mathbb{H}}$, (resp. J and K) on $\mathbb{H}$. We now put the flat metric g on $\mathbb{H}$ arising from the standard Euclidean scalar product on $\mathbb{H} \cong \mathbb{R}^4$ with $1, i, j, k$ as an orthonormal basis. This is called a *hyperkähler metric* because it is a Kähler metric with respect to all three complex structures I, J and K. It means that the differential 2-forms, the so-called *Kähler forms*, given by $\omega_I(X, Y) = g(IX, Y)$ for tangent vectors X and Y, and the analogously defined ω_J and ω_K are closed. For the reader's convenience we write down these Kähler forms in coordinates (x, y, u, v):

$$\begin{aligned} \omega_I &= dx \wedge dy + du \wedge dv, \\ \omega_J &= dx \wedge du + dv \wedge dy, \\ \omega_K &= dx \wedge dv + dy \wedge du. \end{aligned}$$

A special orthogonal transformation, with respect to this metric, is said to preserve the hyperkähler structure if it commutes with all three complex structures I, J and K or equivalently if it preserves the Kähler forms ω_I, ω_J and ω_K. The group of such transformations, the unitary symplectic group $Sp(1)$, is generated by multiplication by unit quaternions from the right. A maximal abelian subgroup $\mathbb{T}_{\mathbb{R}} \cong U(1) \subset Sp(1)$ is thus specified by a choice of a unit quaternion. We break the symmetry between I, J and K and choose the maximal torus generated by multiplication from the right by the unit quaternion i. Thus $U(1)$ acts on $\mathbb{H}$ by sending ξ to $\xi \exp(\phi i)$, for $exp(\phi i) \in U(1) \subset \mathbb{R}\oplus\mathbb{R}i \cong \mathbb{C}$. It follows from (21) that the moment map $\mu_I : \mathbb{H} \to \mathbb{R}$ with respect to the symplectic form ω_I is given by

$$\mu_I(x + yi + uj + vk) = \mu_I(x + yi + (-ui + v)k) = \frac{1}{2}(x^2 + y^2 - u^2 - v^2). \tag{22}$$

Similarly we obtain formulas for μ_J and μ_K by writing down the eigenspace decomposition in the respective complex structures:

$$\mu_J(x + yi + uj + vk) =$$

$$= \mu_J \left[\left(\frac{y+u}{\sqrt{2}} + \frac{-x-v}{\sqrt{2}} j \right) \frac{i+j}{\sqrt{2}} + \left(\frac{y-u}{\sqrt{2}} j + \frac{-x+v}{\sqrt{2}} \right) \frac{k-1}{\sqrt{2}} \right] = yu + xv,$$

$$\mu_K(x+yi+uj+vk) =$$

$$= \mu_K \left[\left(\frac{y+v}{\sqrt{2}} + \frac{-x+u}{\sqrt{2}} k \right) \frac{i+k}{\sqrt{2}} + \left(\frac{y-v}{\sqrt{2}} + \frac{x+u}{\sqrt{2}} k \right) \frac{i-k}{\sqrt{2}} \right] = yv - xu.$$

We now consider the map $\mu_{\mathbb{C}} = \mu_J + i\mu_K$ from $\mathbb{H}$ to $\mathbb{C}$. It can be thought of as the holomorphic moment map for the I-holomorphic action of $\mathbb{T}_{\mathbb{C}} \supset \mathbb{T}_{\mathbb{R}}$ on $\mathbb{H}$ with respect to the I-holomorphic symplectic form $\omega_{\mathbb{C}} = \omega_J + i\omega_K$. If we identify $\mathbb{H}$ with $\mathbb{C} \oplus \mathbb{C}$ by introducing two complex coordinates, $z = x + iy \in \mathbb{R} \oplus \mathbb{R}i \cong \mathbb{C}$ and $w = v - ui \in \mathbb{R} \oplus \mathbb{R}i \cong \mathbb{C}$, then the I-holomorphic moment map $\mu_{\mathbb{C}} : \mathbb{H} \to \mathbb{C}$ is given algebraically by multiplying complex numbers:

$$\mu_{\mathbb{C}}(z,w) \;=\; \mu_J(z,w) \;+\; i\mu_K(z,w) \;=\; yu + xv \;+\; i(yv - xu) \;=\; zw. \tag{23}$$

The discussion in the previous paragraph generalizes in an obvious manner to $\mathbb{H}^n$ for $n > 1$. Indeed, the n-dimensional quaternionic space $\mathbb{H}^n$ has three complex structures I,J and K, given by left multiplication with $i, j, k \in \mathbb{H}$. Putting the flat metric $g_n = g^{\oplus n}$ on $\mathbb{H}^n$ yields a hyperkähler metric, i.e. the differential 2-forms $\omega_I(X,Y) = g_n(IX,Y)$ and similarly ω_J and ω_K are Kähler (meaning closed) forms. The automorphism group of this hyperkähler structure is the unitary symplectic group $Sp(n)$. We fix the maximal torus $\mathbb{T}^n_{\mathbb{R}} = U(1)^n \subset Sp(n)$ given by the following definition. For $\lambda = (\exp(\phi_1 i), \exp(\phi_2 i), \ldots, \exp(\phi_n i)) \in \mathbb{T}^n_{\mathbb{R}}$ and $(\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{H}^n$ we set

$$\lambda(\xi_1, \xi_2, \ldots, \xi_n) \quad = \quad (\xi_1 \exp(\phi_1 i), \xi_2 \exp(\phi_2 i), \ldots, \xi_n \exp(\phi_n i)). \tag{24}$$

As in the $n = 1$ case above, this fixes an isomorphism $\mathbb{H}^n \cong \mathbb{C}^n \oplus \mathbb{C}^n$ where two complex vectors $z, w \in \mathbb{C}^n \cong \mathbb{R}^n \oplus i\mathbb{R}^n$ represent the quaternionic vector $z + wk \in \mathbb{H}^n \cong \mathbb{R}^n \oplus i\mathbb{R}^n \oplus j\mathbb{R}^n \oplus k\mathbb{R}^n$. Expressing vectors in $\mathbb{H}^n$ in these complex coordinates, the torus action (24) translates into

$$\lambda(z,w) \quad = \quad (\lambda z, \lambda^{-1} w) \qquad \text{for } \lambda \in \mathbb{T}^n_{\mathbb{R}} \quad \text{and} \quad (z,w) \in \mathbb{H}^n. \tag{25}$$

The toric hyperkähler manifolds in [3] are constructed by choosing a subtorus $\mathbb{T}^d_{\mathbb{R}} \subset \mathbb{T}^n_{\mathbb{R}}$ and taking the hyperkähler quotient [10] of $\mathbb{H}^n$ by $\mathbb{T}^d_{\mathbb{R}}$. We do this by choosing integer matrices A and B as in (1) and (2). The subtorus $\mathbb{T}^d_{\mathbb{R}}$ of $\mathbb{T}^n_{\mathbb{R}}$ acts on $\mathbb{H}^n$ by (25) preserving the hyperkähler structure. The *hyperkähler moment map* of the action (25) of $\mathbb{T}^d_{\mathbb{R}}$ on $\mathbb{H}^n$ is defined by

$$\mu = (\mu_I, \mu_J, \mu_K) \,:\, \mathbb{H}^n \;\to\; (\mathfrak{t}^d_{\mathbb{R}})^* \otimes \mathbb{R}^3,$$

where μ_I, μ_J and μ_K are the Kähler moment maps with respect to ω_I, ω_J and ω_K respectively. Using the formulas (22) and (23), the components of μ are in complex coordinates as follows:

$$\mu_{\mathbb{R}}(z,w) \;:=\; \mu_I(z,w) \quad = \quad \frac{1}{2} \sum_{i=1}^{n} (|z_i|^2 - |w_i|^2) \cdot a_i \quad \in \quad (\mathfrak{t}^d_{\mathbb{R}})^*, \tag{26}$$

$$\mu_{\mathbb{C}}(z,w) := \mu_J(z,w) + i\mu_K(z,w) = \sum_{i=1}^{n} z_i w_i \cdot a_i \in (\mathfrak{t}^d_{\mathbb{R}})^* \otimes \mathbb{C} \cong (\mathfrak{t}^d_{\mathbb{C}})^*. \quad (27)$$

Here a_i is the i-th column vector of the matrix A. We can also think of $\mu_{\mathbb{C}}$ as the moment map for the I-holomorphic action of $\mathbb{T}^d_{\mathbb{C}}$ on $\mathbb{H}^n$ with respect to $\omega_{\mathbb{C}} = \omega_J + i\omega_K$. Now take

$$\xi = (\xi^1, \xi^2, \xi^3) \in (\mathfrak{t}^d_{\mathbb{R}})^* \otimes \mathbb{R}^3$$

and introduce $\xi_{\mathbb{R}} = \xi^1 \in (\mathfrak{t}^d_{\mathbb{R}})^*$ and $\xi_{\mathbb{C}} = \xi^2 + i\xi^3 \in (\mathfrak{t}^d_{\mathbb{C}})^*$ so we can write $\xi = (\xi_{\mathbb{R}}, \xi_{\mathbb{C}}) \in (\mathfrak{t}^d_{\mathbb{R}})^* \oplus (\mathfrak{t}^d_{\mathbb{C}})^*$. The *hyperkähler quotient* of $\mathbb{H}^n$ by the action (25) of the torus $\mathbb{T}^d_{\mathbb{R}}$ at level ξ is defined as

$$Y(A,\xi) := \mathbb{H}^n ////_{\xi} \mathbb{T}^d_{\mathbb{R}} := \mu^{-1}(\xi)/\mathbb{T}^d_{\mathbb{R}} = \left(\mu_{\mathbb{R}}^{-1}(\xi_{\mathbb{R}}) \cap \mu_{\mathbb{C}}^{-1}(\xi_{\mathbb{C}})\right)/\mathbb{T}^d_{\mathbb{R}}. \quad (28)$$

By a theorem of [10], this quotient has a canonical hyperkähler structure on its smooth locus.

Bielawski and Dancer show in [3] that if $\xi \in (\mathfrak{t}^d_{\mathbb{R}})^* \otimes \mathbb{R}^3$ is generic then $Y(A,\xi)$ is an orbifold, and it is smooth if and only if A is unimodular. Since ξ is generic outside a set of codimension three in $(\mathfrak{t}^d_{\mathbb{R}})^* \otimes \mathbb{R}^3$, they can show that the topology and therefore the cohomology of the toric hyperkähler manifold is independent on ξ. In what follows we consider vectors ξ for which $\xi_{\mathbb{C}} = 0$ in $\mathbb{C}^d$ and $\xi_{\mathbb{R}} = \theta \in \mathbb{Z}^d \subset \mathbb{R}^d \cong (\mathfrak{t}^d_{\mathbb{R}})^*$. The underlying complex manifold in complex structure I of the hyperkähler manifold $Y(A,(\theta, 0_{\mathbb{C}}))$ has a purely algebraic description as explained in the next section.

6 Algebraic construction of toric hyperkähler varieties

The $\mathbb{Z}^d$-graded polynomial ring $\mathbb{C}[z,w] = \mathbb{C}[z_1,\ldots,z_n,w_1,\ldots,w_n]$, with the grading given by $A^{\pm} = [A,-A]$, is the homogeneous coordinate ring of the Lawrence toric variety $X(A^{\pm},\theta)$. By a result of Cox [6], closed subschemes of $X(A^{\pm},\theta)$ correspond to homogeneous ideals in $\mathbb{C}[z,w]$ which are saturated with respect to the irrelevant ideal B_θ in (18). Let us now consider the ideal

$$\mathrm{Circ}(\mathcal{B}) := \langle \sum_{i=1}^{n} a_{ij} z_i w_i \mid j = 1,\ldots,d \rangle \subset \mathbb{C}[z,w], \quad (29)$$

whose generators are the components of the holomorphic moment map $\mu_{\mathbb{C}}$ of (27). The ideal $\mathrm{Circ}(\mathcal{B})$ is clearly homogeneous and it is a complete intersection. We assume that none of the row vectors of the matrix B is zero. Under this hypothesis, the ideal $\mathrm{Circ}(\mathcal{B})$ is a prime ideal.

Definition 6.1 *The* toric hyperkähler variety $Y(A,\theta)$ *is the irreducible subvariety of the Lawrence toric variety* $X(A,\theta)$ *defined by the homogeneous ideal* $\mathrm{Circ}(\mathcal{B})$ *in the coordinate ring* $\mathbb{C}[z,w]$ *of* $X(A,\theta)$.

Proposition 6.2 *If θ is generic then the toric hyperkähler variety $Y(A,\theta)$ is an orbifold. It is smooth if and only if the matrix A is unimodular.*

Proof: It follows from (27) that a point in $\mathbb{C}^{2n}$ has a finite stabilizer under the group $\mathbb{T}^d_{\mathbb{C}}$ if and only if the point is regular for $\mu_{\mathbb{C}}$ of (27), i.e. if the derivative of $\mu_{\mathbb{C}}$ is surjective there. This implies that, for θ generic, the toric hyperkähler variety $Y(A,\theta)$ is an orbifold because then the variety $X(A^{\pm},\theta)$ is an orbifold. For the second statement note that if A is unimodular then $X(A^{\pm},\theta)$ is smooth, consequently $Y(A,\theta)$ is also smooth. However, if A is not unimodular then $X(A^{\pm},\theta)$ has orbifold singularities which lie in the core. Now the core $C(A^{\pm},\theta)$ lies entirely in $Y(A,\theta)$, by Lemma 6.4 below, thus $Y(A,\theta)$ inherits singular points from $X(A^{\pm},\theta)$. □

We can now prove that our toric hyperkähler varieties are biholomorphic to the toric hyperkähler manifolds of the previous section.

Theorem 6.3 *Let $\xi_{\mathbb{R}} = \theta \in \mathbb{Z}^d \subset (\mathfrak{t}^d_{\mathbb{R}})^* \cong \mathbb{R}^d$ for generic θ. Then the toric hyperkähler manifold $Y(A,(\xi_{\mathbb{R}},0))$ with complex structure I is biholomorphic with the toric hyperkähler variety $Y(A,\theta)$.*

Proof: Suppose A is unimodular. The general theory of Kähler quotients (e.g. in [12]) implies that the Lawrence toric variety $X(A^{\pm},\theta)$ and the corresponding Kähler quotient $X(A^{\pm},\xi_{\mathbb{R}}) = \mu_{\mathbb{R}}^{-1}(\xi_{\mathbb{R}})/\mathbb{T}^d_{\mathbb{R}}$ are biholomorphic, where $\mu_{\mathbb{R}}$ is defined in (26) and $\xi_{\mathbb{R}} = \theta \in \mathbb{Z}^d \subset \mathbb{R}^d \cong (\mathfrak{t}^d_{\mathbb{R}})^*$. Now the point is that $\mu_{\mathbb{C}} : \mathbb{H}^n \to \mathbb{C}^d$ is invariant under the action of $\mathbb{T}^d_{\mathbb{R}}$ and therefore descends to a map on $X(A^{\pm},\xi_{\mathbb{R}}) = \mu_{\mathbb{R}}^{-1}(\xi_{\mathbb{R}})/\mathbb{T}^d_{\mathbb{R}}$ and similarly on $X(A^{\pm},\theta)$ making the following diagram commutative:

$$\begin{array}{rccc} \mu^{\xi}_{\mathbb{C}}: & X(A^{\pm},\xi_{\mathbb{R}}) & \to & \mathbb{C}^d \\ & \cong & & \cong \\ \mu^{\theta}_{\mathbb{C}}: & X(A^{\pm},\theta) & \to & \mathbb{C}^d \end{array}.$$

It follows that $Y(A,(\xi_{\mathbb{R}},0)) = (\mu^{\xi}_{\mathbb{C}})^{-1}(0)$ and $Y(A,\theta) = (\mu^{\theta}_{\mathbb{C}})^{-1}(0)$ are biholomorphic. The proof is similar when the spaces have orbifold singularities. □

Recall the affinization map $\pi_X : X(A^{\pm},\theta) \to X(A^{\pm},0)$ from (6), and the analogous map $\pi_Y : Y(A,\theta) \to Y(A,0)$. These fit together in the following commutative diagram:

$$\begin{array}{ccc} Y(A,\theta) & \stackrel{\pi_Y}{\to} & Y(A,0) \\ i_\theta \downarrow & & \downarrow i_0 \\ X(A^{\pm},\theta) & \stackrel{\pi_X}{\to} & X(A^{\pm},0) \\ \mu^{\theta}_{\mathbb{C}} \downarrow & & \downarrow \mu^{0}_{\mathbb{C}} \\ \mathbb{C}^d & \cong & \mathbb{C}^d \end{array},$$

where $i_\theta : Y(A,\theta) \to X(A^{\pm},\theta)$ denotes the natural embedding in Definition 6.1 by the preimage of $\mu^{\theta}_{\mathbb{C}}$ at $0 \in \mathbb{C}^d$. From this we deduce the following lemma:

Lemma 6.4 *The cores of the Lawrence toric variety and of the toric hyperkähler variety coincide, that is, $C(A^{\pm},\theta) = \pi_X^{-1}(0) = \pi_Y^{-1}(0)$.*

Remark. It is shown in [3] that the core of the toric hyperkähler manifold $Y(A,\theta)$ is the preimage of the bounded complex in the hyperplane arrangement $\mathcal{H}(\mathcal{B},\psi)$ by the hyperkähler moment map. We know from Theorem 3.2 that the core of the Lawrence toric variety equals the preimage of P_θ^{bd} under the Kähler moment map. Thus Theorem 4.7 is a combinatorial analogue of Lemma 6.4.

We need one last ingredient in order to prove the theorem stated in the Introduction.

Lemma 6.5 *The embedding of the core $C(A^\pm,\theta)$ in $Y(A,\theta)$ gives an isomorphism in cohomology.*

Proof: Consider the $\mathbb{T}_\mathbb{C}$-action on the Lawrence toric variety $X(A^\pm,\theta)$ defined by the vector $v=\sum_{i=1}^n b_i \in \mathbb{Z}^{n-d}$. This action comes from multiplication by non-zero complex numbers on the vector space $\mathbb{C}^{2n}$. The holomorphic moment map $\mu_\mathbb{C}$ of (27) is homogeneous with respect to multiplication by a non-zero complex number, and consequently $\mu_\mathbb{C}^\theta$ is also homogeneous with respect to the circle action λ_v. It follows that this $\mathbb{T}_\mathbb{C}$-action leaves the toric hyperkähler variety invariant. Moreover, since v is in the interior of $\mathrm{pos}(\mathcal{B})$, all the results in Section 3 are valid for this $\mathbb{T}_\mathbb{C}$-action on $X(A^\pm,\theta)$. Now the proof of Theorem 3.5 can be repeated verbatim to show that the cohomology of $Y(A,\theta)$ agrees with the cohomology of the core. □

Proof of Theorem 1.1: 1.= 3. is a consequence of Lemma 6.4 and Lemma 6.5.
2.= 3. This is a consequence of Theorem 3.5.
1.= 4. is the content of Theorem 4.5. □

Remark. 1. In fact, we could claim more than the isomorphism of cohomology rings in Theorem 1.1. The remark after Theorem 3.5 implies that the spaces $C(A^\pm,\theta)\subset Y(A,\theta)\subset X(A^\pm,\theta)$ are deformation retracts in one another. A similar result appears in [3, Theorem 6.5].
2. The result 2.=4. in the smooth case was proven by Konno in [14].
3. We deduce from Theorem 1.1, Corollary 4.6 and Corollary 4.8 the following formulas for Betti numbers. The second formula is due to Bielawski and Dancer [3, Theorem 6.7].

Corollary 6.6 *The Betti numbers of the toric hyperkähler variety $Y(A,\theta)$ agree with:*

- *the h-numbers of the matroid of $\mathcal{B}$:* $b_{2k}(Y(A,\theta)) = h_k(\mathcal{B})$.
- *the following linear combination of the number of bounded regions of the affine hyperplane arrangement $\mathcal{H}(B,\psi)$:*

$$b_{2k}(Y(A,\theta)) \quad = \quad \sum_{i=k}^{n-d}(-1)^{i-k}\binom{i}{k} f_i(\mathcal{H}^{bd}(B,\psi)). \tag{30}$$

This corollary shows the importance of the combinatorics of the bounded complex $\mathcal{H}^{bd}(B,\psi)$ in the topology of $Y(A,\theta)$ and $X(A^{\pm},\theta)$. This intriguing connection will be more apparent in the next section. Before we get there we infer some important properties of the bounded complex from Corollary 9.1 and Theorem E of [21].

Proposition 6.7 *The bounded complex $\mathcal{H}^{bd}(B,\psi)$ is pure-dimensional. If ψ is generic and $\mathcal{B}$ is coloop-free then every maximal face of $\mathcal{H}^{bd}(B,\psi)$ is an $(n-d)$-dimensional simple polytope.*

A *coloop* of $\mathcal{B}$ is a vector b_i which lies in every column basis of B. This is equivalent to a_i being zero. Note that if A has a zero column then we can delete it to get A', which means that $Y(A,\theta) = Y(A',\theta) \times \mathbb{C}^2$ and similarly for the Lawrence toric variety. Therefore we will assume in the next section that none of the columns of A is zero.

7 Cogenerators of the cohomology ring

There are three natural presentations of the cohomology ring of the toric hyperkähler variety $Y(A,\theta)$ associated with a $d \times n$-matrix A and a generic vector $\theta \in \mathbb{Z}^d$. In these presentations $H^*(Y(A,\theta);\mathbb{Q})$ is expressed as a quotient of the polynomial ring $\mathbb{Q}[x,y]$ in $2n$ variables, as a quotient of the polynomial ring $\mathbb{Q}[x]$ in n variables, or as a quotient of the polynomial ring $\mathbb{Q}[\,t\,] \simeq \mathbb{Q}[x]/\mathrm{Circ}(\mathcal{A})$ in d variables, respectively. In this section we compute systems of cogenerators for $H^*(Y(A,\theta);\mathbb{Q})$ relative to each of the three presentations. As an application we show that the Hard Lefschetz Theorem holds for toric hyperkähler varieties, and we discuss some implications for the combinatorial problem of classifying the h-vectors of matroid complexes.

We begin by reviewing the definition of cogenerators of a homogeneous polynomial ideal. Consider the commutative polynomial ring generated by a basis of derivations on affine m-space:

$$\mathbb{Q}[\partial] \quad = \quad \mathbb{Q}[\partial_1,\partial_2,\ldots,\partial_m].$$

The polynomials in $\mathbb{Q}[\partial]$ act as linear differential operators with constant coefficients on

$$\mathbb{Q}[x] \quad = \quad \mathbb{Q}[x_1,x_2,\ldots,x_m].$$

If Γ is any subset of $\mathbb{Q}[x]$ then its *annihilator* $\mathrm{Ann}(\Gamma)$ is the ideal in $\mathbb{Q}[\partial]$ consisting of all linear differential operators with constant coefficients which annihilate all polynomials in Γ. If I is any zero-dimensional homogeneous ideal in $\mathbb{Q}[\partial]$ then there exists a finite set Γ of homogeneous polynomials in $\mathbb{Q}[x]$ such that $I = \mathrm{Ann}(\Gamma)$. We say that Γ is a set of *cogenerators* of I. If Γ is a singleton, say, $\Gamma = \{p\}$, then $I = \mathrm{Ann}(\Gamma)$ is a *Gorenstein ideal.* In this case, the polynomial $p = p(x)$ which cogenerates I is unique up to scaling. More generally, if all polynomials in Γ are homogeneous of the same degree then

$I = \mathrm{Ann}(\Gamma)$ is a *level ideal*. In this case, the $\mathbb{Q}$-vector space spanned by Γ is unique, and it is desirable for Γ to be a nice basis for this space.

We replace the vector $\psi = (\psi_1, \ldots, \psi_n)$ in Theorem 4.7 by an indeterminate vector $x = (x_1, \ldots, x_n)$ which ranges over a small neighborhood of ψ in $\mathbb{R}^n$. For x in this neighborhood, the polyhedron Q_x remains simple and combinatorially isomorphic to Q_ψ, and the hyperplane arrangement $\mathcal{H}(B, x)$ remains isomorphic to $\mathcal{H}(B, \psi)$. Let $\Delta_1, \ldots, \Delta_r$ denote the maximal bounded regions of $\mathcal{H}(B, x)$. These are $(n-d)$-dimensional simple polytopes, by Proposition 6.7 and our assumption that $\mathcal{B}$ is coloop-free. They can be identified with the maximal bounded faces of the $(2n-d)$-dimensional polyhedron Q_x, by Theorem 4.7. The volume of the polytope Δ_i is a homogeneous polynomial in x of degree $n-d$ denoted

$$V_i(x) \quad = \quad V_i(x_1, \ldots, x_n) \quad = \quad \mathrm{vol}(\Delta_i) \qquad (i = 1, 2, \ldots, r)$$

Theorem 7.1 *The volume polynomials $V_1, \ldots, V_r$ form a basis of cogenerators for the cohomology ring of the Lawrence toric variety $X(A^\pm, \theta)$ and of the toric hyperkähler variety $Y(A, \theta)$:*

$$H^*(Y(A,\theta); \mathbb{Q}) \quad = \quad \mathbb{Q}[\partial_1, \partial_2, \ldots, \partial_n]/\mathrm{Ann}\big(\{V_1, V_2, \ldots, V_r\}\big). \tag{31}$$

Proof: Each simple polytope Δ_i represents an $(n-d)$-dimensional projective toric variety X_i. The core $C(A^\pm, \theta)$ is glued from the toric varieties $X_1, \ldots, X_r$, and it has the same cohomology as $X(A^\pm, \theta)$ and $Y(A, \theta)$ as proved in Theorem 1.1. Hence we get a natural ring epimorphism induced from the inclusion of each toric variety X_i into the core $C(A^\pm, \theta)$:

$$\phi_i \quad : \quad H^*\big(C(A^\pm, \theta); \mathbb{Q}\big) \ \to \ H^*(X_i; \mathbb{Q}). \tag{32}$$

In terms of coordinates, the map ϕ_i is described as follows:

$$\phi_i \,:\, \mathbb{Q}[\partial_1, , \ldots, \partial_n]/\big(M(\mathcal{B}) + \mathrm{Circ}(\mathcal{A})\big) \to \mathbb{Q}[\partial_1, , \ldots, \partial_n]/\big(I_{\Delta_i} + \mathrm{Circ}(\mathcal{A})\big), \tag{33}$$

where I_{Δ_i} is the Stanley-Reisner ring of the simplicial normal fan of the polytope Δ_i. Each facet of Δ_i has the form $\{\, w \in \Delta_i \,:\, b_j \cdot w = \psi_j \,\}$ for some $j \in \{1, 2, \ldots, n\}$. The ideal I_{Δ_i} is generated by all monomials $\partial_{j_1}\partial_{j_2}\cdots\partial_{j_s}$ such that the intersection of the facets $\{\, w \in \Delta_i \,:\, b_{j_\nu} \cdot w = \psi_{j_\nu} \,\}$, for $\nu = 1, 2, \ldots, s$, is the empty set. By the genericity hypothesis on ψ, this will happen if $\{b_{j_1}, b_{j_2}, \ldots, b_{j_s}\}$ is linearly dependent, or, equivalently, if $\partial_{j_1}\partial_{j_2}\cdots\partial_{j_s}$ lies in the matroid ideal $M(\mathcal{B})$. We conclude that $M(\mathcal{B}) \subseteq I_{\Delta_i}$, and the map ϕ_i in (33) is induced by this inclusion.

Proposition 3.4 implies that

$$\ker(\phi_1) \cap \ker(\phi_2) \cap \ \ldots \ \cap \ker(\phi_r) \quad = \quad \{0\}. \tag{34}$$

Here is an alternative proof for this in the toric hyperkähler case. We first note that the top-dimensional cohomology of an equidimensional union of projective varieties equals the direct sum of the pieces:

$$H^{2n-2d}\big(C(A^\pm, \theta); \mathbb{Q}\big) \quad \simeq \quad H^{2n-2d}\big(X_1; \mathbb{Q}\big) \oplus \cdots \oplus H^{2n-2d}\big(X_r; \mathbb{Q}\big), \tag{35}$$

and the restriction of the map ϕ_i to degree $2n-2d$ is the i-th coordinate projection in this direct sum. In particular, (34) holds in the top degree. We now use a theorem of Stanley [18, Theorem III.3.4] which states that the Stanley-Reisner ring of a matroid is level. Using condition (j) in [18, Proposition III.3.2], this implies that the socle of our cohomology ring $H^*\big(C(A^{\pm},\theta);\mathbb{Q}\big)$ consists precisely of the elements of degree $2n-2d$. Suppose that (34) does not hold, and pick a non-zero element $p(\partial)$ of maximal degree in the left hand side. The cohomological degree of $p(\partial)$ is strictly less than $2n-2d$ by (35). For any generator ∂_j of $H^*\big(C(A^{\pm},\theta);\mathbb{Q}\big)$, the product $\partial_j \cdot p(\partial)$ lies in the left hand side of (34) because $\phi_i(\partial_j \cdot p(\partial)) = \phi_i(\partial_j)\cdot\phi(p(\partial)) = 0$. By the maximality hypothesis in the choice of $p(\partial)$, we conclude that $\partial_j \cdot p(\partial) = 0$ in $H^*\big(C(A^{\pm},\theta);\mathbb{Q}\big)$ for all $j=1,2,\ldots,n$. Hence $p(\partial)$ lies in the socle of $H^*\big(C(A^{\pm},\theta);\mathbb{Q}\big)$. By Stanley's Theorem, this means that $p(\partial)$ has cohomological degree $2n-2d$. This is a contradiction and our claim follows.

The result (34) which we just proved translates into the following ideal-theoretic statement:

$$M(\mathcal{B}) + \mathrm{Circ}(\mathcal{A}) \quad = \quad \bigcap_{i=1}^{r} \big(I_{\Delta_i} + \mathrm{Circ}(\mathcal{A}) \big). \tag{36}$$

Since X_i is a projective orbifold, the ring $H^*(X_i;\mathbb{Q})$ is a Gorenstein ring. A result of Khovanskii and Pukhlikov [11] states that its cogenerator is the volume polynomial, i.e.

$$I_{\Delta_i} + \mathrm{Circ}(\mathcal{A}) \quad = \quad \mathrm{Ann}(V_i) \qquad \text{for } i = 1,2,\ldots,r.$$

We conclude that $M(\mathcal{B}) + \mathrm{Circ}(\mathcal{A}) \;=\; \mathrm{Ann}\big(\{V_1,\ldots,V_r\}\big)$, which proves the identity (31). $\square$

Remark. We note that the above proof of (34) is reversible, i.e. Proposition 3.4 actually implies the levelness result of Stanley [18, Theorem III.3.4] for matroids representable over $\mathbb{Q}$.

We next rewrite the result of Theorem 7.1 in terms of the other two presentations of our cohomology ring. From the perspective of the Lawrence toric variety $X(A^{\pm},\theta)$, it is most natural to work in a polynomial ring in $2n$ variables, one for each torus-invariant divisor of $X(A^{\pm},\theta)$.

Corollary 7.2 *The common cohomology ring $H^*(X(A^{\pm},\theta);\mathbb{Q})$ of the Lawrence toric variety and the toric hyperkähler variety has the presentation*

$$\mathbb{Q}[\partial_{x_1},\partial_{x_2},\ldots,\partial_{x_n},\partial_{y_1},\partial_{y_2},\ldots,\partial_{y_n}]/\mathrm{Ann}\big(V_1(x-y),\ldots,V_r(x-y)\big).$$

Proof: The polynomials $V_i(x-y) = V_i(x_1-y_1,x_2-y_2,\ldots,x_n-y_n)$ are annihilated precisely by the annihilators of $V_i(x)$ and by the extra ideal generators $\partial_{x_1}+\partial_{y_1},\ldots,\partial_{x_n}+\partial_{y_n}$. $\square$

This corollary states that the cogenerators of the Lawrence toric variety are the volume polynomials of the maximal bounded faces of the associated polyhedron $Q_\psi = P_\theta$. The same result holds for any semi-projective toric variety, even if the maximal bounded faces of its polyhedron have different dimensions. This can be proved using Proposition 3.4.
The economical presentation of our cohomology ring is as a quotient of a polynomial ring in d variables $\partial_{t_1}, \ldots, \partial_{t_d}$. The matrix A defines a surjective homomorphism of polynomial rings

$$\alpha \ : \ \mathbb{Q}[\partial_{x_1}, \ldots, \partial_{x_n}] \ \rightarrow \ \mathbb{Q}[\partial_{t_1}, \ldots, \partial_{t_d}], \quad \partial_{x_j} \ \mapsto \ \sum_{i=1}^{d} a_{ij} \partial_{t_i},$$

and a dual injective homomorphism of polynomial rings

$$\alpha^* \ : \ \mathbb{Q}[t_1, \ldots, t_d] \ \rightarrow \ \mathbb{Q}[x_1, \ldots, x_n], \quad t_i \ \mapsto \ \sum_{j=1}^{n} a_{ij} x_j.$$

The kernel of α equals $\operatorname{Circ}(\mathcal{A})$ and therefore

$$H^*\big(Y(A,\theta); \mathbb{Q}\big) \quad = \quad \mathbb{Q}[\partial_{t_1}, \ldots, \partial_{t_d}]/\alpha(M(\mathcal{B})). \tag{37}$$

We obtain cogenerators for this presentation of our cohomology ring as follows. Suppose that the indeterminate vector $t = (t_1, \ldots, t_d)$ ranges over a small neighborhood of $\theta = (\theta_1, \ldots, \theta_d)$ in $\mathbb{R}^d$. For t in this neighborhood, the polyhedron P_t remains simple and combinatorially isomorphic to P_θ. The maximal bounded faces of P_t can be identified with $\Delta_1, \ldots, \Delta_r$ as before, but now the volume of Δ_i is a homogeneous polynomial of degree $n-d$ in only d variables:

$$v_i(t) \quad = \quad v_i(t_1, \ldots, t_d) \quad = \quad \operatorname{vol}(\Delta_i) \quad \text{for } i = 1, 2, \ldots, r.$$

The polynomial $v_i(t)$ is the unique preimage of the polynomial $V_i(x)$ under the inclusion α^*.

Corollary 7.3 *The cohomology of the Lawrence toric variety and the toric hyperkähler variety equals*

$$H^*\big(Y(A,\theta); \mathbb{Q}\big) \quad = \quad \mathbb{Q}[\partial_{t_1}, \ldots, \partial_{t_d}]/\operatorname{Ann}\big(\{v_1, \ldots, v_r\}\big).$$

Proof: A differential operator $f = f(\partial_{x_1}, \ldots, \partial_{x_n})$ annihilates $\alpha^*(v)$ for some $v = v(t_1, \ldots, t_d)$ if and only if the operator $\alpha(f)$ annihilates v itself. This is the Chain Rule of Calculus. Hence

$$\begin{aligned} \operatorname{Ann}(\{v_1, \ldots, v_r\}) \quad &= \quad \alpha\big(\operatorname{Ann}(\{V_1, \ldots, V_r\})\big) \\ &= \quad \alpha\big(\operatorname{Circ}(\mathcal{A}) \, + \, M(\mathcal{B})\big) \\ &= \quad \alpha(M(\mathcal{B})). \end{aligned}$$

The claim now follows from equation (37). $\square$

Remark. Since the cohomology ring of $Y(A,\theta)$ does not depend on θ, we get the remarkable fact that the vector space generated by the volume polynomials does not depend on θ either.

We close this section by presenting an application to combinatorics. We use notation and terminology as in [18, Section III.3]. Let M be any matroid of rank $n-d$ on n elements which can be represented over the field $\mathbb{Q}$, say, by a configuration $\mathcal{B} \subset \mathbb{Z}^{n-d}$ as above, and let $h(M) = (h_0, h_1, \ldots, h_k)$ be its h-vector. A longstanding open problem is to characterize the h-vectors of matroids. For a survey see [5] or [18, Section III.3]. We wish to argue that toric hyperkähler geometry can make a valuable contribution to this problem. According to Corollary 6.6 the h-numbers of M are precisely the Betti numbers of the associated toric hyperkähler variety:

$$h_i(M) \quad = \quad \operatorname{rank} H^{2i}\big(Y(A,\theta);\mathbb{Q}\big). \tag{38}$$

As a first step, we prove the injectivity part of the Hard Lefschetz Theorem for toric hyperkähler varieties. The g-vector of the matroid is $g(M) = (g_1, g_2, \ldots, g_{\lfloor \frac{n-d}{2} \rfloor})$ where $g_i = h_i - h_{i-1}$.

Theorem 7.4 *The g-vector of a rationally represented coloop-free matroid is a Macaulay vector, i.e. there exists a graded $\mathbb{Q}$-algebra $R = R_0 \oplus R_1 \oplus \cdots \oplus R_{\lfloor \frac{n-d}{2} \rfloor}$ generated by R_1 and with $g_i = \dim_{\mathbb{Q}}(R_i)$ for all i.*

Proof: Let $[D] \in H^2(Y(A,\theta);\mathbb{Q})$ be the class of an ample divisor. The restriction $D|_{X_j}$ to any component X_j of the core is an ample divisor on the projective toric orbifold X_j. Consider the map

$$L \;:\; H^{2i-2}\big(Y(A,\theta);\mathbb{Q}\big) \;\to\; H^{2i}\big(Y(A,\theta);\mathbb{Q}\big), \tag{39}$$

given by multiplication with $[D]$. We claim that this map is injective for $i = 1, \ldots, \lfloor \frac{n-d}{2} \rfloor$. To see this, let $\alpha \in H^{2i-2}(Y(A,\theta);\mathbb{Q})$ be a nonzero cohomology class. Then according to equation (34), there exists an index $j \in \{1, 2, \ldots, r\}$ such that $\alpha|_{X_j}$ is nonzero. Then the Hard Lefschetz Theorem for the projective toric orbifold X_j implies that $\alpha|_{X_j} \cdot [D|_{X_j}]$ is a non-zero class in $H^{2i}(X_j;\mathbb{Q})$. Its preimage $\alpha \cdot [D]$ under the map ϕ_j is non-zero, and we conclude that the map (39) is injective for $2i \leq n-d$. Consider the quotient algebra $R = H^*(Y(A,\theta);\mathbb{Q})/\langle [D] \rangle$. The injectivity result just established implies that

$$g_i = h_i - h_{i-1} \;=\; \dim_{\mathbb{Q}}\big(H^{2i}(Y(A,\theta);\mathbb{Q})/\langle [D] \rangle\big) \;=\; \dim_{\mathbb{Q}}(R_i).$$

This completes the proof of Theorem 7.4. □

Remark. After the submission of our paper we learned that Swartz [20] has given a different proof of Theorem 7.4 for all coloop-free matroids. The explanation of this theorem in a combinatorial context and a comparision of the two proofs will appear in a forthcoming paper [9].

8 Toric quiver varieties

In this section we discuss an important class of toric hyperkähler manifolds, namely, Nakajima's quiver varieties in the special case when the dimension vector has all coordinates equal to one. Let $Q = (V, E)$ be a directed graph (a *quiver*) with $d+1$ vertices $V = \{v_0, v_1, \ldots, v_d\}$ and n edges $\{e_{ij} : (i,j) \in E\}$. We consider the group of all $\mathbb{Z}$-linear combinations of V whose coefficients sum to zero. We fix the basis $\{v_0 - v_1, \ldots, v_0 - v_d\}$ for this group, which is hence identified with $\mathbb{Z}^d$. We also identify $\mathbb{Z}^n$ with the group of $\mathbb{Z}$-linear combinations $\sum_{ij} \lambda_{ij} e_{ij}$ of the set of edges E. The boundary map of the quiver Q is the following homomorphism of abelian groups

$$A \;:\; \mathbb{Z}^n \to \mathbb{Z}^d, \; e_{ij} \;\mapsto\; v_i - v_j. \tag{40}$$

Throughout this section we assume that the underlying graph of Q is connected. This ensures that A is an epimorphism. The kernel of A consists of all $\mathbb{Z}$-linear combinations of E which represent cycles in Q. We fix an $n \times (n-d)$-matrix B whose columns form a basis for the cycle lattice $\ker(A)$. Thus we are in the situation of (1). The following result is well-known:

Lemma 8.1 *The matrix A representing the boundary map of a quiver Q is unimodular.*

Every edge e_{ij} of Q determines one coordinate function z_{ij} on $\mathbb{C}^n$ and two coordinate functions z_{ij}, w_{ij} on $\mathbb{H}^n$. The action of the d-torus on $\mathbb{C}^n$ and $\mathbb{H}^n$ given by the matrix A equals

$$z_{ij} \mapsto t_i t_j^{-1} \cdot z_{ij}, \quad w_{ij} \mapsto t_i^{-1} t_j \cdot w_{ij}. \tag{41}$$

We are interested in the various quotients of $\mathbb{C}^n$ and $\mathbb{H}^n$ by this action. Since the matrix A represents the quiver Q, we write $X(Q,\theta)$ instead of $X(A,\theta)$, we write $X(Q^{\pm},\theta)$ instead of $X(A^{\pm},\theta)$, and we write $Y(Q,\theta)$ instead of $Y(A,\theta)$. From Corollary 2.9 and Lemma 8.1, we conclude that all of these quotients are manifolds when the parameter vector θ is generic:

Proposition 8.2 *Let θ be a generic vector in the lattice $\mathbb{Z}^d$. Then $X(Q,\theta)$ is a smooth projective toric variety of dimension $n-d$, $X(Q^{\pm},\theta)$ is a non-compact smooth toric variety of dimension $2n-d$, and $Y(Q,\theta)$ is a smooth toric hyperkähler variety of dimension $2(n-d)$.*

We call $Y(Q,\theta)$ a *toric quiver variety*. These are precisely the quiver varieties of Nakajima [17] in the case when the dimension vector has all coordinates equal to one. Altmann and Hille [1] used the term "toric quiver variety" for the projective toric variety $X(Q,\theta)$, which is specified by an oriented quiver. Our toric quiver variety $Y(Q,\theta)$ and its ambient Lawrence toric variety $X(Q^{\pm},\theta)$ incoorporate all orientations of the quiver simultaneously. In view of Theorem 3.2, the Altmann-Hille variety $X(Q,\theta)$ is an irreducible component of the common core of $Y(Q,\theta)$ and $X(Q^{\pm},\theta)$. These manifolds and their core have the same integral cohomology ring, to be described in terms of quiver data in Theorem 8.3.

Fix a vector $\theta \in \mathbb{Z}^d$ and a subset $\tau \subseteq E$ which forms a *spanning tree* in Q. Then there exists a unique linear combination with integer coefficients λ_{ij}^{τ} which represents θ as follows:

$$\theta \quad = \quad \sum_{(i,j)\in\tau} \lambda_{ij}^{\tau} \cdot (v_i - v_j).$$

Note that the vector θ is generic if λ_{ij}^{τ} is non-zero for all spanning trees τ and all $(i,j) \in \tau$.

For every spanning tree τ, we define a subset of the monomials in $T = \mathbb{C}[z_{ij}, w_{ij}]$ as follows.

$$\sigma(\tau,\theta) := \bigl\{\, z_{ij} \,:\, (i,j) \in \tau \text{ and } \lambda_{ij}^{\tau} > 0 \,\bigr\} \,\cup\, \bigl\{\, w_{ij} \,:\, (i,j) \in \tau \text{ and } \lambda_{ij}^{\tau} < 0 \,\bigr\}.$$

Recall that a *cut* of the quiver Q is a collection D of edges which traverses a partition $(W, V\backslash W)$ of the vertex set V. We regard D as a signed set by recording the directions of its edges as follows

$$\begin{aligned} D^- &= \bigl\{\, (i,j) \in E \,:\, i \in V\backslash W \text{ and } j \in W \,\bigr\}, \\ D^+ &= \bigl\{\, (i,j) \in E \,:\, i \in W \text{ and } j \in V\backslash W \,\bigr\}. \end{aligned}$$

We now state our main result regarding toric quiver varieties:

Theorem 8.3 *Let $\theta \in \mathbb{Z}^d$ be generic. The Lawrence toric variety $X(Q^{\pm},\theta)$ is the smooth $(2n-d)$-dimensional toric variety defined by the fan whose $2n$ rays are the columns of $\Lambda(\mathcal{B}) = \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{0} & B^T \end{pmatrix}$ and whose maximal cones are indexed by the sets $\sigma(\tau,\theta)$, where τ runs over all spanning trees of Q. The toric quiver variety $Y(Q,\theta)$ is the $2(n-d)$-dimensional submanifold of $X(Q^{\pm},\theta)$ defined by the equations $\sum_{(i,j)\in D^+} z_{ij}w_{ij} = \sum_{(i,j)\in D^-} z_{ij}w_{ij}$ where D runs over all cuts of Q. The common cohomology ring of these manifolds is the quotient of $\mathbb{Z}[\partial_{ij} : (i,j) \in E]$ modulo the ideal generated by the linear forms in $\partial \cdot B$ and the monomials $\prod_{(i,j)\in D} \partial_{ij}$ where D runs over all cuts of Q.*

A few comments are in place: the variables ∂_{ij}, $(i,j) \in E$, are the coordinates of the row vector ∂, so the entries of $\partial \cdot B$ are a cycle basis for Q. The equations which cut out the toric quiver variety $Y(Q,\theta)$ lie in the Cox homogeneous coordinate ring of the Lawrence toric manifold $X(Q^{\pm},\theta)$. A more compact representation is obtained if we replace "cuts" by "cocircuits". By definition, a *cocircuit* in Q is a cut which is minimal with respect to inclusion. The proof of Theorem 8.3 follows from our general results for integer matrices A.

Corollary 6.6 shows that the Betti numbers of $Y(Q,\theta)$ are the h-numbers of the matroid of B. This is not the usual graphic matroid of Q but it is the *cographic* *matroid* associated with Q. Thus the Betti numbers of the toric quiver variety $Y(Q,\theta)$ are the h-numbers of the cographic matroid of Q. The generating function for the Betti numbers, the h-polynomial of the cographic matroid, is known in combinatorics as the *reliability polynomial* of the graph Q; see [5].

Corollary 8.4 *The Poincaré polynomial of the toric quiver variety $Y(Q,\theta)$ equals the reliability polynomial of the graph Q, which is the h-polynomial of its cographic matroid. In particular, the Euler characteristic of $Y(Q,\theta)$ coincides with the number of spanning trees of Q.*

Lopez [15] gives an explicit enumerative interpretation of the coefficients of the reliability polynomial of a graph and hence of the Betti numbers of a toric quiver variety. In particular, he proves Stanley's longstanding conjecture on h-vectors of matroid complexes [18, Conjecture III.3.6] for the special case of cographic matroids.

9 An example of a toric quiver variety

We shall describe a particular toric quiver variety $Y(K_{2,3},\theta)$ of complex dimension four. Consider the quiver in Figure 1, the complete bipartite graph $K_{2,3}$ given by $d=4$, $n=6$ and $E=\{(0,2),(0,3),(0,4),(1,2),(1,3),(1,4)\}$.

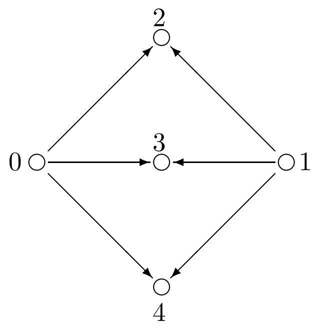


Figure 1: The quiver $K_{2,3}$

The matrix A representing the boundary map (40) is given in Figure 2. The six columns of A span the cone over a triangular prism as depicted in Figure 3. A Gale dual of this configuration is given by the six vectors in the plane in Figure 4. The rows of B^T span the cycle lattice of $K_{2,3}$.
Our manifolds are constructed algebraically from the polynomial rings

$$S = \mathbb{C}[z_{02},z_{03},z_{04},z_{12},z_{13},z_{14}]$$

and

$$A = \begin{bmatrix} 0 & 0 & 0 & -1 & -1 & -1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Figure 2: The matrix A

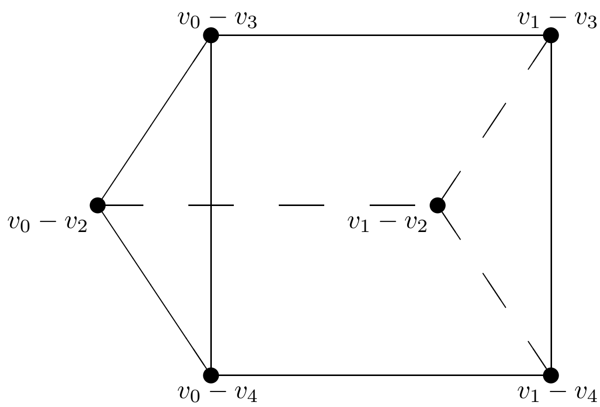


Figure 3: The column vectors of the matrix A

$$T \quad = \quad S[w_{02}, w_{03}, w_{04}, w_{12}, w_{13}, w_{14}],$$

where the degrees of the variables are given by the columns of the matrix $A^{\pm}$:

$$\text{degree}(z_{ij}) \quad = \quad -\text{degree}(w_{ij}) \quad = \quad v_i - v_j. \tag{42}$$

This grading corresponds to the torus action (41) on the polynomial rings S and T. Fix $\theta = (\theta_1, \theta_2, \theta_3, \theta_4) \in \mathbb{Z}^4$. It represents the following linear combination of vertices of $K_{2,3}$:

$$(\theta_1 + \theta_2 + \theta_3 + \theta_4)v_0 - \theta_1 v_1 - \theta_2 v_2 - \theta_3 v_3 - \theta_4 v_4$$

The monomials $z_{02}^{u_{02}} z_{03}^{u_{03}} z_{04}^{u_{04}} z_{12}^{u_{12}} z_{13}^{u_{13}} z_{14}^{u_{14}}$ in the graded component S_θ correspond to the nonnegative 2×3-integer matrices $\begin{pmatrix} u_{02} & u_{03} & u_{04} \\ u_{12} & u_{13} & u_{14} \end{pmatrix}$ with column sums $\theta_2, \theta_3, \theta_4$ and row sums $\theta_1 + \theta_2 + \theta_3 + \theta_4$ and $-\theta_1$. For instance, for $\theta = (-3, 2, 2, 2)$ there are precisely seven monomials in S_θ as shown on Figure 6. Taking "Proj" of the algebra generated by these seven monomials we get

$$B^T = \begin{bmatrix} 1 & 0 & -1 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{bmatrix}$$

Figure 4: Transpose of the matrix B

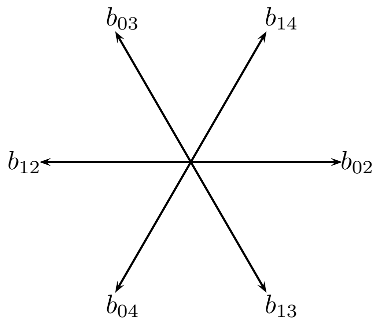


Figure 5: Rows of the matrix B

$$S_\theta = \mathbb{C}\{ \quad \begin{array}{l} z_{02}z_{03}z_{04}z_{12}z_{13}z_{14}, \\ z_{02}^2 z_{04} z_{13}^2 z_{14}, \\ z_{02}^2 z_{03} z_{13} z_{14}^2, \\ z_{02} z_{03}^2 z_{12} z_{14}^2, \\ z_{03}^2 z_{04} z_{12}^2 z_{14}, \\ z_{03} z_{04}^2 z_{12}^2 z_{13}, \\ z_{02} z_{04}^2 z_{12} z_{13}^2 \quad \} \end{array}$$

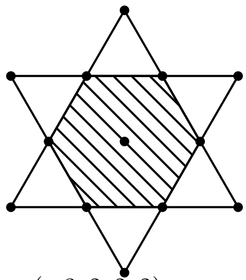


Figure 6: Monomials in multidegree $\theta = (-3, 2, 2, 2)$

a smooth toric surface $X(K_{2,3}, \theta)$ in $\mathbb{P}^6$. This surface is the blow-up of $\mathbb{P}^2$ at three points.

As θ varies, there are eighteen different types of smooth toric surfaces $X(K_{2,3}, \theta)$. They correspond to the eighteen chambers in the triangular prism, or, equivalently, to the eighteen complete fans on $\mathcal{B}$. This picture arises in the *Cremona transformation* of classical algebraic geometry, where the projective plane is blown up at three points and then the lines connecting them are blown down. The eighteen surfaces are the intermediate blow-ups and blow-downs.

We next describe the Lawrence toric varieties $X(K^{\pm}_{2,3}, \theta)$ which are the GIT quotients of $\mathbb{C}^{12}$ by the action (41). First, the (singular) affine quotient $X(K^{\pm}_{2,3}, 0)$ is the spectrum of the algebra

$$\begin{aligned} T_0 = \quad & \mathbb{C}\big[z_{02}w_{02}, z_{03}w_{03}, z_{04}w_{04}, z_{12}w_{12}, z_{13}w_{13}, z_{14}w_{14},\ z_{02}z_{13}w_{12}w_{03}, \\ & z_{03}z_{12}w_{13}w_{02}, z_{02}z_{14}w_{12}w_{04}, z_{04}z_{12}w_{14}w_{02}, z_{03}z_{14}w_{13}w_{04}, z_{04}z_{13}w_{14}w_{03}\big]. \end{aligned}$$

This is the affine toric variety whose fan is the cone over the 7-dimensional Lawrence polytope given by the matrix $\begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{0} & B^T \end{pmatrix}$, where $\mathbf{I}$ is the 6×6-unit matrix. This Lawrence polytope has 160 triangulations, all of which are regular, so there are 160 different types of smooth Lawrence toric varieties $X(K^{\pm}_{2,3}, \theta)$ as θ ranges over the generic points in $\mathbb{Z}^4$. For instance, for $\theta = (-3, 2, 2, 2)$ as in Figure 6, $X(K^{\pm}_{2,3}, \theta)$ is constructed as follows. The graded component T_θ is generated as a T_0-module by 13 monomials: the seven z-monomials in S_θ and the six additional monomials:

$$\begin{array}{c} w_{02}z_{03}^2z_{04}^2z_{12}^3,\ w_{03}z_{02}^2z_{04}^2z_{13}^3,\ w_{04}z_{02}^2z_{03}^2z_{14}^3, \\ w_{12}z_{13}^2z_{14}^2z_{02}^3,\ w_{13}z_{12}^2z_{14}^2z_{03}^3,\ w_{14}z_{12}^2z_{13}^2z_{04}^3. \end{array} \tag{43}$$

The 13 monomial generators of T_θ correspond to the 13 lattice points in the star diagram in Figure 6. The toric variety $X(K^{\pm}_{2,3}, \theta) = \mathrm{Proj}(\oplus_{n \geq 0} T_{n\theta})$ is characterized by its irrelevant ideal in the Cox homogeneous coordinate ring T, which is graded by (42). The irrelevant ideal is the radical of the monomial ideal $\langle T_\theta \rangle$. It is generated by the 12 square-free monomials obtained by erasing exponents of the monomials in (43) and Figure 6. The 7-simplices in the triangulation of the Lawrence polytope are the complements of the supports of these twelve monomials,

We finally come to the toric quiver variety $Y(K_{2,3}, \theta)$, which is smooth and four-dimensional. It is the complete intersection in the Lawrence toric variety $X(K^{\pm}_{2,3}, \theta)$ defined by the equations

$$\begin{aligned} z_{02}w_{02} + z_{03}w_{03} + z_{04}w_{04} \ &= \ z_{02}w_{02} + z_{12}w_{12} \ = \ z_{03}w_{03} + z_{13}w_{13} \ = \\ &= \ z_{04}w_{04} + z_{14}w_{14} \ = \ 0. \end{aligned}$$

These equations are valid for all 160 toric quiver varieties $Y(K_{2,3}, \theta)$. The cores of the manifolds vary greatly. For instance, for $\theta = (-3, 2, 2, 2)$, the core of $Y(K_{2,3}, \theta)$ consists of six copies of the projective plane $\mathbb{P}^2$ which are glued to the blow-up of $\mathbb{P}^2$ at three points. These correspond to the six triangles which are glued to the edges of the hexagon in Figure 6.

The common cohomology ring of the 8-dimensional Lawrence toric varieties $X(K^{\pm}_{2,3}, \theta)$ and the 4-dimensional toric quiver varieties $Y(K_{2,3}, \theta)$ is independent of θ and equals

$$\begin{aligned} &\mathbb{Z}[\partial]/\langle \partial_{03}\partial_{04}\partial_{12}, \partial_{02}\partial_{04}\partial_{13}, \partial_{02}\partial_{03}\partial_{14}, \partial_{13}\partial_{14}\partial_{02}, \partial_{12}\partial_{14}\partial_{03}, \partial_{12}\partial_{13}\partial_{04}, \partial_{02}\partial_{03}\partial_{04}, \\ &\quad \partial_{12}\partial_{13}\partial_{14}, \partial_{02}\partial_{12}, \partial_{03}\partial_{13}, \partial_{04}\partial_{14},\ \partial_{02} - \partial_{03} - \partial_{12} + \partial_{13}, \partial_{02} - \partial_{04} - \partial_{12} + \partial_{14} \rangle. \end{aligned}$$

From this presentation we can compute the Betti numbers as follows:

$$\begin{aligned} H^*(Y(K_{2,3}); \mathbb{Z}) \quad &= \quad H^0(Y(K_{2,3}); \mathbb{Z}) \oplus H^2(Y(K_{2,3}); \mathbb{Z}) \oplus H^4(Y(K_{2,3}); \mathbb{Z}) \\ &= \quad \mathbb{Z}^1 \oplus \mathbb{Z}^4 \oplus \mathbb{Z}^7. \end{aligned}$$

The 7-dimensional space of cogenerators is spanned by the areas of the six triangles in Figure 6, e.g., $V_{\{03,04,12\}}(x) = (x_{03} + x_{04} - x_{12})^2$, together with the area polynomial of the hexagon

$$\begin{aligned} V_{hex}(x) \;=\; & 2x_{03}x_{14} + 2x_{14}x_{02} + 2x_{02}x_{13} + 2x_{13}x_{04} + 2x_{04}x_{12} + 2x_{12}x_{03} - \\ & -x_{02}^2 - x_{03}^2 - x_{04}^2 - x_{12}^2 - x_{13}^2 - x_{14}^2. \end{aligned}$$

10 Which toric varieties are hyperkähler ?

Toric hyperkähler varieties are constructed algebraically as complete intersections in Lawrence toric varieties, but they are generally not toric varieties themselves. What we mean by this is that there does not exist a subtorus of the dense torus of $X(A^{\pm}, \theta)$ such that $Y(A, \theta)$ is an orbit closure of that subtorus. The objective of this section is to characterize and study the rare exceptional cases when $Y(A, \theta)$ happens to be a toric variety. We are particularly interested in the case of manifolds, when A is unimodular. The following is the main result in this section.

Theorem 10.1 *A toric manifold is a toric hyperkähler variety if and only if it is a product of ALE spaces of type A_n if and only if it is a toric quiver variety $X(Q, \theta)$ where Q is a disjoint union of cycles.*

The ALE space of type A_n is denoted $\mathbb{C}^2//\Gamma_n$ where Γ_n is the cyclic group of order n acting on $\mathbb{C}^2$ as the matrix group $\left\{ \begin{pmatrix} \eta & 0 \\ 0 & \eta^{-1} \end{pmatrix} : \eta^n = 1 \right\}$. The name "ALE space" indicates the fact that these varieties are the underlying varieties of Asymptotically Locally Euclidean gravitational instantons, or in other words 4-dimensional hyperkähler manifolds (see [13] for details).
The smooth surface $\mathbb{C}^2//\Gamma_n$ is defined as the unique crepant resolution of the 2-dimensional cyclic quotient singularity

$$\mathbb{C}^2/\Gamma_n \quad = \quad \operatorname{Spec} \mathbb{C}[x, y]^{\Gamma_n} \quad = \quad \operatorname{Spec} \mathbb{C}[x^n, xy, y^n].$$

Equivalently, we can construct $\mathbb{C}^2//\Gamma_n$ as the smooth toric surface whose fan Σ_n consists of the cones $\mathbb{R}_{\geq 0}\{(1, i-1), (1, i)\}$ for $i = 1, 2, \ldots, n$ and whose lattice is the standard lattice $\mathbb{Z}^2$.
Let us start out by showing that the ALE space $\mathbb{C}^2//\Gamma_n$ is indeed a toric quiver variety. Let C_n denote the *n-cycle*. This is the quiver with vertices $V = \{0, 1, \ldots, n-1\}$ and edges

$$E \;=\; \left\{ (0,1), (1,2), (2,3), \ldots, (n-2, n-1), (n-1, 0) \right\}.$$

We prove the following well-know result to illustrate our constructions.

Lemma 10.2 *The affine quiver variety $Y(C_n, 0)$ is isomorphic to $\mathbb{C}^2/\Gamma_n$ and for any generic vector $\theta \in \mathbb{Z}^{n-1}$, the smooth quiver variety $Y(C_n, \theta)$ is isomorphic to the ALE space $\mathbb{C}^2//\Gamma_n$.*

Proof: The boundary map of the n-cycle C_n has the format $\mathbb{Z}^n \to \mathbb{Z}^{n-1}$ and looks like

$$A \quad = \quad \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix}.$$

and its Gale dual is the $1 \times n$-matrix with all entries equal to one:

$$B^T \quad = \quad (\,1 \quad 1 \quad 1 \quad \cdots \quad 1\,). \tag{44}$$

The torus $\mathbb{T}_{\mathbb{C}}^{n-1}$ acts via $A^{\pm}$ on the polynomial ring $T = \mathbb{C}\,[\,z_{i,i+1}, w_{i,i+1} : i = 0, \ldots, n-1\,]$. The affine Lawrence toric variety $X(C_n^{\pm}, 0) = \mathbb{C}^{2n} /\!/_0 \mathbb{T}_{\mathbb{C}}^{n-1}$ is the spectrum of the invariant ring

$$T_0 \quad = \quad \mathbb{C}\,\big[\,z_{01}w_{01}\,,\,\ldots,\,z_{n-1,0}w_{n-1,0}\,,\; z_{01}z_{12}\cdots z_{n-1,0}\,,\; w_{01}w_{12}\cdots w_{n-1,0}\,\big].$$

The common defining ideal of all the quiver varieties $Y(C_n, \theta)$ is the following ideal in T:

$$\mathrm{Circ}(\mathcal{B}) \quad = \quad \langle\, z_{i-1,i}w_{i-1,i} \,-\, z_{i,i+1}w_{i,i+1} \;:\; i = 1, 2, \ldots, n\,\rangle.$$

All indices are considered modulo n. The quiver variety $Y(C_n, 0)$ is the spectrum of $T_0/(T_0 \cap \mathrm{Circ}(\mathcal{B}))$. Dividing T_0 by $T_0 \cap \mathrm{Circ}(\mathcal{B})$ means erasing the double indices of all variables:

$$T_0/(T_0 \cap \mathrm{Circ}(\mathcal{B})) \quad \simeq \quad \mathbb{C}[\,zw,\, z^n,\, w^n\,].$$

Passing to the spectra of these rings proves our first assertion: $Y(C_n, 0) \simeq \mathbb{C}^2/\Gamma_n$.

For the second assertion, we first note that $\theta = (\theta_1, \ldots, \theta_{n-1})$ is generic for $\mathcal{A}^{\pm}$ if and only if all consecutive coordinate sums $\theta_i + \theta_{i+1} + \cdots + \theta_j$ are non-zero. The associated hyperplane arrangement $\Gamma(\mathcal{A})$ is linearly isomorphic to the braid arrangement $\{u_i = u_j\}$. It has $n\,!$ chambers, and the symmetric group acts transitively on the chambers. Hence it suffices to prove $Y(C_n, \theta) \simeq \mathbb{C}^2/\!/\Gamma_n$ for only one vector θ which lies in the interior of any chamber.

We fix the generic vector $\theta = (1, 1, \ldots, 1)$. There are n monomials of degree θ in T, namely,

$$\prod_{j=1}^{i-1} z_{j-1,j}^{i-j} \cdot \prod_{k=i+1}^{n} w_{k-1,k}^{k-i} \qquad\qquad \text{for } i = 1, 2, \ldots, n. \tag{45}$$

The images of these monomials are minimal generators of the $T_0/(T_0 \cap \mathrm{Circ}(\mathcal{B}))$-algebra

$$\bigoplus_{r=0}^{\infty} T_{r\theta}/(T_{r\theta} \cap \mathrm{Circ}(\mathcal{A})(\mathcal{B})).$$

By definition, $Y(C_n,\theta)$ is the projective spectrum of this $\mathbb{N}$-graded algebra. Applying our isomorphism "erasing double indices", the images of our n monomials in (45) translate into

$$z^{\binom{i}{2}} \cdot w^{\binom{n-i+1}{2}} \qquad \text{for } i = 1, 2, \ldots, n. \tag{46}$$

Hence $Y(C_n,\theta)$ is the projective spectrum of the $\mathbb{C}[zw, z^n, w^n]$-algebra generated by (46). It is straightforward to see that this is the toric surface with fan Σ_n, i.e. the ALE space $\mathbb{C}^2//\Gamma_n$. □

It is instructive to write down our presentations for the cohomology ring of the ALE space $Y(C_n,\theta) = \mathbb{C}^2//\Gamma_n$. The circuit ideal of the n-cycle is the principal ideal

$$\mathrm{Circ}(\mathcal{A}) \quad = \quad \langle\, \partial_{01} + \partial_{12} + \partial_{23} + \cdots + \partial_{n-1,0}\,\rangle.$$

The matroid ideal $M(\mathcal{B})$ is generated by all quadratic squarefree monomials in $\mathbb{Z}[\partial]$. It follows that $\mathbb{Z}[\partial]/(\mathrm{Circ}(\mathcal{A}) + M(\mathcal{B}))$ is isomorphic to a polynomial ring in $n-1$ variables modulo the square of the maximal ideal generated by the variables, and hence

$$H^*(Y(C_n,\theta);\mathbb{Z}) \quad = \quad H^0(Y(C_n,\theta);\mathbb{Z}) \oplus H^2(Y(C_n,\theta);\mathbb{Z}) \quad \simeq \quad \mathbb{Z}^1 \oplus \mathbb{Z}^{n-1}.$$

On our way towards proving Theorem 10.1, let us now fix an epimorphism $A : \mathbb{Z}^n \to \mathbb{Z}^d$ and a generic vector $\theta \in \mathbb{Z}^d$. We assume that $\mathcal{A}$ is not a cone, i.e. the zero vector is not in $\mathcal{B}$. We do not assume that A is unimodular. By a *binomial* we mean a polynomial with two terms.

Proposition 10.3 *The following three statements are equivalent:*

(a) The hyperkähler toric variety $Y(A,\theta)$ is a toric subvariety of $X(A^{\pm},\theta)$.

(b) The ideal $\mathrm{Circ}(\mathcal{B})$ is generated by binomials.

(c) The configuration $\mathcal{B}$ lies on $n-d$ linearly independent lines through the origin in $\mathbb{R}^{n-d}$.

Proof: The condition (b) holds if and only if the matrix A can be chosen to have two nonzero entries in each row. This defines a graph $\mathcal{G}$ on $\{1,2,\ldots,n\}$, namely, j and k are connected by an edge if there exists $i \in \{1,\ldots,d\}$ such that $a_{ij} \neq 0$ and $a_{ik} \neq 0$. The graph $\mathcal{G}$ is a disjoint union of $n-d$ trees. Two indices j and k lie in the same connected component of $\mathcal{G}$ if and only if the vectors b_j and b_k are linearly dependent. Thus (b) is equivalent to (c).
Suppose that (b) holds. Then the prime ideal $\mathrm{Circ}(\mathcal{B})$ is generated by the quadratic binomials $a_{ij}z_jw_j + a_{ik}z_kw_k$ indexed by the edges (j,k) of $\mathcal{G}$. The corresponding coefficient-free equations

$$z_jw_j \quad = \quad z_kw_k \qquad \text{for } (j,k) \in \mathcal{G}.$$

define a subtorus $\mathbb{T}$ of the dense torus of the Lawrence toric variety $X(A^{\pm},\theta)$, and the equations

$$a_{ij}z_jw_j \; + \; a_{ik}z_kw_k \quad = \quad 0 \qquad \text{for } (j,k)\in\mathcal{G}.$$

define an orbit of $\mathbb{T}$ in the dense torus of $X(A^{\pm},\theta)$. The solution set of the same equations in $X(A^{\pm},\theta)$ has the closure of that $\mathbb{T}$-orbit as one of its irreducible components. But that solution set is our hyperkähler variety $Y(A,\theta)$. Since $Y(A,\theta)$ is irreducible, we can conclude that it coincides with the closure of the $\mathbb{T}$-orbit. Hence $Y(A,\theta)$ is a toric variety, i.e. (a) holds.
For the converse, suppose that (a) holds. The irreducible subvariety $Y(A,\theta)$ is defined by a homogeneous prime ideal J in the homogeneous coordinate ring T of $X(A^{\pm},\theta)$. Since $Y(A,\theta)$ is a torus orbit closure, the ideal J is generated by binomials. The ideal $\mathrm{Circ}(\mathcal{B})$ has the same zero set as J does, and therefore, by the Nullstellensatz and results of Cox, $\mathrm{rad}\big(\mathrm{Circ}(\mathcal{B}) : B_\theta^\infty\big) = J$. Our hypothesis $0\notin\mathcal{B}$ ensures that $\mathrm{Circ}(\mathcal{B})$ itself is a prime ideal, and therefore we conclude $\mathrm{Circ}(\mathcal{B}) = J$. In particular, this ideal is generated by binomials, i.e. (b) holds. □

Proof of Theorem 10.1: Suppose that Q is a quiver with connected components $Q_1,\ldots,Q_r$. Then its boundary map is given by a matrix with block decomposition

$$A \quad = \quad A_1 \oplus A_2 \oplus \cdots \oplus A_r, \tag{47}$$

where A_i is the boundary map of Q_i. There is a corresponding decomposition of the Gale dual

$$B \quad = \quad B_1 \oplus B_2 \oplus \cdots \oplus B_r. \tag{48}$$

In this situation, the toric hyperkähler variety $Y(A,\theta)$ is the direct product of the toric hyperkähler varieties $Y(A_i,\theta)$ for $i=1,\ldots,r$. For our quiver Q this means

$$Y(Q,\theta) \quad = \quad Y(Q_1,\theta)\times Y(Q_2,\theta)\times\cdots\times Y(Q_r,\theta).$$

Using Lemma 10.2, we conclude that a manifold is a product of ALE spaces of type A_n if and only if it is a toric quiver variety $Y(Q,\theta)$ where Q is a disjoint union of cycles C_{n_i}.
The matrix A in (47) is unimodular if and only if the matrices $A_1,\ldots,A_r$ are unimodular. Hence a product of toric hyperkähler manifolds is a toric hyperkähler manifold. In particular, a product of ALE spaces $\mathbb{C}^2/\Gamma_{n_i}$ is a toric hyperkähler manifold which is also a toric variety.
For the converse, suppose that $Y(A,\theta)$ is a toric hyperkähler manifold which is also a toric variety, so that statement (a) in Proposition 10.3 holds. Statement (c) in Proposition 10.3 says that the matrix B has a decomposition (48) where $r=n-d$ and each B_i is a matrix with exactly one column. We may assume that none of the entries in B_i is zero. The Gale dual A_i of B_i is a unimodular matrix, and hence B_i is unimodular. For a matrix with one column this means that all entries in B_i are either $+1$ or -1. After trivial sign changes, this means $B_i^T = (1\,1\,\ldots\,1)$. Now we are in the situation of (44), which means that $Y(A_i,\theta)$ is an ALE space $\mathbb{C}^2/\Gamma_{n_i}$. □

References

[1] K. Altmann and L. Hille: Strong exceptional sequences provided by quivers, *Algebras and Representation Theory* 2 (1999) 1–17.

[2] A. Bialynicki-Birula: Some properties of the decompositions of algebraic varieties determined by actions of a torus. *Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys.* 24 (1976), no. 9, 667-674.

[3] R. Bielawski and A. Dancer: The geometry and topology of toric hyperkähler manifolds, *Comm. Anal. Geom.* 8 (2000), no. 4, 727–760.

[4] A. Björner, M. Las Vergnas, B. Sturmfels, N. White and G. Ziegler: Oriented Matroids, Cambridge University Press, 1999.

[5] M. Chari and C. Colbourn: Reliability polynomials: a survey. *J. Combin. Inform. System Sci.* 22 (1997) 177–193.

[6] D. Cox: The homogeneous coordinate ring of a toric variety, *Journal of Algebraic Geometry* 4 (1995) 17–50.

[7] D. Cox: Recent developments in toric geometry. *Algebraic Geometry—Santa Cruz 1995*, 389–436, Proc. Sympos. Pure Math., 62, Part 2, Amer. Math. Soc., Providence, RI, 1997.

[8] W. Fulton: *Introduction to Toric Varieties*, Princeton University Press, 1993.

[9] T. Hausel: Quaternionic geometry of matroids, *in preparation.*

[10] N.J. Hitchin, A. Karlhede, U. Lindström and M. Roček: Hyper-Kähler metrics and supersymmetry, *Comm. Math. Phys.*, 108,1987,(4), 535–589.

[11] A.G. Khovanskii and A.V. Pukhlikov: The Riemann-Roch theorem for integrals and sums of quasipolynomials on virtual polytopes. *St. Petersburg Math. J.* 4 (1993) 789–812.

[12] F.C. Kirwan, *Cohomology of Quotients in Symplectic and Algebraic Geometry*, Princeton University Press, Princeton, NJ, 1984.

[13] P.B. Kronheimer, The construction of ALE spaces as hyperkähler quotients, *J. Diff. Geom.* 29 (1989) no. 3, 665–683.

[14] H. Konno: Cohomology rings of toric hyperkähler manifolds, *Internat. J. Math.* 11 (2000), no. 8, 1001–1026.

[15] C.M. Lopez: Chip firing and the Tutte polynomial, *Ann. Combinatorics* 1 (1997) 253–259.

[16] J. Milnor: *Morse Theory*, Princeton University Press, Princeton, NJ, 1969.

[17] H. Nakajima: Quiver varieties and Kac-Moody algebras, *Duke Mathematical Journal* 91 (1998) 515–560.

[18] R.P. Stanley: *Combinatorics and Commutative Algebra*, 2nd ed., Birkhäuser Boston, 1996.

[19] R.P. Stanley: A monotonicity property of h-vectors and h^*-vectors, *European Journal of Combinatorics*, 14, No. 3., May 1993, 251-258.

[20] E.Swartz: g-elements of matroid complexes, *Journal of Combinatorial Theory, Series B*, to appear.

[21] T. Zaslavsky: Facing up to arrangements: face-count formulas for partitions of space by hyperplanes. *Mem. Amer. Math. Soc.* 1 (1975), issue 1, no. 154.

[22] G. Ziegler: *Lectures on Polytopes*, Springer Graduate Texts in Mathematics, 1995.

Tamás Hausel
Department of Mathematics
University of Texas at
Austin TX 78712, USA
hausel@math.utexas.edu

Bernd Sturmfels
Department of Mathematics
University of California at
Berkeley CA 94720, USA
bernd@math.berkeley.edu

The Connection Between May's Axioms for a Triangulated Tensor Product and Happel's Description of the Derived Category of the Quiver D_4

Bernhard Keller and Amnon Neeman

Received: October 21, 2002

Communicated by Peter Schneider

Abstract. In an important recent paper [12], May gave an axiomatic description of the properties of triangulated categories with a symmetric tensor product. The main point of the current article is that there are two other results in the literature which can be used to shed considerable light on May's work. The first is a construction of Verdier's, which appeared in Beilinson, Bernstein and Deligne's [4, Prop. 1.1.11, pp. 24-25]. The second and more important is the beautiful work of Happel, in [9], which can be used to better organise May's axioms.

Keywords and Phrases: derived category, tensor product, quiver

1. Introduction

We should begin with a disclaimer. This article definitely *does not* attempt to give the definitive axiomatic description of tensor products in triangulated categories. In the opinion of the authors, the subject is not ripe for such a treatment. It is only very recently that there has been any real interest in the field. The subject is still at a very formative stage. Time will tell which properties of the tensor product really matter.

Let $\mathcal{T}$ be a triangulated category, and assume it has a (symmetric) tensor product. For example, $\mathcal{T}$ might be the derived category of a commutative ring R, or the homotopy category of spectra. It becomes interesting to know what are the "natural" properties that this tensor product has. In a lovely recent article [12], May made giant steps towards answering this question.

Some properties are obvious, and we do not repeat them here. The interest lies in the following. Given two distinguished triangles

$$x \longrightarrow y \longrightarrow z \longrightarrow \Sigma x$$

$$x' \longrightarrow y' \longrightarrow z' \longrightarrow \Sigma x'$$

one can form the tensor product

$$\begin{tikzcd}
x \otimes x' \arrow[r] \arrow[d] & y \otimes x' \arrow[r] \arrow[d] & z \otimes x' \arrow[r] \arrow[d] & \Sigma x \otimes x' \arrow[d] \\
x \otimes y' \arrow[r] \arrow[d] & y \otimes y' \arrow[r] \arrow[d] & z \otimes y' \arrow[r] \arrow[d] & \Sigma x \otimes y' \arrow[d] \\
x \otimes z' \arrow[r] \arrow[d] & y \otimes z' \arrow[r] \arrow[d] & z \otimes z' \arrow[r] \arrow[d] \arrow[dr, phantom, "(-)"] & \Sigma x \otimes z' \arrow[d] \\
\Sigma x \otimes x' \arrow[r] & \Sigma y \otimes x' \arrow[r] & \Sigma z \otimes x' \arrow[r] & \Sigma^2 x \otimes x'
\end{tikzcd}$$

It is natural to assume that the rows and columns are distinguished triangles. The question is what, if any, are the other reasonable properties one could postulate. It turns out that, at least for reasonable examples of triangulated categories $\mathcal{T}$ with tensor products, in this diagram the diagonal arrows

$$\begin{tikzcd}
x \otimes x' \arrow[r] \arrow[d] & y \otimes x' \arrow[r] \arrow[d] & z \otimes x' \arrow[r] \arrow[d] \arrow[dr] & \Sigma x \otimes x' \arrow[d] \\
x \otimes y' \arrow[r] \arrow[d] & y \otimes y' \arrow[r] \arrow[d] \arrow[dr] & z \otimes y' \arrow[r] \arrow[d] & \Sigma x \otimes y' \arrow[d] \\
x \otimes z' \arrow[r] \arrow[d] \arrow[dr] & y \otimes z' \arrow[r] \arrow[d] & z \otimes z' \arrow[r] \arrow[d] \arrow[dr, phantom, "(-)"] & \Sigma x \otimes z' \arrow[d] \\
\Sigma x \otimes x' \arrow[r] & \Sigma y \otimes x' \arrow[r] & \Sigma z \otimes x' \arrow[r] & \Sigma^2 x \otimes x'
\end{tikzcd}$$

all have a common mapping cone. May's axiom (TC3) describes very well the various diagrams involving this common mapping cone.

Of course, we could look at other diagonal arrows. In the diagram below

$$\begin{tikzcd}
x \otimes x' \arrow[r] \arrow[d] & y \otimes x' \arrow[r] \arrow[d] \arrow[dr, squiggly] & z \otimes x' \arrow[r] \arrow[d] \arrow[dr] & \Sigma x \otimes x' \arrow[d] \\
x \otimes y' \arrow[r] \arrow[d] \arrow[dr, squiggly] & y \otimes y' \arrow[r] \arrow[d] \arrow[dr] & z \otimes y' \arrow[r] \arrow[d] \arrow[dr, dashed] & \Sigma x \otimes y' \arrow[d] \\
x \otimes z' \arrow[r] \arrow[d] \arrow[dr] & y \otimes z' \arrow[r] \arrow[d] \arrow[dr, dashed] & z \otimes z' \arrow[r] \arrow[d] \arrow[dr, phantom, "(-)"] & \Sigma x \otimes z' \arrow[d] \\
\Sigma x \otimes x' \arrow[r] & \Sigma y \otimes x' \arrow[r] & \Sigma z \otimes x' \arrow[r] & \Sigma^2 x \otimes x'
\end{tikzcd}$$

the squiggly arrows have a common mapping cone Σu, the straight arrows a common mapping cone Σv and the broken arrows a common mapping cone Σw. It becomes interesting to describe what relations there should be among u, v and w. It turns out that there are many. May's axioms include one such relation.

In this paper, we will see that May's results are related to earlier work by Verdier and by Happel. We will show that the older approaches lead to new insights; in terms of the above, in general, they lead to infinitely many relations among u, v and w which May missed. We will see that the work of Happel is particularly illuminating.

As we have already said, we do not see this as an attempt to give the definitive foundational treatment. The subject is very young and active. Aside from May's paper there is the totally unrelated work by Balmer [3], and recent talks by Gaitsgory (no manuscript yet) show that his work is also related. At this point, all we want is to advertise widely the fact that the results of Verdier, Happel and May (in chronological order) are related.

Of course, we must also persuade the reader that this relation, among three existing articles in the literature, is interesting. Of most interest is how Happel's work in [9] leads to a better organisation of the theory. To illustrate this, we give examples of new results that can be obtained. We make no attempt to prove the best possible versions of these new results. That is not the point. We settle for weaker-than-optimal statements of our new results, to make transparent how they can be viewed as consequences of Happel's work.

Since we want the article to be accessible to a wide audience, we try not to assume much background knowledge. The experts in representations of quivers will undoubtedly find Section 5 painfully slow and detailed. The experts in topology will undoubtedly wonder why we assume the reader may never have heard of closed model categories. The guiding policy in writing this article was that the presentation should be as free of prerequisites as possible. The unfortunate side effect is that it adds to the length of the article. We ask the experts to be patient with us.

The structure of the article is as follows. Section 2 sets some notation. Section 3 establishes the relation among the three approaches. Sections 4 and 5 apply these to obtain new identities. The main results are

(i) Axiom (TC3) of May produces exactly the same diagram as Verdier found in [4]. (See Theorem 3.5). Haynes Miller noticed this independently.
(ii) The special case of $D^b(\square)$ is universal (Theorem 3.10).
(iii) May's axiom (TC4) follows from (TC3) and the octahedral axiom (Theorem 4.1).
(iv) There is an equivalence of categories $D^b(\square) = D^b(\mathsf{Y})$, where $D^b(\mathsf{Y})$ is the bounded derived category of the category of representations of the quiver D_4. Happel studied this in the special case where the categories are all linear over a field k. In the case of k-linear categories we can therefore glean a great deal of information from Happel's work (Section 5).

2. Notation for the octahedral axiom

An octahedron can be thought of as two pyramids glued together along their square bases. There are three planes along which we can split the octahedron into two pyramids. This gives three squares. For octahedra as in the octahedral axiom, each edge is a morphism in a triangulated category and has a direction. The four squares in an octahedron have arrows as follows

$$
\begin{tikzcd}
{} \arrow[r] \arrow[dr, phantom, "(1)"] & {} \arrow[d] \\
{} \arrow[u] & {} \arrow[l]
\end{tikzcd}
\qquad
\begin{tikzcd}
{} \arrow[r] \arrow[d] \arrow[dr, phantom, "(2)"] & {} \arrow[d] \\
{} \arrow[r] & {}
\end{tikzcd}
\qquad
\begin{tikzcd}
{} \arrow[r] \arrow[d] \arrow[dr, phantom, "(3)"] & {} \arrow[d] \\
{} \arrow[r] & {}
\end{tikzcd}
$$

The convention we adopt is to always write the octahedron as a union of two pyramids, split along the unique square where the arrows cycle around as in (1) above. The octahedron splits into a "top pyramid" and a "bottom pyramid" (of course, it is somewhat arbitrary which pyramid is declared to be "top" and which "bottom").

If we project the top and bottom pyramids to their common base plane, we get diagrams

$$
\begin{tikzcd}
& a \arrow[d] \arrow[dr] & \\
d \arrow[ur] & x \arrow[l] \arrow[r] & b \arrow[dl] \\
& c \arrow[u] \arrow[ul] &
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
& a \arrow[dr] & \\
d \arrow[ur] \arrow[r] & y \arrow[u] \arrow[d] & b \arrow[l] \arrow[dl] \\
& c \arrow[ul] &
\end{tikzcd}
$$

It turns out to be very convenient to twist and torture the octahedron. We wish to switch the positions of c and d. The pyramids become

$$\begin{tikzcd}
 & a \arrow[d] \arrow[dr] & \\
c \arrow[r] \arrow[dr] & x \arrow[r] \arrow[d] & b \\
 & d &
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
 & a \arrow[dr] & \\
c \arrow[r] \arrow[dr] & y \arrow[l] \arrow[u] & b \arrow[l] \\
 & d \arrow[u] &
\end{tikzcd}$$

Of course, it now takes some imagination to see that these are pyramids. There are still four triangles to each pyramid; but two of them project to straight lines (the horizontal and vertical lines). We will frequently write our octahedra in this contorted form.

The octahedral axiom tells us that, in a triangulated category, certain diagrams can be completed to octahedra. The refined octahedral axiom tells us that the two commutative squares, which arise from the two "other" planes splitting the octahedron into pyramids, are homotopy pushouts. In the notation above, there are canonical distinguished triangles

$$x \longrightarrow b \oplus d \longrightarrow y \longrightarrow \Sigma x$$
$$y \longrightarrow a \oplus c \longrightarrow x \longrightarrow \Sigma y.$$

There is a choice of sign here which we do not wish to make explicit, and some of the morphisms are of degree 1. The important thing is that the maps $y \longrightarrow \Sigma x$ and $x \longrightarrow \Sigma y$ are very explicitly given by the octahedron. We will refer to them as the *differentials of the squares.*

3. The relation among the approaches

Suppose $\mathcal{T}$ is a triangulated category with a tensor product. We wish to study when this tensor product is well-behaved. To this end, we make a definition.

DEFINITION 3.1. *We say that the tensor product on $\mathcal{T}$ is* decent *if the following holds.*

(i) *There exists an abelian category $\mathcal{A}$ and a triangulated functor $F : D^b(\mathcal{A}) \longrightarrow \mathcal{T}$. [Here $D^b(\mathcal{A})$ means the bounded derived category of $\mathcal{A}$.]*
(ii) *The category $\mathcal{A}$ has a natural tensor product.*
(iii) *The category $\mathcal{A}$ comes with a collection of* special short exact sequences. *These form a subclass of all the short exact sequences.*
(iv) *For any special short exact sequence*

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

and any $X \in \mathcal{A}$, the two sequences

$$0 \longrightarrow X \otimes A \longrightarrow X \otimes B \longrightarrow X \otimes C \longrightarrow 0$$

$$0 \longrightarrow A \otimes X \longrightarrow B \otimes X \longrightarrow C \otimes X \longrightarrow 0$$

are both exact.

(v) *Any distinguished triangle in $\mathcal{T}$ is isomorphic to*

$$F(A) \longrightarrow F(B) \longrightarrow F(C) \longrightarrow \Sigma F(A)$$

for some special short exact sequence in $\mathcal{A}$

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0.$$

(vi) *With the notation as in (iv) and (v) above, the two triangles*

$$F(X \otimes A) \longrightarrow F(X \otimes B) \longrightarrow F(X \otimes C) \longrightarrow \Sigma F(X \otimes A)$$

$$F(A \otimes X) \longrightarrow F(B \otimes X) \longrightarrow F(C \otimes X) \longrightarrow \Sigma F(A \otimes X)$$

are canonically independent of the choice of the special short exact sequence

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

lifting the triangle

$$F(A) \longrightarrow F(B) \longrightarrow F(C) \longrightarrow \Sigma F(A).$$

(vii) *Suppose X and Y are objects of $\mathcal{A}$, which occur in some special short exact sequences. Then there is a canonical isomorphism*

$$F(X) \otimes F(Y) \quad = \quad F(X \otimes Y).$$

Example 3.2. Let R be a commutative ring and let $\mathcal{T} = D^-(R)$, the derived category of bounded-above chain complexes of R–modules. The tensor product on $\mathcal{T}$ is the derived tensor product. The category $\mathcal{A}$ is defined to be the abelian category of bounded-above chain complexes of R–modules, with the obvious tensor product. The special short exact sequences are the short exact sequences of bounded-above chain complexes of projectives.

Roughly speaking, the idea in May's article [12] is to study decent tensor products in triangulated categories. The existence of the abelian category $\mathcal{A}$ and $F : D^b(\mathcal{A}) \longrightarrow \mathcal{T}$ has many consequences, allowing us to create complicated diagrams in $\mathcal{T}$. What May does is postulate the existence of the diagrams in $\mathcal{T}$ as axioms for the tensor product, even in the absence of any explicit $F : D^b(\mathcal{A}) \longrightarrow \mathcal{T}$.

Remark 3.3. What we said above is slightly inaccurate. May handles a more general framework. Instead of a functor $F : D^b(\mathcal{A}) \longrightarrow \mathcal{T}$, he assumes only that $\mathcal{T}$ has a closed model structure with a compatible tensor product. Since we do not want to assume the reader knows what a closed model is, we have allowed ourselves to restrict to the simplified situation.

Suppose $\mathcal{T}$ is a triangulated category with a decent tensor product. Suppose we are given two distinguished triangles in $\mathcal{T}$. By Definition 3.1(v) these two triangles are the images under F of two special short exact sequences

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

$$0 \longrightarrow A' \longrightarrow B' \longrightarrow C' \longrightarrow 0$$

The tensor product in $\mathcal{A}$ gives a 3×3 diagram with exact rows and columns

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & A \otimes A' \arrow[r] \arrow[d] & B \otimes A' \arrow[r] \arrow[d] & C \otimes A' \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & A \otimes B' \arrow[r] \arrow[d] & B \otimes B' \arrow[r] \arrow[d] & C \otimes B' \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & A \otimes C' \arrow[r] \arrow[d] & B \otimes C' \arrow[r] \arrow[d] & C \otimes C' \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

The central idea in May's [12] is to write down all the distinguished triangles one can deduce from this diagram. It might be simplest to focus on one of May's results. From now until the end of the section, we consider (TC3).

Let X be the quotient of the injective map $A \otimes A' \longrightarrow B \otimes B'$. Then we have two diagrams with exact rows and columns

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & A \otimes A' \arrow[r] \arrow[d, "1"'] & B \otimes A' \arrow[r] \arrow[d] & C \otimes A' \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & A \otimes A' \arrow[r] \arrow[d] & B \otimes B' \arrow[r] \arrow[d] & X \arrow[r] \arrow[d] & 0 \\
 & 0 \arrow[r] & B \otimes C' \arrow[r, "1"] \arrow[d] & B \otimes C' \arrow[r] \arrow[d] & 0 \\
 & & 0 & 0 &
\end{tikzcd}$$

and

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & & \\
0 \arrow[r] & A\otimes A' \arrow[r, "1"] \arrow[d] & A\otimes A' \arrow[r] \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & A\otimes B' \arrow[r] \arrow[d] & B\otimes B' \arrow[r] \arrow[d] & C\otimes B' \arrow[r] \arrow[d, "1"] & 0 \\
0 \arrow[r] & A\otimes C' \arrow[r] \arrow[d] & X \arrow[r] \arrow[d] & C\otimes B' \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

The right column of the first of these diagrams and the bottom row of the second exhibit $F(X)$ as the mapping cone on two more maps, namely

$$\Sigma^{-1}F(B\otimes C') \longrightarrow F(C\otimes A')$$

and

$$\Sigma^{-1}F(C\otimes B') \longrightarrow F(A\otimes C').$$

In other words, the diagonal arrows in the following diagram

$$\begin{tikzcd}
 & & \Sigma^{-1}F(B\otimes C') \arrow[r] \arrow[d] \arrow[dr] & \Sigma^{-1}F(C\otimes C') \arrow[d] \\
 & F(A\otimes A') \arrow[r] \arrow[d] \arrow[dr] & F(B\otimes A') \arrow[r] \arrow[d] & F(C\otimes A') \arrow[d] \\
\Sigma^{-1}F(C\otimes B') \arrow[r] \arrow[d] \arrow[dr] & F(A\otimes B') \arrow[r] \arrow[d] & F(B\otimes B') \arrow[r] \arrow[d] & F(C\otimes B') \arrow[d] \\
\Sigma^{-1}F(C\otimes C') \arrow[r] & F(A\otimes C') \arrow[r] & F(B\otimes C') \arrow[r] & F(C\otimes C')
\end{tikzcd}$$

all have the common mapping cone $F(X)$. Playing around a little gives yet more distinguished triangles. Instead of following May's approach (which the reader can find in [12]), let us see how Verdier found them.

As we have explained, May's approach is based on beginning with a 3×3 diagram in an abelian category $\mathcal{A}$, with exact rows and columns, and reading off induced triangles in the derived category of $\mathcal{A}$. To simplify the notation, let us forget that the diagram arose from a tensor product of two short exact sequences in $\mathcal{A}$. We have a diagram in $\mathcal{A}$

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & x \arrow[r] \arrow[d] & y \arrow[r] \arrow[d] & z \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & x' \arrow[r] \arrow[d] & y' \arrow[r] \arrow[d] & z' \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & x'' \arrow[r] \arrow[d] & y'' \arrow[r] \arrow[d] & z'' \arrow[r] \arrow[d] & 0 \\
 & 0 & 0 & 0 &
\end{tikzcd}$$

with exact rows and columns. It is well known that, for a diagram with exact rows and columns

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & & \\
0 \arrow[r] & x \arrow[r] \arrow[d] & y \arrow[r] \arrow[d] & z \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & x' \arrow[r] \arrow[d] & y' \arrow[r] \arrow[d] & z' \arrow[r] & 0 \\
 & x'' \arrow[r] \arrow[d] & y'' \arrow[d] & & \\
 & 0 & 0 & &
\end{tikzcd}$$

one has:

Lemma 3.4. *the following assertions are equivalent:*

(i) *The map $z \longrightarrow z'$ is a monomorphism.*
(ii) *The map $x'' \longrightarrow y''$ is a monomorphism.*
(iii) *The map from the pushout of*

$$\begin{tikzcd}
x \arrow[r] \arrow[d] & y \\
x' &
\end{tikzcd}$$

to y' is a monomorphism.

The entire 3×3 diagram is, up to canonical isomorphism, entirely determined by the commutative square of monomorphisms

$$\begin{tikzcd} x \arrow[r] \arrow[d] & y \arrow[d] \\ x' \arrow[r] & y' \end{tikzcd}$$

satisfying the condition in Lemma 3.4(iii). Verdier's idea was to build the entire diagram using repeated applications of the octahedral axiom. We remind the reader.
Our commutative square may be viewed as two commutative triangles

$$\begin{tikzcd} x \arrow[r] \arrow[d] \arrow[dr] & y \arrow[d] \\ x' \arrow[r] & y' \end{tikzcd}$$

The two commutative triangles may be completed to two octahedra. In the twisted notation of Section 2, the top pyramids of the octahedra may be written as

$$\begin{tikzcd} & x \arrow[d] \arrow[dr] & \\ \Sigma^{-1}z' \arrow[r] \arrow[dr] & x' \arrow[r] \arrow[d] & y' \\ & x'' & \end{tikzcd} \qquad \begin{tikzcd} & \Sigma^{-1}y'' \arrow[d] \arrow[dr] & \\ x \arrow[r] \arrow[dr] & y \arrow[r] \arrow[d] & z \\ & y' & \end{tikzcd}$$

We remind the reader how this should be read. The horizontal and vertical lines are projections of distinguished triangles. We have, so far, four distinguished triangles. They are the rows and columns below

$$
\begin{tikzcd}
 & & \Sigma^{-1}y'' \arrow[d] & & \\
 & x \arrow[r] \arrow[d] & y \arrow[r] \arrow[d] & z \arrow[r] & \Sigma x \\
\Sigma^{-1}z' \arrow[r] & x' \arrow[r] \arrow[d] & y' \arrow[r] \arrow[d] & z' & \\
 & x'' \arrow[d] & y'' & & \\
 & \Sigma x & & &
\end{tikzcd}
$$

We also have four commutative triangles, as below

$$
\begin{tikzcd}
 & & \Sigma^{-1}y'' \arrow[d] \arrow[dr] & \\
 & x \arrow[r] \arrow[d] \arrow[dr] & y \arrow[r] \arrow[d] & z \\
\Sigma^{-1}z' \arrow[r] \arrow[dr] & x' \arrow[r] \arrow[d] & y' & \\
 & x'' & &
\end{tikzcd}
$$

Our two octahedra also have bottom pyramids. We deduce diagrams

$$
\begin{tikzcd}
 & x \arrow[dr] \arrow[d, phantom, shift left=4ex, "d"] & \\
\Sigma^{-1}z' \arrow[dr] \arrow[r, phantom, shift right=4ex, "d"] & w \arrow[u, "(1)"] \arrow[l, dotted, "(1)"'] & y' \arrow[l] \\
 & x'' \arrow[u, dotted] &
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
 & \Sigma^{-1}y'' \arrow[dr] \arrow[d, phantom, shift left=4ex, "d"] & \\
x \arrow[dr] \arrow[r, phantom, shift right=4ex, "d"] & w \arrow[u, dotted, "(1)"] \arrow[l, "(1)"'] & z \arrow[l, dotted] \\
 & y' \arrow[u] &
\end{tikzcd}
$$

What we mean by this is that we may complete $x \longrightarrow y'$ to a distinguished triangle

$$x \longrightarrow y' \longrightarrow w \longrightarrow \Sigma x.$$

This is our solid triangle. The octahedral lemma allows us to choose the dotted arrows to complete each of the two octahedra. The diagrams depicting the projections of the bottom pyramids exhibit the commutative triangles as straight lines, and the distinguished triangles as triangles. But now we have the bottom pyramid of an octahedron

$$\begin{tikzcd}
 & \Sigma^{-1}y'' \arrow[dr] & \\
\Sigma^{-1}z' \arrow[dr] & w \arrow[l, "(1)"'] \arrow[u, "(1)"] \arrow[ur, phantom, "d"] \arrow[dl, phantom, "d"] & z \arrow[l] \\
 & x'' \arrow[u] &
\end{tikzcd}$$

which we may complete to a top pyramid. We have three octahedra, with top pyramids

$$\begin{tikzcd}
 & & & \Sigma^{-1}y'' \arrow[d] \arrow[dr] & \\
 & & x \arrow[r] \arrow[d] \arrow[dr] & y \arrow[r] \arrow[d] & z \\
 & \Sigma^{-1}z' \arrow[r] \arrow[d] \arrow[dr] & x' \arrow[r] \arrow[d] & y' & \\
\Sigma^{-1}y'' \arrow[r] \arrow[dr] & \Sigma^{-1}z'' \arrow[r] \arrow[d] & x'' & & \\
 & z & & &
\end{tikzcd}$$

The remaining maps, which define the bottom pyramids of the octahedra, can be written as

$$\Sigma^{-1}w \xrightarrow{\begin{pmatrix}\alpha\\ \beta\\ \gamma\end{pmatrix}} \Sigma^{-1}y'' \oplus x \oplus \Sigma^{-1}z' \qquad\qquad z \oplus y' \oplus x'' \xrightarrow{\begin{pmatrix} a & b & c\end{pmatrix}} w$$

The three octahedra tell us, among other things, that the three diagonals have the same mapping cone, namely the object w. But we know more. Each octahedron gives two commutative squares. The "refined" octahedral lemma chooses these commutative squares to be homotopy pushout squares; see Section 2 for details. Using these six homotopy pushout squares, we can extend the above to a commutative diagram

$$(**)\qquad \begin{tikzcd}
 & & \Sigma^{-1}w \arrow[r, "\alpha"] \arrow[d, "\beta"'] \arrow[dr, phantom, "(1)"] & \Sigma^{-1}y'' \arrow[d] \arrow[dr] & \\
 & \Sigma^{-1}w \arrow[r, "\beta"] \arrow[d, "\gamma"'] \arrow[dr, phantom, "(2)"] & x \arrow[r] \arrow[d] \arrow[dr] & y \arrow[r] \arrow[d] \arrow[dr, phantom, "(4)"] & z \arrow[d, "a"] \\
\Sigma^{-1}w \arrow[r, "\gamma"] \arrow[d, "\alpha"'] \arrow[dr, phantom, "(3)"] & \Sigma^{-1}z' \arrow[r] \arrow[d] \arrow[dr] & x' \arrow[r] \arrow[d] \arrow[dr, phantom, "(5)"] & y' \arrow[r, "b"'] \arrow[d, "b"] & w \\
\Sigma^{-1}y'' \arrow[r] \arrow[dr] & \Sigma^{-1}z'' \arrow[r] \arrow[d] \arrow[dr, phantom, "(6)"] & x'' \arrow[r, "c"'] \arrow[d, "c"] & w & \\
 & z \arrow[r, "a"'] & w & &
\end{tikzcd}$$

In this diagram, the six squares labeled $(1)-(6)$ are homotopy pushout squares, and the differential of the square labeled (n) is given by the diagonal of the square labeled $(n+3)$, where we read the labels modulo 6. For example, the differential of (1) is given by the diagonal of $(1+3)=(4)$.

We also have, in the upper part of the three octahedra, six distinguished triangles. These assemble to a diagram

$$\begin{tikzcd}
x \arrow[r] \arrow[d] & y \arrow[r] \arrow[d] & z \arrow[r] \arrow[d] & \Sigma x \arrow[d] \\
x' \arrow[r] \arrow[d] & y' \arrow[r] \arrow[d] & z' \arrow[r] \arrow[d] & \Sigma x' \arrow[d] \\
x'' \arrow[r] \arrow[d] & y'' \arrow[r] \arrow[d] & z'' \arrow[r] \arrow[d] \arrow[dr, phantom, "(-)"] & \Sigma x'' \arrow[d] \\
\Sigma x \arrow[r] & \Sigma y \arrow[r] & \Sigma z \arrow[r] & \Sigma^2 x
\end{tikzcd}$$

The standard sign convention has all the squares commuting except the one at the bottom right, which anticommutes. We put the symbol $(-)$ in the bottom right square to remind ourselves that it anticommutes.
If we begin with a square of monomorphisms in an abelian category, satisfying the condition in Lemma 3.4(iii), then all the choices in the octahedra we constructed are canonically unique. Using only the octahedral axiom the above argument (of Verdier) shows how to extend a commutative square to an elaborate diagram with many distinguished triangles. In the special case where the square is the top left corner of a 3×3 diagram of short exact sequences in $\mathcal{A}$, the extension in $D^b(\mathcal{A})$ is unique. We recover the 3×3 diagram, the mapping cone w on the map $x \longrightarrow y'$, and many distinguished triangles. The first theorem is

Theorem 3.5. *Axiom* (TC3) *of May's is just the assertion that the tensor product of two distinguished triangles comes with* Verdier structure. *By this we mean that there exists an object w, and a diagram* $(**)$ *as on page 547.*

Proof: It needs to be checked that May's list of the properties of the object w precisely coincides with what we obtained above, from the octahedral axiom. We leave this to the interested reader; one needs to compare $(**)$ with May's beautifully drawn diagram on page 49 of [12]. $\square$

Remark 3.6. In the discussion preceding Theorem 3.5 we indicated how, following either May or Verdier, one can prove that a triangulated category $\mathcal{T}$ with a decent tensor product satisfies (TC3).
Note that, both in May's and in Verdier's argument, the tensor product plays a very minor role. What matters is that in the abelian category $\mathcal{A}$ we have a 3×3 diagram with exact rows and columns. The fact that it happens to come from the tensor product of two short exact sequences is largely irrelevant.

Remark 3.7. Haynes Miller independently observed that Verdier's construction yields May's diagram.

So far we have explained the relation between May's work and Verdier's. Now we move to the more interesting observation. We will explain the relation between two approaches we have already seen and Happel's work.
What we have seen so far is the following. We started with a commutative square of monomorphisms in $\mathcal{A}$

$$\begin{tikzcd}
x \arrow[r] \arrow[d] & y \arrow[d] \\
x' \arrow[r] & y'
\end{tikzcd}$$

satisfying the condition in Lemma 3.4(iii). Then, either by pushing out in the abelian category $\mathcal{A}$ or by repeatedly applying the octahedral lemma, we extended to an elaborate diagram of triangles giving May's (TC3). Let k be a noetherian commutative ring. Suppose the category $\mathcal{A}$ is k-linear. (For any abelian category $\mathcal{A}$ we may take $k = \mathbb{Z}$.) A commutative square in $\mathcal{A}$ may be viewed as a k-linear functor $\Box \longrightarrow \mathcal{A}$, where $\Box$ is the k-category presented by the quiver (=oriented graph)

$$\begin{tikzcd}[row sep=small, column sep=small] & 2 \arrow[dr, "\gamma"] & \\ 1 \arrow[ur, "\alpha"] \arrow[dr, "\beta"'] & & 4 \\ & 3 \arrow[ur, "\delta"'] & \end{tikzcd} \qquad \text{with the relation} \quad \gamma\,\alpha = \delta\,\beta \ .$$

In other words, the category $\Box$ has four objects corresponding to the four vertices of the quiver, and its morphisms between two objects are obtained by taking all k-linear combinations of paths between the corresponding vertices of the quiver and dividing out all consequences of the relations. Let $\operatorname{mod} k$ be the category of finitely generated k-modules. Let $\operatorname{mod} \Box = \mathcal{C}at(\Box^{op}, \operatorname{mod} k)$ be the category of all k-linear functors $\Box^{op} \longrightarrow \operatorname{mod} k$. We remind the reader of the well-known

Lemma 3.8. *Up to canonical isomorphism, any k-linear functor $\Box \longrightarrow \mathcal{A}$ may be factored uniquely as*

$$\Box \longrightarrow \operatorname{mod} \Box \xrightarrow{F} \mathcal{A},$$

with F a right exact k-linear functor of k-linear categories.

Proof: Any right exact functor $F : \operatorname{mod} \Box \longrightarrow \mathcal{A}$ is uniquely determined by what it does on projective objects. And each projective object in the functor category $\operatorname{mod} \Box$ is a direct factor of a finite sum of the representable functors $P_i = \operatorname{Hom}_\Box(-, i)$. Since the Yoneda functor is covariant, the representable functors appear in a commutative square

$$(*) \qquad \begin{tikzcd}[row sep=small, column sep=small] & P_2 \arrow[dr, "\gamma_*"] & \\ P_1 \arrow[ur, "\alpha_*"] \arrow[dr, "\beta_*"'] & & P_4 \\ & P_3 \arrow[ur, "\delta_*"'] & \end{tikzcd} \quad .$$

Given a functor $F : \operatorname{mod} \Box \longrightarrow \mathcal{A}$, F must take the commutative square $(*)$ in $\operatorname{mod} \Box$ to a commutative square in $\mathcal{A}$. Conversely, given a commutative square in $\mathcal{A}$, we want a functor F. It is clear how to define F on P_1, P_2, P_3 and P_4. This definition extends by additivity to direct summands of direct sums of the P_i's, that is to all projectives. Finally, to define F on an arbitrary object X, choose a projective presentation for X

$$P \longrightarrow Q \longrightarrow X \longrightarrow 0,$$

and $F(X)$ is defined to be the cokernel of $F(P) \longrightarrow F(Q)$. $\Box$

Taking the left derived functor of the F in Lemma 3.8, we have that any commutative square in $\mathcal{A}$ yields a functor $D^b(\square) \longrightarrow D^b(\mathcal{A})$, where we abbreviate $D^b(\mathrm{mod}\,\square)$ to $D^b(\square)$.

Remark 3.9. For much more detail see [10], [8], [11].

Theorem 3.10. *The relations which hold in $D^b(\square)$ are universal. The same diagram of triangles will exist in any triangulated category $\mathcal{T}$ with a decent tensor product.*

Proof: The commutative square, which we saw in Verdier's construction of the diagram, will give rise to a triangulated functor $D^b(\square) \longrightarrow D^b(\mathcal{A})$. The decency of the tensor product gives a triangulated functor $D^b(\mathcal{A}) \longrightarrow \mathcal{T}$. The composite takes the diagram of triangles in $D^b(\square)$ to $\mathcal{T}$. $\square$

Remark 3.11. The word "universal" is appearing here in an extended, somewhat unusual way. We are not asserting that the category $D^b(\square)$ has a decent tensor product. As far as we know it has no tensor product at all; there does not seem to be a Hopf algebra structure on the quiver algebra.
All we say is the following. Let $\mathcal{T}$ be a triangulated category with a decent tensor product. Then triangles appearing in $D^b(\square)$ will always be reflected in the tensor product of two distinguished triangles in $\mathcal{T}$. As we have already said, the tensor product in $\mathcal{T}$ plays a minor role in the proof, and the category $D^b(\square)$ does not seem to have a tensor product at all.

The real use of Theorem 3.10 is that Happel studied the category $D^b(\square)$ in great detail, in the case where the ground ring k is a field. By appealing to his results we can obtain a great deal of information, at least in the case of k-linear triangulated categories over fields k. In principle, it should not be particularly difficult to generalise Happel's work to the case where $k = \mathbb{Z}$. In this article we chose not to do so. We chose to highlight the idea, not to pursue it to obtain the sharpest results. The main reason is that we wanted to keep the article reasonably brief.
In the next two sections, we will show how the different approaches can yield new results.

4. Consequences of the octahedral axiom

First we establish some notation. Consider the diagram

$$\begin{tikzcd}
 & & x \arrow[r] \arrow[d] \arrow[dr, dashed] & y \arrow[r] \arrow[d] \arrow[dr, squiggly] & z \arrow[r] \arrow[d] \arrow[dr] & \Sigma x \arrow[d] \\
 & \Sigma^{-1}z' \arrow[r] \arrow[d] \arrow[dr, dashed] & x' \arrow[r] \arrow[d] \arrow[dr, squiggly] & y' \arrow[r] \arrow[d] \arrow[dr] & z' \arrow[r] \arrow[d] & \Sigma x' \arrow[d] \\
\Sigma^{-1}y'' \arrow[r] \arrow[d] \arrow[dr, dashed] & \Sigma^{-1}z'' \arrow[r] \arrow[d] \arrow[dr, squiggly] & x'' \arrow[r] \arrow[d] \arrow[dr] & y'' \arrow[r] \arrow[d] & z'' \arrow[r] \arrow[d] & \Sigma x'' \arrow[d] \\
y \arrow[r] & z \arrow[r] & \Sigma x \arrow[r] & \Sigma y \arrow[r] & \Sigma z \arrow[r] & \Sigma^2 x
\end{tikzcd}$$

The axiom (TC3) assigns a common mapping cone w to the three broken arrows. Applying (TC3) to rotations of the triangles, we expect a common mapping cone Σu to the curly arrows, and a common mapping cone Σv to the plain arrows. Needless to say, u, v and w should be related. May found one relation. We will use the different approaches to obtain more.
In this section we will, following Verdier's approach, see what the octahedral lemma buys us. We have:

Theorem 4.1. *May's axiom* (TC4) *is a formal consequence of* (TC3) *and the octahedral axiom. The proof will give us yet another distinguished triangle. It is a triangle May missed, whose existence also follows formally from* (TC3) *and the octahedral axiom.*

Proof: Recall that the octahedra defining w give a homotopy pushout square

$$\begin{tikzcd}
y \arrow[r] \arrow[d] & z \arrow[d] \\
y' \arrow[r] & w
\end{tikzcd}$$

The mapping cone on the diagonal $y \longrightarrow w$ is just the sum of the mapping cones on the horizontal and vertical maps, that is $y'' \oplus \Sigma x$. The triangle

$$\Sigma^{-1}z' \longrightarrow x'' \longrightarrow w \longrightarrow z'$$

gives us a map $w \longrightarrow z'$, and hence a commutative square

$$\begin{tikzcd}
y \arrow[r] \arrow[d] & z \arrow[d] \\
y' \arrow[r] & z'
\end{tikzcd}$$

The object Σu is the mapping cone of the diagonal map $y \longrightarrow z'$ in this square. In other words, Σu is the mapping cone on a composite

$$y \longrightarrow w \longrightarrow z'.$$

We now complete to an octahedron. We know all the objects of the octahedron. In the standard notation, where d stands for a distinguished triangle, $+$ for a commutative one and (1) for an arrow of degree one, we draw the octahedron. The top pyramid is

$$\begin{tikzcd}
y \arrow[rr] \arrow[dr] & & z' \arrow[dd] \\
& w \arrow[ur] \arrow[dl] \arrow[u, phantom, "+"] \arrow[l, phantom, "d"] \arrow[r, phantom, "d"] \arrow[d, phantom, "+"] & \\
y'' \oplus \Sigma x \arrow[uu, "(1)"] & & \Sigma x'' \arrow[ll, "(1)"] \arrow[ul, "(1)"']
\end{tikzcd}$$

The bottom pyramid is

$$\begin{tikzcd}
y \arrow[rr] & & z' \arrow[dd] \arrow[dl] \\
& \Sigma u \arrow[ul, "(1)"'] \arrow[dr] \arrow[u, phantom, "d"] \arrow[l, phantom, "+"] \arrow[r, phantom, "+"] \arrow[d, phantom, "d"] & \\
y'' \oplus \Sigma x \arrow[uu, "(1)"] \arrow[ur] & & \Sigma x'' \arrow[ll, "(1)"]
\end{tikzcd}$$

From this octahedron we deduce two homotopy pushout squares. There are therefore distinguished triangles

$$w \longrightarrow z' \oplus y'' \oplus \Sigma x \longrightarrow \Sigma u \longrightarrow \Sigma w$$

and

$$u \longrightarrow y \oplus x'' \longrightarrow w \longrightarrow \Sigma u$$

The first of these triangles is axiom (TC4) of May's [12]. The second is new. $\square$

In the next section we will see how to better organise all the triangles above, and more.

5. The relation with Happel's work

As we saw in Theorem 3.10, the problem reduces to understanding the category $D^b(\square)$. It helps to introduce an equivalent derived category. We define

Definition 5.1. *Let D_4 be the quiver(=oriented graph)*

(1)
$$\begin{tikzcd} & & 2 \\ 1 \arrow[r, "\lambda"] & 5 \arrow[ur, "\mu"] \arrow[dr, "\nu"'] & \\ & & 3 \end{tikzcd} .$$

Let Y *be the k-category presented by the quiver D_4. Let* $\operatorname{mod}\mathsf{Y}$ *be the category of k-linear functors* $\mathsf{Y}^{op} \longrightarrow \operatorname{mod} k$. *We denote the bounded derived category* $D^b(\operatorname{mod}\mathsf{Y})$ *by* $D^b(\mathsf{Y})$.

The interest in this definition comes from the well-known

Lemma 5.2. *The derived categories $D^b(\Box)$ and $D^b(\mathsf{Y})$ are equivalent, as k-linear triangulated categories.*

Proof: Recall that $\Box$ is the category presented by the quiver

$$\begin{tikzcd} & 2 \arrow[dr, "\gamma"] & \\ 1 \arrow[ur, "\alpha"] \arrow[dr, "\beta"'] & & 4 \\ & 3 \arrow[ur, "\delta"'] & \end{tikzcd} \qquad \text{with the relation} \quad \gamma\,\alpha = \delta\,\beta$$

and $D^b(\Box)$ the bounded derived category of the category $\operatorname{mod}\Box$ of k-linear functors $\Box^{op} \longrightarrow \operatorname{mod} k$. The categories $\operatorname{mod}\mathsf{Y}$ and $\operatorname{mod}\Box$ are related by a natural pair L, R of adjoint functors: If M is in $\operatorname{mod}\Box$, we complete the corresponding diagram of k-modules into

$$\begin{tikzcd} & & M2 \arrow[dll, bend right, "M\alpha"'] \arrow[dl, dotted, "M\mu"] & \\ M1 & M5 \arrow[l, dotted, "M\lambda"'] & & M4 \arrow[ul, "M\gamma"'] \arrow[dl, "M\delta"] \\ & & M3 \arrow[ull, bend left, "M\beta"] \arrow[ul, dotted, "M\nu"'] & \end{tikzcd} ,$$

where $M5$ is the pushout of $M\gamma$ and $M\delta$. We define the object $LM \in \operatorname{mod}\mathsf{Y}$ as the full subdiagram on $M1, M2, M3, M5$. Similarly, if N is in $\operatorname{mod}\mathsf{Y}$, we complete the corresponding diagram by defining $N4$ as the pullback of $N\mu$ and $N\nu$, and we define RN to be the full subdiagram on $N1, N2, N3, N4$. Note that the functors L and R are not equivalences (they take some non zero objects to zero). But the left derived functor of L is easily computed to be quasi-inverse to the right derived functor of R, giving an equivalence between $D^b(\Box)$ and $D^b(\mathsf{Y})$. $\Box$

Remark 5.3. The experts will note that the fact that L and R induce equivalences of derived categories is a special case [1] of tilting theory (cf. e.g. [9], [7], [2]).

Theorem 3.10 tells us that we are reduced to understanding the distinguished triangles in the category $D^b(\Box) = D^b(\mathsf{Y})$. The proof of Theorem 3.10, more specifically Lemma 3.8, tells us that in the category $\operatorname{mod}\Box \subset D^b(\Box)$ we have a commutative square

$$\begin{tikzcd} & P_2 \arrow[dr, "\gamma_*"] & \\ P_1 \arrow[ur, "\alpha_*"] \arrow[dr, "\beta_*"'] & & P_4 \\ & P_3 \arrow[ur, "\delta_*"'] & \end{tikzcd}$$

and everything reduces to understanding the distinguished triangles in which it lies. The equivalence $D^b(\Box) = D^b(\mathsf{Y})$ is explicit enough to be able to work out the image of this commutative square in $D^b(\mathsf{Y})$.

5.1. Happel's description of $D^b(\mathsf{Y})$. From now on, we assume k is a field. We will describe $D^b(\mathsf{Y})$ as a k-linear category, following Happel [9]. This description will also yield a great deal of information on the distinguished triangles of $D^b(\mathsf{Y})$. Note that Happel built on previous work by many researchers, notably Ringel [14], Riedtmann [13], Gabriel [6]. For more information, we refer to the books [15], [2], [7].

Each object of the category $D^b(\mathsf{Y})$ decomposes into a finite sum of indecomposable objects with local endomorphism rings and this decomposition is unique up to permutation and isomorphism. To describe $D^b(\mathsf{Y})$ as a k-linear category, it suffices therefore to describe the full subcategory formed by the indecomposable objects.

We will give a presentation of the category of indecomposables in $D^b(\mathsf{Y})$. Let us first describe its objects: The category $\operatorname{mod}\mathsf{Y}$ is the k-category of representations of a quiver without relations. Therefore, it is an abelian category of global dimension ≤ 1. This entails that in its derived category $D^b(\mathsf{Y})$, each object is (non canonically) isomorphic to the direct sum of its homologies placed in their respective degrees. Each indecomposable of $D^b(\mathsf{Y})$ is therefore concentrated in one degree, i.e. it is a shift of some indecomposable module. Now D_4 is a quiver whose underlying graph is a *Dynkin diagram.* So by Gabriel's theorem [5], there are only finitely many (isomorphism classes of) indecomposable modules; moreover, the indecomposables are in bijection with the twelve positive roots of the corresponding root system (the orientation of the quiver determines the positive cone). The bijection is given by sending each indecomposable M to its *dimension vector* $\underline{\dim}\, M$, i.e. to the function $i \mapsto \dim Mi$. For example, the dimension vector of the module $P_2 : i \mapsto \operatorname{Hom}_{\mathsf{Y}}(i, 2)$ is given by

$$\begin{smallmatrix} & 1 \\ 1 & 1 \\ & 0 \end{smallmatrix}\ .$$

Note that, by definition of P_i, we have

$$Mi = \operatorname{Hom}_{\mathsf{Y}}(P_i, M)\ \ ,\ M \in \operatorname{mod}\mathsf{Y}\,.$$

Thus the i-component of the dimension vector of M is the dimension of the space of morphisms from P_i to M. The map $M \mapsto \underline{\dim}\, M$ induces an isomorphism

$$K_0(\operatorname{mod} \mathsf{Y}) \longrightarrow \mathbf{Z}^4.$$

The simple modules correspond to the vectors of the standard basis.

Caution 5.4. In what follows, we make frequent reference to figure 1 and figure 2. For the reader's convenience, both figures have been placed *on the last page,* after the bibliography.

Let us summarize the above: The indecomposable objects of $D^b(\mathsf{Y})$ are the shifts of the indecomposable modules; the indecomposable modules are determined by their dimension vectors. The positive dimension vectors in figure 1 are precisely the dimension vectors of indecomposable objects of Y. The negative vectors in the figure correspond to shifted indecomposable modules.
We now describe the morphisms between indecomposable objects of $D^b(\mathsf{Y})$. Let U and V be indecomposable. A *radical morphism* from U to V is a non invertible morphism $f : U \longrightarrow V$. Denote by $\operatorname{rad}(U, V)$ the space of radical morphisms from U to V. Clearly, rad is an ideal of the category of indecomposables. Denote its square by rad^2. Thus a morphism $f : U \longrightarrow V$ belongs to $\operatorname{rad}^2(U, V)$ iff it is *reducible*, i.e. we have

$$f = \sum_{i=1}^{n} g_i h_i$$

for some n and for radical morphisms $h_i : U \longrightarrow W_i$ and $g_i : W_i \longrightarrow V$. A morphism is *irreducible* if it is not reducible. The *Auslander-Reiten quiver* of $D^b(\mathsf{Y})$ is the quiver whose vertices are the isomorphism classes $[U]$ of indecomposable objects and which has $\dim \operatorname{rad}(U, V)/\operatorname{rad}^2(U, V)$ arrows from the vertex $[U]$ to the vertex $[V]$.
Happel's theorem [9, Cor. 4.5] yields as a special case that the Auslander-Reiten quiver of $D^b(\mathsf{Y})$ is the quiver R of figure 1.
We will obtain the required presentation of the category of indecomposables of $D^b(\mathsf{Y})$ by dividing the free k-category on the Auslander-Reiten quiver by suitable relations, which we now describe. To do so, we introduce the automorphism $\tau : R \longrightarrow R$ which is the shift by two units to the left. It is called the *Auslander-Reiten translation.* For example, we have

$$\tau\left(\begin{matrix} 0 \\ 0\,1 \\ 1 \end{matrix} \right) = \begin{matrix} 1 \\ 1\,1 \\ 0 \end{matrix} \;,\; \tau\left(\begin{matrix} 1 \\ 1\,2 \\ 1 \end{matrix} \right) = \begin{matrix} 0 \\ 1\,1 \\ 0 \end{matrix} \;.$$

The *mesh relation* associated to a vertex x of R is the relation

$$\sum \alpha\beta = 0 \;,$$

where the sum ranges over all subquivers

$$\tau(x) \xrightarrow{\beta} y \xrightarrow{\alpha} x.$$

Theorem 5.5 (Happel [9]). *There is an equivalence Φ from the k-category presented by the Auslander-Reiten quiver R of figure 1 subject to all mesh relations to the category of indecomposables of $D^b(\mathsf{Y})$. It can be chosen so that for each vertex x of R, $\underline{\dim}\,\Phi(x)$ is the dimension vector associated with x and that for each arrow $\alpha : x \longrightarrow y$ of R, $\Phi(\alpha)$ is an irreducible morphism from $\Phi(x)$ to $\Phi(y)$. Moreover, for each vertex x of R, there is a canonical triangle (called the* Auslander-Reiten triangle*)*

$$\Phi(\tau(x)) \longrightarrow \bigoplus \Phi(y) \longrightarrow \Phi(x) \longrightarrow \Sigma\Phi(\tau(x)) \ ,$$

where the sum ranges over all subquivers

$$\tau(x) \xrightarrow{\beta} y \xrightarrow{\alpha} x.$$

Remark 5.6. Under the equivalence of the theorem, the suspension $U \mapsto \Sigma U$ corresponds to τ^{-3}.

Remark 5.7. The group $S_3 \times \mathbf{Z}$ acts on the quiver R: The factor $\mathbf{Z}$ acts via τ; the factor S_3 fixes the τ-orbit of $\begin{smallmatrix} & 0 \\ 1 & 1 \\ & 0 \end{smallmatrix}$ and simultaneously permutes the vertices

$$\tau^i\left(\begin{smallmatrix} & 1 \\ 1 & 1 \\ & 0 \end{smallmatrix}\right), \quad \tau^i\left(\begin{smallmatrix} & 0 \\ 0 & 1 \\ & 0 \end{smallmatrix}\right), \quad \tau^i\left(\begin{smallmatrix} & 0 \\ 1 & 1 \\ & 1 \end{smallmatrix}\right)$$

for each $i \in \mathbf{Z}$. By Happel's theorem, we obtain an action on $D^b(\mathsf{Y})$. The autoequivalences of $D^b(\mathsf{Y})$ which occur are triangulated functors.

Remark 5.8. Lemma 5.2 gives an equivalence $D^b(\Box) \longrightarrow D^b(\mathsf{Y})$. The natural commutative square in $\operatorname{mod}\Box$ maps via the composite

$$\operatorname{mod}\Box \longrightarrow D^b(\Box) \longrightarrow D^b(\mathsf{Y})$$

to the square formed by the vertices labeled

$$\begin{smallmatrix} & 0 \\ 1 & 0 \\ & 0 \end{smallmatrix}\ , \quad \begin{smallmatrix} & 1 \\ 1 & 1 \\ & 0 \end{smallmatrix}\ , \quad \begin{smallmatrix} & 0 \\ 1 & 1 \\ & 1 \end{smallmatrix}\ , \quad \begin{smallmatrix} & 1 \\ 1 & 1 \\ & 1 \end{smallmatrix}\ .$$

Remark 5.9. Suppose that x is a vertex of R corresponding to a representable functor P_i and y an arbitrary vertex. Then we have

$$\dim \operatorname{Hom}(\Phi(x), \Phi(y)) = \dim \operatorname{Hom}(P_i, \Phi(y)) = (\underline{\dim}\,\Phi(y))_i \ .$$

For an arbitrary vertex x, there is always an $i \in \mathbf{Z}$ so that $\tau^i x$ corresponds to a representable. This allows us to compute

$$\dim \operatorname{Hom}(\Phi(x), \Phi(y)) = \dim \operatorname{Hom}(\Phi(\tau^i x), \Phi(\tau^i y))$$

very easily by inspecting figure 1.

Remark 5.10. Happel's theorem allows us to exhibit many triangles produced by short exact sequences of the module category $\operatorname{mod}\mathsf{Y}$. It is clear that a sequence of modules

$$0 \longrightarrow L \longrightarrow M \longrightarrow N \longrightarrow 0$$

is exact iff it is a complex, the left arrow is a monomorphism, the right arrow an epimorphism and we have $\underline{\dim}\, M = \underline{\dim}\, L + \underline{\dim}\, N$. These conditions are easy to check with the help of the Auslander-Reiten quiver. Using the action of $S_3 \times \mathbf{Z}$ of Remark 5.7 we obtain further triangles.

5.2. Application: Organising the triangles. Suppose that we have a commutative square

$$\begin{array}{ccc} x & \longrightarrow & y \\ \big\downarrow & & \big\downarrow \\ x' & \longrightarrow & y' \end{array}$$

in a k-linear triangulated category $\mathcal{T}$. Suppose that, as in Theorem 3.10, we also have a triangulated functor

$$F : D^b(\square) = D^b(\mathsf{Y}) \longrightarrow \mathcal{T}$$

extending the square. If we compose F with the isomorphism Φ of Happel's theorem, we obtain the mapping suggested by superposing figures 1 and 2. Here we use the notations of Sections 3 and 4, as well as some of the triangles of $D^b(\square) = D^b(\mathsf{Y})$ obtained from Remark 5.10.

Note that, miraculously, the twelve objects of Section 3 and 4 correspond to the twelve orbits of *indecomposable* Y-modules under the action of the group $\Sigma^{\mathbf{Z}}$ generated by Σ and that the 'interesting' objects u, v, w are in the same orbit under the action of the group $\tau^{\mathbf{Z}}$ generated by τ.

Remark 5.11. Perhaps the miracle deserves a small comment. The objects x, x', y and y' in the commutative square

$$\begin{array}{ccc} x & \longrightarrow & y \\ \big\downarrow & & \big\downarrow \\ x' & \longrightarrow & y' \end{array}$$

correspond, in $D^b(\square)$, to P_1, P_2, P_3 and P_4, all of which are (projective) indecomposables in $\square \subset D^b(\square)$. Being indecomposable in $D^b(\square)$, they must remain indecomposable under the equivalence $D^b(\square) \simeq D^b(\mathsf{Y})$ of Lemma 5.2. This much is not surprising.

The miracle, which the authors do not understand, is why the other naturally arising eight objects correspond precisely, up to suspension, to the other classes of indecomposables.

Let us call two triangles *equivalent* if they are obtained from one another by rotations and the action of $S_3 \times \mathbf{Z}$. Then the triangles constructed in Sections 3 and 4 belong to the (distinct) equivalence classes of the following seven triangles (those marked with (∗) are rotations of Auslander-Reiten triangles)

$$\begin{array}{llll}
x \longrightarrow u \longrightarrow \Sigma^{-1}z'' \longrightarrow \Sigma x & & & (*) \\
x \longrightarrow y \longrightarrow z \longrightarrow \Sigma x & & & \\
x \longrightarrow v \longrightarrow z \oplus x'' \longrightarrow \Sigma x & & & \\
x \longrightarrow y' \longrightarrow w \longrightarrow \Sigma x & & & (*) \\
u \longrightarrow x'' \longrightarrow \Sigma x \oplus y'' \longrightarrow \Sigma u & & & \\
u \longrightarrow w \longrightarrow \Sigma x \oplus y'' \oplus z' \longrightarrow \Sigma u & & & (*) \\
u \longrightarrow v \longrightarrow x'' \oplus z' \longrightarrow \Sigma u \ , & & &
\end{array}$$

where the last triangle is equivalent to the new triangle constructed at the end of Section 4. Note that the morphism space $\mathrm{Hom}(u, v)$ is 2-dimensional so that the morphism $u \longrightarrow v$ is not even unique up to a scalar multiple. The morphism $u \longrightarrow v$ occurring in the last triangle is defined to be the composition $u \longrightarrow y \longrightarrow v$. Note that up to the action of $S_3 \times \mathbf{Z}$ this is the only 2-dimensional morphism space between indecomposables.
Let us construct some more triangles: The plane $\mathrm{Hom}(u, v)$ contains three distinguished lines given by the morphisms factoring respectively through x', y and $\Sigma^{-1}z''$. The mapping cone triangle over a morphism lying in one of the lines is equivalent to the last triangle of the list above. However, if we choose a morphism f outside of the three lines, we obtain a new triangle

$$u \xrightarrow{f} v \xrightarrow{g} w \xrightarrow{\varepsilon(f,g)} u$$

by looking at the corresponding short exact sequence of Y-modules. Thus, we obtain a whole new family of isomorphism classes of triangles, parametrized by the projective line over k punctured at 3 points.
We claim that this is the list of all equivalence classes of non-split triangles with two indecomposable vertices. To check this, one proceeds in two steps: (1) classify morphisms between indecomposables of $D^b(\mathsf{Y})$ up to conjugacy under the the group $S_3 \times \mathbf{Z}$; (2) inspect all mapping cone triangles over the morphisms obtained in (1) and eliminate duplicates. We leave the details to the interested reader.

References

[1] Maurice Auslander, María Inés Platzeck, and Idun Reiten, *Coxeter functors without diagrams*, Trans. Amer. Math. Soc. 250 (1979), 1–46.

[2] Maurice Auslander, Idun Reiten, and Sverre Smalø, *Representation theory of Artin algebras*, Cambridge Studies in Advanced Mathematics, vol. 36, Cambridge University Press, 1995 (English).

[3] Paul Balmer, *Presheaves of triangulated categories and reconstruction of schemes,*, To appear in Math. Ann.

[4] Alexander A. Beilinson, Joseph Bernstein, and Pierre Deligne, *Analyse et topologie sur les éspaces singuliers*, Astérisque, vol. 100, Soc. Math. France, 1982 (French).

[5] Peter Gabriel, *Unzerlegbare Darstellungen I*, Manuscripta Math. 6 (1972), 71–103.

[6] ———, *Auslander-Reiten sequences and representation-finite algebras*, Representation theory, I (Proc. Workshop, Carleton Univ., Ottawa, Ont., 1979), Springer, Berlin, 1980, pp. 1–71.

[7] Peter Gabriel and Andrei V. Roiter, *Representations of finite-dimensional algebras*, Encyclopaedia Math. Sci., vol. 73, Springer–Verlag, 1992.

[8] Alexander Grothendieck, *Les dérivateurs*, Manuscript, 1990, electronically edited by M. Künzer, J. Malgoire and G. Maltsiniotis.

[9] Dieter Happel, *On the derived category of a finite-dimensional algebra*, Comment. Math. Helv. 62 (1987), no. 3, 339–389.

[10] Alex Heller, *Homotopy theories*, Mem. Amer. Math. Soc. 71 (1988), no. 383, vi+78.

[11] Bernhard Keller, *Derived categories and universal problems*, Comm. in Algebra 19 (1991), 699–747.

[12] J. Peter May, *The additivity of traces in triangulated categories*, Adv. Math. 163 (2001), no. 1, 34–73.

[13] Christine Riedtmann, *Algebren, Darstellungsköcher, Überlagerungen und zurück*, Comment. Math. Helv. 55 (1980), no. 2, 199–224.

[14] Claus Michael Ringel, *Finite dimensional hereditary algebras of wild representation type*, Math. Z. 161 (1978), no. 3, 235–255.

[15] ———, *Tame algebras and integral quadratic forms*, Lecture Notes in Mathematics, vol. 1099, Springer Verlag, 1984.

Bernhard Keller
Université Paris 7
UFR de Mathématiques
Institut de Mathématiques
UMR 7586 du CNRS
Case 7012
2, place Jussieu
75251 Paris Cedex 05
France
keller@math.jussieu.fr

Amnon Neeman
Centre for Mathematics
and its Applications
Mathematical Sciences Institute
John Dedman Building
The Australian National University
Canberra, ACT 0200
Australia
neeman@maths.anu.edu.au

$$
\begin{tikzcd}[column sep=small]
 & \begin{matrix} & -1 \\ 0 & 0 \\ & 0 \end{matrix} \arrow[dr] & & \begin{matrix} & 1 \\ 1 & 1 \\ & 0 \end{matrix} \arrow[dr] & & \begin{matrix} & 0 \\ 0 & 1 \\ & 1 \end{matrix} \arrow[dr] & & \begin{matrix} & 1 \\ 0 & 0 \\ & 0 \end{matrix} & \\
\cdots & \begin{matrix} & 0 \\ 1 & 0 \\ & 0 \end{matrix} \arrow[r] & \begin{matrix} & 0 \\ 1 & 1 \\ & 0 \end{matrix} \arrow[r] \arrow[ur] \arrow[dr] & \begin{matrix} & 0 \\ 0 & 1 \\ & 0 \end{matrix} \arrow[r] & \begin{matrix} & 1 \\ 1 & 2 \\ & 1 \end{matrix} \arrow[r] \arrow[ur] \arrow[dr] & \begin{matrix} & 1 \\ 1 & 1 \\ & 1 \end{matrix} \arrow[r] & \begin{matrix} & 1 \\ 0 & 1 \\ & 1 \end{matrix} \arrow[r] \arrow[ur] \arrow[dr] & \begin{matrix} & 0 \\ -1 & 0 \\ & 0 \end{matrix} & \cdots \\
 & \begin{matrix} & 0 \\ 0 & 0 \\ & -1 \end{matrix} \arrow[ur] & & \begin{matrix} & 0 \\ 1 & 1 \\ & 1 \end{matrix} \arrow[ur] & & \begin{matrix} & 1 \\ 0 & 1 \\ & 0 \end{matrix} \arrow[ur] & & \begin{matrix} & 0 \\ 0 & 0 \\ & 1 \end{matrix} &
\end{tikzcd}
$$

Figure 1.

$$
\begin{tikzcd}[column sep=small]
 & \Sigma^{-1}z' \arrow[dr] & & y \arrow[dr] & & x'' \arrow[dr] & & z' & \\
\cdots & x \arrow[r] & u \arrow[r] \arrow[ur] \arrow[dr] & \Sigma^{-1}z'' \arrow[r] & v \arrow[r] \arrow[ur] \arrow[dr] & y' \arrow[r] & w \arrow[r] \arrow[ur] \arrow[dr] & \Sigma x & \cdots \\
 & \Sigma^{-1}y'' \arrow[ur] & & x' \arrow[ur] & & z \arrow[ur] & & y'' &
\end{tikzcd}
$$

Figure 2.

Reconstruction Phases for Hamiltonian Systems on Cotangent Bundles

Anthony D. Blaom

Received: October 11, 2000
Revised: August 5, 2002

Communicated by Bernold Fiedler

Abstract. Reconstruction phases describe the motions experienced by dynamical systems whose symmetry-reduced variables are undergoing periodic motion. A well known example is the non-trivial rotation experienced by a free rigid body after one period of oscillation of the body angular momentum vector.

Here reconstruction phases are derived for a general class of Hamiltonians on a cotangent bundle T^*Q possessing a group of symmetries G, and in particular for mechanical systems. These results are presented as a synthesis of the known special cases $Q = G$ and G Abelian, which are reviewed in detail.

2000 Mathematics Subject Classification: 70H33, 53D20.
Keywords and Phrases: mechanical system with symmetry, geometric phase, dynamic phase, reconstruction phase, Berry phase, cotangent bundle.

Contents

Summary of selected notation

Numbers in parentheses refer to the relevant subsection.

$p_\mu : \mathfrak{g}^* \to \mathfrak{g}_\mu^*$	natural projection (dual of inclusion $\mathfrak{g}_\mu \to \mathfrak{g}$) (2.2)
$\mathrm{pr}_\mu : \mathfrak{g} \to \mathfrak{g}_\mu$	orthogonal projection (2.8)
$(\,\cdot\,)_Q$	associated form (3.2, 7.2)
$i_{\mathcal{O}} \in \Omega^0(\mathcal{O}, \mathfrak{g}^*)$	inclusion $\mathcal{O} \hookrightarrow \mathfrak{g}^*$ (3.3)
$\rho^\circ : (\ker \mathrm{T}\rho)^\circ \to \mathrm{T}^*(Q/G)$	map sending $d_q(f \circ \rho)$ to $d_{\rho(q)} f$ (4.1)
$\mathbf{A}' : \mathrm{T}^*Q \to \mathbf{J}^{-1}(0)$	projection along $(\ker \mathbf{A})^\circ$ (4.2)
$\mathrm{T}^*_{\mathbf{A}}\rho : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G)$	Hamiltonian analogue of $\mathrm{T}\rho : \mathrm{T}Q \to \mathrm{T}(Q/G)$ (4.2)
$i_\mu : \mathbf{J}^{-1}(\mu) \hookrightarrow P$	inclusion (5.1)
$E(\mu) \subset \mathrm{T}_\mu \mathfrak{g}^*$	space orthogonal to $\mathrm{T}_\mu(G \cdot \mu)$ (6.1)
$\mathrm{forg}\, E(\mu) \subset \mathfrak{g}^*$	image of $E(\mu)$ under identification $\mathrm{T}_\mu \mathfrak{g}^* \cong \mathfrak{g}^*$ (6.1)
$\iota_\mu : [\mathfrak{g}, \mathfrak{g}_\mu]^\circ \to \mathfrak{g}_\mu^*$	restriction of projection $p_\mu : \mathfrak{g}^* \to \mathfrak{g}_\mu^*$ (6.1)
$\mathrm{id}_{\mathfrak{g}^*} \in \Omega^0(\mathfrak{g}^*, \mathfrak{g}^*)$	the identity map $\mathfrak{g}^* \to \mathfrak{g}^*$ (6.4)

1 Introduction

When the body angular momentum of a free rigid body undergoes one period of oscillation the body itself undergoes some overall rotation in the inertial frame of reference. This rotation is an example of a *reconstruction phase*, a notion one may formulate for an arbitrary dynamical system possessing symmetry, whenever the symmetry-reduced variables are undergoing periodic motion. Interest in reconstruction phases stems from problems as diverse as the control of artificial satellites [8] and wave phenomena [3, 2].

This paper studies reconstruction phases in the context of holonomic mechanical systems, from the Hamiltonian point of view. Our results are quite general in the sense that *non-Abelian* symmetries are included; however certain singularities must be avoided. We focus on so-called *simple* mechanical systems (Hamiltonian='kinetic energy' + 'potential energy') but our results are relevant to other Hamiltonian systems on cotangent bundles T^*Q. The primary prerequisite is invariance of the Hamiltonian with respect to the cotangent lift of a free and proper action on the configuration space Q by the symmetry group G. Our results are deduced as a special case of those in [6].

We do not study phases in the context of mechanical control systems and locomotion generation, as in [17] and [15]; nor do we discuss Hanay-Berry phases for 'moving' mechanical systems (such as Foucault's pendulum), as in [16]. Nevertheless, these problems share many features with those studied here and our results may be relevant to generalizations of the cited works.

1.1 Limiting cases

The free rigid body is a prototype for an important class of simple mechanical systems, namely those for which $Q = G$. Those systems whose symmetry group G is *Abelian* constitute another important class, of which the heavy top is a prototype. Reconstruction phases in these two general classes have been studied before [16], [6]. Our general results are essentially a synthesis of these two cases, but because the synthesis is rather sophisticated, detailed results are formulated after reviewing the special cases in Section 2. This introduction describes the new results informally after pointing out key features of the two prototypes. A detailed outline of the paper appears in 1.5 below.

1.2 The free rigid body

In the free (Euler-Poinsot) rigid body reconstruction phases are given by an elegant formula due to Montgomery [23]. Both the configuration space Q and symmetry group G of the free rigid body can be identified with the rotation group SO(3) (see, e.g., [18, Chapter 15]); here we are viewing the body from an inertial reference frame centered on the mass center. Associated with each state x is a spatial angular momentum $\mathbf{J}(x)$ which is conserved. The *body* representation of angular momentum $\nu \in \mathbb{R}^3$ of a state x with configuration

$q \in \mathrm{SO}(3)$ is

$$\nu = q^{-1}\mathbf{J}(x) \ . \tag{1}$$

The body angular momentum ν evolves according to well known equations of Euler which, in particular, constrain solutions to a sphere $\mathcal{O}$ centered at the origin and having radius $\|\mu_0\|$, where $\mu_0 = \mathbf{J}(x_0)$ is the initial spatial angular momentum. This sphere has a well known interpretation as a co-adjoint orbit of $\mathrm{SO}(3)$.

Solutions to Euler's equations are intersections with $\mathcal{O}$ of level sets of the *reduced Hamiltonian* $h : \mathbb{R}^3 \to \mathbb{R}$, given by $h(\nu) \equiv \frac{1}{2}\nu \cdot \mathbb{I}^{-1}\nu$. Here $\mathbb{I}$ denotes the body inertia tensor; see Fig. 1. Typically, a solution $\nu_t \in \mathcal{O}$ is periodic, in

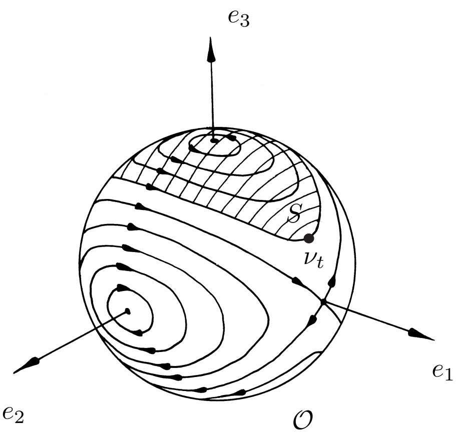


Figure 1: The dynamics of body angular momentum in the free rigid body.

which case (1) implies that $q_T\mu_0 = q_0\mu_0$, where T is the period. This means $q_T = gq_0$ for some rotation $g \in \mathrm{SO}(3)$ about the μ_0-axis. According to [23], the angle $\Delta\theta$ of rotation is given by

$$\Delta\theta = \frac{2Th(\nu_0)}{\|\mu_0\|} - \frac{1}{\|\mu_0\|^2}\int_S dA_{\mathcal{O}} \ , \tag{2}$$

where $S \subset \mathcal{O}$ denotes the region bounded by the curve ν_t (see figure) and $dA_{\mathcal{O}}$ denotes the standard area form on the sphere $\mathcal{O} \subset \mathbb{R}^3$.

Astonishingly, it seems that (2) was unknown to 19th century mathematicians, a vindication of the 'bundle picture' of mechanics promoted in Montgomery's thesis [22].

1.3 The heavy top

Consider a rigid body free to rotate about a point O fixed to the earth (Fig. 2). The configuration space is $Q \equiv \mathrm{SO}(3)$ but full $\mathrm{SO}(3)$ spatial symmetry is broken by gravity (unless O and the center of mass coincide). A residual symmetry group $G \equiv S^1$ acts on Q according to $\theta \cdot q \equiv R^3_\theta q$ ($\theta \in S^1$); here R^3_θ denotes a rotation about the vertical axis e_3 through angle θ.

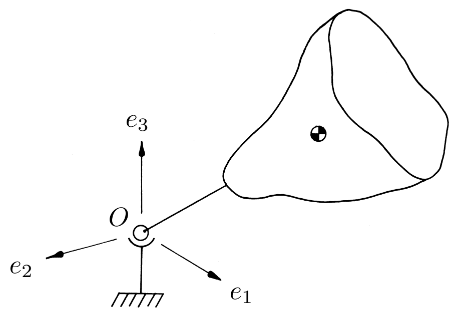


Figure 2: The heavy top.

The quotient space Q/G, known more generally as the *shape space*, is here identifiable with the unit sphere S^2: for a configuration $q \in \mathrm{SO}(3)$ the corresponding 'shape' $r \in S^2$ is the position of the vertical axis viewed in body coordinates:

$$r = q^{-1} e_3 \quad . \tag{1}$$

In place of Euler's rigid body equations one considers the Euler-Poisson heavy top equations [5, (10) & (11), Chapter 1], [18, §15.10]. In the special Lagrange top case these equations are integrable (see, e.g., [4, §30]), but more generally they admit chaotic solutions. In any case, a *periodic* solution to the Euler-Poisson equations determines a periodic solution $r_t \in S^2$ in shape space but the corresponding motion of the body $q_t \in \mathrm{SO}(3)$ need *not* be periodic. However, if T is the period of the given solution to the Euler-Poisson equations, then (1) implies $q_T = R^3_{\Delta\theta} q_0$, for some angle $\Delta\theta$. Assume $r_t \in S^2$ is an embedded curve having T as its *minimal* period. Then

$$\begin{aligned} \Delta\theta &= \int_0^T \frac{dt}{r_t \cdot \mathbb{I} r_t} - \int_S f \, dA_{S^2} \qquad , \\ \text{where} \qquad f(r) &\equiv \frac{\operatorname{Trace} \mathbb{I}}{r \cdot \mathbb{I} r} - \frac{2\mathbb{I} r \cdot \mathbb{I} r}{(r \cdot \mathbb{I} r)^2} \quad . \end{aligned} \tag{2}$$

Here $S \subset S^2$ denotes the region bounded by the curve r_t, dA_{S^2} denotes the standard area form on S^2, and $\mathbb{I}$ denotes the inertia tensor, about O, of the body in its reference configuration ($q = \mathrm{id}$). Equation (2) follows, for instance, from results reviewed in 2.6 and 2.7, together with a curvature calculation along the lines of [16, pp. 48–50].

1.4 General characteristics of reconstruction phases

In both 1.2(2) and 1.3(2) the angle $\Delta\theta$ splits into two parts known as the *dynamic* and *geometric phases*. The dynamic phase amounts to a *time* integral involving the inertia tensor.[1] The geometric phase is a *surface* integral, the integrand depending on the inertia tensor in the case of the heavy top but being independent of system parameters in the case of the free rigid body. Apart from this, an important difference is the *space* in which the phase calculations occur. In the heavy top this is shape space (which is just a point in the free rigid body). In the free rigid body one computes on momentum spheres, i.e., on co-adjoint orbits (which are trivial for the symmetry group S^1 of the heavy top).

As we will show, phases in general mechanical systems are computed in 'twisted products' of shape space Q/G and co-adjoint orbits $\mathcal{O}$, and geometric phases have both a 'shape' and 'momentum' contribution. The source of geometric phases is *curvature*. The 'shape' contribution comes from curvature of a connection $\mathbf{A}$ on Q, bundled over shape space Q/G, constructed using the kinetic energy. This is the so-called *mechanical connection*. The 'momentum' contribution to geometric phases comes from curvature of a connection α_{μ_0} on G, bundled over a co-adjoint orbit $\mathcal{O}$, constructed using an Ad-invariant inner product on the Lie algebra $\mathfrak{g}$ of G. We tentatively refer to this as a *momentum connection*. The mechanical connection depends on the Hamiltonian; the momentum connection is a purely Lie-theoretic object . This explains why system parameters appear explicitly in geometric phases for the heavy top but not in the free rigid body.

In arbitrary simple mechanical systems the dynamic phase is a time integral involving the so-called *locked inertia tensor* $\mathbb{I}$. Roughly speaking, this tensor represents the contribution to the kinetic energy metric coming from symmetry variables. In a system of coupled rigid bodies moving freely through space, it is the inertia tensor about the instantaneous mass center of the rigid body obtained by locking all coupling joints [14, §3.3]

1.5 Paper outline

The new results of this paper are Theorems 3.4 and 3.5 (Section 3). These theorems contain formulas for geometric and dynamic phases in general Hamiltonian systems on cotangent bundles, and in particular for simple mechanical systems. These results are derived as a special case of [6], of which Section 2 is mostly a review. Specifically, Section 2 gives the abstract definition of reconstruction phases, presents a phase formula for systems on arbitrary symplectic manifolds, and surveys the special limiting cases relevant to cotangent bundles. The mechanical connection $\mathbf{A}$, the momentum connection α_{μ_0}, and limiting cases of the locked inertia tensor $\mathbb{I}$ are also defined.

[1] In the free rigid body one has $2Th(\nu_0) = 2Th(\nu_t) = \int_0^T h(\nu_t)\,dt = 2\int_0^T \nu_t \cdot \mathbb{I}^{-1}\nu_t\,dt$.

Section 3 begins by showing how the curvatures of $\mathbf{A}$ and α_{μ_0} can be respectively lifted and extended to structures $\Omega_{\mathbf{A}}$ and Ω_{μ_0} on 'twisted products' of shape space Q/G and co-adjoint orbits $\mathcal{O}$. On these products we also introduce the *inverted locked inertia function* $\xi_{\mathbb{I}}$.

The remainder of the paper is devoted to a proof of Theorems 3.4 and 3.5. Sections 4 and 5 review relevant aspects of cotangent bundle reduction, culminating in an intrinsic formula for symplectic structures on leaves of the Poisson-reduced space $(\mathrm{T}^*Q)/G$. Section 6 builds a natural 'connection' on the symplectic stratification of $(\mathrm{T}^*Q)/G$, and Sections 7 and 8 provide the detailed derivations of dynamic and geometric phases. Appendix A describes the covariant exterior calculus of bundle-valued differential forms, from the point of view of associated bundles.

1.6 Connections to other work

Above what is explicitly cited here, our project owes much to [16]. Additionally, we make crucial use of Cendra, Holm, Marsden and Ratiu's description of reduced spaces in mechanical systems as certain fiber bundle products [9].

In independent work, carried out from the Lagrangian point of view, Marsden, Ratiu amd Scheurle [19] obtain reconstruction phases in mechanical systems with a possibly non-Abelian symmetry group by directly solving appropriate reconstruction equations. Rather than identify separate geometric and dynamic phases, however, their formulas express the phase as a single time integral (no surface integral appears). This integral is along an implicitly defined curve in Q, whereas our formula expresses the phase in terms of 'fully reduced' objects.

The author thanks Matthew Perlmutter for helpful discussions and for making a preliminary version of [24] available.

2 Review

In the setting of Hamiltonian systems on a general symplectic manifold P, reconstruction phases can be expressed by an elegant formula involving derivatives of leaf symplectic structures and the reduced Hamiltonian, these derivatives being computed *transverse* to the symplectic leaves of the Poisson-reduced phase space P/G [6]. This formula, recalled in Theorem 2.3 below, grew out of a desire to 'Poisson reduce' the earlier scheme of Marsden et al. [16, §2A], in which geometric phases were identified with holonomy in an appropriate principal bundle equipped with a connection. Familiarity with this holonomy interpretation is not a prerequisite for understanding and applying Theorem 2.3.

We are ultimately concerned with the special case of cotangent bundles $P = \mathrm{T}^*Q$, and in particular with simple mechanical systems, which are introduced in 2.4. After recalling the definition of the mechanical connection $\mathbf{A}$ in 2.5 we recall the formula for phases in the case of G Abelian (Theorem 2.6 & Addendum 2.7). After introducing the momentum connection α_μ in 2.8 we write down phase formulas for the other limiting case, $Q = G$ (Theorem & Addendum 2.9).

2.1 AN ABSTRACT SETTING FOR RECONSTRUCTION PHASES

Assume G is a connected Lie group acting symplectically from the left on a smooth (C^∞) symplectic manifold (P, ω), and assume the existence of an Ad^*-equivariant momentum map $\mathbf{J} : P \to \mathfrak{g}^*$. (For relevant background, see [14, 1, 18].) Here $\mathfrak{g}$ denotes the Lie algebra of G. Assume G acts freely and properly, and that the fibers of $\mathbf{J}$ are connected. All these hypotheses hold in the case $P = \mathrm{T}^*Q$ when we take G to act by cotangent-lifting a free and proper action on Q and assume Q is connected; details will be recalled in Section 3.
In general, P/G is not a symplectic manifold but merely a Poisson manifold, i.e., a space stratified by lower dimensional symplectic manifolds called *symplectic leaves*; see opi cited. In the free rigid body, for example, one has $P = \mathrm{T}^* \mathrm{SO}(3)$, $G = \mathrm{SO}(3)$, and $P/G \cong \mathfrak{so}(3)^* \cong \mathbb{R}^3$. The symplectic leaves are the co-adjoint orbits, i.e., the spheres centered on the origin.
Let x_t denote an integral curve of the Hamiltonian vector field X_H on P corresponding to some G-invariant Hamiltonian H. Restrict attention to the case that the image curve y_t under the projection $\pi : P \to P/G$ is T-periodic ($T > 0$). Then the associated *reconstruction phase* is the unique $g_{\mathrm{rec}} \in G$ such that $x_T = g_{\mathrm{rec}} \cdot x_0$; see Fig. 3.

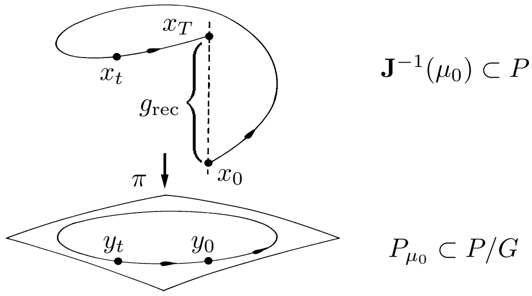


Figure 3: The definition of the reconstruction phase g_{rec}.

Noether's theorem ($\mathbf{J}(x_t) = $ constant) implies that y_t, which is called the *reduced solution*, lies in the *reduced space* P_{μ_0} (see the figure), where

$$P_\mu \equiv \pi(\mathbf{J}^{-1}(\mu)) \subset P/G \qquad (\mu \in \mathfrak{g}^*) \ ,$$

and where $\mu_0 \equiv \mathbf{J}(x_0)$ is the initial momentum. In fact, P_{μ_0} is a symplectic leaf of P/G (see Theorem 5.1) and the Ad^*-equivariance of $\mathbf{J}$ implies $g_{\mathrm{rec}} \in G_{\mu_0}$, where G_{μ_0} is the isotropy of the co-adjoint action at $\mu_0 \in \mathfrak{g}^*$. Invariance of H means $H = h \circ \pi$ for some $h : P/G \to \mathbb{R}$ called the *reduced Hamiltonian*; the reduced solution $y_t \in P_{\mu_0}$ is an integral curve of the Hamiltonian vector field $X_{h_{\mu_0}}$ corresponding to Hamiltonian $h_{\mu_0} \equiv h|P_{\mu_0}$.

2.2 Differentiating across symplectic leaves

We wish to define a kind of derivative in P/G *transverse* to symplectic leaves; these derivatives occur in the phase formula for general Hamiltonian systems to be recalled in 2.3 below. For this we require a notion of infinitesimal transverse. Specifically, if C denotes the characteristic distribution on P/G (the distribution tangent to the symplectic leaves), then a *connection on the symplectic stratification of* P/G is a distribution D on P/G complementary to C: $\mathrm{T}P = C \oplus D$. In that case there is a *canonical two-form* ω_D on P/G determined by D, whose restriction to a symplectic leaf delivers that leaf's symplectic structure, and whose kernel is precisely D.

Below we concern ourselves exclusively with connections D defined in a neighborhood of a nondegenerate symplectic leaf, assuming D to be smooth in the usual sense of constant rank distributions. Then ω_D is smooth also.

Fix a leaf P_μ and assume $D(y)$ is defined for all $y \in P_\mu$. Then at each $y \in P_\mu$ there is, according to the Lemma below, a natural identification of the infinitesimal transverse $D(y)$ with $\mathfrak{g}_\mu^*$, denoted $L(D,\mu,y) : \mathfrak{g}_\mu^* \xrightarrow{\sim} D(y)$.

Now let λ be an arbitrary $\mathbb{R}$-valued p-form on P/G, defined in a neighborhood of P_μ . Then we declare the *transverse derivative* $D_\mu\lambda$ of λ to be the $\mathfrak{g}_\mu$-valued p-form on P_μ defined through

$$\langle \nu, D_\mu\lambda(v_1,\dots,v_p)\rangle = d\lambda\big(L(D,\mu,y)(\nu), v_1,\dots,v_p\big)$$
$$\text{where } \nu \in \mathfrak{g}_\mu^*,\ v_1,\dots,v_p \in \mathrm{T}_y P_\mu \text{ and } y \in P_\mu.$$

Lemma and Definition. *Let $p_\mu : \mathfrak{g}^* \to \mathfrak{g}_\mu^*$ denote the natural projection, and define $\mathrm{T}_{\mathbf{J}^{-1}(\mu)}P \equiv \cup_{x\in\mathbf{J}^{-1}(\mu)}\mathrm{T}_x P$. Fix $y \in P_\mu$ and let $v \in D(y)$ be arbitrary. Then for all $w \in \mathrm{T}_{\mathbf{J}^{-1}(\mu)}P$ such that $\mathrm{T}\pi\cdot w = v$, the value of $p_\mu\langle d\mathbf{J}, w\rangle \in \mathfrak{g}_\mu^*$ is the same. Moreover, the induced map $v \mapsto p_\mu\langle d\mathbf{J}, w\rangle : D(y) \to \mathfrak{g}_\mu^*$ is an isomorphism. The inverse of this isomorphism (which depends on D, μ and y) is denoted by $L(D,\mu,y) : \mathfrak{g}_\mu^* \xrightarrow{\sim} D(y)$.*

We remark that the definition of $L(D,\mu,y)$ is considerably simpler in the case of Abelian G; see [6].

2.3 Reconstruction phases for general Hamiltonian systems

Let $\mathfrak{g}^*_{\mathrm{reg}} \subset \mathfrak{g}^*$ denote the set of regular points of the co-adjoint action, i.e., the set of points lying on co-adjoint orbits of maximal dimension (which fill an open dense subset). If $\mu_0 \in \mathfrak{g}^*_{\mathrm{reg}}$ then $\mathfrak{g}_{\mu_0}$ is Abelian; see Appendix B. In that case G_{μ_0} is Abelian if it is connected.

Now suppose, in the scenario described earlier, that a reduced solution $y_t \in P_{\mu_0}$ bounds a compact oriented surface $\Sigma \subset P_{\mu_0}$.

Theorem (Blaom [6]). *If $\mu_0 \in \mathfrak{g}^*_{\mathrm{reg}}$ and G_{μ_0} is Abelian, then the reconstruc-*

tion phase associated with the periodic solution $y_t \in \partial\Sigma$ is

$$g_{\mathrm{rec}} = g_{\mathrm{dyn}} g_{\mathrm{geom}} \ , \qquad \textit{where:}$$
$$g_{\mathrm{dyn}} = \exp \int_0^T D_{\mu_0} h(y_t)\, dt \ , \qquad g_{\mathrm{geom}} = \exp \int_\Sigma D_{\mu_0} \omega_D \ .$$

Here h denotes the reduced Hamiltonian, D denotes an arbitrary connection on the symplectic stratification of P/G, ω_D denotes the canonical two-form on P/G determined by$D$, and D_{μ_0} denotes the transverse derivative operator determined by D as described above.
The Theorem states that dynamic phases are time integrals of transverse derivatives of the reduced Hamiltonian while geometric phases are surface integrals of transverse derivatives of leaf symplectic structures.
We emphasize that while g_{dyn} and g_{geom} depend on the choice of D, the total phase g_{rec} is, by definition, independent of any such choice.
For the application of the above to non-free actions see [6].

2.4 Simple mechanical systems

Suppose a connected Lie group G acts freely and properly on a connected manifold Q. All actions in this paper are understood to be *left* actions. A Hamiltonian $H : \mathrm{T}^*Q \to \mathbb{R}$ is said to enjoy *G-symmetry* if it is invariant with respect to the cotangent-lifted action of G on T^*Q (see [1, p. 283] for the definition of this action). This action admits an Ad^*-equivariant momentum map $\mathbf{J} : \mathrm{T}^*Q \to \mathfrak{g}^*$ defined through

$$\langle \mathbf{J}(x), \xi \rangle \equiv \langle x, \xi^Q(q) \rangle \qquad (x \in \mathrm{T}^*_q Q,\ q \in Q,\ \xi \in \mathfrak{g}) \ , \tag{1}$$

where ξ^Q denotes the infinitesimal generator on Q corresponding to ξ. A *simple mechanical system* is a Hamiltonian $H : \mathrm{T}^*Q \to \mathbb{R}$ of the form

$$H(x) = \frac{1}{2} \langle\!\langle x, x \rangle\!\rangle_Q^* + V(q) \qquad (x \in \mathrm{T}^*_q Q) \ .$$

Here $\langle\!\langle \cdot\,, \cdot \rangle\!\rangle_Q^*$ denotes the symmetric contravariant two-tensor on Q determined by some prescribed Riemannian metric $\langle\!\langle \cdot\,, \cdot \rangle\!\rangle_Q$ on Q (the *kinetic energy metric*), and V is some prescribed G-invariant function on Q (the *potential energy*). To ensure G-symmetry we are supposing that G acts on Q by $\langle\!\langle \cdot\,, \cdot \rangle\!\rangle_Q$-isometries.

2.5 Mechanical connections

In general, the configuration space Q is bundled in a topologically non-trivial way over shape space Q/G, i.e., there is no global way to separate shape variables from symmetry variables. However, fixing a connection on the bundle allows one to split individual motions. In the case of simple mechanical systems such a connection is determined by the kinetic energy, but in general there is no canonical choice. All the phase formulas we shall present assume some choice has been made.

Under our free and properness assumptions, the projection $\rho : Q \to Q/G$ is a principal G-bundle. So we will universally require that this bundle be equipped with a connection one-form $\mathbf{A} \in \Omega^1(Q, \mathfrak{g})$. If a G-invariant Riemannian metric on Q is prescribed (e.g., the kinetic energy in the case of simple mechanical systems) a connection $\mathbf{A}$ is determined by requiring that the corresponding distribution of horizontal spaces hor $\equiv \ker \mathbf{A}$ are orthogonal to the ρ-fibers (G-orbits). In this context, $\mathbf{A}$ is called the *mechanical connection*; its history is described in [14, §3.3]

As we shall recall in 4.2, a connection $\mathbf{A}$ on $\rho : Q \to Q/G$ allows one to construct a Hamiltonian analogue $\mathrm{T}^*_{\mathbf{A}}\rho : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G)$ for the tangent map $\mathrm{T}\rho : \mathrm{T}Q \to \mathrm{T}(Q/G)$. Thus for a state $x \in \mathrm{T}^*_q Q$ one may speak of the 'generalized momentum' $\mathrm{T}^*_{\mathbf{A}}\rho \cdot x \in \mathrm{T}^*_r(Q/G)$ of the corresponding shape $r = \rho(q) \in Q/G$.

2.6 Phases for Abelian symmetries

Let $H : \mathrm{T}^*Q \to \mathbb{R}$ be an arbitrary Hamiltonian enjoying G-symmetry. When G is Abelian it is known that each reduced space P_μ ($\mu \in \mathfrak{g}^*$, $P = \mathrm{T}^*Q$) is isomorphic to $\mathrm{T}^*(Q/G)$ equipped with the symplectic structure

$$\omega_\mu = \omega_{Q/G} - \langle \mu, (\tau^*_{Q/G})^* \operatorname{curv} \mathbf{A} \rangle \ .$$

It should be emphasized that the identification $P_\mu \cong \mathrm{T}^*(Q/G)$ depends on the choice of connection $\mathbf{A}$. See, e.g., [6] for the details. In the above equation $\omega_{Q/G}$ denotes the canonical symplectic structure on $\mathrm{T}^*(Q/G)$ and $\tau^*_{Q/G} : \mathrm{T}^*(Q/G) \to Q/G$ is the usual projection; $\operatorname{curv} \mathbf{A}$ denotes the curvature of $\mathbf{A}$, viewed as a $\mathfrak{g}$-valued two-form on Q/G (see, e.g., [16, §4]). The value of the reduced Hamiltonian $h_\mu : \mathrm{T}^*(Q/G) \to \mathbb{R}$ at a point $y \in \mathrm{T}^*(Q/G)$ is $H(x)$ where $x \in \mathrm{T}^*Q$ is any point satisfying $\mathbf{J}(x) = \mu$ and $\mathrm{T}^*_{\mathbf{A}}\rho \cdot x = y$.

The Theorem below is implicit in [6]. The special case in Addendum 2.7 is due to Marsden et al [16] (explicitly appearing in [6]).

Theorem. *Let $y_t \in P_{\mu_0} \cong \mathrm{T}^*(Q/G)$ be a periodic reduced solution curve. Let $r_t \equiv \tau^*_{Q/G}(y_t) \in Q/G$ denote the corresponding curve in shape space. Assume $t \mapsto r_t$ bounds a compact oriented surface $S \subset Q/G$. Assume r_t and y_t have the same minimal period T. Then the reconstruction phase associated with y_t is*

$$g_{\mathrm{rec}} = g_{\mathrm{dyn}} g_{\mathrm{geom}} \ , \qquad \textit{where:}$$

$$g_{\mathrm{dyn}} = \exp \int_0^T \frac{\partial h}{\partial \mu}(\mu_0, y_t)\, dt \ , \qquad g_{\mathrm{geom}} = \exp\left(- \int_S \operatorname{curv} \mathbf{A} \right) \ ,$$

and where $\partial h / \partial \mu \,(\mu', y') \in \mathfrak{g}$ is defined through

$$\langle \nu, \frac{\partial h}{\partial \mu}(\mu', y') \rangle = \frac{d}{dt} h_{\mu' + t\nu}(y') \Big|_{t=0} \qquad (\, \nu, \mu' \in \mathfrak{g}^*,\ y' \in \mathrm{T}^*(Q/G)\,) \ .$$

Here $\mathbf{A}$ *denotes an arbitrary connection on* $Q \to Q/G$*.*

2.7 Locked inertia tensor (Abelian case)

In the special case of a simple mechanical system one may be explicit about the dynamic phase. To this end, define for each $q \in Q$ a map $\hat{\mathbb{I}}(q) : \mathfrak{g} \to \mathfrak{g}^*$ through

$$\langle \hat{\mathbb{I}}(q)(\xi), \eta \rangle = \langle\!\langle \xi^Q(q), \eta^Q(q) \rangle\!\rangle_Q \qquad (\xi, \eta \in \mathfrak{g}) \ ,$$

where ξ^Q denotes the infinitesimal generator on Q corresponding to ξ. Varying over all $q \in Q$, one obtains a function $\hat{\mathbb{I}} : Q \to \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}^*)$. When G is Abelian, $\hat{\mathbb{I}}$ is G-invariant, dropping to a function $\mathbb{I} : Q/G \to \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}^*)$ called the *locked inertia tensor* (terminology explained in 1.4). As G acts freely on Q, $\hat{\mathbb{I}}(q) : \mathfrak{g} \to \mathfrak{g}^*$ has an inverse $\hat{\mathbb{I}}(q)^{-1} : \mathfrak{g}^* \to \mathfrak{g}$ leading to functions $\hat{\mathbb{I}}^{-1} : Q \to \mathrm{Hom}(\mathfrak{g}^*, \mathfrak{g})$ and $\mathbb{I}^{-1} : Q/G \to \mathrm{Hom}(\mathfrak{g}^*, \mathfrak{g})$.

Addendum. *When* $H : \mathrm{T}^*Q \to \mathbb{R}$ *is a simple mechanical system and* $\mathbf{A}$ *is the mechanical connection, then the dynamic phase appearing in the preceding Theorem is given by*

$$g_{\mathrm{dyn}} = \int_0^T \mathbb{I}^{-1}(r_t)\mu_0 \, dt \ .$$

In particular, the reconstruction phase g_{rec} *is computed entirely in the shape space* Q/G*.*

2.8 Momentum connections

In the rigid body example discussed in 1.2 ($G = \mathrm{SO}(3)$), the angle $\Delta\theta$ may be identified with an element of $\mathfrak{g}_{\mu_0}$, where $\mu_0 \in \mathfrak{g}^* \cong \mathbb{R}^3$ is the initial spatial angular momentum. This angle is the logarithm of the reconstruction phase $g_{\mathrm{rec}} \in G_{\mu_0}$, there denoted g. Let $\omega^-_{\mathcal{O}}$ denote the 'minus' version of the symplectic structure on $\mathcal{O}$, viewed as co-adjoint orbit (see below). Then Equation 1.2(2) may alternatively be written

$$(1) \qquad \langle \mu_0, \log g_{\mathrm{rec}} \rangle = 2Th(\nu_0) + \int_S \omega^-_{\mathcal{O}} \ .$$

As we shall see, this generalizes to arbitrary groups G, but it refers only to the μ_0-component of the log phase. This engenders the following question, answered in the Proposition below: Of what $\mathfrak{g}_{\mu_0}$-valued two-form on $\mathcal{O}$ is $\omega^-_{\mathcal{O}}$ the μ_0-component?

For an arbitrary connected Lie group G equip $\mathfrak{g}^*$ with the 'minus' Lie-Poisson structure (see, e.g., [14, §2.8]). The symplectic leaves are the co-adjoint orbits; the symplectic structure on an orbit $\mathcal{O} = G \cdot \mu_0$ is $\omega^-_{\mathcal{O}}$, where $\omega^-_{\mathcal{O}}$ is given implicitly by

$$(2) \qquad \omega^-_{\mathcal{O}}\left(\frac{d}{dt} \exp(t\xi_1) \cdot \mu \Big|_{t=0}, \frac{d}{dt} \exp(t\xi_2) \cdot \mu \Big|_{t=0} \right) = -\langle \mu, [\xi_1, \xi_2] \rangle \ ,$$

for arbitrary $\mu \in \mathcal{O}$ and $\xi_1, \xi_2 \in \mathfrak{g}$. The map $\tau_{\mu_0} : G \to \mathcal{O}$ sending g to $g^{-1} \cdot \mu_0$ is a principal G_{μ_0}-bundle. If we denote by $\theta_G \in \Omega^1(G, \mathfrak{g})$ the right-invariant Mauer-Cartan form on G, then (2) may be succinctly written

$$\tau_{\mu_0}^* \omega_{\mathcal{O}}^- = -\langle \mu_0, \frac{1}{2}\theta_G \wedge \theta_G \rangle \ . \tag{3}$$

Assuming $\mathfrak{g}$ admits an Ad-invariant inner product, the bundle $\tau_{\mu_0} : G \to \mathcal{O} \cong G/G_{\mu_0}$ comes equipped with a connection one-form $\alpha_{\mu_0} \equiv \langle \mathrm{pr}_{\mu_0}, \theta_G \rangle$; here $\mathrm{pr}_{\mu_0} : \mathfrak{g} \to \mathfrak{g}_{\mu_0}$ denotes the orthogonal projection. We shall refer to α_{μ_0} as the *momentum connection* on $G \to \mathcal{O} \cong G/G_{\mu_0}$.
For simplicity, assume that μ_0 lies in $\mathfrak{g}^*_{\mathrm{reg}}$ and that G_{μ_0} is Abelian, as in 2.3. Then the curvature of α_{μ_0} may be identified with a $\mathfrak{g}_{\mu_0}$-valued two-form on $\mathcal{O} = G \cdot \mu_0$ denoted $\mathrm{curv}\, \alpha_{\mu_0}$.

Proposition. *Under the above conditions*

$$(\mathrm{curv}\, \alpha_{\mu_0}) \left(\frac{d}{dt} \exp(t\xi_1\,) \cdot \mu \Big|_{t=0}, \frac{d}{dt} \exp(t\xi_2\,) \cdot \mu \Big|_{t=0} \right) = \mathrm{pr}_{\mu_0} g^{-1} \cdot [\xi_1, \xi_2] \ ,$$

where g is any element of G such that $\mu = g \cdot \mu_0$, and $\xi_1, \xi_2 \in \mathfrak{g}$ are arbitrary. In particular, $\omega_{\mathcal{O}}^-$ is a component of curvature: $\omega_{\mathcal{O}}^- = -\langle \mu_0, \mathrm{curv}\, \alpha_{\mu_0} \rangle$.

Proof. Because G_{μ_0} is assumed Abelian, we have $\tau_{\mu_0}^* \mathrm{curv}\, \alpha_{\mu_0} = d\alpha_{\mu_0} = \langle \mathrm{pr}_{\mu_0}, d\theta_G \rangle$. Applying the Mauer-Cartan identity $d\theta_G = \frac{1}{2}\theta_G \wedge \theta_G$, we obtain $\tau_{\mu_0}^* \mathrm{curv}\, \alpha_{\mu_0} = \langle \mathrm{pr}_{\mu_0}, \frac{1}{2}\theta_G \wedge \theta_G \rangle$, which implies both the first part of the Proposition and the identity $\tau_{\mu_0}^* \langle \mu_0, \mathrm{curv}\, \alpha_{\mu_0} \rangle = \langle \mu_0 \circ \mathrm{pr}_{\mu_0}, \frac{1}{2}\theta_G \wedge \theta_G \rangle$. But $\mu_0 \in \mathfrak{g}^*_{\mathrm{reg}}$ implies that the space $\mathfrak{g}_{\mu_0}^\perp$ orthogonal to $\mathfrak{g}_{\mu_0}$ coincides with $[\mathfrak{g}, \mathfrak{g}_{\mu_0}]$ (see Appendix B), implying $\langle \mu_0, \mathrm{pr}_{\mu_0} \xi \rangle = \langle \mu_0, \xi \rangle$ for all $\xi \in \mathfrak{g}$. So $\tau_{\mu_0}^* \langle \mu_0, \mathrm{curv}\, \alpha_{\mu_0} \rangle = \langle \mu_0, \frac{1}{2}\theta_G \wedge \theta_G \rangle \rangle = -\tau_{\mu_0}^* \omega_{\mathcal{O}}^-$, by (3). This implies $\omega_{\mathcal{O}}^- = -\langle \mu_0, \mathrm{curv}\, \alpha_{\mu_0} \rangle$. $\square$

2.9 Phases for $Q = G$

When $Q = G$, the Poisson manifold $P/G = (\mathrm{T}^*G)/G$ is identifiable with $\mathfrak{g}^*$ and the reduced space P_{μ_0} is the co-adjoint orbit $\mathcal{O} \equiv G \cdot \mu_0$, equipped with the symplectic structure $\omega_{\mathcal{O}}^-$ discussed above. Continue to assume that $\mathfrak{g}$ admits an Ad-invariant inner product. As we will show in Proposition 6.1, the restriction $\iota_{\mu_0} : [\mathfrak{g}, \mathfrak{g}_{\mu_0}]^\circ \to \mathfrak{g}^*_{\mu_0}$ of the natural projection $p_{\mu_0} : \mathfrak{g}^* \to \mathfrak{g}^*_{\mu_0}$ is then an isomorphism, assuming $\mu_0 \in \mathfrak{g}^*_{\mathrm{reg}}$. Here $^\circ$ denotes annihilator. The following result is implicit in [6].

Theorem. *Assume $\mu_0 \in \mathfrak{g}^*_{\mathrm{reg}}$ and G_{μ_0} is Abelian. Let $\nu_t \in P_{\mu_0} \cong \mathcal{O} \equiv G \cdot \mu_0$ be a periodic reduced solution curve bounding a compact oriented surface $S \subset \mathcal{O}$. Let $g_t \in G$ be any curve such that $\nu_t = g_t \cdot \mu_0 \equiv \mathrm{Ad}^*_{g_t^{-1}} \mu_0$. Then the* *reconstruction phase associated with ν_t is given by*

$$g_{\text{rec}} = g_{\text{dyn}} g_{\text{geom}} \ , \qquad \textit{where:}$$
$$g_{\text{dyn}} = \exp \int_0^T w(t)\, dt \ , \qquad g_{\text{geom}} = \exp\left(-\int_S \operatorname{curv} \alpha_{\mu_0} \right) \ ,$$

and where $w(t) \in \mathfrak{g}_{\mu_0}$ is defined through

$$\langle \lambda, w(t) \rangle = \frac{d}{d\tau} h\Big(g_t \cdot (\, \mu_0 + \tau \iota_{\mu_0}^{-1}(\lambda)\,)\Big) \Big|_{\tau=0} \qquad (\lambda \in \mathfrak{g}_{\mu_0}^*) \ .$$

Here α_{μ_0} denotes the momentum connection on $G \to \mathcal{O} \cong G/G_{\mu_0}$.

For a simple mechanical system on T^*G the reduced Hamiltonian $h : \mathfrak{g}^* \to \mathbb{R}$ is of the form

$$h(\nu) = \frac{1}{2} \langle \nu, \mathbb{I}^{-1} \nu \rangle \ , \qquad (\nu \in \mathfrak{g}^*)$$

for some isomorphism $\mathbb{I} : \mathfrak{g} \xrightarrow{\sim} \mathfrak{g}^*$, the *inertia tensor*, which we may suppose is symmetric as an element of $\mathfrak{g}^* \otimes \mathfrak{g}^*$.

Addendum ([6]). *Let G act on $\operatorname{Hom}(\mathfrak{g}^*, \mathfrak{g})$ via conjugation, so that $g \cdot \mathbb{I}^{-1} = \operatorname{Ad}_{g^{-1}} \circ \mathbb{I}^{-1} \circ \operatorname{Ad}_{g^{-1}}^*$ $(g \in G)$. Then for a simple mechanical system one has*

$$w(t) = \operatorname{pr}_{\mu_0} \Big((g_t^{-1} \cdot \mathbb{I}^{-1})(\mu_0) \Big) \ ,$$

where $\operatorname{pr}_{\mu_0} : \mathfrak{g} \to \mathfrak{g}_{\mu_0}$ is the orthogonal projection. Moreover the generalization 2.8(1) of Montgomery's rigid body formula holds.

3 Formulation of new results

According to known results reviewed in the preceding section, phases for simple mechanical systems are computed in shape space Q/G when G is Abelian, and on a co-adjoint orbit $\mathcal{O} = G \cdot \mu_0$ when $Q = G$. For the general case, G non-Abelian and $Q \neq G$, we need to introduce the concepts of associated bundles and forms, and the locked inertia tensor for non-Abelian groups (3.1–3.3). In 3.4 and 3.5 we present the main results of the paper, namely explicit formulas for geometric and dynamic phases in Hamiltonian systems on cotangent bundles.

3.1 Associated bundles

Given an arbitrary principal bundle $\rho : Q \to Q/G$ and manifold $\mathcal{O}$ on which G acts, we denote the quotient of $Q \times \mathcal{O}$ under the diagonal action of G by $\mathcal{O}_Q$. This is the total space of a bundle $\rho_{\mathcal{O}} : \mathcal{O}_Q \to Q/G : [q, \nu]_G \mapsto [q]_G$ known as the *associated bundle* for $\mathcal{O}$. As its fibers are diffeomorphic to $\mathcal{O}$, it may be regarded as a 'twisted product' of Q/G and $\mathcal{O}$.

Here the important examples will be the *co-adjoint bundle* $\mathfrak{g}_Q^*$ and the *co-adjoint orbit bundle* $\mathcal{O}_Q \subset \mathfrak{g}_Q^*$, where $\mathcal{O} \subset \mathfrak{g}^*$ is a co-adjoint orbit.

We have seen that log geometric phases are surface integrals of the curvature $\operatorname{curv} \mathbf{A} \in \Omega^2(Q/G, \mathfrak{g})$ of the mechanical connection $\mathbf{A}$, when G is Abelian, and of the curvature $\operatorname{curv} \alpha_{\mu_0} \in \Omega^2(\mathcal{O}, \mathfrak{g}_{\mu_0})$ of the momentum connection α_{μ_0}, when $Q = G$. For simple mechanical systems the log dynamic phase is a time integral of an inverted inertia tensor $\mathbb{I}^{-1}$ in both cases. To elaborate on the claims regarding the general case made in 1.4, we need to see how $\operatorname{curv} \mathbf{A}$, $\operatorname{curv} \alpha_{\mu_0}$ and $\mathbb{I}^{-1}$ can be viewed as objects on $\mathcal{O}_Q$.
A non-Abelian G forces us to regard $\operatorname{curv} \mathbf{A}$ as an element of $\Omega^2(Q/G, \mathfrak{g}_Q)$, i.e., as *bundle*-valued. See, e.g., Note A.6 and A.2(1) for the definition. The pull-back $\rho_{\mathcal{O}}^* \operatorname{curv} \mathbf{A}$ is then a two-form on $\mathcal{O}_Q$, but with values in the pull-back bundle $\rho_{\mathcal{O}}^* \mathfrak{g}_Q$. Pull-backs of bundles and forms are briefly reviewed in Appendix A.
On the other hand, $\operatorname{curv} \alpha_{\mu_0}$ is *vector*-valued because $\mathfrak{g}_{\mu_0}$ is Abelian under the hypothesis $\mu_0 \in \mathfrak{g}_{\mathrm{reg}}^*$. It is a two-form on the model space $\mathcal{O}$ of the fibers of $\rho_{\mathcal{O}} : \mathcal{O}_Q \to Q/G$. Its natural 'extension' to a two-form on $\mathcal{O}_Q$ is the *associated form* $(\operatorname{curv} \alpha_{\mu_0})_Q \in \Omega^2(\mathcal{O}_Q, \mathfrak{g}_{\mu_0})$, which we now define more generally.

3.2 Associated forms

Let $\rho : Q \to Q/G$ be a principal bundle equipped with a connection $\mathbf{A}$, and let $\mathcal{O}$ be a manifold on which G acts. If λ is a G-invariant, $\mathbb{R}$-valued k-form on $\mathcal{O}$ then the *associated form* λ_Q is the $\mathbb{R}$-valued k-form on $\mathcal{O}_Q$ defined as follows: For arbitrary $u_1, \ldots, u_k \in \mathrm{T}_{[q,\nu]_G} \mathcal{O}_Q$, there exist $\mathbf{A}$-horizontal curves $t \mapsto q_i^{\mathrm{hor}}(t) \in Q$ through q, and curves $t \mapsto \nu_i(t) \in Q$ through ν, such that

$$u_i = \frac{d}{dt} [q_i^{\mathrm{hor}}(t), \nu_i(t)]_G \Big|_{t=0} ,$$

in which case λ_Q is well defined by

$$\lambda_Q(u_1, \ldots, u_k) = \lambda \left(\frac{d}{dt} \nu_1(t) \Big|_{t=0}, \ldots, \frac{d}{dt} \nu_k(t) \Big|_{t=0} \right) . \tag{1}$$

When $\mathbb{R}$ is replaced by a general vector space V on which G acts linearly, then the *associated form* λ_Q of a *G-equivariant*, V-valued k-form λ is a certain k-form on $\mathcal{O}_Q$ taking values in the pull-back bundle $\rho_{\mathcal{O}}^* V_Q$. Its definition is postponed to 7.2. In symbols, we have a map

$$\begin{gathered} \lambda \mapsto \lambda_Q \\ \Omega_G^k(\mathcal{O}, V) \to \Omega^k(\mathcal{O}_Q, \rho_{\mathcal{O}}^* V_Q) \ . \end{gathered}$$

The identity $(\lambda \wedge \mu)_Q = \lambda_Q \wedge \mu_Q$ holds. If G acts trivially on V (e.g., $V = \mathbb{R}$ or $\mathfrak{g}_{\mu_0}$), then $\rho_{\mathcal{O}}^* V_Q \cong \mathcal{O}_Q \times V$ and we identify λ_Q with a V-valued form on

$\mathcal{O}_Q$ and (1) holds.

This last remark applies, in particular, to $\operatorname{curv}\alpha_{\mu_0}$.

3.3 Locked inertia tensor (general case)

When G is non-Abelian the map $\hat{\mathbb{I}} : Q \to \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}^*)$ defined in 2.7 is G-equivariant if G acts on $\mathrm{Hom}(\mathfrak{g}, \mathfrak{g}^*)$ via conjugation. It therefore drops to a (bundle-valued) function $\mathbb{I} \in \Omega^0(Q/G, \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}^*)_Q)$, the *locked inertia tensor*:

$$\mathbb{I}([q]_G) \equiv [q, \hat{\mathbb{I}}(q)]_G \ .$$

The inverse $\mathbb{I}^{-1} \in \Omega^0(Q/G, \mathrm{Hom}(\mathfrak{g}^*, \mathfrak{g})_Q)$ is defined similarly.
View the inclusion $i_\mathcal{O} : \mathcal{O} \hookrightarrow \mathfrak{g}^*$ as an element of $\Omega^0(\mathcal{O}, \mathfrak{g}^*)$. Then with the help of the associated form $(i_\mathcal{O})_Q \in \Omega^0(\mathcal{O}_Q, \rho_\mathcal{O}^* \mathfrak{g}_Q^*)$ one obtains a function $\rho_\mathcal{O}^* \mathbb{I}^{-1} \wedge (i_\mathcal{O})_Q$ on $\mathcal{O}_Q$ taking values in $\rho_\mathcal{O}^* \mathfrak{g}_Q$. (Under the canonical identification $\rho_\mathcal{O}^* \mathfrak{g}_Q^* \cong \mathcal{O}_Q \oplus \mathfrak{g}_Q^*$, one has $(i_\mathcal{O})_Q(\eta) = \eta \oplus \eta$.) Here the wedge $\wedge$ implies a contraction $\mathrm{Hom}(\mathfrak{g}^*, \mathfrak{g}) \otimes \mathfrak{g}^* \to \mathfrak{g}$.

3.4 Phases for simple mechanical systems

Before stating our new results, let us summarize with a few definitions. Put

$\Omega_\mathbf{A} \equiv \rho_\mathcal{O}^* \operatorname{curv} \mathbf{A}$:	the *mechanical curvature,*
$\Omega_{\mu_0} \equiv (\operatorname{curv} \alpha_{\mu_0})_Q$:	the *momentum curvature,*
$\xi_\mathbb{I} \equiv \rho_\mathcal{O}^* \mathbb{I}^{-1} \wedge (i_\mathcal{O})_Q$:	the *inverted locked inertia function.*

Recall here that $\mathbf{A}$ denotes a connection on $Q \to Q/G$ (the mechanical connection if H is a simple mechanical system), α_{μ_0} denotes the momentum connection on $G \to \mathcal{O} \cong G/G_{\mu_0}$, $\rho_\mathcal{O} : \mathcal{O}_Q \to Q/G$ denotes the associated bundle projection and $i_\mathcal{O} \in \Omega^0(\mathcal{O}, \mathfrak{g}^*)$ denotes the inclusion $\mathcal{O} \hookrightarrow \mathfrak{g}^*$.
By construction, $\Omega_\mathbf{A}$, Ω_{μ_0} and $\xi_\mathbb{I}$ are all differential forms on $\mathcal{O}_Q$. The momentum curvature Ω_{μ_0} is $\mathfrak{g}_{\mu_0}$-valued, and can therefore be integrated over surfaces $S \subset \mathcal{O}_Q$; the forms $\Omega_\mathbf{A}$ and $\xi_\mathbb{I}$ are $\rho_\mathcal{O}^* \mathfrak{g}_Q$-valued. To make them $\mathfrak{g}_{\mu_0}$-valued requires an appropriate projection:

Definition. Let G act on $\mathrm{Hom}(\mathfrak{g}, \mathfrak{g}_{\mu_0})$ via $g \cdot \sigma \equiv \mathrm{Ad}_g \circ \sigma$ and let $\mathrm{Pr}_{\mu_0} \in \Omega^0(\mathcal{O}, \mathrm{Hom}(\mathfrak{g}, \mathfrak{g}_{\mu_0}))$ denote the unique equivariant zero-form whose value at μ_0 is the orthogonal projection $\mathrm{pr}_{\mu_0} : \mathfrak{g} \to \mathfrak{g}_{\mu_0}$.

With the help of the associated form $(\mathrm{Pr}_{\mu_0})_Q$ and an implied contraction $\mathrm{Hom}(\mathfrak{g}, \mathfrak{g}_{\mu_0}) \otimes \mathfrak{g} \to \mathfrak{g}_{\mu_0}$, we obtain $\rho_\mathcal{O}^*(\mathfrak{g}_{\mu_0})_Q$-valued forms $(\mathrm{Pr}_{\mu_0})_Q \wedge \Omega_\mathbf{A}$ and $(\mathrm{Pr}_{\mu_0})_Q \wedge \xi_\mathbb{I}$. As we declare G to act trivially on $\mathfrak{g}_{\mu_0}$, these forms are in fact identifiable with $\mathfrak{g}_{\mu_0}$-valued forms as required.
For $P = \mathrm{T}^*Q$ and G non-Abelian the reduced space P_{μ_0} can be identified with $\mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$, where $\mathcal{O} \equiv G \cdot \mu_0$. Here $\oplus$ denotes product in the category of

fiber bundles over Q/G (see Notation in 4.2). This observation was first made in the Lagrangian setting by Cendra et al. [9]. We recall details in 4.2 and Proposition 5.1. A formula for the symplectic structure on P_{μ_0} has been given by Perlmutter [24]. We derive the form of it we will require in 5.2. The value of the reduced Hamiltonian $h_{\mu_0} : P_{\mu_0} \to \mathbb{R}$ at $z \oplus [q,\mu]_G \in \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$ is $H(x)$, where $x \in \mathrm{T}^*_q Q$ is any point satisfying $\mathrm{T}^*_{\mathbf{A}}\rho \cdot x = z$ and $\mathbf{J}(x) = \mu$.
In the case of simple mechanical systems one has

$$h_{\mu_0}(z \oplus [q,\mu]_G) = \frac{1}{2}\langle\!\langle z, z\rangle\!\rangle^*_{Q/G} + \frac{1}{2}\langle \mu, \hat{\mathbb{I}}^{-1}(q)\mu\rangle + V_{Q/G}(\rho(q)) \ . \tag{1}$$

Here $V_{Q/G}$ denotes the function on Q/G to which the potential V drops on account of its G-invariance, and $\langle\!\langle \cdot\,,\,\cdot \rangle\!\rangle^*_{Q/G}$ denotes the symmetric contravariant two-tensor on Q/G determined by the Riemannian metric $\langle\!\langle \cdot\,,\,\cdot \rangle\!\rangle_{Q/G}$ that Q/G inherits from the G-invariant metric $\langle\!\langle \cdot\,,\,\cdot \rangle\!\rangle_Q$ on Q. (The second term above may be written intrinsically as $1/2\ ((\mathrm{id}_{\mathfrak{g}^*})_Q \wedge (\rho^*_{\mathfrak{g}^*}\mathbb{I}^{-1} \wedge (\mathrm{id}_{\mathfrak{g}^*})_Q))([q,\mu]_G)$, where $(\mathrm{id}_{\mathfrak{g}^*})_Q$ is defined in 6.4.) The formula (1) is derived in 7.1.

Theorem. *Let $H : \mathrm{T}^*Q \to \mathbb{R}$ be a simple mechanical system, as defined in 2.4. Assume $\mu_0 \in \mathfrak{g}^*_{\mathrm{reg}}$, G_{μ_0} is Abelian, and let $z_t \oplus \eta_t \in P_{\mu_0} \cong \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$ ($\mathcal{O} = G \cdot \mu_0$) denote a periodic reduced solution curve. Assume $z_t \oplus \eta_t$ and η_t have the same minimal period T and assume $t \mapsto \eta_t$ bounds a compact oriented surface $S \subset \mathcal{O}_Q$. Then the corresponding reconstruction phase is*

$$\begin{aligned} g_{\mathrm{rec}} &= g_{\mathrm{dyn}} g_{\mathrm{geom}} \ , \qquad \textit{where} \\ g_{\mathrm{dyn}} &= \exp \int_0^T (\mathrm{Pr}_{\mu_0})_Q \wedge \xi_{\mathbb{I}}\,(\eta_t)\, dt \ , \\ g_{\mathrm{geom}} &= \exp\left(- \int_S (\,\Omega_{\mu_0} + (\mathrm{Pr}_{\mu_0})_Q \wedge \Omega_{\mathbf{A}}\,) \right) \ . \end{aligned}$$

Here $\Omega_{\mathbf{A}}$ is the mechanical curvature, Ω_{μ_0} the momentum curvature, and $\xi_{\mathbb{I}}$ the inverted locked inertia function, as defined above; $\mathbf{A}$ denotes the mechanical connection.

Notice that the phase g_{rec} does not depend on the z_t part of the reduced solution curve (z_t, η_t), i.e., is computed exclusively in the space $\mathcal{O}_Q$.

3.5 Phases for arbitrary systems on cotangent bundles

We now turn to the case of general Hamiltonian functions on T^*Q (not necessarily simple mechanical systems). To formulate results in this case, we need the fact, recalled in Theorem 4.2, that $(\mathrm{T}^*Q)/G$ is isomorphic to $\mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q$, where $\oplus$ denotes product in the category of fiber bundles over Q/G (see Notation 4.2). This isomorphism depends on the choice of connection $\mathbf{A}$ on $\rho : Q \to Q/G$.

Theorem. *Let $H : \mathrm{T}^*Q \to \mathbb{R}$ be an arbitrary G-invariant Hamiltonian and $h : \mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q \to \mathbb{R}$ the corresponding reduced Hamiltonian. Consider a periodic reduced solution curve $z_t \oplus \eta_t \in P_{\mu_0} \cong \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$, as in the Theorem above. Then the conclusion of that Theorem holds, with the dynamic phase now given by*

$$g_{\mathrm{dyn}} = \exp \int_0^T D_{\mu_0} h(z_t \oplus \eta_t)\, dt \ ,$$

where $D_{\mu_0} h(\,\cdot\,) \in \mathfrak{g}_{\mu_0}$ is defined through

$$(1)\ \ \langle \nu, D_{\mu_0} h(z \oplus [q, \mu_0]_G)\rangle = \frac{d}{dt} h(z \oplus [q, \mu_0 + t\iota_{\mu_0}^{-1}(\nu)]_G)\Big|_{t=0} \qquad (\nu \in \mathfrak{g}^*_{\mu_0}) \ .$$

Here $\iota_{\mu_0} : [\mathfrak{g}, \mathfrak{g}_{\mu_0}]^\circ \xrightarrow{\sim} \mathfrak{g}^*_{\mu_0}$ is the isomorphism defined in 2.9.

Theorems 3.4 and 3.5 will be proved in Sections 7 and 8.

4 Symmetry reduction of cotangent bundles

In this section and the next, we revisit the process of reduction in cotangent bundles by describing the symplectic leaves in the associated Poisson-reduced space. For an alternative treatment and a brief history of cotangent bundle reduction, see Perlmutter [24, Chapter 3].

In the sequel G denotes a connected Lie group acting freely and properly on a connected manifold Q, and hence on T^*Q; $\mathbf{J} : \mathrm{T}^*Q \to \mathfrak{g}^*$ denotes the momentum map defined in 2.4(1); $\mathbf{A}$ denotes an arbitrary connection one-form on the principal bundle $\rho : Q \to Q/G$.

4.1 The zero momentum symplectic leaf

The form of an arbitrary symplectic leaf P_μ of $(\mathrm{T}^*Q)/G$ will be described in Section 5.1 using a concrete model for the abstract quotient $(\mathrm{T}^*Q)/G$ described in 4.2 below. However, the structure of the particular leaf $P_0 = \mathbf{J}^{-1}(0)/G$ can be described directly. Moreover, we shall need this description to relate symplectic structures on T^*Q and $\mathrm{T}^*(Q/G)$ (Corollary 4.3).

Since $\rho : Q \to Q/G$ is a submersion, it determines a natural vector bundle morphism $\rho^\circ : (\ker \mathrm{T}\rho)^\circ \to \mathrm{T}^*(Q/G)$ sending $d_q(f \circ \rho)$ to $d_{\rho(q)} f$, for each locally defined function f on Q/G. Here $(\ker \mathrm{T}\rho)^\circ$ denotes the annihilator of $\ker \mathrm{T}\rho$. In fact, 2.4(1) implies that $(\ker \mathrm{T}\rho)^\circ = \mathbf{J}^{-1}(0)$, so that $\mathbf{J}^{-1}(0)$ is a vector bundle over Q, and we have the commutative diagram

$$\begin{tikzcd} \mathbf{J}^{-1}(0) \arrow[r, "\rho^\circ"] \arrow[d] & \mathrm{T}^*(Q/G) \arrow[d] \\ Q \arrow[r, "\rho"'] & Q/G \end{tikzcd} .$$

Notation. We will write $\mathbf{J}^{-1}(0)_q \equiv \mathbf{J}^{-1}(0) \cap \mathrm{T}^*_q Q = (\ker \mathrm{T}_q\rho)^\circ$ for the fiber of $\mathbf{J}^{-1}(0)$ over $q \in Q$.

From the definition of ρ°, it follows that ρ° maps $\mathbf{J}^{-1}(0)_q$ isomorphically onto $\mathrm{T}^*_{\rho(q)}(Q/G)$. In particular, ρ° is surjective.
It is readily demonstrated that the fibers of ρ° are G-orbits so that ρ° determines a diffeomorphism between $\mathrm{T}^*(Q/G)$ and $P_0 = \mathbf{J}^{-1}(0)/G$. Moreover, if $\omega_{Q/G}$ denotes the canonical symplectic structure on $\mathrm{T}^*(Q/G)$ and $i_0 : \mathbf{J}^{-1}(0) \hookrightarrow \mathrm{T}^*Q$ the inclusion, then we have

$$(\rho^\circ)^*\omega_{Q/G} = i_0^*\omega \quad . \tag{1}$$

This formula is verified by first checking the analogous statement for the canonical one-forms on T^*Q and $\mathrm{T}^*(Q/G)$.

4.2 A model for the Poisson-reduced space $(\mathrm{T}^*Q)/G$

Let hor $= \ker \mathbf{A}$ denote the distribution of horizontal spaces on Q determined by $\mathbf{A} \in \Omega^1(Q, \mathfrak{g})$. Then have the decomposition of vector bundles over Q

$$\mathrm{T}Q = \mathrm{hor} \oplus \ker \mathrm{T}\rho \quad , \tag{1}$$

and the corresponding dual decomposition

$$\mathrm{T}^*Q = \mathbf{J}^{-1}(0) \oplus \mathrm{hor}^\circ \quad . \tag{2}$$

If $\mathbf{A}' : \mathrm{T}^*Q \to \mathbf{J}^{-1}(0)$ denotes the projection along hor°, then the composite

$$\mathrm{T}^*_{\mathbf{A}}\rho \equiv \rho^\circ \circ \mathbf{A}' : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G) \tag{3}$$

is a vector bundle morphism covering $\rho : Q \to Q/G$. It the Hamiltonian analogue of the tangent map $\mathrm{T}\rho : \mathrm{T}Q \to \mathrm{T}(Q/G)$.
The momentum map $\mathbf{J} : \mathrm{T}^*Q \to \mathfrak{g}^*$ determines a map $\mathbf{J}' : \mathrm{T}^*Q \to \mathfrak{g}^*_Q$ through

$$\mathbf{J}'(x) \equiv [q, \mathbf{J}(x)]_G \quad \text{for } x \in \mathrm{T}^*_q Q \text{ and } q \in Q \quad .$$

Note that while $\mathbf{J}$ is equivariant, the map $\mathbf{J}'$ is G-*in*variant.

Notation. If M_1, M_2 and B are smooth manifolds and there are maps $f_1 : M_1 \to B$ and $f_2 : M_2 \to B$, then one has the pullback manifold

$$\{(m_1, m_2) \in M_1 \times M_2 \mid f_1(m_1) = f_2(m_2)\} \ ,$$

which we will denote by $M_1 \oplus_B M_2$, or simply $M_1 \oplus M_2$. If f_1 and f_2 are fiber bundle projections then $M_1 \oplus M_2$ is a product in the category of fiber bundles over B. In particular, in the case of *vector* bundles, $M_1 \oplus M_2$ is the Whitney sum of M_1 and M_2. In any case, we write an element of $M_1 \oplus M_2$ as $m_1 \oplus m_2$ (rather than (m_1, m_2)).

Noting that $\mathrm{T}^*(Q/G)$ and $\mathfrak{g}^*_Q$ are both vector bundles over Q/G, we have the following result following from an unravelling of definitions:

Theorem. *The map $\pi : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q$ defined by $\pi(x) \equiv \mathrm{T}^*_{\mathbf{A}}\rho \cdot x \oplus \mathbf{J}'(x)$ is a surjective submersion whose fibers are the G-orbits in T^*Q. In other words, $\mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q$ is a realization of the abstract quotient $(\mathrm{T}^*Q)/G$, the map $\pi : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q$ being a realization of the natural projection $\mathrm{T}^*Q \to (\mathrm{T}^*Q)/G$.*

The above model of $(\mathrm{T}^*Q)/G$ is simply the dual of Cendra, Holm, Marsden and Ratiu's model of $(\mathrm{T}Q)/G$ [9].

4.3 Momentum shifting

Before attempting to describe the symplectic leaves of the Poisson-reduced space $(\mathrm{T}^*Q)/G \cong \mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q$, we should understand the projection $\pi : \mathrm{T}^*Q \xrightarrow{/G} \mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q$ better. In particular, we should understand the map $\mathrm{T}^*_{\mathbf{A}}\rho : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G)$, which means first understanding the projection $\mathbf{A}' : \mathrm{T}^*Q \to \mathbf{J}^{-1}(0)$ along hor°.

Let $x \in \mathrm{T}^*_qQ$ be given and define $\mu \equiv \mathbf{J}(x)$. The restriction of $\mathbf{J}$ to T^*_qQ is a linear map onto $\mathfrak{g}^*$ (by 2.4(1)). The kernel of this restriction is $\mathbf{J}^{-1}(0)_q$ and $\mathbf{J}^{-1}(\mu)_q \equiv \mathbf{J}^{-1}(\mu) \cap \mathrm{T}^*_qQ$ is an affine subspace of T^*_qQ parallel to $\mathbf{J}^{-1}(0)_q$; see Fig. 4.

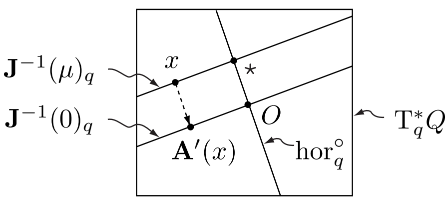


Figure 4: Describing the projection $x \mapsto \mathbf{A}'(x) : \mathrm{T}^*_qQ \to \mathbf{J}^{-1}(0)_q$ along hor°_q.

Since $\mathbf{J}^{-1}(0)_q$ and $\mathbf{J}^{-1}(\mu)_q$ are parallel, it follows from the decomposition 4.2(2) that $\mathbf{J}^{-1}(\mu)_q$ and hor°_q intersect in a single point $*$, as indicated in the figure.

We then have $\mathbf{A}'(x) = x - *$. Indeed, viewing the $\mathbb{R}$-valued one-form $\langle \mu, \mathbf{A} \rangle$ as a section of the cotangent bundle $\mathrm{T}^*Q \to Q$, one checks that the covector $\langle \mu, \mathbf{A} \rangle(q) \in \mathrm{T}^*_q Q$ belongs simultaneously to $\mathbf{J}^{-1}(\mu)$ and hor$^\circ$, so that $* = \langle \mu, \mathbf{A} \rangle(q)$. We have therefore proven the following:

Lemma. *Define the* momentum shift $M_\mu : \mathrm{T}^*Q \to \mathrm{T}^*Q$, *which maps* $\mathbf{J}^{-1}(0)_q$ *onto to* $\mathbf{J}^{-1}(\mu)_q$, *by* $M_\mu(x) \equiv x + \langle \mu, \mathbf{A} \rangle(\tau^*_Q(x))$, *where* $\tau^*_Q : \mathrm{T}^*Q \to Q$ *denotes the cotangent bundle projection. Then*

$$\mathbf{A}'(x) = M^{-1}_{\mathbf{J}(x)}(x) \ .$$

If θ denotes the canonical one-form on T^*Q, then one readily computes $M^*_\mu \theta = \theta + \langle \mu, (\tau^*_Q)^* \mathbf{A} \rangle$. In particular, as $\omega = -d\theta$,

$$M^*_\mu \omega = \omega - \langle \mu, (\tau^*_Q)^* d\mathbf{A} \rangle \ .$$

This identity, Equation 4.1(1), and the above Lemma have the following important corollary, which relates the symplectic structures on the domain and range of the map $\mathrm{T}^*_{\mathbf{A}}\rho : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G)$:

Corollary. *The two-forms* $(\mathrm{T}^*_{\mathbf{A}}\rho)^* \omega_{Q/G}$ *and* $\omega + \langle \mu, (\tau^*_Q)^* d\mathbf{A} \rangle$ *agree when restricted to* $\mathbf{J}^{-1}(\mu)$.

5 Symplectic leaves in Poisson reduced cotangent bundles

In this section we describe the symplectic leaves $P_\mu \subset (\mathrm{T}^*Q)/G$ as subsets of the model described in 4.2. We then describe explicitly their symplectic structures.

5.1 Reduced spaces as symplectic leaves

The following is a specialized version of the symplectic reduction theorem of Marsden, Weinstein and Meyer [20, 21], formulated such that the reduced spaces are realized as symplectic leaves (see, e.g., [7, Appendix E]).

Theorem. *Consider* P, ω, G, $\mathbf{J}$ *and* P_μ, *as defined in 2.1, where* $\mu \in \mathbf{J}(P)$ *is arbitrary. Then:*

(1) P_μ *is a symplectic leaf of* P/G *(which is a smooth Poisson manifold).*

(2) *The restriction* $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ *of* $\pi : P \to P/G$ *is a surjective submersion whose fibers are* G_μ*-orbits in* P, *i.e.,* P_μ *is a realization of the abstract quotient* $\mathbf{J}^{-1}(\mu)/G_\mu$.

(3) *If* ω_μ *is the leaf symplectic structure of* P_μ, *and* $i_\mu : \mathbf{J}^{-1}(\mu) \hookrightarrow P$ *the inclusion, then* $i^*_\mu \omega = \pi^*_\mu \omega_\mu$.

(4) $P_\mu \cap P_{\mu'} \neq \emptyset$ *if and only if* $P_\mu = P_{\mu'}$, *which is true if and only if* μ *and* μ' *lie on the same co-adjoint orbit. Also,* $P/G = \cup_{\mu \in \mathbf{J}(P)} P_\mu$.

(5) $\operatorname{codim} P_\mu = \operatorname{codim} G \cdot \mu$.

Proposition. *Fix $\mu \in \mathfrak{g}^*$. Then, taking $P \equiv \mathrm{T}^*Q$ and identifying P/G with $\mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q$ (Theorem 4.2), one obtains*

$$P_\mu = \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q \ , \quad \textit{where} \quad \mathcal{O} \equiv G \cdot \mu \ .$$

Here $G \cdot \mu$ denotes the co-adjoint orbit through μ and the associated bundle $\mathcal{O}_Q$ is to be viewed as a fiber subbundle of $\mathfrak{g}^*_Q$ in the obvious way.

Proof. Under the given identification, the projection $P \to P/G$ is represented by the map $\pi : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q$ defined in Theorem 4.2. From this definition it easily follows that $P_\mu \equiv \pi(\mathbf{J}^{-1}(\mu))$ is contained in $\mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$. We now prove the reverse inclusion $\mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q \subset \pi(\mathbf{J}^{-1}(\mu))$.
Let $z \oplus [q', \mu']_G$ be an arbitrary point in $\mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$. Then $\mu' \in \mathcal{O}$, so that $\mu' = g \cdot \mu$ for some $g \in G$, giving us $z \oplus [q', \mu']_G = z \oplus [q, \mu]_G$, where $q \equiv g^{-1} \cdot q'$. Now z and $[q, \mu]_G$ necessarily have a common base point in Q/G, which means that $z \in \mathrm{T}^*_{\rho(q)}(Q/G)$. The map $\rho^\circ : \mathbf{J}^{-1}(0) \to \mathrm{T}^*(Q/G)$ of 4.1 maps $\mathbf{J}^{-1}(0)_q \equiv \mathbf{J}^{-1}(0) \cap \mathrm{T}^*_q Q$ isomorphically onto $\mathrm{T}^*_{\rho(q)}(Q/G)$. Therefore there exists $x_0 \in \mathbf{J}^{-1}(0)_q$ such that $\rho^\circ(x_0) = z$. Define $x \equiv M_\mu(x_0) \in \mathbf{J}^{-1}(\mu)$, where M_μ is the momentum shift of Lemma 4.3. Then $\mathrm{T}^*_{\mathbf{A}}\rho \cdot x = z$. We now compute

$$\pi(x) = \mathrm{T}^*_{\mathbf{A}}\rho \cdot x \oplus \mathbf{J}'(x) = z \oplus [\tau^*_Q(x), \mathbf{J}(x)]_G = z \oplus [q, \mu]_G = z \oplus [q', \mu']_G \ .$$

Since x lies in $\mathbf{J}^{-1}(\mu)$ and $z \oplus [q', \mu']_G$ was an arbitrary point of $\mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$, this proves $\mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q \subset \pi(\mathbf{J}^{-1}(\mu))$. □

5.2 The leaf symplectic structures

The remainder of the section is devoted to the proof of the following key result, which is due (in a different form) to Perlmutter [24, Chapter 3]:

Theorem. *Let $\mathcal{O}$ denote the co-adjoint orbit through a point μ in the image of $\mathbf{J}$, let $\omega^-_{\mathcal{O}}$ denotes the 'minus' co-adjoint orbit symplectic structure on $\mathcal{O}$ (see 2.8), and let $i_{\mathcal{O}} \in \Omega^0(\mathcal{O}, \mathfrak{g}^*)$ denote the inclusion $\mathcal{O} \hookrightarrow \mathfrak{g}^*$. Let $(\omega^-_{\mathcal{O}})_Q \in \Omega^2(\mathcal{O}_Q)$ and $(i_{\mathcal{O}})_Q \in \Omega^0(\mathcal{O}_Q, \rho^*_{\mathcal{O}}\mathfrak{g}^*_Q)$ denote the corresponding associated forms; see 3.2. (Under the canonical identification $\rho^*_{\mathcal{O}}\mathfrak{g}^*_Q \cong \mathcal{O}_Q \oplus \mathfrak{g}^*_Q$, one has $(i_{\mathcal{O}})_Q(\eta) = \eta \oplus \eta$.) Then the symplectic structure of the leaf $P_\mu = \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$ is given by*

$$\omega_\mu = \mathrm{pr}_1^* \, \omega_{Q/G} + \mathrm{pr}_2^* \Big((\omega^-_{\mathcal{O}})_Q - (i_{\mathcal{O}})_Q \wedge \rho^*_{\mathcal{O}} \operatorname{curv} \mathbf{A} \Big) \ ,$$

where $\mathrm{pr}_1 : \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q \to \mathrm{T}^*(Q/G)$ *and* $\mathrm{pr}_2 : \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q \to \mathcal{O}_Q$ *denote the projections onto the first and second summands, and* $\operatorname{curv} \mathbf{A} \in \Omega^2(Q/G, \mathfrak{g}_Q)$ *denotes the curvature of* $\mathbf{A}$.

Because the restriction $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ of $\pi : \mathrm{T}^*Q \to \mathrm{T}^*(Q/G) \oplus \mathfrak{g}_Q^*$ is a surjective submersion, by Theorem 5.1(2), to prove the above Theorem it suffices to verify the formula in 5.1(3). Appealing to the definition of π (Theorem 4.2) and Corollary 4.3, we compute

$$(1) \qquad \pi_\mu^* \operatorname{pr}_1^* \omega_{Q/G} = i_\mu^* (\mathrm{T}_{\mathbf{A}}^* \rho)^* \omega_{Q/G} = i_\mu^* \omega + \langle \mu, i_\mu^* (\tau_Q^*)^* d\mathbf{A} \rangle \ .$$

For the next part of the proof we need the following technical result proven at the end:

Lemma. *If $u \in \mathrm{T}_x(\mathbf{J}^{-1}(\mu))$ is arbitrary, then*

$$\mathrm{T}\mathbf{J}' \cdot u = \frac{d}{dt} [q^{\mathrm{hor}}(t), \exp(t\xi) \cdot \mu]_G \Big|_{t=0} \ ,$$

for some $\mathbf{A}$-horizontal curve $t \mapsto q^{\mathrm{hor}}(t) \in Q$, where $\xi \equiv -\mathbf{A}(\mathrm{T}\tau_Q^* \cdot u)$.

Now $\pi_\mu^* \operatorname{pr}_2^* (\omega_{\mathcal{O}}^-)_Q = i_\mu^* (\mathbf{J}')^* (\omega_{\mathcal{O}}^-)_Q$ and, by definition,

$$\omega_{\mathcal{O}}^- \left(\frac{d}{dt} \exp(t\xi) \cdot \mu \Big|_{t=0}, \frac{d}{dt} \exp(t\eta) \cdot \mu \Big|_{t=0} \right) = -\langle \mu, [\xi, \eta] \rangle \qquad (\xi, \eta \in \mathfrak{g}) \ .$$

So it readily follows from the lemma that

$$(2) \qquad \pi_\mu^* \operatorname{pr}_2^* (\omega_{\mathcal{O}}^-)_Q = -\frac{1}{2} \langle \mu, i_\mu^* (\tau_Q^*)^* (\mathbf{A} \wedge \mathbf{A}) \rangle \ .$$

A routine calculation of pullbacks shows that

$$(3) \quad (\pi_\mu^* \operatorname{pr}_2^* (i_{\mathcal{O}})_Q)(x) = [x \oplus \tau_Q^*(x), \mu]_G \in (\rho_{\mathcal{O}} \circ \operatorname{pr}_2 \circ \pi_\mu)^* \mathfrak{g}_Q^* \qquad (x \in \mathbf{J}^{-1}(\mu))$$

and

$$(4) \quad (\pi_\mu^* \operatorname{pr}_2^* \rho_{\mathcal{O}}^* \operatorname{curv} \mathbf{A})(u_1, u_2) = \\ [x \oplus \tau_Q^*(x), i_\mu^* (\tau_Q^*)^* \mathbf{D}\mathbf{A}(u_1, u_2)]_G \in (\rho_{\mathcal{O}} \circ \operatorname{pr}_2 \circ \pi_\mu)^* \mathfrak{g}_Q \ ,$$

for $u_1, u_2 \in \mathrm{T}_x(\mathbf{J}^{-1}(\mu))$, where $\mathbf{D}\mathbf{A} \in \Omega^2(Q, \mathfrak{g})$ denotes the exterior covariant derivative of $\mathbf{A}$. In deriving (4) we have used the fact that $\rho_{\mathcal{O}} \circ \operatorname{pr}_2 \circ \pi_\mu = \rho \circ \tau_Q^* \circ i_\mu$ and that $\rho^* \operatorname{curv} \mathbf{A} \in \Omega^2(Q, \rho^* \mathfrak{g}_Q)$ satisfies the identity

$$(\rho^* \operatorname{curv} \mathbf{A})(v_1, v_2) = [q \oplus q, \mathbf{D}\mathbf{A}(v_1, v_2)]_G \in \rho^* \mathfrak{g}_Q \qquad (v_1, v_2 \in \mathrm{T}_q Q) \ .$$

This identity simply states, in pullback jargon, that $\operatorname{curv} \mathbf{A}$ is the two-form $\mathbf{D}\mathbf{A}$ on Q, viewed as a $\mathfrak{g}_Q$-valued form on the base Q/G.
Carrying out an implied contraction, Equations (3) and (4) deliver

$$(5) \qquad \pi_\mu^* (\operatorname{pr}_2^* ((i_{\mathcal{O}})_Q \wedge \rho_{\mathcal{O}}^* \operatorname{curv} \mathbf{A})) = \langle \mu, i_\mu^* (\tau_Q^*)^* \mathbf{D}\mathbf{A} \rangle \in \Omega^2(\mathbf{J}^{-1}(\mu)) \ .$$

From Equations (2), (5) and the Mauer-Cartan equation $d\mathbf{A} = \mathbf{D}\mathbf{A} + \frac{1}{2} \mathbf{A} \wedge \mathbf{A}$, follows the formula

$$(6) \qquad \pi_\mu^* \operatorname{pr}_2^* ((\omega_{\mathcal{O}}^-)_Q - (i_{\mathcal{O}})_Q \wedge \rho_{\mathcal{O}}^* \operatorname{curv} \mathbf{A}) = -\langle \mu, i_\mu^* (\tau_Q^*)^* d\mathbf{A} \rangle \ .$$

The formula in 5.1(3) follows from (6) and (1), which completes the proof of the theorem.

Proof of the Lemma. We have $u = d/dt\, x(t)\,|_{t=0}$ for some curve $t \mapsto x(t) \in \mathbf{J}^{-1}(\mu)$, in which case

$$\mathrm{T}\mathbf{J}' \cdot u = \frac{d}{dt}\, [q(t), \mu]_G \Big|_{t=0} \ ,$$

where $q(t) \equiv \tau_Q^*(x(t))$. We can write $q(t) = g(t) \cdot q^{\mathrm{hor}}(t)$ for some $\mathbf{A}$-horizontal curve $t \mapsto q^{\mathrm{hor}}(t) \in Q$ and some curve $t \mapsto g(t) \in G$ with $g(0) = \mathrm{id}$ and with

$$\frac{d}{dt}\, g(t) \Big|_{t=0} = \mathbf{A} \left(\frac{d}{dt}\, q(t) \Big|_{t=0} \right) = \mathbf{A}(\mathrm{T}\tau_Q^* \cdot u) = -\xi \ .$$

Then

$$\begin{aligned} \mathrm{T}\mathbf{J}' \cdot u &= \frac{d}{dt}\, [q^{\mathrm{hor}}(t), g(t)^{-1} \cdot \mu]_G \Big|_{t=0} \\ &= \frac{d}{dt}\, [q^{\mathrm{hor}}(t), \exp(t\xi) \cdot \mu]_G \Big|_{t=0} \ , \end{aligned}$$

as required. $\square$

6 A connection on the Poisson-reduced phase space

To apply Theorem 2.3 to the case $P = \mathrm{T}^*Q$ we need to choose a connection D on the symplectic stratification of $P/G \cong \mathrm{T}^*(Q/G) \oplus \mathfrak{g}_Q^*$. Such connections were defined in 2.2. As we shall see, this more-or-less amounts to choosing an inner product on $\mathfrak{g}^*$ (or $\mathfrak{g}$). Life is made considerably easier if this choice is Ad-invariant. (For example, in the case $Q = G$, which we discuss first, one might be tempted to use the inertia tensor $\mathbb{I} \in \mathfrak{g}^* \otimes \mathfrak{g}^*$ to form an inner product. However, this seems to lead to intractable calculations of the phase. It also makes the geometric phase g_{geom} more 'dynamic' and less 'geometric.') Fortunately, we will see that the particular choice of invariant inner product is immaterial.

In 6.3 and 6.4 we discuss details needed to describe explicitly the transverse derivative operator D_μ, and we also compute the canonical two-form ω_D (both these depend on the choice of D). Recall that these will be needed to apply Theorem 2.3.

6.1 The limiting case $Q = G$

When $Q = G$, we have $P/G \cong \mathfrak{g}^*$ and the symplectic leaves are the co-adjoint orbits. A connection on the symplectic stratification of P/G is then distribution on $\mathfrak{g}^*$ furnishing a complement, at each point $\mu \in \mathfrak{g}^*$, for the space $\mathrm{T}_\mu(G \cdot \mu)$ tangent to the co-adjoint orbit $G \cdot \mu$ through μ. As a subspace of $\mathfrak{g}^*$ this tangent space is the annihilator $\mathfrak{g}_\mu^\circ$ of $\mathfrak{g}_\mu$.

Lemma. *Let G be a connected Lie group whose Lie algebra $\mathfrak{g}$ admits an* Ad-*invariant inner product. Then for all $\mu \in \mathfrak{g}^*_{\mathrm{reg}}$ one has*

$$\mathfrak{g}_\mu^\perp = [\mathfrak{g}, \mathfrak{g}_\mu] \ .$$

Here $\mathfrak{g}^*_{\mathrm{reg}}$ denotes the set of regular points of the co-adjoint action

Proof. See Appendix B. □

The following proposition constructs a connection E on the symplectic stratification of $\mathfrak{g}^*$.

Proposition. *Let G be a connected Lie group whose Lie algebra $\mathfrak{g}$ admits an* Ad-*invariant inner product and equip $\mathfrak{g}^*$ with the corresponding* Ad^*-*invariant inner product. Let E denote the connection on the symplectic stratification of $\mathfrak{g}^*$ obtained by orthogonalizing the distribution tangent to the co-adjoint orbits:*

$$E(\mu) \equiv \Big(\mathrm{T}_\mu (G \cdot \mu) \Big)^\perp \ .$$

Let $\operatorname{forg} E(\mu)$ denotes the image of $E(\mu)$ under the canonical identification $\mathrm{T}_\mu \mathfrak{g}^* \cong \mathfrak{g}^*$, i.e., $\operatorname{forg} E(\mu) \subset \mathfrak{g}^*$ is $E(\mu) \subset \mathrm{T}_\mu \mathfrak{g}^*$ with base point 'forgotten.' Then for all $\mu \in \mathfrak{g}^*_{\mathrm{reg}}$:

(1) $\operatorname{forg} E(\mu) = [\mathfrak{g}, \mathfrak{g}_\mu]^\circ$.

(2) *$E(\mu)$ is independent of the particular choice of inner product.*

(3) *The restriction $\iota_\mu : \operatorname{forg} E(\mu) \to \mathfrak{g}^*_\mu$ of the natural projection $p_\mu : \mathfrak{g}^* \to \mathfrak{g}^*_\mu$ is an isomorphism.*

(4) *The orthogonal projection $\mathrm{pr}_\mu : \mathfrak{g} \to \mathfrak{g}_\mu$ is independent of the choice of inner product and satisfies the identity*

$$\langle \iota_\mu^{-1}(\nu), \xi \rangle = \langle \nu, \mathrm{pr}_\mu \xi \rangle \qquad (\nu \in \mathfrak{g}^*_\mu,\ \xi \in \mathfrak{g}) \ .$$

(5) *The complementary projection $\mathrm{pr}_\mu^\perp \equiv \mathrm{id} - \mathrm{pr}_\mu$ satisfies the identity*

$$\mathrm{pr}_\mu [\mathrm{pr}_\mu^\perp \xi, \mathrm{pr}_\mu^\perp \eta] = \mathrm{pr}_\mu [\xi, \eta] \qquad (\xi, \eta \in \mathfrak{g}) \ .$$

(6) *There exists a subspace $V \subset \mathfrak{g}^*$ containing μ and an open neighborhood $S \subset V$ of μ such that $\mathrm{T}_s S = E(s)$ for all $s \in S$.*

Remark. One can choose the V in (6) to be G_μ-invariant (see the proof below), so that S (suitably shrunk) is a slice for the co-adjoint action. This is provided, of course, that G has *closed* co-adjoint orbits. Although we do *not* assume that these orbits are closed, the reader may nevertheless find it helpful to think of S as a slice. We do not use (6) until Section 8.

Proof. In fact (3) is true for *any* space $E(\mu)$ complementary to $\mathrm{T}_\mu(G \cdot \mu)$, for this means

$$\mathrm{T}_\mu \mathfrak{g}^* = E(\mu) \oplus \mathrm{T}_\mu(G \cdot \mu) \ , \tag{7}$$

which, on identifying the spaces with subspaces of $\mathfrak{g}^*$, delivers the decomposition

$$\mathfrak{g}^* = \operatorname{forg} E(\mu) \oplus \mathfrak{g}_\mu^\circ \ .$$

Since $\mathfrak{g}_\mu^\circ$ is the kernel of the linear surjection $p_\mu : \mathfrak{g}^* \to \mathfrak{g}_\mu^*$, (3) must be true. The identity in (4) is an immediate corollary.
Because taking annihilator and orthogonalizing are commutable operations, we deduce from the above Lemma the formula $(\mathfrak{g}_\mu^\circ)^\perp = [\mathfrak{g}, \mathfrak{g}_\mu]^\circ$. Since $\mathfrak{g}_\mu^\circ = \operatorname{forg} \mathrm{T}_\mu(G \cdot \mu)$, (1) holds. Claim (2) follows.
Regarding (5), we have

$$\begin{aligned} \mathrm{pr}_\mu[\mathrm{pr}_\mu^\perp \xi, \mathrm{pr}_\mu^\perp \eta] &= \mathrm{pr}_\mu[\xi - \mathrm{pr}_\mu \xi, \eta - \mathrm{pr}_\mu \eta] \\ &= \mathrm{pr}_\mu([\xi, \eta] + [\mathrm{pr}_\mu \xi, \mathrm{pr}_\mu \eta] - [\xi, \mathrm{pr}_\mu \eta] + [\eta, \mathrm{pr}_\mu \xi]) \ . \end{aligned}$$

The second term in parentheses vanishes because $\mathfrak{g}_\mu$ is Abelian (since $\mu \in \mathfrak{g}_{\mathrm{reg}}^*$). The third and fourth terms vanish because they lie in $[\mathfrak{g}, \mathfrak{g}_\mu]$, which is the kernel of pr_μ, on account of the Lemma. This kernel is evidently independent of the choice of inner product, which proves the first part of (4).
To prove (6), take

$$V \equiv [\mathfrak{g}, \mathfrak{g}_\mu]^\circ = \{\nu \in \mathfrak{g}^* \mid \mathfrak{g}_\nu \subset \mathfrak{g}_\mu\} \ ,$$

which clearly contains μ. Since $\dim \mathfrak{g}_\mu = \dim \mathfrak{g}_\nu$ if and only if $\nu \in \mathfrak{g}_{\mathrm{reg}}^*$, we conclude that

$$V \cap \mathfrak{g}_{\mathrm{reg}}^* = \{\nu \in \mathfrak{g}^* \mid \mathfrak{g}_\nu = \mathfrak{g}_\mu\} \ .$$

Since, $\mathfrak{g}_{\mathrm{reg}}^* \subset \mathfrak{g}^*$ is an open set (see Appendix B), it follows that μ has a neighborhood $S \subset V$ of μ such that $S \subset \mathfrak{g}_{\mathrm{reg}}^*$ and $\mathfrak{g}_s = \mathfrak{g}_\mu$ for all $s \in S$. For any $s \in S$ we then have

$$\operatorname{forg}(E(s)) = [\mathfrak{g}, \mathfrak{g}_s]^\circ = [\mathfrak{g}, \mathfrak{g}_\mu]^\circ = V = \operatorname{forg}(\mathrm{T}_s S) \ , \tag{8}$$

where the first equality follows from (1). Equation (8) implies that $E(s) = \mathrm{T}_s S$, as required. □

Henceforth E denotes the connection on the symplectic stratification of $\mathfrak{g}^*$ defined in the above Proposition.

6.2 The general case $Q \neq G$

In general, a connection D on the symplectic stratification of $(\mathrm{T}^*Q)/G \cong \mathrm{T}^*(Q/G) \oplus \mathfrak{g}_Q^*$ is given by

$$D(z \oplus [q, \mu]_G) \equiv \left\{ \frac{d}{dt} z \oplus [q, \mu + t\delta]_G \Big|_{t=0} \;\middle|\; \delta \in \operatorname{forg} E(\mu) \right\} \tag{1}$$
$$\left(z \in \mathrm{T}^*_{\rho(q)}(Q/G),\, q \in Q,\, \mu \in \mathfrak{g}^* \right) \ .$$

If $[q',\mu']_G = [q,\mu]_G$, then the right-hand side of (1) is unchanged by a substitution by primed quantities, because E is G-invariant. This shows that the distribution D is well defined. It is a connection on the symplectic stratification of $\mathrm{T}^*(Q/G)\oplus\mathfrak{g}^*_Q$ because E is a connection on the symplectic stratification of $\mathfrak{g}^*$, and because the symplectic leaf through a point $z\oplus[q,\mu]_G$ is $\mathrm{T}^*(Q/G)\oplus\mathcal{O}_Q$, where $\mathcal{O}\equiv G\cdot\mu$.

6.3 Transverse derivatives.

To determine the transverse derivative operator D_μ determined by D in the special case of cotangent bundles (needed to apply Theorem 2.3), we will need an explicit expression for the isomorphism $L(D,y,\mu):\mathfrak{g}^*_\mu\to D(y)$ defined in 2.2.

Lemma. *Fix $\mu\in\mathfrak{g}^*_{\mathrm{reg}}$. Then:*

(1) *Each $y\in P_\mu$ is of the form $y = z\oplus[q,\mu]_G$ for some $q\in Q$ and $z\in\mathrm{T}^*_{\rho(q)}(Q/G)$.*

(2) *For each such y one has*

$$L(D,\mu,y)(\nu) = \frac{d}{dt}\,z\oplus[q,\mu+t\iota_\mu^{-1}(\nu)]_G\Big|_{t=0}\ ,$$

where ι_μ is defined by 6.1(3)*.*

Proof. That each $y\in P_\mu$ is of the form given in (1) follows from an argument already given in the proof of Proposition 5.1. Moreover, that proof shows that there exists $x_0\in\mathbf{J}^{-1}(0)_q$ such that $\rho^\circ(x_0)=z$. We prove (2) by first computing the natural isomorphism $D(y)\xrightarrow{\sim}\mathfrak{g}^*_\mu$ in Lemma 2.2 whose inverse defines $L(D,\mu,y)$. Define $x\equiv M_\mu(x_0)$, where M_μ is the momentum shift defined in 4.3. Then $x\in\mathbf{J}^{-1}(\mu)$. According to (1), an arbitrary vector $v\in D(y)$ is of the form

$$v = \frac{d}{dt}\,z\oplus[q,\mu+t\delta]_G\Big|_{t=0}\ ,$$

for some $\delta\in\operatorname{forg}E(\mu)$. We claim that the vector

$$w\equiv\frac{d}{dt}\,M_{\mu+t\delta}(x_0)\Big|_{t=0}\in\mathrm{T}_xP\subset\mathrm{T}_{\mathbf{J}^{-1}(\mu)}P\qquad(P=\mathrm{T}^*Q)$$

is a valid choice for the corresponding vector w in Lemma 2.2. Indeed, one has

$$\begin{aligned}\mathrm{T}\pi\cdot w = \frac{d}{dt}\,\pi(M_{\mu+t\delta}(x_0))\Big|_{t=0} &= \frac{d}{dt}\,\mathrm{T}^*_{\mathbf{A}}\rho\cdot M_{\mu+t\delta}(x_0)\oplus\mathbf{J}'(M_{\mu+t\delta}(x_0))\Big|_{t=0}\\ &= \frac{d}{dt}\,\rho^\circ(x_0)\oplus[\tau^*_Q(M_{\mu+t\delta}(x_0)),\mu+t\delta]_G\Big|_{t=0}\\ &= \frac{d}{dt}\,z\oplus[q,\mu+t\delta]_G\Big|_{t=0} = v\ ,\end{aligned}$$

as required. We now compute

$$p_\mu \langle d\mathbf{J}, w\rangle = p_\mu \frac{d}{dt}\, \mathbf{J}(M_{\mu+t\delta}(x_0))\Big|_{t=0} = p_\mu \delta \ .$$

The natural isomorphism $D(y) \xrightarrow{\sim} \mathfrak{g}^*_\mu$ is therefore given by

$$\frac{d}{dt}\, z \oplus [q, \mu + t\delta]_G \Big|_{t=0} \mapsto p_\mu \delta \qquad (\delta \in \operatorname{forg} E(\mu)) \ .$$

Since $L(D, y, \mu)$ is the inverse of this map, this proves (2). □

6.4 The canonical two-form determined by D

We now determine the canonical two-form ω_D determined by D in the cotangent bundle case.
According to Theorem 5.2, the symplectic structure of the leaf $P_\mu = \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q$ ($\mathcal{O} \equiv G \cdot \mu$) is given by

$$\omega_\mu = \mathrm{pr}_1^*\, \omega_{Q/G} + \mathrm{pr}_2^* \Big((\omega^-_{\mathcal{O}})_Q - (i_{\mathcal{O}})_Q \wedge \rho^*_{\mathcal{O}} \operatorname{curv} \mathbf{A} \Big) \ , \tag{1}$$

where $\mathrm{pr}_1 : \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q \to \mathrm{T}^*(Q/G)$ and $\mathrm{pr}_2 : \mathrm{T}^*(Q/G) \oplus \mathcal{O}_Q \to \mathcal{O}_Q$ are the canonical projections. We claim that the canonical two-form $\omega_D \in \Omega^2(\mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q)$ determined by D (see 2.2) is given by

$$\omega_D = \mathrm{pr}_1^*\, \omega_{Q/G} + \mathrm{pr}_2^* \Big((\omega_E)_Q - (\mathrm{id}_{\mathfrak{g}^*})_Q \wedge \rho^*_{\mathfrak{g}^*} \operatorname{curv} \mathbf{A} \Big) \ . \tag{2}$$

Here pr_1 and pr_2 denote the canonical projections $\mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q \to \mathrm{T}^*(Q/G)$ and $\mathrm{T}^*(Q/G) \oplus \mathfrak{g}^*_Q \to \mathfrak{g}^*_Q$. The form ω_E denotes the canonical two-form on $\mathfrak{g}^*$ determined by E. The zero-form $(\mathrm{id}_{\mathfrak{g}^*})_Q \in \Omega^0(\mathfrak{g}^*_Q, \rho^*_{\mathfrak{g}^*}\mathfrak{g}^*_Q)$ denotes the form associated with the identity map $\mathrm{id}_{\mathfrak{g}^*} : \mathfrak{g}^* \to \mathfrak{g}^*$, viewed as an element of $\Omega^0(\mathfrak{g}^*, \mathfrak{g}^*)$. (If one makes the identification $\rho^*_{\mathfrak{g}^*}\mathfrak{g}^*_Q \cong \mathfrak{g}^*_Q \oplus \mathfrak{g}^*_Q$, then $(\mathrm{id}_{\mathfrak{g}^*})_Q(\eta) = \eta \oplus \eta$.) Recall that $\rho_{\mathfrak{g}^*} : \mathfrak{g}^*_Q \to Q/G$ denotes associated bundle projection.
The formula in (2) is easily verified by checking that $\omega_D(v, \cdot\,) = 0$ for $v \in D$, and by checking that the restriction of ω_D to a leaf P_μ coincides with the two-form on the right-hand side of (1).

7 The dynamic phase

For general G-invariant Hamiltonians $H : \mathrm{T}^*Q \to \mathbb{R}$ the formula for g_{dyn} in Theorem 3.5 follows from Theorem 2.3, Lemma 6.3, and the definition of D_{μ_0} given in 2.2. In this section we deduce the form taken by this phase in simple mechanical systems, as reported in Theorem 3.4.

7.1 The reduced Hamiltonian

The kinetic energy metric $\langle\!\langle \cdot\,,\,\cdot \rangle\!\rangle_Q$ induces an isomorphism $\mathrm{T}Q \xrightarrow{\sim} \mathrm{T}^*Q$ sending hor $\equiv \ker \mathbf{A}$ to $\mathbf{J}^{-1}(0)$ and $\ker \mathrm{T}\rho$ to hor° (see 4.2(1) and 4.2(2)). Since $\mathbf{J}^{-1}(0)_q = (\ker \mathrm{T}_q\rho)^\circ$, it is not too difficult to see that

$$(1) \qquad x \in \mathbf{J}^{-1}(0)_q \quad \Rightarrow \quad \langle\!\langle x, x \rangle\!\rangle_Q^* = \langle\!\langle \rho^\circ(x), \rho^\circ(x) \rangle\!\rangle_{Q/G}^* \ ,$$

where $\langle\!\langle \cdot\,,\,\cdot \rangle\!\rangle_{Q/G}^*$ is defined in 3.4 and ρ° is defined in 4.1.
If instead, $x \in \mathrm{hor}_q^\circ$, then x is the image under the isomorphism $\mathrm{T}Q \xrightarrow{\sim} \mathrm{T}^*Q$ of $\xi^Q(q)$, for some $\xi \in \mathfrak{g}$. For such ξ, and arbitrary $\eta \in \mathfrak{g}$, we compute

$$\langle \mathbf{J}(x), \eta \rangle = \langle x, \eta^Q(q) \rangle = \langle\!\langle \xi^Q(q), \eta^Q(q) \rangle\!\rangle_Q = \langle \hat{\mathbb{I}}(q)(\xi), \eta \rangle \ ,$$

where the first equality follows from 2.4(1). Since $\eta \in \mathfrak{g}$ is arbitrary, it follows that $\xi = \hat{\mathbb{I}}^{-1}(q)(\mathbf{J}(x))$. We now conclude that

$$(2) \quad x \in \mathrm{hor}_q^\circ \quad \Rightarrow \quad \langle\!\langle x, x \rangle\!\rangle_Q^* = \langle\!\langle \xi^Q(q), \xi^Q(q) \rangle\!\rangle_Q = \langle \mathbf{J}(x), \hat{\mathbb{I}}^{-1}(q)(\mathbf{J}(x)) \rangle \ .$$

An arbitrary element $x \in \mathrm{T}_q^*Q$ decomposes into unique parts along $\mathbf{J}^{-1}(0)_q$ and hor_q°, the first component being $\mathbf{A}'(x)$. From (1) and (2) one deduces

$$(3) \ \ \langle\!\langle x, x \rangle\!\rangle_Q^* = \langle\!\langle \mathrm{T}_{\mathbf{A}}^*\rho \cdot x, \mathrm{T}_{\mathbf{A}}^*\rho \cdot x \rangle\!\rangle_{Q/G}^* + \langle \mathbf{J}(x), \hat{\mathbb{I}}^{-1}(q)(\mathbf{J}(x)) \rangle \qquad (x \in \mathrm{T}_q^*Q) \ .$$

Define $h : \mathrm{T}^*(Q/G) \oplus \mathfrak{g}_Q^* \to \mathbb{R}$ by

$$(4) \qquad h(z \oplus [q, \mu]_G) = \frac{1}{2}\langle\!\langle z, z \rangle\!\rangle_{Q/G}^* + \frac{1}{2}\langle \mu, \hat{\mathbb{I}}^{-1}(q)\mu \rangle + V_{Q/G}(\rho(q)) \ ,$$

where $V_{Q/G}$ denotes the function on Q/G to which V drops on account of its G-invariance. With the help of (3), one checks that $H = h \circ \pi$, i.e., h is the Poisson-reduced Hamiltonian. Substituting (4) into 3.5(1) delivers the formula

$$(5) \qquad D_{\mu_0} h(z \oplus [q, \mu_0]_G) = \mathrm{pr}_{\mu_0} \hat{\mathbb{I}}^{-1}(q)\mu_0 \ ,$$

where $\mathrm{pr}_{\mu_0} : \mathfrak{g} \to \mathfrak{g}_{\mu_0}$ denotes the orthogonal projection.
To establish the formula for g_{dyn} in Theorem 3.4 it remains to show that

$$(6) \qquad \Big((\mathrm{Pr}_{\mu_0})_Q \wedge \xi_{\mathbb{I}} \Big)([q, \mu_0]_G) = \mathrm{pr}_{\mu_0} \hat{\mathbb{I}}^{-1}(q)\mu_0 \ ,$$

where $\xi_{\mathbb{I}} \equiv \rho_{\mathcal{O}}^* \mathbb{I}^{-1} \wedge (i_{\mathcal{O}})_Q$. We will be ready to do so after providing the general definition of associated forms alluded to in 3.2.

7.2 Associated forms (general case)

Let V be a real vector space on which G acts linearly and $\mathcal{O}$ an arbitrary manifold on which G acts smoothly. Let λ be a V-valued k-form on $\mathcal{O}$. For the sake of clarity, we will suppose $k = 1$; the extension to general k will be obvious.

Assuming that $\lambda \in \Omega^1(\mathcal{O}, V)$ is equivariant in the sense that

$$\lambda(g \cdot u) = g \cdot \lambda(u) \qquad (g \in G,\, u \in \mathrm{T}\mathcal{O}) \ ,$$

we will construct a *bundle*-valued differential form $\lambda_Q \in \Omega^1(\mathcal{O}_Q, \rho_{\mathcal{O}}^* V_Q)$ called the *associated form*. Recall that $\rho_{\mathcal{O}} : \mathcal{O}_Q \to Q/G$ denotes the projection of the associated bundle $\mathcal{O}_Q \equiv (Q \times \mathcal{O})/G$, and $\rho_{\mathcal{O}}^*$ denotes pullback. As always, we assume $\rho : Q \to Q/G$ is equipped with a connection one-form $\mathbf{A}$.

We begin by noting that an arbitrary vector tangent to $\rho_{\mathcal{O}}^* Q \equiv \mathcal{O}_Q \oplus_{Q/G} Q$ is of the form

$$\frac{d}{dt}\, [q^{\mathrm{hor}}(t), \nu(t)]_G \oplus \exp(t\xi) \cdot q^{\mathrm{hor}}(t) \Big|_{t=0} \ , \tag{1}$$

for some $\xi \in \mathfrak{g}$, some $\mathbf{A}$-horizontal curve $t \mapsto q^{\mathrm{hor}}(t) \in Q$, and some curve $t \mapsto \nu(t) \in \mathcal{O}$. Define $\Lambda \in \Omega^1(\rho_{\mathcal{O}}^* Q, V)$ by

$$\Lambda \left(\frac{d}{dt}\, [q^{\mathrm{hor}}(t), \nu(t)]_G \oplus \exp(t\xi) \cdot q^{\mathrm{hor}}(t) \Big|_{t=0} \right) \equiv \lambda \left(\frac{d}{dt}\, \nu(t) \Big|_{t=0} \right) \ .$$

As the reader is left to verify, the equivariance of λ ensures that Λ is well defined. Now $\rho_{\mathcal{O}}^* Q \equiv \mathcal{O}_Q \oplus_{Q/G} Q$ is a principal G-bundle (G acts according $g \cdot (\eta \oplus q) \equiv \eta \oplus (g \cdot q)$) and we claim that Λ is tensorial.

Proof that Λ is tensorial. The (tangent-lifted) action of G on $\mathrm{T}(\rho_{\mathcal{O}}^* Q)$ is given by

$$\begin{aligned} & g \cdot \frac{d}{dt}\, [q^{\mathrm{hor}}(t), \nu(t)]_G \oplus \exp(t\xi) \cdot q^{\mathrm{hor}}(t) \Big|_{t=0} \\ =& \frac{d}{dt}\, [q^{\mathrm{hor}}(t), \nu(t)]_G \oplus g \exp(t\xi) \cdot q^{\mathrm{hor}}(t) \Big|_{t=0} \\ =& \frac{d}{dt}\, [g \cdot q^{\mathrm{hor}}(t), g \cdot \nu(t)]_G \oplus \exp(tg \cdot \xi) \cdot (g \cdot q^{\mathrm{hor}}(t)) \Big|_{t=0} \ . \end{aligned}$$

Since $t \mapsto g \cdot q^{\mathrm{hor}}(t)$ is $\mathbf{A}$-horizontal, it follows that

$$\begin{aligned} & \Lambda \left(g \cdot \frac{d}{dt}\, [q^{\mathrm{hor}}(t), \nu(t)]_G \oplus \exp(t\xi) \cdot q^{\mathrm{hor}}(t) \Big|_{t=0} \right) \\ & \qquad = \lambda \left(\frac{d}{dt}\, g \cdot \nu(t) \Big|_{t=0} \right) = g \cdot \lambda \left(\frac{d}{dt}\, \nu(t) \Big|_{t=0} \right) \ , \end{aligned}$$

where the second quality follows from the equivariance of λ. What we have just shown is that

$$\Lambda(g \cdot u) = g \cdot \Lambda(u) \qquad (g \in G)$$

for arbitrary $u \in \mathrm{T}(\rho_{\mathcal{O}}^{*}Q)$, i.e., Λ is equivariant. Also, the generic tangent vector in (1) is vertical (in the principal bundle $\rho_{\mathcal{O}}^{*}Q \to \mathcal{O}_Q$) if and only if $d/dt\,[q^{\mathrm{hor}}(t), \nu(t)]_G\,|_{t=0} = 0$. This is true if and only if $d/dt\,\nu(t)\,|_{t=0} = 0$. It follows that Λ vanishes on vertical vectors. This fact and the forementioned equivariance establishes that Λ is tensorial. $\square$

Because $\Lambda \in \Omega^1(\rho_{\mathcal{O}}^{*}Q, V)$ is tensorial, it drops to an element of $\Omega^1(\mathcal{O}_Q, \rho_{\mathcal{O}}^{*}V_Q)$, which is the sought after associated form λ_Q. By construction one has the implicit formula

$$(2) \qquad \lambda_Q\left(\frac{d}{dt}\,[q^{\mathrm{hor}}(t), \nu(t)]_G\,\Big|_{t=0}\right) = \left[[q,\nu]_G \oplus q, \lambda\left(\frac{d}{dt}\,\nu(t)\,\Big|_{t=0}\right)\right]_G \quad ,$$

where $q \equiv q^{\mathrm{hor}}(0)$ and $\nu \equiv \nu(0)$.

Formula (2) is for a one-form λ. From the zero-form analogue of (2), one deduces

$$(3) \qquad (\mathrm{Pr}_{\mu_0})_Q([q,\mu_0]_G) = [[q,\mu_0]_G \oplus q, \mathrm{pr}_{\mu_0}]_G$$

$$(4) \qquad (i_{\mathcal{O}})_Q([q,\mu_0]_G) = [[q,\mu_0]_G \oplus q, \mu_0]_G \quad .$$

Since

$$(\rho_{\mathcal{O}}^{*}\mathbb{I}^{-1})([q,\mu_0]_G) = [[q,\mu_0]_G \oplus q, \hat{\mathbb{I}}(q)]_G \quad ,$$

we deduce

$$(\mathrm{Pr}_{\mu_0})_Q(\rho_{\mathcal{O}}^{*}\mathbb{I}^{-1} \wedge (i_{\mathcal{O}})_Q)([q,\mu_0]_G) = \mathrm{pr}_{\mu_0}\,\hat{\mathbb{I}}^{-1}(q)\mu_0 \quad ,$$

which proves 7.1(6).

8 The geometric phase

This section derives the formula for g_{geom} reported in Theorem 3.4. We will carry out several computations, some of them somewhat involved. However, our objective throughout is clear: To apply the formula for g_{geom} in 2.3 we must calculate the transverse derivative $D_{\mu_0}\omega_D$ of the leaf symplectic structures $\omega_\mu = \omega_D|P_\mu$. To do so we must first compute $d\omega_D$. Our preference for a coordinate free proof leads us to lift the computation to a bigger space, which we do with the help of the 'slice' S for the co-adjoint action delivered by 6.1(6).

Using the fact that d is an antiderivation, that d commutes with pullbacks, and that $d\omega_{Q/G} = 0$, we obtain from 6.4(2)

$$(1) \quad d\omega_D = \mathrm{pr}_2^{*}(\,d(\omega_E)_Q - d(\mathrm{id}_{\mathfrak{g}^*})_Q \wedge \rho_{\mathfrak{g}^*}^{*}\,\mathrm{curv}\,\mathbf{A} - (\mathrm{id}_{\mathfrak{g}^*})_Q \wedge \rho_{\mathfrak{g}^*}^{*}\,d\,\mathrm{curv}\,\mathbf{A}\,) \quad .$$

Note here that we are using the exterior derivative in the generalized sense of *bundle*-valued forms, as defined with respect to the connection $\mathbf{A}$; see A.5, Appendix A. The last term in parentheses is immediately dispensed with, for one has Bianchi's identity[2]

$$d\,\mathrm{curv}\,\mathbf{A} = 0 \ . \tag{2}$$

To write down formulas for other terms in (1), it will be convenient to have an appropriate representation for vectors tangent to $\mathfrak{g}^*_Q$. Indeed, as the reader will readily verify, each such vector is of the form

$$\frac{d}{dt}\,[q^{\mathrm{hor}}(t),\mu(t)]_G\,\Big|_{t=0} \ ,$$

for some $\mathbf{A}$-horizontal curve $t \mapsto q^{\mathrm{hor}}(t) \in Q$ and some curve $t \mapsto \mu(t) \in \mathfrak{g}^*$. On occasion, and without loss of generality, we will take $\mu(t)$ to be of the form

$$\mu(t) = \exp(t\xi)\cdot(\mu + tv) \ ,$$

for some $\xi \in \mathfrak{g}$, $\mu \in \mathfrak{g}^*$ and $v \in \mathrm{forg}\,E(\mu)$ (see Proposition 6.1).
A straightforward computation gives

$$d(\mathrm{id}_{\mathfrak{g}^*})_Q\left(\frac{d}{dt}\,[q^{\mathrm{hor}}(t),\mu(t)]_G\,\Big|_{t=0}\right) = [[q^{\mathrm{hor}}(0),\mu(0)]_G \oplus q^{\mathrm{hor}}(0),\dot\mu(0)]_G \in \rho^*_{\mathfrak{g}^*}\mathfrak{g}^*_Q,$$

where $\dot\mu(0) \equiv d/dt\,\mu(t)\,|_{t=0} \in \mathfrak{g}^*$. From this follows the formula

(3)

$$\begin{aligned}
(d(\mathrm{id}_{\mathfrak{g}^*})_Q \wedge \rho^*_{\mathfrak{g}^*}\,\mathrm{curv}\,\mathbf{A})&\left(\frac{d}{dt}\,[q_1^{\mathrm{hor}}(t),\mu_1(t)]_G\,\Big|_{t=0},\ldots,\frac{d}{dt}\,[q_3^{\mathrm{hor}}(t),\mu_3(t)]_G\,\Big|_{t=0}\right)\\
&= \Big\langle\, \dot\mu_1(0),\, \mathbf{DA}(\dot q_2^{\mathrm{hor}}(0),\dot q_3^{\mathrm{hor}}(0))\,\Big\rangle\\
&+ \Big\langle\, \dot\mu_2(0),\, \mathbf{DA}(\dot q_3^{\mathrm{hor}}(0),\dot q_1^{\mathrm{hor}}(0))\,\Big\rangle\\
&+ \Big\langle\, \dot\mu_3(0),\, \mathbf{DA}(\dot q_1^{\mathrm{hor}}(0),\dot q_2^{\mathrm{hor}}(0))\,\Big\rangle \ ,
\end{aligned}$$

where $\mathbf{D}$ denotes exterior covariant derivative and $\dot q_j^{\mathrm{hor}}(0) \equiv d/dt\,q_j^{\mathrm{hor}}(t)\,|_{t=0}$.
To compute $d(\omega_E)_Q$ is not so straightforward.[3] The difficulty lies partly in the fact that the co-adjoint orbit symplectic structures, which ω_E 'collects together,' are defined *implicitly* in terms of the infinitesimal generators of the co-adjoint action, and this action is generally not free. We overcome this by pulling $(\omega_E)_Q$ back to a 'bigger' space where we can be explicit. We compute

[2]Perhaps the better known form of this identity is $\mathbf{D}(\mathbf{DA}) = 0$, where $\mathbf{D}$ denotes exterior covariant derivative (see, e.g., [12, Theorem II.5.4]). Since, in the notation of Appendix A, $\mathbf{DA} = (\mathrm{curv}\,\mathbf{A})^\frown$, it follows that $(d\,\mathrm{curv}\,\mathbf{A})^\frown = 0$, which in turn implies (2).

[3]The exterior derivative d does *not* commute with the formation of associated forms!

the derivative in the bigger space and then drop to $\mathfrak{g}^*_Q$. Here is the formula we will derive:

$$
\begin{aligned}
(4)\qquad d(\omega_E)_Q \bigg(& \frac{d}{dt}\,[q_1^{\mathrm{hor}}(t), \exp(t\xi_1)\cdot(\mu+tv_1)]_G\Big|_{t=0}, \\
& \frac{d}{dt}\,[q_2^{\mathrm{hor}}(t), \exp(t\xi_2)\cdot(\mu+tv_2)]_G\Big|_{t=0}, \\
& \frac{d}{dt}\,[q_3^{\mathrm{hor}}(t), \exp(t\xi_3)\cdot(\mu+tv_3)]_G\Big|_{t=0}\bigg) \\
= & -\langle v_1,[\xi_2,\xi_3]\rangle - \langle v_2,[\xi_3,\xi_1]\rangle - \langle v_3,[\xi_1,\xi_2]\rangle \\
& - \langle \mu, [\xi_1, \mathbf{DA}(\dot q_2^{\mathrm{hor}}(0), \dot q_3^{\mathrm{hor}}(0))]_G\rangle \\
& - \langle \mu, [\xi_2, \mathbf{DA}(\dot q_3^{\mathrm{hor}}(0), \dot q_1^{\mathrm{hor}}(0))]_G\rangle \\
& - \langle \mu, [\xi_3, \mathbf{DA}(\dot q_1^{\mathrm{hor}}(0), \dot q_2^{\mathrm{hor}}(0))]_G\rangle \\
& \Big(\xi_j\in\mathfrak{g},\ \mu\in\mathfrak{g}^*_{\mathrm{reg}},\ v_j\in \operatorname{forg} E(\mu)\Big)\ ,
\end{aligned}
$$

where $q_1^{\mathrm{hor}}(0) = q_2^{\mathrm{hor}}(0) = q_3^{\mathrm{hor}}(0) \equiv q \in Q$. Note that we insist that μ lies in $\mathfrak{g}^*_{\mathrm{reg}}$. In other words, (4) is a formula for $(\omega_E)_Q$ on the open dense set $(\mathfrak{g}^*_{\mathrm{reg}})_Q \subset \mathfrak{g}^*_Q$.

Derivation of (4). With $\mu \in \mathfrak{g}^*_{\mathrm{reg}}$ fixed, let $S \subset V \subset \mathfrak{g}^*$ denote the corresponding 'slice' furnished by Proposition 6.1(6). Define the map

$$
\begin{aligned}
b : Q\times G\times S &\to \mathfrak{g}^*_Q \\
(q,g,s) &\mapsto [q, g\cdot s]_G\ .
\end{aligned}
$$

At each point $(q,g,s)\in Q\times G\times S$ we define, for each $(u,\eta,\xi,v)\in \mathrm{T}_{\rho(q)}(Q/G)\times \mathfrak{g}\times\mathfrak{g}\times V$, the tangent vector

$$
\begin{aligned}
&\langle u,\eta,\xi,v;\, q,g,s\rangle \equiv \\
&\quad \frac{d}{dt}\Big(\exp(t\eta)\cdot q^{\mathrm{hor}}(t),\ \exp(t\xi)\cdot g,\ s+tv\Big)\Big|_{t=0} \in \mathrm{T}_{(q,g,s)}(Q\times G\times S)\ ,
\end{aligned}
$$

where $t\mapsto q^{\mathrm{hor}}(t)\in Q$ is any $\mathbf{A}$-horizontal curve satisfying

$$
\begin{aligned}
& q^{\mathrm{hor}}(0) = q \\
\text{and}\qquad & \frac{d}{dt}\,\rho(q^{\mathrm{hor}}(t))\Big|_{t=0} = u\ .
\end{aligned}
$$

Note that every vector tangent to $Q\times G\times S$ is of the above form, and that

$$
\begin{aligned}
(5)\quad \mathrm{T}b\cdot\langle u,\eta,\xi,v;\, q,g,s\rangle &= \frac{d}{dt}\,[\exp(t\eta)\cdot q^{\mathrm{hor}}(t),\ \exp(t\xi)g\cdot(s+tv)]_G\Big|_{t=0} \\
(6)\qquad\qquad &= \frac{d}{dt}\,[q^{\mathrm{hor}}(t),\ \exp(-t\eta)\exp(t\xi)g\cdot(s+tv)]_G\Big|_{t=0}\ .
\end{aligned}
$$

From (6) and the definition of associated forms 3.2(1), we obtain

$$\begin{aligned}(7)\qquad & b^*(\omega_E)_Q\Big(\langle u_1,\eta_1,\xi_1,v_1;\,q,g,s\rangle,\,\langle u_2,\eta_2,\xi_2,v_2;\,q,g,s\rangle\Big)\\ &=\omega_E\left(\frac{d}{dt}\,\exp(-t\eta_1)\exp(t\xi_1)g\cdot(s+tv_1)\Big|_{t=0},\right.\\ &\qquad\left.\frac{d}{dt}\,\exp(-t\eta_2)\exp(t\xi_2)g\cdot(s+tv_2)\Big|_{t=0}\right)\ .\end{aligned}$$

Now ω_E is the canonical two-form on $\mathfrak{g}^*$ determined by E and according to 6.1(6), we have

$$\frac{d}{dt}\,s+tv_j\Big|_{t=0}\in \mathrm{T}_sS=E(s)\qquad (j=1,2)\ .$$

It follows from (7) that

$$\begin{aligned}(8)\quad & b^*(\omega_E)_Q\Big(\langle u_1,\eta_1,\xi_1,v_1;\,q,g,s\rangle,\,\langle u_2,\eta_2,\xi_2,v_2;\,q,g,s\rangle\Big)\\ &=-\Big\langle\, g\cdot s,\,[\xi_1-\eta_1,\xi_2-\eta_2]\,\Big\rangle\qquad (u_j\in \mathrm{T}_{\rho(q)}(Q/G);\,\eta_j,\xi_j\in\mathfrak{g};\,v_j\in V)\ .\end{aligned}$$

It is now that we see the reason for pulling $(\omega_E)_Q$ back to $Q\times G\times S$. For if we define natural projections

$$\begin{aligned}\pi_Q &: Q\times G\times S\to Q:(q,g,s)\mapsto q\\ \pi_G &: Q\times G\times S\to G:(q,g,s)\mapsto g\\ \pi_{\mathfrak{g}^*} &: Q\times G\times S\to \mathfrak{g}^*:(q,g,s)\mapsto g\cdot s\end{aligned}$$

and denote by $\theta_G\in\Omega^1(G,\mathfrak{g})$ the right-invariant Mauer-Cartan form on G, then (8) may be written intrinsically as

$$b^*(\omega_E)_Q=-\frac{1}{2}\pi_{\mathfrak{g}^*}\wedge\Big((\pi_G^*\theta_G-\pi_Q^*\mathbf{A})\wedge(\pi_G^*\theta_G-\pi_Q^*\mathbf{A})\Big)\ ,$$

where we view $\pi_{\mathfrak{g}^*}:Q\times G\times S\to\mathfrak{g}^*$ as an element of $\Omega^0(Q\times G\times S,\mathfrak{g}^*)$. We can now take d of both sides, obtaining

$$\begin{aligned}(9)\quad & b^*d(\omega_E)_Q=\\ &-\frac{1}{2}d\pi_{\mathfrak{g}^*}\wedge\Big((\theta'_G-\mathbf{A}'')\wedge(\theta'_G-\mathbf{A}'')\Big)+\pi_{\mathfrak{g}^*}\wedge\Big((\theta'_G-\mathbf{A}'')\wedge d(\theta'_G-\mathbf{A}'')\Big)\ ,\end{aligned}$$

where a single prime indicates pullback by π_G, and a double prime indicates pullback by π_Q. We expand and simplify (9) by invoking the following identities:

$$(10)\qquad d\theta'_G=\frac{1}{2}\theta'_G\wedge\theta'_G\ ,$$

$$(11)\qquad d\mathbf{A}''=(\mathbf{DA})''+\frac{1}{2}\mathbf{A}''\wedge\mathbf{A}''\ ,$$

$$(12)\qquad \theta'_G\wedge(\theta'_G\wedge\theta'_G)=0\ ,$$

$$(13)\qquad \mathbf{A}''\wedge(\mathbf{A}''\wedge\mathbf{A}'')=0\ .$$

If the primes are suppressed, then (10) and (11) are the Mauer-Cartan equations for G and the principal bundle Q resp., while (12) and (13) follow from Jacobi's identity. That we may add the primes follows from the fact that d commutes with pullbacks, and that pullbacks distribute over wedge products. After some manipulation, Equation (9) becomes

$$\begin{aligned} b^* d(\omega_E)_Q = &-\frac{1}{2} d\pi_{\mathfrak{g}^*} \wedge (\mathbf{A}'' \wedge \mathbf{A}'') - \frac{1}{2} d\pi_{\mathfrak{g}^*} \wedge (\theta'_G \wedge \theta'_G) \\ &+ \pi_{\mathfrak{g}^*} \wedge \Big(\mathbf{A}'' \wedge (\mathbf{DA})'' \Big) - \pi_{\mathfrak{g}^*} \wedge \Big(\theta'_G \wedge (\mathbf{DA})'' \Big) \\ &- \frac{1}{2} \pi_{\mathfrak{g}^*} \wedge \Big(\mathbf{A}'' \wedge (\theta'_G \wedge \theta'_G) \Big) - \frac{1}{2} \pi_{\mathfrak{g}^*} \wedge \Big(\theta'_G \wedge (\mathbf{A}'' \wedge \mathbf{A}'') \Big) \ . \end{aligned} \tag{14}$$

For future reference, we note here the easily computed formula

$$d\pi_{\mathfrak{g}^*}(\langle u, \eta, \xi, v;\, q, g, s \rangle) = -\operatorname{ad}^*_\xi (g \cdot s) + g \cdot v \ . \tag{15}$$

By (5), we have

$$\frac{d}{dt} \left[q^{\mathrm{hor}}(t), \exp(t\xi) \cdot (\mu + tv) \right]_G \Big|_{t=0} = \mathrm{T}b \cdot \langle u, 0, \xi, v;\, q, \mathrm{id}, \mu \rangle \ ,$$

so that

$$\begin{aligned} d(\omega_E)_Q \bigg(&\frac{d}{dt} [q_1^{\mathrm{hor}}(t), \exp(t\xi_1) \cdot (\mu + tv_1)]_G \Big|_{t=0}, \\ &\frac{d}{dt} [q_2^{\mathrm{hor}}(t), \exp(t\xi_2) \cdot (\mu + tv_2)]_G \Big|_{t=0}, \\ &\frac{d}{dt} [q_3^{\mathrm{hor}}(t), \exp(t\xi_3) \cdot (\mu + tv_3)]_G \Big|_{t=0} \bigg) \\ = b^* d(\omega_E)_Q(&\langle u_1, 0, \xi_1, v_1;\, q, \mathrm{id}, \mu \rangle, \langle u_2, 0, \xi_2, v_2;\, q, \mathrm{id}, \mu \rangle, \langle u_3, 0, \xi_3, v_3;\, q, \mathrm{id}, \mu \rangle), \end{aligned}$$

We now substitute the formula for $b^* d(\omega_E)_Q$ in (14). In fact, since

$$\mathbf{A}''(\langle u_j, 0, \xi_j, v_j;\, q, \mathrm{id}, \mu \rangle) = 0 \qquad (j = 1,\ 2 \text{ or } 3) \ ,$$

the only part on the right-hand side of (14) with a nontrivial contribution is

$$-\frac{1}{2} d\pi_{\mathfrak{g}^*} \wedge (\theta'_G \wedge \theta'_G) - \pi_{\mathfrak{g}^*} \wedge \Big(\theta'_G \wedge (\mathbf{DA})'' \Big)$$

and we obtain, with the help of (15),

$$\begin{aligned}
d(\omega_E)_Q \biggl(& \frac{d}{dt}\, [q_1^{\rm hor}(t), \exp(t\xi_1)\cdot(\mu + tv_1)]_G \Big|_{t=0}, \\
& \frac{d}{dt}\, [q_2^{\rm hor}(t), \exp(t\xi_2)\cdot(\mu + tv_2)]_G \Big|_{t=0}, \\
& \frac{d}{dt}\, [q_3^{\rm hor}(t), \exp(t\xi_3)\cdot(\mu + tv_3)]_G \Big|_{t=0} \biggr) \\
= {} & \langle \mathrm{ad}^*_{\xi_1} \mu - v_1, [\xi_2, \xi_3] \rangle \;+\; \text{cyclic terms} \\
& - \; \langle \mu, [\xi_1, \mathbf{DA}(\dot q_2^{\rm hor}(0), \dot q_3^{\rm hor}(0))] \rangle \;-\; \text{cyclic terms} \\
= {} & \langle \mu, [\xi_1, [\xi_2, \xi_3]] \;+\; \text{cyclic terms} \rangle \\
& - \; \langle v_1, [\xi_2, \xi_3] \rangle \;-\; \text{cyclic terms} \\
& - \; \langle \mu, [\xi_1, \mathbf{DA}(\dot q_2^{\rm hor}(0), \dot q_3^{\rm hor}(0))] \rangle \;-\; \text{cyclic terms}\ .
\end{aligned}$$

The term appearing in the third last row vanishes, by Jacobi's identity, and what is left amounts to Equation (4). □

One computes, using Lemma 6.3 and the definition of D_{μ_0} in 2.2,

$$\begin{aligned}
\Bigl\langle \nu, D_{\mu_0}\omega_D \Bigl(& \frac{d}{dt}\, z_1(t) \oplus [q_1^{\rm hor}(t), \exp(t\xi_1)\cdot \mu_0]_G \Big|_{t=0}, \\
& \frac{d}{dt}\, z_2(t) \oplus [q_2^{\rm hor}(t), \exp(t\xi_2)\cdot \mu_0]_G \Big|_{t=0} \Bigr) \Bigr\rangle \\
= d\omega_D \Bigl(& \frac{d}{dt}\, z \oplus [q, \mu_0 + t\iota_{\mu_0}^{-1}(\nu)]_G \Big|_{t=0}, \\
& \frac{d}{dt}\, z_1(t) \oplus [q_1^{\rm hor}(t), \exp(t\xi_1)\cdot \mu_0]_G \Big|_{t=0}, \\
& \frac{d}{dt}\, z_2(t) \oplus [q_2^{\rm hor}(t), \exp(t\xi_2)\cdot \mu_0]_G \Big|_{t=0} \Bigr) \\
= \quad & - \Bigl\langle \iota_{\mu_0}^{-1}(\nu),\ [\xi_1, \xi_2] \Bigr\rangle - \Bigl\langle \iota_{\mu_0}^{-1}(\nu),\ \mathbf{DA}(\dot q_1^{\rm hor}(0), \dot q_2^{\rm hor}(0)) \Bigr\rangle \\
= \quad & \Bigl\langle \nu, -\mathrm{pr}_{\mu_0} \Bigl([\xi_1, \xi_2] + \mathbf{DA}(\dot q_1^{\rm hor}(0), \dot q_2^{\rm hor}(0) \Bigr) \Bigr\rangle\ .
\end{aligned}$$

The second equality follows from Equations (1)–(4) derived above; the last equality follows from 6.1(4). Since $\nu \in \mathfrak{g}^*_{\mu_0}$ in this computation is arbitrary, we conclude that

$$\begin{aligned}
D_{\mu_0}\omega_D \Bigl(& \frac{d}{dt}\, z_1(t) \oplus [q_1^{\rm hor}(t), \exp(t\xi_1)\cdot \mu_0]_G \Big|_{t=0}, \\
& \frac{d}{dt}\, z_2(t) \oplus [q_2^{\rm hor}(t), \exp(t\xi_2)\cdot \mu_0]_G \Big|_{t=0} \Bigr) \\
= \quad & -\mathrm{pr}_{\mu_0}[\xi_1, \xi_2] - \mathrm{pr}_{\mu_0} \mathbf{DA}(\dot q_1^{\rm hor}(0), \dot q_2^{\rm hor}(0))\ .
\end{aligned} \tag{16}$$

Proposition 2.8 and the definition 3.2(1) of associated forms delivers the formula

$$(17)\qquad (\operatorname{curv}\alpha_{\mu_0})_Q\Big(\frac{d}{dt}\,[q_1^{\mathrm{hor}}(t),\exp(t\xi_1)\cdot\mu_0]_G\,\Big|_{t=0},\\ \frac{d}{dt}\,[q_2^{\mathrm{hor}}(t),\exp(t\xi_2)\cdot\mu_0]_G\,\Big|_{t=0}\Big)=\operatorname{pr}_{\mu_0}[\xi_1,\xi_2]\ .$$

On the other hand, we have

$$(\rho_{\mathcal{O}}^*\operatorname{curv}\mathbf{A})\Big(\frac{d}{dt}\,[q_1^{\mathrm{hor}}(t),\exp(t\xi_1)\cdot\mu_0]_G\,\Big|_{t=0},\\ \frac{d}{dt}\,[q_2^{\mathrm{hor}}(t),\exp(t\xi_2)\cdot\mu_0]_G\,\Big|_{t=0}\Big)\\ =[[q,\mu_0]_G\oplus q,\mathbf{DA}(\dot q_1^{\mathrm{hor}}(0),\dot q_2^{\mathrm{hor}}(0))]_G\in\rho_{\mathcal{O}}^*\mathfrak{g}_Q\ .$$

Combining this with 7.2(3) gives

$$(18)\qquad ((\operatorname{Pr}_{\mu_0})_Q\wedge\rho_{\mathcal{O}}^*\operatorname{curv}\mathbf{A})\Big(\frac{d}{dt}\,[q_1^{\mathrm{hor}}(t),\exp(t\xi_1)\cdot\mu_0]_G\,\Big|_{t=0},\\ \frac{d}{dt}\,[q_2^{\mathrm{hor}}(t),\exp(t\xi_2)\cdot\mu_0]_G\,\Big|_{t=0}\Big)\\ =\operatorname{pr}_{\mu_0}\mathbf{DA}(\dot q_1^{\mathrm{hor}}(0),\dot q_2^{\mathrm{hor}}(0))\ .$$

Comparing the right-hand side of (16) with the right-hand sides of (17) and (18), we deduce the intrinsic formula

$$(19)\qquad D_{\mu_0}\omega_D=-\operatorname{pr}_2^*\Big((\operatorname{curv}\alpha_{\mu_0})_Q+(\operatorname{Pr}_{\mu_0})_Q\wedge\rho_{\mathcal{O}}^*\operatorname{curv}\mathbf{A}\Big)\\ =-\operatorname{pr}_2^*(\Omega_{\mu_0}+(\operatorname{Pr}_{\mu_0})_Q\wedge\Omega_{\mathbf{A}})\ .$$

The curve $t\mapsto\eta_t\in\mathcal{O}_Q$ in Theorem 3.4 is a closed embedded curve because it bounds the surface S. Because $z_t\oplus\eta_t$ and η_t have the same minimal period, it follows that there exists a smooth map $s:\partial S\to\mathrm{T}^*(Q/G)\oplus\mathcal{O}_Q$ such that $s(\eta_t)=z_t\oplus\eta_t$. As $\operatorname{pr}_2:\mathrm{T}^*(Q/G)\oplus\mathcal{O}_Q\to\mathcal{O}_Q$ is a vector bundle, the map s can be extended to a global section $s:\mathcal{O}_Q\to\mathrm{T}^*(Q/G)\oplus\mathcal{O}_Q$ of pr_2. This follows, for example, from [12, Theorem I.5.7]. Define $\Sigma\equiv s(S)$, so that $\operatorname{pr}_2(\Sigma)=S$ and $t\mapsto z_t\oplus\eta_t$ is the boundary of Σ. Appealing to Theorem 2.3 and (19), we obtain

$$g_{\mathrm{geom}}=\exp\int_\Sigma D_{\mu_0}\omega_D=\exp\Big(-\int_\Sigma\operatorname{pr}_2^*\,(\Omega_{\mu_0}+(\operatorname{Pr}_{\mu_0})_Q\wedge\Omega_{\mathbf{A}}\,)\Big)\\ =\exp\Big(-\int_S(\,\Omega_{\mu_0}+(\operatorname{Pr}_{\mu_0})_Q\wedge\Omega_{\mathbf{A}}\,)\Big)\ ,$$

which is the form of g_{geom} given in Theorem 3.4.

A On bundle-valued differential forms

The exterior calculus of differential forms taking values in a vector bundle is ordinarily constructed via Koszul (or 'affine') connections. See, for example, [10, Chap. 9] or [13, Chap. 17]. On the other hand, given an associated vector bundle V_Q (see 3.1 for notation), one can model the exterior calculus of V_Q-valued forms on the exterior covariant calculus of tensorial V-valued forms on Q. In place of a Koszul connection, one prescribes a principal connection on Q. (The corresponding Koszul connection ∇ appears as the $p = 0$ case of Lie derivatives of bundle-valued p-forms; see A.7.) When the vector bundle at hand is realized as an associated bundle, this latter approach, while equivalent to the former, is better suited to explicit computations. As we are unaware of a readily accessible account of it, we outline the basics here.

A.1 Notation

Let $\xi : E \to B$ be a vector bundle with base B and consider the (Abelian) category of real vector bundles over B, restricting attention to morphisms covering the identity on B. Denote by $\mathrm{Alt}^p(\mathrm{T}B, E)$ the bundle over B of all alternating p-linear bundle morphisms from $\mathrm{T}B \oplus \cdots \oplus \mathrm{T}B$ into E . Then an *E-valued differential p-form* is a smooth section of $\mathrm{Alt}^p(\mathrm{T}B, E) \to B$. The space of all such forms is denoted $\Omega^p(B, E)$.

A.2 Bundle-valued forms as tensorial vector-valued forms

Let $\rho : Q \to B$ be a principal G-bundle equipped with a connection one-form $\mathbf{A}$ and let V be a real vector space on which G acts linearly. Let $\Omega^p_{\mathrm{tens}}(Q, V)$ denote the space of tensorial V-valued forms on Q (for a definition of tensorial forms see, e.g., [12, Section II.5]). Here, as elsewhere, all actions are understood to be *left* actions (contrary to the convention adopted in [12]). As is well known, one has an isomorphism

$$\begin{gathered} \lambda \mapsto \hat\lambda \\ \Omega^p(B, V_Q) \xrightarrow{\sim} \Omega^p_{\mathrm{tens}}(Q, V) \end{gathered}$$

defined implicitly through the formula

$$\lambda(\mathrm{T}\rho \cdot u_1, \ldots, \mathrm{T}\rho \cdot u_p) = [q, \hat\lambda(u_1, \ldots, u_p)]_G \qquad (u_j \in \mathrm{T}_q Q,\ q \in Q) \ . \tag{1}$$

A.3 Pullbacks

If $f : B' \to B$ is a smooth map, then the pullback f^*Q of the principal bundle Q is defined by

$$f^*Q \equiv B' \oplus_B Q$$

(see 4.2 for notation). The manifold f^*Q is itself a principal G-bundle; its base space is B', the bundle projection is $b' \oplus q \to b'$, and G acts according to

$g \cdot (b' \oplus q) \equiv (b' \oplus g \cdot q)$. One defines a map $\hat{f} : f^*Q \to Q$ by $\hat{f}(b' \oplus q) \equiv q$ and has the commutative diagram

$$\begin{tikzcd} f^*Q \arrow[r, "\hat{f}"] \arrow[d] & Q \arrow[d, "\rho"] \\ B' \arrow[r, "f"'] & B \end{tikzcd} .$$

A connection one-form for $f^*Q \to B'$ is $\hat{f}^*\mathbf{A}$.
If $g : B'' \to B'$ is a second map, then a natural isomorphism $(f \circ g)^*Q \cong g^*(f^*Q)$ is given by

$$\begin{gathered} B'' \oplus_B Q \xrightarrow{\sim} B'' \oplus_{B'} (B' \oplus_B Q) \\ b'' \oplus q \to b'' \oplus (g(b'') \oplus q) \ . \end{gathered}$$

The pullback f^*V_Q of an associated vector bundle V_Q can be defined analogously but we will define it in a way making the pullback itself an associated bundle:

$$f^*V_Q \equiv V_{f^*Q} \ .$$

By the above we have $(f \circ g)^*V_Q \cong g^*(f^*V_Q)$. This definition of f^*V_Q is equivalent to the forementioned alternative, for we have an isomorphism

$$\begin{gathered} f^*V_Q \xrightarrow{\sim} B' \oplus_B V_Q \\ [b' \oplus q, v]_G \mapsto b' \oplus [q, v]_G \ . \end{gathered}$$

The map $f : B' \to B$ defines a pullback operator on forms $f^* : \Omega^p(B, V_Q) \to \Omega^p(B', f^*V_Q)$ defined through

$$(f^*\lambda)\widehat{\ }= \hat{f}^*\hat{\lambda} \ ,$$

where the pullback on the right-hand side is the usual one for vector-valued forms. Making the identification $(f \circ g)^*V_Q \cong g^*(f^*V_Q)$ indicated above, we have $(f \circ g)^* = g^* \circ f^*$.

A.4 Wedge products

The wedge product $\lambda \wedge \mu \in \Omega^{p+q}(B, (U \otimes V)_Q)$ of forms $\lambda \in \Omega^p(B, U_Q)$ and $\mu \in \Omega^q(B, V_Q)$ is defined through

$$(\lambda \wedge \mu)\widehat{\ }= \hat{\lambda} \wedge \hat{\mu} \ .$$

Suppose there is a natural, bilinear pairing $(u, v) \mapsto \langle u, v \rangle : U \times V \to W$ that is equivariant in the sense that $\langle g \cdot u, g \cdot v \rangle = g \cdot \langle u, v \rangle$. Then there is a G-invariant homomorphism $U \otimes V \to W$ allowing one to identify $\hat{\lambda} \wedge \hat{\mu}$ with an element of $\Omega^{p+q}_{\mathrm{tens}}(Q,W)$; $\lambda\wedge\mu$ is correspondingly identified with an element of $\Omega^{p+q}(B,W_Q)$. In the special case that G acts trivially on W (e.g., $W=\mathbb{R}$), one has $W_Q\cong W\times Q$ and there is a further identification $\Omega^{p+q}(B,W_Q)\cong\Omega^{p+q}(B,W)$.

A.5 Exterior derivatives

The exterior derivative $d\lambda\in\Omega^{p+1}(B,V_Q)$ of a form $\lambda\in\Omega^p(B,V_Q)$ is defined through

$$(d\lambda)\widehat{\ }=\mathbf{D}\hat{\lambda}\ ,$$

where $\mathbf{D}$ denotes exterior covariant derivative with respect to the connection $\mathbf{A}$ (see [12]).

A.6 Curvature

We next define the *curvature form* $\mathbf{B}_V$, which measures the degree to which Poincaré's identity $d^2=0$ fails for V_Q-valued differential forms.

By its equivariance, a tensorial zero-form $F\in\Omega^0_{\mathrm{tens}}(Q,V)$ satisfies the identity

$$dF\left(\frac{d}{dt}\exp(t\xi)\cdot q\Big|_{t=0}\right)=\mathrm{ad}^V_\xi F(q)\qquad(\xi\in\mathfrak{g},\ q\in Q)\ ,$$

where ad^V_ξ denotes the infinitesimal generator of the linear action of G on V along ξ, viewed as an element of $\mathrm{Hom}(V,V)$. From the definition of exterior covariant derivative, one deduces the identity $\mathbf{D}F=dF-\mathbf{A}_V\wedge F$, where $\mathbf{A}_V\in\Omega^1(Q,\mathrm{Hom}(V,V))$ is defined by

$$\mathbf{A}_V(u)\equiv\mathrm{ad}^V_{\mathbf{A}(u)}\qquad(u\in\mathrm{T}Q)\ .$$

It follows that $\mathbf{D}^2F=-\mathbf{D}\mathbf{A}_V\wedge F$. Note that by the linearity of $\xi\mapsto\mathrm{ad}^V_\xi$, we have

$$\mathbf{D}\mathbf{A}_V(u_1,u_2)=\mathrm{ad}^V_{\mathbf{D}\mathbf{A}(u_1,u_2)}\qquad(u_1,u_2\in\mathrm{T}_qQ,\ q\in Q)\ .$$

The two-form $\mathbf{D}\mathbf{A}_V$ is tensorial (with G acting on $\mathrm{Hom}(V,V)$ by conjugation), and so defines a two-form $\mathbf{B}_V\in\Omega^2(B,\mathrm{Hom}(V,V)_Q)$ through[4]

$$\hat{\mathbf{B}}_V=-\mathbf{D}\mathbf{A}_V\ ,$$

allowing us to write $\mathbf{D}^2F=\hat{\mathbf{B}}_V\wedge F$. Moreover, one can show that F in this identity can be replaced by an arbitrary, tensorial, V-valued p-form. One does so using the fact that such a form is an $\mathbb{R}$-linear combination of products of the form $\omega\wedge F$, for some $\omega\in\Omega^p_{\mathrm{tens}}(Q,\mathbb{R})$ and $F\in\Omega^0_{\mathrm{tens}}(Q,V)$. In particular, replacing F by $\hat{\lambda}$ ($\lambda\in\Omega^p(B,V_Q)$), one deduces the important identity

$$d^2\lambda=\mathbf{B}_V\wedge\lambda\qquad(\lambda\in\Omega^p(B,V_Q))\ .\tag{1}$$

[4]We have inserted a minus sign in the formula defining $\mathbf{B}_V$ to ensure that the identity (1) conforms with the case of *right* principal bundles, as well as the theory as developed via Koszul connections.

Notice that $d^2 = 0$ if and only if $\mathbf{B}_V = 0$, which is true if and only if G acts trivially on V or $\mathbf{A}$ is a flat connection.

Note. The two-form $\operatorname{curv}\mathbf{A} \in \Omega^2(B, \mathfrak{g}_Q)$ defined through $(\operatorname{curv}\mathbf{A})\widehat{\ } = \mathbf{DA}$ is known as the *curvature of* $\mathbf{A}$. It is related to $\mathbf{B}_\mathfrak{g}$ in the following way: $\mathbf{B}_\mathfrak{g}$ is the image of $\operatorname{curv}\mathbf{A}$ under the natural map $\Omega^2(B, \mathfrak{g}_Q) \to \Omega^2(B, \operatorname{Hom}(\mathfrak{g}, \mathfrak{g})_Q)$ induced by $\xi \mapsto -\operatorname{ad}_\xi : \mathfrak{g} \to \operatorname{Hom}(\mathfrak{g}, \mathfrak{g})$.

A.7 Lie derivatives and interior products

If X is a vector field on B and X^h denotes its $\mathbf{A}$-horizontal lift to a vector field on Q, then the (covariant) Lie derivative $\mathcal{L}_X : \Omega^p(B, V_Q) \to \Omega^p(B, V_Q)$ (often denoted ∇_X) is defined through

$$(\mathcal{L}_X \lambda)\widehat{\ } = \mathcal{L}_{X^h} \hat{\lambda} \qquad (\lambda \in \Omega^p(B, V_Q)) \ ,$$

where $\mathcal{L}_{X^h} : \Omega^p(Q, V) \to \Omega^p(Q, V)$ is the standard Lie derivative along X^h. The interior product (contraction) $X \lrcorner \lambda$ of a vector field X on B with a V_Q-valued p-form λ satisfies the identity

$$(X \lrcorner \lambda)\widehat{\ } = X^h \lrcorner \hat{\lambda} \ .$$

A.8

Many familiar identities generalize to the vector bundle case, despite the fact that $d^2 \neq 0$. These are derived by simply dropping the appropriate identity for tensorial forms. For example, one has the well known identity

$$\frac{d}{dt} \exp\Big(t\mathbf{DA}(X^h(q), Y^h(q)) \Big) \cdot q \,\Big|_{t=0} = [X, Y]^h(q) - [X^h, Y^h](q) \qquad (q \in Q) \ ,$$

which, for an arbitrary zero-form $F \in \Omega^0(B, V_Q)$, implies

$$\mathbf{DA}_V(X^h, Y^h) \wedge \hat{F} = \mathcal{L}_{[X,Y]^h} \hat{F} - \mathcal{L}_{X^h} \mathcal{L}_{Y^h} \hat{F} + \mathcal{L}_{Y^h} \mathcal{L}_{X^h} \hat{F} \ .$$

Dropping to B, we obtain

$$\mathbf{B}_V(X, Y) \wedge F = \mathcal{L}_X \mathcal{L}_Y F - \mathcal{L}_Y \mathcal{L}_X F - \mathcal{L}_{[X,Y]} F \ ,$$

which characterizes $\mathbf{B}_V$ in terms of (covariant) Lie derivatives.
Similarly, just as $\mathbf{D}$ is an antiderivation on tensorial vector-valued forms, d is an antiderivation on bundle-valued forms, i.e.,

$$d(\lambda \wedge \mu) = d\lambda \wedge \mu + (-1)^p \lambda \wedge d\mu \qquad (\lambda \in \Omega^p(B, V_Q),\ \mu \in \Omega^q(B, V_Q)) \ .$$

In particular, applying d to both sides of A.6(1) gives

$$d^2(d\lambda) = d\mathbf{B}_V \wedge \lambda + \mathbf{B}_V \wedge d\lambda \ .$$

Replacing λ in A.6(1) by $d\lambda$ and substituting into the above equation yields

$$\begin{aligned} & \mathbf{B}_V \wedge d\lambda = d\mathbf{B}_V \wedge \lambda + \mathbf{B}_V \wedge d\lambda \\ \Rightarrow \qquad & d\mathbf{B}_V \wedge \lambda = 0 \ . \end{aligned}$$

Since $\lambda \in \Omega^p(B, V_Q)$ is arbitrary, we conclude that $d\mathbf{B}_V = 0$ (Bianchi's identity).

B On regular points of the co-adjoint action

This appendix is devoted to the proof of Lemma 6.1. While this could be done using standard structure theory, we opt for a 'direct' proof based on the following well known fact:[5]

Theorem (Duflo-Vergne [11]). *Let G be any finite-dimensional Lie group. Then $\mathfrak{g}^*_{\mathrm{reg}}$ is open and dense in $\mathfrak{g}^*$. Furthermore, for all $\mu \in \mathfrak{g}^*_{\mathrm{reg}}$, $\mathfrak{g}_\mu$ is Abelian.*

Fix an Ad-invariant inner product on $\mathfrak{g}$ and equip $\mathfrak{g}^*$ with the corresponding Ad^*-invariant inner product. The product on $\mathfrak{g}$ defines an equivariant isomorphism $\rho : \mathfrak{g} \xrightarrow{\sim} \mathfrak{g}^*$, establishing an equivalence between the adjoint and co-adjoint representations. We therefore begin by examining the adjoint action, where computations are easier.
Put $\lambda \equiv \rho^{-1}(\mu)$ ($\mu \in \mathfrak{g}^*_{\mathrm{reg}}$). By invariance of the inner product on $\mathfrak{g}$, the map $\xi \mapsto [\xi, \lambda]$ is skew-symmetric. Its kernel $\mathfrak{g}_\lambda$ and image $[\mathfrak{g}, \lambda]$ are therefore orthogonal:

$$\mathfrak{g}_\lambda^\perp = [\mathfrak{g}, \lambda] \ . \tag{1}$$

In addition, we claim that

$$[\mathfrak{g}, \lambda] = [\mathfrak{g}, \mathfrak{g}_\lambda] \ . \tag{2}$$

Proof of (2). Equation (1) implies that $\mathfrak{g} = \mathfrak{g}_\lambda + [\mathfrak{g}, \lambda]$ so that for an arbitrary $\eta \in \mathfrak{g}_\lambda$ we have

$$[\mathfrak{g}, \eta] = [\mathfrak{g}_\lambda + [\mathfrak{g}, \lambda], \eta] = [[\mathfrak{g}, \lambda], \eta] \ . \tag{3}$$

The second equality holds because $\mathfrak{g}_\lambda$ is Abelian (by the Theorem and the equivalence of the adjoint and co-adjoint representations). It follows from (3) that an arbitrary element of $[\mathfrak{g}, \eta]$ is of the form $[[\xi, \lambda], \eta]$, for some $\xi \in \mathfrak{g}$. But, by Jacobi's identity, we have

$$\begin{aligned} & [[\xi, \lambda], \eta] = -[[\eta, \xi], \lambda] - [[\lambda, \eta], \xi] \\ & \qquad\qquad = -[[\eta, \xi], \lambda] \ , \quad \text{since } [\lambda, \eta] = 0 \ (\mathfrak{g}_\lambda \text{ is Abelian}) \\ \Rightarrow \quad & [[\xi, \lambda], \eta] \in [\mathfrak{g}, \lambda] \ . \end{aligned}$$

So $[\mathfrak{g}, \eta] \subset [\mathfrak{g}, \lambda]$. Since $\eta \in \mathfrak{g}_\lambda$ was arbitrary, we conclude that $[\mathfrak{g}, \mathfrak{g}_\lambda] \subset [\mathfrak{g}, \lambda]$. As the reverse containment is obvious, Equation (2) is established. □

[5]Proofs of this theorem in English are given in [18] and [25].

Together (1) and (2) imply that $\mathfrak{g}_\lambda^\perp = [\mathfrak{g}, \mathfrak{g}_\lambda]$. By the equivalence of the adjoint and co-adjoint representations, we have $\mathfrak{g}_\mu^\perp = [\mathfrak{g}, \mathfrak{g}_\mu]$, as claimed.

References

[1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Addison-Wesley Publishing Co., Reading, Massachusetts, 2nd edition, 1978.

[2] M. S. Alber, G. G. Luther, J. E. Marsden, and J. M. Robbins. Geometric phases, reduction and Lie-Poisson structure for the resonant three-wave interaction. *Phys. D*, 123:271–290, 1998.

[3] M. S. Alber and J. E. Marsden. On geometric phases for soliton equations. *Comm. Math. Phys*, 149:217–240, 1992.

[4] V. I. Arnold. *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 1989.

[5] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. *Dynamical Systems III.* Encycl. Math. Sci. Springer, New York, 1988.

[6] A. D. Blaom. Reconstruction phases via Poisson reduction. *Differential Geom. Appl.*, 12(3):231–252, 2000.

[7] A. D. Blaom. A geometric setting for Hamiltonian perturbation theory. *Mem. Amer. Math. Soc.*, 153(727):1–112, 2001.

[8] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, and G. Sánchez de Alvarez. Stabilization of rigid body dynamics by internal and external torques. *Automatica J. IFAC*, 28:745–756, 1992.

[9] H. Cendra, D. D. Holm, J. E. Marsden, and T. S. Ratiu. Lagrangian reduction, the Euler-Poincaré equations, and semidirect products. *Amer. Math. Soc. Transl. Ser. 2*, 186:1–25, 1998.

[10] R. W. R. Darling. *Differential Forms and Connections*. Cambridge University Press, 1994.

[11] M. Duflo and M. Vergne. Une proprieté de la représentation coadjointe d'une algébre de Lie. *C. R. Acad. Sci. Paris Sér. A-B*, 268:583–585, 1969.

[12] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry, Volume I.* Wiley, New York, 1963.

[13] I. Madsen and J. Tornehave. *From Calculus to Cohomology.* Cambridge University Press, 1997.

[14] J. E. Marsden. *Lectures on Mechanics*, volume 174 of *London Mathematical Society Lecture Notes Series.* Cambridge University Press, 1992.

[15] J. E. Marsden. Park City lectures on mechanics, dynamics, and symmetry. In *Symplectic geometry and topology (Park City, UT, 1997)*, pages 335–430. Amer. Math. Soc., Providence, RI, 1999.

[16] J. E. Marsden, R. Montgomery, and T. Ratiu. Reduction, symmetry, and phases in mechanics. *Mem. Amer. Math. Soc.*, 88(436):1–110, 1990.

[17] J. E. Marsden and J. Ostrowski. Symmetries in motion: Geometric foundations of motion control. *Nonlinear Sci. Today*, January, 1998.

[18] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer, 1994.

[19] J. E. Marsden, T. S. Ratiu, and J. Scheurle. Reduction theory and the Lagrange-Routh Equations. *J. Math. Phys.*, 41:3379–3429, 2000.

[20] J. E. Marsden and A. Weinstein. Reduction of symplectic manifolds with symmetry. *Rep. Math. Phys.*, 5:121–130, 1974.

[21] K. R. Meyer. Symmetries and Integrals in Mechanics. In M. Peixoto, editor, *Dynamical Systems*, pages 259–273. Academic Press, New York, 1973.

[22] R. Montgomery. *The Bundle Picture in Mechanics*. PhD thesis, University of California, Berkeley, 1986.

[23] R. Montgomery. How much does a rigid body rotate? A Berry's phase from the 18th century. *Amer. J. Phys.*, 59:394–398, 1991.

[24] M. Perlmutter. *Symplectic Reduction by Stages*. PhD thesis, University of California, Berkeley, 1999.

[25] A. Weinstein. The local structure of Poisson manifolds. *J. Differential Geom.*, 18:523–557, 1983.

Anthony D. Blaom
Washington, D. C. and
Princeton, New Jersey
Current address:
Department of Mathematics
University of Auckland
Private Bag 92019
Auckland, New Zealand
a.blaom@auckland.ac.nz
www.math.auckland.ac.nz/~blaom

Group C^*-Algebras as Compact Quantum Metric Spaces

Marc A. Rieffel [1]

Received: September 26, 2002
Revised: December 20, 2002

Communicated by Joachim Cuntz

Abstract. Let ℓ be a length function on a group G, and let M_ℓ denote the operator of pointwise multiplication by ℓ on $\ell^2(G)$. Following Connes, M_ℓ can be used as a "Dirac" operator for $C_r^*(G)$. It defines a Lipschitz seminorm on $C_r^*(G)$, which defines a metric on the state space of $C_r^*(G)$. We investigate whether the topology from this metric coincides with the weak-$*$ topology (our definition of a "compact quantum metric space"). We give an affirmative answer for $G = \mathbb{Z}^d$ when ℓ is a word-length, or the restriction to $\mathbb{Z}^d$ of a norm on $\mathbb{R}^d$. This works for $C_r^*(G)$ twisted by a 2-cocycle, and thus for non-commutative tori. Our approach involves Connes' cosphere algebra, and an interesting compactification of metric spaces which is closely related to geodesic rays.

2000 Mathematics Subject Classification: Primary 47L87; Secondary 20F65, 53C23, 58B34
Keywords and Phrases: Group C^*-algebra, Dirac operator, quantum metric space, metric compactification, boundary, geodesic ray, Busemann point.

0. Introduction

The group C^*-algebras of discrete groups provide a much-studied class of "compact non-commutative spaces" (that is, unital C^*-algebras). In [11] Connes showed that the "Dirac" operator of an unbounded Fredholm module over a unital C^*-algebra provides in a natural way a metric on the state space of the algebra. Unbounded Fredholm modules (i.e. spectral triples) also provide smooth structure, important homological data and much else. In the subsequent years Connes has been strongly advocating this use of Dirac operators as the way to deal with the Riemannian geometry of non-commutative spaces [12], [15], [14], [13]. The class of examples most discussed in [11] consists

[1]The research reported here was supported in part by National Science Foundation Grant DMS–99–70509.

of the group C^*-algebras of discrete groups, with the Dirac operator coming in a simple way from a length function on the group. Connes obtained in [11] strong relationships between the growth of a group and the summability of Fredholm modules over its group C^*-algebra. However he did not explore much the metric on the state space.

In [39], [40] I pointed out that, motivated by what happens for ordinary compact metric spaces, it is natural to desire that the topology from the metric on the state space coincides with the weak-$*$ topology (for which the state space is compact). This property was verified in [39] for certain examples, notably the non-commutative tori, with "metric" structure coming from a different construction. (See [40], [41], [42] for further developments.) But in general I have found this property to be difficult to verify for many natural examples.

The main purpose of this paper is to examine this property for Connes' initial class of examples, the group C^*-algebras with the Dirac operator coming from a length function. To be more specific, let G be a countable (discrete) group, and let $C_c(G)$ denote the convolution $*$-algebra of complex-valued functions of finite support on G. Let π denote the usual $*$-representation of $C_c(G)$ on $\ell^2(G)$ coming from the unitary representation of G by left translation on $\ell^2(G)$. The norm-completion of $\pi(C_c(G))$ is by definition the reduced group C^*-algebra, $C_r^*(G)$, of G. We identify $C_c(G)$ with its image in $C_r^*(G)$, so that it is a dense $*$-subalgebra.

Let a length function ℓ be given on G. We let M_ℓ denote the (usually unbounded) operator on $\ell^2(G)$ of pointwise multiplication by ℓ. Then M_ℓ will serve as our "Dirac" operator. One sees easily [11] that the commutators $[M_\ell, \pi_f]$ are bounded operators for each $f \in C_c(G)$. We can thus define a seminorm, L_ℓ, on $C_c(G)$ by $L_\ell(f) = \|[M_\ell, \pi_f]\|$.

In general, if L is a seminorm on a dense $*$-subalgebra A of a unital C^*-algebra $\bar{A}$ such that $L(1) = 0$, we can define a metric, ρ_L, on the state space $S(\bar{A})$ of $\bar{A}$, much as Connes did, by

$$\rho_L(\mu, \nu) = \sup\{|\mu(a) - \nu(a)| : a \in A,\ L(a) \leq 1\}.$$

(Without further hypotheses ρ_L may take value $+\infty$.) In [40] we define L to be a *Lip-norm* if the topology on $S(\bar{A})$ from ρ_L coincides with the weak-$*$ topology. We consider a unital C^*-algebra equipped with a Lip-norm to be a compact quantum metric space [40].

The main question dealt with in this paper is whether the seminorms L_ℓ coming as above from length functions on a group are Lip-norms. In the end we only have success in answering this question for the groups $\mathbb{Z}^d$. The situation there is already somewhat complicated because of the large variety of possible length-functions. But we carry out our whole discussion in the slightly more general setting of group C^*-algebras twisted by a 2-cocycle (definitions given later), and so this permits us to treat successfully also the non-commutative tori [38]. The main theorem of this paper is:

Main Theorem 0.1. *Let ℓ be a length function on $\mathbb{Z}^d$ which is either the word-length function for some finite generating subset of $\mathbb{Z}^d$, or the restriction* *to $\mathbb{Z}^d$ of some norm on $\mathbb{R}^d$. Let c be a 2-cocycle on $\mathbb{Z}^d$, and let π be the regular representation of $C^*(\mathbb{Z}^d, c)$ on $\ell^2(\mathbb{Z}^d)$. Then the seminorm L_ℓ defined on $C_c(\mathbb{Z}^d)$ by $L_\ell(f) = \|[M_\ell, \pi_f]\|$ is a Lip-norm on $C^*(\mathbb{Z}^d, c)$.*

The path which I have found for the proof of this theorem is somewhat long, but it involves some objects which are of considerable independent interest, and which may well be useful in treating more general groups. Specifically, we need to examine Connes' non-commutative cosphere algebra [14] for the examples which we consider. This leads naturally to a certain compactification which one can construct for any locally compact metric space. We call this "the metric compactification". Actually, this compactification had been introduced much earlier by Gromov [24], but it is different from the famous Gromov compactification for a hyperbolic metric space, and it seems not to have received much study. Our approach gives a new way of defining this compactification. We also need to examine the strong relationship between geodesic rays and points in the boundary of this compactification, since this will provide us with enough points of the boundary which have finite orbits. For word-length functions on $\mathbb{Z}^d$ this is already fairly complicated.

The contents of the sections of this paper are as follows. In Section 1 we make more precise our notation, and we make some elementary observations showing that on any separable unital C^*-algebra there is an abundance of Lip-norms, and that certain constructions in the literature concerning groups of "rapid decay" yield natural Lip-norms on $C_r^*(G)$. In Section 2 we begin our investigation of the Dirac operators for $C_r^*(G)$ coming from length functions. In Section 3 we examine Connes' cosphere algebra for our situation. We show in particular that if the action of the group on the boundary of its metric compactification is amenable, then the cosphere algebra has an especially simple description. Then in Section 4 we study the metric compactification in general, with attention to the geodesic rays.

In Section 5 we begin our study of specific groups by considering the group $\mathbb{Z}$. This is already interesting. (Consider a generating set such as $\{\pm 3, \pm 8\}$.) The phenomena seen there for $\mathbb{Z}$ indicate some of the complications which we will encounter in trying to deal with $\mathbb{Z}^d$. In Section 6 we study the metric compactification of $\mathbb{R}^d$ for any given norm, and then in Section 7 we apply this to prove the part of our Main Theorem for length functions on $\mathbb{Z}^d$ which are the restrictions of norms on $\mathbb{R}^d$. In Section 8 we study the metric compactification of $\mathbb{Z}^d$ for word-length functions, and in Section 9 we apply this to prove the remaining part of our Main Theorem. We conclude in Section 10 with a brief examination of the free (non-Abelian) group on two generators, to see both how far our approach works, and where we become blocked from proving for it the corresponding version of our Main Theorem.

Last-minute note: I and colleagues believe we have a proof that the Main Theorem is also true for word-hyperbolic groups with word-length functions, using techniques which are entirely different from those used here, and which do not seem to apply to the case of $\mathbb{Z}^d$ treated here.

A substantial part of the research reported here was carried out while I visited the Institut de Mathématique de Luminy, Marseille, for three months. I would like to thank Gennady Kasparov, Etienne Blanchard, Antony Wasserman, and Patrick Delorme very much for their warm hospitality and their mathematical stimulation during my very enjoyable visit.

1. An abundance of Lip-norms

In this section we establish some of our notation, and show that on any separable unital C^*-algebra there is an abundance of Lip-norms. In the absence of further structure these Lip-norms appear somewhat artificial. But we then show that some known constructions for group C^*-algebras yield somewhat related but more natural Lip-norms.

Our discussion in the next few paragraphs works in the greater generality of order-unit spaces which was used in [40]. But we will not use that generality in later sections, and so the reader can have in mind just the case of dense unital *-subalgebras of unital C^*-algebras, with the identity element being the order unit. We recall that a (possibly discontinuous) seminorm L on an order-unit space is said to be lower semicontinuous if $\{a \in A : L(a) \leq r\}$ is norm-closed for any $r > 0$.

Proposition 1.1. *Let A be an order-unit space which is separable. For any countable subset E of A there are many lower semicontinuous Lip-norms on A which are defined and finite on E.*

Proof. The proof is a minor variation on the fact that the weak-$*$ topology on the unit ball of the dual of a separable Banach space is metrizable (theorem V.5.1 of [16]). We scale each non-zero element of E so that it is in the unit ball of A (and $\neq 0$), and we incorporate E into a sequence, $\{b_n\}$, of elements of A which is dense in the unit ball of A. Let $\{\omega_n\}$ be any sequence in $\mathbb{R}$ such that $\omega_n > 0$ for each n and $\Sigma\omega_n < \infty$. Define a norm, M, on the dual space A' of A by

$$M(\lambda) = \Sigma\omega_n|\lambda(b_n)|.$$

The metric from this norm, when restricted to the unit ball of A', gives the weak-$*$ topology, because it is easily checked that if a net in the unit ball of A' converges for the weak-$*$ topology then it converges for the metric from M, and then we can apply the fact that the unit ball is weak-$*$ compact.

We let $S(A)$ denote the state space of A. Since $S(A)$ is a subset of the unit ball of A', the restriction to $S(A)$ of the metric from the norm M gives $S(A)$ the weak-$*$ topology. Let L_M denote the corresponding Lipschitz seminorm on $C(S(A))$ from this metric, allowing value $+\infty$. View each element of A as a function on $S(A)$ in the usual way. Then $L_M(b_n) \leq \omega_n^{-1} < \infty$ for each n, because if $\mu, \nu \in S(A)$ then

$$|b_n(\mu) - b_n(\nu)| = |(\mu - \nu)(b_n)| \leq \omega_n^{-1}M(\mu - \nu) = \omega_n^{-1}\rho_M(\mu, \nu).$$

Let B denote the linear span of $\{b_n\}$ together with the order unit. Then B is a dense subspace of A containing the order-unit, and L_M restricted to B is a seminorm which can be verified to be lower semicontinuous. The inclusion of A into $C(S(A))$ is isometric (on self-adjoint elements if A is a C^*-algebra) and since L_M comes from an ordinary metric, it follows that L_M on A is a Lip-norm. (For example, use theorem 1.9 of [39].) □

The considerations above are close to those of theorem 9.8 of [40]. Let me take advantage of this to mention here that Hanfeng Li showed me by clever counterexample that theorem 9.8 of [40] is not correct as presented, because A may not be big enough. However, if A is taken to be norm-complete, then there is no difficulty. Theorem 9.11 needs to be adjusted accordingly. But this change does not affect later sections of [40] nor the subsequent papers [41], [42].

We now turn to (twisted) group C^*-algebras, and we use a different approach, which takes advantage of the fact that the group elements provide a natural "basis" for the group C^*-algebras. Thus let G be a countable discrete group, and let c be a 2-cocycle [47] on G with values in the circle group $\mathbb{T}$. We assume that c is normalized so that $c(x,y) = 1$ if $x = e$ or $y = e$. We let $C^*(G,c)$ denote the full c-twisted group C^*-algebra of G, and we let $C^*_r(G,c)$ denote the reduced c-twisted group C^*-algebra [47], [35] coming from the left regular representation, π, on $\boldsymbol{\ell}^2(G)$. Both C^*-algebras are completions of $C_c(G)$, the space of finitely supported $\mathbb{C}$-valued functions on G, with convolution twisted by c. Our conventions, following [47], are that

$$\begin{aligned}(f*g)(x) &= \Sigma f(y)g(y^{-1}x)c(y,y^{-1}x),\\ f^*(x) &= \bar{f}(x^{-1})\bar{c}(x,x^{-1}).\end{aligned}$$

The left regular representation is given by the same formula as the above twisted convolution, but with g viewed as an element of $\boldsymbol{\ell}^2(G)$. Then $C^*_r(G,c)$ is the completion of $C_c(G)$ for the operator norm coming from the left regular representation. We will often set $A = C^*_r(G,c)$. We note that when π is restricted to G we have

$$(\pi_y\xi)(x) = \xi(y^{-1}x)c(y,y^{-1}x)$$

for $\xi \in \boldsymbol{\ell}^2(G)$ and $x,y \in G$. In particular $\pi_y\pi_z = c(y,z)\pi_{yz}$.

There is a variety of norms on $C_c(G)$ which have been found to be useful in addition to the C^*-norms. These other norms are not necessarily algebra norms. To begin with, there is the $\boldsymbol{\ell}^1$-norm, as well as the $\boldsymbol{\ell}^p$-norms for $1 < p \le \infty$. But let ℓ be a length function on G, so that $\ell(xy) \le \ell(x) + \ell(y)$, $\ell(x^{-1}) = \ell(x)$, $\ell(x) \ge 0$, and $\ell(x) = 0$ exactly if $x = e$, the identity element of G. Then in connection with groups of "rapid decay" (such as word-hyperbolic groups) one defines norms on $C_c(G)$ of the following form [30], [29], [27], [28]:

$$\|f\|_{p,k} = (\Sigma(|f(x)|(1+\ell(x))^k)^p)^{1/p}.$$

These norms clearly have the properties that

1) $\|f\|_{p,k} \le \||f|\|_{p,k}$ (actually $=$),
2) if $|f| \le |g|$ then $\||f|\|_{p,k} \le \||g|\|_{p,k}$.

Their interest lies in the fact that for a rapid-decay group and an appropriate choice of p and k depending on the group, one has (see the first line of the proof of theorem 1.3 of [27], combined, in the case of nontrivial cocycle, with proposition 3.10b of [28]):

3) There is a constant, K, such that $\|f\|_{C_r^*} \leq K\|f\|_{p,k}$.

Notice also that if the cocycle c is trivial and if G is amenable [35] then the C^*-norm itself satisfies the above three properties, because from the trivial representation we see that for $f \in C_c(G)$ we have

$$\|f\|_{C^*(G)} \leq \|f\|_1 = \||f|\|_{C^*(G)},$$

while if $|f| \leq |g|$ then

$$\||f|\|_{C^*(G)} = \|f\|_1 \leq \|g\|_1 = \||g|\|_{C^*(G)}.$$

Finally, for any group and any cocycle we always have at least the ℓ^1-norm which satisfies the above three properties.

With these examples in mind, we make

Definition 1.2. *Let $\|\cdot\|_A$ denote the C^*-norm on $A = C_r^*(G,c)$. We will say that a norm, $\|\cdot\|$, on $C_c(G)$ is* order-compatible *with $\|\cdot\|_A$ if for all $f, g \in C_c(G)$ we have:*

1) $\|f\| \leq \||f|\|$.
2) *If $|f| \leq |g|$ then $\||f|\| \leq \||g|\|$.*
3) *There is a constant, K, such that $\|f\|_A \leq K\|f\|$.*

We remark that these conditions are a bit weaker than those required for a "good norm" in [32].

Suppose now that ω is a real-valued function on G such that $\omega(e) = 0$ and $\omega(x) > 0$ for $x \neq e$. Fix an order-compatible norm $\|\cdot\|$ on $C_c(G)$, and set

$$L(f) = \|\omega|f|\|.$$

It is clear that L is a seminorm which is 0 only on the span of the identity element of the convolution algebra $C_c(G,c)$. (Thus L is a Lipschitz seminorm as defined in [40].) In the way discussed in the introduction, L defines a metric, ρ_L, on $S(C_r^*(G,c))$ by

$$\rho_L(\mu,\nu) = \sup\{|\mu(f) - \nu(f)| : L(f) \leq 1\},$$

which may take value $+\infty$. Denote $C_r^*(G,c)$ by A, and its C*-norm by $\|\cdot\|_A$, as above.

Lemma 1.3. *Suppose that there is a constant $s > 0$ such that $\omega(x) \geq s$ for all $x \neq e$. Then ρ_L gives $S(A)$ finite radius. (In particular, ρ_L does not take the value $+\infty$.)*

Proof. Let $f \in C_c(G)$, and assume that $f(e) = 0$. Let K be the constant in the definition of "order-compatible". Then

$$\|f\|_A \leq K\|f\| \leq K\||f|\| \leq Ks^{-1}\|\omega|f|\| = Ks^{-1}L(f).$$

The desired conclusion then follows from proposition 2.2 of [40]. $\square$

Lemma 1.4. *Suppose that $\omega(x) = 0$ only if $x = e$ and that the function ω is "proper", in the sense that for any n the set $\{x \in G : \omega(x) \leq n\}$ is finite (so, in particular, there exists a constant s as in the above lemma). Then the topology from the metric ρ_L on $S(A)$ coincides with the weak-$*$ topology. Thus L is a Lip-norm.*

Proof. We apply theorem 1.9 of [39]. As in that theorem, we set

$$\mathcal{B}_1 = \{f \in C_c(G) : \|f\|_A \leq 1 \text{ and } L(f) \leq 1\}.$$

The theorem tells us that it suffices to show that $\mathcal{B}_1$ is totally bounded for $\|\cdot\|_A$. So let $\varepsilon > 0$ be given. Adjust K if necessary so that $K \geq 1$, and set

$$E = \{x \in G : \omega(x) \leq 3K/\varepsilon\}.$$

Then E is a finite set because ω is proper. Set $A^E = \{f \in C_c(G) : f(x) = 0 \text{ for } x \notin E\}$, so that A^E is a finite-dimensional subspace of $C_c(G)$. In particular, $A^E \cap \mathcal{B}_1$ is totally bounded.

Let $f \in \mathcal{B}_1$. Then $f = g + h$ where $g \in A^E$ and $h(x) = 0$ for $x \in E$. Now $|h| \leq |f|$, and $\omega(x) \geq 3K/\varepsilon$ on the support of h, and so

$$\begin{aligned} \|h\|_A &\leq K\|h\| \leq K\||h|\| \leq K(\varepsilon/3K)\|\omega|h|\| \\ &\leq (\varepsilon/3)\|\omega|f|\| = (\varepsilon/3)L(f) \leq \varepsilon/3. \end{aligned}$$

Thus $\|f - g\|_A = \|h\|_A \leq \varepsilon/3$. In particular, $\|g\|_A \leq 1 + (\varepsilon/3)$. Note also that $L(g) = \|\omega|g|\| \leq \|\omega|f|\| = L(f) \leq 1$. Thus upon scaling g by $(1 + \varepsilon/3)^{-1}$ if necessary to obtain an element of $\mathcal{B}_1$, we see that f is within distance $2\varepsilon/3$ of $\mathcal{B}_1 \cap A^E$. Thus a finite subset of $\mathcal{B}_1 \cap A^E$ which is $\varepsilon/3$ dense in $\mathcal{B}_1 \cap A^E$ will be ε-dense in $\mathcal{B}_1$. $\square$

Lemma 1.5. *Even without ω being proper, or satisfying the condition of Lemma 1.3, the seminorm L is lower semicontinuous (with respect to $\|\cdot\|_A$).*

Proof. Let $\{f_n\}$ be a sequence in $C_c(G)$ which converges to $g \in C_c(G)$ for $\|\cdot\|_A$, and suppose that there is an $r \in \mathbb{R}$ such that $L(f_n) \leq r$ for all n. Now $\pi_f\delta_0 = f$ where on the right f is viewed as an element of $\boldsymbol{\ell}^2$ and δ_0 is the "delta-function" at 0. Consequently $\|f\|_A \geq \|f\|_2 \geq \|f\|_\infty$. Thus f_n converges uniformly on G to g. Let S denote the support of g, and let χ_S be its characteristic function. Then the sequence $\omega\chi_S|f_n|$ converges uniformly to $\omega|g|$. But all norms on a finite-dimensional vector space are equivalent, and so $\omega\chi_S|f_n|$ converges to $\omega|g|$ for $\|\cdot\|$. This says that $L(\chi_S f_n)$ converges to $L(g)$. But $L(\chi_S f_n) = \|\omega\chi_S f_n\| \leq L(f) \leq r$. Thus $L(g) \leq r$. $\square$

We combine the above lemmas to obtain:

Proposition 1.6. *Let ω be a proper non-negative function on G such that $\omega(x) = 0$ exactly if $x = e$. Let $\|\cdot\|$ be an order-compatible norm on $C_c(G)$, and set*

$$L(f) = \|\omega|f|\|$$

for $f \in C_c(G)$. Then L is a lower semicontinuous Lip-norm on $C_r^*(G)$.

We remark that when ω is a length function on G and when $\|\cdot\| = \|\cdot\|_1$, it is well-known and easily seen that L satisfies the Leibniz rule with respect to $\|\cdot\|_1$, that is

$$L(f * g) \leq L(f)\|g\|_1 + \|f\|_1 L(g).$$

But there seems to be no reason why many of the above Lip-norms should satisfy the Leibniz rule with respect to $\|\cdot\|_A$. And it is not clear to me what significance the Leibniz rule has for the metric properties which we are examining.

2. Dirac operators from length functions

In this section we make various preliminary observations about the seminorms L which come from using length functions on a group as "Dirac" operators, as described in the introduction. We also reformulate our main question as concrete questions concerning $C_r^*(G)$ itself.

We use the notation of the previous section, and we let M_ℓ denote the (usually unbounded) operator on $\ell^2(G)$ of pointwise multiplication by the length function ℓ. We recall from [11] why the commutators $[M_\ell, \pi_f]$ are bounded for $f \in C_c(G)$. Let $y \in G$ and $\xi \in \ell^2(G)$. Then we quickly calculate that

$$([M_\ell, \pi_y]\xi)(x) = (\ell(x) - \ell(y^{-1}x))\xi(y^{-1}x)c(y, y^{-1}x).$$

From the triangle inequality for ℓ we know that $|\ell(x) - \ell(y^{-1}x)| \leq \ell(y)$, and so $\|[M_\ell, \pi_y]\| \leq \ell(y)$. In fact, this observation indicates the basic property of ℓ which we need for the elementary part of our discussion, namely that, although ℓ is usually unbounded, it differs from any of its left translates by a bounded function.

This suggests that we work in the more general context of functions having just this latter property, as this may clarify some aspects. Additional motivation for doing this comes from the importance which Connes has demonstrated for examining the effect of automorphisms of the C^*-algebra as gauge transformations, and the resulting effect on the metric. In Connes' approach the inner automorphisms play a distinguished role, giving "internal fluctuations" of the metric [9], [10] (called "internal perturbations" in [15]). However, in our setting we usually do not have available the "first order" condition which is crucial in Connes' setting. We discuss this briefly at the end of this section.

Anyway, in our setting the algebra $C_r^*(G, c)$ has some special inner automorphisms, namely those coming from the elements of G. The automorphism corresponding to $z \in G$ is implemented on $\ell^2(G)$ by conjugating by π_z. When this automorphism is composed with the representation, the effect is to change $D = M_\ell$ to $M_{\alpha_z(\ell)}$, where $\alpha_z(\ell)$ denotes the left translate of ℓ by z. But $\alpha_z(\ell)$ need not again be a length function, although it is translation bounded. (In order to try to clarify contexts, we will from now on systematically use α to denote ordinary left translation of functions, especially when those functions are not to be viewed as being in $\ell^2(G)$. Our convention is that $(\alpha_z\ell)(x) = \ell(z^{-1}x)$.)

We will make frequent use of the easily-verified commutation relation that

$$\pi_y M_h = M_{\alpha_y(h)}\pi_y$$

for any function h on G and any $y \in G$, as long as the domains of definitions of the product operators are respected. This commutation relation is what we used above to obtain the stated fact about the effect of inner automorphisms.

In what follows we will only use real-valued functions to define our Dirac operators, so that the latter are self-adjoint. But much of what follows generalizes easily to complex-valued functions, or to functions with values in C^*-algebras such as Clifford algebras. These generalizations deserve exploration.

To formalize our discussion above we make:

Definition 2.1. *We will say that a (possibly unbounded) real-valued function, ω, on G is (left)* translation-bounded *if $\omega - \alpha_y\omega$ is a bounded function for every $y \in G$. For $y \in G$ we set $\varphi_y = \omega - \alpha_y(\omega)$. So the context must make clear what ω is used to define φ. For each $y \in G$ we set $\ell^\omega(y) = \|\varphi_y\|_\infty$.*

Thus every length-function on G is translation-bounded. Any group homomorphism from G into $\mathbb{R}$ is translation bounded. (E.g., the homomorphism $\omega(n) = n$ from $\mathbb{Z}$ to $\mathbb{R}$ which is basically the Fourier transform of the usual Dirac operator on $\mathbb{T}$.) Linear combinations of translation-bounded functions are translation bounded. In particular, the sum of a translation-bounded function with any bounded function is translation bounded. (As a more general context one could consider any faithful unitary representation $(\pi, \mathcal{H})$ of G together with an unbounded self-adjoint operator D on $\mathcal{H}$ such that $D - \pi_z D\pi_z^*$ is densely defined and bounded for each $z \in G$, and D satisfies suitable non-triviality conditions. Our later discussion will indicate why one may also want to require that the $(\pi_z D\pi_z^*)$'s all commute with each other.)

It is simple to check that the φ_y's satisfy the 1-cocycle identity

$$\varphi_{yz} = \varphi_y + \alpha_y(\varphi_z). \tag{2.2}$$

We will make use of this relation a number of times. This type of relation occurs in various places in the literature in connection with dynamical systems.

Simple calculations show that ℓ^ω satisfies the axioms for a length function except that we may have $\ell^\omega(x) = 0$ for some $x \neq e$. Notice also that if ω is already a length function, then $\ell^\omega = \omega$. We also remark that in general we can always add a constant function to ω without changing the corresponding φ_y's, ℓ^ω, or the commutators $[M_\ell, \pi_y]$. In particular, we can always adjust ω in this way so that $\omega(e) = 0$ if desired.

We now fix a translation-bounded function, ω, on G, and we consider the operator, M_ω, of pointwise multiplication on $\ell^2(G)$. It is self-adjoint. We use it as a "Dirac operator". The calculation done earlier becomes

$$[M_\omega, \pi_y] = M_{\varphi_y}\pi_y.$$

From this we see that for each $y \in G$ we have

$$\|[M_\omega, \pi_y]\| = \ell^\omega(y).$$

For any $f \in C_c(G)$ we have

$$[M_\omega, \pi_f] = \Sigma f(y) M_{\varphi_y} \pi_y,$$

and consequently we have

$$\|[M_\omega, \pi_f]\| \leq \|\ell^\omega f\|_1,$$

where $\ell^\omega f$ denotes the pointwise product. We set

$$L^\omega(f) = \|[M_\omega, \pi_f]\|.$$

Then L^ω is a seminorm on $C_c(G) \subseteq C_r^*(G, c)$, and L^ω is lower semicontinuous by proposition 3.7 of [40]. A calculation above tells us that $L^\omega(\delta_x) = \ell^\omega(x)$ for all $x \in G$. In particular, $L^\omega(\delta_e) = 0$, with δ_e the identity element of the convolution algebra $C_c(G)$.

If we view δ_z as the usual basis element at z for $\ell^2(G)$, then for any $f \in C_c(G)$ we have

$$[M_\omega, \pi_f]\delta_z = \Sigma f(y) M_{\varphi_y} c(y, z) \delta_{yz}$$

for each z. From this we easily obtain:

Proposition 2.3. *Let $f \in C_c(G)$. Then $L^\omega(f) = 0$ exactly if $\varphi_y = 0$ for each y in the support of f, that is, exactly if $\ell^\omega f = 0$. Thus if $\ell^\omega(x) > 0$ for all $x \neq e$, then L^ω is a Lipschitz seminorm in the sense that its null space is spanned by δ_e.*

We would like to know when L^ω is a Lip-norm. Of course, L^ω defines, as earlier, a metric on the state space $S(C_r^*(G, c))$, which may take value $+\infty$. We denote this metric by ρ_ω. As a first step, we would like to know whether ρ_ω gives $S(C_r^*(G, c))$ finite radius. We recall from proposition 2.2 of [40] that this will be the case if there is an $r \in \mathbb{R}$ such that $\|f\|^\sim \leq rL(f)$ for all $f \in C_c(G)$, where $\|f\|^\sim = \inf\{\|f - \alpha\delta_e\| : \alpha \in \mathbb{C}\}$. Officially speaking we should work with self-adjoint f's, but by the comments before definition 2.1 of [41] we do not need to make this restriction because clearly $L^\omega(f^*) = L^\omega(f)$ for each f. However we find it convenient to use the following alternative criterion for finite radius, which is natural in our situation because we have a canonical tracial state:

Proposition 2.4. *Let L be a Lipschitz seminorm on an order-unit space A, and let μ be a state of A. If the metric ρ_L from L gives $S(A)$ finite radius r, then $\|a\| \leq 2rL(a)$ for all $a \in A$ such that $\mu(a) = 0$. Conversely, if there is a constant k such that*

$$\|a\| \leq kL(a)$$

for all $a \in A$ such that $\mu(a) = 0$, then ρ_L gives $S(A)$ radius no greater than k.

Proof. Suppose the latter condition holds. For any given $a \in A$ set $b = a - \mu(a)e$. (Here e is the order-unit.) Then $\mu(b) = 0$, and so $\|a - \mu(a)e\| \leq kL(a)$. It follows that $\|a\|^\sim \leq kL(a)$, so that the ρ_L-radius of $S(A)$ is no greater than k.

Suppose conversely that $\|a\|^{\sim} \leq rL(a)$ for all a. Let $a \in A$ with $\mu(a) = 0$. There is a $t \in \mathbb{R}$ such that $\|a - te\| \leq rL(a)$. Then

$$|t| = |\mu(a) - t| = |\mu(a - te)| \leq \|a - te\| \leq rL(a).$$

Thus

$$\|a\| \leq \|a - te\| + \|te\| \leq 2rL(a).$$

So for $k = 2r$ we have $\|a\| \leq kL(a)$ if $\mu(a) = 0$. $\square$

We see that the constant k is not precisely related to the radius. But for our twisted group algebras there is a very natural state to use, namely the tracial state τ defined by $\tau(f) = f(e)$, which is the vector state for $\delta_e \in \boldsymbol{\ell}^2(G)$.

Suppose now that ρ_ω gives $S(C_r^*(G, c))$ finite radius, so that as above, if $\tau(f) = 0$ then $\|\pi(f)\| \leq 2rL(f)$. Let $x \in G$ with $x \neq e$. Then $\tau(\delta_x) = 0$, and so

$$1 = \|\pi(\delta_x)\| \leq 2rL^\omega(\delta_x) = 2r\ell^\omega(x).$$

We thus obtain:

Proposition 2.5. *If ρ_ω gives $S(C_r^*(G, c))$ finite radius r, then $\ell^\omega(x) \geq (2r)^{-1}$ for all $x \neq e$.*

Thus, for example, if θ is an irrational number, then neither the (unbounded) length function ℓ defined on $\mathbb{Z}^2$ by $\ell(m, n) = |m + n\theta|$, nor the homomorphism $\omega(m, n) = m + n\theta$, will give metrics for which $S(C^*(\mathbb{Z}^2))$ has finite radius.

But the condition of Proposition 2.5 is not at all sufficient for finite radius. For example, for any G we can define a length function ℓ by $\ell(x) = 1$ if $x \neq e$. Then it is easily checked that if $f = f^*$ then

$$L^\ell(f) = \|f - \tau(f)\delta_e\|_2.$$

If L^ℓ gives $S(C^*(G))$ finite radius, so that there is a constant k such that $\|\pi_f\| \leq kL^\ell(f)$ if $f(e) = 0$, then it follows that $\|\pi_f\| \leq 2k\|f\|_2$ when $f(e) = 0$. Since for any f we have $|f(e)| \leq \|f\|_2$, it follows that $\|\pi_f\| \leq (2k+1)\|f\|_2$, so that for any $g \in C_c(G)$ we have

$$\|f * g\|_2 \leq (2k+1)\|f\|_2\|g\|_2.$$

This quickly says that the norm on $\boldsymbol{\ell}^2(G)$ can be normalized so that $\boldsymbol{\ell}^2(G)$ forms an H^*-algebra, as defined in section 27 of [34]. But our algebra is unital, and the theory of H^*-algebra in [34] shows that G must then have finite-dimensional square-integrable unitary representations. But Weil pointed out on page 70 of [46] that this means that G is compact (so finite), because if $x \to U_x$ is the unitary matrix representation for a finite-dimensional square integrable representation, then the matrix coefficients of

$$x \mapsto I = U_x U_x^*$$

are integrable.

But beyond these elementary comments it is not clear to me what happens even for word-length functions. Thus we have the basic:

Question 2.6. *For which finitely generated groups G with cocycle c does the word-length function ℓ corresponding to a finite generating subset give a metric ρ_ℓ which gives $S(C_r^*(G,c))$ finite diameter? That is, when is there a constant, k, such that if $f \in C_c(G)$ and $f(e) = 0$ then*

$$\|\pi(f)\| \le k\|[M_\ell, \pi(f)]\|?$$

(Is the answer independent of the choice of the generating set?)

I do not know the answer to this question when the cocycle c is trivial and, for example, G is the discrete Heisenberg group, or the free group on two generators. In later sections we will obtain some positive answers for $G = \mathbb{Z}^d$, but even that case does not seem easy.

Even less do I know answers to the basic:

Question 2.7. *For which finitely generated groups G with 2-cocycle c does the word-length function ℓ corresponding to a finite generating subset give a metric ρ_ℓ which gives $S(C_r^*(G,c))$ the weak-$*$ topology. That is* [39]*, given that ρ_ℓ does give $S(C_r^*(G,c))$ finite diameter, when is*

$$\mathcal{B}_1 = \{f \in C_c(G) : \|\pi_f\| \le 1 \text{ and } L_\ell(f) \le 1\}$$

a totally-bounded subset of $C_r^*(G)$?

But we now make some elementary observations about this second question.

Proposition 2.8. *Let L be a Lip-norm on an order-unit space A. If L is continuous for the norm on A, then A is finite-dimensional.*

Proof. Much as just above we set

$$\mathcal{B}_1 = \{a \in A : \|\alpha\| \le 1 \text{ and } L(a) \le 1\}.$$

Since L is a Lip-norm, $\mathcal{B}_1$ is totally bounded by theorem 1.9 of [39]. But if L is also norm-continuous, then there is a constant $k \ge 1$ such that $L(a) \le k\|a\|$ for all $a \in A$. Consequently $\{a : \|a\| \le k^{-1}\} \subseteq \mathcal{B}_1$. It follows that the unit ball for the norm is totally bounded, and so the unit ball in the completion of A is compact. But it is well-known that the unit ball in a Banach space is not norm-compact unless the Banach space is finite-dimensional. □

Corollary 2.9. *Let A be an order-unit space which is represented faithfully as operators on a Hilbert space $\mathcal{H}$. Let D be a self-adjoint operator on $\mathcal{H}$, and set $L(a) = \|[D,a]\|$. Assume that L is (finite and) a Lip-norm on A. If D is a bounded operator, then A is finite-dimensional.*

From this we see that in our setting of $D = M_\omega$ for $C_r^*(G,c)$, if we want L^ω to be a Lip-norm, then we must use unbounded ω's unless G is finite. But it is not clear to me whether ω must always be a proper function, that is, whether $\{x : |\omega(x)| \le k\}$ must be finite for every k. However, the referee has pointed out to me that if ω is actually a length function, then ω must be proper if L^ω is to be a Lip-norm. For if it is not proper, then there is a constant, r, with

$0 < r \leq 1$, such that $S = \{x : \omega(x) \leq r^{-1}\}$ is infinite. But if ω is a length function then $L(\delta_x) = \omega(x)$. Thus $\{r\delta_x : x \in S\}$ is a norm-discrete subset of

$$\mathcal{B}_1 = \{f \in C_c(G) : \|f\|_A \leq 1 \text{ and } L(f) \leq 1\},$$

so that $\mathcal{B}_1$ can not be totally bounded. (See the first three sentences of the proof of Lemma 1.4.)

Finally, we will examine briefly three of Connes' axioms for a non-commutative Riemannian geometry [15]. We begin first with the axiom of "reality" (axiom $7'$ on page 163 of [15] and condition 4 on page 483 of [23]). For any C^*-algebra A with trace τ there is a natural and well-known "charge-conjugation" operator, J, on the GNS Hilbert space for τ, determined by $Ja = a^*$. We are in that setting, and so our J is given by

$$(J\xi)(x) = \bar{\xi}(x^{-1})$$

for $\xi \in \ell^2(G)$. For any $f \in C_c(G)$ one checks that $J\pi_f J$ is the operator of *right*-convolution by f^*, where $f^*(x) = \bar{f}(x^{-1})$. In particular, $J\pi_f J$ will commute with any π_g for $g \in C_c(G)$. This means exactly that the axiom of reality is true if one considers our geometry to have dimension 0.

With the axiom of reality in place, Connes requires that D be a "first-order operator" (axiom $2'$ of [15], or condition 5 on page 484 of [23], where the terminology "first order" is used). This axiom requires that $[D, a]$ commutes with JbJ for all $a, b \in A$. For our situation, let ρ_z denote right c-twisted translation on $\ell^2(G)$ by $z \in G$, so that $J\pi_z^* J = \rho_z$. Then in terms of the notation we have established, the first-order condition requires that ρ_z commutes with M_{φ_y} for each z and y. This implies that for each $x \in G$ we have

$$\omega(x) - \omega(y^{-1}x) = \omega(xz) - \omega(y^{-1}xz).$$

If we choose $z = x^{-1}$ and rearrange, we obtain

$$\omega(x) + \omega(y^{-1}) = \omega(y^{-1}x) + \omega(e).$$

This says that if we subtract the constant function $\omega(e)$, then ω is a group homomorphism from G into $\mathbb{R}$. Thus the first-order condition is rarely satisfied in our context. In fact, if we want ω to give $S(C_r^*(G))$ finite radius then it follows from Proposition 2.5 that $G \cong \mathbb{Z}$ or is finite.

Lastly, we consider the axiom of smoothness (axiom 3 on page 159 of [15], or condition 2 on page 482 of [23], where it is called "regularity" rather than "smoothness"). This requires that a and $[D, a]$ are in the domains of all powers of the derivation $T \mapsto [|D|, T]$. In our context $|D| = M_{|\omega|}$. But

$$||\omega(x)| - |\omega(z^{-1}x)|| \leq |\omega(x) - \omega(z^{-1}x)|,$$

so that $|\omega|$ is translation-bounded when ω is. From this it is easily seen that the axiom of smoothness is always satisfied in our setting.

3. The cosphere algebra

We now begin to establish some constructions which will permit us to obtain positive answers to Questions 2.6 and 2.7 for the groups $\mathbb{Z}^d$, and which may eventually be helpful in dealing with other groups.

Connes has shown (section 6 of [13], [22]) how to construct for each spectral triple $(A, \mathcal{H}, D)$ a certain C^*-algebra, denoted S^*A. He shows that if $A = C^\infty(\mathcal{M})$ where $\mathcal{M}$ is a compact Riemannian spin manifold, and if $(\mathcal{H}, D)$ is the corresponding Dirac operator, then S^*A is canonically isomorphic to the algebra of continuous functions on the unit cosphere bundle of $\mathcal{M}$. Thus in the general case it seems reasonable to call S^*A the cosphere algebra of $(A, \mathcal{H}, D)$. (In [22] S^*A is called the "unitary cotangent bundle".) In this section we will explore what this cosphere algebra is for our (almost) spectral triples of form $(C_c(G), \ell^2(G), M_\omega)$. (I thank Pierre Julg for helpful comments about this at an early stage of this project.)

We now review the general construction. But for our purposes we do not need the usual further hypothesis of finite summability for D. Thus we just require that we have $(A, \mathcal{H}, D)$ such that $[D, a]$ is bounded for all $a \in A$. But, following Connes, we also make the smoothness requirement that $[|D|, a]$ be bounded for all $a \in A$. We saw in the previous section that this latter condition is always satisfied in our setting where $D = M_\omega$.

Connes' construction of the algebra S^*A is as follows. (See also the introduction of [22].) Form the strongly continuous one-parameter unitary group $U_t = \exp(it|D|)$. Let $\mathcal{C}_D$ be the C^*-algebra of operators on $\mathcal{H}$ generated by the algebra $\mathcal{K}$ of compact operators on $\mathcal{H}$ together with all of the algebras $U_t A U_{-t}$ for $t \in \mathbb{R}$. (Note that usually $U_t A U_{-t} \not\subseteq A$.) Clearly the action of conjugating by U_t carries $\mathcal{C}_D$ into itself. We denote this action of $\mathbb{R}$ on $\mathcal{C}_D$ by η. Because of the requirement that $[|D|, a]$ be bounded, the action η is strongly continuous on $\mathcal{C}_D$. (See the first line of the proof of corollary 10.16 of [23].) Since $\mathcal{K}$ is an ideal (η-invariant) in $\mathcal{C}_D$, we can form $\mathcal{C}_D/\mathcal{K}$. Then by definition $S^*A = \mathcal{C}_D/\mathcal{K}$. The action η drops to an action of $\mathbb{R}$ on S^*A, which Connes calls the "geodesic flow".

We now work out what the above says for our case in which we have $(C_r^*(G, c), \ell^2(G), M_\omega)$. We will write $\mathcal{C}_\omega$ instead of $\mathcal{C}_D$. Since only $|\omega|$ is pertinent, we assume for a while that $\omega \geq 0$. Set $u_t(x) = \exp(it\omega(x))$ for $t \in \mathbb{R}$, so that the U_t of the above construction becomes M_{u_t}. Then for each $y \in G$ our algebra $\mathcal{C}_\omega$, defined as above, must contain

$$U_t \pi_y U_t^* = M_{u_t} M_{\alpha_y(u_t^*)} \pi_y = M_{u_t \alpha_y(u_t^*)} \pi_y.$$

But $\mathcal{C}_\omega$ must also contain $(\pi_y)^{-1}$, and thus it contains each $u_t \alpha_y(u_t^*)$, where for notational simplicity we omit M. But

$$(u_t \alpha_y(u_t^*))(x) = \exp(it(\omega(x) - \omega(y^{-1}x))) = \exp(it\varphi_y(x)).$$

Since φ_y is bounded, the derivative of $U_t \alpha_y(U_t^*)$ at $t = 0$ will be the norm-limit of the difference quotients. Thus we see that also $\varphi_y \in \mathcal{C}_\omega$ for each $y \in G$. But $\mathcal{C}_\omega \supseteq \mathcal{K}$, and so $\mathcal{C}_\omega \supseteq C_\infty(G)$, the space of continuous functions vanishing at infinity, where the elements of $C_\infty(G)$ are here viewed as multiplication operators. Note also that $\mathcal{C}_\omega$ contains the identity element.

All of this suggests that we consider, inside the algebra $C_b(G)$ of bounded functions on G, the unital norm-closed subalgebra generated by $C_\infty(G)$ together with $\{\varphi_y : y \in G\}$. We denote this subalgebra by E_ω. Let $\bar{G}^\omega$ denote the maximal ideal space of E_ω, with its compact topology, so that $E_\omega = C(\bar{G}^\omega)$. Note that G sits in $\bar{G}^\omega$ as a dense open subset because $E_\omega \supseteq C_\infty(G)$. That is, $\bar{G}^\omega$ is a compactification of the discrete set G. We will call it the *ω-compactification* of G. Note that $C(\bar{G}^\omega)$ is separable because G is countable and so there is only a countable number of φ_y's. Thus the compact topology of $\bar{G}^\omega$ has a countable base.

The action α of G on $C_b(G)$ by left translation clearly carries E_ω into itself. From this we obtain an induced action on $\bar{G}^\omega$ by homeomorphisms. We denote this action again by α.

Of course $C(\bar{G}^\omega)$ is faithfully represented as an algebra of pointwise multiplication operators on $\ell^2(G)$. This representation, M, together with the representation π of G on $\ell^2(G)$ form a covariant representation [35], [47] of $(C(\bar{G}^\omega), G, \alpha, c)$. We have already seen earlier several instances of the covariance relation $\pi_x M_f = M_{\alpha_x(f)}\pi_x$. The integrated form of this covariant representation, which we denote again by π, gives then a representation on $\ell^2(G)$ of the full twisted crossed product algebra $C^*(G, C(\bar{G}^\omega), \alpha, c)$. It is clear from the above discussion that our algebra $\mathcal{C}_\omega$ contains $\pi(C^*(G, C(\bar{G}^\omega), \alpha, c))$. But for any $y \in G$ and $t \in \mathbb{R}$ we have $\exp(it\varphi_y) \in C(\bar{G}^\omega)$. From our earlier calculation this means that $\pi(C^*(G, C(\bar{G}^\omega), \alpha, c))$ contains $U_t \pi_y U_t^*$. Thus it also contains $U_t\pi(C_c(G))U_t^*$. Consequently:

Lemma 3.1. *We have* $\mathcal{C}_\omega = \pi(C^*(G, C(\bar{G}^\omega), \alpha, c))$.

Now $C(\bar{G}^\omega)$ contains $C_\infty(G)$ as an α-invariant ideal. The following fact must be known, but I have not found a reference for it.

Lemma 3.2. *With notation as above,*

$$C^*(G, C_\infty(G), \alpha, c) \cong \mathcal{K}(\ell^2(G)),$$

the algebra of compact operator on $\ell^2(G)$, with the isomorphism given by π.

Proof. If we view elements of $C_c(G, C_\infty(G))$ as functions on $G \times G$, and if for $f \in C_c(G, C_\infty(G))$ we set $(\Phi f)(x,y) = f(x,y)c(x, x^{-1}y)$, then

$$\Phi(f *_c g) = (\Phi f) * (\Phi g),$$

where only here we let $*_c$ denote convolution (in the crossed product) twisted by c, while $*$ denotes ordinary convolution. The verification requires using the 2-cocycle identity to see that

$$c(y, y^{-1}z)c(y^{-1}x, x^{-1}z) = c(y, y^{-1}x)c(x, x^{-1}z).$$

The untwisted crossed product $C_\infty(G) \times_\alpha G$ is well-known to be carried onto $\mathcal{K}(\ell^2(G))$ by π. (See [37].) (For non-discrete groups one must be more careful, because cocycles are often only measurable, not continuous.) $\square$

Because $\mathcal{K}(\boldsymbol{\ell}^2(G))$ is simple, it follows that the reduced C^*-algebra $C_r^*(G, C_\infty(G), \alpha, c)$ coincides with the full twisted crossed product, even when G is not amenable. Anyway, the consequence of this discussion is:

Proposition 3.3. *With notation as above, the cosphere algebra is*

$$S^*A = \pi(C^*(G, C(\bar{G}^\omega), \alpha, c))/\mathcal{K}(\boldsymbol{\ell}^2(G)).$$

For an element of $\pi(C^*(G, C(\bar{G}^\omega), \alpha, c)$ it is probably appropriate to call its image in S^*A its "symbol", in analogy with the situation for pseudodifferential operators.

We can use recently-developed technology to obtain a simpler picture in those cases in which the action α of G on $\bar{G}^\omega$ is amenable [1], [3], [2], [26], [25]. This action will always be amenable if G itself is amenable, which will be the case when we consider $\mathbb{Z}^d$ in detail later. So the following comments will only be needed there for that case. But we will see in Section 10 that the action can be amenable also in some situations for which G is not amenable, namely for the free group on two generators and its standard word-length function.

Let $\partial_\omega G = \bar{G}^\omega \setminus G$. It is reasonable to call $\partial_\omega G$ the "ω-boundary" of G. Notice that α carries $\partial_\omega G$ into itself. Suppose that the action α of G on $\partial_\omega G$ is amenable [2], [3]. One of the equivalent conditions for amenability of α (for discrete G) is that the quotient map from $C^*(G, C(\partial_\omega G))$ onto $C_r^*(G, C(\partial_\omega G))$ is an isomorphism (theorem 4.8 of [1] or theorem 3.4 of [2]). (No cocycle c is involved here.) In proposition 2.4 of [31] it is shown that for situations like this the amenability of the action on $\partial_\omega G$ is equivalent to amenability of the action on $\bar{G}^\omega$. (I thank Claire Anantharaman–Delaroche for bringing this reference to my attention, and I thank both her and Jean Renault for helpful comments on related matters.) The proof in [31] uses the characterization of amenability of the action in terms of nuclearity of the crossed product. Here is another argument which does not use nuclearity. Following remark 4.10 of [36], we consider the exact sequence of full crossed products

$$0 \to C^*(G, C_\infty(G), \alpha) \to C^*(G, C(\bar{G}^\omega), \alpha) \to C^*(G, C(\partial_\omega G), \alpha) \to 0$$

and its surjective maps onto the corresponding sequence of reduced crossed products (which initially is not known to be exact). A simple diagram-chase shows that if the quotient map onto $C_r^*(G, C(\partial_\omega G), \alpha)$ is in fact an isomorphism, then the sequence of reduced crossed products is in fact exact. Also, as discussed above, $C^*(G, C_\infty(G), \alpha)$ is the algebra of compact operators, so simple, and so the quotient map from it must be an isomorphism. A second simple diagram-chase then shows that the quotient map from $C^*(G, C(\bar{G}^\omega), \alpha)$ must be an isomorphism, so that the action α of G on $\bar{G}^\omega$ is amenable. (The verification that if the action on $\bar{G}^\omega$ is amenable then so is that on $\partial_\omega G$ follows swiftly from the equivalent definition of amenability in terms of maps whose values are probability measures on G. This definition is given further below and in example 2.2.14(2) of [3].)

For our general functions ω it is probably not reasonable to hope to find a nice criterion for amenability of the action. But in the case in which ω is a length-function ℓ (in which case we write $\partial_\ell G$ instead of $\partial_\omega G$), we will obtain in the next sections considerable information about $\partial_\ell G$, and so it is reasonable to pose:

Question 3.4. *Let G be a finitely generated group, and let ℓ be the word-length function for some finite set of generators. Under what conditions will the action of G on $\partial_\ell G$ be amenable? For which class of groups will there exist a finite set of generators for which the action is amenable? For which class of groups will this amenability be independent of the choice of generators?*

It is known that if G is a word-hyperbolic group, then its action on its Gromov boundary is amenable. See the appendix of [3], written by E. Germain, and the references given there. We would have a positive answer to Question 3.4 for word-hyperbolic groups if we had a positive answer to:

Question 3.5. *Is it the case that for any word-hyperbolic group G and any word-length function on G for a finite generating set, there is an equivariant continuous surjection from $\partial_\ell G$ onto the Gromov boundary of G?*

This seems plausible in view of our discussion of geodesic rays in the next section, since the Gromov boundary considers geodesic rays which stay a finite distance from each other to be equivalent.

We now explore briefly the consequences of the action being amenable. The first consequence is that the full and reduced twisted crossed products coincide. We have discussed the case of a trivial cocycle c above. I have not seen the twisted case stated in the literature, but it follows easily from what is now known. We outline the proof. To every 2-cocycle there is associated an extension, E, of G by $\mathbb{T}$. As a topological space $E = \mathbb{T} \times G$, and the product is given by $(s,x)(t,y) = (stc(x,y), xy)$. (See III.5.12 of [20].) We can compose the evident map from E onto G with α to obtain an action, α, of E on $\bar{G}^\omega$. Let W be any compact space on which G acts, with the action denoted by α. If α is amenable, then by definition (example 2.2.14(2) of [3], [2], [26], [25]) there is a sequence $\{m_j\}$ of weak-$*$ continuous maps from W into the space of probability measures on G such that, for α denoting also the corresponding action on probability measures, we have for every $x \in G$

$$\lim_j \sup_{w \in W} \|\alpha_x(m_j(w)) - m_j(\alpha_x(w))\|_1 = 0.$$

Let h denote normalized Haar measure on $\mathbb{T}$, and for each j and each $w \in W$ let $n_j(w)$ be the product measure $h \otimes m_j(w)$ on E. Thus each $n_j(w)$ is a probability measure on E. It is easily verified that the function $w \mapsto n_j(w)$ is weak-$*$ continuous. Furthermore, a straight-forward calculation shows that

$$\alpha_{(s,x)}(n_j) = h \otimes \alpha_x(m_j)$$

for each $(s,x) \in E$ and each j. Now E is not discrete. But from this calculation it is easily seen that the action of E on W is amenable, where now we use definition 2.1 of [2]. Then from theorem 3.4 of [2] (which is a special case of proposition 6.1.8 of [3]), it follows that $C^*(E, C(W), \alpha)$ coincides with $C^*_r(E, C(W), \alpha)$.

Now let p be the function on $\mathbb{T}$ defined by $p(t) = \exp(2\pi i t)$, where here we identify $\mathbb{T}$ with $\mathbb{R}/\mathbb{Z}$. Since $\mathbb{T}$ is an open subgroup of E, we can view p as a function on E by giving it value 0 off of $\mathbb{T}$. Since $\mathbb{T}$ is central in E, and α is trivial on $\mathbb{T}$, and $C(W)$ is unital, it follows that p is a central projection in $C^*(E, C(W), \alpha)$. From this it follows that the cut-down algebras $pC^*(E, C(W), \alpha)$ and $pC^*_r(E, C(W), \alpha)$ coincide. But it is easily seen (see page 84 of [18] or page 144 of [19]) that $pC^*(E, C(W), \alpha) = C^*(G, C(W), \alpha, c)$, and similarly for C^*_r. In this way we obtain:

Proposition 3.6. *Let G be a discrete group, let α be an action of G on a compact space W, and let c be a 2-cocycle on G. If the action α is amenable, then $C^*(G, C(W), \alpha, c)$ coincides with $C^*_r(G, C(W), \alpha, c)$.*

With some additional care the above proposition can be extended to the case in which W is only locally compact. In that case the projection p is only in the multiplier algebras of the twisted crossed products.

We now return to the case in which G acts on $\bar{G}^\omega$ and $\partial_\omega G$. From the above proposition it follows that if G acts amenably on $\partial_\omega G$, and so on $\bar{G}^\omega$, then we can view π as a representation of the reduced crossed product $C^*_r(G, C(\bar{G}^\omega), \alpha, c)$. This has the benefit that we can apply corollary 4.19 of [47] to conclude that π is a faithful representation of $C^*_r(G, C(\bar{G}^\omega), \alpha, c)$. The hypotheses of this corollary 4.19 are that M be a faithful representation of $C(\bar{G}^\omega)$, which is clearly true, and that M be G-almost free (definition 1.12 of [47]). This latter means that for any non-zero subrepresentation N of M and any $x \in G$ with $x \neq e$ there is a non-zero subrepresentation P of N whose composition with the inner automorphism from x is disjoint from P. But subrepresentations of M correspond to non-empty subsets of G, and for P we can take any one-point subset of a given subset. Thus our algebra $\mathcal{C}^\omega$ coincides (under π) with $C^*(G, C(\bar{G}^\omega), \alpha, c)$.

Now from Lemma 3.2 we know that $C^*(G, C_\infty(G), \alpha, c)$ coincides with $\mathcal{K}(\boldsymbol{\ell}^2(G))$, and the process of forming full twisted crossed products preserves short exact sequences. (See the top of page 149 of [47].) Thus from Proposition 3.3, and on removing our requirement that $\omega \geq 0$, we obtain:

Theorem 3.7. *Let ω be a translation bounded function on G such that the action of G on $\partial_{|\omega|}G$ is amenable. Then the cosphere algebra $S^*_\omega A$ for $(C^*_r(G, c), \boldsymbol{\ell}^2(G), M_\omega)$ is (naturally identified with)*

$$S^*_\omega A = C^*(G, C(\partial_\omega G), \alpha, c) = C^*_r(G, C(\partial_\omega G), \alpha, c).$$

4. The metric compactification

The purpose of this section is to show that when ω is a length-function on G then geodesic rays in G for the metric on G from ω give points in the compactification $\bar{G}^\omega$. This will be a crucial tool for us in dealing with $\mathbb{Z}^d$, since it will supply us with a sufficient collection of points in the boundary which have finite orbits. We will also see that $\bar{G}^\omega$ is then a special case of a compactification of complete locally compact metric spaces introduced by Gromov [24] some time ago. (This is probably related to the comment which Connes makes about nilpotent groups in the second paragraph after the end of the proof of proposition 2 of section 6 of [14].) Gromov's definition appears fairly different from that which we gave in the previous section, and so our treatment here can also be viewed as showing how to define Gromov's compactification as the maximal ideal space of a unital commutative C^*-algebra. We will refrain from using here the terms "Gromov compactification" and "Gromov boundary", since these terms seem already reserved in the literature for use with hyperbolic spaces, where they have a different meaning and give objects which depend only on the coarse quasi-isometry class of the metric. (See IIIH3 of [6].) We will instead use the terms "metric compactification" and "metric boundary", and our notation will often show the dependence on the metric. We will see in Example 5.2 that for a hyperbolic metric space the metric boundary and the Gromov boundary can fail to be homeomorphic.

Let (X, ρ) be a metric space, and let $C_b(X)$ denote the algebra of continuous bounded functions on X, equipped with the supremum norm $\|\cdot\|_\infty$. Motivated by the observations in the previous section, we define $\varphi_{y,z}$ on X for $y, z \in X$ by

$$\varphi_{y,z}(x) = \rho(x, y) - \rho(x, z).$$

Then the triangle inequality tells us that $\|\varphi_{y,z}\|_\infty \leq \rho(y, z)$, so that $\varphi_{y,z} \in C_b(X)$. But on setting $x = z$ we see that, in fact, $\|\varphi_{y,z}\|_\infty = \rho(y, z)$. Let H_ρ denote the linear span in $C_b(X)$ of $\{\varphi_{y,z} : y, z \in X\}$. Suppose that we fix some base point $z_0 \in X$. Then it is easily checked that $\varphi_{y,z} = \varphi_{z_0,z} - \varphi_{z_0,y}$. Thus H_ρ is equally well the linear span of $\{\varphi_{z_0,y} : y \in X\}$, but is independent of the choice of z_0. (It will be useful to us that we can change base-points at will.) We often find it convenient to fix z_0, and to set $\varphi_y = \varphi_{z_0,y}$, so that H_ρ is the linear span of the φ_y's. When X is a group, it is natural to choose $z_0 = e$. We were implicitly doing this in the previous section. We note that $\|\varphi_y\|_\infty = \rho(y, z_0)$.

Much as above, we have $\varphi_y - \varphi_z = \varphi_{z,y}$, and so $\|\varphi_y - \varphi_z\|_\infty = \|\varphi_{z,y}\|_\infty = \rho(y, z)$. Thus the mapping $y \mapsto \varphi_y$ is an isometry from (X, ρ) into $C_b(X)$. The latter space is complete, and so this isometry extends to the completion of X.

We desire to obtain a compactification of X to which all of the functions φ_y extend as continuous functions. We want X to be an open subset of the compactification, and so we must require that X is locally compact. Then the various compactifications of X in which X is open are just the maximal-ideal spaces of the various unital closed *-subalgebras of $C_b(X)$ which contain $C_\infty(X)$. Thus we set:

Definition 4.1. *Let (X, ρ) be a metric space whose topology is locally compact. Let $\mathcal{G}(X, \rho)$ be the norm-closed subalgebra of $C_b(X)$ which is generated by $C_\infty(X)$, the constant functions, and H_ρ. Let $\bar{X}^\rho$ denote the maximal ideal space of $\mathcal{G}(X, \rho)$. We call $\bar{X}^\rho$ the* metric compactification *of X for ρ.*

Then, essentially by construction, $\bar{X}^\rho$ is a compactification of X (within which X is open). We remark that if, instead, we take the norm-closed subalgebra of $C_b(X)$ generated by all of the bounded Lipschitz functions, then we obtain the algebra of all bounded uniformly continuous (for ρ) functions on X. (See the bottom of page 23 of [45].)

It is natural to think of $X_\rho \backslash X$ as a boundary at infinity for X. But from a metric standpoint this is not always reasonable. Suppose that X is not complete. Each of the functions φ_y is a Lipschitz function, and so extends to the completion $\hat{X}^\rho$ of X. Each $f \in C_\infty(X)$ extends continuously to $\hat{X}^\rho$ by setting it equal to 0 off X. The constant functions obviously extend to $\hat{X}^\rho$. Thus the algebraic algebra generated by H_ρ, $C_\infty(X)$ and the constant functions extends to an algebra of functions on $\hat{X}^\rho$, and the supremum norm is preserved under this extension. Thus our completed algebra $\mathcal{G}(X,\rho)$ can be viewed as a unital subalgebra of $C_b(\hat{X}^\rho)$. It is easily seen that this algebra separates the points of $\hat{X}^\rho$. (E.g., use the fact that ρ extends to the completion.) Thus we obtain a (continuous) injection of $\hat{X}^\rho$ into $\bar{X}^\rho$. But there is no reason that $\hat{X}^\rho$ should be open in $\bar{X}^\rho$, notably if the completion is not locally compact. Even if $\hat{X}^\rho$ is locally compact, the points of $\hat{X}^\rho \backslash X$ will all be of finite distance from the points of X, and so are not "at infinity". For this reason it seems best to define the "boundary" only for *complete* locally compact metric spaces. Thus we make:

Definition 4.2. *Let (X,ρ) be a metric space which is complete and locally compact. Then its* metric boundary *is $\bar{X}^\rho \backslash X$. We will denote the metric boundary by $\partial_\rho X$.*

We now show that the metric compactification and the metric boundary which we have defined above coincide with those constructed by Gromov [24] in a somewhat different way. Gromov proceeds as follows. (See also 3.1 of [5], II.1 of [4] and II.8.12 of [6].) Let (X,ρ) be a complete locally compact metric space, let $C(X)$ denote the vector space of all continuous (possibly unbounded) functions on X, and equip X with the topology of uniform convergence on compact subsets of X. Let $C_*(X)$ denote the quotient of $C(X)$ by the subspace of constant functions, with the quotient topology. For $f \in C(X)$ denote its image in $C_*(X)$ by $\bar{f}$. For $y \in X$ set $\psi_y(x) = \rho(x,y)$. Then $x \mapsto \psi_x$ is an embedding of X into $C(X)$. Let ι denote the corresponding embedding of X into $C_*(X)$, and let $\mathcal{C}\ell(X)$ be the closure of $\iota(X)$ in $C_*(X)$. Then $\mathcal{C}\ell(X)$ can be shown to be compact, and $\iota(X)$ can be shown to be open in $\mathcal{C}\ell(X)$, so that $\mathcal{C}\ell(X)\backslash X$ is a boundary at infinity for X.

We now explain the relationship between this construction of Gromov and our construction given earlier in this section. Fix a base point z_0. For any given $u \in \bar{X}^\rho$ define the function g_u by $g_u(x) = -\varphi_x(u)$, where φ_x is now viewed as a function on $\bar{X}^\rho$. If $u \in X$ then $g_u(x) = \rho(u,x) - \rho(u,z_0)$. Since $\rho(u,z_0)$ is constant in x, the image of g_u in $C_*(X)$ is exactly Gromov's $\iota(u)$. On the other hand, suppose that $u \in \partial_\rho X$. Because X is dense in $\bar{X}^\rho$, there is a net $\{y_\alpha\}$ of elements of X which converges to u. Then for each $x \in X$ we have

$$g_u(x) = -\varphi_x(u) = -\lim \varphi_x(y_\alpha) = \lim g_{y_\alpha}(x).$$

That is, g_{y_α} converges to g_u pointwise on X. But each g_y for $y \in X$ is clearly a Lipschitz function of Lipschitz constant 1, and pointwise convergence of a net of functions of bounded Lipschitz constant implies uniform convergence on compact sets. Thus g_{y_α} converges uniformly to g_u on compact subsets of X, so that $\bar{g}_u \in \mathcal{C}\ell(X)$. (In the literature cited above, g_u would be called a *horofunction* if $u \in \partial_\rho X$.) In this way we obtain a mapping, $u \mapsto \bar{g}_u$, from $\bar{X}^\rho$ to $\mathcal{C}\ell(X)$. If $\bar{g}_u = \bar{g}_v$ for some $u, v \in \bar{X}^\rho$, then there is a constant, k, such that $\varphi_x(u) = \varphi_x(v) + k$ for all $x \in X$. From this it is easily seen that $u = v$. Thus the mapping $u \mapsto \bar{g}_u$ is injective on $\bar{X}^\rho$. Finally, if $\{u_\alpha\}$ is a net in $\bar{X}^\rho$ which converges to $u \in \bar{X}^\rho$, then, much as above, g_{u_α} converges to g_u pointwise, and so uniformly on compact sets. Thus the mapping $u \mapsto \bar{g}_u$ is continuous from $\bar{X}^\rho$ into $\mathcal{C}\ell(X)$. Since $\bar{X}^\rho$ is compact, it follows that this mapping is a homeomorphism onto its image. But the image of X in $\mathcal{C}\ell(X)$ is dense, and so the mapping is a homeomorphism from $\bar{X}^\rho$ onto $\mathcal{C}\ell(X)$, and so from $\partial_\rho X$ to $\mathcal{C}\ell(X)\backslash X$, as desired.

For our later purposes it is important for us to examine the relationship between geodesics and points of $\partial_\rho X$. Much of the content of the next paragraphs appears in some form in various places in the literature [5], [4], [6], though usually not in the generality we consider here. And here we reformulate it in terms of our approach to the construction of $\partial_\rho X$.

We will not assume that our metric spaces are connected. For example, we will later consider $\mathbb{Z}^d$ with its Euclidean metric from $\mathbb{R}^d$. Every ray (half-line) in $\mathbb{R}^d$ should give a direction toward infinity for $\mathbb{Z}^d$. But if the direction involves irrational angles, the ray may not meet $\mathbb{Z}^d$ at an infinite number of points. So we need a slight generalization of geodesic rays. For perspective we also include a yet weaker definition.

Definition 4.3. *Let (X, ρ) be a metric space, let T be an unbounded subset of $\mathbb{R}^+$ which contains 0, and let γ be a function from T into X. We will say that:*

a) *γ is a* geodesic ray *if $\rho(\gamma(t), \gamma(s)) = |t - s|$ for all $t, s \in T$.*

b) *γ is an* almost-geodesic ray *if it satisfies the condition:*

For every $\varepsilon > 0$ there is an integer N such that if $t, s \in T$ and $t \geq s \geq N$, then

$$|\rho(\gamma(t), \gamma(s)) + \rho(\gamma(s), \gamma(0)) - t| < \varepsilon.$$

c) *γ is a* weakly-geodesic ray *if for every $y \in X$ and every $\varepsilon > 0$ there is an integer N such that if $s, t \geq N$ then*

$$|\rho(\gamma(t), \gamma(0)) - t| < \varepsilon$$

and

$$|\rho(\gamma(t), y) - \rho(\gamma(s), y) - (t - s)| < \varepsilon.$$

It is evident that any geodesic ray is an almost-geodesic ray. (I thank Simon Wadsley for pointing out to me that my definition of weakly-geodesic rays in the first version of this paper was defective.)

Lemma 4.4. *Let γ be an almost-geodesic ray, and let (ε, N) be as in Definition 4.3b. Then for $t \geq s \geq N$ we have:*

a) $|\rho(\gamma(t), \gamma(0)) - t| < \varepsilon$.
b) $|\rho(\gamma(t), \gamma(s)) - (t-s)| < 2\varepsilon$.
c) $\rho(\gamma(t), \gamma(s)) < \rho(\gamma(t), \gamma(0)) - \rho(\gamma(s), \gamma(0)) + 2\varepsilon$.

Proof. For a) set $s = t$ in the condition of Definition 4.3b. For b) we have

$$\begin{aligned}&|\rho(\gamma(t)), \gamma(s)) - (t-s)|\\ &\quad = |(\rho(\gamma(t), \gamma(s)) + \rho(\gamma(s), \gamma(0)) - t) - (\rho(\gamma(s), \gamma(0)) - s)| < 2\varepsilon.\end{aligned}$$

Finally, for c) we have

$$\begin{aligned}&\rho(\gamma(t), \gamma(s))\\ &\quad = (\rho(\gamma(t), \gamma(s)) + \rho(\gamma(s), \gamma(0)) - t) - \rho(\gamma(s), \gamma(0))\\ &\quad +\rho(\gamma(t), \gamma(0)) - (\rho(\gamma(t), \gamma(0)) - t)\\ &\quad < \rho(\gamma(t), \gamma(0)) - \rho(\gamma(s), \gamma(0)) + 2\varepsilon.\end{aligned}$$

□

Lemma 4.5. *Any almost-geodesic ray is weakly geodesic. Let γ be a weakly-geodesic ray. Take $\gamma(0)$ as the base-point for defining φ_y for any $y \in X$. Then $\lim_{t\to\infty} \varphi_y(\gamma(t))$ exists for every $y \in X$. If γ is actually a geodesic ray, then $t \mapsto \varphi_y(\gamma(t))$ is a non-decreasing (bounded) function.*

Proof. To motivate the rest of the proof, suppose first that γ is a geodesic ray. We show that $t \mapsto \varphi_y(\gamma(t))$ is a non-decreasing function (so has a limit). For $t \geq s$ we have

$$\begin{aligned}\varphi_y(\gamma(t)) \quad &- \quad \varphi_y(\gamma(s))\\ &= \quad \rho(\gamma(t), \gamma(0)) - \rho(\gamma(t), y) - \rho(\gamma(s), \gamma(0)) + \rho(\gamma(s), y)\\ &= \quad t - s + \rho(\gamma(s), y) - \rho(\gamma(t), y)\\ &= \quad \rho(\gamma(t), \gamma(s)) + \rho(\gamma(s), y) - \rho(\gamma(t), y) \geq 0\end{aligned}$$

by the triangle inequality.

Next, let γ be an almost-geodesic ray. It is useful and instructive to first see why $\lim_{t\to\infty} \varphi_y(\gamma(t))$ exists. Given $\varepsilon > 0$, take N as in Definition 4.3b. We will show first that if $t \geq s \geq N$ then $\varphi_y(\gamma(t)) > \varphi_y(\gamma(s)) - 3\varepsilon$. In fact,

$$\begin{aligned}\varphi_y(\gamma(t)) \quad &- \quad \varphi_y(\gamma(s))\\ &= \quad \rho(\gamma(t), \gamma(0)) - \rho(\gamma(t), y) - \rho(\gamma(s), \gamma(0)) + \rho(\gamma(s), y)\\ &\geq \quad -\rho(\gamma(t), \gamma(s)) + \rho(\gamma(t), \gamma(0)) - \rho(\gamma(s), \gamma(0)) > -3\varepsilon,\end{aligned}$$

by part c) of Lemma 4.4.

Now let $m = \overline{\lim}\varphi_y(\gamma(t))$. Since $\varphi_y(x) \leq \rho(y, \gamma(0))$ for all $x \in X$, we must have $m \leq \rho(y, \gamma(0))$. Now there is an $s_0 \geq N$ such that $\varphi_y(\gamma(s_0)) \geq m - \varepsilon$.

Set $M = s_0$. Then for $t \geq M$ we must have $m \geq \varphi_y(\gamma(t)) \geq m - 4\varepsilon$ according to the previous paragraph. It follows that $\lim \varphi_y(\gamma(t)) = m$.

We can now show that γ is weakly-geodesic. Given $\varepsilon > 0$, choose N and $M \geq N$ as above. Then for $t \geq s \geq M$ the first condition of Definition 4.3c is satisfied by Lemma 4.4a, while for the second condition we have from above

$$\begin{aligned} |\rho(\gamma(t), y) \;\; - & \;\; \rho(\gamma(s), y) - (t-s)| \\ \leq & \;\; |\rho(\gamma(t), y) - \rho(\gamma(t), \gamma(0)) - \rho(\gamma(s), y) + \rho(\gamma(s), \gamma(0))| \\ + & \;\; |\rho(\gamma(t), \gamma(0)) - t| + |\rho(\gamma(s), \gamma(0)) - s| \\ \leq & \;\; |\varphi_y(\gamma(t)) - \varphi_y(\gamma(s))| + 2\varepsilon < 6\varepsilon. \end{aligned}$$

Finally, suppose that γ is a weakly-geodesic ray. For any $y \in X$ we show that $\{\varphi_y(\gamma(t))\}$ is a Cauchy net. Let ε and N be as in Definition 4.3c. Then for $t, s \geq N$ we have

$$\begin{aligned} |\varphi_y(\gamma(t)) \;\; - & \;\; \varphi_y(\gamma(s))| \\ = & \;\; |\rho(\gamma(t), \gamma(0)) - \rho(\gamma(t), y) - \rho(\gamma(s), \gamma(0)) + \rho(\gamma(s), y)| \\ \leq & \;\; |\rho(\gamma(s), y) - \rho(\gamma(t), y) - (s-t)| \\ + & \;\; |\rho(\gamma(t), \gamma(0)) - t| + |s - \rho(\gamma(s), \gamma(0))| < 3\varepsilon. \end{aligned}$$

□

For the next theorem we will need:

Proposition 4.6. *Let (X, ρ) be a locally compact metric space. If the topology of X has a countable base, then so do the topologies of $\bar{X}^\rho$ and $\partial_\rho X$.*

Proof. If (X, ρ) is a locally compact metric space whose topology has a countable base, then $C_\infty(X)$ has a countable dense set. Also, X has a countable dense set, and the corresponding φ_y's can be used to construct a countable dense subset of H_ρ. Thus the C^*-algebra $\mathcal{G}(X, \rho)$ will have a countable dense set, and so the underlying spaces will have countable bases for their topologies. □

We recall that a metric is said to be *proper* if every closed ball of finite radius is compact.

Theorem 4.7. *Let (X, ρ) be a complete locally compact metric space, and let γ be a weakly-geodesic ray in X. Then $\lim_{t \to \infty} f(\gamma(t))$ exists for every $f \in \mathcal{G}(X, \rho)$, and defines an element of $\partial_\rho X$. Conversely, if ρ is proper and if the topology of (X, ρ) has a countable base, then every point of $\partial_\rho X$ is determined as above by a weakly-geodesic ray.*

Proof. It is clear that the limit exists for the constant functions. From the definition of a weakly geodesic ray we see that γ must leave any compact set. Thus the limit exists and is 0 for all $f \in C_\infty(X)$. Choose $\gamma(0)$ as the base-point in defining φ_y for any $y \in X$. Then from Lemma 4.5 we know that $\lim \varphi_y(\gamma(t))$ exists for all $y \in X$.

Let $\tilde{\mathcal{G}}(X,\rho)$ denote the subalgebra of $C_b(X)$ generated by $C_\infty(X)$, the constant functions, and the φ_y's, before taking the norm-closure. It is clear from the above that $\lim f(\gamma(t))$ exists for every $f \in \tilde{\mathcal{G}}(X,\rho)$, and that $|\lim f(\gamma(t))| \le \|f\|_\infty$. Thus the limit defines a homomorphism from $\tilde{\mathcal{G}}(X,\rho)$ to $\mathbb{C}$ which is norm-continuous, and so extends to all of $\mathcal{G}(X,\rho)$ by continuity. It thus defines a point, say u, of $\bar{X}^\rho$. But because γ leaves any compact subset of X, the point defined by the limit must be in $\partial_\rho X$. It is easy to check now that $\lim f(\gamma(t))$ exists and equals $f(u)$ for all $f \in \mathcal{G}(X,\rho)$.

Suppose now that the topology of (X,ρ) has a countable base, and that ρ is proper. Let $u \in \partial_\rho X$. Then we can apply Proposition 4.6 to conclude that there is a sequence, $\{w_n\}$, in X which converges in $\bar{X}^\rho$ to u. Since $u \notin X$ and ρ is proper, the sequence $\{w_n\}$ must be unbounded. Thus we can find a subsequence, which we denote again by $\{w_n\}$, such that if $n > m$ then $\rho(w_n, w_0) > \rho(w_m, w_0)$. Let T denote the set of $\rho(w_n, w_0)$'s, and for any $t \in T$ with $t = \rho(w_n, w_0)$ set $\gamma(t) = w_n$. Then $\lim \gamma(t) = u$. We show that γ is weakly-geodesic. Notice that by construction $\rho(\gamma(t), \gamma(0)) = t$ for each $t \in T$, so that the first condition of Definition 4.3c is satisfied. Let $y \in X$. Use $\gamma(0)$ as the base-point for defining φ_y. Now $\varphi_y(\gamma(t))$ converges to $\varphi_y(u)$, and so, given $\varepsilon > 0$, we can find an N such that whenever $s, t \in T$ with $s, t \ge N$ then $|\varphi_y(t) - \varphi_y(s)| \le \varepsilon$. Then for such s, t we have

$$|\rho(\gamma(t), y) - \rho(\gamma(s), y) - (t - s)| = |\varphi_y(\gamma(t)) - \varphi_y(\gamma(s))| \le \varepsilon.$$

□

In view of the history of these ideas (see 1.2 of [24]), we make:

Definition 4.8. *A point of $\partial_\rho X$ which is defined as above by an almost-geodesic ray γ will be called a* Busemann point *of $\partial_\rho X$, and we will denote the point by b_γ.*

For any (X,ρ) it is an interesting question as to whether every point of $\partial_\rho X$ is a Busemann point. This is known to be the case for CAT(0) spaces (corollary II.8.20 of [6]). But in the next section we will need to deal with metric spaces which are not CAT(0). We will also see there by example that two metrics ρ_1 and ρ_2 on X which are Lipschitz equivalent, in the sense that there are positive constants k, K such that

$$k\rho_1 \le \rho_2 \le K\rho_1,$$

can give metric boundaries for X which are not homeomorphic.

Here is an example of a complete locally compact non-compact metric space X which has no geodesic rays, but for which every point of $\partial_\rho X$ is a Busemann point. Let X be the subset $X = \{(n, 1/n) : n \ge 1\}$ of $\mathbb{R}^2$, with the restriction to it of the Euclidean metric on $\mathbb{R}^2$. This suggests the usefulness of almost-geodesic rays. Just before Proposition 5.4 we will give an example of a proper metric on $\mathbb{Z}$ for which there are no almost-geodesic rays, so no Busemann points (but there are sufficiently many weakly-geodesic rays).

We will later need:

Proposition 4.9. *Let $z_0 \in X$ and let γ and γ' be almost-geodesic rays from z_0 (i.e., $\gamma(0) = z_0 = \gamma'(0)$). If for any positive integer N and any $\varepsilon > 0$ we can find s and t in the domains of γ and γ' respectively such that $s, t \geq N$ and $\rho(\gamma(s), \gamma'(t)) < \varepsilon$, then $b_\gamma = b_{\gamma'}$.*

Proof. Each φ_y has Lipschitz constant ≤ 2, so

$$|\varphi_y(\gamma(s)) - \varphi_y(\gamma'(t))| \leq 2\rho(\gamma(s), \gamma'(t)).$$

The desired result follows quickly from this. □

We now briefly consider isometries. Suppose that α is an isometry of (X, ρ) onto itself. Then for $y, z \in X$ we have $\varphi_{y,z} \circ \alpha^{-1} = \varphi_{\alpha(y),\alpha(z)}$. Thus H_ρ is carried onto itself by α. Clearly so are $C_\infty(X)$ and the constant functions, and so α gives an automorphism of the algebra $\mathcal{G}(X, \rho)$. It follows that α gives a homeomorphism of $\bar{X}^\rho$ onto itself which extends α on X. This homeomorphism carries $\partial_\rho X$ onto itself. Thus:

Proposition 4.10. *Every isometry of a complete locally compact metric space (X, ρ) extends uniquely to a homeomorphism of $\bar{X}^\rho$ onto itself which carries $\partial_\rho X$ onto itself.*

Later we will need to consider (cartesian) products of metric spaces. There are many ways to define a metric on a product. One of these ways meshes especially simply with the construction of the metric compactification. If (X, ρ_X) and (Y, ρ_Y) are metric spaces, we define ρ on $X \times Y$ by

$$\rho((x_1, y_1), (x_2, y_2)) = \rho_X(x_1, x_2) + \rho_Y(y_1, y_2).$$

We will call ρ the "sum of metrics".

Proposition 4.11. *Let (X, ρ_X) and (Y, ρ_Y) be locally compact metric spaces, and let ρ be the sum of metrics on $X \times Y$. Then*

$$(X \times Y)^{-\rho} = (\bar{X}^{\rho_X}) \times (\bar{Y}^{\rho_Y}).$$

Proof. We need to show that the evident map from $X \times Y$ to $(\bar{X}^{\rho_X}) \times (\bar{Y}^{\rho_Y})$ extends to a homeomorphism from $(X \times Y)^{-\rho}$. For this it suffices to show that the restriction map from $C((\bar{X}^{\rho_X}) \times (\bar{Y}^{\rho_Y}))$ to $C_b(X \times Y)$ maps into $C((X \times Y)^{-\rho})$ and is onto. Let x_0, y_0 be base-points in X and Y respectively, and use (x_0, y_0) as a base-point for $X \times Y$. Then for $(u, v) \in X \times Y$ we have

$$\begin{aligned}
\varphi_{(u,v)}(x, y) &= \rho((x, y), (x_0, y_0)) - \rho((x, y), (u, v)) \\
&= \rho_X(x, x_0) - \rho_X(x, u) + \rho_Y(y, y_0) - \rho_Y(y, v) \\
&= \varphi_u(x) + \varphi_v(y).
\end{aligned}$$

In particular, $\varphi_{(u,y_0)} = \varphi_u \otimes 1_Y$ and $\varphi_{(x_0,v)} = 1_X \otimes \varphi_v$. Thus the restrictions of $\varphi_u \otimes 1_Y$ and $1_X \otimes \varphi_v$ are in $C((X \times Y)^{-\rho})$. The same is true for any $f \otimes 1_Y$ and $1_X \otimes g$ where $f \in C_c(X)$ and $g \in C_c(Y)$, or for constant functions. Thus the range of the restriction map is in $C((X \times Y)^{-\rho})$. But from the calculation above we also see that any $\varphi_{(u,v)}$ is in the range of the restriction map, and from this it is easily seen that the restriction map is onto $C((X \times Y)^{-\rho})$. □

5. The case of $G = \mathbb{Z}$

In this section we will see how the constructions of the previous sections can be used to deal with Questions 2.6 and 2.7 when $G = \mathbb{Z}$. This case already reveals some phenomena which we will have to deal with later for the case $G = \mathbb{Z}^d$.

Example 5.1. We examine first the case in which ℓ is the standard length function on $G = \mathbb{Z}$ defined by $\ell(n) = |n|$, so that $\rho(m,n) = |m-n|$. Note that ℓ is the word-length function for the generating set $S = \{\pm 1\}$. We determine $\partial_\ell G$. For any $k \in \mathbb{Z}$ we have

$$\varphi_k(n) = |n| - |n-k|.$$

In particular,

$$\varphi_k(n) = \begin{cases} k & \text{for } n \geq 0 \text{ and } n \geq k \\ -k & \text{for } n \leq 0 \text{ and } n \leq k. \end{cases}$$

From this it is clear that $\bar{\mathbb{Z}}^\ell$ is just $\mathbb{Z}$ with the points $\{\pm\infty\}$ adjoined in the traditional way. The action α of $\mathbb{Z}$ on $\bar{\mathbb{Z}}^\ell$ is by translation leaving the points at infinity fixed. Thus $\partial_\ell Z = \{\pm\infty\}$ with the trivial action α of $\mathbb{Z}$.

Now let $f \in C_c(\mathbb{Z})$ be given. Since $\mathbb{Z}$ is amenable, we know that $[M_\ell, \pi(f)]$ is in $C(\bar{\mathbb{Z}}^\ell) \times_\alpha \mathbb{Z}$, and that this crossed product is faithfully represented on $\boldsymbol{\ell}^2(\mathbb{Z})$, as discussed in Section 3. We can factor by $\mathcal{K} = C_\infty(\mathbb{Z}) \times_\alpha Z$, and so look at the image of $[M_\ell, \pi_f]$ in the cosphere algebra S^*A, which by the discussion of Section 3 is exactly $C(\partial_\ell \mathbb{Z}) \times_\alpha \mathbb{Z}$. This latter is isomorphic to two copies of $C^*(\mathbb{Z})$. The image of $\Sigma f(y) M_{\varphi_y} \pi_y$ in the copy at $+\infty$ will be $\{k \mapsto kf(k)\}$, while the image in the copy at $-\infty$ will be $\{k \to -kf(k)\}$. Let us take here the convention that the Fourier series for any $g \in C_c(\mathbb{Z})$ is given by $\hat{g}(t) = \Sigma g(k) e^{ikt}$, so that $\hat{g}'(t) = i\Sigma k g(k) e^{ikt}$. Then we see from just above that

$$L(f) = \|\Sigma f(y) M_{\varphi_y} \pi_y\| \geq \|\hat{f}'\|_\infty.$$

But $\|\hat{f}'\|_\infty$ agrees with the standard Lip-norm on $C^*(\mathbb{Z}) = C(\mathbb{T})$ which gives the circle a circumference of 2π. From the comparison lemma 1.10 of [39] it follows that L is a Lip-norm, and that it gives $\mathbb{T}$ (and so the state space $S(C^*(Z))$) radius no larger than π.

Example 5.2. Again we take $G = \mathbb{Z}$, but now we take the word-length function ℓ corresponding to the generating set $\{\pm 1, \pm 2\}$. Then ℓ is given by

$$\ell(n) = [|n|/2],$$

where $[\cdot]$ denotes "least integer not less than". Thus for any $k \in \mathbb{Z}$

$$\varphi_k(n) = [|n|/2] - [|n-k|/2].$$

From this one finds that if k is even then

$$\varphi_k(n) = \begin{cases} k/2 & \text{for } n \geq 0 \text{ and } n \geq k \\ -k/2 & \text{for } n \leq 0 \text{ and } n \leq k, \end{cases}$$

whereas if k is odd then

$$\varphi_k(n) = \begin{cases} \left.\begin{array}{l} (k-1)/2 \text{ for } n \text{ even} \\ (k+1)/2 \text{ for } n \text{ odd} \end{array}\right\} & \text{for } n \geq 0 \text{ and } n \geq k \\ \left.\begin{array}{l} -(k+1)/2 \text{ for } n \text{ even} \\ -(k-1)/2 \text{ for } n \text{ odd} \end{array}\right\} & \text{for } n \leq 0 \text{ and } n \leq k. \end{cases}$$

From this it is easily seen that $\partial_\ell \mathbb{Z}$ will consist of 4 points, two at $+\infty$ and two at $-\infty$, which we can label "even" and "odd". The action of $\mathbb{Z}$ on $\partial_\ell \mathbb{Z}$ will at each end be that of $\mathbb{Z}$ on $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$. In particular, the boundary contains no fixed-points for this action.

We learn several things from comparing this example with the one just before. First, two word-length metrics on a given group can give metric boundaries which are not homeomorphic. But it is well-known (e.g., proposition 8.3.18 of [8]) and easily seen that if G is a finitely-generated group and if ℓ_1 and ℓ_2 are the word-length functions for two finite generating sets, then the corresponding left-invariant metrics are (Lipschitz) equivalent in the sense defined in the previous section. Thus we see that equivalent metrics which give (the same) locally compact topologies (even discrete) and for which the set is complete, can give metric boundaries which are not homeomorphic.

Next, $\mathbb{Z}$ is an example of a hyperbolic group [21], and so for the metric from either of these generating sets it is a hyperbolic metric space. But the Gromov boundary of a hyperbolic space is independent of the metrics as long as the metrics are equivalent, or at least coarsely equivalent. The Gromov boundary for $\mathbb{Z}$ is just $\{\pm\infty\}$. One way of viewing what is happening is that for the metric of the Example 5.2 the maps $m \mapsto 2m$ and $m \mapsto 2m+1$ are geodesic rays which determine Busemann points in the boundary which are our two points at $+\infty$. But for the Gromov boundary any two geodesic rays which stay a bounded distance from each other define the same point at infinity. In particular, our present example shows that for a given hyperbolic metric space the metric boundary and the Gromov boundary can fail to be homeomorphic.

For our next observation, let (X, ρ) be a proper metric space with basepoint z_0, and let $T \subset \mathbb{R}^+$ be a fixed domain for geodesic rays, so that $0 \in T$ and T is unbounded. On the set of geodesic rays from z_0 whose domain is T we put the topology of pointwise convergence (which, because geodesic rays are Lipschitz maps of Lipschitz constant 1, is equivalent to the topology of uniform convergence on bounded subsets of T). This is done in various places in the literature. Because ρ is proper, it is easy to see that the set of all such geodesic rays is compact. For groups G with a word-length ℓ (or for graphs in general) it is natural to take $T = \mathbb{Z}^+$. It is reasonable to wonder then whether $\partial_\ell G$ is the quotient of this compact set of geodesics, with the quotient topology. If it were, then for each $y \in G$ the function which assigns to each such geodesic ray γ from e the number $\lim \varphi_y(\gamma(t))$ should be a continuous function on this compact set. But this already fails for Example 5.2. For each $k \geq 1$ let γ^k be the geodesic ray from 0 defined by

$$\gamma^k(n) = \begin{cases} 2n & \text{if } n \le k \\ 2n-1 & \text{if } n \ge k+1. \end{cases}$$

Then γ^k converges pointwise to the geodesic ray defined by $\gamma^\infty(n) = 2n$ for all n. But it is easy to see that b_{γ^∞} is the even point at $+\infty$ while b_{γ^k} is the odd point at $+\infty$ for all k. We also remark that in our present example there is no geodesic line which joins the two points at $+\infty$ (so this example fails to have the property of "visibility" [21]).

Our Example 5.2 also shows that the metric compactification is not in general well-related to the Higson compactification, as defined in 5.4 of [43]. For that definition let (X, ρ) be a proper metric space. For any $r > 0$ we define the variational function, $V_r f$, of any function f by

$$(V_r f)(x) = \sup\{|f(x) - f(y)| : \rho(y, x) \le r\}.$$

The Higson compactification is the maximal ideal space of the unital commutative C^*-algebra of all bounded continuous functions on X such that for each $r > 0$ the function $V_r f$ vanishes at infinity. For Example 5.2 let us consider $V_2\varphi_1$. Easy calculation shows that for any $n \ge 1$ we have $\varphi_1(2k) = 0$ while $\varphi_1(2k+1) = 1$. But $\rho(2k, 2k+1) = \ell(1) = 1$ for all k. Thus $(V_2\varphi_1)(k) \ge 1$ for all k. Consequently φ_1 does not extend to the Higson compactification. More generally, if a complete locally compact metric space (X, ρ) has geodesic rays which determine distinct Busemann points of $\partial_\rho X$ and yet stay a finite distance from each other, then $\bar{X}^\rho$ is not a quotient of the Higson compactification. Indeed, since the φ_y's separate the points of $\partial_\rho X$, there will be some y such that its φ_y separates the two Busemann points, and $V_r\varphi_y$ will not vanish at infinity if r is larger than the distance between the two rays.

The situation becomes yet more interesting when we consider generating sets such as $\{\pm 3, \pm 8\}$. But the proof given above that we obtain a Lip-norm when we use the generating set $\{\pm 1\}$ extends without too much difficulty to the case of arbitrary finite generating sets for $\mathbb{Z}$. We do not include this proof here since in Section 9 we will treat the general case of $\mathbb{Z}^d$ by similar techniques, though the details are certainly more complicated.

However we will discuss here another approach for the case of $G = \mathbb{Z}$ which uses a classical argument which was pointed out to me by Michael Christ. (I thank him for his guidence in this matter). This second approach seems less likely to generalize to more complicated groups, but it gives a stronger result for $\mathbb{Z}$. For any β with $0 < \beta \le 1$ and any metric ρ on a set, ρ^β will again be a metric, because $t \to |t|^\beta$ is a length function on $\mathbb{R}$. In particular, if we set $\ell_\beta(n) = |n|^\beta$ then ℓ_β is a length function on $\mathbb{Z}$.

Theorem 5.3. *Let ω be a translation-bounded function on $\mathbb{Z}$ such that $\omega(0) = 0$. If ℓ_β/ω is a bounded function (ignoring $n = 0$) for some β with $1/2 < \beta \le 1$, then L_ω is a Lip-norm on $C^*(\mathbb{Z}) = C(\mathbb{T})$.*

Proof. For any group G and any ω we have

$$[M_\omega, \pi_f]\delta_e = \Sigma f(y)\varphi_y(y)\delta_y = \Sigma\omega(y)f(y)\delta_y,$$

where $\{\delta_y\}$ here denotes the standard basis for $\boldsymbol{\ell}^2(G)$. Thus

$$\|\omega f\|_2 \le \|[M_\omega, \pi_f]\| = L_\omega(f).$$

What is special about $\mathbb{Z}$ is that $\|\omega f\|_2$ can control the norms we need. (This is related to our discussion of "rapid decay" in Section 1.) For this we need that $\ell_\beta^{-1} \in \boldsymbol{\ell}^2(\mathbb{Z})$, which happens exactly for $\beta > 1/2$. (Here and below we ignore $n = 0$ or set $\ell_\beta^{-1}(0) = 0$.) Let $f \in C_c(\mathbb{Z})$, with $\hat{f}$ its Fourier transform on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, viewed as a periodic function on $\mathbb{R}$. For $s, t \in \mathbb{R}$ with $s < t$ and $|t - s| < 1$ let $\chi_{[s,t]}$ denote the characteristic function of the interval $[s, t]$, extended by periodicity. Then

$$|\hat{f}(s) - \hat{f}(t)| = \left| \int_s^t \hat{f}'(r)dr \right| = |\langle \hat{f}', \chi_{[s,t]}\rangle| = |\langle (\hat{f}')^\vee, (\chi_{[s,t]})^\vee \rangle|.$$

But $(\chi_{[s,t]})^\vee(n) = (1/i2\pi n)(e(nt) - e(ns))$ if we set $e(r) = e^{2\pi i n r}$, while $(\hat{f}')^\vee(n) = -2\pi i n f(n)$. Thus if we set $g_{s,t}(n) = (e(nt) - e(ns))$, the above becomes $|\langle f, g_{s,t}\rangle|$ as a pairing between functions in $\boldsymbol{\ell}^1(\mathbb{Z})$ and $\boldsymbol{\ell}^\infty(\mathbb{Z})$. But (with $\omega^{-1}(0) = 0$) we can rewrite this as

$$|\langle \omega f, \omega^{-1} g_{s,t}\rangle| \le \|\omega f\|_2 \|\omega^{-1} g_{s,t}\|_2,$$

and notice that

$$\begin{aligned} \|\omega^{-1} g_{s,t}\|_2 &= \|(\ell_\beta/\omega)\ell_\beta^{-1} g_{s,t}\|_2 \\ &\le \|\ell_\beta/\omega\|_\infty \|\ell_\beta^{-1} g_{s,t}\|_2 < \infty, \end{aligned}$$

since $\ell_\beta^{-1} \in \boldsymbol{\ell}^2(\mathbb{Z})$. Set $m(s,t) = \|\ell_\beta^{-1} g_{s,t}\|_2$. Then putting the above together, we obtain

$$|\hat{f}(t) - \hat{f}(s)| \le m(s,t)\|\ell_\beta/\omega\|_\infty \|L_\omega(f)\|.$$

A simple estimate using the fact that $\ell_\beta^{-1} \in \boldsymbol{\ell}^2(\mathbb{Z})$ shows that for each $\varepsilon > 0$ there is a $\delta > 0$ such that if $|t - s| < \delta$ then $m(s,t) < \varepsilon$. From this we see that the set of $\hat{f}$'s for which $L_\omega(f) \le 1$ and $f(0) = 0$ forms a bounded subset of $C(\mathbb{T})$ which is equicontinuous, so totally bounded by the Arzela–Ascoli theorem. From this it is clear that L_ω gives finite radius and, by theorem 1.9 of [39], that it is a Lip-norm. $\square$

I suspect that when $\beta < 1/2$ then L_{ℓ_β} fails to be a Lip-norm, but I have not found a proof of this.

Notice that Theorem 5.3 applies if $|\omega(n)| \ge 1$ for $n \ne 0$ and if there are positive constants c and K such that $|\omega - c\ell^\beta| \le K$, for then $|\ell/\omega| \le (K+1)c$. This is the situation which occurs for the various word-length functions on $\mathbb{Z}$ (for $\beta = 1$).

It is interesting to see what the metric compactification of $\mathbb{Z}$ is when $\beta < 1$. For any $p \in \mathbb{Z}$ we have

$$\varphi_p(n) = |n|^\beta - |n-p|^\beta = \int_{|n-p|}^{|n|} \beta t^{\beta-1}\, dt.$$

Since $t^{\beta-1} \to 0$ at $+\infty$ because $\beta < 1$, it follows that $\varphi_p(n) \to 0$ as $n \to \pm\infty$. Thus $\varphi_p \in C_\infty(\mathbb{Z})$, and so the metric compactification is just the one-point compactification of $\mathbb{Z}$. Note also that $[M_{\ell_\beta}, \pi_f]$ is a compact operator for each $f \in C_c(\mathbb{Z})$. Thus the cosphere algebra for $(C^*(\mathbb{Z}), \boldsymbol{\ell}^2(\mathbb{Z}), M_{\ell_\beta})$ is $C^*(\mathbb{Z})$, and the image of $[M_{\ell_\beta}, \pi_f]$ in it is 0. We also remark that it is easily verified that if we set $\gamma(n^\beta) = n$, then γ is a weakly-geodesic ray, but that there are no almost-geodesic rays in $\mathbb{Z}$ for this metric, since by parts a) and b) of Lemma 4.4 if γ were such a ray we would have, for any fixed big r, that $|\gamma(t)|^\beta - |\gamma(t-r)|^b$ would be approximately r as $t \to \infty$, contradicting our observation above that it must go to 0.

We conclude this section with the following observation, which applies to our more general case of $\mathbb{Z}^d$.

Proposition 5.4. *Let ω be a translation-bounded function on a countable discrete Abelian group G, let L_ω on $C_c(G)$ be defined as earlier by $L_\omega(f) = \|[M_\omega, \pi_f]\|$, and let ρ_ω be the corresponding metric on $\hat{G}$ (which may not give the usual topology of $\hat{G}$). Then ρ_ω is invariant under translation on $\hat{G}$.*

Proof. Let us denote the pairing between G and $\hat{G}$ by $\langle m, t\rangle$. Then translation on $\hat{G}$ corresponds to the dual action, β, of $\hat{G}$ on $C^*(G)$ given on $C_c(G)$ by $(\beta_t(f))(m) = \langle m, t\rangle f(m)$. This is unitarily implemented in $\boldsymbol{\ell}^2(G)$ by M_t, where $(M_t\xi)(m) = \langle m, t\rangle \xi(m)$. Then

$$\begin{aligned} [M_\omega, \beta_t(\pi_f)] &= [M_\omega, M_t \pi_f M_t^*] \\ &= M_t[M_\omega, \pi_f]M_t^*, \end{aligned}$$

so that $L_\omega(\beta_t(f)) = L_\omega(f)$. (In other words, β is an action by isometries as defined in [41].) ☐

From Theorem 5.3 one begins to see that $\mathbb{T}^d$ has a bewildering variety of translation invariant metrics which give its topology. For example, if ρ is such a metric then so is ρ^r for any r with $0 < r < 1$, as is any convex function of ρ. The sum of two metrics and the supremum of two metrics are again metrics. More generally, the "$\boldsymbol{\ell}^p$-sum" of two metrics is a metric. These operations all preserve translation invariance. For the case of $\mathbb{T}$, any strictly increasing continuous function ℓ on $[0, 1/2]$ such that $\ell(0) = 0$ and $\ell(s+t) \le \ell(s) + \ell(t)$ if $s + t \le 1/2$ gives in an evident way a continuous length function on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and all continuous length functions on $\mathbb{T}$ arise in this way. It would be interesting to determine which generating sets for $\mathbb{Z}$ determine which length functions on $\mathbb{T}$, but I have not investigated this question.

6. The metric compactification for norms on $\mathbb{R}^d$

One of our eventual aims is to show that when ℓ is a length function on $\mathbb{Z}^d$ which is the restriction to $\mathbb{Z}^d$ of a norm on $\mathbb{R}^d$, then L_ℓ is a Lip-norm. In preparation for this we examine here the metric compactification of $\mathbb{R}^d$ for any given norm. We begin by considering the usual $\boldsymbol{\ell}^1$-norm, both because it is simple to treat and displays some interesting phenomena, and also because its restriction to $\mathbb{Z}^d$ gives the word-length function for the standard generating set. Following up on Example 5.1, we set $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ in the usual way, with the action of $\mathbb{R}$ fixing the points $\pm\infty$.

Proposition 6.1. *The metric compactification of* $(\mathbb{R}^d, \|\cdot\|_1)$ *is just* $(\bar{\mathbb{R}})^d$ *with its product action of* $\mathbb{R}^d$. *Thus the metric boundary is the set of* $(\tilde{x}_j) \in (\bar{\mathbb{R}})^d$ *such that at least one entry is* $+\infty$ *or* $-\infty$.

Proof. The metric from $\|\cdot\|_1$ on $\mathbb{R}^d$ is easily seen to be the sum of the metrics on $\mathbb{R}$ in the sense used in Proposition 4.11. Thus we just need to apply that proposition a number of times. □

We note that now there are orbits in the boundary which are not finite, but there are also fixed points (only a finite number of them).

We now investigate what happens for other norms on $\mathbb{R}^d$. It is notationally convenient for us just to consider a finite-dimensional vector space V with some given norm $\|\cdot\|$. We will denote the corresponding metric boundary simply by $\partial_\ell V$, where $\ell(x) = \|x\|$ for all $x \in V$.

For any $v \in V$ with $\|v\| = 1$ it is evident that the function $\gamma(t) = tv$ for $t \in T = [0, \infty)$ is a geodesic ray, and so from our earlier discussion it will determine a Busemann point, b_v, in $\partial_\ell V$. We now convert to this picture some of the known elementary facts about tangent functionals of convex sets, as explained for example in section V.9 of [17]. There is at least one linear functional, say σ, on V such that $\|\sigma\| = 1 = \sigma(v)$. We call such a σ a "support functional" at v. Then for any $y \in V$ we have

$$\varphi_y(\gamma(t)) = \|tv\| - \|tv - y\| \leq t - \sigma(tv - y) = \sigma(y).$$

In particular, $\varphi_{-y}(\gamma(t)) \leq -\sigma(y)$. On letting t go to $+\infty$ we find that

$$-\varphi_{-y}(b_v) \geq \sigma(y) \geq \varphi_y(b_v).$$

But theorem 5 of section V.9 of [17] (which uses the Hahn–Banach theorem) tells us that for any real number r such that $-\varphi_{-y}(b_v) \geq r \geq \varphi_y(b_v)$ there is a support functional σ at v such that $\sigma(y) = r$. To see that theorem V.9.5 really applies here, we note that if we set $s = t^{-1}$ then

$$\|tv\| - \|tv - y\| = (\|v\| - \|v - sy\|)/s,$$

and that $s \to +0$ as $t \to +\infty$. From this viewpoint we are thus looking at the negative of the tangent functional to the unit ball at v in the direction of $-y$, which fits the setting of theorem V.9.5.

The point v is called a *smooth* point of the unit sphere if there is only one support functional σ at v. We denote this unique σ by σ_v. Then the above considerations tell us that if v is smooth then $\varphi_y(b_v) = -\varphi_{-y}(b_v)$. On combining this with the inequalities found above, we obtain:

Proposition 6.2. *Let v be a smooth point of the unit sphere of V. Then*

$$\varphi_y(b_v) = \sigma_v(y)$$

for all $y \in V$.

For us the following proposition will be of considerable importance. We consider the action of V on itself by translation, and the corresponding action on $\partial_\ell V$.

Proposition 6.3. *Let v be a smooth point of the unit sphere of V. Then b_v is a fixed point under the action of V on $\partial_\ell V$.*

Proof. We use the 1-cocycle relation 2.2 and Proposition 6.2 to calculate that for any $x, y \in V$ we have

$$\begin{aligned} (\alpha_x\varphi_y)(b_v) &= \varphi_{x+y}(b_v) - \varphi_x(b_v) \\ &= \sigma_v(x+y) - \sigma_v(x) = \sigma_v(y) = \varphi_y(b_v). \end{aligned}$$

□

Finally, we note that theorem 8 of section V.9 of [17] says that, for any norm, the set of smooth points of the unit sphere is dense in the unit sphere. This does not imply that there are infinitely many fixed points in $\partial_\ell V$, as the next example shows. But we will see later that it does show that there are enough for our purposes.

Example 6.4. We examine the case of $\mathbb{R}^2$ with $\|\cdot\|_1$, whose metric compactification is described by Proposition 6.1. Let us see how our considerations concerning geodesics fit this example. We identify the dual space V' in the usual way with $\mathbb{R}^2$ with the norm

$$\|(r,s)\|_\infty = \max\{|r|, |s|\}.$$

All but 4 points of the unit sphere of V are smooth. However, for any $v = (a,b)$ with $0 < a$, $0 < b$ and $a + b = 1$ we see that $\sigma_v = (1,1) \in V'$. Thus all these different v's determine the same Busemann point of $\partial_\ell V$. This accords with Proposition 6.2 and the fact that $\partial_\ell V$ has only 4 fixed-points for the action of $\mathbb{R}^2$.

If instead we let v be the non-smooth point $(1,0)$ and let γ be the corresponding geodesic ray, then for any $y = (p,q) \in \mathbb{R}^2$ we have

$$\begin{aligned} \varphi_y(\gamma(t)) &= \|\gamma(t)\| - \|\gamma(t) - (p,q)\| \\ &= |t| - |t-p| - |q|. \end{aligned}$$

The limit as $t \to +\infty$ is clearly $p - |q|$, so that

$$\varphi_{(p,q)}(b_v) = p - |q| = \varphi_p(+\infty) + \varphi_q(0),$$

where φ_p and φ_q are for $\mathbb{R}$. Thus $b_v = (+\infty, 0)$ in the description of $\partial_\ell V$ given by Proposition 6.1. Clearly b_v is not given by an element of V'. It is easily seen that this b_v is not invariant under translation.

We see in this way that the *linear* geodesic rays from 0, corresponding to the points of the unit sphere, determine only 8 Busemann points of $\partial_\ell V$. But we can show that every point of $\partial_\ell V$ is determined by at least one (possibly non-linear) geodesic ray from 0. For example, if we consider $(+\infty, s) \in \partial_\ell V$ for some fixed $s \in \mathbb{R}$, we can pick any $t_0 \geq 0$ and let γ consist of the unit-speed straight-line path from $(0,0)$ to $(t_0, 0)$, followed by that from $(t_0, 0)$ to (t_0, s), followed by the linear ray from (t_0, s) in the direction $(1, 0)$. (We deal here with the "Manhattan metric".) It is easy to check that γ is a geodesic ray whose Busemann point corresponds to $(+\infty, s)$. We see in this way that every point of $\partial_\ell V$ is a Busemann point. It is also easy to see that for each of the 4 points $(\pm\infty, 0)$ and $(0, \pm\infty)$ of $\partial_\ell V$ there is only one geodesic ray to them from 0, but that for every other point of $\partial_\ell V$ there are uncountably many geodesic rays to it from 0.

Question 6.5. *Is it true that, for every finite-dimensional vector space and every norm on it, every point of $\partial_\ell V$ is a Busemann point?*

One says that $(V, \|\cdot\|)$ is *smooth* if every point of the unit sphere, S, of V is a smooth point. Let S' denote the unit sphere of V'. Then our earlier mapping $v \mapsto \sigma_v$ is defined on all of S. Furthermore it is onto S', because V, being finite dimensional, is reflexive. This mapping σ can also be seen to be continuous. This is essentially the fact that, as remarked at the bottom of page 60 of [33], a compactness argument shows that smoothness implies uniform smoothness. However, if S has "flat spots" then σ will not be injective. It is not difficult to show that for $(V, \|\cdot\|)$ smooth, $\partial_\ell V$ can be naturally identified with S', glued at ∞ using σ. In this case each point of $\partial_\ell V$ will be fixed by the action of V.

Question 6.6. *For a general $(V, \|\cdot\|)$ is there an attractive description of $\partial_\ell V$ and of the action of V on it?*

We have seen in Example 6.4 that the number of support functionals σ_v coming from smooth points v of the unit sphere can be finite. The reason that they nevertheless are adequate for our later purposes is given by the following proposition (which must be already known):

Proposition 6.7. *Let $\|\cdot\|$ be a norm on a finite-dimensional vector space V. Let $w \in V$, and suppose that $|\sigma_v(w)| \leq r$ for all smooth points v of the unit sphere. Then $\|w\| \leq r$. Furthermore, the closed convex hull of $\{\sigma_v : v$ is a smooth point$\}$ is the unit ball in the dual space V' for the dual norm $\|\cdot\|'$.*

Proof. Let $\|w\| = s$. Because the smooth points are dense in the unit sphere by theorem 8 of section V.9 of [17], for any $\varepsilon > 0$ we can find a smooth point v such that $\|w - sv\| < \varepsilon$. Then $|\sigma_v(w) - s| = |\sigma_v(w - sv)| < \varepsilon$. Since $|\sigma_v(w)| \leq r$ and ε is arbitrary, it follows that $\|w\| = s \leq r$.

Suppose now that $\tau \in V'$ and that $\tau \notin \bar{co}\{\sigma_v : v \text{ smooth}\}$. Then by the Hahn–Banach theorem there is a $w \in V$ and an $r \in \mathbb{R}$ such that $|\sigma_v(w)| \leq r < \tau(w)$ for all smooth v. But we have just seen that then $\|w\| \leq r$. Thus $\|\tau\|' > 1$. □

7. Restrictions of norms to $\mathbb{Z}^d$

In this section we will examine what happens when norms on $V = \mathbb{R}^d$ are restricted to $\mathbb{Z}^d$. We begin with the case of the norm $\|\cdot\|_1$. Following up on Example 6.4 we set $\bar{\mathbb{Z}} = \mathbb{Z} \cup \{\pm\infty\}$ in the usual way, with its action of $\mathbb{Z}$ leaving fixed the points at infinity. The proof of the following proposition is basically the same as that of Proposition 6.1.

Proposition 7.1. *For $\ell = \|\cdot\|_1$, the metric compactification of $(\mathbb{Z}^d, \ell)$ is $(\bar{\mathbb{Z}})^d$ with its product action of $\mathbb{Z}^d$. The metric boundary is the set of $(\tilde{n}_j) \in (\bar{\mathbb{Z}})^d$ such that at least one entry is $+\infty$ or $-\infty$.*

Suppose now that $\ell = \|\cdot\|$ is any norm on $V = \mathbb{R}^d$, and that we restrict it to $\mathbb{Z}^d$. For any $y \in \mathbb{Z}^d$ the function φ_y clearly extends to $\bar{V}^\ell$, and then restricts to the closure of $\mathbb{Z}^d$ in $\bar{V}^\ell$. It is not evident to me whether the φ_y's for $y \in \mathbb{Z}^d$ separate the points of this closure. But even if they did, it is not clear to me that we could then use this to apply the results of the previous section to show that there are sufficient fixed-points in $\partial_\ell \mathbb{Z}^d$ for the action of $\mathbb{Z}^d$. It is this supply of fixed-points which we need later. So we take a more direct tack. We show that every *linear* geodesic ray in V can be approximated by an almost-geodesic ray in $\mathbb{Z}^d$. The following lemma is closely related to Kronecker's theorem [7], so we just sketch the proof.

Lemma 7.2. *Let $v \in V$ with $\|v\| = 1$. Then there is an unbounded strictly increasing sequence $\{s_n\}$ of positive real numbers such that for every $\varepsilon > 0$ there is an N such that if $s_n > N$ then there is an $x \in \mathbb{Z}^d$ for which $\|x - s_n v\| < \varepsilon$.*

Proof. If there is an $r \in \mathbb{R}^+$ with $rv \in \mathbb{Z}^d$ then we simply take $s_n = nr$. Suppose now that no such r exists. Consider the image of $\mathbb{R}v$ in $V/\mathbb{Z}^d$. Its closure is a connected subgroup, and so is a torus. The dimension of this torus must be ≥ 2 for otherwise there would be an r as above. But for any finite closed interval I of $\mathbb{R}$ the image of Iv is compact, and so must stay away from 0 except at 0. Thus for any neighborhood of 0 there must be a t outside of I such that the image of tv is in that neighborhood. □

Let $\{s_n\}$ be as in the lemma. Then we can find a subsequence, $\{t_k\}$, of the sequence $\{s_n\}$, and for each k we can choose a $x_k \in \mathbb{Z}^d$, such that $\|x_k - t_k v\| < 1/k$ for all k.

Lemma 7.3. *For v, $\{t_k\}$ and $\{x_k\}$ as above, define γ by $\gamma(0) = 0$ and $\gamma(t_k) = x_k$. Then γ is an almost-geodesic ray in V which determines the same Busemann point in $\partial_\ell V$ as does the ray $t \mapsto tv$.*

Proof. Given $\varepsilon > 0$, choose N such that $1/N < \varepsilon/3$. Then for $t_n \geq t_m \geq N$ we have from the triangle inequality

$$\begin{aligned} |\|x_n - x_m\| \quad + \quad \|x_m\| - t_n| &= |\|x_n - x_m\| - \|(t_n - t_m)v\| + \|x_m\| - t_m| \\ &\leq \|(x_n - t_n v) - (x_m - t_m v)\| + \|x_m - t_m v\| < \varepsilon. \end{aligned}$$

From this it follows that γ is an almost-geodesic ray. The fact that it determines the same Busemann point as does v now follows from Proposition 4.9. □

Proposition 7.4. *Let v be a smooth point of the unit sphere of V, with support functional σ_v. Then there is a Busemann point $b_v \in \partial_\ell \mathbb{Z}^d$ such that for any $y \in \mathbb{Z}^d$ we have*

$$\varphi_y(b_v) = \sigma_v(y).$$

Furthermore, b_v is a fixed-point for the action of $\mathbb{Z}^d$ on $\partial_\ell \mathbb{Z}^d$.

Proof. Let γ be an almost-geodesic ray associated with v as in the above lemmas. By Proposition 6.2 we know that

$$\lim \varphi_y(x_k) = \sigma_v(y)$$

for all $y \in V$. But γ is equally well an almost-geodesic ray in $\mathbb{Z}^d$, and so defines a Busemann point $b_\gamma \in \partial \mathbb{Z}^d$. But for $y \in \mathbb{Z}^d$ its φ_y for $\mathbb{Z}^d$ is just the restriction to $\mathbb{Z}^d$ of its φ_y for V. Thus $\varphi_y(b_\gamma) = \sigma_v(y)$ for $y \in \mathbb{Z}^d$. The proof that b_γ is a fixed-point for the action is the same as that for Proposition 6.3. □

We remark that, just as for V, different smooth points v may have the same σ_v, and so determine the same Busemann point of $\partial_\ell \mathbb{Z}^d$, and so it can happen that only a finite number of points of $\partial_\ell \mathbb{Z}^d$ arise from smooth points v.

We are now ready to prove one part of our Main Theorem 0.1, namely:

Theorem 7.5. *Let ℓ on $\mathbb{Z}^d$ be defined by $\ell(x) = \|x\|$ for a norm $\|\cdot\|$ on $\mathbb{R}^d$. Let L_ℓ be defined on $C_c(\mathbb{Z}^d, c)$ as before by*

$$L_\ell(f) = \|[M_\ell, \pi_f]\|.$$

Then L_ℓ is a Lip-norm on $C^*(\mathbb{Z}^d, c)$.

Proof. Let v be a smooth point of the unit sphere of V for $\|\cdot\|$. Let σ_v denote its support functional, and b_v its corresponding Busemann point as above in $\partial_\ell \mathbb{Z}^d$. Since b_v is a fixed-point, it determines a homomorphism from the cosphere algebra $C^*(G, C(\partial_\ell G), \alpha, c)$ onto $C^*(G, c)$ which takes M_{φ_y} to the constant $\sigma_v(y)$. (We use here the amenability of $\mathbb{Z}^d$.) Then under this homomorphism $[M_\ell, \pi_f]$ is sent to the operator

$$\Sigma f(y) \varphi_y(b_v) \pi_y = \Sigma f(y) \sigma_v(y) \pi_y$$

in $C^*(\mathbb{Z}^d, c)$. Let us denote this operator, and the corresponding function, by $X_v f$. Of course $\|X_v f\| \leq L_\ell(f)$.

We let β denote the usual dual action [35] of the dual group $\hat{G}$ on $C^*(G, c)$ determined by

$$(\beta_p(f))(x) = \langle x, p \rangle f(x)$$

for $f \in C_c(\mathbb{Z}^d)$ and $p \in \hat{G}$, where $\langle \cdot, \cdot \rangle$ denotes the pairing of G and $\hat{G}$. Each $\tau \in V'$ determines an element of $\hat{G}$ by $\langle x, \tau \rangle = \exp(i\tau(x))$ for $x \in \mathbb{Z}^d$. Let Γ denote the lattice in V' consisting of elements which on $\mathbb{Z}^d$ take values in $2\pi\mathbb{Z}$. Then we can identify $\hat{G}$ with the torus V'/Γ, and then V' is identified with the Lie algebra of $\hat{G}$, so that the exponential mapping is just the quotient map from V' to V'/Γ. The action β has an infinitesimal version which is a Lie algebra homomorphism from the (Abelian) Lie algebra V' into the Lie algebra of derivations on $C^*(G, c)$. We denote it by $d\beta$, and it is determined by

$$(d\beta_\tau(f))(x) = i\tau(x)f(x).$$

Each $f \in C_c(G)$ then determines a linear mapping, $\tau \mapsto d\beta_\tau(f)$, from V' into $C^*(G, c)$, which we denote by df, much as done for theorem 3.1 of [39].

In terms of the notation just introduced, we see that for any smooth point v we have

$$iX_v f = d\beta_{\sigma_v} f = df(\sigma_v).$$

With this notation our earlier inequality becomes

$$\|df(\sigma_v)\| \leq L_\ell(f).$$

Now V' has the dual norm $\| \cdot \|'$, and $C^*(G, c)$ has its C^*-norm. So the norm of the linear map df between them is well-defined. We denote it by $\|df\|$. But by Proposition 6.7 the closed convex hull of the set of σ_v's is the unit ball in V'. It follows that

$$\|df\| \leq L_\ell(f).$$

But in theorem 3.1 of [39] it is shown that $f \mapsto \|df\|$ is a Lip-norm. Thus we can apply comparison lemma 1.10 of [39] to conclude that L_ℓ is a Lip-norm as well. $\square$

8. The boundary of $(\mathbb{Z}^d, S)$

Let S be a finite generating subset of $G = \mathbb{Z}^d$ such that $S = -S$ and $0 \notin S$. Let ℓ denote the corresponding word-length function on G. I do not know how to give a concrete description of $\partial_\ell G$. (But note that $\partial_\ell G$ is totally disconnected since each φ_y takes only integer values, in contrast to what happens if ℓ comes, for example, from the Euclidean norm on $\mathbb{R}^d$.) We will show here how to construct a substantial supply of geodesic rays. (Somewhat related considerations appear in [44], but geodesic rays and compactifications are not considered there.) In the next section we will show that our supply is sufficient to prove that when M_ℓ is used as the Dirac operator for $C^*(G, c)$, then the corresponding metric on the state space of $C^*(G, c)$ gives the weak-$*$ topology.

Our construction is motivated by several features which we found in Sections 6 and 7. For convenience we view $G = \mathbb{Z}^d$ as embedded in $\mathbb{R}^d$. We let $K = K_S$ denote the (closed) convex hull in $\mathbb{R}^d$ of S. Because K is balanced (since $S = -S$), it determines a norm, $\| \cdot \|_S$, on $\mathbb{R}^d$, for which it is the unit ball. (In fact, $(\mathbb{R}^d, \| \cdot \|_S)$ is the "asymptotic cone" of $(\mathbb{Z}^d, \ell)$—see exercise 8.2.12 of [8].) We will see later that this norm is relevant. The set of extreme points of K_S is a subset of S, which we will denote by S^e. The faces of K_S (of all dimensions) will have certain subsets of S^e as their extreme points, and will intersect S in certain subsets F. Such an F is characterized by the fact that there is a linear functional σ on $\mathbb{R}^n$ (not necessarily unique) such that $\sigma(s) \leq 1$ for all $s \in S$ and $F = \{s \in S : \sigma(s) = 1\}$. We call any such σ a *support functional* for F. Note that $|\sigma(s)| \leq 1$ for all $s \in S$. By abuse of terminology we will refer to F itself as a face of K_S, and we will not distinguish between F and the usual face which F determines.

Lemma 8.1. *Let σ be a support functional for a face F of K_S. Then*

$$|\sigma(x)| \leq \ell(x)$$

for all $x \in G$.

Proof. Suppose that $x = \Sigma q(s)s$ for some function q from S to $\mathbb{Z}$. Then

$$\sigma(x) = \Sigma q(s)\sigma(s) \leq \Sigma |q(s)|.$$

On considering the minimum for all such q, we see that $\sigma(x) \leq \ell(x)$. But this holds for $-x$ as well, which gives the desired result. $\square$

Let F be a face of K_S. Any function γ from $\mathbb{Z}^+$ to G which consists of successively adding elements of F (i.e., $\gamma(n+1) - \gamma(n) \in F$ for $n \geq 0$) is a geodesic ray. In fact, for any support functional σ for F the above lemma tells us that we have $n \geq \ell(\gamma(n)) \geq \sigma(\gamma(n)) = n$. Since F is finite, some (perhaps all) elements of F will have to be added in an infinite number of times. One can see that if the order in which the elements of F are added-in is changed, but the number of times they ultimately appear is the same, then one obtains an equivalent geodesic ray. A class of such geodesic rays can be specified by a function on F which has values either in $\mathbb{Z}^+$ or $+\infty$. But it seems to be tricky to decide when two such functions (possibly for different faces) determine the same Busemann point. For our present purposes we do not need to concern ourselves with this issue. It is sufficient for us to associate a canonical geodesic ray to each face. This will be a special case of forming geodesic rays by successively adding elements of the semigroup generated by F (so that the domain of the ray may be a proper subset of $\mathbb{Z}$).

Notation 8.2. *For a face F of K_S set $z_f = \Sigma\{s : s \in F\}$, and let γ_F denote the geodesic ray whose domain is $|F|\mathbb{Z}^+$ (where $|F|$ denotes the number of elements of F) and which is defined by $\gamma(|F|n) = nz_F$. We denote by b_F the corresponding Busemann point. We denote by G_F the subgroup of G generated by F.*

Again Lemma 8.1 quickly shows that the above ray is geodesic. The following proposition is analogous to Proposition 6.2.

Proposition 8.3. *Let σ be a support functional for a face F of K_S. For every $u \in G_F$ we have*

$$\varphi_u(b_F) = \sigma(u).$$

Proof. Since $u \in G_F$, there is a positive integer N such that whenever $n \geq N$ then $nz_F - u$ can be expressed as a sum of elements of F, so that $\ell(nz_F - u) = \sigma(nz_F - u)$. Of course $\ell(nz_F) = \sigma(nz_F)$. Thus for $n \geq N$

$$\varphi_u(nz_F) = \sigma(nz_F) - \sigma(nz_F - u) = \sigma(u).$$

$\square$

Proposition 8.4. *Let F and σ be as above. For any $y \in G$ and $u \in G_F$ we have*

$$\varphi_{y+u}(b_F) = \varphi_y(b_F) + \sigma(u).$$

Proof. Consider the set of u's such that this equation holds for all $y \in G$. It is easy to verify that this set is a subsemigroup of G. But for u in this set we have

$$\varphi_{y-u}(\beta_F) = \varphi_{(y-u)+u}(\beta_F) - \sigma(u) = \varphi_y(\beta_F) + \sigma(-u),$$

so that this set is a group. It thus suffices to verify the above equation for each $u = s \in F$.

So let $s \in S$. Since $n \mapsto \varphi_y(nz_F)$ is integer-valued, non-decreasing by Lemma 4.5, and bounded, we can find a positive integer N such that

$$\varphi_y(b_F) = \ell((N+m)z_F) - \ell((N+m)z_F - y)$$

for all $m \geq 0$. We can find a larger N such that also

$$\varphi_{y+s}(b_F) = \ell((N+m)z_F) - \ell((N+m)z_F - (y+s))$$

for all $m \geq 0$. Since $\sigma(s) = 1$ it is then clear that we need to show that

$$\ell((N+m)z_F - (y+s)) = \ell((N+m)z_F - y) - 1$$

for some $m \geq 0$. Let $\bar{y} = y - Nz_F$. Then what we need becomes

$$\ell(mz_F - (\bar{y}+s)) = \ell(mz_F - \bar{y}) - 1$$

for some $m \geq 0$. Note that $\ell(mz_F) - \ell(mz_F - \bar{y})$ is independent of $m \geq 0$ because of our choice of N, and similarly for $\bar{y}+s$ instead of $\bar{y}$.

Since $S = -S$ and $0 \notin S$, we can find a subset, S^+, such that $S^+ \cup (-S^+) = S$ and $S^+ \cap (-S^+) = \emptyset$. Since $F \cap (-F) = \emptyset$, we can require that $F \subseteq S^+$. Index the elements of S^+ in such a way that $s_1 = s$, and $F = \{s_1, \dots, s_{|F|}\}$, where $|F|$ denote the number of elements in F. Since S generates G, we can express $\bar{y}$ as $\bar{y} = \Sigma n_j s_j$ where $n_j \in \mathbb{Z}$ for each j. Then $\ell(\bar{y})$ will be the minimum of the sums $\Sigma|n_j|$ over all such expressions for $\bar{y}$. We make a specific choice of such a minimizing set $\{n_j\}$. (It need not be unique.)

Since $\ell(mz_F) = m|F|$ by Lemma 8.1, the stability described earlier says that $m|F| - \ell(mz_F - \bar{y})$ is independent of $m \geq 0$. We combine this for $m = 0$ and $m = 1$ to obtain $-\ell(-\bar{y}) = |F| - \ell(z_F - \bar{y})$. We use this to calculate

$$\begin{aligned} |F| \quad + \quad \Sigma|n_j| &= |F| + \ell(-\bar{y}) = \ell(z_F - \bar{y}) \\ &= \ell\left(\sum_{j \leq |F|} (1-n_j)s_j + \sum_{j > |F|} n_j s_j\right) \leq \sum_{j \leq |F|} |1-n_j| + \sum_{j>|F|} |n_j|. \end{aligned}$$

On comparing the two ends, we see that we must have $n_j \le 0$ for $j \le |F|$, and that the two ends must be equal. Thus

$$\ell(z_F - \bar{y}) = \sum_{j \le |F|} (1 - n_j) + \sum_{j > |F|} |n_j|.$$

Now

$$\begin{aligned} z_F - \bar{y} - s &= \sum_{j \le |F|} (1 - n_j) s_j + \sum_{j > |F|} n_j s_j - s_1 \\ &= -n_1 s_1 + \sum_{2}^{|F|} (1 - n_j) s_j + \sum_{j > |F|} n_j s_j. \end{aligned}$$

From the fact that $n_j \le 0$ for $j \le |F|$ it follows that

$$\begin{aligned} \ell(z_F - \bar{y} - s) &\le -n_1 + \sum_{2}^{|F|} (1 - n_j) + \sum_{j > |F|} |n_j| \\ &= -1 + \ell(z_F - \bar{y}). \end{aligned}$$

From the triangle inequality and the fact that $\ell(s) = 1$ it follows that

$$\ell(z_F - \bar{y} - s) = -1 + \ell(z_F - \bar{y}),$$

as needed. □

COROLLARY 8.5. *For any $y, z \in G$ and any $u \in G_F$, and for any support functional σ for F, we have*

$$\varphi_{y+u}(\alpha_z(b_F)) = \varphi_y(\alpha_z(b_F)) + \sigma(u).$$

Proof. Using the 1-cocycle identity 2.2 and Proposition 8.4 we obtain

$$\begin{aligned} \varphi_{y+u}(\alpha_z(b_F)) &= (\alpha_{-z}\varphi_{y+u})(b_F) = \varphi_{y-z+u}(b_F) - \varphi_{-z}(b_F) \\ &= \varphi_{y-z}(b_F) + \sigma(u) - \varphi_{-z}(b_F) \\ &= (\alpha_{-z}\varphi_y)(b_F) + \sigma(u) = \varphi_y(\alpha_z(b_F)) + \sigma(u). \end{aligned}$$

□

PROPOSITION 8.6. *Let F be a face of K. For each $u \in G_F$ the homeomorphism α_u of $\bar{G}^\ell$ leaves fixed each point of the α-orbit of b_F. That is, for each $z \in G$ we have*

$$\alpha_u(\alpha_z(b_F)) = \alpha_z(b_F).$$

Proof. Because G is Abelian, it suffices to show that $\alpha_u(b_F) = b_F$. For this we must verify that $f(\alpha_u(b_F)) = f(b_F)$ for all $f \in C(\bar{G}^\ell)$. It suffices to verify this for $f = \varphi_y$ for each $y \in G$. But from the 1-cocycle identity 2.2 and Proposition 8.4 we have

$$\begin{aligned} \varphi_y(\alpha_u(b_F)) &= (\alpha_{-u}\varphi_y)(b_F) = \varphi_{y-u}(b_F) - \varphi_{-u}(b_F) \\ &= \varphi_y(b_F) + \sigma(-u) + \sigma(u) = \varphi_y(b_F). \end{aligned}$$

□

We will also need the following fact:

Proposition 8.7. *If $y \notin G_F$ then φ_y is not constant on the G-orbit of b_F, and in fact there is an $s \in S$ such that $s \notin F$ and*

$$\varphi_y(\alpha_s(b_F)) = \varphi_y(b_F) + (1 - \varphi_{-s}(b_F)),$$

with $\varphi_{-s}(b_F) = 0$ or -1.

Proof. Let S^+ and the indexing $\{s_j\}$ be as in the proof of Proposition 8.4. Much as in that proof, we can find a large enough N that $\varphi_{y\pm s_j}((N+m)z_F)$ is constant for $m \geq 0$ for all $\pm s_j$ simultaneously, as is $\varphi_y((N+m)z_F)$. Set $\bar{y} = y - Nz_F$. For this $\bar{y}$ choose $\{n_j\}$ as before so that $\bar{y} = \Sigma n_j s_j$ and $\ell(\bar{y}) = \Sigma|n_j|$. Since $y \notin G_F$, also $\bar{y} \notin G_F$, and so there is a $k > |F|$ such that $n_k \neq 0$. Suppose that $n_k \geq 1$. Then

$$\bar{y} - s_k = \sum_{j \neq k} n_j s_j + (n_k - 1)s_k,$$

so that

$$\ell(\bar{y} - s_k) \leq \sum_{j \neq k} |n_j| + n_k - 1 = \ell(\bar{y}) - 1.$$

From the triangle inequality we then obtain $\ell(\bar{y} - s_k) = \ell(\bar{y}) - 1$, that is,

$$\ell(Nz_F - y + s_k) = \ell(Nz_F - y) - 1.$$

From our choice of N (and with $m = 0$) we then get

$$\begin{aligned}
\varphi_{y-s_k}(b_F) &= \varphi_{y-s_k}(Nz_F) \\
&= N|F| - \ell(Nz_F - y + s_k) = N|F| - \ell(N|F| - y) + 1 \\
&= \varphi_y(b_F) + 1.
\end{aligned}$$

We combine this with the 1-cocycle identity 2.2 to obtain

$$\begin{aligned}
\varphi_y(\alpha_{s_k}(b_F)) &= (\alpha_{-s_k}\varphi_y)(b_F) \\
&= \varphi_{y-s_k}(b_F) - \varphi_{-s_k}(b_F) \\
&= \varphi_y(b_F) + (1 - \varphi_{-s_k}(b_F)).
\end{aligned}$$

Since φ_{-s_k} takes only the values 0, ± 1, the desired conclusion is then obtained from:

Lemma 8.8. *If $s \in S$ and $\varphi_s(b_F) = 1$ then $s \in F$.*

Proof. If $\varphi_s(b_F) = 1$, then for large n, and for a support functional σ for F, we have

$$\begin{aligned}
n|F| - 1 &= \ell(nz_F) - 1 = \ell(nz_F - s) \\
&\geq \sigma(nz_F - s) = n|F| - \sigma(s),
\end{aligned}$$

so that $1 \leq \sigma(s)$, and so $s \in F$. □

The above argument for the proof of Proposition 8.7 was under the assumption that $n_k \geq 1$. If instead we have $n_k \leq -1$, then we carry out a similar argument using $-s_k$ instead of s_k. This concludes the proof of Proposition 8.7. □

9. Word-length functions give Lip-norms on $C^*(\mathbb{Z}^d, c)$

We will now see how to use the results of the previous section to prove the part of our Main Theorem 0.1 concerning word-length functions. We use the notation of the previous section, and in particular, the norm $\|\cdot\|_S$ determined by $K = K_S$. Here we will consider the (proper) faces of K of maximal dimension, namely of dimension $d-1$. We will call them "facets" of K, as is not infrequently done. The interior points of the facets are the smooth points of the unit sphere for $\|\cdot\|_S$. Again our terminology and notation will not distinguish between facets as intersections of K with hyperplanes, and as the corresponding subsets of S. Because K has only a finite number of extreme points, every point of the boundary of K is contained in at least one facet, and there are only a finite number of facets. Each facet F has a unique support functional, which we denote by σ_F. Furthermore, F contains a basis for $\mathbb{R}^d$, and consequently G_F is of finite index in G. This has the crucial consequence for us that the orbit, $\mathcal{O}_F$, of b_F in $\partial_\ell G$ under the action α, is finite. (Apply Proposition 8.6.) We consider the restriction map from $C(\partial_\ell G)$ onto $C(\mathcal{O}_F)$. Since it is α-equivariant, it gives an algebra homomorphism, Π_F, from $C^*(G, C(\partial_\ell G), \alpha, c)$ onto $C^*(G, C(\mathcal{O}_F), \alpha, c)$. If we let π and M denote also the corresponding homomorphisms of G and $C(\mathcal{O}_F)$ into this latter algebra, and if for each $y \in G$ we let ψ_y denote the restriction of φ_y to $\mathcal{O}_F$, then

$$\Pi_F([M_\ell, \pi_f]) = \Sigma f(y) M_{\psi_y} \pi_y.$$

Let Q be a set of coset representatives for G_F in G containing 0. Then we can express the above as

$$\Sigma_{q\in Q}(\Sigma_{u\in G_F} f(u+q) M_{\psi_{u+q}} \bar{c}(u,q)\pi_u)\pi_q.$$

From Corollary 8.5 we see that $\psi_{u+q} = \psi_q + \sigma_F(u)$. For each q let g^q be the function on G_F defined by $g^q(u) = f(u+q)\bar{c}(u,q)$. We can also view g^q as a function on G by giving it value 0 off G_F. Then we can rewrite our previous expression for $\Pi_F([M_\ell, \pi_f])$ as

$$\Sigma_q(\Sigma_u g^q(u)(\sigma_F(u) + M_{\psi_q})\pi_u)\pi_q.$$

As in Section 7 let $\hat{G} = \mathbb{T}^d$ be the dual group of G, and denote the pairing between G and $\hat{G}$ by $\langle x, s\rangle$. Let β now denote the usual dual action of $\hat{G}$ on $C^*(G, C(\mathcal{O}_F), \alpha, c)$, so that

$$\beta_s(M_\psi \pi_x) = \langle x, s\rangle M_\psi \pi_x.$$

Then the finite group $(G/G_F)^\wedge$ can be identified with the set of characters on G which take value 1 on G_F. We can thus restrict β to $(G/G_F)^\wedge$ and average over $(G/G_F)^\wedge$. This gives a projection of norm 1 onto the subalgebra of elements supported on G_F, and this projection on functions on G is just restriction of functions to G_F. If for each fixed q we apply this projection to the product with π_q^* of the above expression for $\Pi_F([M_\ell, \pi_f])$, we find that

$$\|[M_\ell, \pi_f]\| \geq \|\Sigma_u g^q(u)(\sigma_F(u) + M_{\psi_q})\pi_u\|$$

for each q. The norm on the right is that of $C^*(G, C(\mathcal{O}_F), \alpha, c)$. But section 2.27 of [47] tells us that $C^*(G_F, C(\mathcal{O}_F), \alpha, c)$ is a C^*-subalgebra of $C^*(G, C(\mathcal{O}_F), \alpha, c)$ under the evident identification of functions. Thus we can view the operator on the right as being in $C^*(G_F, C(\mathcal{O}_F), \alpha, c)$, where we are here restricting α and c to G_F. But from Proposition 8.6 we know that the action α of G_F on $\mathcal{O}_F$ is trivial. Thus we have the decomposition

$$C^*(G_F, C(\mathcal{O}_F), \alpha, c) \cong C(\mathcal{O}_F) \otimes C^*(G_F, c).$$

Let $a_q = \Sigma g^q(u)\pi_u$ and $b_q = \Sigma g^q(u)\sigma_F(u)\pi_u$. Then in terms of the above decomposition we are looking at $I \otimes b_q + \psi_q \otimes a_q$. From Proposition 8.7 we know that ψ_q is not constant on $\mathcal{O}_F$ for $q \neq 0$. Note that $\psi_0 \equiv 0$. For given $q \neq 0$ let m_j for $j = 1, 2$ be two distinct values of ψ_q. Upon evaluating at the points where ψ_q takes these values, and using our earlier inequality, we see that

$$\|b_q + m_j a_q\| \leq \|[M_\ell, \pi_f]\| = L_\ell(f)$$

for $j = 1, 2$. Upon writing the inequalities as

$$\|m_j^{-1} b_q + a_q\| \leq |m_j|^{-1} L_\ell(f)$$

and using the triangle inequality to eliminate a_q, and simplifying, we find that

$$\|b_q\| \leq (|m_1| + |m_2|)/|m_1 - m_2| L_\ell(f).$$

(If either m_j is 0 the path is simpler.) Of course m_1 and m_2 depend on q. Thus we see that we have found a constant, k_q, such that $\|b_q\| \leq k_q L_\ell(f)$. For $q = 0$ we have the same inequality with $k_0 = 1$ since $\psi_0 = 0$. Much as in Section 7 set $X_F f = \Sigma \sigma_F(x) f(x) \pi_x$. Then

$$\begin{aligned} X_F f &= \Sigma \sigma_F(x) f(x) \pi_x = \Sigma_q (\Sigma_{u \in G_F} \sigma_F(u+q) f(u+q) \bar{c}(u, q) \pi_u) \pi_q \\ &= \Sigma_q \sigma_F(q) (\Sigma_u \sigma_F(u) g^q(u) \pi_u) \pi_q = \Sigma \sigma_F(q) b_q \pi_q. \end{aligned}$$

When we combine this with the inequality obtained earlier for $\|b_q\|$, we obtain

$$\|X_F f\| \leq (\Sigma |\sigma_F(q)| k_q) L_\ell(f).$$

Observe that the $\sigma_F(q)$'s and k_q's do not depend on f, but only on F and the choice Q of coset representatives. Thus for each facet F we have obtained a constant, k_F, such that

$$\|X_F f\| \leq k_F L_\ell(f)$$

for all $f \in C_c(G)$. Note that knowing that k_F is finite is the crucial place where we use that the number of coset representatives in Q is finite.

Just as toward the end of Section 7, we have the dual action β of $\mathbb{T}^d$ on $C^*(G, c)$, and the corresponding differential df of any $f \in C_c(G)$, such that $df(\sigma_F) = iX_F f$. Then our inequality above gives, much as in Section 7,

$$\|df(\sigma_F)\| \leq k_F L_\ell(f).$$

Recall now the norm $\|\cdot\|_S$ determined by $K = K_S$. The σ_F's are exactly the support functionals corresponding to the smooth points of the unit sphere for $\|\cdot\|_S$. Let $\|df\|_S$ denote the norm of the linear map df using the dual norm

$\|\cdot\|'_S$. Also let $k = \max\{k_F : F \text{ is a facet}\}$. Then from Proposition 6.7 we conclude, much as in Section 7, that

$$\|df\|_S \leq kL_\ell(f).$$

Then just as in Section 7 we conclude that L_ℓ is a Lip-norm. This concludes the proof of Main Theorem 0.1.

□

Since the norm $\|\cdot\|'_S$ on V' does not come from an inner product, and V' can be thought of as the analogue of the tangent space at the non-existent points of the quantum space $C^*(G,c)$, we can consider that we have here a non-commutative Finsler geometry (as also in section 3 of [39]). The metric geometry from L_ℓ also, in a vague way, seems Finsler-like.

I imagine that the above considerations can be generalized so that the Main Theorem can be extended to weighted-word-length functions, where each generator has been assigned a weight. I imagine that they can also be generalized to deal with extensions of $\mathbb{Z}^d$ by finite groups. But I have not explored these possibilities.

Since our estimates for the proof of the Main Theorem depend just on the behavior of the φ_y's on the boundary, the conclusions of the Main Theorem will also be valid if ℓ is replaced by the translation-bounded function $\ell + h$ where h is any function in $C_\infty(\mathbb{Z}^d)$.

10. The free group

We briefly discuss here how the ideas developed earlier apply to the free (non-Abelian) group on two generators, $G = F_2$. Denote the two generators by a and b, and take them and their inverses as our generating set S. Let ℓ denote the corresponding length function. It is well-known [21] that F_2 is a hyperbolic group, and that its Gromov boundary, $\partial_h G$, is described as the set of all infinite (to the right) reduced words in the elements of S. (The "h" in $\partial_h G$ is for "hyperbolic"—it does not denote a length function.) The action of G on $\partial_h G$ is the evident one by "left concatenation" (and then reduction). We can obtain the topology of $\partial_h G$ and of the compactification of G as follows. (See comment ii) on page 104 of [21].) To include the elements of G we need a "stop" symbol. We denote it by p. We let S' denote S with p added, and we let $\prod^\infty S'$ denote the set of sequences with values in S', with its compact topology of "index-wise" convergence.

Notation 10.1. *Let $\bar{G}^h$ be the subset of $\prod^\infty S'$ consisting of all sequences such that*

1) *If p occurs in the sequence then all subsequent letters in that sequence are p.*
2) *The sequence is reduced, in the sense that a and a^{-1} are never adjacent entries, and similarly for b and b^{-1}.*

It is easily seen that $\bar{G}^h$ is a closed subset of $\prod^{\infty} S'$, so compact. We identify the elements of G with the words containing p (and in particular, we identify the identity element of G with the constant sequence with value p). With this understanding, it is easily seen that G is an open dense subset of $\bar{G}^h$. We identify $\partial_h G$ with the infinite words which do not contain p.

The group G again acts on $\bar{G}^h$ by left concatenation. It is easily seen that this action is by homeomorphisms. Consider the function φ_a on G. For any word w we have $\ell(a^{-1}w) = \ell(w) + 1$ if w begins with the letters a^{-1}, b or b^{-1}, or is the identity element, while $\ell(a^{-1}w) = \ell(w) - 1$ if w begins with the letter a. Thus $\varphi_a(w) = \ell(w) - \ell(a^{-1}w)$ has value 1 if w begins with the letter a, and value -1 otherwise. But we can extend φ_a to $\bar{G}^h$ by exactly this same prescription, and it is easily seen that this extended φ_a is continuous on $\bar{G}^h$. We do the same with φ_b, $\varphi_{a^{-1}}$ and $\varphi_{b^{-1}}$. By using the 1-cocycle identity 2.2 inductively, we see that each φ_x for $x \in G$ extends to a continuous function on $\bar{G}^h$ (in a unique way since G is dense). Of course the functions in $C_\infty(G)$ extend by giving them value 0 on $\partial_h G$, and the constant functions also extend. In this way we identify $C(\bar{G}^\ell)$ with a unital subalgebra of $C(\bar{G}^h)$.

Let us see now that the subalgebra $C(\bar{G}^\ell)$ separates the points of $\bar{G}^h$. Because the subalgebra contains $C_\infty(G)$, it is clear that we only need to treat the points of $\partial_h G$. Let $v, w \in \partial_h G$ with $v \neq w$. Then there must be a first entry where they differ. That is, we can write them as $v = x\tilde{v}$, $w = x\tilde{w}$ where x is a finite word while $\tilde{v}$ and $\tilde{w}$ differ in their first entry. Suppose the first entry of $\tilde{v}$ is a while the first entry of $\tilde{w}$ is not a. Then from what we saw above

$$(\alpha_x \varphi_a)(v) = \varphi_a(x^{-1}v) = \varphi_a(\tilde{v}) = 1,$$

while in the same way $(\alpha_x \varphi_a)(w) = -1$. Thus the subalgebra $C(\bar{G}^\ell)$ separates the points of $\bar{G}^h$, and so by the Stone–Weierstrass theorem $C(\bar{G}^\ell) = C(\bar{G}^h)$, so that $\bar{G}^\ell = \bar{G}^h$. Thus in this case the metric and hyperbolic boundaries coincide. (The referee has pointed out that if instead we take as generating set $\{a^{\pm 1}, a^{\pm 2}, b^{\pm 1}\}$, then the resulting metric compactification will be different from that described above, because just as in Example 5.2 we will obtain two "parallel" geodesic rays, namely $(e, a^2, a^4, \dots)$ and $(a, a^3, a^5, \dots)$, which will give different Busemann points.)

Each $w \in \partial_h G$ specifies a unique geodesic ray to it from e, namely $e, w_1, w_1 w_2, w_1 w_2 w_3, \dots$. Thus every point of $\partial_\ell G$ is a Busemann point. It is well-known [2] that the action of G on $\partial_h G$ is amenable. If one uses the definition of amenability in terms of maps from $\partial_h G$ to probability measures on G which was stated in Section 3, then this is seen by letting the n-th map, m_n, be the map which assigns to $w \in \partial_h G$ the probability measure which gives mass $1/n$ to the first n points of the geodesic ray from e to w [2]. In view of Theorem 3.7 this implies that the cosphere algebra $S^*_\ell A$ for the spectral triple $(A = C^*_r(F_2), \boldsymbol{\ell}^2(F_2), M_\ell)$ is $C^*(G, C(\partial_h F_2), \alpha)$.

However, the action α on $\partial_h F_2$ does not have any finite orbits, and so I do not see how to continue along the lines of the previous section to determine whether the metric on the state space $S(C^*_r(F_2))$ coming from the above spectral triple gives the state space the weak-$*$ topology, or even just finite diameter. The difficulty remains: What information can one obtain about $\|\pi_f\|$ if one knows that $\|[M_\ell, \pi_f]\| \leq 1$?

References

[1] C. Anantharaman-Delaroche. Systèmes dynamiques non commutatifs et moyennabilité. *Math. Ann.*, 279(2):297–315, 1987.

[2] C. Anantharaman-Delaroche. Amenability and exactness for dynamical systems and their C^*-algebras. arXiv:math.OA/0005014.

[3] C. Anantharaman-Delaroche and J. Renault. *Amenable groupoids.* L'Enseignement Mathématique, Geneva, 2000.

[4] W. Ballmann. *Lectures on spaces of nonpositive curvature.* Birkhäuser Verlag, Basel, 1995.

[5] W. Ballmann, M. Gromov, and V. Schroeder. *Manifolds of nonpositive curvature.* Birkhäuser Boston Inc., Boston, MA, 1985.

[6] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature.* Springer-Verlag, Berlin, 1999.

[7] T. Bröcker and T. tom Dieck. *Representations of compact Lie groups.* Springer-Verlag, New York, 1995.

[8] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry.* American Mathematical Society, Providence, RI, 2001.

[9] A. H. Chamseddine and A. Connes. Universal formula for noncommutative geometry actions: unification of gravity and the standard model. *Phys. Rev. Lett.*, 77(24):4868–4871, 1996.

[10] A. H. Chamseddine and A. Connes. The spectral action principle. *Comm. Math. Phys.*, 186(3):731–750, 1997. arXiv:hep-th/9606001

[11] A. Connes. Compact metric spaces, Fredholm modules, and hyperfiniteness. *Ergodic Theory Dynamical Systems*, 9(2):207–220, 1989.

[12] A. Connes. *Noncommutative Geometry.* Academic Press Inc., San Diego, CA, 1994.

[13] A. Connes. Geometry from the spectral point of view. *Lett. Math. Phys.*, 34(3):203–238, 1995.

[14] A. Connes. Noncommutative geometry and reality. *J. Math. Phys.*, 36(11):6194–6231, 1995.

[15] A. Connes. Gravity coupled with matter and the foundation of noncommutative geometry. *Comm. Math. Phys.*, 182(1):155–176, 1996. arXiv:hep-th/9603053.

[16] J. B. Conway. *A course in functional analysis.* Springer-Verlag, New York, second edition, 1990.

[17] N. Dunford and J. T. Schwartz. *Linear operators. Part I.* John Wiley & Sons Inc., New York, 1988. Reprint of the 1958 original, A Wiley-Interscience Publication.

[18] J. M. G. Fell. *An extension of Mackey's method to Banach *- algebraic bundles.* American Mathematical Society, Providence, R.I., 1969.

[19] J. M. G. Fell. *Induced representations and Banach *-algebraic bundles.* Springer-Verlag, Berlin, 1977. Lecture Notes in Mathematics, Vol. 582.

[20] J. M. G. Fell and R. S. Doran. *Representations of *-algebras, locally compact groups, and Banach *-algebraic bundles. Vol. 1.* Academic Press Inc., Boston, MA, 1988.

[21] É. Ghys and P. de la Harpe, editors. *Sur les groupes hyperboliques d'après Mikhael Gromov.* Birkhäuser Boston Inc., Boston, MA, 1990.

[22] F. Golse and E. Leichtnam. Applications of Connes' geodesic flow to trace formulas in noncommutative geometry. *J. Funct. Anal.*, 160(2):408–436, 1998.

[23] J. M. Gracia-Bondía, J. C. Várilly, and H. Figueroa. *Elements of noncommutative geometry.* Birkhäuser Boston Inc., Boston, MA, 2001.

[24] M. Gromov. Hyperbolic manifolds, groups and actions. In *Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference*, pages 183–213, Princeton Univ. Press, Princeton, N.J., 1981.

[25] N. Higson. Bivariant K-theory and the Novikov conjecture. *Geom. Funct. Anal.*, 10(3):563–581, 2000.

[26] N. Higson and J. Roe. Amenable group actions and the Novikov conjecture. *J. Reine Angew. Math.*, 519:143–153, 2000.

[27] R. Ji. Smooth dense subalgebras of reduced group C^*-algebras, Schwartz cohomology of groups, and cyclic cohomology. *J. Funct. Anal.*, 107(1):1–33, 1992.

[28] R. Ji and L. B. Schweitzer. Spectral invariance of smooth crossed products, and rapid decay locally compact groups. *K-Theory*, 10(3):283–305, 1996.

[29] P. Jolissaint. Rapidly decreasing functions in reduced C^*-algebras of groups. *Trans. Amer. Math. Soc.*, 317(1):167–196, 1990.

[30] P. Jolissaint. K-theory of reduced C^*-algebras and rapidly decreasing functions on groups. *K-Theory*, 2(6):723–735, 1989.

[31] J. Kaminker and G. Yu. Boundary amenability of groups and positive scalar curvature. *K-Theory*, 18(1):93–97, 1999.

[32] V. Lafforgue. Une démonstration de la conjecture de Baum-Connes pour les groupes réductifs sur un corps p-adique et pour certains groupes discrets possédant la propriété (T). *C. R. Acad. Sci. Paris Sér. I Math*, 327:439–444, 1998.

[33] J. Lindenstrauss and L. Tzafriri. *Classical Banach spaces. II Function spaces.* Springer-Verlag, Berlin, 1979.

[34] L. H. Loomis. *An introduction to abstract harmonic analysis.* D. Van Nostrand Company, Inc., Toronto-New York-London, 1953.

[35] G. K. Pedersen. *C^*-algebras and their automorphism groups.* Academic Press Inc. London, 1979.

[36] J. Renault. The ideal structure of groupoid crossed product C^*-algebras. *J. Operator Theory*, 25(1):3–36, 1991.

[37] M. A. Rieffel. On the uniqueness of the Heisenberg commutation relations. *Duke Math. J.*, 39:745–752, 1972.

[38] M. A. Rieffel. Noncommutative tori—a case study of noncommutative differentiable manifolds. In *Geometric and topological invariants of elliptic operators*, pages 191–211. Amer. Math. Soc., Providence, RI, 1990.

[39] M. A. Rieffel. Metrics on states from actions of compact groups. *Doc. Math.*, 3:215–229 (electronic), 1998. arXiv:math.OA/9807084.

[40] M. A. Rieffel. Metrics on state spaces. *Doc. Math.*, 4:559–600 (electronic), 1999. arXiv:math.OA/9906151.

[41] M. A. Rieffel. Gromov-Hausdorff distance for quantum metric spaces. 2000. to appear Memoirs Amer. Math. Soc., arXiv:math.OA/0011063.

[42] M. A. Rieffel. Matrix algebras converge to the sphere for quantum Gromov-Hausdorff distance. 2001. to appear Memoirs Amer. Math. Soc., arXiv:math.OA/0108005.

[43] J. Roe. Coarse cohomology and index theory on complete Riemannian manifolds. *Mem. Amer. Math. Soc.*, 104(497), 1993.

[44] M. Shapiro. Pascal's triangles in abelian and hyperbolic groups. *J. Austral. Math. Soc. Ser. A*, 63(2):281–288, 1997.

[45] N. Weaver. *Lipschitz Algebras*. World Scientific, Singapore, 1999.

[46] A. Weil. *L'intégration dans les groupes topologiques et ses applications, Second Edition*. Hermann et Cie., Paris, 1953.

[47] G. Zeller-Meier. Produits croisés d'une C^*-algèbre par un groupe d'automorphismes. *J. Math. Pures Appl. (9)*, 47:101–239, 1968.

Marc Rieffel
Department of Mathematics
University of California
Berkeley, CA 94720-3840, USA
rieffel@math.berkeley.edu

Asymptotics of Complete Kähler-Einstein Metrics – Negativity of the Holomorphic Sectional Curvature

Georg Schumacher

Received: April 18, 2002
Revised: December 12, 2002

Communicated by Thomas Peternell

Abstract. We consider complete Kähler-Einstein metrics on the complements of smooth divisors in projective manifolds. The estimates proven earlier by the author [5] imply that in directions parallel to the divisor at infinity the metric tensor converges to the Kähler-Einstein metric on the divisor. Here we show that the holomorphic sectional curvature is bounded from above by a negative constant near infinity.

2000 Mathematics Subject Classification: 32Q05, 32Q20, 53C55
Keywords and Phrases: Complete Kähler-Einstein metrics, negative sectional curvature

1. Introduction and Statement of the Result

Complete Kähler-Einstein metrics ds^2_X of constant Ricci curvature on quasi-projective varieties are of interest in various geometric situations. The existence of a unique complete Kähler-Einstein metric ds^2_X of constant Ricci curvature -1 on a manifold X of the form $\overline{X}\backslash C$, where $\overline{X}$ is a compact complex manifold and C a smooth divisor is guaranteed under the condition

$$K_{\overline{X}} + C > 0 \qquad (N)$$

(cf. [2, 3, 7]).

In [5], asymptotic properties of ds^2_X (with Kähler form ω_X) were investigated. In the special asymptotic situation, it is possible to prove estimates for the curvature tensor based on constant negative Ricci curvature. We obtain the following theorem.

Theorem 1. *Let $\overline{X}$ be a compact complex surface, $C \subset \overline{X}$ a smooth divisor satisfying (N). Then the holomorphic sectional curvature of the complete Kähler-Einstein metric on $X = \overline{X} \backslash C$ is bounded from above by a negative constant near the compactifying divisor. The sectional and holomorphic bisectional curvatures are bounded on X.*

For $X = \mathbb{P}_2$, the assumption of the theorem is satisfied, if the degree of the curve C is at least four. So the statement of the above theorem appears to be related to the Kobayashi conjecture about the algebraic degeneracy of entire holomorphic curves and the hyperbolicity of complements of plane curves.

2. Properties of the Complete Kähler-Einstein Metric

The above condition (N) also implies the existence of a Kähler-Einstein metric ω_C on C. Using a canonical section σ of $[C]$ as a local coordinate function on $\overline{X}$ one can restrict the complete Kähler-Einstein metric ω_X to the locally defined sets $C_{\sigma_0} = \{\sigma = \sigma_0\}$. The notion of locally uniform convergence of $\omega_X|C_{\sigma_0}$, where $\sigma_0 \to 0$, makes sense.

Theorem 2 ([5]). *The Kähler-Einstein metric ω_X converges to the Kähler-Einstein metric ω_C, when restricted to directions parallel to C.*

The result is a precise analytic version of the adjunction formula: Let h be an hermitian metric on $[C]$ and Ω^∞ a volume form of class C^∞ on $\overline{X}$, such that the restriction of Ω^∞/h to C is the Kähler-Einstein volume form on C.
In terms of the Hölder spaces $C^{k,\lambda}(X)$ and $C^{k,\lambda}(W)$ for open subsets $W \subset X$ depending on quasi-coordinates used in [1] by Cheng and Yau (cf. [2, 3]), the above statement follows from the estimate below [5]: *There exists a number $0 < \alpha \leq 1$ such that for all $k \in \mathbb{N}$ and all $0 < \lambda < 1$ the volume form of the complete Kähler-Einstein metric is of the form*

$$\frac{2\Omega^\infty}{\|\sigma\|^2 \log^2(1/\|\sigma\|^2)} \left(1 + \frac{\nu}{\log^\alpha(1/\|\sigma\|^2)}\right) \text{ with } \nu \in C^{k,\lambda}(W) \tag{1}$$

For any given $\nu \in C^{k,\lambda}(W)$ according to [5, 2] there exists some $\mu \in C^{k-1,\lambda}(W)$ such that

$$\frac{\partial \nu}{\partial \sigma} = \frac{\mu}{\sigma \log(1/|\sigma|^2)}.$$

From now on we assume that $\dim_{\mathbb{C}}(X) = 2$.

Let (σ, w) be local coordinates near $[C]$, and $g_{\sigma\overline{\sigma}}, g_{\sigma\overline{w}}$ etc. the coefficients of the metric tensor of the Kähler-Einstein metric ω_X.

$$g_{\sigma\overline{\sigma}} = \frac{2}{|\sigma|^2 \log^2(1/|\sigma|^2)} \left(1 + \frac{g^0_{\sigma\overline{\sigma}}}{\log^\alpha(1/|\sigma|^2)}\right) \tag{2}$$

$$g_{\sigma\overline{w}} = \frac{g^0_{\sigma\overline{w}}}{\sigma \log^{1+\alpha}(1/|\sigma|^2)} \tag{3}$$

$$g_{w\overline{\sigma}} = \frac{g^0_{w\overline{\sigma}}}{\overline{\sigma} \log^{1+\alpha}(1/|\sigma|^2)} \tag{4}$$

$$g_{w\overline{w}} = g^\infty_{w\overline{w}} \left(1 + \frac{g^0_{w\overline{w}}}{\log^\alpha(1/|\sigma|^2)}\right) \tag{5}$$

where $g^0_{\sigma\overline{\sigma}}$, $g^0_{\sigma\overline{w}}$, $g^0_{w\overline{\sigma}}$, $g^0_{w\overline{w}}$ are in $C^{k,\lambda}(W)$, whereas $g^\infty_{w\overline{w}}$ is of class C^∞, and $\omega_C = \sqrt{-1}(g^\infty_{w\overline{w}}|C) dw \wedge \overline{dw}$ has constant curvature -1. We observe that in the determinant of the components of the metric tensor the diagonal terms dominate the rest. Moreover:

PROPOSITION 1.

$$g^{\overline{\sigma}\sigma} \sim |\sigma|^2 \log^2(1/|\sigma|^2) \tag{6}$$

$$g^{\overline{\sigma}w}, g^{\overline{w}\sigma} = O\left(|\sigma| \log^{1-\alpha}(1/|\sigma|^2)\right) \tag{7}$$

$$g^{\overline{w}w} \sim 1 \tag{8}$$

where $p \sim q$ denotes the existence of some $C > 1$ such that $(1/C) \cdot p \leq q \leq C \cdot p$.

3. ASYMPTOTICS OF THE CURVATURE TENSOR

The order of the arguments is critical. We begin with the off-diagonal terms of the curvature tensor, which require special attention. In the sequel we need the volume form Ω_X in our local coordinates (σ, w), and we set $D := g_{\sigma\overline{\sigma}} \cdot g_{w\overline{w}} - g_{\sigma\overline{w}} \cdot g_{w\overline{\sigma}}$, i.e.

(9)

$$D = \frac{2g^\infty_{w\overline{w}}}{|\sigma|^2 \log^2(1/|\sigma|^2)} \left(1 + \frac{g^0_{\sigma\overline{\sigma}} + g^0_{w\overline{w}}}{\log^\alpha(1/|\sigma|^2)} + \frac{g_{\sigma\overline{\sigma}} \cdot g_{w\overline{w}} - (g^0_{\sigma\overline{w}} \cdot g^0_{w\overline{\sigma}}/2g^\infty_{w\overline{w}})}{\log^{2\alpha}(1/|\sigma|^2)}\right).$$

We estimate

$$R_{\sigma\overline{w}\sigma\overline{w}} = O\left(1/\log^{2+\alpha}(1/|\sigma|^2)\right) \tag{10}$$

Proof of (10). We compute $-D \cdot R_{\sigma\overline{w}\sigma\overline{w}} = D \cdot \frac{\partial^2 g_{\sigma\overline{w}}}{\partial\sigma\overline{\partial w}} - \frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma} g_{w\overline{w}} \frac{\partial g_{\sigma\overline{w}}}{\overline{\partial w}} + \frac{\partial g_{\sigma\overline{w}}}{\partial\sigma} g_{w\overline{\sigma}} \frac{\partial g_{\sigma\overline{w}}}{\overline{\partial w}} + \frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma} g_{\sigma\overline{w}} \frac{\partial g_{w\overline{w}}}{\overline{\partial w}} - \frac{\partial g_{\sigma\overline{w}}}{\partial\sigma} g_{\sigma\overline{\sigma}} \frac{\partial g_{w\overline{w}}}{\overline{\partial w}}$. We gather the first three terms: $D \cdot \frac{\partial^2 g_{\sigma\overline{w}}}{\partial\sigma\overline{\partial w}} = \frac{-2g^\infty_{w\overline{w}} \cdot \frac{\partial g^0_{\sigma\overline{w}}}{\overline{\partial w}}}{|\sigma|^2\sigma^2 \log^{3+\alpha}(1/|\sigma|^2)} \cdot \left(1 + \frac{g^0_{\sigma\overline{\sigma}} + g^0_{w\overline{w}}}{\log^\alpha(1/|\sigma|^2)} + \frac{g_{\sigma\overline{\sigma}} \cdot g_{w\overline{w}} - (g^0_{\sigma\overline{w}} \cdot g^0_{w\overline{\sigma}}/2g^\infty_{w\overline{w}})}{\log^{2\alpha}(1/|\sigma|^2)}\right) \cdot \left(1 + O\left(1/\log^1(1/|\sigma|^2)\right)\right)$. Next $-\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma} g_{w\overline{w}} \frac{\partial g_{\sigma\overline{w}}}{\overline{\partial w}} = \frac{-2g^\infty_{w\overline{w}} \cdot \frac{\partial g^0_{\sigma\overline{w}}}{\overline{\partial w}}}{|\sigma|^2\sigma^2 \log^{3+\alpha}(1/|\sigma|^2)} \cdot$

$\left(1+\frac{g^0_{\sigma\overline{\sigma}}+g^0_{w\overline{w}}}{\log^\alpha(1/|\sigma|^2)}+\frac{g^0_{\sigma\overline{\sigma}}\cdot g^0_{w\overline{w}}}{\log^{2\alpha}(1/|\sigma|^2)}\right)\cdot\left(1+O\left(1/\log^1(1/|\sigma|^2)\right)\right)$, and $\frac{\partial g_{\sigma\overline{w}}}{\partial\sigma}g_{w\overline{\sigma}}\frac{\partial g_{\sigma\overline{w}}}{\partial w}=$ $\frac{-g_{\sigma\overline{\sigma}}\cdot g_{w\overline{w}}\cdot\frac{\partial g^0_{\sigma\overline{w}}}{\partial w}}{|\sigma|^2\sigma^2\log^{3+\alpha}(1/|\sigma|^2)}\cdot\left(1+O\left(1/\log^1(1/|\sigma|^2)\right)\right)$.

Hence the sum of the first three terms is of the form $O\left(1/|\sigma|^4\log^{4+3\alpha}(1/|\sigma|^2)\right)$. Concerning the sum of the last two terms $\left(\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma}\cdot g_{\sigma\overline{w}}-\frac{\partial g_{\sigma\overline{w}}}{\partial\sigma}\cdot g_{\sigma\overline{\sigma}}\right)\cdot\frac{\partial g_{w\overline{w}}}{\partial w}$ we observe that both $\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma}\cdot g_{\sigma\overline{w}}$, and $\frac{\partial g_{\sigma\overline{w}}}{\partial\sigma}\cdot g_{\sigma\overline{\sigma}}$ are of the form

$\frac{-2g^0_{\sigma\overline{w}}}{\sigma^2|\sigma|^2\log^{3+\alpha}(1/|\sigma|^2)}\cdot\left(1+\frac{g^0_{\sigma\overline{\sigma}}}{\log^\alpha(1/|\sigma|^2)}\right)\cdot\left(1+O\left(1/\log^1(1/|\sigma|^2)\right)\right)$.

Hence again the sum is of order $O\left(1/|\sigma|^4\log^{4+3\alpha}(1/|\sigma|^2)\right)$. □

Next, we claim

$$R_{\sigma\overline{\sigma}w\overline{w}}=O\left(1/\log^{2+\alpha}(1/|\sigma|^2)\right) \tag{11}$$

Proof. We compute $-D\cdot R_{\sigma\overline{\sigma}w\overline{w}}=D\cdot\frac{\partial^2 g_{\sigma\overline{\sigma}}}{\partial w\partial\overline{w}}-\frac{\partial g_{\sigma\overline{\sigma}}}{\partial w}g_{w\overline{w}}\overline{\frac{\partial g_{\sigma\overline{\sigma}}}{\partial w}}+\frac{\partial g_{\sigma\overline{w}}}{\partial w}g_{w\overline{\sigma}}\overline{\frac{\partial g_{\sigma\overline{\sigma}}}{\partial w}}+\frac{\partial g_{\sigma\overline{\sigma}}}{\partial w}g_{\sigma\overline{w}}\overline{\frac{\partial g_{w\overline{\sigma}}}{\partial w}}-\frac{\partial g_{\sigma\overline{w}}}{\partial w}g_{\sigma\overline{\sigma}}\overline{\frac{\partial g_{w\overline{\sigma}}}{\partial w}}$. It follows immediately that all summands are of the class $O\left(1/|\sigma|\log^{4+\alpha}(1/|\sigma|^2)\right)$. □

The remaining estimates are shown in several cycles.

STEP 1. *The following estimates hold for the components of the curvature tensor:*

$$R_{\sigma\overline{\sigma}\sigma\overline{\sigma}}=O\left(1/|\sigma|^4\log^2(1/|\sigma|^2)\right) \tag{12}$$

$$R_{\sigma\overline{\sigma}\sigma\overline{w}}=O\left(1/|\sigma|^3\log^2(1/|\sigma|^2)\right) \tag{13}$$

$$R_{\sigma\overline{w}w\overline{w}}=O(1/|\sigma|\log^{1+\alpha}(1/|\sigma|^2)) \tag{14}$$

$$R_{w\overline{w}w\overline{w}}=O(1) \tag{15}$$

Proof. For (12) we estimate $\frac{\partial^2 g_{\sigma\overline{\sigma}}}{\partial\sigma\partial\overline{\sigma}}$, $\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma}g^{\overline{\sigma}\sigma}\overline{\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma}}=O\left(1/|\sigma|^4\log^2(1/|\sigma|^2)\right)$, $\frac{\partial g_{\sigma\overline{w}}}{\partial\sigma}g^{\overline{w}\sigma}\overline{\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma}}$, $\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma}g^{\overline{\sigma}w}\overline{\frac{\partial g_{w\overline{\sigma}}}{\partial\sigma}}=O\left(1/|\sigma|^4\log^{2+2\alpha}(1/|\sigma|^2)\right)$, $\frac{\partial g_{\sigma\overline{w}}}{\partial\sigma}g^{\overline{w}w}\overline{\frac{\partial g_{w\overline{\sigma}}}{\partial\sigma}}=O\left(1/|\sigma|^4\log^{1+\alpha}(1/|\sigma|^2)\right)$.

We consider (13): $\frac{\partial^2 g_{\sigma\overline{\sigma}}}{\partial\sigma\partial\overline{w}}$, $\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma}g^{\overline{\sigma}\sigma}\overline{\frac{\partial g_{\sigma\overline{\sigma}}}{\partial w}}=O(1/|\sigma|^3\log^2(1/|\sigma|^2))$, $\frac{\partial g_{\sigma\overline{w}}}{\partial\sigma}g^{\overline{w}\sigma}\overline{\frac{\partial g_{\sigma\overline{\sigma}}}{\partial w}}$, $\frac{\partial g_{\sigma\overline{\sigma}}}{\partial\sigma}g^{\overline{\sigma}w}\overline{\frac{\partial g_{w\overline{\sigma}}}{\partial w}}$, $\frac{\partial g_{\sigma\overline{w}}}{\partial\sigma}g^{\overline{w}w}\overline{\frac{\partial g_{w\overline{\sigma}}}{\partial w}}=O(1/|\sigma|^3\log^{2+2\alpha}(1/|\sigma|^2))$

In the same way we arrive at the estimates (14), (15). □

Some of these estimates need to be improved in a second step.

STEP 2.

$$R_{\sigma\overline{\sigma}\sigma\overline{\sigma}}=O(1/|\sigma|^4\log^{3+\alpha}(1/|\sigma|^2)) \tag{16}$$

$$R_{\sigma\overline{\sigma}\sigma\overline{w}}=O(1/|\sigma|^3\log^{3+\alpha}(1/|\sigma|^2)) \tag{17}$$

Proof. Concerning (16) we consider the equation

$$-g_{\sigma\overline{\sigma}} = R_{\sigma\overline{\sigma}\sigma\overline{\sigma}}g^{\overline{\sigma}\sigma} + R_{\sigma\overline{\sigma}\sigma\overline{w}}g^{\overline{w}\sigma} + R_{\sigma\overline{\sigma}w\overline{\sigma}}g^{\overline{\sigma}w} + R_{\sigma\overline{\sigma}w\overline{w}}g^{\overline{w}w}$$

According to Step 1 and Proposition 1 the term $R_{\sigma\overline{\sigma}\sigma\overline{\sigma}}g^{\overline{\sigma}\sigma}$ can be estimated from above and below by $1/|\sigma|^2$ whereas the remaining terms are at least of the class $O\left(1/|\sigma|^2\log^{1+\alpha}(1/|\sigma|^2)\right)$. This proves (16).
Next

$$-g_{\sigma\overline{w}} = R_{\sigma\overline{w}\sigma\overline{\sigma}}g^{\overline{\sigma}\sigma} + R_{\sigma\overline{w}\sigma\overline{w}}g^{\overline{w}\sigma} + R_{\sigma\overline{w}w\overline{\sigma}}g^{\overline{\sigma}w} + R_{\sigma\overline{w}w\overline{w}}g^{\overline{w}w}$$

is of the class $O\left(1/|\sigma|\log^{1+\alpha}(1/|\sigma|^2)\right)$, and on the right-hand side all terms but the first are a priori at least in $O\left(1/|\sigma|\log^{1+\alpha}(1/|\sigma|^2)\right)$, whereas $R_{\sigma\overline{w}\sigma\overline{\sigma}}g^{\overline{\sigma}\sigma}$ so far is in $O\left(1/|\sigma|\right)$. This shows (17). □
We need to do (16) again.

Step 3.

$$R_{\sigma\overline{\sigma}\sigma\overline{\sigma}} = O\left(1/|\sigma|^4\log^4(1/|\sigma|^2)\right) \tag{18}$$

We consider once again

$$-g^2_{\sigma\overline{\sigma}} = R_{\sigma\overline{\sigma}\sigma\overline{\sigma}}g^{\overline{\sigma}\sigma}g_{\sigma\overline{\sigma}} + R_{\sigma\overline{\sigma}\sigma\overline{w}}g^{\overline{w}\sigma}g_{\sigma\overline{\sigma}} + R_{\sigma\overline{\sigma}w\overline{\sigma}}g^{\overline{\sigma}w}g_{\sigma\overline{\sigma}} + R_{\sigma\overline{\sigma}w\overline{w}}g^{\overline{w}w}g_{\sigma\overline{\sigma}}.$$

The last three terms on the right-hand side are at least in $O\left(1/|\sigma|^4\log^{4+\alpha}(1/|\sigma|^2)\right)$, whereas $g^2_{\sigma\overline{\sigma}}$ is in $O\left(1/|\sigma|^4\log^4(1/|\sigma|^2)\right)$. This shows (18), and moreover $-R_{\sigma\overline{\sigma}\sigma\overline{\sigma}} \sim g^2_{\sigma\overline{\sigma}}$. □

Let us conclude with a refinement of (15)

Step 4.

$$-R_{w\overline{w}w\overline{w}} = g^2_{w\overline{w}}\left(1 + O\left(1/\log^\alpha(1/|\sigma|^2)\right)\right)$$

Proof. We regard $-g_{w\overline{w}} = R_{w\overline{w}w\overline{w}}g^{\overline{w}w} + R_{w\overline{\sigma}w\overline{w}}g^{\overline{\sigma}w} + R_{w\overline{w}\sigma\overline{w}}g^{\overline{w}\sigma} + R_{w\overline{\sigma}\sigma\overline{w}}g^{\overline{\sigma}\sigma}$. The first summand is in $O(1)$, whereas the remaining three terms are at least in $O\left(1/\log^\alpha(1/|\sigma|^2)\right)$. □

We summarize our estimates in the following way.

Proposition 2. *In a neighborhood of the divisor at infinity we have*

$$-R_{\sigma\overline{\sigma}\sigma\overline{\sigma}} = g^2_{\sigma\overline{\sigma}}(1 + O\left(1/\log^\alpha(1/|\sigma|^2)\right)) \tag{19}$$

$$-R_{w\overline{w}w\overline{w}} = g^2_{w\overline{w}}(1 + O\left(1/\log^\alpha(1/|\sigma|^2)\right)) \tag{20}$$

$$R_{\sigma\overline{\sigma}\sigma\overline{w}} = O\left(1/|\sigma|^3\log^{3+\alpha}(1/|\sigma|^2)\right) \tag{21}$$

$$R_{\sigma\overline{\sigma}w\overline{w}} = O\left(1/|\sigma|^2\log^{2+\alpha}(1/|\sigma|^2)\right) \tag{22}$$

$$R_{\sigma\overline{w}\sigma\overline{w}} = O\left(1|\sigma|^2\log^{2+\alpha}(1/|\sigma|^2)\right) \tag{23}$$

$$R_{\sigma\overline{w}w\overline{w}} = O\left(1/|\sigma|\log^{1+\alpha}(1/|\sigma|^2)\right) \tag{24}$$

Proof of Theorem 1. The above Proposition 2 implies that the curvature tensor is dominated by the diagonal terms (19) and (20). We determine a domain of negative holomorphic sectional curvature. Let

$$t = a \cdot |\sigma| \log(1/|\sigma|^2) \frac{\partial}{\partial\sigma} + b \cdot \frac{\partial}{\partial w}$$

with $a, b \in \mathbb{C}$. Then

$$\begin{aligned} \|t\|^2 &= g_{\sigma\overline{\sigma}} |a|^2 \log^2(1/|\sigma|^2) + (g_{\sigma\overline{w}} a\overline{b} + g_{w\overline{\sigma}} b\overline{a})|\sigma| \log(1/|\sigma|^2) + g_{w\overline{w}}|b|^2 \\ &= 2|a|^2 + g_{w\overline{w}}|b|^2 + (|a|^2 + |b|^2) \cdot O\left(1/\log^\alpha(1/|\sigma|^2)\right). \end{aligned}$$

According to Proposition 2 we have

$$-R(t, \overline{t}, t, \overline{t}) = 4|a|^4 + |b|^4 + (|a|^2 + |b|^2)^2 \cdot O\left(1/\log^\alpha(1/|\sigma|^2)\right)$$

Now we pick a small upper bound for $\|\sigma\|^2$ which yields a negative upper bound for the holomorphic sectional curvature. The sectional and holomorphic bisectional curvatures are dealt with in a similar way. □

References

[1] Cheng, S.-Y, Yau, S.T.: On the existence of a complete Kähler metric in non-compact complex manifolds and the regularity of Fefferman's equation, Comm. Pure Appl. Math. 33, 507–544 (1980)

[2] Kobayashi, R.: Einstein-Kähler metrics on open algebraic surfaces of general type, Tohoku Math. J. 37, 43–77 (1985)

[3] Kobayashi, R.: Kähler-Einstein metric on an open algebraic manifold, Osaka J. Math. 21, 399–418 (1984)

[4] Mok, N. and Yau S.-T.: Completeness of the Kähler-Einstein metric on bounded domains and characterization of domains of holomorphy by curvature conditions. In: The Mathematical Heritage of Henri Poincaré. Proc. Symp. Pure Math. 39 (Part I), 41-60 (1983),

[5] Schumacher, G.: Asymptotics of Kähler-Einstein metrics on quasi-projective manifolds and an extension theorem on holomorphic maps, to appear 1998 in Math. Ann.

[6] Tsuji, H.: An inequality of Chern numbers for open algebraic varieties. Math. Ann. 277 , 483–487 (1987)

[7] Tian, G., Yau, S.-T.: Existence of Kähler-Einstein metrics on complete Kähler manifolds and their applications to algebraic geometry. Mathematical aspects of string theory (San Diego, Calif., 1986), 574–628, Adv. Ser. Math. Phys., 1, World Sci. Publishing, Singapore (1987)

Georg Schumacher
Fachbereich Mathematik
Philipps-Universität
D-35032 Marburg
Germany
schumac@mathematik.uni-marburg.de