Pacific Journal of Mathematics

Volume 327 No. 2 December 2023

PACIFIC JOURNAL OF MATHEMATICS

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

msp.org/pjm

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

The subscription price for 2023 is US $605/year for the electronic version, and $820/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 1945-5844 electronic, 0030-8730 printed) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 327, No. 2, 2023
https://doi.org/10.2140/pjm.2023.327.183

CO-HOPFIAN AND BOUNDEDLY ENDO-RIGID MIXED ABELIAN GROUPS

MOHSEN ASGHARZADEH, MOHAMMAD GOLSHANI AND SAHARON SHELAH

For a given cardinal λ and a torsion abelian group K of cardinality less than λ, we present, under some mild conditions (for example, $\lambda = \lambda^{\aleph_0}$), boundedly endo-rigid abelian group G of cardinality λ with $\operatorname{tor}(G) = K$. Essentially, we give a complete characterization of such pairs (K, λ). Among other things, we use a twofold version of the black box. We present an application of the construction of boundedly endo-rigid abelian groups. Namely, we turn to the existence problem of co-Hopfian abelian groups of a given size, and present some new classes of them, mainly in the case of mixed abelian groups. In particular, we give useful criteria to detect when a boundedly endo-rigid abelian group is co-Hopfian and completely determine cardinals $\lambda > 2^{\aleph_0}$ for which there is a co-Hopfian abelian group of size λ.

1. Introduction

By a torsion (resp. torsion-free) group we mean an abelian group such that all its nonzero elements are of finite (resp. infinite) order. A mixed group G contains

Golshani's research has been supported by a grant from IPM (no. 1401030417).

Shelah would like to thank the Israel Science Foundation (ISF) for partially supporting this research (grant 1838/19). His research partially supported by the grant "Independent Theories" NSF-BSF NSF 2051825 (BSF 3013005232). Shelah is grateful to an individual who prefers to remain anonymous for providing typing services that were used during the work on the paper. This is publication 1232 of Shelah.

MSC2020: 03E75, 16S50, 20K30.

Keywords: black boxes, bounded endomorphisms, co-Hopfian groups, endomorphism algebras, mixed abelian groups, p-groups, set theoretical methods in algebra.

both nonzero elements of finite order and elements of infinite order, and these are connected via the celebrated short exact sequence

$$0 \to \mathrm{tor}(G) \to G \to \frac{G}{\mathrm{tor}(G)} \to 0. \tag{$*$}$$

Despite the importance of $(*)$, there are series of questions concerning how to glue the issues from torsion and torsion-free parts and put them together to check the desired properties for mixed groups.

Reinhold Baer [2; 3] was interested to find an interplay between abelian groups and rings. In this regard, he raised the following general problem:

Problem 1.1. Which ring can be the endomorphism ring of a given abelian group G?

There are a lot of interesting research papers and books that study this problem, see, for example, [11; 16]. According to Fuchs [15], for mixed groups, only very little can be said. As an achievement, we cite the works of Corner and Göbel [8] and Franzen and Goldsmith [12].

For any group G, by $E_f(G)$ we mean the ideal of $\mathrm{End}(G)$ consisting of all elements of $\mathrm{End}(G)$ whose image is finitely generated. Corner [7] has constructed an abelian group $G := (M, +)$, for some ring R and an R-module M, such that any of its endomorphisms is of the form multiplication by some $r \in R$ plus a distinguished function from $E_f(G)$. One can allow such a distinguished function ranges over other classes such as finite-range, countable-range, inessential range or even small homomorphism, and there are a lot of work trying to clarify such situations. As a short list, we may mention Corner and Göbel [8], Dugas and Göbel [10], Corner [7], Thomé [30] and Pierce [21].

Here, by a bounded group, we mean a group G such that $nG = 0$ for some fixed $0 < n \in \mathbb{N}$. By a theorem of Baer and Prüfer, a bounded group is a direct sum of cyclic groups. The converse is not true. However, there is a partial converse for countable p-groups. For more details, see Fuchs [15]. A homomorphism $h \in G_1 \to G_2$ of abelian groups is called bounded if $\mathrm{Rang}(h)$ is bounded.

Definition 1.2. An abelian group G is *boundedly rigid* when every endomorphism of it has the form $\mu_n + h$, where μ_n is multiplication by $n \in \mathbb{Z}$ and h has bounded range. By $\mathrm{E_b}(G)$ we mean the ideal of $\mathrm{End}(G)$ consisting of all elements of $\mathrm{End}(G)$ whose image is bounded.

Let us explain some motivation. The concept of a rigid system of torsion-free groups has a natural analogue for the class of separable p-primary groups: a family $\{G_i : i \in I\}$ of separable p-primary groups is called rigid-like if for all $i \neq j \in I$ every homomorphism $G_i \to G_j$ is small, and also for all $i \in I$, every endomorphism of G_i is the sum of a small endomorphism and multiplication by a p-adic integer. Shelah [23] confirmed a conjecture of Pierce [21] by showing that if μ is an uncountable strong limit cardinal, then there is a rigid-like system $\{G_i : i \in I\}$ of separable p-primary groups such that $|G_i| = \mu$ and $|I| = 2^{\mu}$, see also [25] for more results in this direction.

Let us now state our main results. Section 2 contains the preliminaries, basic definitions and notations that we need. The reader may skip it, and come back to it when needed later. In Section 3, and as a main result, we prove the following.

Theorem 1.3. *Given a cardinal λ such that $\lambda = \lambda^{\aleph_0} > 2^{\aleph_0}$ and a torsion group K of cardinality less than λ, there is a boundedly rigid abelian group G of cardinality λ with* $\mathrm{tor}(G) = K$.

To prove this, we introduce a series of definitions and present several claims. The first one is the rigidity context, denoted by $\boldsymbol{k}$, see Definition 3.1. Also, the main technical tool is a variation of "Shelah's black box", and we refer to it as *twofold black box*. For its definition (resp. its existence), see Definition 3.13 (resp. Lemma 3.15). It may be worth to mention that the black boxes were introduced by Shelah in [26], where he showed that they follow from ZFC (here, ZFC means the Zermelo–Fraenkel set theory with the axiom of choice). We can consider black boxes as general methods to generate a class of diamond-like principles provable in ZFC. Then, we continue by introducing the approximation blocks, denoted by AP, see Definition 3.18 for more precise definition. There is a distinguished object $\boldsymbol{c}$ in AP that we call it full. The twofold black box helps us to find such distinguished objects, see Lemma 3.30. Here, one may define the group $G := G_{\boldsymbol{c}}$. Let $h \in \mathrm{End}(G)$. In order to show that h is boundedly rigid, we apply a couple of reductions (see Lemmas 3.35–3.43), to reduce to the case that h factors throughout $G \to \mathrm{tor}(G)$. Finally, in Lemma 3.31 we handle this case, by showing that any map $G \to \mathrm{tor}(G)$ is indeed boundedly rigid.

In the course of the proof of Theorem 1.3, we develop a general method which allows us to prove that $0 \to \mathbb{Z} \to \mathrm{End}(G) \to \frac{\mathrm{End}(G)}{\mathrm{E_b}(G)} \to 0$ is exact, and also enables us to present a connection to Problem 1.1. In order to display the connection, let R be a ring coming from the rigidity context. For the propose of the introduction, we may assume that $(R, +)$ is cotorsion-free, see Definition 2.8 (with the convenience that the argument becomes easier if we work with $R := \mathbb{Z}$, or even $(R, +)$ is $\aleph_1$-free). Following our construction, every endomorphism of G has the form $\mu_r + h$, where μ_r is a multiplication by $r \in R$ and h has bounded range, i.e., the sequence

$$0 \to R \to \mathrm{End}(G) \to \frac{\mathrm{End}(G)}{\mathrm{E_b}(G)} \to 0$$

is exact.

Definition 1.4. A group G is called *Hopfian* (resp. *co-Hopfian*) if its surjective (resp. injective) endomorphisms are automorphisms.

Essentially, we give complete characterization of the pairs (K, λ) by relating our work with the recent works of Paolini and Shelah, see [19; 20]. To this end, first we recall the following folklore problem:

Problem 1.5. Construct co-Hopfian groups of a given size.

Baer [4] was the first to investigate Problem 1.5 for abelian groups. A torsion-free abelian group is co-Hopfian if and only if it is divisible of finite rank, and hence the problem naturally reduces to the torsion and mixed cases. Beaumont and Pierce [5] proved that if G is co-Hopfian, then $\mathrm{tor}(G)$ is of size at most continuum, and further that G cannot be a p-groups of size $\aleph_0$. This naturally left open the problem of the existence of co-Hopfian p-groups of uncountable size $\leq 2^{\aleph_0}$, which was later solved by Crawley [9] who proved that there exist co-Hopfian p-groups of size $2^{\aleph_0}$. Braun and Strüngmann [6] showed that the existence of three types of infinite abelian p-groups of size $\aleph_0 < |G| < 2^{\aleph_0}$ are independent of ZFC:

(a) Both Hopfian and co-Hopfian.

(b) Hopfian but not co-Hopfian.

(c) Co-Hopfian but not Hopfian.

Also, they proved that the above three types of groups of size $2^{\aleph_0}$ exist in ZFC. So, in light of Theorem 1.3, the remaining part is $2^{\aleph_0} < \lambda < \lambda^{\aleph_0}$. Very recently, and among other things, Paolini and Shelah [19] proved that there is no co-Hopfian group of size λ for such a λ. As an application, in Section 4, we determine cardinals $\lambda > 2^{\aleph_0}$ for which there is a co-Hopfian group of size λ. For the precise statement, see Corollary 4.13.

Let us recall a connection between the concepts boundedly endo-rigid groups and (co-)Hopfian groups. First, recall from the seminal paper [22], for any λ less than the first beautiful cardinal, Shelah proved that there is an endo-rigid torsion-free group of cardinality λ. By definition, for any $f \in \mathrm{End}(G)$ there is $m_f \in \mathbb{Z}$ such that $f(x) = m_f x$. So, f is onto if and only if $m_f = \pm 1$. In other words, G is Hopfian. This naturally motives us to detect co-Hopfian property by the help of some boundedly endo-rigid groups. This is what we want to do in Section 4. Namely, our first result on co-Hopfian groups is stated as follows.

Construction 1.6. Let $K := \oplus\left\{\frac{\mathbb{Z}}{p^n\mathbb{Z}} : p \in \mathbb{P} \text{ and } 1 \leq n < m\right\}$, where $m < \omega$, and $\mathbb{P}$ is the set of prime numbers. Let G be a boundedly endo-rigid abelian group such that $\mathrm{tor}(G) = K$. Then G is co-Hopfian.

We may recall from Theorem 1.3 that such a group exists for any $\lambda = \lambda^{\aleph_0} > 2^{\aleph_0}$. In fact, the size of G is λ.

Let h be a natural number. One of the tools that we use is the h-power torsion subgroup of G:

$$\Gamma_h(G) := \{g \in G : \exists n \in \mathbb{N} \text{ such that } h^n g = 0\}.$$

The assignment $G \mapsto \Gamma_h(G)$ defines a functor from the category of abelian groups to itself. It may be worth to mention that, in the style of Grothendieck, this is called section functor and some authors use $\mathrm{Tor}_h(-)$ to denote it.

In our study of the co-Hopfian property of G, the following subset of prime numbers appears:

$$S_G := \{p \in \mathbb{P} : G/\,\Gamma_p(G) \text{ is not } p\text{-divisible}\}.$$

The set S_G helps us to present a useful criterion to detect when a boundedly endo-rigid abelian group is co-Hopfian:

Proposition 1.7. *Assume $\lambda > 2^{\aleph_0}$ and G is a boundedly endo-rigid abelian group of size λ. Then G is co-Hopfian if and only if*:

(a) *S_G is a nonempty set of primes.*

(b) (b_1) $\Gamma_p(G) \neq G$.
  (b_2) *If $p \in S_G$, then $\Gamma_p(G)$ is not bounded.*
  (b_3) *If $\Gamma_p(G)$ is bounded, then it is finite.*

Let G be an abelian group. In order to show that G is (not) co-Hopfian, and also to see a connection to bounded morphisms, we introduce a useful set $\mathrm{NQr}_{(\mathrm{m,n})}(\mathrm{G})$ consisting of those bounded $h \in \mathrm{End}(\Gamma_n(G))$ such that

(1) $h' := m \cdot \mathrm{id}_{\Gamma_n(G)} + h \in \mathrm{End}(\Gamma_n(G))$ is 1-to-1,

(2) h' is not onto or $m > 1$ and $G/\,\Gamma_n(G)$ is not m-divisible.

In a series of nontrivial cases we check $\mathrm{NQr}_{(\mathrm{m,n})}(\mathrm{G})$ and its negation. This enables us to present some new classes of co-Hopfian and non-co-Hopfian groups (see below, items 4.4–4.11).

See Eklof and Mekler [11] and Göbel and Trlifaj [16] for all unexplained definitions from set theoretic algebra. Also, for unexplained definitions from the group theory, see the books of Fuchs [13; 14; 15].

2. Preliminaries

In this paper all groups are abelian, otherwise specialized. In this section we recall some basic definitions and facts that will be used in the later sections of the paper.

Definition 2.1. An abelian group G is called $\aleph_1$-free if every countable subgroup of G is free. More generally, an abelian group G is called λ-free if every subgroup of G of cardinality $< \lambda$ is free.

Definition 2.2. Let κ be a regular cardinal. An abelian group G is said to be strongly κ-free if there is a set $\mathcal{S}$ of $< \kappa$-generated free subgroups of G containing 0 such that for any subset S of G of cardinality $< \kappa$ and any $N \in \mathcal{S}$, there is an $L \in \mathcal{S}$ such that $S \cup N \subset L$ and L/N is free.

A group G is pure in an abelian group H if $G \subseteq H$ and $nG = nH \cap G$ for every $n \in \mathbb{Z}$. The common notation for this notion is $G \subseteq_* H$.

Fact 2.3. *Suppose G is a torsion-free group. Then the intersection of pure subgroups of G is again pure. In particular, for every $S \subset G$, there exists a minimal pure subgroup of G containing S. The common notation for this subgroup is $\langle S \rangle_G^*$.*

Fact 2.4 (see [17, Theorem 7]). *Let G be an abelian group and H a pure and bounded subgroup of G. Then H is a direct summand of G.*

The notation $\mathrm{tor}(G)$ stands for the full torsion subgroup of G. There is a natural connection with the functor $\mathrm{Tor}_1^{\mathbb{Z}}(-, \sim)$:

$$\mathrm{tor}(G) = \mathrm{Tor}_1^{\mathbb{Z}}(\mathbb{Q}/\mathbb{Z}, G).$$

Fact 2.5 (see [17, Theorem 8]). *Let G be an abelian group and $T \subseteq_* \mathrm{tor}(G)$. If T is the direct sum of a divisible group and a group of bounded exponent, then T is a direct summand of G. The same result holds if $T \subseteq_* G$.*

Fact 2.6 (see [5]). (i) *Let G be a countable p-group. Then G is co-Hopfian if and only if G is finite.*

(ii) *If a group G is co-Hopfian, then* $\mathrm{tor}(G)$ *is of size at most continuum, and further that G cannot be a p-groups of size $\aleph_0$.*

Fact 2.7 (see [13, Theorem 17.2]). *If G is a p-group of bounded exponent, then G is a direct sum of (finitely many, up to isomorphism) finite cyclic groups.*

Definition 2.8. (i) An abelian group G is called *cotorsion* if $\mathrm{Ext}(J, G) = 0$ for all torsion-free abelian groups J.

(ii) An abelian group G is called *cotorsion-free* if it has no nonzero co-torsion subgroup.

In other words, G is cotorsion provided that it is a direct summand of every abelian group H containing G with the property that H/G is torsion-free. Here, we recall a useful source to produce a cotorsion-free group:

Fact 2.9 (see [11, Corollary 2.10(ii)]). *Any $\aleph_1$-free group is cotorsion-free.*

The p-torsion parts of a group G are important sources to produce pure subgroups.

Notation 2.10. Let $\mathbb{P}$ denote the set of all prime numbers.

(i) Let $p \in \mathbb{P}$. The p-power torsion subgroup of G is

$$\Gamma_p(G) := \{g \in G : \exists n \in \mathbb{N} \text{ such that } p^n g = 0\}.$$

(ii) For each $1 \leq m < \omega$, we let $\Gamma_m(G) := \bigoplus\{\Gamma_p(G) : p \mid m\}$.

Recall that the assignment $G \mapsto \Gamma_h(G)$ defines a functor from the category of abelian groups to itself, which is also called section functor. It has the following important property. Suppose $f : G \to H$ is a homomorphism of abelian groups. Then the following diagram of natural short exact sequences is commutative:

$$\begin{tikzcd}
0 \arrow[r] & \Gamma_h(H) \arrow[r, "\subseteq"] & H \arrow[r] & H/\,\Gamma_h(H) \arrow[r] & 0 \\
0 \arrow[r] & \Gamma_h(G) \arrow[r, "\subseteq"] \arrow[u, "f\restriction"] & G \arrow[r] \arrow[u, "f"] & G/\,\Gamma_h(H) \arrow[r] \arrow[u, "\bar{f}"] & 0
\end{tikzcd}$$

where $\bar{f}(g + \Gamma_h(G)) := f(g) + \Gamma_h(H)$.

The connection from p-power torsion functors and the classical torsion functor is read as

$$\operatorname{Tor}_1^{\mathbb{Z}}(\mathbb{Q}/\mathbb{Z}, G) = \operatorname{tor}(G) = \bigoplus_{p\in\mathbb{P}} \Gamma_p(G).$$

Notation 2.11. In this paper, by $\operatorname{End}(-)$ we mean $\operatorname{End}_{\mathbb{Z}}(-)$ where $(-)$ is at least an abelian group, otherwise we specify it.

The following notion of boundedness plays an important role in establishing the main theorems.

Definition 2.12. Let G be an abelian group of size λ. We say G is *boundedly endo-rigid* when for every $f \in \operatorname{End}(G)$ there is $m \in \mathbb{Z}$ such that the map $x \mapsto f(x) - mx$ has bounded range.

The next fact follows from the definition.

Fact 2.13. *An abelian group G is boundedly endo-rigid if and only if for every $f \in \operatorname{End}(G)$ there is $m \in \mathbb{Z}$ and bounded $h \in \operatorname{End}(G)$ such that $f(x) = mx + h(x)$.*

Fact 2.14. *Let K be a bounded torsion abelian group and let $G \subseteq_* H$. There is $h \in \operatorname{Hom}(H, K)$ extending g if $g \in \operatorname{Hom}(G, K)$. This property is conveniently summarized by the subjoined diagram*:

$$\begin{tikzcd}
0 \arrow[r] & G \arrow[r, "\subseteq_*"] \arrow[d, "g"'] & H \arrow[dl, "\exists h"] \\
 & K &
\end{tikzcd}$$

Fact 2.15. *Let G be abelian group and suppose that G is not bounded, then the bounded endomorphisms of G (i.e., those $f \in \operatorname{End}(G)$ with bounded range) form an ideal of the ring* $\operatorname{End}(G)$, *we denote this ideal by* $\mathrm{E_b}(G)$. *With respect to this terminology, G is boundedly rigid if and only if the quotient ring* $\operatorname{End}(G)/\,\mathrm{E_b}(G) \cong \mathbb{Z}$.

Remark 2.16. Recall that torsion subgroups are pure. Let f be a bounded endomorphism of $\mathrm{tor}(G)$. By Fact 2.14, we have

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{tor}(G) \arrow[r, "\subseteq_*"] \arrow[d, "f"'] & G \arrow[dl, "\exists h"] \\ & \mathrm{tor}(G) & \end{tikzcd}$$

Let $\hat{f} : G \xrightarrow{h} \mathrm{tor}(G) \xrightarrow{\subseteq} G$. In sum, f extends to an endomorphisms $\hat{f}$ of G with the same range:

$$\begin{tikzcd} \mathrm{tor}(G) \arrow[r, "f"] \arrow[d, "\subseteq"'] & \mathrm{tor}(G) \arrow[d, "\subseteq"] \\ G \arrow[r, "\hat{f}"'] & G \end{tikzcd}$$

Hence, the notion of boundedly rigid is really the right notion of endo-rigidity for mixed groups (for G torsion-free abelian group, we say that G is endo-rigid when $\mathrm{End}(G) \cong \mathbb{Z}$). For instance, we look at

$$K = \bigoplus\left\{ \frac{\mathbb{Z}}{p^{\ell+1}\mathbb{Z}} : \ell < m \right\}$$

for some $m < \omega$, and recall that this has many bounded endomorphisms. The same will happen for any G extending it.

In what follows we will use the concept of reduced group several times. Let us recall its definition.

Definition 2.17. Let G be an abelian group.

(a) G is called reduced if it contains no divisible subgroup other than 0.

(b) G is called injective if for any inclusion $G_1 \subseteq G_2$ of abelian groups, any morphism $f : G_1 \to G$ can be extended into G_2:

$$\begin{tikzcd} 0 \arrow[r] & G_1 \arrow[r, "\subseteq"] \arrow[d, "f"'] & G_2 \arrow[dl, "\exists h"] \\ & G & \end{tikzcd}$$

Fact 2.18 (see [15]). *An abelian group G is divisible if and only if it is injective.*

Here, we recall a connection between reduced and co-torsion-free abelian groups.

Fact 2.19 (see [11, Theorem V.2.9]). *An abelian group G is cotorsion-free if and only if it is reduced and torsion-free and does not contain a subgroup isomorphic to $\hat{\mathbb{Z}}_p$ for any prime p.*

Recall that $\hat{\mathbb{Z}}_p$ means completion of $\mathbb{Z}$ in the p-adic topology. Here, we collect more basic facts about injective groups that we need:

Discussion 2.20. Let $p \in \mathbb{P}$ be a prime number.

(i) (See [11, page 11].) By the structure theorem for an injective abelian group I, we mean the following decomposition:

$$I = \bigoplus_{p \in \mathbb{P}} \mathbb{Z}(p^\infty)^{\oplus x_p} \oplus \mathbb{Q}^{\oplus x},$$

where x_p and x are index sets.

(ii) (See [18, Theorem 3.7].) Let $p, q \in \mathbb{P}_0 := \mathbb{P} \cup \{0\}$ and set $\mathbb{Z}(0^\infty) := \mathbb{Q}$. Then

$$\operatorname{Hom}(\mathbb{Z}(p^\infty), \mathbb{Z}(q^\infty)) = \begin{cases} \hat{\mathbb{Z}}_p & \text{if } p = q, \\ 0, & \text{otherwise,} \end{cases}$$

with the convenience that $\hat{\mathbb{Z}}_0 = \mathbb{Q}$.

(iii) Combining (i) and (ii) we get the following well-known formula:

$$\operatorname{End}(I) = \prod_{p \in \mathbb{P}_0} \hat{\mathbb{Z}}_p^{\oplus x_p},$$

where $x_0 := x$.

3. The ZFC construction of boundedly rigid mixed groups

In this section we show that for any cardinal $\lambda = \lambda^{\aleph_0} > 2^{\aleph_0}$ and any torsion abelian group K of size less than λ, there exists a boundedly rigid abelian group G with $\operatorname{tor}(G) = K$, see Theorem 3.11.

To this end, we define the notion of rigidity context $\boldsymbol{k}$ which in particular codes a torsion group K, and assign to it a collection of objects $\boldsymbol{m}$, which among other things have a group G with $\operatorname{tor}(G) = K$. We show that under the above assumptions on λ and K, we can always find such an $\boldsymbol{m}$ that the associated group G is boundedly rigid.

Definition 3.1. (1) We say a tuple $\boldsymbol{k}$ is a *rigidity context* when

$$\boldsymbol{k} = (K_{\boldsymbol{k}}, R_{\boldsymbol{k}}, \phi_r^{\boldsymbol{k}}, \Psi_{r,s}^{\boldsymbol{k}}, \Psi_{(r,s)}^{\boldsymbol{k}}, S_{\boldsymbol{k}})_{r,s \in R_{\boldsymbol{k}}} = (K, R, \phi_r, \Psi_{r,s}, \Psi_{(r,s)}, S)_{r,s \in R},$$

where:

(a) K is a reduced torsion abelian group.

(b) R is a ring.

(c) S is a set of prime numbers, $S_{\boldsymbol{k}}^{\perp} = \mathbb{P} \setminus S$ is its complement, and R is $S_{\boldsymbol{k}}^{\perp}$-divisible. This means that R is divisible for any $p \in S_{\boldsymbol{k}}^{\perp}$.

(d) For $r \in R$, the map $\phi_r \in \operatorname{End}(K)$ has bounded range.

(e) If $r, s \in R$, then $\Psi_{r,s} = \phi_r + \phi_s - \phi_{r+s} \in \mathrm{End}(K)$.

(f) If $r, s \in R$, then $\Psi_{(r,s)} \in \mathrm{End}(K)$ has bounded range and, letting $t = rs$, for $x \in K$ we have
$$\Psi_{(r,s)}(x) = \phi_r(\phi_s(x)) - \phi_t(x).$$

(2) We say $\boldsymbol{k}$ is nontrivial when for some prime $p \in S_{\boldsymbol{k}}$ the p-torsion $\Gamma_p(K)$ is infinite, or the set
$$\{p \in S_{\boldsymbol{k}} : \Gamma_p(K) \neq 0\}$$
is infinite.

(3) By $\mathbb{Z}_{\boldsymbol{k}}$ we mean the subring of $\mathbb{Q}$ generated by $\{1\} \cup \left\{\frac{1}{p} : p \in S_{\boldsymbol{k}}^{\perp}\right\}$.

Observation 3.2. *Suppose $(R_{\boldsymbol{k}}, +)$ is cotorsion-free as an abelian group. Then $S_{\boldsymbol{k}} \neq \varnothing$.*

Proof. Suppose on the way of contradiction that $S_{\boldsymbol{k}} = \varnothing$. In other words, $S_{\boldsymbol{k}}^{\perp}$ is the set of prime numbers. By Definition 3.1(1)(c), R is $S_{\boldsymbol{k}}^{\perp}$-divisible. This means that $\mathbb{Q} \subseteq R_{\boldsymbol{k}}$. It turns out from Fact 2.19 that $(R_{\boldsymbol{k}}, +)$ is not cotorsion-free, a contradiction. $\square$

Definition 3.3. Let $\boldsymbol{k}$ be a rigidity context. By $\boldsymbol{M_k}$ we mean the family of all tuples:
$$\boldsymbol{m} = (\boldsymbol{k_m}, G_{\boldsymbol{m}}, F_r^{\boldsymbol{m}}, F_{r,s}^{\boldsymbol{m}}, F_{(r,s)}^{\boldsymbol{m}})_{r,s \in R_{\boldsymbol{k_m}}} = (\boldsymbol{k}, G, F_r, F_{r,s}, F_{(r,s)})_{r,s \in R_{\boldsymbol{k}}},$$
where:

(a) G is an abelian group.

(b) $\mathrm{tor}(G) = K_{\boldsymbol{k}}$.

(c) For $r \in R_{\boldsymbol{k}}$, F_r is an endomorphism of G extending $\phi_r^{\boldsymbol{k}}$:

$$\begin{array}{ccc} K & \xrightarrow{\phi_r} & K \\ \subseteq\big\downarrow & & \big\downarrow\subseteq \\ G & \xrightarrow[F_r]{} & G \end{array}$$

(d) For $r, s \in R_{\boldsymbol{k}}$, $F_{r,s} \in \mathrm{End}(G)$ extends $\Psi r, s$:

$$\begin{array}{ccc} K & \xrightarrow{\Psi r,s} & K \\ \subseteq\big\downarrow & & \big\downarrow\subseteq \\ G & \xrightarrow[F_{r,s}]{} & G \end{array}$$

and they have the same range $F_{r,s}[G] = \Psi_{r,s}[K]$.

(e) For $r, s \in R_{\boldsymbol{k}}$, $F_{(r,s)} \in \mathrm{End}(G)$ extends $\Psi^{\boldsymbol{k}}_{(r,s)}$:

$$\begin{array}{ccc} K & \xrightarrow{\Psi(r,s)} & K \\ {\scriptstyle\subseteq}\big\downarrow & & \big\downarrow{\scriptstyle\subseteq} \\ G & \xrightarrow[F_{(r,s)}]{} & G \end{array}$$

and thereby they have the same range $F_{(r,s)}[G] = \Psi_{(r,s)}[K]$.

(f) If $r, s, t \in R$ and $t = r + s$, then for $x \in G$,

$$F_{r,s}(x) = F_r(x) + F_s(x) - F_t(x),$$

(g) If $r, s, t \in R$ and $t = rs$, then for $x \in G$,

$$F_{(r,s)}(x) = F_r(F_s(x)) - F_t(x).$$

Definition 3.4. Adopt the previous notation, and let

$$\boldsymbol{M} = \bigcup \{\boldsymbol{M_k} : \boldsymbol{k} \text{ is a rigidity context}\}.$$

(1) We define $\leq_{\boldsymbol{M}}$ as the partial order on $\boldsymbol{M}$. Namely, $\boldsymbol{m} \leq_{\boldsymbol{M}} \boldsymbol{n}$ if and only if
   (a) $\boldsymbol{m}, \boldsymbol{n} \in \boldsymbol{M}$,
   (b) $\boldsymbol{k_m} = \boldsymbol{k_n}$,
   (c) $G_{\boldsymbol{m}} \subseteq G_{\boldsymbol{n}}$,
   (d) $F^{\boldsymbol{m}}_r \subseteq F^{\boldsymbol{n}}_r$.

(2) By $\leq_{\boldsymbol{M_k}}$ we mean $\leq_{\boldsymbol{M}} \restriction \boldsymbol{M_k}$.

Notation 3.5. Let $r \in R$ and $x \in G_{\boldsymbol{m}}$. By rx we mean $rx := F^{\boldsymbol{m}}_r(x) \in G_{\boldsymbol{m}}$.

Definition 3.6. Suppose $\boldsymbol{k}$ is a rigidity context and $\boldsymbol{m} \in \boldsymbol{M_k}$.

(1) We say $\boldsymbol{m}$ is *boundedly rigid* when for every $f \in \mathrm{End}(G_{\boldsymbol{m}})$ there are $r \in R$ and $h \in \mathrm{End}_b(G_{\boldsymbol{m}})$[1] and

$$x \in G_{\boldsymbol{m}} \Longrightarrow f(x) = rx + h(x).$$

(2) We say $\boldsymbol{m}$ is *free* when it has a base B which means that the set $\{x + K_{\boldsymbol{k}} : x \in B\}$ is a free base of the abelian group $G_{\boldsymbol{m}}/K$.

(3) We say $\boldsymbol{m}$ is λ*-free* when $G_{\boldsymbol{m}}/K$ is.

(4) We say $\boldsymbol{m}$ is *strongly* λ*-free* when $G_{\boldsymbol{m}}/K$ is.

(5) Let $M_{\boldsymbol{m}}$ be the R-module obtained by expanding $G_{\boldsymbol{m}}/K$ such that for $x, y \in G_{\boldsymbol{m}}$ and $r \in R$

$$rx + K = y + K \Longleftrightarrow F^{\boldsymbol{m}}_r(x) = y.$$

The next easy lemma shows that $M_{\boldsymbol{m}}$ as defined above is well defined.

[1] so, h has a bounded range.

Lemma 3.7. *Suppose $\boldsymbol{k}$ is a rigidity context and $\boldsymbol{m} \in \boldsymbol{M_k}$. Then $M_{\boldsymbol{m}}$ can be turn to an R-module structure.*

Proof. Since $M_{\boldsymbol{m}}$ is an expansion of $G_{\boldsymbol{m}}/K$, it is an abelian group. Let $r \in R$ and $m := g + K \in M_{\boldsymbol{m}}$ where $g \in G$. The assignment

$$(r, m) \mapsto rm := F_r^{\boldsymbol{m}}(g) + K \in G_{\boldsymbol{m}}/K = M_{\boldsymbol{m}}$$

defines the desired module structure on $M_{\boldsymbol{m}}$. □

Lemma 3.8. *Suppose $\boldsymbol{k}$ is a rigidity context and $\boldsymbol{m} \in \boldsymbol{M_k}$. Then*:

(1) *Suppose $R_{\boldsymbol{k}} = \mathbb{Z}$ (so, $S_{\boldsymbol{k}}^{\perp} = \varnothing$). Then $\boldsymbol{m}$ is boundedly rigid if and only if $G_{\boldsymbol{m}}$ is boundedly rigid.*

(2) *Let $R_{\boldsymbol{k}} = \mathbb{Z}_{\boldsymbol{k}}$ (see* Definition 3.1(3)). *Then $\boldsymbol{m}$ is boundedly rigid if and only if $G_{\boldsymbol{m}}$ is boundedly rigid.*

(3) *If $\phi_r^{\boldsymbol{k}}$ is zero for every $r \in R$, then $G_{\boldsymbol{m}}$ is an R-module.*

Proof. (1) and (2) are trivial and follow from the definitions.

(3) For each $x \in G_{\boldsymbol{m}}$ and $r \in R$, we set $rx := F_r^{\boldsymbol{m}}(x)$. It is straightforward to furnish the following three properties.

- The identity $r(x + y) = rx + ry$ follows from Definition 3.1(1)(c).
- The equality $(r + s)x = rx + sx$ follows from Definition 3.1(1)(d).
- The equality $r(sm) = (rs)m$ follows from (e) and (f) of Definition 3.1(1).

From these, $G_{\boldsymbol{m}}$ is equipped with an R-module structure. □

In what follows, the notation $\lg(-)$ stands for the length function.

Definition 3.9. Let $\alpha \in \mathrm{Ord}$.

(1) By $\Lambda_\omega[\alpha]$ we mean

$$\big\{\eta : \lg(\eta) = \omega \text{ and } \eta(n) = (\eta(n,1), \eta(n,2)) \text{ for } \eta(n,1) \le \eta(n,2) < \eta(n+1,1) < \alpha\big\}.$$

(2) For each $\eta \in \Lambda_\omega[\alpha]$, we let $\boldsymbol{j}(\eta) = \bigcup\{\eta(n, 1) : n < \omega\}$.

(3) $\Lambda_{<\omega}[\alpha] := \{\langle\ \rangle\} \cup \bigcup_{k<\omega} \Lambda_k[\alpha]$, where $\Lambda_k[\alpha]$ is the set of all η furnished with the properties:

(a) $\lg(\eta) = k + 1$.

(b) $\eta(k) < \alpha$.

(c) For any $\ell < k$ we suppose $\eta(\ell)$ is furnished with a pairing property in the sense that:

(i) $\eta(\ell) = (\eta(\ell, 1), \eta(\ell, 2))$, where $\eta(\ell, 1) \le \eta(\ell, 2) < \alpha$.

(ii) Additionally, let $\ell+1 < k$, we may and do assume that $\eta(\ell, 2) < \eta(\ell+1, 1)$.

(d) If $\ell < k$, then $\eta(\ell, 1) = \eta(\ell, 2) \Leftrightarrow \ell = 0$.

(4) $\Lambda[\alpha] := \Lambda_\omega[\alpha] \cup \Lambda_{<\omega}[\alpha]$.

(5) For any $\eta \in \Lambda[\alpha]$ and $k+1 < \lg(\eta)$, we set

(a) $\eta \restriction_L k := \langle(\eta(\ell, 1), \eta(\ell, 2)) : \ell < k\rangle^\frown\langle\eta(k, 1)\rangle$ and

(b) $\eta \restriction_R k := \langle(\eta(\ell, 1), \eta(\ell, 2)) : \ell < k\rangle^\frown\langle\eta(k, 2)\rangle$.

Note that $\eta \restriction_L k$ and $\eta \restriction_R k$ belong to $\Lambda_{k+1}[\alpha]$.

(6) We say $\Lambda \subseteq \Lambda[\alpha]$ is *downward closed* while for each $\eta \in \Lambda$ and $k+1 < \lg(\eta)$ we have $\eta \restriction_L k, \eta \restriction_R k \in \Lambda$.

We next define when a subset of $\Lambda_\omega[\alpha]$ is free.

Definition 3.10. Suppose $\alpha \in \mathrm{Ord}$ and $\Lambda \subseteq \Lambda_\omega[\alpha]$.

(1) We say Λ is free whenever there is a function $h : \lambda \to \omega$ such that the sequence

$$\langle\{\eta\restriction_L n, \eta\restriction_R n : h(\eta) \leq n < \omega\} : \eta \in \Lambda\rangle$$

is a sequence of pairwise disjoint sets.

(2) We say Λ is μ-free when every $\Lambda' \subseteq \Lambda$ of cardinality $< \mu$ is free.

We can now state the main result of this section.

Theorem 3.11. *Let $\lambda = \lambda^{\aleph_0} > 2^{\aleph_0}$. Let $\boldsymbol{k}$ be a nontrivial rigidity context such that $K := K_{\boldsymbol{k}}$ and $R := R_{\boldsymbol{k}}$ are of cardinality $\leq \lambda$. Then there exists an abelian group G such that* $\mathrm{tor}(G) = K$ *and G is boundedly rigid. In particular*, *the sequence*

$$0 \to R \to \mathrm{End}(G) \to \frac{\mathrm{End}(G)}{\mathrm{E_b}(G)} \to 0$$

is exact.

The rest of this section is devoted to the proof of the above theorem.

Definition 3.12. For any ordinal γ, a sequence $\eta \in \Lambda[\lambda]$ and a family $\Lambda \subseteq \Lambda[\lambda]$, we define:

(1) S_γ is the closure of $\omega \cup \gamma$ under finite subsets, so including finite sequences.

(2) $\gamma(\eta) = \eta(0, 1)$.

(3) $\Lambda_\gamma = \{\eta \in \Lambda : \gamma(\eta) < \gamma\}$.

(4) We set $\Lambda_{<\omega} = \Lambda \cap \Lambda_{<\omega}[\alpha]$ and $\Lambda_\omega = \Lambda \cap \Lambda_\omega[\alpha]$.

In order to prove Theorem 3.11, we need a twofold version of the black box, that we now introduce. On simple black boxes, see [24; 27; 28]. The presentation here is a special case of the n-fold λ-black box from [29], when $n = 2$.

Definition 3.13. We say $\boldsymbol{b}$ is a *twofold λ-black box* when it consists of:

(1) $\bar{g} = \langle g_\eta : \eta \in \Lambda_\omega[\lambda]\rangle$, where g_η is a function from ω into S_λ.

(2) Suppose $g : \Lambda_{<\omega}[\lambda] \to S_\lambda$ is a function and $f : \Lambda_{<\omega}[\lambda] \to \gamma$ where $\gamma < \lambda$. Then, for some $\eta \in \Lambda_\omega[\lambda]$,
   (a) $\gamma(\eta) > \gamma$,
   (b) $g_\eta(0) = g(\langle\,\rangle)$,
   (c) $g_\eta(n+1) = \big(g(\eta \restriction_L n), g(\eta \restriction_R n)\big)$,
   (d) $\eta(n, 1) < \eta(n, 2)$ and $f(\eta \restriction_L n) = f(\eta \restriction_R n)$ for all $1 \le n < \omega$.

Hypothesis 3.14. For the rest of this section we adopt the following hypotheses, otherwise specializes:

- $\lambda = \lambda^{\aleph_0} > 2^{\aleph_0}$.
- $\boldsymbol{k}$ is a rigidity context as in Definition 3.1.
- $K = K_{\boldsymbol{k}}$ and $R = R_{\boldsymbol{k}}$ are of cardinality $< \lambda$. Without loss of generality, we may assume that the set of elements of K and R are subsets of λ.
- $(R, +)$ is cotorsion-free.
- $\boldsymbol{b}$ is a twofold λ-black box.

The following result was proved in [29, Lemma 1.14], with a setting more general than here. As this plays a crucial ingredient, we sketch its proof.

Lemma 3.15. *There exists a twofold λ-black box.*

Proof. For notational simplicity, we set $S := S_\lambda$, and look at the fixed partition of λ into λ-many sets, each of cardinality λ:

$$\langle W_{s_1,s_2} : s_1, s_2 \in S\rangle.$$

For each $\eta \in \Lambda_\omega[\lambda]$, we define $g_\eta(n) \in S$, by induction on $n < \omega$.

To start, set

$$(*_1) \qquad g_\eta(0) = s \iff \eta(0, 1) = \eta(0, 2) \in W_{s,s}.$$

Now suppose that $n < \omega$ and $g_\eta \restriction (n+1)$ is defined. We are going to define $g_\eta(n+1)$. It is enough to note that

$$(*_2) \qquad g_\eta(n+1) = (s_1, s_2) \iff \eta(n+1, 1) \in W_{s_1,s_2}.$$

We show that $\bar{g} = \langle g_\eta : \eta \in \Lambda_\omega[\lambda]\rangle$ is as required. Suppose that $g : \Lambda_{<\omega}[\lambda] \to S_\lambda$ is a function and $f : \Lambda_{<\omega}[\lambda] \to \gamma$ where $\gamma < \lambda$. We define $\eta \in \Lambda_\omega[\lambda]$, by defining $\eta(n)$, by induction on n.

Let $\eta(0) := \langle \eta(0, 1), \eta(0, 2)\rangle$, where

$$(*_3) \qquad \gamma < \eta(0, 1) = \eta(0, 2) \in W_{g(\langle\,\rangle),g(\langle\,\rangle)}.$$

Now, suppose that $n < \omega$ and we have defined $\eta \restriction n + 1$. We define

$$\eta(n+1) = \langle \eta(n+1, 1), \eta(n+1, 2)\rangle.$$

Set

(a) $s_1 := g(\eta \restriction_L n)$,

(b) $s_2 := g(\eta \restriction_R n)$, and

(c) $\boldsymbol{c}_n : W_{s_1,s_2} \to \gamma$ is defined via the assignment

$$\boldsymbol{c}_n(\alpha) := f\big((\eta \restriction n+1)^\frown\langle\alpha\rangle\big). \tag{+}$$

As $\gamma < \lambda$ and W_{s_1,s_2} has size λ, we can find an unbounded subset W_n of W_{s_1,s_2} such that $\boldsymbol{c}_n \restriction W_n$ is constant. Let $\eta(n+1,1) < \eta(n+1,2)$ be such that

$$\eta(n,2) < \eta(n+1,1), \quad \eta(n+1,2) \in W_n \subseteq W_{g(\eta\restriction_L n), g(\eta\restriction_R n)}. \tag{$*_4$}$$

We claim that the η we constructed as above, satisfies the required conditions of Definition 3.13(2). Indeed, thanks to our construction, $\gamma(\eta) = \eta(0,1) > \gamma$. We also have

$$g_\eta(0) = g(\langle\,\rangle) \iff \eta(0,1) = \eta(0,2) \in W_{g(\langle\,\rangle), g(\langle\,\rangle)},$$

which is true by $(*_3)$. We also have

$$g_\eta(n+1) = \big(g(\eta \restriction_L n), g(\eta \restriction_R n)\big) \iff \eta(n+1,1) \in W_{g(\eta\restriction_L n), g(\eta\restriction_R n)},$$

which is again true by $(*_4)$. Finally note that, clearly $f(\eta \restriction_L 1) = f(\eta \restriction_R 1)$, and for all n,

$$\begin{aligned} f(\eta \restriction_L n+2) &= f(\eta \restriction n+1^\frown\langle\eta(n+1,1)\rangle) \\ &\overset{(+)}{=} \boldsymbol{c}_n(\eta(n+1,1)) \\ &\overset{(*_4)}{=} \boldsymbol{c}_n(\eta(n+1,2)) \\ &\overset{(+)}{=} f(\eta \restriction n+1^\frown\langle\eta(n+1,2)\rangle) = f(\eta \restriction_R n+2). \end{aligned}$$

The lemma follows. □

Assuming hypotheses beyond ZFC, we can get stronger versions of twofold λ-black box (see again [29]).

Observation 3.16. *Assume $\lambda = \mathrm{cf}(\lambda) \geq \aleph_1$. Let*

$$S \subseteq \{\alpha < \lambda : \mathrm{cf}(\alpha) = \aleph_0\}$$

be a stationary and nonreflecting subset of λ such that the principle $\diamondsuit_S$ holds. Then there is a λ-free twofold λ-black box $\boldsymbol{b}$ such that $\Lambda_{\boldsymbol{b}} = \{\eta_\delta : \delta \in S\}$ and $\boldsymbol{j}(\eta_\delta) = \delta$ for every $\delta \in S$.

Recall that Jensen's diamond principle $\diamondsuit_S$ is a kind of prediction principle whose truth is independent of ZFC. The point in the above proof is that if $\Lambda_{\boldsymbol{b}} = \{\eta_\delta : \delta \in S\}$ and $\boldsymbol{j}(\eta_\delta) = \delta$ for every $\delta \in S$, then as S does not reflect, the set $\Lambda_{\boldsymbol{b}}$ is λ-free.

Remark 3.17. Recall from [6] that a (co-)Hopfian group of size $\lambda = 2^{\aleph_0}$ exists in ZFC. We can also deal with the case of $\lambda = 2^{\aleph_0}$, but all is known in this case, so we just concentrate on the case $\lambda = \lambda^{\aleph_0} > 2^{\aleph_0}$.

Definition 3.18. Let $\mathrm{AP} := \mathrm{AP}_{\boldsymbol{k},\lambda}$ be the set of all quintuples

$$\boldsymbol{c} = (\Lambda_{\boldsymbol{c}}, \boldsymbol{m}_{\boldsymbol{c}}, \Gamma_{\boldsymbol{c}}, X_{\boldsymbol{c}}, \langle a^{\boldsymbol{c}}_{\eta,n} : \eta \in \Lambda_{\boldsymbol{c}}, n < \omega\rangle)$$

such that:

(a) $\Lambda_{\boldsymbol{c}} \subseteq \Lambda[\lambda]$ is downward closed.

(b) $\boldsymbol{m}_{\boldsymbol{c}} \in \boldsymbol{M}_{\boldsymbol{k}}$. We may write $G_{\boldsymbol{c}}$, $M_{\boldsymbol{c}}$ instead of $G_{\boldsymbol{m}_{\boldsymbol{c}}}$, $M_{\boldsymbol{m}_{\boldsymbol{c}}}$ respectively, etc.

(c) $X_{\boldsymbol{c}}$ is the set

$$\{rx_\nu : r \in R, \nu \in \Lambda_{\boldsymbol{c},<\omega}\} \cup \{ry_{\eta,n} : r \in R, \eta \in \Lambda_{\boldsymbol{c},\omega}, n < \omega\}.$$

(d) $G_{\boldsymbol{c}}$ is generated, as an abelian group, by the sets K and $X_{\boldsymbol{c}}$. The relations presented in (f), see below.

(e) For any ordinal α, let $G_{\boldsymbol{c},\alpha}$ be the subgroup of $G_{\boldsymbol{c}}$ generated by the set K and

$$\{rx_\nu : r \in R, \nu \in \Lambda_{\boldsymbol{c},<\omega} \cap \Lambda[\alpha]\} \cup \{ry_{\rho,n} : r \in R, \rho \in \Lambda_{\boldsymbol{c},\omega} \cap \Lambda[\alpha], n < \omega\}.$$

(f) $M_{\boldsymbol{c}}$, as an R-module, is generated by $X_{\boldsymbol{c}} \cup K$, freely except the following set $\Gamma_{\boldsymbol{c}}$ of equations:

$$y_{\eta,n} = a^{\boldsymbol{c}}_{\eta,n} + (n!)\, y_{\eta,n+1} + (x_{\eta\restriction_L n} - x_{\eta\restriction_R n}),$$

where $a^{\boldsymbol{c}}_{\eta,n} \in G_{\boldsymbol{c},\eta(0,1)}$.

The following is clear:

Lemma 3.19. *Suppose $\boldsymbol{c} \in \mathrm{AP}_{\boldsymbol{k},\lambda}$. Then $G_{\boldsymbol{c}}$ is of size $\lambda^{\aleph_0}$.*

Definition 3.20. For any $\boldsymbol{c} \in \mathrm{AP}_{\boldsymbol{k},\lambda}$, we define:

(1) $\gamma_{\boldsymbol{c}} := \min\{\gamma \leq \lambda : \Lambda_{\boldsymbol{c}} \subseteq \Lambda[\gamma]\}$.

(2) Let $\Omega_{\boldsymbol{c}} := \Lambda_{\boldsymbol{c},<\omega} \cup (\Lambda_{\boldsymbol{c},\omega} \times \omega)$ and define $\langle x_\rho : \rho \in \Omega_{\boldsymbol{c}}\rangle$ by the following rules:
  (a) If $\rho \in \Lambda_{\boldsymbol{c},<\omega}$, then x_ρ is defined as in Definition 3.18(c).
  (b) If $\rho = (\eta, n) \in \Lambda_{\boldsymbol{c},\omega} \times \omega$, we define $x_\rho := y_{\eta,n}$.

(3) For $b \in G_{\boldsymbol{c}}$ choose the sequence

$$\langle r_{b,\ell}, \eta_{b,\ell}, m_{b,\ell} : \ell < n_b\rangle$$

such that

$$b - \sum_{\ell < n_b} r_{b,\ell}\, y_{\eta_{b,\ell}, m_{b,\ell}} \in \sum_{\rho \in \Lambda_{\boldsymbol{c},<\omega}} Rx_\rho + K,$$

where $r_{b,\ell} \in R \setminus \{0\}$ and $(\eta_{b,\ell}, m_{b,\ell}) \in \Lambda_{\boldsymbol{c},\omega} \times \omega$.

(4) By $\mathrm{supp}_\circ(b)$ we mean $\{\eta_{b,\ell} : \ell < n_b\}$.

Definition 3.21. Suppose $\boldsymbol{c} \in \mathrm{AP}_{\boldsymbol{k},\lambda}$ and let $a \in G_{\boldsymbol{c}}$.

(a) There is a finite set $\Lambda_a \subseteq \Lambda_{\boldsymbol{c}}$, a sequence $S := \langle r_\rho : \rho \in \Lambda_a \rangle$ of nonzero elements of R, an $n(a) < \omega$ and $d_a \in K$ such that

$$a = \sum_{\eta \in \Lambda_{a,<\omega}} r_\eta\, x_\eta + \sum_{\nu \in \Lambda_{a,\omega}} r_\nu\, y_{\nu,n(a)} + d_a,$$

where $\Lambda_{a,<\omega} = \Lambda_a \cap \Lambda_{\boldsymbol{c},<\omega}$ and $\Lambda_{a,\omega} = \Lambda_a \cap \Lambda_{\boldsymbol{c},\omega}$.

(b) Let $\mathrm{supp}_{\boldsymbol{c}}(a) = \mathrm{supp}(a)$ be the minimal set $\Lambda \subseteq \Lambda_{\boldsymbol{c}}$ with respect to the following two properties:

(i) $\Lambda_a \subseteq \Lambda$.

(ii) If $\nu \in \Lambda_a \cap \Lambda_{\boldsymbol{c},\omega}$ and $n < \omega$, then $\Lambda_{a^{\boldsymbol{c}}_{\nu,n}} \subset \Lambda$ and $\eta{\restriction_L} n, \eta{\restriction_R} n \in \Lambda$.

Remark 3.22. Adopt the previous notation, and $a \in G_{\boldsymbol{c}}$. Then $\mathrm{supp}_{\boldsymbol{c}}(a)$ is the minimal set $\Lambda \subseteq \Lambda_{\boldsymbol{c}}$ such that

$$a \in \langle \{x_\eta, y_{\nu,n} : \eta \in \Lambda(L, R), \nu \in \Lambda, n < \omega\} \cup K \rangle^*_{G_{\boldsymbol{c}}}.$$

Remark 3.23. Adopt the previous notation.

(1) The set $\mathrm{supp}_{\boldsymbol{c}}(a)$ is countable.

(2) If $a = x_\nu$ for some $\nu \in \Lambda_{\boldsymbol{c}}$, then

$$\mathrm{supp}(a) \setminus S_{\eta(\nu,1)} = \{\nu\} \cup \{\nu{\restriction_L}, n, \nu{\restriction_R}, n : n < \omega\}.$$

Definition 3.24. Let $\leq_{\mathrm{AP}}$ be the following partial order on $\mathrm{AP} = \mathrm{AP}_{\boldsymbol{k},\lambda}$. For any $\boldsymbol{c}, \boldsymbol{d} \in \mathrm{AP}$ we say $\boldsymbol{c} \leq_{\mathrm{AP}} \boldsymbol{d}$ when:

(a) $\Lambda_{\boldsymbol{c}} \subseteq \Lambda_{\boldsymbol{d}}$.

(b) $\boldsymbol{m}_{\boldsymbol{c}} \leq_{\boldsymbol{M}} \boldsymbol{m}_{\boldsymbol{d}}$, and hence $G_{\boldsymbol{c}} \subseteq G_{\boldsymbol{d}}$, etc.

(c) $a^{\boldsymbol{c}}_{\eta,\ell} = a^{\boldsymbol{d}}_{\eta,\ell}$ for $\eta \in \Lambda_{\boldsymbol{c}},\ \ell < \omega$.

(d) $x^{\boldsymbol{c}}_\eta = x^{\boldsymbol{d}}_\eta$ for $\eta \in \Lambda_{\boldsymbol{c},<\omega}$.

(e) $y^{\boldsymbol{c}}_{\eta,\ell} = y^{\boldsymbol{d}}_{\eta,\ell}$ for $\eta \in \Lambda_{\boldsymbol{c},\omega}$ and $\ell < \omega$.

Lemma 3.25. (1) *$\leq_{\mathrm{AP}}$ is indeed a partial order,*

(2) *If $\bar{\boldsymbol{c}} = \langle \boldsymbol{c}_\alpha : \alpha < \delta \rangle$ is $\leq_{\mathrm{AP}}$-increasing, then there exists $\boldsymbol{c}_\delta = \bigcup_{\alpha<\delta} \boldsymbol{c}_\alpha$ in* AP *which is the $\leq_{\mathrm{AP}}$-least upper bound of the sequence $\bar{\boldsymbol{c}}$.*

Proof. Clause (1) is clear. For (2), let

$$\boldsymbol{c}_\delta := (\Lambda, \boldsymbol{m}, \Gamma, X, \langle a_{\eta,n} : \eta \in \Lambda, n < \omega \rangle),$$

where $\Lambda := \bigcup_{\alpha<\delta} \Lambda_{\boldsymbol{c}_\alpha}$, $\boldsymbol{m} =: (G, F_r, F_{r,s}, F_{(r,s)})$, with

$$G := \bigcup_{\alpha<\delta} G_{\boldsymbol{c}_\alpha}, \quad F_r := \bigcup_{\alpha<\delta} F^{\boldsymbol{c}_\alpha}_r, \quad F_{r,s} := \bigcup_{\alpha<\delta} F^{\boldsymbol{c}_\alpha}_{r,s}, \quad F_{(r,s)} := \bigcup_{\alpha<\delta} F^{\boldsymbol{c}_\alpha}_{(r,s)},$$

$\Gamma := \bigcup_{\alpha<\delta} \Gamma_{\boldsymbol{c}_\alpha}$, $X := \bigcup_{\alpha<\delta} X_{\boldsymbol{c}_\alpha}$, and for $\eta \in \Lambda_\omega$ and $n < \omega$, we have $a_{\eta,n} = a^{\boldsymbol{c}_\alpha}_{\eta,n}$, for some and hence any $\alpha < \delta$ such that $\eta \in \Lambda_{\boldsymbol{c}_\alpha,\omega}$.

It is easily seen that $\boldsymbol{c}_\delta$ is as required. □

An R-module M is called $\aleph_1$-free if every countably generated submodule of M is contained in a free submodule of M. Similarly, μ-free can be defined. For more details, see [11, Chapter IV, Definition 1.1].

Lemma 3.26. *Let* $\boldsymbol{c} \in \mathrm{AP}$.

(1) $\mathrm{tor}(G_{\boldsymbol{c}}) = K$.

(2) *The group*

$$G_{\boldsymbol{c}}/\langle K \cup \{r x_\nu : r \in R, \nu \in \Lambda_{\boldsymbol{c},<\omega}\}\rangle$$

is divisible and torsion-free. Also, the parallel result holds for the R*-module*:

$$M_{\boldsymbol{c}}/\langle K \cup \{r x_\nu : r \in R, \nu \in \Lambda_{\boldsymbol{c},<\omega}\}\rangle.$$

(3) *The following three properties are satisfied*:

(a) $\Lambda_{\boldsymbol{c}}$ *is* $\aleph_1$*-free.*

(b) *If* $\Lambda_{\boldsymbol{c}}$ *is* μ*-free, then* $M_{\boldsymbol{c}}$ *is* μ*-free.*

(c) *If* $\Lambda_{\boldsymbol{c}}$ *is* μ*-free and* $(R, +)$ *is* μ*-free, then* $G_{\boldsymbol{c}}/K$ *is a* μ*-free abelian group.*

(4) *If* $\gamma \le \gamma_{\boldsymbol{c}}$ *and* $\Lambda \subseteq \Lambda_{\boldsymbol{c}}$, *then there exists a unique* $\boldsymbol{d} \in \mathrm{AP}$ *such that*

$$\Lambda_{\boldsymbol{d}} = \Lambda \cap \Lambda[\gamma] \quad \text{and} \quad G_{\boldsymbol{d}} \subseteq G_{\boldsymbol{c}}.$$

Such a unique object is denoted by $\boldsymbol{d} := \boldsymbol{c} \restriction (\gamma, \Lambda)$.

(5) *Assume* $\eta \in \Lambda_\omega[\lambda] \setminus \Lambda_{\boldsymbol{c}}$, $\ell < \omega$ *and* $a_\ell \in G_{\boldsymbol{c}}$ *are such that* $a_\ell \in G_{\boldsymbol{c},\eta(0,1)}$ *for each* ℓ. *Then there is* $\boldsymbol{d} \in \mathrm{AP}$ *equipped with the following three properties*:

(a) $\Lambda_{\boldsymbol{d}} = \Lambda_{\boldsymbol{c}} \cup \{\eta\} \cup \{\eta\restriction_L n, \eta\restriction_R n : n < \omega\}$.

(b) $\boldsymbol{c} \le_{\mathrm{AP}} \boldsymbol{d}$ *and so* $G_{\boldsymbol{c}} \subseteq G_{\boldsymbol{d}}$.

(c) $a^{\boldsymbol{d}}_{\eta,\ell} = a_\ell$ *for* $\ell < \omega$.

(6) *The group* $G_{\boldsymbol{c}}$ *is of size* λ.

Proof. (1)–(2) These are easy.

(3)(a) Let $\Lambda \subseteq \Lambda_{\boldsymbol{c},\omega}$ be countable, and let $\{\eta_n : n < \omega\}$ be an enumeration of it. Define the maps h_1 and h_2 from Λ to ω as

$$h_1(\eta_n) := \min\big\{k : \forall j < n, \forall \ell, r \in \{L, R\} \text{ we have } \eta_j \restriction_\ell k \ne \eta_n \restriction_r k\big\},$$

$$h_2(\eta_n) := \min\{k : \eta_n\restriction_L k \ne \eta_n \restriction_R k\}.$$

Finally, we set

$$h(\eta_m) := \max\{h_1(\eta_n), h_2(\eta_n)\} + 1.$$

Having Definition 3.10 in mind, we are going to show h is as required. Let $j < i < \omega$ and let

$$h(\eta_j) \leq n_j < \omega \quad \text{and} \quad h(\eta_i) \leq n_i < \omega.$$

We will show that $\eta_j \restriction_\ell n_i \neq \eta_i \restriction_r n_j$, where $\ell, r \in \{L, R\}$. To see this, we note that there is nothing to prove if $n_i \neq n_j$. So, we may and do assume that $n := n_i = n_j$. Thus, $h(\eta_j), h(\eta_i) \leq n$. We look at $m := h_1(\eta_i)$. According to the definition of h_1, we know that $\eta_j \restriction_\ell m \neq \eta_i \restriction_r m$. As $m \leq n$ one has

$$\eta_i \restriction_\ell n \neq \eta_j \restriction_r n.$$

Also given any $i < \omega$, if $n \geq h(\eta_i)$, then by the definition of h_2 and as $n \geq h_2(\eta_i)$, we have

$$\eta_i \restriction_L n \neq \eta_i \restriction_R n.$$

It follows that the sequence

$$\langle \{\eta \restriction_L n, \eta \restriction_R n : h(\eta) \leq n < \omega\} : \eta \in \Lambda \rangle$$

is a sequence of pairwise disjoint sets. By definition, $\Lambda_{\boldsymbol{c}}$ is $\aleph_1$-free.

(3)(b) For simplicity, we present the proof when $\mu := \aleph_1$. Let $X \subseteq M_{\boldsymbol{c}}$ be countable. We are going to show that it is included into a countably generated free R-submodule of $M_{\boldsymbol{c}}$. As X countable, we have

$$\exists \Lambda \subseteq \Lambda_{\boldsymbol{c},\omega} \text{ countable}, \quad \exists \Lambda_* \subseteq \Lambda_{\boldsymbol{c},<\omega} \text{ countable}$$

such that

$$X \subseteq \sum \{R y_{\eta,n} : \eta \in \Lambda \text{ and } n < \omega\} + \sum \{R x_\rho : \rho \in \Lambda_*\}.$$

As $\Lambda_{\boldsymbol{c}}$ is $\aleph_1$-free and Λ is countable, there is a function $h : \Lambda \to \omega$ such that

$$\langle \{\eta \restriction_L n, \eta \restriction_R n : h(\eta) \leq n < \omega\} : \eta \in \Lambda \rangle$$

is a sequence of pairwise disjoint sets. Now, we note the following two properties:

(b_1) The R-module $M_\Lambda := \langle x_{\eta \restriction_L n}, x_{\eta \restriction_R n}, y_{\eta,n} : \eta \in \Lambda : h(\eta) \leq n < \omega \rangle$ is free.

(b_2) Set $M_{\Lambda \cup \Lambda_*} := \langle M_\Lambda \cup \{x_\nu : \nu \in \Lambda_*\} \rangle$. Then the R-module $M_{\Lambda \cup \Lambda_*}/M_{\Lambda_*}$ is free.

In view of (b_2) the short exact sequence

$$0 \to M_\Lambda \to M_{\Lambda \cup \Lambda_*} \to M_{\Lambda \cup \Lambda_*}/M_\Lambda \to 0,$$

splits. Combining this along with (b_1), we observe that $M_{\Lambda \cup \Lambda_*}$ is free. Since it includes X, we get the desired claim.

(3)(c) Now, suppose $(R, +)$ is μ-free. Let H be a subset of $(G_{\boldsymbol{c}}/K, +)$ of size $< \mu$. There is a free R-module F such that $H \subset F$. There is a subset S of R of size $< \mu$ such that any element of H can be written from a linear combination from F with coefficients taken from S. As $(R, +)$ is μ-free, there is a free subgroup $(T, +)$ of it containing S. In other words, we have

$$H \subseteq T * F := \Big\langle \sum \{t_i f_i : t_i \in T,\ f_i \in F\} \Big\rangle.$$

Since $(T * F, +)$ is free as an abelian group, we get the desired claim.

(4) Let $\boldsymbol{d}$ be such that:

(i) $\Lambda_{\boldsymbol{d}} = \Lambda \cap \Lambda[\gamma]$.

(ii) $X_{\boldsymbol{d}}$ is defined using $\Lambda_{\boldsymbol{d}}$ naturally.

(iii) For $\nu \in \Lambda_{\boldsymbol{d},\omega}$ and $n < \omega$, $a^{\boldsymbol{d}}_{\nu,n} = a^{\boldsymbol{c}}_{\nu,n}$.

(iv) $\Gamma_{\boldsymbol{d}}$ is defined naturally as the set of equations in (1), but only for $\eta \in \Lambda_{\boldsymbol{d},\omega}$.

This is straightforward to check that $\boldsymbol{d}$ is as required.

(5) Let $\boldsymbol{d}$ be defined in the natural way, so that:

(i) $\Lambda_{\boldsymbol{d}} = \Lambda_{\boldsymbol{c}} \cup \{\eta\} \cup \{\eta{\restriction_L} n, \eta{\restriction_R} n : n < \omega\}$.

(ii) $X_{\boldsymbol{d}} = X_{\boldsymbol{c}} \cup \{x_{\eta\restriction_L n}, x_{\eta\restriction_R n} : n < \omega\} \cup \{y_{\eta,n} : n < \omega\}$.

(iii) For $\nu \in \Lambda_{\boldsymbol{c},\omega}$ and $n < \omega$, $a^{\boldsymbol{d}}_{\nu,n} = a^{\boldsymbol{c}}_{\nu,n}$.

(iv) $a^{\boldsymbol{d}}_{\eta,n} = a_n$ for $n < \omega$.

(v) In addition to the equations displayed in $\Gamma_{\boldsymbol{c}}$, $\Gamma_{\boldsymbol{d}}$ contains equations of the forms

$$y_{\eta,n} = a_n + (n!)\, y_{\eta,n+1} + (x_{\eta\restriction_L n} - x_{\eta\restriction_R n}),$$

where $n < \omega$.

The assertion is now obvious by the above definition of $\boldsymbol{d}$.

(6) In view of Lemma 3.19, the group $G_{\boldsymbol{c}}$ is of size $\lambda^{\aleph_0}$. Recall from Hypothesis 3.14 that $\lambda^{\aleph_0} = \lambda$. So, the desired claim is clear. □

Lemma 3.27. *Let $\boldsymbol{c} \in$ AP. Then the abelian group $G_{\boldsymbol{c}}/K$ is reduced.*

Proof. Suppose on the way of contradiction that $G_{\boldsymbol{c}}/K$ is not reduced. Then it has a divisible direct summand, say I. By Fact 2.18, I is injective. We apply the structure theorem for injective abelian groups (see Discussion 2.20(i)) to find the decomposition

$$I = \bigoplus_{p \in \mathbb{P}} \mathbb{Z}(p^{\infty})^{\oplus x_p} \oplus \mathbb{Q}^{\oplus x},$$

where x_p and x are index sets. Since $G_{\boldsymbol{c}}/K$ is torsion-free, I is torsion-free. So, I has no p-torsion part. This shows that $x_p = \varnothing$ for all $p \in \mathbb{P}$. In other words, $I = \mathbb{Q}^{\oplus x}$. Since I is nonzero, $x \neq \varnothing$. This yields that $(\mathbb{Q}, +)$ is a directed summand of $G_{\boldsymbol{c}}/K$. Thanks to Lemma 3.26(3)(a), $\Lambda_{\boldsymbol{c}}$ is $\aleph_1$-free. We combine this with Lemma 3.26(3)(b) to deduce that $M_{\boldsymbol{c}}$ is $\aleph_1$-free as an R-module.

We have two possibilities: (1) $\boldsymbol{k}$ is trivial and (2) $\boldsymbol{k}$ is nontrivial.

(1) $\boldsymbol{k}$ is trivial. Then $R := \mathbb{Z}$. Recall that $M_{\boldsymbol{c}} = G_{\boldsymbol{c}}/K$ is $\aleph_1$-free. Since $(\mathbb{Q}, +)$ is countable, it should be free, a contradiction.

(2) $\boldsymbol{k}$ is nontrivial. Recall that R is $S^{\perp}_{\boldsymbol{k}}$-divisible. Since the context is nontrivial, there is $p \in S^{\perp}_{\boldsymbol{k}}$ such that $\{1/p^n : n \gg 0\} \subseteq R$. For simplicity, we assume that $\{1/p^n : n > 0\} \subseteq R$. Since $M_{\boldsymbol{c}}$ is $\aleph_1$-free and that $\{1/p^n : n > 0\} \subseteq \mathbb{Q} \subseteq M_{\boldsymbol{c}}$, there is a free R-module $F \subseteq M_{\boldsymbol{c}}$ such that $\{1/p^n : n > 0\} \subseteq F$. Let $F = \bigoplus R$. So, the desired contraction follows by

$$\begin{aligned}\{r/p^n : n > 0, r \in R\} &= \bigcap_{\ell>0} p^{\ell}\{r/p^n : n > 0, r \in R\}\\ &\subseteq \bigcap_{\ell>0} p^{\ell} F = \bigoplus\Big(\bigcap_{\ell>0} p^{\ell} R\Big) \subseteq \bigoplus\Big(\bigcap_{\ell>0} \ell R\Big) = 0,\end{aligned}$$

where the last equality comes from the fact that $(R, +)$ is cotorsion-free. In fact, by Fact 2.19, the abelian group $(R, +)$ is reduced, and so $\bigcap_{\ell>0} \ell R = 0$. The proof is now complete. □

Lemma 3.28. *Let $\boldsymbol{c} \in \mathrm{AP}_{\boldsymbol{k},\lambda}$. Then*

$$y^{\boldsymbol{c}}_{\eta,0} = \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\, a^{\boldsymbol{c}}_{\eta,i} + \Big(\prod_{i=1}^{n} i!\Big)\, y^{\boldsymbol{c}}_{\eta,n+1} + \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)(x^{\boldsymbol{c}}_{\eta\restriction_L i} - x^{\boldsymbol{c}}_{\eta\restriction_R i})$$

is valid for any $n < \omega$.

Proof. We proceed by induction on n. The desired claim is clearly holds for $n = 0$. Suppose inductively that it holds for n. We are going to show the claim for $n+1$. To this end, we apply the induction assumption along with the relation

$$y^{\boldsymbol{c}}_{\eta,n+1} = a^{\boldsymbol{c}}_{\eta,n+1} + (n+1)!\, y^{\boldsymbol{c}}_{\eta,n+2} + (x^{\boldsymbol{c}}_{\eta\restriction_L n+1} - x^{\boldsymbol{c}}_{\eta\restriction_R n+1})$$

to deduce

$$\begin{aligned} y^{\boldsymbol{c}}_{\eta,0} &= \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\, a^{\boldsymbol{c}}_{\eta,i} + \Big(\prod_{i=1}^{n} i!\Big)\, y^{\boldsymbol{c}}_{\eta,n+1} + \sum_{i=0}^{n+1}(x^{\boldsymbol{c}}_{\eta\restriction_L i} - x^{\boldsymbol{c}}_{\eta\restriction_R i})\\ &= \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\, a^{\boldsymbol{c}}_{\eta,i} + \Big(\prod_{i=0}^{n} i!\Big)\, a^{\boldsymbol{c}}_{\eta,n+1} + \Big(\prod_{i=1}^{n} i!\Big)(n+1)!\, y^{\boldsymbol{c}}_{\eta,n+2}\\ &\qquad + \Big(\prod_{i=0}^{n} i!\Big)(x^{\boldsymbol{c}}_{\eta\restriction_L n+1} - x^{\boldsymbol{c}}_{\eta\restriction_R n+1}) + \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)(x^{\boldsymbol{c}}_{\eta\restriction_L i} - x^{\boldsymbol{c}}_{\eta\restriction_R i})\\ &= \sum_{i=0}^{n+1}\Big(\prod_{j<i} j!\Big)\, a^{\boldsymbol{c}}_{\eta,i} + \Big(\prod_{i=1}^{n+1} i!)\, y^{\boldsymbol{c}}_{\eta,n+2} + \sum_{i=0}^{n+1}\Big(\prod_{j<i} j!\Big)(x^{\boldsymbol{c}}_{\eta\restriction_L i} - x^{\boldsymbol{c}}_{\eta\restriction_R i}).\end{aligned}$$

Thus the claim holds for $n+1$ as well. □

There are some distinguished and useful objects in $\mathrm{AP}_{\boldsymbol{k},\lambda}$.

Definition 3.29. We say $\boldsymbol{c} \in \mathrm{AP}_{\boldsymbol{k},\lambda}$ is full when:

(a) $\Lambda_{\boldsymbol{c}} \supseteq \Lambda_{<\omega}[\lambda]$.

(b) If $a_n \in G_{\boldsymbol{c}}$ for $n < \omega$ and $f : \Lambda_{<\omega}[\lambda] \to \gamma$, where $\gamma < \lambda$, then for some $\eta \in \Lambda_{\boldsymbol{c}}$ and all $n < \omega$ we have $a^{\boldsymbol{c}}_{\eta,n} = a_n$ and $f(\eta \restriction_L n) = f(\eta \restriction_R n)$.

Now, we study the existence problem for fullness in AP.

Lemma 3.30. *Adopt the notation from* Hypothesis 3.14. *Then there are some full* $\boldsymbol{c} \in \mathrm{AP}_{\boldsymbol{k},\lambda}$.

Proof. Let $\boldsymbol{b}$ be a twofold λ-black box, which exists by Lemma 3.15. We look at

$$\Omega := \Lambda_{<\omega}[\lambda] \cup (\Lambda_{\omega}[\lambda] \times \omega),$$

and for each ordinal $\alpha < \lambda$ we set

$$\Omega_\alpha := \Lambda_{<\omega}[\alpha] \cup (\Lambda_{\omega}[\alpha] \times \omega).$$

Fix a bijection map

$$h : S_\lambda \xrightarrow{\cong} (\oplus_{\rho\in\Omega} Rx_\rho) \oplus K$$

such that for each ordinal $\alpha < \lambda$ one has

$$(*) \qquad h''[S_\alpha] \subseteq (\oplus_{\rho\in\Omega_\alpha} Rx_\rho) \oplus K.$$

This is possible, as for each α,

$$|S_\alpha| \leq \aleph_0 + |\alpha| \leq |(\oplus_{\rho\in\Omega_\alpha} Rx_\rho) \oplus K| < \lambda.$$

Let $\boldsymbol{c}$ be defined by:

(1) $\Lambda_{\boldsymbol{c}} = \Lambda_\omega[\lambda] \cup \Lambda_{<\omega}[\lambda]$.

(2) $X_{\boldsymbol{c}}$ is the set

$$\{rx_\nu : r \in R, \nu \in \Lambda_{\boldsymbol{c},<\omega}\} \cup \{ry_{\eta,n} : r \in R, \eta \in \Lambda_{\boldsymbol{c},\omega}, n < \omega\}.$$

(3) $a^{\boldsymbol{c}}_{\eta,n} = h(g^{\boldsymbol{b}}_\eta(n+1))$, where $g^{\boldsymbol{b}}_\eta$ is given by the twofold λ-black box.

(4) $G_{\boldsymbol{c}}$ is generated, as an abelian group, freely by the sets K and $X_{\boldsymbol{c}}$ except the set of relations

$$y_{\eta,n} = a^{\boldsymbol{c}}_{\eta,n} + (n!)\, y_{\eta,n+1} + (x_{\eta\restriction_L n} - x_{\eta\restriction_R n}),$$

with the convenience that $a^{\boldsymbol{c}}_{\eta,n}$ is regarded as an element of $G_{\boldsymbol{c}}$ via the quotient map

$$\left(\bigoplus_{\rho\in\Omega} Rx_\rho\right) \oplus K \twoheadrightarrow G_{\boldsymbol{c}}.$$

From this identification and $(*)$, we have $a^{\boldsymbol{c}}_{\eta,n} \in G_{\boldsymbol{c},\eta(0,1)}$.

(5) $\Gamma_{\boldsymbol{c}}$ is defined naturally as in Definition 3.18.

Let us show that $\boldsymbol{c}$ is as required. It clearly satisfies (a) of Definition 3.29. To show that (b) of Definition 3.29 is satisfied, let $\langle a_n : n < \omega \rangle \in {}^{\omega}G_{\boldsymbol{c}}$ and $f : \Lambda_{<\omega}[\lambda] \to \gamma$, where $\gamma < \lambda$. Let $g : \Lambda_{<\omega}[\lambda] \to S_\lambda$ be defined such that for all $\nu \in \Lambda_{<\omega}[\lambda] \setminus \{\langle\ \rangle\}$,

$$(+) \qquad h(g(\nu)) = a_{\lg(\nu)-1}.$$

We are going to apply the twofold λ-black box $\boldsymbol{b}$. According to its properties, there is an $\eta \in \Lambda_\omega[\lambda]$ such that:

(6) $\gamma(\eta) > \gamma$,

(7) $g^{\boldsymbol{b}}_\eta(0) = g(\langle\ \rangle)$,

(8) $g^{\boldsymbol{b}}_\eta(n+1) = g(\eta \restriction_L n)$,[2]

(9) $\eta(n, 1) < \eta(n, 2)$ and $f(\eta \restriction_L n) = f(\eta \restriction_R n)$ for all $1 \leq n < \omega$.

Applying h to the both sides of (8), one has

$$a^{\boldsymbol{c}}_{\eta,n} \overset{(3)}{=} h(g^{\boldsymbol{b}}_\eta(n+1)) = h(g(\eta\restriction_L n)) \overset{(+)}{=} a_n,$$

thereby completing the proof. □

Lemma 3.31. *Assume $\boldsymbol{c} \in$ AP is full and let $h \in \operatorname{Hom}(G_{\boldsymbol{c}}, K)$ be unbounded. Then there is a sequence*

$$\langle a_n : n < \omega \rangle \in {}^{\omega}\mathrm{Rang}(h)$$

such that the following set of equations Γ has no solution, not only in $G_{\boldsymbol{c}}$, but in any $G_{\boldsymbol{d}}$ with $\boldsymbol{c} \leq \boldsymbol{d} \in$ AP, where

$$\Gamma := \{z_n = a_n + n!\, z_{n+1} : n < \omega\}.$$

Proof. We have two possibilities. First, suppose for some prime number p, the group $\Gamma_p(\mathrm{Rang}(h))$ is infinite, and let p be the first such prime number. Also, let $p_n = p$ for all $n < \omega$. Otherwise, we let

$$p_n \in \{p : \Gamma_p(\mathrm{Rang}(h)) \neq 0\}$$

be a strictly increasing sequence of prime numbers. We refer this as a second possibility.

In the first part of the proof, we argue for both possibilities at the same time. Then, we consider each scenario separately.

Since h is not bounded, we can find by induction on n, the pair (H_n, a_n) such that:

(+) (a) $H_0 = \mathrm{Rang}(h)$.

(b) $H_n = a_n \mathbb{Z} \oplus H_{n+1}$.

[2]Here we are using a modified version of the twofold λ-black box $\boldsymbol{b}$, which can be easily obtained from the original one.

(c) a_n has order $p_n^{\boldsymbol{l}_n}$.

(d) For $n = m + 1$ we have

$$(d_n) : \boldsymbol{l}_n > \boldsymbol{l}_m + \left(\prod_{i=0}^{n+1} i!\right).$$

To see this, let $H_0 := \mathrm{Rang}(h)$ and let $a_0 \in \Gamma_{p_0}[\mathrm{Rang}(h)]$ be any nonzero element. Now, suppose inductively that $n > 0$ and we have defined $\langle H_i : i \leq n\rangle$ and $\langle a_i : i < n\rangle$ satisfying the above items. We shall now define a_n and H_{n+1}. By our induction assumption, we have

$$\mathrm{Rang}(h) = \left(\bigoplus_{i<n} a_i\, \mathbb{Z}\right) \oplus H_n.$$

In particular, H_n is torsion. Using Fact 2.5 (and also Fact 2.7 in the second possibility case), we can find for some ℓ_n and an element a_n such that a_n has order $p_n^{\boldsymbol{l}_n}$ and $a_n\, \mathbb{Z}$ is a direct summand of H_n. We may further suppose that

$$\boldsymbol{l}_n > \boldsymbol{l}_m + \left(\prod_{i=0}^{n+1} i!\right).$$

Since (a_n) is a direct summand of H_n, there is an abelian group H_{n+1} so that $H_n = a_n\, \mathbb{Z} \oplus H_{n+1}$.

To prove that the sequence $\langle a_n : n < \omega\rangle$ is as required, assume towards a contradiction that there is $\boldsymbol{c} \leq \boldsymbol{d} \in \mathrm{AP}$ such that $\langle c_n : n < \omega\rangle$ is a solution of Γ in $G_{\boldsymbol{d}}$. So

$$(*) \qquad G_{\boldsymbol{d}} \models \bigwedge_{n<\omega} (c_n = a_n + n!\, c_{n+1}).$$

Since for each n, $a_n \in K$, it follows that

$$G_{\boldsymbol{d}}/K \models \bigwedge_{n<\omega} (c_n + K = n!\, c_{n+1} + K).$$

By Lemma 3.27, $G_{\boldsymbol{c}}/K$ is reduced, and hence necessarily,

$$\bigwedge_{n<\omega} (c_n + K = 0 + K).$$

In other words, $c_n \in K$ for all $n < \omega$.

We now show that for each n,

$$(**) \qquad \left(\prod_{i<n} i!\right) c_n \in H_n$$

This is true for $n=0$, because $c_0 \in K = H_0$. Suppose it holds for n. Then multiplying both sides of $(*)$ into $\prod_{i<n} i!$ we get

$$\Big(\prod_{i<n} i!\Big) c_n = \Big(\prod_{i<n} i!\Big) a_n + \Big(\prod_{i<n+1} i!\Big) c_{n+1}.$$

Using the induction hypothesis and $(+)(b)$ we get

$$\Big(\prod_{i<n+1} i!\Big) c_{n+1} \in H_{n+1},$$

as requested.

By an easy induction, for each n we have

$$(***_n) \qquad c_0 = a_0 + \sum_{\ell \le n} \Big(\prod_{i=1}^{\ell} i!\Big) a_\ell + \Big(\prod_{i=1}^{n} i!\Big) c_{n+1}.$$

Indeed this is true for $n=0$, as $c_0 = a_0 + c_1$. Suppose it holds for n, then using $(*)$ and the induction hypothesis, we get

$$\begin{aligned} c_0 &= a_0 + \sum_{\ell \le n} \Big(\prod_{i=1}^{\ell} i!\Big) a_\ell + \Big(\prod_{i=1}^{n} i!\Big) c_{n+1} \\ &= a_0 + \sum_{\ell \le n} \Big(\prod_{i=1}^{\ell} i!\Big) a_\ell + \Big(\prod_{i=1}^{n} i!\Big)\big(a_{n+1} + (n+1)!\, c_{n+2}\big) \\ &= a_0 + \sum_{\ell \le n+1} \Big(\prod_{i=1}^{\ell} i!\Big) a_\ell + \Big(\prod_{i=1}^{n+1} i!\Big) c_{n+2}. \end{aligned}$$

We are now ready to complete the proof. Let $m(*)$ be the order of c_0.

Now, we consider each case separately.

Case 1. *$p_n = p$ for all n.*

Let t be an integer such that

$$m(*) = tp^{\ell(*)} > 1,$$

where $\ell(*) \ge 0$s and $(p, t) = 1$, i.e., p does not divide t. Let k be the least natural number such that $\boldsymbol{l}_k > \ell(*)$. By multiplying both sides of $(***)_{k+1}$ into $tp^{\boldsymbol{l}_k}$, we get to

$$tp^{\boldsymbol{l}_k} c_0 = tp^{\boldsymbol{l}_k} a_0 + tp^{\boldsymbol{l}_k} \sum_{\ell \le k+1} \Big(\prod_{i=1}^{\ell} i!\Big) a_\ell + tp^{\boldsymbol{l}_k} \Big(\prod_{i=1}^{k+1} i!\Big) c_{k+2}.$$

Since the sequence $\langle \boldsymbol{l}_\ell : \ell \leq k \rangle$ is increasing, we have $p^{\boldsymbol{l}_k} a_\ell = 0$ for all $\ell \leq k$. Consequently,

$$0 = tp^{\boldsymbol{l}_k} \left(\prod_{i=1}^{k+1} i! \right) a_{k+1} + tp^{\boldsymbol{l}_k} \left(\prod_{i=1}^{k+1} i! \right) c_{k+2} \tag{$\dagger$}$$

According to $(+)$(b), we know $a_{k+1}\,\mathbb{Z} \cap H_{k+2} = 0$, and by using $(**)$ along with $(\dagger)$ we get that

$$tp^{\boldsymbol{l}_k} \left(\prod_{i=1}^{k+1} i! \right) a_{k+1} = 0.$$

Recall that the order of a_{k+1} is a power of p. We apply this along with the equality $(p, t) = 1$ to get that

$$p^{\boldsymbol{l}_k} \left(\prod_{i=1}^{k+1} i! \right) a_{k+1} = 0.$$

Moreover,

$$p^{\boldsymbol{l}_{k+1}} = \operatorname{ord}(a_{k+1}) \leq p^{\boldsymbol{l}_k} \left(\prod_{i=1}^{k+1} i! \right) \leq p^{\boldsymbol{l}_k + (\prod_{i=1}^{k+1} i!)}.$$

Taking $\log_p(-)$ from both sides, we have $\boldsymbol{l}_{k+1} \leq \boldsymbol{l}_k + \left(\prod_{i=1}^{k+1} i! \right)$. But, this contradicts $(d_{\boldsymbol{l}_{k+1}})$. The result follows.

Thereby, without loss of generality we deal with:

Case 2. *Otherwise.*

The sequence $\langle p_n : n < \omega \rangle$ is strictly increasing. If k is the least integer, then

$$p_{k+1} > m(*) \times \left(\prod_{i=1}^{k+1} i! \right). \tag{$\dagger\dagger$}$$

By multiplying both sides of $(***)_{k+1}$ into $m(*) \times \left(\prod_{i=1}^{k} p_i^{\boldsymbol{l}_i} \right)$ we get

$$\begin{aligned} 0 &= m(*) \times \left(\prod_{i=1}^{k} p_i^{\boldsymbol{l}_i} \right) c_0 \\ &= m(*) \times \left(\prod_{i=1}^{k} p_i^{\boldsymbol{l}_i} \right) a_0 + m(*) \times \left(\prod_{i=1}^{k} p_i^{\boldsymbol{l}_i} \right) \sum_{\ell \leq k+1} \left(\prod_{i=1}^{\ell} i! \right) a_\ell \\ &\qquad + m(*) \times \left(\prod_{i=1}^{k} p_i^{\boldsymbol{l}_i} \right) \left(\prod_{i=1}^{k+1} i! \right) c_{k+2}. \end{aligned}$$

We have that $m(*) \times \left(\prod_{i=1}^{k} p_i^{\boldsymbol{l}_i} \right) a_0 = 0$ and

$$m(*) \times \left(\prod_{i=1}^{k} p_i^{\boldsymbol{l}_i} \right) \left(\prod_{i=1}^{\ell} i! \right) a_\ell = 0 \quad \text{for all } \ell \leq k.$$

Thus

$$0 = m(*) \times \left(\prod_{i=1}^{k} p_i^{l_i}\right)\left(\prod_{i=1}^{k+1} i!\right) a_{k+1} + m(*) \times \left(\prod_{i=1}^{k} p_i^{l_i}\right)\left(\prod_{i=1}^{k+1} i!\right) c_{k+2}.$$

Again, according to $(+)(\mathrm{b})$, we know $a_{k+1}\,\mathbb{Z} \cap H_{k+2} = 0$, and by using $(**)$ along with the previous formula, we lead to the following vanishing formula:

$$m(*) \times \left(\prod_{i=1}^{k} p_i^{l_i}\right)\left(\prod_{i=1}^{k+1} i!\right) a_{k+1} = 0.$$

As the order of a_{k+1} is a power of p_{k+1} and it is different from all p_ℓ's, for $\ell \leq k$, we have

$$m(*) \times \left(\prod_{i=1}^{k+1} i!\right) a_{k+1} = 0.$$

So,

$$p_{k+1} < p_{k+1}^{l_{k+1}} = \mathrm{ord}(a_{k+1}) \leq m(*) \times \left(\prod_{i=1}^{k+1} i!\right).$$

But this contradicts $(\dagger\dagger)$. The result follows. □

To prove the endo-rigidity property, we first deal with the following special case, and then we reduce things to this situation.

Lemma 3.32. *Let $\boldsymbol{c} \in \mathrm{AP}$ be full. Then every $h \in \mathrm{Hom}(G_{\boldsymbol{c}}, K)$ is bounded.*

Proof. Towards a contradiction assume $h \in \mathrm{Hom}(G_{\boldsymbol{c}}, K)$ is not bounded. In view of Lemma 3.31, this implies that there is a sequence

$$\langle a_n : n < \omega \rangle \in {}^{\omega}\mathrm{Rang}(h)$$

such that the set of equations

$$\Gamma := \{z_n = a_n + n!\, z_{n+1} : n < \omega\}$$

has no solutions in $G_{\boldsymbol{c}}$. Let $\gamma = |K|$, and define $f : \Lambda_{<\omega}[\lambda] \to \gamma$ such that

$$(*) \qquad f(\eta) = f(\nu) \Longleftrightarrow h(x_\eta) = h(x_\nu)$$

Since $a_n \in \mathrm{Rang}(h)$ there is b_n such that

$$(+) \qquad \forall n < \omega, \quad a_n = h(b_n)$$

As $\boldsymbol{c}$ is full, we can find some η such that

$$f(\eta{\restriction_L} n) = f(\eta{\restriction_R}\, n) \quad \text{and} \quad a^{\boldsymbol{c}}_{\eta,n} = b_n \ \text{ for each } n.$$

Let us combining $(*)$ and (1). This yields that

$$(\dagger) \qquad \forall n < \omega, \quad h(x_{\eta{\restriction_L} n}) = h(x_{\eta{\restriction_R}\, n}).$$

Moreover, by applying h to the both sides of the equation

$$y_{\eta,n} = a^{\boldsymbol{c}}_{\eta,n} + (n!)\, y_{\eta,n+1} + (x_{\eta\restriction_L n} - x_{\eta\restriction_R n}),$$

we lead to the following equation:

$$\begin{aligned} h(y_{\eta,n}) &= h(a^{\boldsymbol{c}}_{\eta,n}) + n!\, h(y_{\eta,n+1}) + \big(h(x_{\eta\restriction_L n}) - h(x_{\eta\restriction_R n})\big) \\ &\overset{(2)}{=} h(b_n) + n!\, h(y_{\eta,n+1}) + \big(h(x_{\eta\restriction_L n}) - h(x_{\eta\restriction_R n})\big) \\ &\overset{(\dagger)}{=} h(b_n) + (n!)\, h(y_{\eta,n+1}) \overset{(+)}{=} a_n + (n!)\, h(y_{\eta,n+1}). \end{aligned}$$

In other words, $h(y_{\eta,n})$ is a solution for

$$\Gamma = \{z_n = a_n + n!\, z_{n+1} : n < \omega\}.$$

This is a contradiction with the choice of the sequence $\langle a_n : n < \omega\rangle$. □

Notation 3.33. Suppose $\boldsymbol{c} \in \mathrm{AP}$. For each $n < \omega$, we define

$$G_n := \frac{G_{\boldsymbol{c}}}{K + \big(\prod_{i=1}^{n} i!\big)\, G_{\boldsymbol{c}}}.$$

Also, the notation π_n stands for the natural projection $G_{\boldsymbol{c}} \twoheadrightarrow G_n$.

Fact 3.34. *Adopt the above notation, let $n < \omega$ and $g \in G_{\boldsymbol{c}}$.*

(a) *The abelian group G_n is a torsion abelian group with the following minimal generating set*

$$\{x_\rho : \rho \in \Lambda_{\boldsymbol{c},<\omega}\} \cup \{y_{\eta,k} : \eta \in \Lambda_{\boldsymbol{c},\omega} \text{ and } k \geq n+2\}.$$

(b) *Similar to* Definition 3.20, *we can define* $\mathrm{supp}_\circ(\pi_n(g))$ *with respect to generating set presented in* (a).

(c) *According to its definition, it is easy to see that* $\mathrm{supp}_\circ(\pi_n(g)) \subseteq \mathrm{supp}_\circ(g)$.

(d) *Recall from* Lemma 3.27 *that G_c/K is reduced. This in turns gives us an integer $m_n > n$ such that* $\mathrm{supp}_\circ(g) \subseteq \mathrm{supp}_\circ(\pi_{m_n}(g))$.

Proof. This is straightforward. □

Lemma 3.35. *Suppose $\boldsymbol{c} \in \mathrm{AP}$ is full and $h \in \mathrm{End}(G_{\boldsymbol{c}})$. Then for some countable $\Lambda_h \subseteq \Omega_{\boldsymbol{c}}$ we have*

$$r \in R, \quad \nu \in \Omega_{\boldsymbol{c}} \setminus \Lambda_h \Longrightarrow \mathrm{supp}_\circ(h(r x_\nu)) \subseteq \{\nu\} \cup \Lambda_h.$$

Proof. Towards contradiction assume $h \in \mathrm{End}(G_{\boldsymbol{c}})$ but there is no Λ_h as promised.

We define a sequence

$$\langle (\eta_i, Y_i, \nu_i, r_i) : i < \omega_1\rangle,$$

by induction on $i < \omega_1$, such that

$(*)$ (a) $\eta_i \in \Omega_{\boldsymbol{c}}$ and $r_i \in R \setminus \{0\}$,

(b) $Y_i = \bigcup\{\mathrm{supp}_\circ(h(r_j\, x_{\eta_j})) : j < i\} \cup \{\eta_j : j < i\}$,

(c) $\nu_i \in \mathrm{supp}_\circ(h(r_i\, x_{\eta_i}))$ but $\nu_i \neq \eta_i$, $\nu_i \notin Y_i$.

To this end, suppose that $i < \omega_1$ and we have defined $\langle(\eta_j, Y_j, \nu_j, r_j) : j < i\rangle$. Set

$$Y_i = \bigcup\{\mathrm{supp}_\circ(h(r_j\, x_{\eta_j})) : j < i\} \cup \{\eta_j : j < i\}.$$

Following its definition, we know Y_i is at most countable. Thus, due to our assumption, we can find some $\eta_i \in \Omega_{\boldsymbol{c}} \setminus Y_i$ and $r_i \in R \setminus \{0\}$ such that

$$\mathrm{supp}_\circ(h(r_i\, x_{\eta_i})) \nsubseteq (\{\eta_i\} \cup Y_i).$$

This allows us to define ν_i, namely, it is enough to take ν_i be any element of $\mathrm{supp}_\circ(h(r_i\, x_{\eta_i})) \setminus (\{\eta_i\} \cup Y_i)$. This completes the definition of $(\eta_i, Y_i, \nu_i, r_i)$.

Combining the facts $\nu_i \in \mathrm{supp}_\circ(h(r_i\, x_{\eta_i}))$ and $\nu_i \notin (Y_i \cup \{\eta_i\})$ along with the finiteness of $\mathrm{supp}_\circ(h(x_{\eta_i}))$ we are able to find a subset $W \subseteq \omega_1$ of cardinality ω_1 such that $\nu_j \notin \mathrm{supp}_\circ(h(r_i\, x_{\eta_i}))$ when $i \neq j \in W$.

Without loss of generality we may and do assume that $W = \omega_1$. Let $a_i = r_i\, x_{\eta_i}$. We can find

$$f : \Lambda_{\boldsymbol{c},<\omega} \to |R| + \aleph_0 < \lambda$$

such that if $b \in G_{\boldsymbol{c}}$,[3] then from $f(b)$ we can compute

$$\langle n_b, \{(\ell, m_{b,\ell}, r_{b,\ell}) : \ell < n_b\}\rangle.$$

Recall that $\boldsymbol{c}$ is full, and that $\mathrm{Rang}(f)$ has size less than λ. From these, there is some $\eta \in \Lambda_{\boldsymbol{c},\omega}$ furnished with two properties:

(1) $f(\eta \restriction_L n) = f(\eta \restriction_R n)$ for $n < \omega$,

(2) $a^{\boldsymbol{c}}_{\eta,n} = a_n$ for all $n < \omega$.

Now, we bring a claim.

Claim. $\nu_i \in \mathrm{supp}_0(h(y_{\eta 0}))$ *for all* $i < \omega$.

Note that this will give us the desired contradiction, as $\mathrm{supp}_0(h(y_{\eta 0}))$ is finite.

Proof of Claim. By Lemma 3.28 we first observe that

$$y_{\eta,0} = \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\, r_i\, x_{\eta_i} + \Big(\prod_{i=1}^{n} i!\Big)\, y_{\eta,n+1} + \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)(x_{\eta\restriction_L i} - x_{\eta\restriction_R i}).$$

Let ℓ be any integer. We are going to use the notation presented in Notation 3.33 for $n = m_\ell$. Applying $\pi_n h(-)$ to it yields that

[3] Recall that we have chosen $b - \sum_{\ell<n_b} r_{b,\ell}\, y_{\eta_{b,\ell},m_{b,\ell}} \in \sum_{\rho\in\Lambda_{\boldsymbol{c},<\omega}} Rx_\rho + K$.

$$(3)\quad \pi_n(h(y_{\eta,0})) = \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\,\pi_n\, h(r_i\, x_{\eta_i}) + \Big(\prod_{i=1}^{n} i!\Big)\,\pi_n\, h(y_{\eta,n+1}) + \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\,\pi_n\, h(x_{\eta\restriction_L i} - x_{\eta\restriction_R i})$$
$$= \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\,\pi_n\, h(r_i\, x_{\eta_i}) + \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\,\pi_n\, h(x_{\eta\restriction_L i} - x_{\eta\restriction_R i}),$$

where the last equality follows by Notation 3.33. Now, we recall from the construction $(*)$ that

$$\nu_i \in \operatorname{supp}_\circ(h(r_i\, x_{\eta_i})), \quad \nu_i \neq \eta_i,\ \nu_i \notin Y_i.$$

Thanks to Fact 3.34(d) we have

$$(4)\qquad \nu_i \in \operatorname{supp}_\circ(\pi_n h(r_i\, x_{\eta_i})).$$

By clause (1) above, $\operatorname{supp}_\circ(h(x_{\eta\restriction_L i} - x_{\eta\restriction_R i})) = \varnothing$. In view of Fact 3.34(c), we deduce that

$$(5)\qquad \operatorname{supp}_\circ\big(\pi_n(h(x_{\eta\restriction_L i} - x_{\eta\restriction_R i}))\big) = \varnothing.$$

First, we plug items (4) and (5) in the clause (3), then we use $(*)$. These enable us to observe that

$$\nu_i \in \operatorname{supp}_\circ\Bigg(\sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\,\pi_n\, h(r_i\, x_{\eta_i}) + \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)\,\pi_n\, h(x_{\eta\restriction_L i} - x_{\eta\restriction_R i})\Bigg) = \operatorname{supp}_\circ(\pi_n\, h(y_{\eta,0})).$$

Another use of Fact 3.34(c), shows that $\nu_i \in \operatorname{supp}_\circ(h(y_{\eta,0}))$. This completes the proof of the claim. □

The lemma follows. □

Lemma 3.36. *Let $\boldsymbol{c} \in \mathrm{AP}$ be full and $h \in \mathrm{End}(G_{\boldsymbol{c}})$. Let $Y_0 \subseteq \Omega_{\boldsymbol{c}}$ be the downward closure of Λ_h, where Λ_h is as in Lemma 3.35 and set*

$$K^+ := K + \sum_{\rho\in Y_0\cap\Lambda_{\boldsymbol{c},<\omega}} R x_\rho + \sum_{\rho\in Y_0\cap\Lambda_{\boldsymbol{c},\omega} n<\omega} R y_{\rho,n}.$$

If $b \in G_{\boldsymbol{c}}$, then there are choices

- $\bar r_b := \langle r^2_{b,\rho} : \rho \in \Lambda_b\rangle$, *and*
- $\Lambda_{\boldsymbol{b}} \subseteq \Lambda_{\boldsymbol{c},<\omega} \setminus Y_0$ *finite*

such that

$$b - \sum_{\rho\in\Lambda_{\boldsymbol{b}}} r^2_{b,\rho}\, x_\rho \in K^+.$$

Proof. This is straightforward. □

Hypothesis 3.37. For the rest of this section, we fix a well-ordering $\prec$ of the large enough part of the universe, and for each:

- $\boldsymbol{c} \in \mathrm{AP}$ which is full,
- $h \in \mathrm{End}(G_{\boldsymbol{c}})$, and
- $b \in G_{\boldsymbol{c}}$,

we let $\bar{r}_b := \langle r^2_{b,\rho} : \rho \in \Lambda_b \rangle$ be the $\prec$-least sequence satisfying the conclusions of Lemma 3.36.

Notation 3.38. Suppose $\boldsymbol{c} \in \mathrm{AP}$ and $\Lambda \subseteq \Lambda_{\boldsymbol{c}}$. By $G_{\boldsymbol{c},\Lambda}$ we mean

$$G_{\boldsymbol{c},\Lambda} := G_\Lambda := \langle \{ r x_\nu, r y_{\eta,n} : r \in R, \nu \in \Lambda_{<\omega}, \eta \in \Lambda_\omega \text{ and } n < \omega \} \rangle.$$

We have the following observation, but as we do not use it, we leave its proof.

Observation 3.39. *Suppose $\Lambda \subseteq \Lambda[\lambda]$ is downward closed. Then $G_{\boldsymbol{c},\Lambda}$ is a pure subgroup of $G_{\boldsymbol{c}}$.*

Lemma 3.40. *Let $\boldsymbol{c} \in \mathrm{AP}$ be full, and $h \in \mathrm{End}(G_{\boldsymbol{c}})$. Then for some countable $\Lambda_h \subseteq \Lambda[\lambda]$ we have*

$$r \in R, \quad \nu \in \Omega_{\boldsymbol{c}} \setminus \Lambda_h \Longrightarrow h(r x_\nu) \in G_{\boldsymbol{c},\Lambda_h \cup \{\nu\}} + K.$$

Proof. Suppose on the way of contradiction that the lemma fails. Let Y_0 be as Lemma 3.36. We define a sequence

$$\langle (Y_i, \nu_i, \rho_i, r_i) : i < \omega_1 \rangle,$$

by induction on $i < \omega_1$, such that

($\natural$) (a) $r_i \in R \setminus \{0\}$,
(b) $Y_i = \bigcup \{ \mathrm{supp}(h(r_j x_{\nu_j})) : j < i \} \cup \{ \rho_j : j < i \} \cup Y_0$,
(c) $\nu_i \in \Omega_{\boldsymbol{c}} \setminus Y_i$,
(d) $h(r_i \nu_i) \notin G_{\boldsymbol{c}, Y_i \cup \{\nu_i\}} + K$,
(e) let $b_i := h(r_i \nu_i)$, and let $\bar{r}_{b_i} := \langle r^2_{b_i,\rho} : \rho \in \Lambda_i \rangle$ be as Lemma 3.36 applied to b_i. Then $\rho_i \in \Lambda_i \setminus (Y_i \cup \{\nu_i\})$, and even

$$r^2_{b_i,\rho_i} x_{\rho_i} \notin G_{\boldsymbol{c}, Y_i \cup \{\nu_i\}} + K.$$

To construct this, suppose $i < \omega$ and we have constructed the sequence up to i. Now, ($\natural$)(b) gives the definition of Y_i. Since we assume that the lemma fails, there is an $r_i \in R$ and $\nu_i \in \Omega_{\boldsymbol{c}} \setminus Y_i$ such that $h(r_i x_{\nu_i}) \notin G_{\boldsymbol{c},\Lambda_h \cup \{\nu\}} + K$. Now, we define

$b_i := h(r_i \nu_i)$. Thanks to Lemma 3.36, there is a finite set $\Lambda_i \subseteq \Lambda_{\boldsymbol{c},<\omega} \setminus Y_i$ and a sequence $\langle r^2_{b_i,\rho} : \rho \in \Lambda_i \rangle$ such that

$$b_i - \sum_{\rho \in \Lambda_i} r^2_{b_i,\rho}\, x_\rho \in K^+.$$

As $b_i \notin G_{\boldsymbol{c},Y_i \cup \{\nu_i\}} + K$ and due to the following containment

$$b_i - \sum_{\rho \in \Lambda_i} r^2_{b_i,\rho}\, x_\rho \in K^+ \subseteq G_{\boldsymbol{c},Y_i \cup \{\nu_i\}} + K,$$

there is $\rho_i \in \Lambda_i$ such that $\rho_i \notin (Y_i \cup \{\nu_i\})$, and indeed

$$r^2_{b_i,\rho_i} x_{\rho_i} \notin G_{\boldsymbol{c},Y_i \cup \{\nu_i\}} + K.$$

This completes the proof of construction. By shrinking the sequence, we may and do assume in addition that $\rho_j \notin \Lambda_i$ for all $i \neq j < \omega_1$.

Let $a_n := r_n\, x_{\nu_n}$ and define

$$f : \Lambda_{\boldsymbol{c},<\omega} \to |R| + |K| + \aleph_0 < \lambda$$

be such that for any $\rho \in \Lambda_{\boldsymbol{c},<\omega}$, $f(\rho)$ codes

- $\langle r^2_{b,\rho} : \rho \in \Lambda_b \rangle$, and
- $b - \sum_{\nu \in \Lambda_i} r^2_{b,\nu}\, x_\nu$,

where $b := h(x_\rho)$. To see such a function f exists, first we define:

- $f_1 : R^{<\omega} \times K^+ \to |R| + |K| + \aleph_0$ is a bijection, and
- $f_2 : \Lambda_{\boldsymbol{c},<\omega} \to R^{<\omega} \times K^+$ is defined as

$$f_2(b) = \left(\langle r^2_{b,\rho} : \rho \in \Lambda_b \rangle, b - \sum_{\nu \in \Lambda_i} r^2_{b,\nu}\, x_\nu \right).$$

Then, we set $f := f_1 \circ f_2$. Suppose $\rho_1, \rho_2 \in \Lambda_{\boldsymbol{c},<\omega}$ are such that $f(\rho_1) = f(\rho_2)$. We claim that $h(x_{\rho_1}) = h(x_{\rho_2})$. To see this, it is enough to apply $f(\rho_1) = f(\rho_2)$, and conclude that

(1) $\langle r^2_{b_1,\nu} : \nu \in \Lambda_{b_1} \rangle = \langle r^2_{b_2,\nu} : \nu \in \Lambda_{b_2} \rangle$

(2) $b_1 - \sum_{\nu \in \Lambda_{b_1}} r^2_{b,\nu}\, x_\nu = b_2 - \sum_{\nu \in \Lambda_{b_2}} r^2_{b,\nu}\, x_\nu$,

where $b_i = h(x_{\rho_i})$. But, then we have

$$b_1 = b_1 - \sum_{\nu \in \Lambda_{b_1}} r^2_{b,\nu}\, x_\nu + \left(\sum_{\nu \in \Lambda_{b_1}} r^2_{b,\nu}\, x_\nu \right) \overset{(2)}{=} b_2 - \sum_{\nu \in \Lambda_{b_2}} r^2_{b,\nu}\, x_\nu + \left(\sum_{\nu \in \Lambda_{b_2}} r^2_{b,\nu}\, x_\nu \right) = b_2,$$

i.e., $h(x_{\rho_1}) = h(x_{\rho_2})$.

Since $\boldsymbol{c}$ is full, and in light of Definition 3.29(b), we are able to find an $\eta \in \Lambda_{\boldsymbol{c},\omega}$ such that

(3) $a_n = a^{\boldsymbol{c}}_{\eta,n}$, and

(4) $f(\eta\restriction_L n) = f(\eta\restriction_R n)$,

for all $n < \omega$. Thanks to the previous paragraph and clause (4) we deduce

$$(\sharp) \qquad h(x_{\eta\restriction_L n}) = h(x_{\eta\restriction_R n})$$

By applying h to the both sides of the equation

$$y_{\eta,0} = \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big) r_i\, x_{\nu_i} + \Big(\prod_{i=1}^{n} i!\Big) y_{\eta,n+1} + \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big)(x_{\eta\restriction_L i} - x_{\eta\restriction_R i}),$$

we get

$$(+) \qquad \begin{aligned} h(y_{\eta,0}) &= \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big) h(r_i\, x_{\nu_i}) + \Big(\prod_{i=1}^{n} i!\Big) h(y_{\eta,n+1}) \\ &\qquad + \Big(\prod_{j<i} j!\Big)\big(h(x_{\eta\restriction_L n}) - h(x_{\eta\restriction_R n})\big) \\ &\overset{(\sharp)}{=} \sum_{i=0}^{n}\Big(\prod_{j<i} j!\Big) h(r_i\, x_{\nu_i}) + \Big(\prod_{i=1}^{n} i!\Big) h(y_{\eta,n+1}). \end{aligned}$$

For each $i < \omega_1$, let $b_i = h(r_i\, x_{\nu_i})$. Let also $b = h(y_{\eta,0})$ and let Λ_b be as in Lemma 3.36. As Λ_b is finite, for some large enough n, we have

$$\{\rho_i : i < n\} \setminus \Lambda_b \neq \varnothing.$$

Let $i < n$ be such that $\rho_i \notin \Lambda_b$. Here, we apply the arguments presented in items (3)–(4) in the proof of Lemma 3.35 to the displayed formula $(+)$. So, on the one hand, it turns out that

$$\rho_i \in \Lambda_i \subseteq \Lambda_b.$$

On the other hand by the choice of i, $\rho_i \notin \Lambda_b$. This is a contraction that we searched for it. $\square$

Lemma 3.41. *Let $\boldsymbol{c} \in \mathrm{AP}$ be full, and $h \in \mathrm{End}(G_{\boldsymbol{c}})$. Then for some $m_* \in R$ and some countable $\Lambda_h = \mathrm{cl}(\Lambda_h) \subseteq \Lambda[\lambda]$ we have*

$$r \in R, \quad \nu \in \Omega_{\boldsymbol{c}} \setminus \Lambda_h \Longrightarrow h(rx_\nu) - m_* x_\nu \in G_{\Lambda_h} + K.$$

Proof. In view of Lemma 3.40, there is some countable downward closed subset Λ of $\Lambda_{\boldsymbol{c}}$ such that for every $r \in R$ and $\eta \in \Omega_{\boldsymbol{c}} \setminus \Lambda$, we have $h(rx_\eta) \in G_{\Lambda\cup\{\nu\}} + K$. Thus, for such r and η, there are $m^r_\eta \in R$ and b^r_η satisfying the following two properties:

$$h(rx_\eta) = m^r_\eta\, x_\eta + b^r_\eta \quad \text{and} \quad b^r_\eta \in G_\Lambda + K.$$

Suppose on the way of contradiction that the desired conclusion fails. By induction on $i < \omega_1$ we define a sequence

$$\langle Y_i, r_{i,1}, r_{i,2}, \eta_{i,1}, \eta_{i,2} : i < \omega_1 \rangle$$

such that:

(†) (a) $Y_i = \Lambda \cup \{\eta_{j,\ell} : j < i, \ell \in \{1,2\}\}$,
(b) $r_{i,1}, r_{i,2} \in R \setminus \{0\}$,
(c) $\eta_{i,\ell} \in \Omega_{\boldsymbol{c}} \setminus Y_i$ for $\ell \in \{1,2\}$,
(d) $m^{r_{i,1}}_{\eta_{i,1}} \neq m^{r_{i,2}}_{\eta_{i,2}}$.

The construction is easy, but we elaborate. Let us start with the case $i = 0$. We set $Y_0 = \Lambda$ and then choose $r_{0,1}, r_{0,2} \in R \setminus \{0\}$ and $\eta_{0,1}, \eta_{0,2} \in \Lambda_{<\omega}[\lambda] \setminus \Lambda_h$ such that $m^{r_{0,1}}_{\eta_{0,1}} \neq m^{r_{0,2}}_{\eta_{0,2}}$. Now suppose $i < \omega_1$ and we have define the sequence for all $j < i$. Define Y_i as in clause (†)(a). By our assumption, we can find

(i) $r_{i,1}, r_{i,2} \in R \setminus \{0\}$ and

(ii) $\eta_{i,1}, \eta_{i,2} \in \Omega_{\boldsymbol{c}} \setminus Y_i$,

so that $m^{r_{i,1}}_{\eta_{i,1}} \neq m^{r_{i,2}}_{\eta_{i,2}}$. This completes the induction construction.

Let

$$f : \Lambda_{\boldsymbol{c},<\omega} \to |R| + |K| + \aleph_0 < \lambda$$

be such that if $r \in R$ and $\eta \in \Omega_{\boldsymbol{c}}$, then $f(rx_\eta)$ is defined in a way that one can compute m^r_η and b^r_η. Again we can define f as

$$f = f_1 \circ f_2 \circ f_3,$$

where

- $f_1 : R \times (G_\Lambda + K) \to | R | + | K | + \aleph_0$ is a bijection,
- $f_2 : R \times \Lambda_{\boldsymbol{c},<\omega} \to R \times (G_\Lambda + K)$ is defined as $f_2(r, \eta) = (m^r_\eta, b^r_\eta)$,
- $f_3 : \Lambda_{\boldsymbol{c},<\omega} \to R \times \Lambda_{\boldsymbol{c},<\omega}$ is a bijection.

For each $n < \omega$, we set

$$a_n := r_{n,1}\, x_{\eta_{n,1}} - r_{n,2}\, x_{\eta_{n,2}}.$$

Applying h to it yields

$$(+) \qquad h(a_n) = m^{r_{n,1}}_{\eta_{n,1}} x_{\eta_{n,1}} - m^{r_{n,2}}_{\eta_{n,2}}\, x_{\eta_{n,2}} + b_n,$$

where $b_n := b^{r_{n,1}}_{\eta_{n,1}} - b^{r_{n,1}}_{\eta_{n,1}}$. Since $\boldsymbol{c}$ is full, there is an $\eta \in \Lambda_{\boldsymbol{c},\omega}$ such that

(1) $a_n = a^{\boldsymbol{c}}_{\eta,n}$, and

(2) $f(\eta{\restriction}_L n) = f(\eta{\restriction}_R n)$

for all $n < \omega$. By clause (2) we deduce:

(3) $\mathrm{supp}_\circ(h(x_{\eta{\restriction}_L n} - x_{\eta{\restriction}_R n})) = \varnothing$ for all $n < \omega$.

Applying h to

$$y_{\eta,0} = \sum_{i=0}^{n} a_i + \left(\prod_{i=1}^{n} i!\right) y_{\eta,n+1} + \sum_{i=0}^{n}\left(\prod_{j<i} j!\right)(x_{\eta\restriction_L i} - x_{\eta\restriction_R i}),$$

yields that

$$(\natural)\quad h(y_{\eta,0}) = \sum_{i=0}^{n} h(a_i) + \left(\prod_{i=1}^{n} i!\right) h(y_{\eta,n+1}) + \left(\prod_{j<i} j!\right)\big(h(x_{\eta\restriction_L n}) - h(x_{\eta\restriction_R n})\big)$$
$$\overset{(3)}{=} \sum_{i=0}^{n} h(a_i) + \left(\prod_{i=1}^{n} i!\right) h(y_{\eta,n+1})$$
$$\overset{(+)}{=} \sum_{i=0}^{n} (m_{\eta_{n,1}}^{r_{n,1}} x_{\eta_{n,1}} - m_{\eta_{n,2}}^{r_{n,2}} x_{\eta_{n,2}} + b_n) + \left(\prod_{i=1}^{n} i!\right) h(y_{\eta,n+1}).$$

Let $n < \omega$ be large enough. Here, we are going to apply the arguments taken from items (3)–(4) in the proof of Lemma 3.35 to the displayed formula $(\natural)$. Then,

(4) $\mathrm{supp}_\circ(h(y_{\eta,0})) \supseteq \mathrm{supp}_\circ(h(a_n))$, and

(5) $\mathrm{supp}_\circ(h(a_n)) \cap \{\eta_{n,1}, \eta_{n,2}\} \neq \varnothing$.

Without loss of generality, assume that for each $n < \omega$, $\eta_{n,1} \in \mathrm{supp}_\circ((h(a_n))$. So,

$$\{\eta_{n,1} : n < \omega\} \subseteq \mathrm{supp}_\circ(h(y_{\eta,0})),$$

which is infinite. This is a contraction. □

Lemma 3.42. *Assume $\Lambda = \mathrm{cl}(\Lambda) \subseteq \Lambda_{\boldsymbol{c}}$ is countable and $h \in \mathrm{Hom}(G_{\boldsymbol{c}}, G_\Lambda + K)$. Then h is bounded.*

Proof. Towards a contradiction we assume that h is unbounded. It follows from Lemma 3.32 that $\mathrm{Rang}(h) \nsubseteq K$. Let $b_* \in \mathrm{Rang}(h) \setminus K$. Then, for some $d_* \in K$, a finite set Λ_* and two sequences $\langle r_\eta \in R \setminus \{0\} : \eta \in \Lambda_*\rangle$ and $\langle m_\eta \in \omega : \eta \in \Lambda_*\rangle$, we can represent b_* as

$$b_* = \sum\{r_\eta x_\eta : \eta \in \Lambda_* \cap \Lambda_{<\omega}\} + \sum\{r_\eta y_{\eta,m(\eta)} : \eta \in \Lambda_* \cap \Lambda_\omega\} + d_*.$$

Let

(1) $J_0 = G_\Lambda + K$,

(2) $J_1 = J_0/K$, which is torsion free.

So, $b_* \in J_0$. Let $\pi : J_0 \to J_1$ be the natural map defined by the assignment $b \mapsto \pi(b) := b + K$. Since $b_* \in \mathrm{Rang}(h) \setminus K$, we have $\pi(b_*) \neq 0$.

Suppose on the way of contradiction that for any sequence $\langle e_n : n < \omega\rangle \in {}^{\omega}\mathbb{Z}$ the following system of equations

$$\Gamma := \{y_n = n!\, y_{n+1} + e_n b_* : n < \omega\}$$

is solvable in J_1. Say, for example, $\{y_n : n < \omega\}$ is such a solution.

Thanks to Lemma 3.26(3)(a) we find that $\Lambda_{\boldsymbol{c}}$ is $\aleph_1$-free. We combine this with Lemma 3.26(3)(b) to deduce that $M_{\boldsymbol{c}}$ is $\aleph_1$-free as an R-module. Now, since J_1 is countably generated, we can find a solution to

$$\Gamma = \{y_n = n!\, y_{n+1} + e_n\, \bar{b}_* : n < \omega\}$$

in R. Since R is cotorsion-free, a such system of equations has no solution the ring. So, there is a sequence $\langle e_n : n < \omega\rangle \in {}^{\omega}\mathbb{Z}$ the following equations:

$$\Gamma = \{y_n = n!\, y_{n+1} + e_n\, b_* : n < \omega\}$$

is not solvable in J_1.

Let $a_* \in G_{\boldsymbol{c}}$ be such that $b_* = h(a_*)$. Let also $f : \Lambda_{\boldsymbol{c},<\omega} \to \omega$ be such that for all $\nu, \rho \in \Lambda_{\boldsymbol{c},<\omega}$,

$$f(\nu) = f(\rho) \Longleftrightarrow \pi \circ h(x_\nu) = \pi \circ h(x_\rho).$$

As $\boldsymbol{c}$ is full, there is some $\eta \in \Lambda_{\boldsymbol{c},\omega}$ such that:

(3) $a^{\boldsymbol{c}}_{\eta,n} = e_n\, a_*$, for all $n < \omega$, and

(4) $f(\eta \restriction_L n) = f(\eta \restriction_R n)$, for $n < \omega$.

Thanks to (4), one has

$$(+) \qquad \forall n < \omega, \quad \pi \circ h(x_{\eta\restriction_L n}) = \pi \circ h(x_{\eta\restriction_R n})$$

By applying $\pi \circ h$ into the equation

$$y_{\eta,n} = a^{\boldsymbol{c}}_{\eta,n} + n!\, y_{\eta,n+1} + (x_{\eta\restriction_L n} - x_{\eta\restriction_R n}),$$

and using clause (3) and (+) we get

$$\pi \circ h(y_{\eta,n}) = e_n\, \pi(b_*) + n!\, \pi \circ h(y_{\eta,n+1}).$$

This clearly gives a contradiction, as then

$$J_1 \models y_n = n!\, y_{n+1} + e_n\, b''_*,$$

where $y_n = \pi \circ h(y_{\eta,n})$. □

Lemma 3.43. *Let* $\boldsymbol{c}$ *be full and* $h \in \mathrm{End}(G_{\boldsymbol{c}})$. *Then* $\mathrm{Rang}(h)$ *is bounded.*

Proof. Suppose not, it follows that for some countable $\Lambda = \mathrm{cl}(\Lambda) \subseteq \Lambda_{\boldsymbol{c}}$,

$$h \restriction G \in \mathrm{Hom}(G, G_\Lambda + K)$$

is unbounded, where G is the subgroup of $G_{\boldsymbol{c}}$ generated by $h^{-1}[G_\Lambda + K]$. This contradicts Lemma 3.42. □

Now, we are ready to prove:

Theorem 3.44. *Adopt the notation from Hypothesis 3.14. Then there is some* $\boldsymbol{c}$ *such that the abelian group* $G_{\boldsymbol{c}}$ *is boundedly rigid. In particular, there is an abelian group* G *equipped with the following properties*:

(1) $\mathrm{tor}(G) = K$.

(2) *G is of size* λ.

(3) *The sequence*

$$0 \to R_{\boldsymbol{c}} \to \mathrm{End}(G) \to \frac{\mathrm{End}(G)}{\mathrm{E_b}(G)} \to 0$$

is exact.

Proof. According to Lemma 3.30, there is a full $\boldsymbol{c} \in \mathrm{AP}$. This allows us to apply Lemma 3.43, and deduce that $G := G_{\boldsymbol{c}}$ is boundedly rigid. By definition, this completes the proof. □

4. Co-Hopfian and boundedly endo-rigid abelian groups

As stated in [15], it is difficult to construct an infinite Hopfian–co-Hopfian p-group. What about mixed groups? In this section, we answer this question. We start by recalling that a group G is called:

(i) *Hopfian* if its surjective endomorphisms are automorphisms.

(ii) *co-Hopfian* if its injective endomorphisms are automorphisms.

In what follows we will use the following two items.

Fact 4.1. (i) *Any direct summand of a co-Hopfian abelian group is again co-Hopfian.*

(ii) *Suppose* $2^{\aleph_0} < \lambda < \lambda^{\aleph_0}$. *Then there is no co-Hopfian abelian group of size* λ (*see* [19, Theorem 1.2]).

Here, we introduce a useful criterion.

Definition 4.2. Let G be an abelian group of size λ and $m, n \geq 1$ be such that $m \mid n$.

(1) $\mathrm{NQr_{(m,n)}(G)}$ means that there is an (m, n)-antiwitness h such that
   (a) $h \in \mathrm{End}(\Gamma_n(G))$,
   (b) $\mathrm{Rang}(h)$ is a bounded group,
   (c) $h' := m \cdot \mathrm{id}_{\Gamma_n(G)} + h \in \mathrm{End}(\Gamma_n(G))$ is 1-to-1,
   (d) h' is not onto or $m > 1$ and $G/\Gamma_n(G)$ is not m-divisible.

(2) $\mathrm{NQr_m(G)}$ means $\mathrm{NQr_{(m,n)}(G)}$ for some $n \geq 1$.

(3) $\mathrm{NQr(G)}$ means $\mathrm{NQr_m(G)}$ for some $m \geq 1$.

Definition 4.3. Adopt the previous notation.

(1) $\mathrm{Qr(G)}$ means the negation of $\mathrm{NQr(G)}$.

(2) $\mathrm{Qr}_*(\mathrm{G})$ means Qr(G) and in addition that $\Gamma_p(G)$ is unbounded, for at least one $p \in \mathbb{P}$.

In items 4.4–4.11 we check $\mathrm{NQr}_{(\mathrm{m,n})}(\mathrm{G})$ and its negation. This enables us to present some new classes of co-Hopfian and non-co-Hopfian groups.

Lemma 4.4. *Let G be an abelian group such that the property* NQr(G) *holds. Then G is not co-Hopfian. Furthermore, let $h \in \mathrm{Hom}(G, \Gamma_n(G))$ be such that $h \restriction \Gamma_n(G)$ is an (m, n)-antiwitness. Then $m \cdot \mathrm{id}_G + h$ witnesses that G is not co-Hopfian.*

Proof. Suppose that G admits an (m, n)-antiwitness $h_0 \in \mathrm{End}(\Gamma_n(G))$ as in Definition 4.2. As h_0 is bounded, by Fact 2.14 we extend h_0 to $h_1 \in \mathrm{Hom}(G, \Gamma_n(G))$. So, the following diagram commutes:

$$\begin{tikzcd} 0 \arrow[r] & \Gamma_n(G) \arrow[r, "\subseteq_*"] \arrow[d, "h_0"'] & G \arrow[dl, "\exists h_1"] \\ & \Gamma_n(G) & \end{tikzcd}$$

We claim that $f = m \cdot \mathrm{id_G} + \mathrm{h_1} \in \mathrm{End(G)}$ is 1-to-1 but not onto.

$(*_1)$ f is one-to-one.

To see this, suppose $x \in G$ in nonzero and we want to show that $f(x) \neq 0$. Suppose first we deal with the case $x \in \Gamma_n(G) \setminus \{0\}$. According to Definition 4.2(1)(c), we have

$$f(x) = mx + h_1(x) = m \cdot \mathrm{id}_{\Gamma_n(G)}(x) + h_0(x) \Rightarrow f(x) \neq 0.$$

Now, suppose that $x \in G \setminus \Gamma_n(G)$. Recall from Definition 4.2 that m divides n. As $m \mid n$, we have $mx \in G \setminus \Gamma_n(G)$. If $f(x) = 0$, we have $mx + h_1(x) = 0$, thus

$$h_1(x) = -mx \in G \setminus \Gamma_n(G).$$

But, $\mathrm{Rang}(h_1) \subseteq \Gamma_n(G)$, which is impossible. Thus f is 1-to-1, as wanted.

$(*_2)$ f is not onto.

For this, we consider two cases.

Case 1. *h_0 is not onto.*

By the case assumption, there is

$$y \in \Gamma_n(G) \setminus \mathrm{Rang}\big(\mathrm{id}_{\Gamma_n(G)} + (\mathrm{h_0} \restriction \Gamma_n(G))\big)$$

and it is easy to see that such a y is also a witness for f to be not onto.

Case 2. *h_0 is onto.*

By Definition 4.2(1)(d), we must have $m > 1$ and $G/\Gamma_n(G)$ is not m-divisible. Let $z \in G$ be such that $z + \Gamma_n(G)$ is not divisible by m in $G/\Gamma_m(G)$. Clearly, z does not belong to $\mathrm{Rang}(f)$.

The lemma follows. □

Lemma 4.5. *Let K be an abelian p-group. The following claims are valid: If* $\mathrm{NQr}(K)$ *holds, then K is infinite.*

Proof. By definition, there are m and n such that $m \mid n$ and that $\mathrm{NQr}_{(m,n)}(K)$ holds. Thanks to Definition 4.2(1), there is $h \in \mathrm{End}(\Gamma_n(G))$ satisfying the following properties:

(a) $\mathrm{Rang}(h)$ is a bounded group.

(b) $h' := m \cdot (\mathrm{id}_{\Gamma_n(K)}) + \mathrm{h} \in \mathrm{End}(\Gamma_n(K))$ is 1-to-1.

(c) h' is not onto or $m > 1$ and $K / \Gamma_n(K)$ is not m-divisible.

We have two possibilities: (1) $p \nmid n$ and (2) $p \mid n$.

(1) Suppose first that $p \nmid n$. As K is a p-group, $\Gamma_n(K) = \{0\}$. This means that h is constantly zero and is onto, as well as h'. Thanks to clause (c) it follows that $m > 1$ and K is not m-divisible. Since $m \mid n$ we deduce that $p \nmid m$. Now, we consider the map $m \cdot \mathrm{id}_K : K \to K$. Since K is not m-divisible, this map is not surjective. Let us show that it is 1-to-1. To this end, let $x \in K$ be such that $mx = 0$. Let ℓ be the order of x so that $p^\ell x = 0$. As $(p^\ell, m) = 1$, we can find r, s such that $rp^\ell + sm = 1$. By multiplying both sides with x, we obtain

$$x = rp^\ell x + smx = 0 + 0 = 0.$$

It follows that $m \cdot \mathrm{id}_K : K \to K$ is 1-to-1 and not onto, hence K is infinite.

(2) Suppose $p \mid n$. As K is a p-group, this implies that $\Gamma_n(K) = K$. Therefore, in the above item (c), the case "$K / \Gamma_n(K)$ is not m-divisible" does not occur. This is in turn implies that h' is not onto K. We proved that the map $h' \in \mathrm{End}(K)$ is 1-to-1 and not onto. Hence K is infinite. □

Discussion 4.6. Keep the notation of Fact 2.5. One cannot replace "divisible" with "reduced" and drives a similar result, as some easy examples suggest this. Here, we consider this as an application of the construct of co-Hopfian groups.

(1) Suppose on the way of contradiction that the replacement is valid.

(2) Let G be a co-Hopfian group such that its reduced part is unbounded (recall from the introduction that a such group exists, see [9]).

(3) Here, we drive a contradiction by showing from that G is not co-Hopfian. Indeed, let K_2 be the maximal divisible subgroup of K. Recall from Fact 2.18 that K_2 is injective. Since it is injective, we know K_2 is a directed summand. Let us write K as $K = K_1 \oplus K_2$. Due to the maximality of K_2 one may know that K_1 is reduced. We show that K_1 is not co-Hopfian, and hence by Fact 4.1(i), K is not co-Hopfian. Thus by replacing K by K_1 if necessary, we may assume without loss

of generality that K is reduced and unbounded. For $\ell < \omega$, we choose by induction H_ℓ, y_ℓ and z_ℓ such that:

(a) $H_0 = K$.

(b) If $\ell = k+1$, then $H_k = H_\ell \oplus \mathbb{Z}z_\ell$.

(c) For $z_\ell \in (\mathbb{Z}y_\ell)_*$, recall that $(\mathbb{Z}y_\ell)_*$ denotes the pure closure of $\mathbb{Z}y_\ell$.

(d) $y_{\ell+1} \in H_\ell$.

(e) The order of z_i is $\geq p^\ell$.

[Why? For $\ell = 0$, we set $H_0 = K$ and let $y_0 \in K$ be arbitrary. Then $(\mathbb{Z}y_0)_*$ is a pure subgroup of K of bounded exponent. Thanks to Fact 2.5, we know that $(\mathbb{Z}y_0)_*$ is a direct summand of K. In view of Fact 2.7 we can find z_0 such that $\mathbb{Z}z_0$ is a direct summand of $(\mathbb{Z}y_0)_*$. In other words, $\mathbb{Z}z_0$ is a direct summand of $H_0 = K$ as well. Consequently, we have $H_0 = H_1 \oplus \mathbb{Z}z_0$ for some H_1. Having defined inductively $\{H_\ell, y_\ell, z_\ell\}$, let $y_{\ell+1} \in H_\ell$. Let χ be a regular cardinal, large enough, so that $H_\ell \in \mathscr{H}(\chi)$. The notation $\mathscr{B}$ stands for $(\mathscr{H}(\chi), \in)$. Let $\mathscr{B}_\ell$ be countable such that $H_\ell \in \mathscr{B}_\ell$. Now, we look at

$$\mathcal{L}_\ell := \mathscr{B}_\ell \cap H_\ell.$$

We find easily that $\mathcal{L}_\ell$ is an unbounded countable abelian p-group. Hence it is of the form $\oplus_i \mathbb{Z}z_{\ell,i}$ where $z_{\ell,i}$ is of order $p^{m(\ell,i)}$. As $\mathcal{L}_\ell$ is unbounded, we may and do assume that $m(\ell, i) > \ell$. This implies that $\mathbb{Z}z_{\ell,i}$ is a pure subgroup of $\mathcal{L}_\ell$, and hence H_ℓ. Consequently, $\mathbb{Z}z_{\ell,i}$ is a direct summand of H_ℓ as well. By definition, we have $H_\ell = H_{\ell+1} \oplus \mathbb{Z}z_{\ell+1}$ for some abelian subgroup $H_{\ell+1}$ of H_ℓ.]

For each $i < \omega$, we let $\ell(i) > 1$ be such that z_i is of order $p^{\ell(i)}$. Following (e), clearly we can find some infinite $u \subseteq \omega$ such that the sequence $\langle \ell(i) : i \in u\rangle$ is increasing. For any $j < \omega$, we clearly have $\bigoplus_{i \in u \cap j} \mathbb{Z}z_i \subseteq_* K$, and hence $\bigoplus_{i \in u} \mathbb{Z}z_i \subseteq_* K$. In light of part (i), $\bigoplus_{i \in u} \mathbb{Z}z_i$ is a direct summand of K. Thus there is some K_3 such that $K = \bigoplus_{i \in u} \mathbb{Z}z_i \oplus K_3$. Assume that $\langle j(k) : k < \omega\rangle$ lists u in an increasing order, and define $h \in \mathrm{End}(K)$ such that

- $h \restriction K_3 = \mathrm{id}_{K_3}$,
- $h(z_{j(k)}) = p^{\ell(k+1)-1} z_{j(\ell+1)}$.

It is easy to check that h is a well-defined endomorphism of K and it satisfies

- h is injective,
- h is not surjective.

In sum, h witnesses that K is not co-Hopfian, a contradiction we searched for.

Corollary 4.7. *Let G be a p-group such that its reduced part is unbounded and its countable pure subgroups are directed summand. Then G is not co-Hopfian.*

Lemma 4.8. *Let G be an abelian group of size λ and $m \geq 1$. Suppose there is a bounded $h \in \mathrm{End}(G)$ such that $f := m \cdot \mathrm{id}_G + h \in \mathrm{End}(G)$ is 1-to-1 not onto.*[4] *Then for some $n \geq 1$ we have*:

(i) $\mathrm{NQr}_{(m,n)}(G)$.

(ii) *Letting $h_0 = h \restriction \Gamma_n(G)$, h_0 is an (m, n)-antiwitness for $\Gamma_n(G)$.*

Proof. Let f and h be as above. As $\mathrm{Rang}(h)$ is bounded, for some $n \geq 1$ we have $\mathrm{Rang}(h) \leq \Gamma_n(G)$ and without loss of generality $m \mid n$. Possibly, replacing n with nm, which is possible as $n_1 \mid n_2$ implies that $\Gamma_{n_1}(G) \leq \Gamma_{n_2}(G)$. Notice that:

$(*_1)$ (a) f maps $\Gamma_n(G)$ into itself.
(b) If $x \in G \setminus \Gamma_n(G)$, then $f(x) \notin \Gamma_n(G)$.

Clause (a) clearly holds as by the choice of n we have $\mathrm{Rang}(h) \leq \Gamma_n(G)$. For (b), we suppose by contradiction that $f(x) = mx + h(x) \in \Gamma_n(G)$. It follows that $mx = f(x) - h(x) \in \Gamma_n(G)$, and hence as $m \mid n$, $x \in \Gamma_n(G)$, a contradiction.

Now let $h_0 = h \restriction \Gamma_n(G)$. Then we have:

$(*_2)$ (a) $h_0 \in \mathrm{End}(\Gamma_n(G))$.
(b) h_0 is bounded.
(c) Since f is 1-to-1, so is $f_0 = m \cdot \mathrm{id}_{\Gamma_n(G)} + h_0 \in \mathrm{End}(\Gamma_n(G))$.

We are left to show that h_0 is an (m, n)-antiwitness. By $(*_2)$ it suffices show that f_0 is not onto or $G/\Gamma_n(G)$ is not m-divisible. Suppose on the contrary that f_0 is onto and $G/\Gamma_n(G)$ is m-divisible. We are going to show that f is onto, which contradicts our assumption. To this end, let $x \in G$. Since $G/\Gamma_n(G)$ is m-divisible, we can find some $y \in G$ such that

$$x - my \in \Gamma_n(G).$$

We look at

$$w := x - my - h_0(y) \in \Gamma_n(G).$$

As f_0 is onto, we can find some $z \in \Gamma_n(G)$ such that $f_0(z) = w$. So,

$$x - my - h_0(y) = w = f_0(z) = mz + h_0(z).$$

Using this equation, and the additivity of h_0, we observe that

$$x = m(y + z) + h_0(y + z) = f(y + z).$$

In other words, f is onto. This is a contradiction. □

Notation 4.9. Let κ and μ be infinite cardinals. The infinitary language $\mathcal{L}_{\mu,\kappa}(\tau)$ is defined so as its vocabulary is the same as τ, it has the same terms and atomic formulas as in τ, but we also allow conjunction and disjunction of length less than μ,

[4] Thus f witnesses non-co-Hopfianity of G.

i.e., if ϕ_j, for $j < \beta < \mu$ are formulas, then so are $\bigvee_{j<\beta} \phi_j$ and $\bigwedge_{j<\beta} \phi_j$. Also, quantification over less than κ many variables.

Lemma 4.10. *Let G be a reduced abelian group of size λ such that*

(1) $\lambda > 2^{\aleph_0}$,

(2) *G is co-Hopfian.*

Then the property $\mathrm{Qr}_*(G)$ *is valid.*

Proof. Thanks to Lemma 4.4 we know $\mathrm{Qr}(G)$ is satisfied, so it is enough to show that $\Gamma_p(G)$ is not bounded for some prime p. Towards a contradiction, we suppose that $\Gamma_p(G)$ is bounded for every prime $p \in \mathbb{P}$.

Here, we are going to show that the pure subgroup $\Gamma_p(G)$ is finite. Suppose on the way of contradiction that $\Gamma_p(G)$ is infinite. Recall that p-torsion subgroups are pure. According to Fact 2.4, $\Gamma_p(G)$ is a direct summand of G, as we assumed that it is bounded. Also, following Fact 2.7 we know that $\Gamma_p(G)$ is a direct summand of cyclic groups. In sum, we observed that $\Gamma_p(G)$ has a direct summand K which is a countably infinite p-group. In view of Fact 2.6(i), we may and do assume that K is not co-Hopfian. Recall that any direct summand of co-Hopfian, is co-Hopfian. This means that G is not co-Hopfian as well, which contradicts our assumption. Thus, it follows that for every $p \in \mathbb{P}$, the group $\Gamma_p(G)$ is finite and therefore a direct summand of G, and hence there is a projection h_p from G onto $\Gamma_p(G)$. Recall that $p \in \mathbb{P}$ and also $h_p \restriction \Gamma_p(G) \in \mathrm{End}(\Gamma_p(G))$ is essentially equal to the identity map, so is one-to-one, and hence onto, as $\Gamma_p(G)$ is finite. Since $\mathrm{Qr}(G)$ is satisfied, it follows from Definition 4.2(1)(d) that $G/\Gamma_p(G)$ is p-divisible.

Now, we take χ be a regular cardinal, large enough, such that $G \in \mathscr{H}(\chi)$ and let

$$M \prec_{\mathcal{L}_{\aleph_1,\aleph_1}} (\mathscr{H}(\chi), \in)$$

be such that

- M has cardinality $2^{\aleph_0}$,
- $G, \mathrm{tor}(G) \in M$,
- $2^{\aleph_0} + 1 \subseteq M$.

In light of Fact 2.6(ii), we may and do assume that $|\mathrm{tor}(G)| = \mu \leq 2^{\aleph_0}$. Recall that $2^{\aleph_0} + 1 \subseteq M$ and $\mathrm{tor}(G) \in M$. These imply that $\mathrm{tor}(G) \subseteq M$. Now, as $G/\Gamma_p(G)$ is p-divisible, then so is

$$\frac{G/\Gamma_p(G)}{(G \cap M)/\Gamma_p(G)},$$

which by the third isomorphism theorem, is canonically isomorphic to $G/G \cap M$. As $\mathrm{tor}(G) \subseteq M$, $G/(G \cap M)$ is torsion-free, it is divisible. Let $x \in G \setminus M$ and define the sequence $(x_n : n < \omega)$ such that

- $x_0 = x$,
- if $n = m + 1$ then

$$G/(G \cap M) \models n!\, x_n + (G \cap M) = x_m + (G \cap M)''.$$

So, letting $a_0 = 0$ and for $n = m + 1 < \omega$,

$$a_n = n!\, x_n - x_m \in G \cap M,$$

we have that $(a_n : n < \omega) \in M^\omega \subseteq M$ and so, as

$$M \prec_{\mathcal{L}_{\aleph_1,\aleph_1}} (\mathscr{H}(\chi), \in),$$

we can find

$$\bar{y} = (y_n : n < \omega) \in (G \cap M)^\omega$$

such that $a_n = n!\, y_n - y_m$, but then for every $m < \omega$ we have

$$G \models m!\,(x_{m+1} - y_{m+1}) = x_m - y_m''.$$

Hence,

$$\bigcup\{\mathbb{Z}(x_m - y_m) : m < \omega\}$$

is a nontrivial divisible subgroup of G, contradicting the assumption that G is reduced. So we have proved the desired claim. $\square$

Proposition 4.11. *Let $G \in$ be a boundedly endo-rigid abelian group. The following assertions are valid*:

(1) *G is co-Hopfian if and only if* $\mathrm{Qr}(G)$.

(2) *If $|G| > 2^{\aleph_0}$, then G is co-Hopfian if and only if* $\mathrm{Qr}_*(G)$.

Proof. (1) If G is co-Hopfian, then by Lemma 4.4, $\mathrm{Qr}(G)$ holds. For the other direction, suppose that G is boundedly rigid and $\mathrm{Qr}(G)$ holds. Let $f \in \mathrm{End}(G)$ be 1-to-1, we want to show that f is onto. As G is boundedly rigid we have m, h and L such that

(a) $m \in \mathbb{Z},\ h \in \mathrm{End(G)}$,

(b) $f(x) = mx + h(x)$,

(c) $L = \mathrm{Rang}(h)$ is a bounded subgroup of G (and so of $\mathrm{tor}(G)$).

If f is not onto, then by Lemma 4.8, there is $n \geq 1$ such that $\mathrm{NQr}_{(m,n)}(G)$ holds, which is not possible (as we are assuming $\mathrm{Qr}(G)$). Thus f is onto as required.

(2) It follows from clause (1) and Lemma 4.10. $\square$

Construction 4.12. Let $K := \oplus\{\frac{\mathbb{Z}}{p^n\mathbb{Z}} : p \in \mathbb{P} \text{ and } 1 \leq n < m\}$, where $m < \omega$, and $\mathbb{P}$ is the set of prime numbers. Let G be a boundedly endo-rigid abelian group such that $\operatorname{tor}(G) = K$.[5] Then G is co-Hopfian.

Proof. For any $p_1 \in \mathbb{P}$ and $n_1 < m$, let us define

$$(x_{(p_1,n_1)})_{(p,n)} = \begin{cases} 1 + p^n\mathbb{Z} & \text{if } (p,n) = (p_1,n_1), \\ 0, & \text{otherwise.} \end{cases}$$

For simplicity, we abbreviate it by $x_{(p_1,n_1)}$. Assume towards a contradiction that there exists $f \in \operatorname{End}(G)$ such that f is 1-to-1 and not onto. As G is boundedly endo-rigid, there are $m \in \mathbb{Z}$ and $h \in \mathrm{E_b}(G)$ such that $f = m \cdot \mathrm{id}_G + h$. As f is 1-to-1 and K has no infinite bounded subgroup, we can conclude that $m \neq 0$.

$(*_1)$ $m \in \{1, -1\}$.

To see $(*_1)$, suppose on the contrary that there is $p \in \mathbb{P}$ such that $p \mid m$ and let m_1 be such that $m = m_1 p$. Now, as $\operatorname{Rang}(h)$ is bounded, there is $k \geq 1$ such that

$$p^k(\operatorname{Rang}(h)) \cap \Gamma_p(G) = \{0\}.$$

Let $n \geq k + 1$, then

$$\begin{aligned} f(p^{n-1}x_{(p,n)}) &= mp^{n-1}x_{(p,n)} + h(p^{n-1}x_{(p,n)}) \\ &= m_1pp^{n-1}x_{(p,n)} + p^kh(p^{n-1-k}x_{(p,n)}) = 0, \end{aligned}$$

which contradicts the fact that f is 1-to-1. This completes the argument of $m \in \{1, -1\}$ and without loss of generality we may assume that $m = 1$. Thus $f = \mathrm{id}_G + h$.

$(*_2)$ f maps $G \setminus \operatorname{tor}(G)$ into itself.

This is because f is 1-to-1. Indeed let $x \in G \setminus \operatorname{tor}(G)$. If $f(x) \in \operatorname{tor}(G)$, then $f(kx) = kf(x) = 0$ for some k, thus $kx = 0$, i.e., $x \in \operatorname{tor}(G)$ which contradicts $x \in G \setminus \operatorname{tor}(G)$.

$(*_3)$ $f \restriction \operatorname{tor}(G) \in \operatorname{End}(\operatorname{tor}(G))$ is 1-to-1 not onto.

Clearly $f \restriction \operatorname{tor}(G) \in \operatorname{End}(\operatorname{tor}(G))$, and since f is 1-to-1, $f \restriction \operatorname{tor}(G)$ is 1-to-1 as well. Now, suppose by contradiction that $f \restriction \operatorname{tor}(G)$ is onto. Then

(1) $\operatorname{tor}(G) \subseteq \operatorname{Rang}(f)$,

(2) $x \in G \Rightarrow f(x) = x + h(x) \in \operatorname{tor}(G)$.

Recall that $h(x) \in \operatorname{tor}(G)$. Apply this along with (1), we deduce that $h(x) \in \operatorname{Rang}(f)$. Also, recall that $\operatorname{Rang}(f)$ is a group. Let $x \in G$. Thanks to (2), we observe that

$$x = f(x) - h(x) \in \operatorname{Rang}(f).$$

[5] In light of our main result, such a group exists for any $\lambda = \lambda^{\aleph_0} > 2^{\aleph_0}$ and the size of G should be λ.

In other words, f is onto, a contradiction. So, $f \restriction \mathrm{tor}(G)$ is not onto.

$(*_4)$ (a) For every $p \in \mathbb{P}$, f maps $\Gamma_p(G)$ into itself and so $f \restriction \Gamma_p(G)$ is 1-to-1.
(b) For some $p \in \mathbb{P}$, $f \restriction \Gamma_p(G)$ is not onto.

Item (a) above is simply because f is 1-to-1. To see (b) holds, note that if $f \restriction \Gamma_p(G)$ is onto for all prime number p, then so is $f \restriction \mathrm{tor}(G)$, which contradicts $(*_3)$.

Thus, let us fix some prime $p \in \mathbb{P}$ such that $f \restriction \Gamma_p(G)$ is not onto and let $h_p = h \restriction \Gamma_p(G)$. Then by the above observations, it equipped with the following properties:

$(*_5)$ (a) $h_p \in \mathrm{End}(\Gamma_\mathrm{p}(\mathrm{G}))$.
(b) $\mathrm{Rang}(h_p)$ is bounded.
(c) $h'_p = m \cdot \mathrm{id}_{\Gamma_p(G)} + h_p = \mathrm{id}_{\Gamma_p(G)} + h_p$ is 1-to-1.
(d) h'_p is not onto.

In light of Definition 4.2 and $(*_5)$ we observe that

$(*_6)$ h_p is a $(1, p)$-antiwitness for $\Gamma_p(G)$ and so $\mathrm{NQr}(\Gamma_p(G))$.

Thanks to Lemma 4.5, $\Gamma_p(G)$ is infinite. But,

$$\Gamma_p(G) = \Gamma_p(K) = \bigoplus \left\{ \frac{\mathbb{Z}}{p^n\mathbb{Z}} : 1 \leq n < m \right\},$$

which is finite. Thus we get a contradiction, and hence f is onto. It follows that G is co-Hopfian and the lemma follows. □

Corollary 4.13. *For any cardinals $\lambda > 2^{\aleph_0}$, there is a co-Hopfian abelian group G of size λ if and only if $\lambda = \lambda^{\aleph_0}$.*

Proof. Let $\lambda > 2^{\aleph_0}$ be given. Suppose first that $\lambda < \lambda^{\aleph_0}$. In other words, $2^{\aleph_0} < \lambda < \lambda^{\aleph_0}$. According to Fact 4.1(ii), there is no co-Hopfian abelian group of size λ. Now, assume that $\lambda = \lambda^{\aleph_0}$. Let

$$K := \oplus \left\{ \frac{\mathbb{Z}}{p^n\mathbb{Z}} : p \in \mathbb{P} \text{ and } 1 \leq n < m \right\},$$

where $m < \omega$. In light of Theorem 3.11, there exists a boundedly endo-rigid abelian group G with $\mathrm{tor}(G) = K$. By Construction 4.12, G is co-Hopfian. □

Lemma 4.14. *Let $G = G_1 \oplus G_2$ be a boundedly endo-rigid abelian group. Then G_1 is boundedly endo-rigid.*

Proof. Let $f_1 \in \mathrm{End}(G_1)$. Then $f_1 \oplus \mathrm{id}_{G_2} \in \mathrm{End}(G)$. Since G is boundedly endo-rigid there is $m \in \mathbb{Z}$ such that the map $x \mapsto f(x) - mx$ has bounded range. In other words,

$$(f_1 - m \cdot \mathrm{id}_{G_1}) \oplus 0 \subseteq (f_1 - m \cdot \mathrm{id}_{G_1}) \oplus (\mathrm{id}_{G_2} - m \cdot \mathrm{id}_{G_2}) = (f - m \cdot \mathrm{id}_G)$$

has bounded range. By definition, G_1 is boundedly endo-rigid. □

Notation 4.15 (Harrison). For each group G, we set

$$S := S_G := \{p \in \mathbb{P} : G/\Gamma_p(G) \text{ is not } p\text{-divisible}\}.$$

Now, we are ready to present the following promised criteria:

Proposition 4.16. *Let $\lambda > 2^{\aleph_0}$, and suppose G is a boundedly endo-rigid abelian group of size λ. Then G is co-Hopfian if and only if*:

(a) *S_G is a nonempty set of primes.*

(b) (b_1) $\mathrm{tor}(G) \neq G$.
  (b_2) *If $p \in S$, then $\Gamma_p(G)$ is not bounded.*
  (b_3) *If $\Gamma_p(G)$ is bounded, then it is finite* (*and $p \notin S_G$*).

Proof. Let $K := \mathrm{tor}(G)$, and for each prime number p, we set $K_p := \Gamma_p(G)$.

First, we assume that G is co-Hopfian, and we are going to show items (a) and (b) are valid. As G is co-Hopfian, and recall from the introduction that Beaumont and Pierce (see [5]) proved that for the co-Hopfian group G, we know $\mathrm{tor}(G)$ is of size at most continuum. In other words, $|\mathrm{tor}(G)| \leq 2^{\aleph_0}$. We combine this along with our assumption $|G| = \lambda > 2^{\aleph_0}$, and conclude that $K = \mathrm{tor}(G) \neq G$, as claimed by ($b_1$).

To prove (b_2), let $p \in S$ and suppose by contradiction that K_p is bounded. As K_p is pure in G, and following Fact 2.4, the boundedness property guarantees that K_p is a direct summand of G. By definition, there is G_p such that $G = K_p \oplus G_p$. Now, we look at $\mathrm{id}_{K_p} + p \cdot \mathrm{id}_{G_p} \in \mathrm{End}(G)$. Let

$$(k, g) \in \mathrm{Ker}(\mathrm{id}_{K_p} + p \cdot \mathrm{id}_{G_p}).$$

Following definition, we have

$$(0, 0) = (\mathrm{id}_{K_p} + p \cdot \mathrm{id}_{G_p})(k, g) = (k, pg).$$

In other words, $k = 0$ and as G_p is p-torsion-free, $g = 0$. This means that

$$\mathrm{Ker}(\mathrm{id}_{K_p} + p \cdot \mathrm{id}_{G_p}) = 0,$$

and hence $\mathrm{id}_{K_p} + p \cdot \mathrm{id}_{G_p}$ is 1-to-1. Since $p \in S$, $G_p := G/\Gamma_p(G)$ is not p-divisible, thus there is g in G_p such that $g \notin \mathrm{Rang}(p \cdot \mathrm{id}_{G_p})$. Consequently, $\mathrm{id}_{K_p} + p \cdot \mathrm{id}_{G_p}$ is 1-to-1 not onto. This is in contradiction with the co-Hopfian assumption, so K_p is not bounded and (b_2) follows.

In order to check (b_3), suppose $K_p = \Gamma_p(G)$ is bounded. Then it is a direct summand of G, say $G = K_p \oplus G_p$. Since G is co-Hopfian, and in view of Fact 4.1, we observe that K_p is co-Hopfian. Thanks to Fact 2.6 K_p is finite.

Lastly, we check clause (a). Suppose on the way of contradiction that S is empty. Let $G_1 \prec_{\mathcal{L}_{\aleph_1,\aleph_1}} G$ be of cardinality $2^{\aleph_0}$ containing $\mathrm{tor}(G)$, recalling $|\mathrm{tor}(G)| \leq 2^{\aleph_0}$, so G/G_1 is divisible of cardinality λ.

As $G_1 \neq G$, there is $x_0 \in G \setminus G_1$, and note that $x \notin \operatorname{tor}(G)$. Now as $G/\operatorname{tor}(G)$ is divisible, we can choose the sequence $\langle x_n : n \geq 1 \rangle$ of elements of G, by induction on n, such that $x_0 = x$ and for each n,

$$G/\operatorname{tor}(G) \models n!\, x_{n+1} + \operatorname{tor}(G) = x_n + \operatorname{tor}(G)''.$$

Set

$$a_n := n!\, x_{n+1} - x_n \in \operatorname{tor}(G).$$

Note that $\langle a_n : n < \omega \rangle \in G_1$, thus as $G_1 \prec_{\mathcal{L}_{\aleph_1,\aleph_1}} G$, we can find elements $y_n \in G_1$ for $n < \omega$ such that

$$n!\, y_{n+1} = y_n + a_n.$$

Subtracting the last two displayed formulas, shows that the group

$$L = \bigcup \{\mathbb{Z}(x_n - y_n) : n < \omega\}$$

is a nonzero divisible subgroup of G. Recall from Fact 2.18 that L is injective. Since it is injective, we know L is a directed summand of its extensions. In sum, the sequence

$$0 \to L \xrightarrow{g} G \to \operatorname{coker}(g) \to 0,$$

splits. Recall from Discussion 2.20 that

$$\operatorname{End}(I) = \prod_{p \in \mathbb{P}_0} \hat{\mathbb{Z}}_p^{\oplus x_p},$$

where $\mathbb{P}_0 := \mathbb{P} \cup \{0\}$ and x_p are some index sets. This turns out that I is not boundedly endo-rigid, provided it is nonzero. Recall from Lemma 4.14 that the property of boundedly endo-rigid behaves well with respect to direct summand, it obviously implies G is not boundedly endo-rigid. This contradiction implies that S is not empty. So clause (a) holds. All together, we are done proving the left-right implication.

For the right-left implication, assume items (a) and (b) hold, and we show that G is co-Hopfian. Suppose on the way of contradiction that there exists $f \in \operatorname{End}(G)$ such that f is 1-to-1 and not onto. As G is boundedly endo-rigid, there are $m \in \mathbb{Z}$ and $h \in \mathrm{E_b}(G)$ such that $f = m \cdot \mathrm{id}_G + h$.

$(*_1)$ $m \neq 0$.

To see $(*_1)$, suppose $m = 0$. Then $f = h$, and since $\operatorname{Rang}(h)$ is bounded and f is 1-to-1, we can conclude that G is bounded and therefor $G = \operatorname{tor}(G)$. This contradicts clause (b_1).

$(*_2)$ If $\Gamma_p(G)$ is infinite, then $p \nmid m$.

In order to see $(*_2)$, first note that $\operatorname{tor}(G)$ is unbounded, as otherwise $\Gamma_p(G)$ is also bounded, and hence by (b_3) it is finite, contradicting our assumption. Suppose on the way of contradiction that $p \mid m$. Then there is m_1 such that $m = m_1 p$. Now, as Rang(h) is bounded, there exists $k \geq 1$ such that

$$p^k\big(\mathrm{Rang}(h)\restriction\Gamma_p(G)\big) = \{0\}.$$

Recall that K_p is unbounded. This gives us an element $x \in \Gamma_p(G)$ of order p^n for some $n \geq k+1$. But then

$$f(p^{n-1}x) = mp^{n-1}x + h(p^{n-1}x) = m_1 pp^{n-1}x + p^k h(p^{n-1-k}x) = 0,$$

which contradicts the fact that f is 1-to-1.

As before, we have the following properties:

($*_3$) f maps $G \setminus \mathrm{tor}(G)$ into itself.

($*_4$) $f \restriction \mathrm{tor}(G) \in \mathrm{End}(\mathrm{tor}(G))$ is 1-to-1 not onto.

($*_5$) (a) For every $p \in \mathbb{P}$, f maps $\Gamma_p(G)$ into itself and so $f \restriction \Gamma_p(G)$ is 1-to-1.
(b) For some $p \in \mathbb{P}$, $f \restriction \Gamma_p(G)$ is not onto.

Fix $p \in \mathbb{P}$ such that $f \restriction \Gamma_p(G)$ is not onto. Then $h_p := h \restriction \Gamma_p(G)$ is equipped with the following properties:

($*_6$) (a) $h_p \in \mathrm{End}(\Gamma_\mathrm{p}(\mathrm{G}))$.
(b) Rang(h_p) is bounded.
(c) $h'_p = m \cdot \mathrm{id}_{\Gamma_p(G)} + h_p = \mathrm{id}_{\Gamma_p(G)} + h_p$ is 1-to-1.
(d) h'_p is not onto.

In light of its definition, h_p is a $(1, p)$-antiwitness and so $\mathrm{NQr}(\Gamma_p(G))$ holds. Thanks to Lemma 4.5:

($*_7$) $\Gamma_p(G)$ is infinite.

This is in contradiction with ($*_2$). □

In [1] we studied absolutely co-Hopfian abelian groups. Recall that an abelian group is absolutely co-Hopfian if it is co-Hopfian in any further generic extension of the universe. Also, see [20] for the existence of absolutely Hopfian abelian groups of any given size. Similarly, one may define absolutely endo-rigid groups. Despite its simple statement, one of the most frustrating problems in the theory infinite abelian groups is as follows:

Problem 4.17. Are there absolutely endo-rigid abelian groups of arbitrary large cardinality?

Acknowledgement

The authors thank the referees for reading the paper thoroughly and providing valuable comments.

References

[1] M. Asgharzadeh, M. Golshani, and S. Shelah, "Expressive power of infinitary logic and absolute co-Hopfianity", preprint, 2023. arXiv 2309.16997

[2] R. Baer, "Types of elements and characteristic subgroups of Abelian groups", *Proc. London Math. Soc.* (2) **39**:7 (1935), 481–514. MR Zbl

[3] R. Baer, "Automorphism rings of primary abelian operator groups", *Ann. Math.* (2) **44** (1943), 192–227. Zbl

[4] R. Baer, "Groups without proper isomorphic quotient groups", *Bull. Amer. Math. Soc.* **50** (1944), 267–278. MR Zbl

[5] R. A. Beaumont and R. S. Pierce, "Partly transitive modules and modules with proper isomorphic submodules", *Trans. Amer. Math. Soc.* **91** (1959), 209–219. MR Zbl

[6] G. Braun and L. Strüngmann, "The independence of the notions of Hopfian and co-Hopfian Abelian p-groups", *Proc. Amer. Math. Soc.* **143**:8 (2015), 3331–3341. MR Zbl

[7] A. L. S. Corner, "On endomorphism rings of primary abelian groups", *Quart. J. Math. Oxford Ser.* (2) **20** (1969), 277–296. MR Zbl

[8] A. L. S. Corner and R. Göbel, "Prescribing endomorphism algebras, a unified treatment", *Proc. London Math. Soc.* (3) **50**:3 (1985), 447–479. MR Zbl

[9] P. Crawley, "An infinite primary abelian group without proper isomorphic subgroups", *Bull. Amer. Math. Soc.* **68** (1962), 463–467. MR Zbl

[10] M. Dugas and R. Göbel, "Endomorphism rings of separable torsion-free abelian groups", *Houston J. Math.* **11**:4 (1985), 471–483. MR

[11] P. C. Eklof and A. H. Mekler, *Almost free modules*: *set-theoretic methods*, North-Holland Mathematical Library **65**, North-Holland Publishing Co., Amsterdam, 2002. MR Zbl

[12] B. Franzen and B. Goldsmith, "On endomorphism algebras of mixed modules", *J. London Math. Soc.* (2) **31**:3 (1985), 468–472. MR Zbl

[13] L. Fuchs, *Infinite abelian groups*, *I*, Pure and Applied Mathematics **36**, Academic Press, New York, 1970. MR Zbl

[14] L. Fuchs, *Infinite abelian groups*, *II*, Pure and Applied Mathematics **36-II**, Academic Press, New York, 1973. MR Zbl

[15] L. Fuchs, *Abelian groups*, Springer, Cham, 2015. MR Zbl

[16] R. Göbel and J. Trlifaj, *Approximations and endomorphism algebras of modules*, De Gruyter Expositions in Mathematics **41**, Walter de Gruyter, Berlin, 2012. MR

[17] I. Kaplansky, *Infinite abelian groups*, University of Michigan Press, Ann Arbor, MI, 1954. MR Zbl

[18] E. Matlis, "Injective modules over Noetherian rings", *Pacific J. Math.* **8** (1958), 511–528. MR Zbl

[19] G. Paolini and S. Shelah, "On the existence of uncountable Hopfian and co-Hopfian Abelian groups", preprint, 2021. arXiv 2107.11290

[20] G. Paolini and S. Shelah, "Torsion-free abelian groups are Borel complete", preprint, 2021. arXiv 2102.12371

[21] R. S. Pierce, "Homomorphisms of primary abelian groups", pp. 215–310 in *Topics in abelian groups* (New Mexico State Univ., 1962), Scott Foresman, Chicago, IL, 1963. MR

[22] S. Shelah, "Infinite abelian groups, Whitehead problem and some constructions", *Israel J. Math.* **18** (1974), 243–256. MR Zbl

[23] S. Shelah, “Existence of rigid-like families of abelian p-groups”, pp. 384–402 in *Model theory and algebra*, Lecture Notes in Math. **498**, Springer, Berlin, 1975. MR Zbl

[24] S. Shelah, “Constructions of many complicated uncountable structures and Boolean algebras”, *Israel J. Math.* **45**:2-3 (1983), 100–146. MR Zbl

[25] S. Shelah, “A combinatorial principle and endomorphism rings, I: On p-groups”, *Israel J. Math.* **49**:1-3 (1984), 239–257. MR Zbl

[26] S. Shelah, “A combinatorial theorem and endomorphism rings of abelian groups, II”, pp. 37–86 in *Abelian groups and modules* (Udine, 1984), edited by R. Göbel et al., CISM Courses and Lect. **287**, Springer, Vienna, 1984. MR Zbl

[27] S. Shelah, “Universal classes”, pp. 264–418 in *Classification theory* (Chicago, IL, 1985), Lecture Notes in Math. **1292**, Springer, 1987. MR Zbl

[28] S. Shelah, “Black boxes”, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **65** (2022), 69–130. MR Zbl

[29] S. Shelah, “Building complicated index models and Boolean algebras”, preprint, 2023, available at https://shelah.logic.at/files/230169/511.pdf.

[30] B. Thomé, “$\aleph_1$-separable groups and Kaplansky’s test problems”, *Forum Math.* **2**:3 (1990), 203–212. MR Zbl

Received February 23, 2023. Revised October 8, 2023.

MOHSEN ASGHARZADEH
HAKIMIYEH
TEHRAN
IRAN
mohsenasgharzadeh@gmail.com

MOHAMMAD GOLSHANI
INSTITUTE FOR RESEARCH IN FUNDAMENTAL SCIENCES (IPM)
TEHRAN
IRAN
golshani.m@gmail.com

SAHARON SHELAH
EINSTEIN INSTITUTE OF MATHEMATICS
THE HEBREW UNIVERSITY OF JERUSALEM
JERUSALEM
ISRAEL
and
DEPARTMENT OF MATHEMATICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NJ
UNITED STATES
shelah@math.huji.ac.il

PACIFIC JOURNAL OF MATHEMATICS
Vol. 327, No. 2, 2023
https://doi.org/10.2140/pjm.2023.327.233

THE FUNDAMENTAL SOLUTION TO $\Box_b$ ON QUADRIC MANIFOLDS WITH NONZERO EIGENVALUES AND NULL VARIABLES

ALBERT BOGGESS AND ANDREW RAICH

We prove sharp pointwise bounds on the complex Green operator and its derivatives on a class of embedded quadric manifolds of high codimension. In particular, we start with the class of quadrics that we previously analyzed (*Trans. Amer. Math. Soc. Ser. B* 10 (2023), 507–541) — ones whose directional Levi forms are nondegenerate, and add in null variables. The null variables do not substantially affect the estimates or analysis at the form levels for which $\Box_b$ is solvable and hypoelliptic. In the nonhypoelliptic degrees, however, the estimates and analysis are substantially different. In the earlier paper, when hypoellipticity of $\Box_b$ failed, so did solvability. Here, however, we show that if there is at least one null variable, $\Box_b$ is always solvable, and the estimates are qualitatively different than in the other cases. Namely, the complex Green operator has blow-ups off of the diagonal. We also characterize when a quadric M whose Levi form vanishes on a complex subspace admits a $\Box_b$-invariant change of coordinates so that M presents with a null variable.

1. Introduction

A quadric submanifold of $\mathbb{C}^n \times \mathbb{C}^m$ is a CR manifold that can be written as a graph of a scalar- or vector-valued Hermitian symmetric quadratic form, ϕ, i.e.,

$$M = \{(z, w) \in \mathbb{C}^n \times \mathbb{C}^m : \operatorname{Im} w = \phi(z, z)\}.$$

For a hypersurface ($m = 1$), the analysis of the Kohn Laplacian, $\Box_b$, and the complex Green operator (the relative inverse of $\Box_b$) is well understood and has a long history. The motivating example is the Heisenberg group where $\phi(z, z) = |z|^2$. Its group structure can be exploited to construct explicit convolution kernels to invert the sub-Laplacian as well as the Kohn Laplacian in degree $(0, q)$, $1 \le q \le n-1$, the cases where $\Box_b$ is invertible [Folland and Stein 1974a; 1974b; Hulanicki 1976;

This work was supported by a grant from the Simons Foundation (707123, ASR).
MSC2020: primary 32V20, 32W10, 35R03; secondary 42B37, 43A80.
Keywords: quadric submanifolds, higher codimension, nonzero eigenvalues, complex Green operator, null variables, zero eigenvalues.

Gaveau 1977; Beals et al. 2000; Boggess and Raich 2009]. Estimates of these kernels then show that the Green operator as well as some of its derivatives are continuous operators on $L^p(M)$ as well as in other normed topologies.

For higher codimension quadrics, i.e., $m \geq 2$, much less is understood about the behavior of the Green operator. Part of the difficulty has to do with the fact that the Levi form, ϕ, is vector valued instead of scalar valued as is the case for a hypersurface. Thus, one must consider the *directional Levi form* for each normal direction (see (2) for a precise definition). A breakthrough result came when Peloso and Ricci [2003] characterized the solvability and hypoellipticity for the $\Box_b$-equation on quadrics based on the inertias of the directional Levi forms. This result provided the impetus for much of our research. In [Boggess and Raich 2023], we analyzed the pointwise estimates and L^p regularity of the complex Green operator on $(0, q)$-forms under the assumption that the eigenvalues of each directional Levi form are nonvanishing. In particular, we showed that the complex Green operator in this setting possesses all the same regularity properties as that of the Heisenberg group. On the other hand, there are simple examples of quadrics (see [Boggess and Raich 2021]) where some of the directional Levi forms are degenerate (i.e., have vanishing eigenvalues) and for which the estimates on the complex Green operator have no known parallel with that of any quadric hypersurface. The goal of this paper is to introduce degeneracy into the Levi form in a controlled manner. We do this by adding what we call *null variables* and catalog the effect on the solvability of the $\Box_b$-equation as well as providing sharp estimates for the complex Green operator. As an added dividend, our techniques yield a new result on estimates for the complex Green operator for a hypersurface with null directions in its Levi form.

Analyzing the $\Box_b$-operator on quadrics is a problem that mathematicians have been working on for the past 50 years. Hans Lewy [1957] discovered his famous counterexample of the Cauchy–Kowalevsky theorem in the C^∞ category while investigating the associated $\bar\partial_b$-operator on the Heisenberg group. Regardless of the hypotheses on the Levi form, $\Box_b$ is neither elliptic nor constant coefficient and this makes the function theory difficult. The additional tools provided by the Lie group structure of quadrics permits analysis that is currently unavailable in the general case, especially in the higher codimension setting. For additional background on the $\bar\partial_b$- and $\Box_b$-operators, please see [Boggess 1991; Chen and Shaw 2001; Biard and Straube 2017]. For detailed analysis of the $\Box_b$-operator on quadric manifolds, please see [Boggess 1991; Peloso and Ricci 2003; Boggess and Raich 2011; 2013; 2020; 2022b] and especially [Boggess and Raich 2023].

As mentioned above, in [Boggess and Raich 2023] we analyzed the estimates on the Green operator for a quadric in $\mathbb{C}^n \times \mathbb{C}^m$ where the codimension, m, is at least 2 and where all the directional Levi forms are nondegenerate. As detailed below, this assumption implies that n must be even. In this paper, we add null directions. Therefore, our setting is as follows: let $n' \geq 1$, $n'' \geq 0$, and $n = 2n' + n''$. Let $\phi_0 : \mathbb{C}^{2n'} \times \mathbb{C}^{2n'} \to \mathbb{C}^m$ be a Hermitian symmetric quadratic form; define $\phi : \mathbb{C}^{2n'+n''} \times \mathbb{C}^{2n'+n''}$ by

$$\phi((z', z''), (\tilde{z}', \tilde{z}'')) = \phi_0(z', \tilde{z}').$$

Here, z'' is a *null variable* whereby we mean that ϕ is independent of z''. We let $z = (z', z'')$ so that $z'_k = z_k$ for $1 \leq k \leq 2n'$ and $z''_j = z_j$ for $j = 2n'+1, \ldots, 2n'+n''$. Our main focus is on quadric submanifolds of the form

$$M_\phi = M = \{(z', z'', w) \in \mathbb{C}^{2n'} \times \mathbb{C}^{n''} \times \mathbb{C}^m : \operatorname{Im} w = \phi(z', z')\}. \tag{1}$$

For each unit vector $\nu \in S^{m-1} \subset \mathbb{R}^m$, we define the *directional Levi form* $\phi^\nu : \mathbb{C}^{2n'+n''} \times \mathbb{C}^{2n'+n''} \to \mathbb{C}$ by

$$\phi^\nu(z, \tilde{z}) = \phi(z, \tilde{z}) \cdot \nu = (\tilde{z}')^* A_\nu z', \tag{2}$$

where A_ν is a Hermitian symmetric matrix, depending linearly on the parameter $\nu \in S^{m-1}$. We define the eigenvalues and eigenvectors of the directional Levi forms to be the eigenvalues and eigenvectors of A_ν, and let $n^\pm(\nu)$ be the number of positive/negative eigenvalues of A_ν. When M is a hypersurface, there are directional Levi forms in only two directions: $\nu = 1$ and $\nu = -1$ since S^0 has two points. In codimension $m \geq 2$, ν belongs to the unit sphere S^{m-1}, a connected set. As shown in [Boggess and Raich 2023], the connectivity of S^{m-1}, $m \geq 2$, implies that $n^+(\nu) = n^-(\nu) = n'$ whereas this is not necessarily true for the hypersurface case ($m = 1$).

Peloso and Ricci [2003] found that $\Box_b$ is solvable (resp. hypoelliptic) on $(0, q)$-forms on M_ϕ if and only if there does not exist $\nu \in \mathbb{R}^m \setminus \{0\}$ so that $n^+(\nu) = q$ (resp. $n^+(\nu) \leq q$) and $n^-(\nu) = 2n' + n'' - q$ (resp. $n^-(\nu) \leq 2n' + n'' - q$). For the $m \geq 2$ and $n'' = 0$ case studied in [Boggess and Raich 2023], $n^+(\nu) = n' = n^-(\nu)$, and hence $\Box_b$ is solvable and hypoelliptic for all $q \neq n'$ and neither solvable nor hypoelliptic when $q = n'$. The lack of solvability is related to the fact that $\ker \Box_b \neq \{0\}$ when $q = n'$. After subtracting the orthogonal projection onto $\ker \Box_b$ in the case $q = n'$, the complex Green operator satisfies estimates analogous to those for the Heisenberg group, that is, estimates that are completely governed by the control metric for M. We know, however, that when the eigenvalues of the directional Levi forms are not bounded away from zero, the control distance does not always suffice to control estimates on $N_{0,q}$. This occurs both for hypersurfaces as well as higher codimension quadrics [Machedon 1988; Nagel and Stein 2006; Boggess and Raich 2021].

As mentioned above, z'' are null variables, and we henceforth assume that $n'' \geq 1$. Given this assumption and the fact that for all $\nu \in \mathbb{R}^m \setminus \{0\}$, $n^+(\nu) = n^-(\nu) = n'$, it follows that $\Box_b$ is solvable on M_ϕ for all $0 \leq q \leq 2n' + n''$. Additionally,

$\Box_b$ fails to be hypoelliptic if q satisfies $n' \le q$ and $n' \le 2n' + n'' - q$, that is, $n' \le q \le n' + n''$. Interestingly, adding in null variables improves the solvability of $\Box_b$ while leaving alone the number of hypoelliptic degrees. The estimate for $N_{0,q}$ in the nonhypoelliptic cases is qualitatively worse than in the hypoelliptic cases. The sharp bound is no longer controlled solely by the control distance and the integral kernel has singularities off of the diagonal. Detailed results are stated in Section 2.

In contrast, the class of hypersurfaces we study are of the form

$$M = \{(z', z'', w) \in \mathbb{C}^{n'} \times \mathbb{C}^{n''} \times \mathbb{C} : \operatorname{Im} w = \phi(z', z')\},$$

where $\phi : \mathbb{C}^{n'} \times \mathbb{C}^{n'} \to \mathbb{C}$ is a Hermitian symmetric, scalar-valued, quadratic form. We write $\phi(z', z') = (z')^* A z'$, where A is a nondegenerate, Hermitian symmetric matrix. Suppose that A has n^+ positive eigenvalues and n^- negative eigenvalues. Here, we are *not* assuming $n^+ = n^-$. Solvability always holds because solvability fails if and only if there is a direction for which the sum of the positive eigenvalues and negative eigenvalues is n. However, this never happens with A or $-A$ as this sum equals $n' < n$. Additionally, hypoellipticity fails if $n^+ \le q \le n - n^-$ or $n^- \le q \le n - n^+$ and holds otherwise. Since $n' = n^+ + n^-$, hypoellipticity fails if and only if $n^+ \le q \le n^+ + n''$ or $n - n^+ - n'' \le q \le n - n^+$. Detailed estimates on the Green operator for a hypersurface with null variables are stated in Section 2.

As with many past researchers (e.g., Folland and Stein [1974a], Nagel et al. [2001], and Nagel and Stein [2006]), our approach to computing a working formula for the Green operator, for $m \ge 1$, involves the integral of the fundamental solution to the heat equation associated to $\Box_b$ in the time variable. However, in the case of a one-dimensional null space ($n'' = 1$), the heat kernel is not integrable in the time variable, and we therefore develop a new technique to obtain the Green operator in this case. The resulting kernel and its estimates are stated in Section 2. Proofs of the theorems stated in Section 2 are given in Sections 3, 4, and 5. In Section 6, we show that the estimates given in our theorems are sharp.

2. Notation and main results

Notation for null variables. Define the projection $\pi : \mathbb{C}^{2n'+n''} \times \mathbb{C}^m \to \mathbb{C}^{2n'+n''} \times \mathbb{R}^m$ by $\pi(z, t+is) = (z, t)$. Given a quadric $M \subset \mathbb{C}^{2n'+n''} \times \mathbb{C}^m$, the projection π induces CR and Lie group structures on $\mathbb{C}^{2n'+n''} \times \mathbb{R}^m$, and we call this Lie group G. Since the projection is a CR isomorphism, we primarily work on G but use the same notation interchangeably for objects on M and their pushfowards/pullbacks on G.

The group structure for G is

$$(z, t) * (\zeta, u) = (z + \zeta, t + u - 2 \operatorname{Im} \phi(z, \zeta)) \quad \text{for } (z, t), (\zeta, u) \in G, \tag{3}$$

and this group operation can easily be lifted to M.

Denote the set of increasing q-tuples by

$$\mathcal{I}_q = \{K = (k_1, \dots, k_q) \in \mathbb{N}^q : 1 \le k_1 < k_2 < \dots < k_q \le 2n' + n''\}.$$

Definition 2.1. Given an index $K \in \mathcal{I}_q$, we say a current $N_K = \sum_{L\in\mathcal{I}_q} \tilde{N}_{K,L}(z,t)\, d\bar{z}_L$ is a *fundamental solution to* $\Box_b$ on forms spanned by $d\bar{z}_K$ if $\Box_b N_K = \delta_0(z,t)\, d\bar{z}_K$. A *fundamental solution* $N_{0,q}$ to $\Box_b$ acting on $(0, q)$-forms is then given by

$$N_{0,q} f = \sum_{K\in\mathcal{I}_q} N_K\{f_K\, d\bar{z}_K\}.$$

In higher codimension ($m \ge 2$) a fundamental solution to $\Box_b$ on forms spanned by $d\bar{z}_K$ usually involves terms spanned by $d\bar{z}_L$ for $L \ne K$ in addition to $L = K$.

N_K acts on smooth forms with compact support by componentwise convolution with respect to the group structure on G, that is, if $f = f_0\, d\bar{z}^K$, then $N_K * f = \sum_{L\in\mathcal{I}_q} \tilde{N}_{K,L} * f_0\, d\bar{z}_L$. Thus

$$\tilde{N}_{K,L} * f_0(z,t) = \int_{(\zeta,u)\in G} N_{K,L}((z,t)*(\zeta,u)^{-1}) f_0(\zeta,u)\, dv(z)\, dt,$$

where $dv(z)\, dt$ is the usual volume form for G, and (using (3))

$$(z,t)*(\zeta,u)^{-1} = (z-\zeta,\, t-u+2\,\mathrm{Im}\,\phi(z,\zeta)).$$

Recall that $\delta_0 * f = f$. Therefore, if N_K is a fundamental solution to $\Box_b$ and $f = f\, d\bar{z}_K$ is a smooth form with compact support, then $\Box_b\{N_K * f\} = f$. As mentioned in the introduction, Peloso and Ricci [2003] showed that solvability in our context is possible in all degrees, i.e., $0 \le q \le n = 2n' + n''$. They also showed that solvability is equivalent to the triviality of the L^2 null space of $\Box_b$. We therefore conclude that if $n'' > 0$, then any two fundamental solutions to $\Box_b$ must differ by a non-L^2 current.

For a multiindex $I = (I_1, I_2, I_3) \in \mathbb{N}_0^{4n'+2n''+m}$, the multiindex $I_1 \in \mathbb{N}_0^{4n'}$ records the differentiation in the z' and $\bar{z}'$ variables, $I_2 \in \mathbb{N}_0^{2n''}$ records the differentiation in the z'' and $\bar{z}''$ variables, and $I_3 \in \mathbb{N}_0^m$ records the t-derivatives. Given such a multiindex I, define the *weighted order* of I by $\langle I\rangle = |I_1| + |I_2| + 2|I_3|$ and the *order* of I by $|I| = |I_1| + |I_2| + |I_3|$.

As a consequence of the discussion in Section 1, we assume the following when the codimension, m, is at least 2:

- For each $\nu \in S^{m-1}$, there are n' positive eigenvalues μ^ν_j for j in some index set P^ν of cardinality n from the set $\{1, 2, \dots, 2n'\}$ and n' negative eigenvalues μ^ν_k for $k \in (P^\nu)^c$, the complement of P^ν in $\{1, 2, \dots, 2n'\}$.

Remark 2.2. Given that our nonzero eigenvalues stay bounded away from 0 independently of $\nu \in S^{m-1}$, we may arrange the indices so that $P^\nu = P$ is independent of ν.

Recall the set of increasing q-tuples is denoted by

$$\mathcal{I}_q = \{K = (k_1, \dots, k_q) \in \mathbb{N}^q : 1 \le k_1 < k_2 < \cdots < k_q \le 2n' + n''\}.$$

Also set

$$\mathcal{I}'_{q'} = \{K' = (k_1, \dots, k_{q'}) \in \mathbb{N}^{q'} : 1 \le k_1 < k_2 < \cdots < k_{q'} \le 2n'\},$$
$$\mathcal{I}''_{q''} = \{K'' = (k_1, \dots, k_{q''}) \in \mathbb{N}^{q''} : 2n'+1 \le k_1 < k_2 < \cdots < k_{q''} \le 2n' + n''\}.$$

Given $K \in I_q$, we can always decompose $K = (K', K'')$ where $K' \in \mathcal{I}'_{q'}$ and $K'' \in \mathcal{I}''_{q''}$ for some q', q'' with $q' + q'' = q$. Our notation follows [Boggess and Raich 2022b]. For $\lambda \in \mathbb{R}^m \setminus \{0\}$, set $\nu = \lambda/|\lambda| \in S^{m-1}$. We write $z' \in \mathbb{C}^{n'}$ in terms of the unit eigenvectors of ϕ_λ which means that $(z')^\lambda_j = (z')^\nu_j$ is given by

$$(z')^\nu := Z(\nu, z') := U(\nu)^* \cdot z',$$

where $U(\nu)$ is the matrix whose columns are the eigenvectors v^ν_k, $1 \le k \le 2n'$, of the directional Levi form ϕ^ν, and $\cdot$ represents matrix multiplication with z' written as a column vector. Note that the corresponding orthonormal basis of (0, 1)-covectors for this basis is

$$d\bar{Z}_j(\nu, z'), \quad 1 \le j \le 2n',$$

where $d\bar{Z}(\nu, z') = U(\nu)^T \cdot d\bar{z}'$, $d\bar{z}'$ is written as a column vector of (0, 1)-forms, and the superscript T stands for transpose. Note that $(z')^\nu = Z(\nu, z')$ depends smoothly on $z' \in \mathbb{C}^{n'}$ but only is *locally integrable* as a function of $\nu \in S^{m-1}$ [Rainer 2011]. The coordinates for the remaining n'' variables, $z'' = (z_{2n'+1}, \dots, z_n)$, do not depend on ν. Denote by $\mathbb{I}_{n''}$ the $n'' \times n''$ identity matrix. We write

$$z^\nu = (z', z'')^\nu = (z^\nu, z'') = Z(\nu, z) = (Z(\nu, z'), z'') = (U(\nu)^* \oplus \mathbb{I}_{n''})(z', z''),$$

where $(A \oplus B)(z', z'') := (A(z'), B(z''))$ for any $n' \times n'$ matrix, A, and any $n'' \times n''$ matrix, B. Also,

$$d\bar{Z}(\nu, z) = (d\bar{Z}(\nu, z'), d\bar{z}'') = (U(\nu)^T \oplus \mathbb{I}_{n''}) \cdot (d\bar{z}', d\bar{z}'').$$

We will need to express $d\bar{z}_K$ in terms of $d\bar{Z}(\nu, z)_L$ for $L \in \mathcal{I}_q$. We have

$$d\bar{z}_K = d\bar{z}'_{K'} \wedge d\bar{z}''_{K''} = \sum_{L' \in \mathcal{I}'_{q'}} \det(\bar{U}(\nu)_{K',L'})\, d\bar{Z}(\nu, z')_{L'} \wedge d\bar{z}''_{K''}, \tag{4}$$

where $\bar{U}(\nu)_{K',L'}$ is the $q' \times q'$ minor of $\bar{U}(\nu)$ comprised of elements in the rows K' and columns L'. Note that if $q = 2n' + n''$, then the above sum only has one term and $\det \bar{U}(\nu)_{K,K} = 1$. In addition, when $q = 0$, $\mathcal{I}_0 = \varnothing$ and the sum (4) does not appear. Similarly,

$$d\bar{Z}(\nu, z')_L = \sum_{J \in \mathcal{I}'_{q'}} \det(U(\nu)^T_{L',J})\, d\bar{z}'_J \wedge d\bar{z}''_{L''}. \tag{5}$$

Throughout the paper, we use the function

$$A(r, \nu, z) = A(r, \nu, z', z'') = \frac{2}{|\log r|}|z''|^2 + \sum_{j=1}^{2n'} |\mu^\nu_j| \left(\frac{1 + r^{|\mu^\nu_j|}}{1 - r^{|\mu^\nu_j|}}\right) |z^\nu_j|^2,$$

where μ^ν_j are the nonzero eigenvalues for A_ν and the dimensional constant is

$$K_{2n'+n'',m} = \frac{4^{2n'+n''}(2n' + n'' + m - 2)!}{2(2\pi)^{2n'+n''+m}}.$$

Main results for codimension ≥ 2. Our first theorem provides a formula for the fundamental solution to $\Box_b$ in the case where the null variable dimension satisfies $n'' \geq 2$.

Theorem 2.3. *Let $M \subset \mathbb{C}^{2n'+n''} \times \mathbb{C}^m$, with $m \geq 2$, $n' \geq 1$, and $n'' \geq 2$, be a quadric submanifold defined by* (1) *with associated projection G, and assume that there exists a Hermitian symmetric quadratic form $\phi_0 : \mathbb{C}^{n'} \times \mathbb{C}^{n'} \to \mathbb{C}^m$ such that*

(1) *$\phi(z, \tilde{z}) = \phi_0(z', \tilde{z}')$ for all $z \in \mathbb{C}^{2n'+n''}$ and*

(2) *the eigenvalues of the directional Levi forms of ϕ_0 are nonzero.*

For any $0 \leq q \leq 2n' + n''$, there is a fundamental solution $N = N_{0,q}$ to $\Box_b$ on $(0, q)$-forms given by convolution with the kernel

$$\begin{aligned} N_K(z, t) = K_{2n'+n'',m} &\sum_{L \in \mathcal{I}_q} \int_{\nu \in S^{m-1}} \det(\bar{U}(\nu)_{K,L})\, d\bar{Z}(\nu, z)^L \\ &\times \int_{r=0}^1 \frac{1}{|\log r|^{n''}} \Biggl(\prod_{\substack{j \in L^c \cap P \\ j \in L \cap P^c}} \frac{r^{|\mu^\nu_j|}|\mu^\nu_j|}{1 - r^{|\mu^\nu_j|}} \prod_{\substack{k \in L \cap P \\ k \in L^c \cap P^c}} \frac{|\mu^\nu_k|}{1 - r^{|\mu^\nu_k|}} \Biggr) \\ &\qquad \times \frac{1}{(A(r, \nu, z', z'') - i\nu \cdot t)^{2n+m-1}} \frac{dr\, d\nu}{r}, \end{aligned} \tag{6}$$

where $d\nu$ is surface measure on the unit sphere S^{m-1}.

This theorem follows directly from Theorem 2.3 in [Boggess and Raich 2022b], and the formula is similar to the corresponding one in the same work, where $n'' = 0$ (no $\log r$ term appears). The formula for N is the s-integral over $0 \leq s < \infty$ of the partial Fourier transform of the $\Box_b$ heat kernel $\tilde{H}_K(s, z, \hat{\lambda})$; see (16) (where s represents time). For this derivation, we require that this heat kernel is integrable in

s over $0 \le s < \infty$, which, as we shall see below, holds whenever $\Box_b$ is hypoelliptic or $n'' \ge 2$. However, when $n'' = 1$ in the nonhypoelliptic case, this heat kernel fails to be integrable in s and, consequently, the factor $1/(r\,|\log r|^{n''})$ appearing in (6) is *not* integrable in r near $r = 0$ when $n'' = 1$. The numerator is nonvanishing at $r = 0$ when $L = P$. In Theorem 2.4, below, we derive a fundamental solution for $\Box_b$ when $n'' = 1$ and $L = P$ by modifying our earlier construction to ensure greater decay in the time variable s without disturbing the approximation of the identity behavior as $s \to 0$. This kernel requires a genuinely new idea that is not anticipated in [Boggess and Raich 2022b].

Theorem 2.4. *Let $M \subset \mathbb{C}^n \times \mathbb{C}^m$ be a quadric submanifold as in Theorem 2.5 but with $n'' = 1$ (and $n = 2n' + 1$). Let $K \in \mathcal{I}_q$ where $q = n'$ or $q = n' + 1$ and $K' \in \mathcal{I}'_{n'}$. Then $\tilde{H}_K(s, z, \hat{\lambda})$ is not integrable on $(0, n')$- or $(0, n' + 1)$-forms, and a fundamental solution to $\Box_b$ on forms spanned by $d\bar{z}_K$ is given by*

$$
\begin{aligned}
(7)\quad & N_K(z,t) \\
&= K_{2n'+1,m} \sum_{\substack{L\in\mathcal{I}'_{q'}\\ L\neq P}} \int_{\nu\in S^{m-1}} \det(\bar{U}(\nu)_{K',L})\, d\bar{Z}(\nu,z')_L \wedge d\bar{z}''_{K''} \\
&\quad\times \int_{r=0}^{1} \Biggl(\prod_{\substack{j\in(L')^c\cap P\\ j\in L'\cap P^c}} \frac{r^{|\mu_j^\nu|}|\mu_j^\nu|}{1-r^{|\mu_j^\nu|}} \prod_{\substack{k\in L'\cap P\\ k\in(L')^c\cap P^c}} \frac{|\mu_k^\nu|}{1-r^{|\mu_k^\nu|}} \Biggr) \\
&\qquad\qquad\qquad\qquad\times \frac{1}{(A(r,\nu,z)-i\nu\cdot t)^{2n'+m}}\, \frac{dr\,d\nu}{|\log r|r} \\
&\quad+ \int_{\nu\in S^{m-1}} \det(\bar{U}(\nu)_{K',P})\, d\bar{Z}(\nu,z')_P \wedge d\bar{z}''_{K''} |\det A_\nu| \\
&\quad\times \int_{r=0}^{\frac12} \Biggl[\Biggl(\prod_{j=1}^{2n'} \frac{1}{1-r^{|\mu_j^\nu|}} \Biggr) \frac{1}{(A(r,\nu,z)-i\nu\cdot t)^{2n'+m}} \\
&\qquad\qquad\qquad - \frac{1}{(A(0,\nu,z',0)-i\nu\cdot t)^{2n'+m}} \Biggr] \frac{dr\,d\nu}{|\log r|r} \\
&\quad+ \int_{\nu\in S^{m-1}} \det(\bar{U}(\nu)_{K',P})\, d\bar{Z}(\nu,z')_P \wedge d\bar{z}''_{K''} |\det A_\nu| \\
&\quad\times \int_{r=\frac12}^{1} \Biggl(\prod_{j=1}^{2n'} \frac{1}{1-r^{|\mu_j^\nu|}} \Biggr) \frac{1}{(A(r,\nu,z)-i\nu\cdot t)^{2n'+m}}\, \frac{dr\,d\nu}{|\log r|r}.
\end{aligned}
$$

When $L = P$ in the above formula for N, the term inside the large brackets, $[\,\cdot\,]$, in the integrand of (7) vanishes sufficiently quickly at $r = 0$, and thus this term is integrable in r over $0 \le r \le \frac{1}{2}$.

Our main theorem regarding pointwise bounds on the kernel for the fundamental solution of $\Box_b$ is the following:

Theorem 2.5. *Let $M \subset \mathbb{C}^{2n'+n''} \times \mathbb{C}^m$, with $m \geq 2$ and $n', n'' \geq 1$, be a quadric submanifold defined by* (1) *with associated projection G, and assume that there exists a Hermitian symmetric quadratic form $\phi_0 : \mathbb{C}^{n'} \times \mathbb{C}^{n'} \to \mathbb{C}^m$ so that*

(1) *$\phi(z, \tilde{z}) = \phi_0(z', \tilde{z}')$ for all $z \in \mathbb{C}^{2n'+n''}$ and*

(2) *the eigenvalues of the directional Levi forms of ϕ_0 are nonzero.*

Let $N = N_{0,q}$.

- *Suppose that $0 \leq q < n'$ or $q > n' + n''$. For any multiindex $I \in \mathbb{N}_0^{4n'+2n''+m}$, there exists a constant $C_I > 0$ so that*

$$|D^I N(z,t)| \leq \frac{C_I}{(|z|^2+|t|)^{2n'+n''+m-1+\frac{1}{2}\langle I\rangle}}. \tag{8}$$

- *Suppose that $n' \leq q \leq n' + n''$ and $n'' \geq 2$. Then there exists a constant $C_I > 0$ so that*

$$|D^I N(z,t)| \leq \frac{C_I}{(|z|^2+|t|)^{n''-1+\frac{1}{2}|I_2|}(|z'|^2+|t|)^{2n'+m+\frac{1}{2}|I_1|+|I_3|}}. \tag{9}$$

- *Finally, suppose that $n' \leq q \leq n' + n''$ and $n'' = 1$. Then there exists a constant $C_I > 0$ so that*

$$|D^I N(z,t)| \leq C_I \begin{cases} \dfrac{\log\left(1+\frac{|z_{n'+1}|^2}{|z'|^2+|t|}\right)}{(|z'|^2+|t|)^{2n'+m}} & \text{if } I = 0, \\[2ex] \dfrac{1}{(|z|^2+|t|)^{\frac{1}{2}|I_2|}(|z'|^2+|t|)^{2n'+m+\frac{1}{2}|I_1|+|I_3|}} & \text{if } I \neq 0. \end{cases} \tag{10}$$

These estimates are sharp.

In this paper, we only provide the proof for the case $I = 0$. The proof in the $I \neq 0$ case provides no additional insights, though we do discuss later how derivatives affect the estimates. Keeping track of higher derivatives requires some bookkeeping, which is thoroughly explained and carried out in [Boggess and Raich 2023].

In the case where $0 \leq q < n'$ or $q > n' + n''$, the estimate in (8) implies that N_q is locally integrable in G and more can be said about the regularity of N_q as an operator using the theory of homogeneous groups. Let $W^{k,p}(M)$ denote the Sobolev space of forms on M with z-, $\bar{z}$- and t-derivatives of order k in $L^p(M)$. Following the approach of [Boggess and Raich 2022a, Section 7.3], we can view G (and hence M) as a homogeneous group with norm function $\rho(z,t) = |z| + |t|^{1/2}$. From (8), it follows that the integration kernel of $N_{0,q}$ and its derivatives have the appropriate pointwise decay (analogous to that in the case of nonzero eigenvalues handled in [Boggess and Raich 2023]). A second consequence of (8) is that $N_{0,q}$ is a tempered distribution, and combining this fact with the natural dilation structure and that $D^I N_{0,q}$ is a convolution operator shows that $D^I N_{0,q}$ is uniformly bounded on normalized bump functions. This is exactly what is required to establish the L^p boundedness, $1 < p < \infty$. The convolution operator $D^I N_{0,q}$ extends to a bounded operator on $W^{k,p}(\mathbb{C}^n \times \mathbb{R}^m)$, and we state this as a corollary to Theorem 2.5.

Corollary 2.6. *Let $M \subset \mathbb{C}^{2n'+n''} \times \mathbb{C}^m$ be a quadric submanifold satisfying the hypothesis of Theorem 2.5. Assume $0 \le q < n'$ or $q > n' + n''$. Given a multiindex $I \in \mathbb{N}_0^{4n+m}$ such that $\langle I \rangle = 2$, the operator $D^I N_{0,q}$ is exactly regular on $W^{k,p}(M)$ for all $k \ge 0$ and all $1 < p < \infty$. In other words, $D^I N_{0,q}$ extends to a bounded operator on $W^{k,p}(M)$. In particular, $D^I N_{0,q}$ is a hypoelliptic operator.*

The regularity properties of $N_{(0,q)}$ are not yet known for $n' \le q \le n' + n''$.

Results for hypersurfaces. Even though our focus is mostly on the higher codimension case, our technique provides a new result in the hypersurface case as well. When M is a hypersurface, M is of the form

$$M = \{(z', z'', w) \in \mathbb{C}^{n'} \times \mathbb{C}^{n''} \times \mathbb{C} : \operatorname{Im} w = \phi_0(z', z')\}, \tag{11}$$

where $\phi_0(z', z') = (z')^* A z'$ and A is a nondegenerate Hermitian matrix. Since A is Hermitian, we can choose coordinates in which A is diagonal. In these coordinates (which we still call (z', z'')),

$$\phi(z, z) = \sum_{j=1}^{n'} \mu_j |z_j|^2,$$

where $\mu_1, \dots, \mu_{n'}$ are the nonzero eigenvalues of A. In the hypersurface case, there is not a requirement that n' is even or $n^+ = n^-$. Also, $\Box_b$ acts diagonally in these coordinates. This means if $f = \sum_{J \in \mathcal{I}_q} f_J \, d\bar{z}_J$, then $\Box_b f = \sum_{J \in \mathcal{I}_q} \Box_J f_J \, d\bar{z}_J$. Consequently, to invert $\Box_b$, we need only to invert the $\Box_J$-operators which is simpler than the higher codimension cases handled in the previous subsection. We continue to let P denote the indices of the positive eigenvalues of A. For the theorems in this section, we need the following notation. Let

$$A(r, z) = \frac{2}{|\log r|} |z''|^2 + \sum_{j=1}^{n'} \frac{1 + r^{|\mu_j|}}{1 - r^{|\mu_j|}} |\mu_j| \, |z_j|^2$$

and

$$\varepsilon_{j,L} = \begin{cases} \operatorname{sgn}(\mu_j), & j \in L, \\ -\operatorname{sgn}(\mu_j), & j \notin L. \end{cases}$$

The proof of Theorem 2.4 is easily adapted to prove the following result.

Theorem 2.7. *Let $M \subset \mathbb{C}^{n'} \times \mathbb{C}^{n''} \times \mathbb{C}$ be a quadric hypersurface described by* (11). *Fix $0 \le q \le n$, where $n = n' + n''$, and let $L \in \mathcal{I}_q$.*

(1) *If $n'' \ge 2$ or $n'' = 1$ and L' is neither P nor P^c, then the fundamental solution to the $\Box_L$-equation given by the inverse Fourier transform in t of $\int_0^\infty e^{-s\Box_L}\, ds$ is*

$$N_L(z,t) = \frac{2^{2n-1}(n-1)!}{(2\pi)^{n+1}} |\det A| \Bigg(\int_0^1 \prod_{j=1}^{n'} \frac{r^{\frac{1}{2}(1-\varepsilon_{j,L})|\mu_j|}}{1-r^{|\mu_j|}} \frac{1}{(A(r,z)-it)^n} \frac{dr}{r\,|\log r|^{n''}} + \int_0^1 \prod_{j=1}^{n'} \frac{r^{\frac{1}{2}(1+\varepsilon_{j,L})|\mu_j|}}{1-r^{|\mu_j|}} \frac{1}{(A(r,z)+it)^n} \frac{dr}{r\,|\log r|^{n''}} \Bigg)$$

(2) *If $n'' = 1$ and $L' = P$, then there is a fundamental solution to the $\Box_L$-equation given by*

$$N_L(z,t) = \frac{2^{2n-1}(n-1)!}{(2\pi)^{n+1}} |\det A| \Bigg(\int_0^1 \prod_{j=1}^{n'} \frac{r^{|\mu_j|}}{1-r^{|\mu_j|}} \frac{1}{(A(r,z)+it)^n} \frac{dr}{r\,|\log r|} + \int_{\frac{1}{2}}^1 \prod_{j=1}^{n'} \frac{1}{1-r^{|\mu_j|}} \frac{1}{(A(r,z)-it)^n} \frac{dr}{r\,|\log r|} + \int_0^{\frac{1}{2}} \bigg(\prod_{j=1}^{n'} \frac{1}{1-r^{|\mu_j|}} \frac{1}{(A(r,z)-it)^n} - \frac{1}{(A(0,z)-it)^n} \bigg) \frac{dr}{r|\log r|} \Bigg).$$

(3) *If $n'' = 1$ and $L' = P^c$, then*

$$N_P(z,-t) = \overline{N_P(z,t)}$$

is a fundamental solution to the $\Box_L$-equation.

The form of the solutions from Theorem 2.7 are simpler versions than in Theorem 2.4 in the $n'' = 1$ case and (20) in the $n'' \ge 2$ case. The analysis in the higher codimension case shows that the size comes from the r-integral and there is no cancellation in the ν-integral. Consequently, the proof of Theorem 2.5 proves the following theorem as well.

Theorem 2.8. *Let $M \subset \mathbb{C}^{n'} \times \mathbb{C}^{n''} \times \mathbb{C}$ be a quadric hypersurface described by* (11). *Fix $0 \le q \le n$, where $n = n' + n''$, and $L \in \mathcal{I}_q$. For any multiindex $I \in \mathbb{N}_0^{2n+1}$, there exists a constant $C_I > 0$ so that the following hold.*

- *If L' is neither P nor P', then*

$$|D^I N(z,t)| \le \frac{C_I}{(|z|^2 + |t|)^{n+\frac{1}{2}\langle I\rangle}}.$$

This case includes the q for which $\Box_b$ is hypoelliptic.

- *If $n'' \geq 2$ and $L' = P$ or $L' = P^c$, then*

$$|D^I N(z,t)| \leq \frac{C_I}{(|z|^2+|t|)^{n''-1+\frac{1}{2}|I_2|}(|z'|^2+|t|)^{n'+1+\frac{1}{2}|I_1|+|I_3|}}. \tag{12}$$

- *Finally, suppose that $n'' = 1$ and $L' = P$ or $L' = P^c$. Then*

$$|D^I N(z,t)| \leq C_I \begin{cases} \dfrac{\log\bigl(1+\frac{|z_n|^2}{|z'|^2+|t|}\bigr)}{(|z'|^2+|t|)^{n'+1}} & \text{if } I = 0, \\ \dfrac{1}{(|z|^2+|t|)^{\frac{1}{2}|I_2|}(|z'|^2+|t|)^{n'+1+\frac{1}{2}|I_1|+|I_3|}} & \text{if } I \neq 0. \end{cases} \tag{13}$$

These estimates are sharp.

Corollary 2.9. *Suppose M is a quadric hypersurface in $\mathbb{C}^n$ satisfying the hypotheses of Theorem 2.8. Fix $0 \leq q \leq n$, where $n = n' + n''$ and $L \in \mathcal{I}_q$. If L' is neither P nor P', then for any multiindex $I \in \mathbb{N}_0^{4n+m}$ with $\langle I \rangle = 2$, the operator $D^I N_L$ extends to a bounded operator on $W^{k,p}(M)$. In particular, $D^I N_L$ is a hypoelliptic operator.*

Remark 2.10. The estimates in (9), (10), (12), and (13) suggest that we investigate N from the point of view of flag kernels, à la Nagel, Ricci, and Stein [2001]. N is the wrong degree to be a flag kernel as it inverts second-order differential operators, just as the Newtonian potential is the wrong degree to be a Calderón–Zygmand operator. The are four types of second-order derivatives (two derivatives in z' variables, two derivatives in z'' variables, one derivative each in z' and z'' variables, and one derivative in a t variable), and only applying two derivatives in z'' variables to N produces a kernel with the correct order of decay. Even in this case, it is currently unclear if the kernel is a flag kernel. It would be an interesting project to understand the complete mapping properties of N and its second-order derivatives.

Vanishing variables. Our above assumption is that z'' is a null variable. There is a more general concept that we call a *vanishing variable* which is defined as follows: z'' is a *vanishing variable* for ϕ if $\phi(z,z) = 0$ whenever $z = (0, z'')$, $z'' \in \mathbb{C}^{n''}$. A null variable is also a vanishing variable but the converse is not true, as illustrated by the example below. We briefly discuss vanishing variables since the techniques in this paper only apply to null variables. We expect that the analysis of estimates for fundamental solutions in the case of vanishing variables will be more complicated.

Here is an example in $\mathbb{C}^3$ where z_3 is a vanishing variable but not a null variable:

$$\begin{aligned} \phi_1(z,z) &= |z_1|^2 - |z_2|^2, \\ \phi_2(z,z) &= \sqrt{2}\,\mathrm{Re}(z_3\bar{z}_1 + z_3\bar{z}_2), \\ \phi_3(z,z) &= \sqrt{2}\,\mathrm{Re}(iz_3\bar{z}_1 - iz_3\bar{z}_2). \end{aligned} \tag{14}$$

Note that z_3 is a vanishing variable but not a null variable for ϕ due to ϕ's dependence on z_3. There is no $\Box_b$-invariant change of coordinates that will make z_3 a null variable for ϕ. Here, a $\Box_b$-invariant change of coordinates between two quadrics M and M' in $\mathbb{C}^n \times \mathbb{C}^m$ is a nonsingular, complex linear map $T : \mathbb{C}^n \times \mathbb{C}^m \to \mathbb{C}^n \times \mathbb{C}^m$ with $T(M) = M'$ and $T^*(\Box_b f) = \Box_b(T^*(f))$ for all $(0, q)$-forms on M'. As shown in [Boggess and Raich 2020], a $\Box_b$-invariant change of variables requires a unitary change of coordinates in the z variables, i.e., $\hat{z} = U(z)$ where U is a unitary matrix. However, in order to preserve the independence of z_3 for ϕ_1, U must map the copy of $\mathbb{C}^2$ spanned by the z_1 and z_2 axes to itself. Since U is unitary, the orthogonal complement of this set (namely the z_3 axis) must remain invariant under U. Therefore U has the form

$$U = \begin{pmatrix} U_2 & 0 \\ 0 & 1 \end{pmatrix},$$

where U_2 is a 2×2 unitary matrix. A change of variables involving this U cannot remove the dependence of ϕ_2 or ϕ_3 on z_3.

This example illustrates the following point: if z'' is a null variable, then ϕ only depends on the variable z', which is the coordinate for the orthogonal complement of the space spanned by the null variables. This observation and the analysis in the previous paragraph leads to the following theorem.

Theorem 2.11. *Suppose L is a complex subspace of $\mathbb{C}$-dimension n'' in $\mathbb{C}^n$ ($n'' \le n$), and suppose $\phi(z, z) = 0$ for all $z \in L$. Then there exists a $\Box_b$-invariant change of variables so that $z'' \in \mathbb{C}^{n''}$ is a null variable for ϕ if and only if for each $1 \le j \le n$, the map $z \in \mathbb{C}^n \to A_j z$ preserves $L^\perp$ (the orthogonal complement of L in $\mathbb{C}^n$), where A_j are the Hermitian matrices corresponding to the directional Levi forms of the standard basis vectors, E_j, $1 \le j \le m$, in $\mathbb{R}^m$, that is, $\phi_j(z, z) = z^* A_j z$.*

Proof. The proof is clear — if there is a unitary change of variables mapping L to a space spanned by the null variable z'', then the matrices A_j, $1 \le j \le n$, in the new variables must preserve the directions spanned by the z' variables. Since U is unitary, in the original coordinates, A_j must map $L^\perp$ to itself. The converse is similar. $\Box$

From a practical point of view, finding a null variable or vanishing variable for a given ϕ can proceed as follows. First, establish whether all the A_j have a common kernel. If the common kernel is trivial, then there are no vanishing or null variables. If there is a nontrivial common kernel, then diagonalize the matrix representing one of the coordinate functions, say A_1. At least one of the variables, say z_n, is a vanishing variable (representing an eigenvector corresponding to the zero eigenvalue of A_1). Next, see if the other component functions are independent of z_n. If so, then z_n is also a null variable. If not, then z_n is a vanishing variable but not a null variable. There may be additional vanishing and/or null variables depending on the dimension of the common kernel.

3. The $\Box_b$-heat equation and the proof of Theorem 2.4

$\Box_b$ and the partial Fourier transform. The operator $\Box_b$ is translation invariant in t, and so we introduce the *partial Fourier transform* of a function $f(z,t)$ by

$$f(z,\hat{\lambda}) = \frac{1}{(2\pi)^{m/2}} \int_{\mathbb{R}^m} f(t) e^{-i\lambda\cdot t}\, dt$$

with ˆ appearing over the transform variables. As is shown in [Peloso and Ricci 2003], for a fixed $\lambda \in \mathbb{R}^m$ (with $\nu = \lambda/|\lambda|$), the coordinates $Z(\nu, z')$ that diagonalize A_ν also diagonalize $\widehat{\Box}_b$. On the transform side, we treat λ as a parameter and write the transformed operator as $\widehat{\Box}_b^\lambda$. Fix $K \in \mathcal{I}_q$. Note that if $f(z,t) = f_K d\bar{z}_K$ and $q' = |K'|$, then

$$f(z,\hat{\lambda}) = f_K(z,\hat{\lambda})\, d\bar{z}'_{K'} \wedge d\bar{z}''_{K''} = \sum_{L\in\mathcal{I}'_{q'}} f_K(z,\hat{\lambda}) \det(\bar{U}(\nu)_{K',L})\, d\bar{Z}(\nu,z')_L \wedge d\bar{z}''_{K''}.$$

One of the reasons for using the $Z(\nu, z')$ coordinates is that $\widehat{\Box}_b^\lambda$ acts diagonally in these coordinates (see [Boggess and Raich 2022b]). Specifically,

$$(\Box_b f)(z,\hat{\lambda}) = \widehat{\Box}_b^\lambda\{f(z,\hat{\lambda})\} = \sum_{L\in\mathcal{I}'_{q'}} \widehat{\Box}_L^\lambda\{f_K(z,\hat{\lambda}) \det(\bar{U}(\nu)_{K',L})\}\, d\bar{Z}(\nu,z')_L \wedge d\bar{z}''_{K''}$$

where

$$\widehat{\Box}_L^\lambda = -\tfrac{1}{4}\Delta_z + 2i\sum_{k=1}^n \mu_k^\lambda \operatorname{Im}\{z_k^\nu \partial_{z_k^\nu}\} + \sum_{k=1}^n (\mu_k^\lambda)^2 |z_k^\nu|^2 - \Bigl(\sum_{k\in L}\mu_k^\lambda - \sum_{k\notin L}\mu_k^\lambda\Bigr)$$

and Δ_z is the ordinary Laplacian in the indicated variables. Our approach to solving the $\Box_b$-equation is via the $\Box_b$-heat equation. Given the diagonalization of $\widehat{\Box}_b$, it is enough to solve the $\widehat{\Box}_L^\lambda$ equations

$$\begin{aligned} &\Bigl(\frac{\partial}{\partial s} + \widehat{\Box}_L^\lambda\Bigr)\{\tilde{H}_L(s,z,\hat{\lambda})\} = 0 \quad \text{for } s>0, \\ &\tilde{H}_L(s=0,z,\hat{\lambda}) = (2\pi)^{-m/2}\delta_0(z)\otimes 1_\lambda, \end{aligned} \tag{15}$$

where $\delta_0(z)$ is the Dirac-delta function centered at the origin in the z variables and 1_λ is the function which is identically 1 for all $\lambda \in \mathbb{R}^m$. The function $\tilde{H}_L(s,z,\hat{\lambda})$ is called the heat kernel and is given by (see [Boggess and Raich 2011])

$$\tilde{H}_L(s,z,\hat{\lambda}) = \frac{2^n}{(2\pi)^{m/2+n}} \frac{e^{-|z''|^2/s}}{s^{n''}} \prod_{j=1}^{2n'} \frac{e^{s\varepsilon^\nu_{j,L}|\mu_j^\lambda|}|\mu_j^\lambda|}{\sinh(s|\mu_j^\lambda|)} e^{-|\mu_j^\lambda|\coth(s|\mu_j^\lambda|)|Z_j(\nu,z')|^2} \tag{16}$$

where

$$\varepsilon^{\alpha}_{j,L} = \begin{cases} \operatorname{sgn}(\mu^{\lambda}_j) & \text{if } j \in L, \\ -\operatorname{sgn}(\mu^{\lambda}_j) & \text{if } j \notin L. \end{cases}$$

Integrability in s over $0 \le s < \infty$ holds when $n'' \ge 2$ or when $L \ne P$. However, integrability fails when $L = P$ and $n'' = 1$ since

$$\tilde{H}_P(s, z, \hat{\lambda}) = \frac{2^n}{(2\pi)^{m/2+n}} \frac{e^{-|z''|^2/s}}{s} \prod_{j=1}^{2n'} \frac{e^{s|\mu^{\lambda}_j|}|\mu^{\lambda}_j|}{\sinh(s|\mu^{\lambda}_j|)} e^{-|\mu^{\lambda}_j|\coth(s|\mu^{\lambda}_j|)\,|Z_j(\nu,z')|^2}$$

and so $\tilde{H}_P(s, z, \hat{\lambda})$ decays like $\frac{1}{s}$ as $s \to \infty$. Consequently, the harmonic projection onto $\ker \widehat{\Box}^{\lambda}_L$ is 0 yet the “formula"

$$(\widehat{\Box}^{\lambda}_P)^{-1} = \int_0^{\infty} e^{-s\widehat{\Box}^{\lambda}_P}\, ds$$

fails to hold because the integral on the right-hand side diverges.

Proof of Theorem 2.4. Set $\delta_{L,P} = 1$ if $L = P$ and $\delta_{L,P} = 0$ otherwise. Define

$$\begin{aligned} \tilde{S}_{L,P}(z', \hat{\lambda}) &= \lim_{s\to\infty} \frac{2^n}{(2\pi)^{m/2+n}} \prod_{j=1}^{2n'} \frac{e^{s|\mu^{\lambda}_j|}|\mu^{\lambda}_j|}{\sinh(s|\mu^{\lambda}_j|)} e^{-|\mu^{\lambda}_j|\coth(s|\mu^{\lambda}_j|)\,|Z_j(\nu,z')|^2} \delta_{L,P} \\ &= \frac{2^{n+2n'}}{(2\pi)^{m/2+n}} |\det A_{\lambda}| \prod_{j=1}^{2n'} e^{-|\mu^{\lambda}_j|\,|Z_j(\nu,z')|^2} \delta_{L,P}. \end{aligned}$$

Let χ be an indicator function on the ray $[b, \infty)$ where $b > 0$ is to be determined later. Set

$$\tilde{N}_L(z, \hat{\lambda}) = \int_0^{\infty} \tilde{H}_L(s, z, \hat{\lambda}) - \frac{\chi(s|\lambda|)}{s} \tilde{S}_{L,P}(z', \hat{\lambda})\, ds. \tag{17}$$

The integral defining $\tilde{N}_L$ converges because

$$\frac{e^{s|\mu^{\lambda}_j|}}{\sinh(s|\mu^{\lambda}_j|)} e^{-|\mu^{\lambda}_j|\coth(s|\mu^{\lambda}_j|)\,|Z_j(\nu,z')|^2} - 2e^{-|\mu^{\lambda}_j|\,|Z_j(\nu,z')|^2}$$

decays exponentially in s (and the integral kernel is $\partial \tilde{H}_L/\partial s$ near 0). Not coincidentally, $\tilde{S}_P(z', \hat{\lambda})$ is the integral kernel of the harmonic projection onto $\ker\{\widehat{\Box}^{\lambda, M_0}_P\}$ on the quadric M_0. Since $\widehat{\Box}^{\lambda}_P = -\triangle_{z''} + \widehat{\Box}^{\lambda, M_0}_P$, it follows that $\widehat{\Box}^{\lambda}_L \tilde{S}_{L,P} = 0$ for all L.

Consequently,

$$\begin{aligned}\widehat{\Box}_L^\lambda \tilde{N}_L(z,\hat{\lambda}) &= \int_0^\infty \widehat{\Box}_L^\lambda \tilde{H}_L(s,z,\hat{\lambda}) - \frac{\chi(s|\lambda|)}{s}\widehat{\Box}_L^\lambda \tilde{S}_{L,P}(z',\hat{\lambda})\,ds \\ &= -\int_0^\infty \frac{\partial \tilde{H}_L(s,z,\hat{\lambda})}{\partial s}\,ds = \delta_0(z)\otimes 1_\lambda \quad \text{by (15)},\end{aligned}$$

as desired. The latter integral converges as $s\to\infty$ because $\partial\tilde{H}_L(s,z,\hat{\lambda})/\partial s$ decays at least as fast as s^{-2}. We can now construct a solution to invert $\Box_b$ using the modified $\tilde{N}_L(z,\hat{\lambda})$ functions. Following the argument of [Boggess and Raich 2022b, Proposition 3.2], we have the following solution. In the following statement $\mathcal{F}_\lambda^{-1}$ denotes the inverse partial Fourier transform in λ.

Proposition 3.1. *For given indices $K\in\mathcal{I}_q$ and $L\in\mathcal{I}'_{q'}$, define*

$$N_{K,L}(z,\hat{\lambda}) = \det(\bar{U}(\nu)_{K',L})\tilde{N}_L((z',z''),\hat{\lambda})\,d\bar{Z}(z',\nu)_L\wedge d\bar{z}_{K''}, \tag{18}$$

where $\tilde{N}_L(z',z'',\hat{\lambda})$ is defined by (17). *Then there is a fundamental solution to $\Box_b$ on M applied to a form spanned by $d\bar{z}_K$ given by*

$$N_K(z,t) = \mathcal{F}_\lambda^{-1}\Big\{\sum_{L\in\mathcal{I}'_{q'}} N_{K,L}(z,\hat{\lambda})\Big\}(t). \tag{19}$$

We now continue with the proof of Theorem 2.4. If $L\neq P$, then $\tilde{S}_{L,P}(z',\hat{\lambda})=0$ in (17). Recalling that $n''=1$, the calculation in Section 4 of [Boggess and Raich 2022b] shows

$$\begin{aligned}&\mathcal{F}_\lambda^{-1}\{N_{K,L}(z,\hat{\lambda})\}(t) \\ &\quad = K_{n,m}\int_{\nu\in S^{m-1}} \det(\bar{U}(\nu)_{K',L})\,d\bar{Z}(\nu,z')_L\wedge d\bar{z}''_{K''} \\ &\qquad \times\int_{r=0}^1\Bigg(\prod_{\substack{j\in(L')^c\cap P\\ j\in L'\cap P^c}}\frac{r^{|\mu_j^\nu|}|\mu_j^\nu|}{1-r^{|\mu_j^\nu|}}\prod_{\substack{k\in L'\cap P\\ k\in(L')^c\cap P^c}}\frac{|\mu_k^\nu|}{1-r^{|\mu_k^\nu|}}\Bigg) \\ &\qquad\qquad \times\frac{1}{(A(r,\nu,z)-i\nu\cdot t)^{2n'+m}}\frac{dr\,d\nu}{|\log r|r}.\end{aligned} \tag{20}$$

This establishes the terms in (7) where $L\neq P$.

When $L=P$, the $S_{P,P}$ term is present in $\tilde{N}_P$ (see (17)) and we compute the inverse Fourier transform in λ by switching to polar coordinates, $\lambda=\tau\nu$, $\tau\geq 0$,

$\nu \in S^{m-1}$. We have

$$
\begin{aligned}
(21)\quad & \mathcal{F}_\lambda^{-1}\{N_{K,P}(z,\hat\lambda)\}(t)\\
&= \frac{1}{(2\pi)^{m/2}} \int_{\lambda\in\mathbb{R}^m} e^{i\lambda\cdot t}\{\det(\bar U(\nu)_{K',P})\tilde N_P(z,\hat\lambda)\, d\bar Z(z,\nu)_P \wedge d\bar z''_{K''}\}\, d\lambda\\
&= \frac{1}{(2\pi)^{m/2}} \int_{\nu\in S^{m-1}} \det(\bar U(\nu)_{K',P})\, d\bar Z(z,\nu)_P \wedge d\bar z''_{K''}\\
&\quad \times \int_{\tau=0}^{\infty} e^{i\tau\nu\cdot t} \int_{s=0}^{\infty} \Big(\tilde H_P(s,z,\widehat{\tau\nu}) - \frac{\chi(s\tau)}{s} S_{P,P}(z',\widehat{\tau\nu})\Big)\tau^{m-1}\, ds\, d\tau\, d\nu,
\end{aligned}
$$

where $d\nu$ is surface measure on the unit sphere S^{m-1}. Now we insert the heat kernel, $\tilde H_P$, from (16) and focus on the above s,τ-integral in (21), denoted by I_ν. Note that

$$\mu_j^\lambda = \tau\mu_j^\nu \quad \text{and} \quad \det A_\lambda = \tau^{2n'} \det A_\nu.$$

We scale in s by replacing $s\tau$ by s and then integrate we in τ. With $C_{m,n} = 2^n/(2\pi)^{m/2+n}$, we have

$$
\begin{aligned}
I_\nu &= C_{m,n}|\det A_\nu| \int_{s=0}^{\infty}\int_{\tau=0}^{\infty} \Bigg(e^{-\tau|z''|^2/s} \prod_{j=1}^{2n'} \frac{e^{s|\mu_j^\nu|}}{\sinh(s|\mu_j^\nu|)} e^{-|\mu_j^\nu|\coth(s|\mu_j^\nu|)\tau|z_j^\nu|^2}\\
&\qquad - 2^{2n'}\chi(s)\prod_{j=1}^{2n'} e^{-|\mu_j^\nu|\tau|z_j^\nu|^2}\Bigg) e^{it\cdot\nu\tau}\tau^{2n'+m-2}\, d\tau\, \frac{ds}{s}\\
&= C_{m,n}|\det A_\nu| \int_{s=0}^{\infty}\int_{\tau=0}^{\infty} \Bigg(\prod_{j=1}^{2n'} \frac{e^{s|\mu_j^\nu|}}{\sinh(s|\mu_j^\nu|)} e^{-\tau\left(|z''|^2/s+\sum_{j=1}^{2n'}|\mu_j^\nu|\coth(s|\mu_j^\nu|)|z_j^\nu|^2 - i\nu\cdot t\right)}\\
&\qquad - 2^{2n'}\chi(s) e^{-\tau\left(\sum_{j=1}^{2n'}|\mu_j^\nu||z_j^\nu|^2 - it\cdot\nu\right)}\Bigg)\tau^{2n'+m-1}\, d\tau\, \frac{ds}{s}\\
&= (2n'+m-1)!\, C_{m,n}|\det A_\nu|\\
&\quad \times \int_{s=0}^{\infty}\Bigg(\Bigg(\prod_{j=1}^{2n'}\frac{e^{s|\mu_j^\nu|}}{\sinh(s|\mu_j^\nu|)}\Bigg)\frac{1}{\left(\frac{|z''|^2}{s}+\sum_{j=1}^{2n'}|\mu_j^\nu|\coth(s|\mu_j^\nu|)|z_j^\nu|^2 - i\nu\cdot t\right)^{2n'+m}}\\
&\qquad - 2^{2n'}\chi(s)\frac{1}{\left(\sum_{j=1}^{2n'}|\mu_j^\nu||z_j^\nu|^2 - it\cdot\nu\right)^{2n'+m}}\Bigg)\frac{ds}{s},
\end{aligned}
$$

where the last equality uses the formula

$$\int_0^\infty \tau^p e^{-\alpha\tau}\, d\tau = \frac{p!}{\alpha^{p+1}} \quad \text{for } \operatorname{Re}\alpha > 0.$$

We use the substitution $r = e^{-2s}$ in the remaining s-integral (and so $ds/s = -dr/(r\,|\log r|)$ and the oriented r-limits of integration become 1 to 0) to obtain

$$I_\nu = K_{m,n} |\det A_\nu| \int_{r=0}^{1} \Biggl(\biggl(\prod_{j=1}^{2n'} \frac{1}{1 - r^{|\mu_j^\nu|}} \biggr) \frac{1}{(A(r,\nu,z',z'') - i\nu\cdot t)^{2n'+m}} \\ - \chi\bigl(\tfrac{1}{2}|\log r|\bigr) \frac{1}{(A(0,\nu,z',0) - i\nu\cdot t)^{2n'+m}} \Biggr) \frac{dr}{r\,|\log r|}.$$

We choose $b = \frac{1}{2}\log 2$ so that $\chi(\frac{1}{2}|\log r|)$ is the characteristic function of $\bigl[0, \frac{1}{2}\bigr]$. From (21), observe that

$$\mathcal{F}_\lambda^{-1}\{N_{K,P}(z,\hat\lambda)\}(t) = \int_{\nu\in S^{m-1}} I_\nu \det(\bar U(\nu)_{K',P})\, d\bar Z(z,\nu)_P \wedge d\bar z''_{K''}\, d\nu,$$

which equals the term in (7) with $L = P$. Therefore, the proof of Theorem 2.4 is complete.

4. Proof of Theorem 2.5, $|t| \ge |z|^2$

In [Boggess and Raich 2023], the case when $|t| \ge |z|^2$ is the most delicate for the proof of the estimates. In our current manuscript, when $n'' \ge 2$, the case $|t| \ge |z|^2$ is handled by adapting the argument from the corresponding argument in [Boggess and Raich 2023]. Here we only sketch this argument with details on the modifications needed to handle the null variables (z''). We then provide complete details when $n'' = 1$ since new ideas are involved.

The primary new term is $(2\,|z''|^2)/|\log r|$ that appears in $A(r,\nu,z',z'')$. However, the series expansion for $1/|\log r|$ around $r = 1$ has leading term $1/(1-r)$, so the effect of the null directions on the estimates near $r = 1$ is the same as for the nonnull directions. Some bookkeeping is required but the estimates in our context here are very similar to the estimates presented in detail in [Boggess and Raich 2023].

The first step of the analysis is to factor out $|t|^{2n'+n''+m-1}$ from the denominator and rotate in ν via an orthogonal matrix M_t chosen so that $M_t(t/|t|)$ is the unit vector in the ν_1 direction (so in the new coordinates, $\nu\cdot t = \nu_1|t|$). We also set $\nu^t = M_t^{-1}\nu$,

$$p = (p', p'') = \frac{z}{|t|^{1/2}} \in \mathbb{C}^{2n'+n''},$$

$$Q(\nu^t, p) = \frac{Z(\nu^t, z)}{|t|^{1/2}} = \frac{(Z(\nu^t, z'), z'')}{|t|^{1/2}} = \frac{(U(\nu^t)^*\cdot z', z'')}{|t|^{1/2}}.$$

Note that $|Q(\nu^t, p)|^2 = |p|^2$ since U_{ν^t} is unitary.

We obtain

$$N_K(z,t) = |t|^{-(2n'+n''+m-1)} \sum_{L' \in \mathcal{I}'_{q'}} N_{KL'}(p)$$

where

$$N_{K,L'}(p) = \int_{\nu^t \in S^{m-1}} \int_{r=0}^{1} \frac{\det(\bar{U}(\nu^t)_{K',L'}) B_{L'}(r,\nu^t)\, d\bar{Z}(\nu^t,z')_{L'} \wedge d\bar{z}''_{K''}}{(A(r,\nu^t,p) - i\nu_1)^{2n'+n''+m-1}} \frac{d\nu\, dr}{r|\log r|^{n''}} \tag{22}$$

if $L' \neq P$ or $n'' \geq 2$ and where

$$B_{L'}(r,\nu) = \prod_{\substack{j \in (L')^c \cap P \\ j \in L' \cap P^c}} \frac{r^{|\mu_j^\nu|}|\mu_j^\nu|}{1 - r^{|\mu_j^\nu|}} \prod_{\substack{k \in L' \cap P \\ k \in (L')^c \cap P^c}} \frac{|\mu_k^\nu|}{1 - r^{|\mu_k^\nu|}}, \tag{23}$$

$$A(r,\nu,p) = \frac{2}{|\log r|}|p''|^2 + \sum_{j=1}^{2n'} |\mu_j^\nu| \left(\frac{1 + r^{|\mu_j^\nu|}}{1 - r^{|\mu_j^\nu|}} \right) |Q_j(\nu,p')|^2. \tag{24}$$

If $L' = P$ and $n'' = 1$, then

$$\begin{aligned} N_{K,P}(p) = &\int_{\nu^t \in S^{m-1}} \det(\bar{U}(\nu^t)_{K',P})\, d\bar{Z}(\nu^t,z')_P \wedge d\bar{z}''_{K''} |\det A_{\nu^t}| \\ &\times \int_{r=0}^{\frac{1}{2}} \left(\left(\prod_{j=1}^{n-1} \frac{1}{1 - r^{|\mu_j^{\nu^t}|}} \right) \frac{1}{(A(r,\nu^t,p) - i\nu_1)^{2n'+m}} \right. \\ &\qquad \left. - \frac{1}{(A(0,\nu^t,p',0) - i\nu_1)^{2n'+m}} \right) \frac{dr\, d\nu^t}{|\log r| r} \\ &+ \int_{\nu^t \in S^{m-1}} \det(\bar{U}(\nu^t)_{K',P})\, d\bar{Z}(\nu^t,z')_P \wedge d\bar{z}''_{K''} |\det A_{\nu^t}| \\ &\times \int_{r=\frac{1}{2}}^{1} \left(\prod_{j=1}^{n-1} \frac{1}{1 - r^{|\mu_j^{\nu^t}|}} \right) \frac{1}{(A(r,\nu^t,p) - i\nu_1)^{2n'+m}} \frac{dr\, d\nu}{|\log r| r}. \end{aligned}$$

To prove Theorem 2.5 in the case that $|t| \geq |z|^2$ and $0 \leq q \leq 2n' + n''$, it suffices to prove the following theorem.

Theorem 4.1. *There is a uniform constant $C > 0$ so that $|N_{K,L}(p)| \leq C$ for all $p \in \mathbb{C}^{2n'+n''}$ with $|p| \leq 1$ and all $K, L \in I_q$ with $0 \leq q \leq 2n' + n''$.*

We first sketch the estimate of the kernel near $r = 1$ using the ideas from [Boggess and Raich 2023].

Subcase: $|t| \geq |z|^2$ and $\frac{1}{2} < r < 1$. We prove Theorem 4.1. We start with a key result — Lemma 5.2 in [Boggess and Raich 2023], which we restate here.

Lemma 4.2. *Let*

$$B(r,\nu)=B_{\varnothing}(r,\nu)=\prod_{j\in P}\frac{r^{|\mu_j^\nu|}|\mu_j^\nu|}{1-r^{|\mu_j^\nu|}}\prod_{k\in P^c}\frac{|\mu_k^\nu|}{1-r^{|\mu_k^\nu|}}. \tag{25}$$

Then

$$\sum_{L'\in\mathcal{I}'_{q'}}\det(\bar U(\nu)_{K',L'})\,d\bar Z(\nu,z')_{L'}B_{L'}(r,\nu)=\sum_{J'\in\mathcal{I}'_{q'}}\det([r^{-\bar A_\nu}]_{K',J'})B(r,\nu)\,d\bar z_{J'} \tag{26}$$

is real analytic in $\nu\in S^{m-1}$ *and* $0<r<1$.

Remark 4.3. The real content of this lemma is the real analyticity in ν of the expression in (26), especially in view of the fact that the eigenvalues μ_j^ν are not necessarily real analytic or even smooth in the parameter ν. As shown in [Boggess and Raich 2023], the expression $B(r,\nu)$ is real analytic in ν due to the fact that the positive eigenvalues are bounded away from the negative eigenvalues. In addition, $r^{-\bar A_\nu}$ is real analytic in ν since A_ν depends linearly on ν.

Using Lemma 4.2, a typical term for $N_{K,L}(p)$ in (22) — with $\frac12\le r<1$ for the domain of integration — is

$$N^u_{K,J}(p)=\int_{\nu\in S^{m-1}}\int_{r=\frac12}^{1}\frac{\det([r^{-\bar A_\nu}]_{K',J'})B(r,\nu)}{(A(r,\nu^t,p)-i\nu_1)^{2n'+n''+m-1}}\frac{d\nu\,dr}{r\,|\log r|^{n''}}. \tag{27}$$

The superscript u refers to the fact that the integral is over the "upper" piece of the r-interval. Our goal in this section is to establish the following lemma.

Lemma 4.4. *There is a uniform constant C such that*

$$|N^u_{K,J}(p)|\le C$$

for all $p\in\mathbb{C}^{2n'+n''}$ *with* $|p|\le 1$.

As in [Boggess and Raich 2023], we use the change of variable

$$r=r(s)=\frac{s-1}{s+1}\qquad\text{or equivalently}\qquad s=\frac{r+1}{1-r}\quad\text{with}\quad\frac{dr}{r}=\frac{2\,ds}{s^2-1} \tag{28}$$

and observe that $\frac12\le r<1$ transforms to $s\ge 3$. We obtain

$$\begin{aligned}&N^u_{K,J}(p)\\&=2\int_{\nu\in S^{m-1}}\int_{s=3}^{\infty}\frac{\det[r(s)^{-\bar A_\nu}]_{K',J'}}{(A(r(s),\nu^t,p)-i\nu_1)^{2n'+n''+m-1}}\frac{B(r(s),\nu)r'(s)}{r(s)}\frac{ds\,d\nu}{|\log r(s)|^{n''}}.\end{aligned} \tag{29}$$

We then expand the various components of the integrand defining $N^u_{K,J}(p)$ on the last line of (29) about $s=\infty$. We briefly outline the main steps in Sections 5, 6, and 7 in [Boggess and Raich 2023] and point out the differences needed to deal with the factor of $|\log r(s)|^{n''}$ in the denominator. From Proposition 5.4 in [Boggess and Raich 2023], we have

$$\frac{B(r(s),\nu)r'(s)}{r(s)} = \frac{2}{2^{2n'}\left(1-\frac{1}{s^2}\right)}\left(\sum_{\ell=0}^{2n'-1} P_\ell(\nu)s^{2n'-\ell-2} + \frac{O(s,\nu)}{s^2}\right), \tag{30}$$

(31) a typical monomial in $P_\ell(\nu) = \nu^{\ell-e}$, where e is even with $0 \le e \le \ell$.

Here, $P_\ell(\nu)$ is a polynomial in $\nu = (\nu_1, \dots \nu_m) \in S^{m-1}$ of total degree ℓ. By an abuse of notation, the term $\nu^{\ell-e}$ in (31) stands for a monomial in the coordinates of ν of total degree $\ell - e$. Also note that the term $(1-s^{-2})^{-1}$ on the right-hand side of (30) only has even powers of $1/s$ in its expansion about $s = \infty$.

Next, we use the second part of Proposition 5.4 in [Boggess and Raich 2023] to expand $\det[r(s)^{-\bar{A}_\nu}]_{K,J}$ around $s = \infty$. The result is a sum of terms of the form

$$\frac{\nu^{\ell'-e'}}{s^{\ell'}}, \tag{32}$$

where $\ell' \ge 1$, e' is an even integer with $0 \le e' \le \ell'$, and $\nu^{\ell'-e'}$ is a monomial of degree $\ell' - e'$ in the coordinates of $\nu \in S^{m-1}$.

Now, we expand $|\log r(s)|^{-n''}$ about $s = \infty$ and obtain

$$\frac{1}{|\log r(s)|^{n''}} = \frac{s^{n''}}{2^{n''}} + \sum_{k=1}^{\infty} c_{k,n''}s^{n''-2k}. \tag{33}$$

Finally, we have the following expansion of the terms involving $A(r(s), \nu^t, p)$ from equation (36) in [Boggess and Raich 2023] (with $I_1, I_2, I_3 = \varnothing$):

$$\begin{aligned} &\frac{1}{(A(r(s),\nu^t,p) - i\nu_1)^{2n'+n''+m-1}} \\ &= \frac{1}{(s|p|^2 - i\nu_1)^{2n'+n''+m-1}}\left(1 + \sum_{j=1}^{\infty}\frac{\alpha_j\left(\sum_{k=1}^{\infty}\frac{p^*\cdot\left(p_{2k}(A_{\nu^t})\oplus\frac{1}{2}c_{k,1}\mathbb{I}_{n''}\right)\cdot p}{s^{2k-1}}\right)^j}{(s|p|^2 - i\nu_1)^j}\right). \end{aligned} \tag{34}$$

Now we assemble a typical term in the expansion of the integrand in (29) by multiplying the typical terms from (31), (32), (33), and (34). We summarize a typical term from each of the components that comprise (29) in the following chart:

term	typical term	notes
$\det[r(s)^{-\bar{A}_\nu}]_{K,J}$	$\frac{\nu^{\ell'-e_1}}{s^{\ell'}}$	$\ell' \ge 1$, e_1 is even, and $0 \le e_1 \le \ell'$
$\frac{B(r(s),\nu)r'(s)}{r(s)}$	$\nu^{\ell-e_2}s^{2n'-\ell-e_3-2}$	e_2 and e_3 are even, $0 \le e_2 \le \ell$
$\frac{1}{\lvert\log r(s)\rvert^{n''}}$	$s^{n''-e_4}$	e_4 is even
(34)	$\frac{1}{(s\lvert p\rvert^2 - i\nu_1)^{2n'+n''+m-1+j}}\,\frac{\nu^{j(2k-e_5)}}{s^{j(2k-1)}}$	$j, k \ge 1$, e_5 is even, and $0 \le e_5 \le k$

The typical terms of (29) that require the most care are those involving powers of s which are greater than -2. The remaining terms comprise the "remainder term" and will be handled later. From the above chart, we see that a typical term from the integrand of (29) is of the form

$$C(p,\bar{p})^{2j}\frac{s^{N_j-2-\ell_j-2k_j}\nu^{\ell_j-e_j}}{(s|p|^2-i\nu_1)^{N_j+m-1}}, \tag{35}$$

where the integers N_j, ℓ_j, e_j, k_j satisfy

$$N_j = 2n'+n''+j, \quad e_j > 0 \text{ is even}, \quad 0 < e_j \le \ell_j, \quad \text{and} \quad k_j \ge 0. \tag{36}$$

What is relevant for the proof of Lemma 4.5 below is that a typical term in the expansion satisfies

$$\text{exponent(denominator)} - \text{exponent}(s) - \text{exponent}(\nu) = m+1+E, \tag{37}$$

where E is an even, nonnegative integer.

In view of Lemma 4.2, the remainder term is analytic in $\nu \in S^{m-1}$ and $s > 3$. In addition, the typical term is

$$\frac{O(p^{\alpha'})O(\nu,s)}{(s|p|^2-i\nu_1)^{\alpha}s^{\beta}}, \tag{38}$$

where $O(\nu,s)$ is real analytic in $\nu \in S^{m-1}$ and $s \ge 3$, bounded in s, and $\beta \ge 2$.

Analysis of typical term in (35). We will now show that the integral (over $\nu \in S^{m-1}$ and $s \ge 1$) of the typical term in (35) is bounded in p. We will also show the same for the remainder term in (38).

As to the first task, let $\hat{r} = |p|^2 > 0$ and define

$$H_{N,\ell,m,e,k}(\hat{r},s,\nu) = \frac{s^{N-2-\ell-2k}\nu^{\ell-e}}{(s\hat{r}-i\nu_1)^{N+m-1}}.$$

To establish Lemma 4.4 over the region $\frac{1}{2} \le r < 1$, we need to show that for each $\ell \ge 0$, there is a uniform constant C such that

$$\left|\int_{\nu\in S^{m-1}}\int_{s=3}^{\infty} H_{N,\ell,m,e,I_3,k}(\hat{r},s,\nu)\,ds\,d\nu\right| \le C \tag{39}$$

for all $\hat{r} > 0$ near zero.

As discussed at the end of Section 7 in [Boggess and Raich 2023], we can assume the monomial $\nu^{\ell-e}$ depends on ν_1 only (by writing $\nu = (\nu_1,\nu')$ and noting that integrals of odd powers of monomials in ν' over $\nu' \in S^{m-2}$ are zero). We let $x = \nu_1$, and then the surface measure on the unit sphere in S^{m-1} can be written as

$$d\nu = (1-x^2)^{(m-3)/2}\,dx\,d\nu'$$

where $d\nu'$ is the surface measure on S^{m-2}.

The desired estimate in (39) will follow from the next lemma.

Lemma 4.5. *For any nonnegative integers N, m and ℓ with $m \geq 2$ and any even integer E with $0 \leq E \leq |\ell|$, let*

$$A^{\ell,E}_{N,m,k}(\hat{r}) = \int_{x=-1}^{1} \int_{s=3}^{\infty} \frac{(1-x^2)^{(m-3)/2} s^{N-2-\ell-2k} x^{\ell-E}\, ds\, dx}{(s\hat{r} - ix)^{N+m-1}}.$$

Then $A^{\ell,E}_{N,m,k}(\hat{r})$ is a smooth function of $\hat{r} > 0$ up to $\hat{r} = 0$.

This lemma is almost identical to Lemma 8.1 in [Boggess and Raich 2023] (the difference is in the exponent of s). Below, we give a short argument to reduce our lemma to Lemma 8.1 in [Boggess and Raich 2023].

Proof of Lemma 4.5. First write

$$A^{\ell,E}_{N,m,k}(\hat{r}) = C_{N,\ell} D_{\hat{r}}^{N-(2+\ell+2k)} \{B^{\ell,E}_m(\hat{r})\},$$

where $C_{N,\ell}$ is a constant and

$$B^{\ell,E,2k}_m(\hat{r}) = \int_{x=-1}^{1} \int_{s=3}^{\infty} \frac{(1-x^2)^{(m-3)/2} x^{\ell-E}\, ds\, dx}{(s\hat{r} - ix)^{m+\ell+2k+1}}.$$

Here, $D^j_{\hat{r}}$ indicates the j-th derivative with respect to $\hat{r}$. The index j is allowed to be negative in which case this means the $|j|$-th antiderivative with respect to $\hat{r}$ (with a particular initial condition specified at a fixed value of $\hat{r} = \hat{r}_0 > 0$).

Note, $B^{\ell,E,2k}_m(\hat{r})$ is identical to the corresponding expression in the proof of [Boggess and Raich 2023, Lemma 8.1] except that the exponent in the denominator differs by the even integer $2k \geq 0$. The rest of the proof proceeds exactly as the proof of Lemma 8.1 to show that $B^{\ell,E,2k}_m(\hat{r})$ is smooth for $\hat{r} > 0$ up to $\hat{r} = 0$. □

Analysis of Remainder Term in (38). The remainder term in (38) is

$$\frac{O(\nu, s)}{(s|p|^2 - i\nu_1)^{\alpha} s^{\beta}} \quad \text{with } \beta \geq 2 \text{ and } \alpha \geq 2.$$

As above, we set $x = \nu_1$. Since $s^{-\beta}$ is integrable over $\{s \geq 3\}$ and since $O(\nu', \nu_1, x)$ is real analytic (and hence uniformly bounded) in $\nu' \in \sqrt{1-x^2}S^{m-2}$ (($m-2$)-dimensional sphere of radius $\sqrt{1-x^2}$), the following lemma will finish the proof of Theorem 4.1 for the integral over the region $\frac{1}{2} \leq r < 1$ (and in the case $|t| \geq |z|^2$ and $0 \leq q \leq 2n' + n''$).

Lemma 4.6. *For $m \geq 2$, let*

$$R(s, \hat{r}, \nu') = \int_{x=-1}^{1} \frac{(1-x^2)^{(m-3)/2} O(\nu', x, s)\, dx}{(s\hat{r} - ix)^{\alpha}}.$$

Then $R(s, \hat{r}, \nu')$ is uniformly bounded for $s \geq 3$, $\hat{r} \geq 0$, and $\nu' \in \sqrt{1-x^2}S^{m-2}$.

This lemma is identical to Lemma 9.1 in [Boggess and Raich 2023]. The basic idea is to use Cauchy's theorem to deform the contour of integration into the upper half plane and away from $x = 0$.

***Subcase:* $|t| \geq |z|^2$ *and* $0 < r < \frac{1}{2}$.** We first assume that $n'' \geq 2$ or $n'' = 1$ and $J' \neq P$. We start with the lower r version of (27). In this case, however, we stick with the r variable, $0 \leq r \leq \frac{1}{2}$ (instead of changing to s). We rewrite this term here:

$$N^{\ell}_{K,J}(p) = \int_{\nu \in S^{m-1}} \int_{r=0}^{\frac{1}{2}} \frac{\det([r^{-\bar{A}_\nu}]_{K',J'})B(r,\nu)}{(A(r,\nu^t,p) - i\nu_1)^{2n+m-1}} \frac{d\nu\, dr}{r\,|\log r|^{n''}}. \tag{40}$$

The ℓ superscript indicates that we are working on the lower half of the r-interval. $N^{\ell}_{K,J}(p)$ is the coefficient of the $d\bar{z}_{J'}$ component of

$$\int_{\nu^t \in S^{m-1}} \int_{r=0}^{\frac{1}{2}} \frac{\det(\bar{U}(\nu^t)_{K',J'})\, d\bar{Z}(\nu^t, z)_{L'} \wedge d\bar{z}''_{K''} B_{L'}(r,\nu^t)}{(A(r,\nu^t,p) - i\nu_1)^{2n+m-1}} \frac{d\nu\, dr}{r\,|\log r|^{n''}} \tag{41}$$

Our goal is to prove the following:

Lemma 4.7. *We have*

$$|N^{\ell}_{K,J}(p)| \leq C \quad \text{for all } p = \frac{z}{|t|^{1/2}} \in \mathbb{C}^{2n'+n''}, \tag{42}$$

where C is a uniform constant.

Proof. The proof is nearly identical to the proof of Lemma 10.1 in [Boggess and Raich 2023] with the only difference being the presence of the log-terms. We give a quick outline. We are in a case where at least one of $L \cap P^c$ or $L^c \cap P$ is nonempty. In view of (23), there must be a positive power of r in the numerator of $B_{L'}(r,\nu)$. Therefore

$$\frac{|B_{L'}(r,\nu)|}{r\,|\log r|^{n''}} \leq \frac{C r^{c_0}}{r\,|\log r|^{n''}}, \tag{43}$$

where C and c_0 are uniform positive constants. Having a positive power of r in the numerator turns out to be one of the most useful terms for offsetting enough of the blow-up of $1/r$ as $r \to 0$ to guarantee integrability in r near 0. We repeatedly use this fact in both the $|t|$ large and $|z|$ large cases. In fact, as soon as there is a factor of r^{c_0} for some $c_0 > 0$ in the numerator, we can use a straightforward size argument to bound the integrand.

For $|t| \geq |z|^2$, the presence of a positive power of r allows for the following. First, the integrand of $N^{\ell}_{K,J}$ is integrable over the interval $0 < r < \frac{1}{2}$. Therefore, the integral on the right-hand side of (41) over the set $\{0 \leq r \leq \frac{1}{2}\} \times \{|\nu_1| \geq \frac{1}{2}\}$ is uniformly bounded for $p \in C^{2n'+n''}$. Thus, we turn our attention to the integral over $\{0 \leq r \leq \frac{1}{2}\} \times \{|\nu_1| \leq \frac{1}{2}\}$.

The idea is to integrate by parts in ν_1 over the integral in (40) over the interval $\{|\nu_1| \le \frac{1}{2}\}$ to reduce the power of $(A(r, \nu_t, p) - i\nu_1)$ in the denominator where $A(r, \nu_t, p)$ is defined in (24). As shown in Section 10 in [Boggess and Raich 2023], $A(r, \nu_t, p)$ is analytic in $\nu \in S^{m-1}$.

Let

$$X(r, \nu, p) := \left(\frac{\partial}{\partial \nu_1}\{A(r, \nu^t, p)\} - i\right)^{-1}$$

and note that

$$X(r, \nu, p) D_{\nu_1}\left\{\frac{-i(2n' + n'' - 2)^{-1}}{(A(r, \nu^t, p) - i\nu_1)^{2n'+n''-2}}\right\} = \frac{1}{(A(r, \nu^t, p) - i\nu_1)^{2n'+n''-1}}.$$

When integrating by parts with $X(r, \nu, p) D_{\nu_1}$ over $\{|\nu_1| \le \frac{1}{2}\}$, there will be terms involving the ν_1-derivatives of $X(r, \nu, p)$, $r^{-\bar{A}^\nu}$ and $B(r, \nu)$ that occur in the integrand of (40). These derivatives produce additional powers of $|\log r|$ which do not affect the integrability in r over $0 \le r \le \frac{1}{2}$. In addition, there are boundary terms at $|\nu_1| = \frac{1}{2}$ and these terms are uniformly integrable on $\{0 \le r \le \frac{1}{2}\} \times \{|\nu_1| = \frac{1}{2}\}$.

This process of integration by parts with $X(r, \nu, p) D_{\nu_1}$ can be repeated until the integrand in (40) involves only $\log(A(r, \nu^t, p) - i\nu_1)$ (using the principle branch of log since the A term is positive). This log-term is uniformly integrable on $\{0 \le r \le \frac{1}{2}\} \times \{|\nu_1| \le \frac{1}{2}\}$, and thus Lemma 4.7 is proved. For more details, see Section 10 of [Boggess and Raich 2023] (where z-, $\bar{z}$-, and t-derivatives are also handled in full generality).

The remaining case is $n'' = 1$ and $J' = P$ where the relevant term to estimate is given by (7) with the r-interval of integration restricted to $0 \le r \le \frac{1}{2}$. We first recall [Boggess and Raich 2023, Lemma 12.3].

Lemma 4.8. *The following functions are analytic as a function of $\nu \in S^{m-1}$:*

- $\nu \to |\det A_\nu|$.
- $\nu \to A(0, \nu, p) = \sum_{j=1}^{2n'} |\mu_j^\nu||p_j^\nu|^2$.
- $\nu \to \det(\bar{U}(\nu)_{K,P})\, d\bar{Z}(p, \nu)^P = \sum_{J \in \mathcal{I}_n} \det(\bar{U}(\nu)_{K,P}) \det[U(\nu)_{P,J}]^T\, d\bar{z}^J$.

Therefore, the functions to estimate in (7) with the r-interval of integration restricted to $0 \le r \le \frac{1}{2}$ are of the form

$$\begin{aligned} N_{K,J}^\ell(p) = \int_{\nu^t \in S^{m-1}} & \det(\bar{U}(\nu)_{K,P}) \det[U(\nu)_{P,J}]^T |\det A_{\nu^t}| \\ & \times \int_{r=0}^{\frac{1}{2}} \left(\left(\prod_{j=1}^{2n'} \frac{1}{1 - r^{|\mu_j^{\nu^t}|}} \right) \frac{1}{(A(r, \nu^t, p) - i\nu_1)^{2n'+m}} \right. \\ & \qquad \left. - \frac{1}{(A(0, \nu^t, p', 0) - i\nu_1)^{2n'+m}} \right) \frac{dr\, d\nu^t}{|\log r| r}. \end{aligned}$$

By writing

$$\frac{1}{1-r^{|\mu_j^{\nu^t}|}} = 1 + \frac{r^{|\mu_j^{\nu^t}|}}{1-r^{|\mu_j^{\nu^t}|}} \quad \text{and} \quad \frac{1+r^{|\mu_j^{\nu^t}|}}{1-r^{|\mu_j^{\nu^t}|}} = 1 + \frac{2r^{|\mu_j^{\nu^t}|}}{1-r^{|\mu_j^{\nu^t}|}}, \tag{44}$$

we can write

$$\begin{aligned} N^{\ell}_{K,J}(p) = \int_{\nu^t \in S^{m-1}} & \det(\bar{U}(\nu)_{K,P}) \det[U(\nu)_{P,J}]^T |\det A_{\nu^t}| \\ & \times \int_{r=0}^{\frac{1}{2}} \Bigg(\frac{1}{\big(\frac{2}{|\log r|}|q''|^2 + A(0,\nu^t,p) - i\nu_1\big)^{2n'+m}} \\ & \qquad - \frac{1}{(A(0,\nu^t,p',0)-i\nu_1)^{2n'+m}} \Bigg) \frac{dr\, d\nu^t}{|\log r| r} + \text{OK}, \end{aligned} \tag{45}$$

where the OK term is comprised of terms with r^{c_0} in the numerator for values $c_0 > 0$ and the discussion after (43) applies. Focusing on the integral in r, we let $s = -2/\log r$ so that $ds/s = dr/(|\log r| r)$ so that

$$\begin{aligned} & \int_{r=0}^{\frac{1}{2}} \Bigg(\frac{1}{\big(\frac{2}{|\log r|}|p''|^2 + A(0,\nu^t,p',0) - i\nu_1\big)^{2n'+m}} - \frac{1}{(A(0,\nu^t,p',0)-i\nu_1)^{2n'+m}} \Bigg) \frac{dr}{r\,|\log r|} \\ & \quad = \int_{s=0}^{\frac{2}{\log 2}} \Bigg(\frac{1}{(s|p''|^2 + A(0,\nu^t,p',0) - i\nu_1)^{2n'+m}} - \frac{1}{(A(0,\nu^t,p',0)-i\nu_1)^{2n'+m}} \Bigg) \frac{ds}{s}. \end{aligned}$$

By Lemma A.1, with $a = |p''|^2$, $b = A(0,\nu^t,p',0) - i\nu_1$, and $\gamma = 2/\log 2$,

$$\begin{aligned} & \int_{s=0}^{\frac{2}{\log 2}} \Bigg(\frac{1}{(s|p''|^2 + A(0,\nu^t,p',0) - i\nu_1)^{2n'+m}} - \frac{1}{(A(0,\nu^t,p',0)-i\nu_1)^{2n'+m}} \Bigg) \frac{ds}{s} \\ & \quad = \frac{1}{(A(0,\nu^t,p',0)-i\nu_1)^{2n'+m}} \log\Bigg(1 + \frac{2}{\log 2} \frac{|z''|^2}{A(0,\nu^t,z',0) - i\nu_1|t|}\Bigg) \\ & \qquad + E_{2n'+m}(|p''|^2, A(0,p',0) - i\alpha_1). \end{aligned} \tag{46}$$

To complete the proof of Lemma 4.7, we use Lemma 4.8 and shift the contour in ν_1 to avoid $\nu_1 = 0$. By doing this,

$$|A(0,\nu^t,p',0) - i\nu_1| \sim |p'|^2 + 1$$

on the new contour and basic size estimates now suffice. □

5. Proof of Theorem 2.5, $|z|^2 \geq |t|$

Subcase: $|z|^2 \geq |t|$ ***and*** $0 < r < \frac{1}{2}$. Analogous to the case when $|t| \geq |z|^2$, we investigate the terms in (41) but with the term $|t|^{2n'+n''+m-1}$ inserted back into the denominator of the integrand. Using (5) we are led to estimate the term

$$
\begin{aligned}
(47)\quad N^{\ell}_{K,L,J}(z,t) = \int_{\nu\in S^{m-1}} & \det(\bar U(\nu)_{K',L}) \det(U(\nu)_{L,J'}) |\det A_\nu| \\
&\times \int_{r=0}^{\frac12} \Bigg(\prod_{\substack{j\in(L')^c\cap P\\ j\in L'\cap P^c}} \frac{r^{|\mu^\nu_j|}}{1-r^{|\mu^\nu_j|}} \prod_{\substack{k\in L'\cap P\\ k\in(L')^c\cap P^c}} \frac{1}{1-r^{|\mu^\nu_k|}} \Bigg) \\
&\times \frac{1}{(A(r,\nu,z)-i\nu\cdot t)^{2n'+n''+m-1}} \frac{dr\,d\nu}{|\log r|^{n''} r},
\end{aligned}
$$

when $n''\ge 2$ or $L'\ne P$, and

$$
\begin{aligned}
(48)\quad N^{\ell}_{K,P,J}(z,t) = \int_{\nu^t\in S^{m-1}} & \det(\bar U(\nu)_{K,P}) \det[U(\nu)_{P,J}]^T |\det A_{\nu^t}| \\
&\times \int_{r=0}^{\frac12} \Bigg(\Bigg(\prod_{j=1}^{2n'} \frac{1}{1-r^{|\mu^{\nu^t}_j|}} \Bigg) \frac{1}{(A(r,\nu^t,z)-i\nu\cdot t)^{2n'+m}} \\
&\qquad - \frac{1}{(A(0,\nu^t,z',0)-i\nu\cdot t)^{2n'+m}} \Bigg) \frac{dr\,d\nu^t}{|\log r|\, r},
\end{aligned}
$$

when $n''=1$ and $L'=P$.

We start with the case $n''=1$ and $L'=P$ because the analysis of (48) is virtually identical to that of (45). The same reductions and equalities hold, and factoring $|t|$ back into (46) is the calculation that we need. The size estimates are more straightforward than the $|t|$ large case because we do not have to shift the contour.

We now focus on (47). We first assume $|z'|^2\ge|z''|^2$. The upper bound estimates in this case will follow directly from size estimates. Since $|A(r,\nu,z)-i\nu\cdot t|\ge c|z'|^2$ and either $1/(r|\log r|^{n''})$ is integrable near $r=0$ ($n''\ge 2$) or there is an r^{c_0} term in the numerator ($n''=1$ and $J'\ne P$), we use size estimates to establish

$$|N^{\ell}_{K,L,J}(z,t)| \le \frac{C}{|z'|^{2(2n'+n''+m-1)}}.$$

The $|z''|\ge|z'|$ estimate requires more care. In the case that there is a factor of r^{c_0} in the numerator, the estimate is straightforward with size estimates, as bounding $(1+r^{\mu})/(1-r^{\mu})$ by $|\log r|$ shows that

$$
\begin{aligned}
\frac{r^{c_0}}{|A(r,\nu,z)-i\nu\cdot t|^{2n'+n''+m-1}|\log r|^{n''} r} &\le \frac{r^{c_0}}{\big(\frac{|z|^2}{|\log r|}\big)^{2n'+n''+m-1}|\log r|^{n''} r} \\
&= \frac{1}{|z|^{2(2n'+n''+m-1)}} r^{c_0-1}|\log r|^{2n'+m-1}
\end{aligned}
$$

is integrable at 0, and the estimate

$$|N^{\ell}_{K,L,J}(z,t)| \le C|z|^{-2(2n'+n''+m-1)} \tag{49}$$

holds. A factor r^{c_0} will always be present whenever N is hypoelliptic, that is, when $0 \le q < n'$ or $n'+n'' < q \le n$. Additionally, it will also be present when $n' \le q \le n'+n''$ as long as $L \ne P$ and (49) holds, a better estimate than (9).

It remains to analyze

$$N^{\ell}_{K,P,J}(z,t) = \int_{\nu\in S^{m-1}} \det(\bar{U}(\nu)_{K',P})\det(U(\nu)_{P,J'})|\det A_\nu| \\ \times \int_{r=0}^{\frac12} \prod_{k=1}^{2n'} \frac{1}{1-r^{|\mu^\nu_k|}} \frac{1}{(A(r,\nu,z)-i\nu\cdot t)^{2n'+n''+m-1}} \frac{dr\,d\nu}{|\log r|^{n''} r},$$

when $n'' \ge 2$. As we have seen, once we have a positive power of r in the numerator, we can use size estimates to obtain the estimates in (8). This is relevant for the error estimates when $|z'|^2 \ge |t|$ in two ways. First, we can apply (44) to replace $\prod_{k=1}^{2n'} 1/(1-r^{|\mu^\nu_k|})$ by 1 and an OK term. Second, since

$$A(r,\nu,z) = \frac{2}{|\log r|}|z''|^2 + \sum_{j=1}^{2n'} |\mu^\nu_j||z^\nu_j|^2 + \sum_{j=1}^{2n'} \frac{2r^{|\mu^\nu_j|}}{1-r^{|\mu^\nu_j|}}|z^\nu_j|^2,$$

we can write

$$\frac{1}{(A(r,\nu,z)-i\nu\cdot t)^{2n'+n''+m-1}} \\ = \frac{1}{(A_0(r,\nu,z)+i\nu\cdot t)^{2n'+n''+m-1}} + \frac{O(r^{c_0})}{(A_0(r,\nu,z)+i\nu\cdot t)^{2n'+n''+m-1}} \\ + \frac{O(r^{c_0}|z'|^2)}{(A_0(r,\nu,z)+i\nu\cdot t)^{2n'+n''+m}},$$

where

$$A_0(r,\nu,z) = \frac{2}{|\log r|}|z''|^2 + \sum_{j=1}^{2n'} |\mu^\nu_j||z^\nu_j|^2$$

and $c_0 > 0$. The first error term arises from estimating $B_P(r,\nu)$ by $|\det(A_\nu)|$. The second error term uses the expansion

$$\frac{1}{(V+\zeta)^{2n'+n''+m-1}} = \frac{1}{V^{2n'+n''+m-1}} + \sum_{j=1}^{\infty} \alpha_j \frac{\zeta^j}{V^{2n'+n''+m-1+j}}$$

and therefore has $A_0(r,\nu,z)$ raised to one higher power than in the main term. When integrated, however, the estimate from the extra degree in the denominator is offset by the additional factor of $|z|^2$ in the numerator.

This means that the remaining term to analyze is

$$\int_{\nu\in S^{m-1}} \det(\overline{U(\nu)_{K',P}})\det(U(\nu)^T_{P,J'})|\det A_\nu| \\ \times\int_0^{\frac{1}{2}} \frac{1}{(A_0(r,\nu,z)-i\nu\cdot t)^{2n'+n''+m-1}}\frac{dr\,d\nu}{r|\log r|^{n''}}.$$

We factor out $2|z''|^2$ from the denominator and let

$$a=\sum_{j=1}^{2n'}|\mu_j^\nu|\frac{|z_j^\nu|^2}{|z''|^2}-i\nu\cdot\frac{t}{|z''|^2}.$$

Note that $1/\log 2+a=O(1)$. By (53), we compute

$$\begin{aligned}\int_0^{\frac{1}{2}} &\frac{1}{(A_0(r,\nu,z)-i\nu\cdot t)^{2n'+n''+m-1}}\frac{dr\,d\nu}{r|\log r|^{n''}}\\ &=\frac{1}{|z''|^{2(2n'+n''+m-1)}}\int_0^{\frac{1}{\log 2}}\frac{s^{n''-2}}{(s+a)^{2n'+n''+m-1}}\,ds\\ &=\frac{1}{|z''|^{2(2n'+n''+m-1)}}\sum_{\ell=0}^{n''-2}\binom{n''-2}{\ell}\frac{(-1)^{n''-\ell}}{2n'+n''+m-1-\ell-1}\left(\frac{1}{a^{2n'+m}}+O(1)\right)\\ &=\frac{C}{|z''|^{2(n''-1)}}\frac{1}{\left(\sum_{j=1}^{2n'}|\mu_j^\nu||z_j^\nu|^2-i\nu\cdot t\right)^{2n'+m}}+O(|z''|^{-2(2n'+n''+m-1)}).\end{aligned}$$

If $|z'|^2\ge|t|$, then $\sum_{j=1}^{2n'}|\mu_j^\nu||z_j^\nu|^2-i\nu\cdot t=O(|z'|^2)$, and size estimates produce $O(|z|^{-2(n''-1)}|z'|^{-2(2n'+m)})$, the desired estimate. If, on the other hand, $|t|\ge|z'|^2$, then we treat the integral similarly to the large $|t|$ case, rotating in ν and factoring out $|t|$ to produce the integral

$$\frac{C}{|z''|^{2(n''-1)}|t|^{2n'+m}}\int_{\nu\in S^{m-1}}\det(\overline{U(\nu)_{K',P}})\det(U(\nu)^T_{P,J'})\frac{|\det A_\nu|}{\left(\sum_{j=1}^{2n'}|\mu_j^\nu||q_j^\nu|^2-i\nu_1\right)^{2n'+m}}\,d\nu$$

where $q_j^\nu=z_j^\nu/|t|^{1/2}$. The integrand in the above integral is $O(1)$ when $|\nu_1|\ge\frac{1}{2}$. In the case $|\nu_1|\le\frac{1}{2}$, we handled this exact type of integral in [Boggess and Raich 2023, (68)] and showed that the above integral is bounded by $C/(|z''|^{2(n''-1)}|t|^{2n'+m})$ (in fact, this bound is sharp).

Subcase: $|z|^2\ge|t|$ and $\frac{1}{2}<r<1$. We are finally in a position to finish the proof of the estimates in Theorem 2.5. As with the previous subsection, we include the term $|t|^{-(2n'+n''+m-1)}$ in the integrand. Define $N^u_{K,J}(z,t)$ analogously to $N^u_{K,J}(p)$ in (27), with the r-integral over $\left[\frac{1}{2},1\right]$ and including the term $|t|^{-(2n'+n''+m-1)}$ in the integrand. We follow the analysis of the $|t|$ large case through (35) to obtain

$$\begin{aligned} N^u_{K,J}(z,t) &:= \int_{s=3}^{\infty}\int_{\nu\in S^{m-1}} \frac{\text{typical term in } N_{K,J}(p)}{|t|^{2n'+n''+m-1}}\,d\nu\,ds \\ &= \int_{s=3}^{\infty}\int_{\nu\in S^{m-1}} C(z,\bar z)^{2j}\,\frac{s^{N_j-2-\ell_j-K_j}\nu^{\ell_j-e_j}}{(s|z|^2-i\nu_1|t|)^{N_j+m-1}}\,d\nu\,ds, \end{aligned}$$

where $N_j = 2n'+n''+j$ and $\ell_j, K_j \ge 0$, $m\ge 2$. Since $|z|^2 \ge |t|$ and $|\nu|=1$, we use size estimates and drop the t-term in the denominator to obtain

$$\begin{aligned} |N^u_{K,J}(z,t)| &\le \int_{s=3}^{\infty}\int_{\nu\in S^{m-1}} \frac{C\,|z|^{2j}s^{N_j-1-\ell_j-K_j}}{(s|z|^2)^{N_j+m-1}}\,d\nu\,ds \\ &\le \int_{s=3}^{\infty}\int_{\nu\in S^{m-1}} \frac{C}{|z|^{2(2n'+n''+m-1)}}\cdot\frac{1}{s^3}\,d\nu\,ds \end{aligned}$$

after taking into account the constraints on ℓ_j, K_j, m. Therefore

$$|N^u_{K,J}(z,t)| \le \frac{C}{|z|^{2(2n'+n''+m-1)}}, \tag{50}$$

and we have established the estimates in Theorem 2.5.

Higher derivatives. As mentioned in the introduction, we will refer the reader to [Boggess and Raich 2023] for details on how to handle the estimates for higher derivatives. Here is the basic idea on how to obtain the estimates for derivatives. Note that z and $\bar z$ appear quadratically in $A(r,\nu,z)$ and t only appears in the $\nu\cdot t$ term. Thus, differentiating (20) once with a z' or $\bar z'$ derivative adds one more factor of $A(r,\nu,z)-i\nu\cdot t$ to the denominator along with a linear z' or $\bar z'$ term in the numerator. The overall estimate in (8) changes by a factor of $(|z|^2+|t|)^{-1/2}$. By contrast, a t-derivative of (20) also adds a factor of $A(r,\nu,z)-i\nu\cdot t$ to the denominator but with no compensating factor of z', $\bar z'$ or t in the numerator. Thus the overall estimate in (8) changes by a factor of $(|z|^2+|t|)^{-1}$. The z''- and $\bar z''$-derivatives behave similarly. This is the basic idea behind why there is a $\frac{1}{2}$ in front of the exponents $|I_1|$ and $|I_2|$, which represent z- or $\bar z$-derivatives, and not in front of $|I_3|$, which represents t-derivatives.

6. Conclusion of the proof Theorem 2.5 — sharpness of the estimates

We will show the dominant term in (9) is nonzero for the index $K=P$ provided the eigenvectors of A_ν depend continuously on ν.

We focus on the $d\bar z'_P$ component of N_P (here, the value of n'' is not important because we are focusing on the integral in ν). Ignoring the power of $|z'|$ out front, this term is

$$N_P = \int_{\nu \in S^{m-1}} |\det U(\nu)_{P,P}|^2 \frac{|\det A_\nu|}{\left(\frac{2}{|\log r|}|z''|^2 + \sum_{j=1}^{2n'} |\mu_j^\nu||z_j^\nu|^2 - i\nu\cdot t\right)^{2n'+m}}\, d\nu.$$

Consider the case when $|t|$ is smaller than $|z'|^2 < |z''|^2$. We factor out $|z'|^{2(2n'+m)}$ from the denominator and obtain $N_P = |z'|^{-2(2n'+m)}\tilde{N}_P$ where

$$\tilde{N}_P = \int_{\nu \in S^{m-1}} |\det U(\nu)_{P,P}|^2 \frac{|\det A_\nu|}{\left(\frac{2}{|\log r|}|q''|^2 + \sum_{j=1}^{2n'} |\mu_j^\nu||q_j^\nu|^2 - i\nu\cdot q_t\right)^{2n'+m}}\, d\nu$$

with $q'' = z''/|z'|$, $q_j^\nu = |z_j^\nu|^2/|z'|^2$ and $q_t = t/|z'|^2$.

Now take a limit as $q_t \to 0$ and we obtain

$$\lim_{q_t\to 0} \tilde{N}_P = \int_{\nu \in S^{m-1}} |\det U(\nu)_{P,P}|^2 \frac{|\det A_\nu|}{\left(\frac{2}{|\log r|}|q''|^2 + \sum_{j=1}^{2n'} |\mu_j^\nu||q_j^\nu|^2\right)^{2n'+m}}\, d\nu. \tag{51}$$

Now $\mu_j^\nu \neq 0$ for $j = 1, \dots, 2n'$; and $\sum_{j=1}^{2n'} q_j^\nu = 1$; and $\det A_\nu \neq 0$ for all $\nu \in S^{m-1}$. So if the integral on the right-hand side of (51) vanishes, then we conclude that $\det U(\nu)_{P,P} = 0$ for all $\nu \in S^{m-1}$ except for a set of zero measure in ν. Thus, to conclude the proof of Theorem 2.5, we have only to show

$$\int_{\nu \in S^{m-1}} |\det U(\nu)_{[P,P]}|^2\, d\nu > 0, \tag{52}$$

where $U(\nu)$ is the unitary matrix which diagonalizes A_ν and where P is the set of indices corresponding to the positive eigenvalues of A_ν and $U(\nu)_{[P,P]}$ is the $P\times P$ minor matrix of $U(\nu)$. We may assume $P = \{1, 2, \dots, n'\}$ and $P^c = \{n'+1, \dots, 2n'\}$ where here, the eigenvalues are counted with multiplicity. We also let $N = 2n'$.

Define

- $N_0 =$ essential sup{the number of distinct eigenvalues of $A_\nu : \nu \in S^{m-1}$},
- $S_0 = \{\nu \in S^{m-1}$: the number of distinct eigenvalues of $A_\nu = N_0\}$,
- $\lambda_j(\nu)$, $1 \le j \le N_0$, are the distinct eigenvalues of A_ν for $\nu \in S_0$,
- $E_j(\nu)$ equals the eigenspace of $\lambda_j(\nu)$ in $\mathbb{C}^N$ for $\nu \in S_0$.

Note that N_0 is an even number between 1 and $N = 2n'$. The set S_0 has positive measure by the definition of essential sup. Since there are only a finite number of choices for $\dim_{\mathbb{C}}\{E_j(\nu)\}$, we can shrink S_0, but still with positive measure, so that $\dim_{\mathbb{C}}\{E_j(\nu)\}$ is constant in $\nu \in S_0$ for each $1 \le j \le N_0$.

Although we are not assuming the eigenvalues are continuous in $\nu \in S^{m-1}$, the $\lambda_j(\cdot)$ are measurable functions that are locally integrable on S^{m-1}. Using the usual row and column operations together with Gram–Schmidt, we can find an orthonormal set of eigenvectors for the eigenspace $E_j(\nu)$ of the form

$$U_j^k(\nu), \quad 1 \le j \le N_0,\ 1 \le k \le \dim_{\mathbb{C}}\{E_j(\nu)\},$$

where these C^N-valued functions are measurable and integrable in $\nu \in S^{m-1}$. Now let $U(\nu)$ be the unitary matrix with column vectors $U_j^k(\nu)$.

By removing a set of measure zero from S_0, we can assume that every point in S_0 lies in the Lebesgue set of each $\lambda_j(\,\cdot\,)$ and $U_j^k(\,\cdot\,)$ as well as all n-fold products of the component entries of $U_j^k(\,\cdot\,)$. Now fix any $\nu_0 \in S_0$ and choose coordinates for $\mathbb{C}^N$ which diagonalize A_{ν_0} where the first n diagonal entries correspond to the positive eigenvalues of A_{ν_0}. Note that in these coordinates, $U_{[P,P]}(\nu_0)$ is the identity matrix.

Now, for $\varepsilon > 0$, define

$$B(\nu_0, \varepsilon) = \{\nu \in S^{m-1} : |\nu - \nu_0| < \varepsilon\}.$$

From the Lebesgue differentiation theorem,

$$\lim_{\varepsilon\to 0} \frac{1}{|B(\nu_0,\varepsilon)|} \int_{\nu\in B(\nu_0,\varepsilon)} |\det U_{[P,P]}(\nu)|^2\, d\nu \to |\det U_{[P,P]}(\nu_0)|^2 = 1,$$

where $|B(\nu_0, \varepsilon)|$ is the Lebesgue measure of $B(\nu_0, \varepsilon)$ relative to S^{m-1}. We conclude that, for small enough $\varepsilon > 0$,

$$\int_{\nu\in B(\nu_0,\varepsilon)} |\det U_{[P,P]}(\nu)|^2\, d\nu > 0,$$

and this implies (52).

Appendix: Calculus computations

Lemma A.1. *Suppose that $a, \gamma > 0$, $b \neq 0$, and $k \in \mathbb{N}$. Then*

$$\int_0^\gamma \frac{1}{s(as+b)^k} - \frac{1}{sb^k}\, ds = \frac{1}{b^k} \log\left(1 + \gamma\frac{a}{b}\right) + E(a,b),$$

where $E_k(a, b, \gamma)$ is comprised of a sum of terms of the form

$$E_k(a,b,\gamma) = \sum_{\ell=0}^{k} \frac{c_\ell}{b^\ell(a\gamma+b)^{k-\ell}}$$

for some constants c_ℓ.

Proof. The proof is a computation using a partial fraction decomposition, recognizing that the $1/s$ terms cancel (so that the integral converges). $\square$

In that vein, we also have the following. We compute

$$\begin{aligned}\int_0^{\frac{1}{\log 2}} \frac{s^{k-2}}{(s+a)^\ell}\,ds &= \int_0^{\frac{1}{\log 2}} \frac{(s+a-a)^{k-2}}{(s+a)^\ell}\,ds \\ &= \sum_{i=0}^{k-2}\binom{k-2}{i}(-1)^{k-2-i}\int_0^{\frac{1}{\log 2}} \frac{a^{k-2-i}}{(s+a)^{\ell-i+2}}\,ds \\ &= \sum_{i=0}^{k-2}\binom{k-2}{i}\frac{(-1)^{k-1-i}}{(-\ell+i+1)}\left(\frac{1}{a^{\ell-k+1}} - \frac{a^{k-2-i}}{\left(\frac{1}{\log 2}+a\right)^{\ell-i-1}}\right). \qquad (53)\end{aligned}$$

Acknowledgements

The authors thank the anonymous referee for the careful reading and thoughtful commentary.

References

[Beals et al. 2000] R. Beals, B. Gaveau, and P. C. Greiner, "Hamilton–Jacobi theory and the heat kernel on Heisenberg groups", *J. Math. Pures Appl.* (9) **79**:7 (2000), 633–689. MR Zbl

[Biard and Straube 2017] S. Biard and E. J. Straube, "L^2-Sobolev theory for the complex Green operator", *Internat. J. Math.* **28**:9 (2017), art. id. 1740006. MR Zbl

[Boggess 1991] A. Boggess, *CR manifolds and the tangential Cauchy–Riemann complex*, CRC Press, Boca Raton, FL, 1991. MR Zbl

[Boggess and Raich 2009] A. Boggess and A. Raich, "A simplified calculation for the fundamental solution to the heat equation on the Heisenberg group", *Proc. Amer. Math. Soc.* **137**:3 (2009), 937–944. MR Zbl

[Boggess and Raich 2011] A. Boggess and A. Raich, "The $\Box_b$-heat equation on quadric manifolds", *J. Geom. Anal.* **21**:2 (2011), 256–275. MR Zbl

[Boggess and Raich 2013] A. Boggess and A. Raich, "Fundamental solutions to $\Box_b$ on certain quadrics", *J. Geom. Anal.* **23**:4 (2013), 1729–1752. MR Zbl

[Boggess and Raich 2020] A. Boggess and A. Raich, "The fundamental solution to $\Box_b$ on quadric manifolds, part 2: L^p regularity and invariant normal forms", *Complex Anal. Synerg.* **6**:2 (2020), art. id. 13. MR Zbl

[Boggess and Raich 2021] A. Boggess and A. Raich, "The fundamental solution to $\Box_b$ on quadric manifolds, part 3: Asymptotics for a codimension 2 case in $\mathbb{C}^4$", *J. Geom. Anal.* **31**:11 (2021), 11529–11583. MR Zbl

[Boggess and Raich 2022a] A. Boggess and A. Raich, "Analysis on quadrics", *Complex Anal. Synerg.* **8**:4 (2022), art. id. 18. MR Zbl

[Boggess and Raich 2022b] A. Boggess and A. Raich, "The fundamental solution to $\Box_b$ on quadric manifolds, part 1: General formulas", *Proc. Amer. Math. Soc. Ser. B* **9** (2022), 186–203. MR Zbl

[Boggess and Raich 2023] A. Boggess and A. Raich, "The fundamental solution to $\Box_b$ on quadric manifolds with nonzero eigenvalues", *Trans. Amer. Math. Soc. Ser. B* **10** (2023), 507–541. MR Zbl

[Chen and Shaw 2001] S.-C. Chen and M.-C. Shaw, *Partial differential equations in several complex variables*, AMS/IP Studies in Advanced Mathematics **19**, American Mathematical Society, Providence, RI, 2001. MR Zbl

[Folland and Stein 1974a] G. B. Folland and E. M. Stein, "Estimates for the $\bar{\partial}_b$ complex and analysis on the Heisenberg group", *Comm. Pure Appl. Math.* **27** (1974), 429–522. MR Zbl

[Folland and Stein 1974b] G. B. Folland and E. M. Stein, "Parametrices and estimates for the $\bar{\partial}_b$ complex on strongly pseudoconvex boundaries", *Bull. Amer. Math. Soc.* **80** (1974), 253–258. MR Zbl

[Gaveau 1977] B. Gaveau, "Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents", *Acta Math.* **139**:1-2 (1977), 95–153. MR Zbl

[Hulanicki 1976] A. Hulanicki, "The distribution of energy in the Brownian motion in the Gaussian field and analytic-hypoellipticity of certain subelliptic operators on the Heisenberg group", *Studia Math.* **56**:2 (1976), 165–173. MR Zbl

[Lewy 1957] H. Lewy, "An example of a smooth linear partial differential equation without solution", *Ann. of Math.* (2) **66** (1957), 155–158. MR Zbl

[Machedon 1988] M. Machedon, "Estimates for the parametrix of the Kohn Laplacian on certain domains", *Invent. Math.* **91**:2 (1988), 339–364. MR Zbl

[Nagel and Stein 2006] A. Nagel and E. M. Stein, "The $\overline{\partial}_b$-complex on decoupled boundaries in $\mathbb{C}^n$", *Ann. of Math.* (2) **164**:2 (2006), 649–713. MR

[Nagel et al. 2001] A. Nagel, F. Ricci, and E. M. Stein, "Singular integrals with flag kernels and analysis on quadratic CR manifolds", *J. Funct. Anal.* **181**:1 (2001), 29–118. MR Zbl

[Peloso and Ricci 2003] M. M. Peloso and F. Ricci, "Analysis of the Kohn Laplacian on quadratic CR manifolds", *J. Funct. Anal.* **203**:2 (2003), 321–355. MR Zbl

[Rainer 2011] A. Rainer, "Quasianalytic multiparameter perturbation of polynomials and normal matrices", *Trans. Amer. Math. Soc.* **363**:9 (2011), 4945–4977. MR Zbl

Received August 26, 2023. Revised December 24, 2023.

ALBERT BOGGESS
SCHOOL OF MATHEMATICAL AND STATISTICAL SCIENCES
ARIZONA STATE UNIVERSITY
TEMPE, AZ
UNITED STATES
boggess@asu.edu

ANDREW RAICH
DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF ARKANSAS
FAYETTEVILLE, AR
UNITED STATES
araich@uark.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 327, No. 2, 2023
https://doi.org/10.2140/pjm.2023.327.267

MORE PROPERTIES OF OPTIMAL POLYNOMIAL APPROXIMANTS IN HARDY SPACES

RAYMOND CHENG AND CHRISTOPHER FELDER

We study optimal polynomial approximants (OPAs) in the classical Hardy spaces on the unit disk, H^p ($1 < p < \infty$). For fixed $f \in H^p$ and $n \in \mathbb{N}$, the OPA of degree n associated to f is the polynomial which minimizes the quantity $\|qf - 1\|_p$ over all complex polynomials q of degree less than or equal to n. We begin with some examples which illustrate, when $p \neq 2$, how the Banach space geometry makes the above minimization problem interesting. We then weave through various results concerning limits and roots of these polynomials, including results which show that OPAs can be witnessed as solutions of certain fixed-point problems. Finally, using duality arguments, we provide several bounds concerning the error incurred in the OPA approximation.

1. Introduction

This paper concerns a minimization problem in classical Hardy spaces on the unit disk $\mathbb{D}$,

$$H^p := \left\{ f \in \mathrm{Hol}(\mathbb{D}) : \sup_{0 \le r < 1} \int_0^{2\pi} |f(re^{i\theta})|^p \, d\theta < \infty \right\},$$

where $\mathrm{Hol}(\mathbb{D})$ denotes the collection of holomorphic functions on $\mathbb{D}$. As is standard, for $1 \le p < \infty$, we define the norm of $f \in H^p$ as

$$\|f\|_p := \left(\sup_{0 \le r < 1} \int_0^{2\pi} |f(re^{i\theta})|^p \, d\theta \right)^{1/p}.$$

MSC2020: primary 30E10; secondary 46E30.
Keywords: optimal polynomial approximant, Pythagorean inequality, duality, fixed point.

When $p = \infty$, we have the set of bounded analytic functions

$$H^\infty := \{f \in \mathrm{Hol}(\mathbb{D}) : \sup_{z \in \mathbb{D}} |f(z)| < \infty\},$$

with corresponding norm

$$\|f\|_\infty := \sup_{z \in \mathbb{D}} |f(z)|.$$

We will frequently view these spaces as subspaces of the Lebesgue spaces $L^p := L^p(\mathbb{T}, dm)$, where dm is normalized Lebesgue measure on the unit circle $\mathbb{T}$.

Our main objects of study are *optimal polynomial approximants* (OPAs) in Hardy spaces; these are solutions to the minimization problem

$$\inf_{q \in \mathscr{P}_n} \|qf - 1\|_p,$$

where $f \in H^p$ and $\mathscr{P}_n$ is the set of complex polynomials of degree less than or equal to n. We point out that the infimum above is actually a minimum. In our context, this means the problem of finding a degree n OPA can be restated as finding the solution to

$$\inf_{h \in f\mathscr{P}_n} \|h - 1\|_p,$$

which is given by the *metric projection* of 1 on the subspace $f\mathscr{P}_n$. A priori, the minimizing argument may not be unique. However, when $1 < p < \infty$, it is well known that there is, in fact, a unique minimizing polynomial due to the uniform convexity of the space; for in a such a space, any closed subspace enjoys a unique nearest-point property. When $p \neq 2$, the projection is *nonlinear*, which starkly contrasts with the Hilbert space setting.

For $p \neq 2$, this problem was originally studied by Centner [10], and considered again in an additional paper by Centner and the authors [11]. We will give some background now, but point the reader to [10; 11] for more thorough exposition, and to [5; 6; 7; 8; 9; 10; 15; 17; 18] for relevant work on OPAs in various Hilbert spaces.

When $p = 2$, this problem was first studied by engineers in work related to digital filter design. The problem reemerged later as a potential way to study cyclic vectors for the forward shift (see [4] for historical discussion). This renewed interest is evidenced by many papers over the last decade (again, see [4], as well as [1; 3] for recent results in the weighted and noncommutative settings, respectively). Other than the work in [14], these results concern only Hilbert spaces, where the geometry makes computation of OPAs an explicit (but nontrivial!) linear algebra exercise. For example, in H^2, the coefficients (say, $a_0, \dots, a_n$) of the OPA of degree n associated to a function $f \in H^2$ can be found via the linear system

$$(\langle S^j f, S^k f\rangle_{H^2})_{0 \le j,k \le n}(a_0, \dots, a_n)^T = (\overline{f(0)}, 0, \dots, 0)^T, \tag{1.0.1}$$

where S is the forward shift operator, given by $f(z) \mapsto zf(z)$ (see, e.g., [15, Theorem 2.1]).

In the Banach space setting (for example, H^p, $p \neq 2$), there is not a direct analogue of this exercise, and the nonlinearity of the metric projection makes explicit calculation of OPAs a highly nontrivial task. In the next section, we will state precisely the definition of optimal polynomial approximant. Before moving there, let us give an outline of the paper:

- Section 2 will formally introduce the OPA problem, give some background information concerning the geometry of Banach spaces, and provide some examples illustrating how this geometry differs from that of Hilbert spaces.
- The results of Section 3 are broken into three parts:
  - convergence of OPAs under variance of the parameters of the OPA problem (e.g., n, p, and f),
  - the location of roots of OPAs,
  - constant and linear OPAs as solutions to a fixed-point problem.
- Using duality, Section 4 establishes various bounds for the error $\|qf - 1\|_p$.

2. Preliminaries and geometric oddities

We begin here by providing some background material concerning the geometry of Banach spaces, followed by several examples in H^p which illustrate some oddities that arise when $p \neq 2$.

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be vectors belonging to a normed linear space $\mathscr{X}$. We say that $\boldsymbol{x}$ is *orthogonal* to $\boldsymbol{y}$ in the Birkhoff–James sense [2; 16] if

$$\|\boldsymbol{x} + \beta \boldsymbol{y}\|_{\mathscr{X}} \geqslant \|\boldsymbol{x}\|_{\mathscr{X}} \tag{2.0.1}$$

for all scalars β. In this situation we write $\boldsymbol{x} \perp_{\mathscr{X}} \boldsymbol{y}$. In the case $\mathscr{X} = L^p$, let us write $\perp_p$ instead of $\perp_{L^p}$, and similarly for $\mathscr{X} = H^p$.

For $1 < p < \infty$, there is also a function-theoretic test for p-orthogonality, which we note now.

Theorem 2.0.2 (James [16]). *Suppose $1 < p < \infty$. Then for f and g belonging to L^p, we have*

$$f \perp_p g \iff \int_{\mathbb{T}} |f|^{p-2} \bar{f} g \, dm = 0, \tag{2.0.3}$$

where any occurrence of "$|0|^{p-2}0$" in the integrand is interpreted as zero.

In light of (2.0.3) we define, for a measurable function f and any $s > 0$,

$$f^{\langle s \rangle} := |f|^{s-1} \bar{f}. \tag{2.0.4}$$

If $f \in L^p$, then $f^{\langle p-1\rangle} \in L^q$, where q is the classical Hölder conjugate to p, satisfying $\frac{1}{p} + \frac{1}{q} = 1$. For $g \in L^p$ and $f \in L^q$, we use the standard notation for the dual pairing

$$\langle f, g\rangle = \int_{\mathbb{T}} f\bar{g}\, dm,$$

and from (2.0.3), we have

$$f \perp_p g \iff \langle g, f^{\langle p-1\rangle}\rangle = 0. \tag{2.0.5}$$

Consequently, the relation $\perp_p$ is linear in its second argument when $1 < p < \infty$, and it then makes sense to speak of a vector being orthogonal to a subspace. We use this now to formally define OPAs.

Definition 2.0.6 (OPA). Let $1 < p < \infty$ and let $f \in H^p \setminus \{0\}$. Given a nonnegative integer n, the *n-th optimal polynomial approximant* to $1/f$ in H^p is the polynomial solving the minimization problem

$$\min_{q \in \mathscr{P}_n} \|qf - 1\|_p,$$

where $\mathscr{P}_n$ is the set of complex polynomials of degree less than or equal to n. This polynomial exists, is unique, and will be denoted by

$$q_{n,p}[f].$$

Given f, n and p, we refer to the problem of finding the corresponding OPA as the *OPA problem*.

Considering previous discussion on the metric projection, it is immediate that

$$1 - q_{n,p}[f]f \perp_p \bigvee\{f, zf, z^2 f, \ldots, z^n f\}.$$

We will use the notation $z^k f$ and $S^k f$ interchangeably if there is no risk of confusion. We will also use the notation $[f]_p$ to denote the closure of $\bigvee\{f, zf, z^2 f, z^3 f, \ldots\}$ in H^p, i.e.,

$$[f]_p := \overline{\bigvee\{f, zf, z^2 f, z^3 f, \ldots\}}^{H^p}.$$

In order to avoid trivialities, we will also often ask that $f(0) \neq 0$; in the case that $f(0) = 0$, this is equivalent to $1 \perp_p f$, and so the metric projection of 1 onto $f\mathscr{P}_n$ is identically zero.

In connection with Birkhoff–James orthogonality, there is a version of the Pythagorean theorem for L^p. This theorem takes the form of a family of inequalities relating the lengths of orthogonal vectors with that of their sum [12, Corollary 3.4].

Theorem 2.0.7. *Suppose that $x \perp_p y$ in L^p. If $p \in (1, 2]$, then*

$$\|x+y\|_p^p \leqslant \|x\|_p^p + \frac{1}{2^{p-1}-1}\|y\|_p^p,$$
$$\|x+y\|_p^2 \geqslant \|x\|_p^2 + (p-1)\|y\|_p^2.$$

If $p \in [2, \infty)$, then

$$\|x+y\|_p^p \geqslant \|x\|_p^p + \frac{1}{2^{p-1}-1}\|y\|_p^p,$$
$$\|x+y\|_p^2 \leqslant \|x\|_p^2 + (p-1)\|y\|_p^2.$$

These Pythagorean inequalities enable us to obtain bounds and estimates when $p \neq 2$, in lieu of exact calculations possible in the Hilbert space case.

The following examples illustrate some of the ways the geometry of H^p $(p \neq 2)$ can run counterintuitive to experience in Hilbert space. Although these examples may not be immediately surprising to the Banach space enthusiast, we relay them for the general functional analyst, especially working in linear approximation problems, as interesting observations related to natural geometric questions.

Example 2.0.8. In a Hilbert space, an orthogonal projection is always a contraction. However, when $p \neq 2$, the norm of the metric projection of a vector can exceed the length of the vector itself.

Consider the linear OPA for $f(z) = 1 + 0.5z$ in H^4. Numerically, we find that $Q(z) := q_{1,4}[f] \approx 0.9771018 - 0.4339644z$, and thus

$$\|Qf\|_4^4 \approx 1.10294 > 1.$$

Example 2.0.9. In H^2, it is simple to verify that if $F(0) = 0$, then, for $c \in \mathbb{C}$, the quantity

$$\|c+F\|_2$$

is minimized when $c = 0$. However, this is not the case when $p \neq 2$.

For example, let $p = 4$ and $F(z) = z + 2z^2$. Note that since F has real coefficients, the value of c which minimizes $\|c+F\|_p$ must also be real, or else uniqueness of nearest points is violated. In turn, we have

$$\|c+F\|_4^4 = 33 + 8c + 20c^2 + c^4.$$

Numerically, this is minimized when $c \approx -0.199209$. In particular, the value of the minimizing argument can be nonzero when $p \neq 2$.

Example 2.0.10. Notice for $f \in H^2$ and any $n > 0$, we have

$$1 - q_{n,2}[f]f \perp_2 zf \quad \text{and} \quad 1 \perp_2 zf.$$

Using linearity, we have

$$q_{n,2}[f]f = 1 - (1 - q_{n,2}[f]f) \perp_2 zf.$$

It is natural to ask if this is true when $p \neq 2$, that is, is it true in general that

$$q_{n,p}[f]f \perp_p zf?$$

Let us take $p = 4$, $n = 1$, and $f(z) = 1 + 2z + z^8$. Numerically, one can find that

$$\int_{\mathbb{T}} (q_{1,p}[f]f)^{\langle p-1\rangle} zf \, dm \approx 0.00355837,$$

which is nonzero, and so orthogonality fails. This illustrates that for $p \neq 2$, the relation $\perp_p$ fails to be linear in its first argument, and so $q_{n,p}[f]f$ is not necessarily orthogonal to zf.

Example 2.0.11. For $f, g \in H^2$, an exercise shows that if $\|f\|_2 \leqslant \|g\|_2$, then $\|1 + zf\|_2 \leqslant \|1 + zg\|_2$. Might a similar statement hold for $p \neq 2$? The following example shows that the answer is *no*.

Let $p = 4$ and choose

$$\begin{aligned} f(z) &= 0.9(1 + z + z^2), \\ g(z) &= -1 - z - z^2. \end{aligned}$$

It is immediate that $\|f\|_4 < \|g\|_4$. However, numerically, we find

$$\begin{aligned} \|1 + zf\|_4^4 &\approx 31.9339, \\ \|1 + zg\|_4^4 &\approx 20.0000. \end{aligned}$$

With these examples in hand, it may now be reasonable to suspect that OPAs have a dependence on p which is highly nonlinear. In general, this is true. Let us demonstrate this with what we describe as the OPA "error" — the quantity $\|q_{n,p}[f]f - 1\|_p$. We use this as motivation in Section 3, where we study the p-dependence of OPAs.

Example 2.0.12. For m a positive integer, consider $f(z) = 1 + 2z^m$. Let us show

$$\|q_{0,2}[f]f - 1\|_2 \neq \|q_{0,4}[f]f - 1\|_4.$$

For $p = 2$ and any scalar $a \in \mathbb{R}$, we have

$$\begin{aligned} \|af - 1\|_2^2 &= \int_0^{2\pi} ([a-1]^2 + 2ae^{im\theta})([a-1]^2 + 2ae^{-im\theta}) \frac{d\theta}{2\pi} \\ &= (a-1)^2 + 4a^2. \end{aligned}$$

Notice that the result of the integration is the extraction of the constant OPA (again, this constant must be real, as the coefficients of f are real). Minimizing this expression (by differentiating with respect to a), we find

$$a = \tfrac{1}{5}.$$

This yields

$$\|af - 1\|_2 = \sqrt{\left(\tfrac{1}{5} - 1\right)^2 + 4\left(\tfrac{1}{5}\right)^2} = \sqrt{\tfrac{4}{5}} \approx 0.894427.$$

For $p = 4$, we have

$$\|af - 1\|_4^4 = \int_0^{2\pi} (a + 2ae^{im\theta} - 1)^2 (a + 2ae^{-im\theta} - 1)^2 \, \frac{d\theta}{2\pi},$$

and one may extract the constant term as

$$(a-1)^4 + 16a^2(a-1)^2 + 16a^4.$$

Next, one may numerically find that $a \approx 0.121991$ minimizes the above expression.

Finally, this yields

$$\|q_{0,4}[f]f - 1\|_4 \approx \sqrt[4]{0.781388} \approx 0.940192.$$

In addition to the error, one may also notice that the OPAs themselves vary with p. For example, letting $f(z) = 1 + 2z + z^8$, one may numerically find that

$$\begin{aligned} q_{0,4}[f] &\approx 0.0970262, \\ q_{0,6}[f] &\approx 0.0674066. \end{aligned}$$

However, this is not always the case(!). The following example is a generalization of [10, Example 6.1], which showed that the constant OPAs for $f(z) = 1 - z$ do not vary with p.

Example 2.0.13. Let $1 < p < \infty$ and let $f \in H^p$. Let $\lambda \in \mathbb{C}$ and let $h = 1 + f$. Suppose that $|f(e^{it})| = 1$ a.e. on $\mathbb{T}$ and $\overline{f(e^{-it})} = f(e^{it})$ (i.e., the Fourier coefficients of f are real).

Putting $a = q_{0,p}[h]$, we observe

$$\begin{aligned} \|a(1+f(e^{it}))-1\|_p &= \|af(e^{it})+a-1\|_p \\ &= \|a+\overline{f(e^{it})}(a-1)\|_p && \text{(multiply by } \bar{f}\text{, inner)} \\ &= \|a+f(e^{-it})(a-1)\|_p && \text{(real coefficients)} \\ &= \|a+f(e^{it})(a-1)\|_p && (t \mapsto -t) \\ &= \|-a+1-1-f(e^{it})(a-1)\|_p && \text{(multiply by } -1 \text{ and add 0)} \\ &= \|(1-a)(1+f(e^{it}))-1\|_p. \end{aligned}$$

This tells us that $a = \frac{1}{2}$, which is independent of p. Note that if f is any Blaschke product with real zeros, the hypotheses above are satisfied.

3. Limits and continuity

In this section, we provide results which relate to varying the parameters in the OPA problem (i.e., the degree n, the value of p, and the function f). We first deal with this directly. Then, as a corollary, the first subsection below discusses the possible set of roots for OPAs. In the final subsection, we show that OPAs (in certain cases) are solutions to a fixed-point problem. All of these results enable us to make estimates concerning OPAs, knowing that exact computation is difficult when $p \neq 2$.

We begin by recording, without proof, a known result about metric projections (see, e.g., [13, Proposition 4.8.3]).

Proposition 3.0.1. *Let* $1 < p < \infty$. *Let* $f \in H^p$ *with* $f(0) \neq 0$ *and let* h *be the metric projection of* 1 *onto* $[f]_p$. *Then, in norm,*

$$q_{n,p}[f]f \to h \quad \text{as } n \to \infty.$$

In the following proposition, for $1 < p < \infty$ and $f \in H^p$, we write $Q_\infty f$ for the metric projection of 1 onto $[f]_p$, understanding that Q_∞ need not be a bona fide H^p function. The next result tells us something about the error incurred by approximating $q_{n,p}[f]$ using the Taylor polynomials of Q_∞, when the (rather strict) assumption of norm convergence holds.

Proposition 3.0.2. *Let* $1 < p < \infty$, *and* $f \in H^p$. *Suppose that the representation*

$$Q_\infty(z)f(z) = \sum_{k=0}^{\infty} \alpha_k z^k f(z)$$

converges in norm. Then there exist a positive constant C *and an index* N *such that*

$$\|q_{n,p}[f]f - Q_{(n)}f\|_p^r \leqslant C\,\|Q_\infty f - Q_{(n)}f\|_p$$

for all $n \geqslant N$, *where* $Q_{(n)}(z) = \sum_{k=0}^{n} \alpha_k z^k$, *and* r *and* K *are the applicable Pythagorean parameters.*

Proof. From the orthogonality relation

$$1 - q_{n,p}[f]f \perp_p q_{n,p}[f]f - Q_{(n)}f,$$

the Pythagorean inequality gives

$$\|1 - q_{n,p}[f]f\|_p^r + K\,\|q_{n,p}[f]f - Q_{(n)}f\|_p^r \leqslant \|1 - Q_{(n)}f\|_p^r.$$

Rearrange and estimate to get

$$\begin{aligned}
K\,\|q_{n,p}[f]f - Q_{(n)}f\|_p^r &\leqslant \|1 - Q_{(n)}f\|_p^r - \|1 - q_{n,p}[f]f\|_p^r \\
&\leqslant \|1 - Q_{(n)}f\|_p^r - \|1 - Q_\infty f\|_p^r \\
&\leqslant r\,\|1 - Q_{(n)}f\|_p^{r-1}(\|1 - Q_{(n)}f\|_p - \|1 - Q_\infty f\|_p) \\
&\leqslant r\,\|1 - Q_{(n)}f\|_p^{r-1}\|Q_\infty f - Q_{(n)}f\|_p \\
&\leqslant 2r\,\|1 - Q_\infty f\|_p^{r-1}\|Q_\infty f - Q_{(n)}f\|_p,
\end{aligned}$$

for n sufficiently large. In the third step we applied the elementary inequality

$$a^r - b^r \leqslant r a^{r-1}(a-b),$$

for $0 < b < a$ and $r > 1$.

This verifies the claim, with $C = 2r\,\|1 - Q_\infty f\|_p^{r-1}/K$. □

The previous proposition can be applied when f is any polynomial; we record that result now.

Proposition 3.0.3. *Suppose that* $1 < p < \infty$. *Let* $z_1, z_2, \ldots, z_N$ *be a sequence of nonzero points of* $\mathbb{D}$, *and define*

$$f(z) := \left(1 - \frac{z}{z_1}\right)\left(1 - \frac{z}{z_2}\right)\cdots\left(1 - \frac{z}{z_N}\right).$$

Set $r = 2$ *if* $1 < p \leqslant 2$, *and set* $r = p$ *if* $2 < p < \infty$. *Then the metric projection* h *of the unit constant function* 1 *onto the subspace* $[f]_p$ *of* H^p *has a norm convergent representation*

$$h(z) = \sum_{k=0}^{\infty} b_k z^k f(z),$$

and there exists a positive constant C *such that*

$$\left\| q_{n,p}[f]f - \sum_{k=0}^{n} b_k z^k f \right\|_p^r \leqslant C \left\| \sum_{k=n+1}^{\infty} b_k z^k f \right\|_p \tag{3.0.4}$$

for all positive integers n.

We omit the proof here, but note that the metric projection h must vanish at the zeros of f; in turn, boundedness of the difference-quotient operator, given by

$$(B_w f)(z) = \frac{f(z) - f(w)}{z - w}, \quad z, w \in \mathbb{D},$$

(applied where $f(w) = 0$) then ensures the norm convergent representation.

3.1. ***Continuity.*** As discussed earlier, OPAs generally vary with p. We discuss this variance here, first showing that when f is a bounded function, $q_{n,p}[f]$ varies continuously with respect to p.

Lemma 3.1.1. *Let $f \in H^\infty$ with $f(0) \neq 0$ and let $d \in \mathbb{N}$. If $(p_k)_k \subseteq (1, \infty)$ with $p_k \to p \in (1, \infty)$, then $q_{d,p_k}[f]$ converges to $q_{d,p}[f]$ uniformly as $k \to \infty$.*

Proof. Let us write

$$\begin{aligned} f(z) &= \sum_{j=0}^{\infty} f_j z^j, \\ q_{d,p}[f](z) &= a_0 + a_1 z + a_2 z^2 + \cdots + a_d z^d, \\ q_{d,p_k}[f](z) &= a_0^{(k)} + a_1^{(k)} z + a_2^{(k)} z^2 + \cdots + a_d^{(k)} z^d. \end{aligned}$$

Since $f \in H^\infty$, Hölder yields

$$\left| \int_{\mathbb{T}} f z^{-j}\, dm \right| \leqslant \|f\|_\infty$$

for all $j \geq 0$. Hence, all of the coefficients f_k are bounded by $\|f\|_\infty$.

Letting p_k' be the dual exponent of p_k, we observe

$$|a_0^{(k)} f_0 - 1| = \left| \int_{\mathbb{T}} (q_{d,p_k}[f] f - 1)\, dm \right| \leqslant \|q_{d,p_k}[f] f - 1\|_{p_k} \|1\|_{p_k'} \leqslant 1,$$

and so the sequence $\{a_0^{(k)}\}$ is bounded.

Further, since

$$|a_0^{(k)} f_j + a_1^{(k)} f_{j-1} + \cdots + a_j^{(k)} f_0| \leqslant \left| \int_{\mathbb{T}} (q_{d,p_k}[f] f - 1) z^{-j}\, dm \right| \leqslant 1,$$

it follows

$$|a_j^{(k)}| \leqslant \frac{|a_0^{(k)} f_j + a_1^{(k)} f_{j-1} + \cdots + a_{j-1}^{(k)} f_1| + 1}{|f_0|},$$

for all $k \in \mathbb{N}$ and $1 \leqslant j \leqslant d$. That is, $\{a_j^{(k)}\}_{k=1}^{\infty}$ is also a bounded sequence for $1 \leqslant j \leqslant d$. By passing to a subsequence and relabeling, we can assume that $\{q_{n,p_k}[f]\}_{k=1}^{\infty}$ is a uniformly convergent sequence of polynomials, which converges to some polynomial, say, $A(z) = a_0 + a_1 z + \cdots + a_d z^d \in \mathscr{P}_d$.

Now, for $0 \leqslant j \leqslant d$, recall the orthogonality equations

$$\int_{\mathbb{T}} [q_{d,p_k}[f] f - 1]^{\langle p_k - 1 \rangle} z^j f\, dm = 0.$$

Taking $k \to \infty$ and invoking uniform convergence, we find that

$$\int_{\mathbb{T}} [Af - 1]^{\langle p_k - 1\rangle} z^j f \, dm = 0,$$

for $0 \leqslant j \leqslant d$ (the taking of $\langle p_k - 1\rangle$ powers also being well behaved). By uniqueness of the optimal polynomial, it must be that

$$A(z) = q_{d,p}[f](z).$$

Since every subsequence of the originally given sequence $\{q_{d,p_k}[f]\}_{k=1}^{\infty}$ has a further subsequence that converges to the same limit $q_{d,p}[f]$, it must be that

$$q_{d,p_k}[f] \to q_{d,p}[f]$$

uniformly. □

We now present another continuity result — continuity in f. In particular, if $f_k \to f$ in H^p, then $q_{n,p}[f_k] \to q_{n,p}[f]$. Before establishing this result, we need a couple of lemmas.

Lemma 3.1.2. *Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. If $\varphi_k \to \varphi$ in L^p, then $\varphi_k^{\langle p-1\rangle} \to \varphi^{\langle p-1\rangle}$ in L^q.*

Proof. First, we check that

$$\int_{\mathbb{T}} |\varphi^{\langle p-1\rangle}|^q \, dm = \int_{\mathbb{T}} |\varphi|^{(p-1)q} \, dm = \int_{\mathbb{T}} |\varphi|^p \, dm,$$

and so $\varphi^{\langle p-1\rangle} \in L^q$; similarly $\varphi_k^{\langle p-1\rangle} \in L^q$.

Next, we apply the generalized dominated convergence theorem, using the sequential bound

$$|\varphi_k^{\langle p-1\rangle} - \varphi^{\langle p-1\rangle}|^q \leqslant 2^{q-1}(|\varphi_k|^p + |\varphi|^p) \quad \text{a.e.-}dm,$$

with the Carleson–Hunt theorem supplying pointwise convergence almost everywhere. The conclusion is

$$\int_{\mathbb{T}} |\varphi_k^{\langle p-1\rangle} - \varphi^{\langle p-1\rangle}|^q \, dm \to 0,$$

as claimed. □

Below, we use the standard notation $\hat{f}(n)$ to denote the n-th Fourier coefficient of a function $f \in L^p$.

Lemma 3.1.3. *Let $1 < p < \infty$. If $\varphi \in H^p$, then*

$$\|\varphi\|_p^r \geqslant |\hat{\varphi}(0)|^r + K|\hat{\varphi}(1)|^r + K^2|\hat{\varphi}(2)|^r + \cdots,$$

where r and K are the lower Pythagorean parameters.

Proof. This follows immediately from the orthogonality relations

$$z^k \perp_p z^m H^p \quad \text{for all } m > k \geqslant 0,$$

and repeated application of the lower Pythagorean inequality. □

We are now prepared to prove the aforementioned result.

Theorem 3.1.4. *Suppose that $1 < p < \infty$ and $n \in \mathbb{N}$. Let $f_k \in H^p$ and let $Q_k := q_{n,p}[f_k]$ for each $k \in \mathbb{N}$. If $f_k \to f$ in H^p, and $f(0) \neq 0$, then $Q_k \to Q := q_{n,p}[f]$.*

Proof. Let us first handle the case $n = 1$, and write $Q_k(z) = a_k + b_k z$ for $q_{1,p}[f]$.

Since $f_k(0) \to f(0)$, and $f(0) \neq 0$, there is no harm in assuming that there exists $c > 0$ such that $|f_k(0)| \geqslant c$ for all k.

From the relation

$$1 \perp_p zH^p$$

we see that

$$1 \geqslant \|1 - Q_k f_k\|_p^r \geqslant |1 - a_k f_k(0)|^r + K \, \|Q_k f_k - a_k f_k(0)\|_p^r,$$

where r and K are the lower Pythagorean parameters. It follows that

$$1 \geqslant |1 - a_k f_k(0)|,$$

implying that

$$|a_k| \leqslant \frac{2}{c}$$

for all k. Thus $\{a_k\}$ is a bounded complex sequence, from which we can extract a convergent subsequence, which for now we relabel as the original sequence.

Next, subharmonicity and the triangle inequality yield

$$c\,|b_k| \leqslant |b_k f_k(0)| \leqslant \|b_k z f_k\|_p \leqslant \|1 - (a_k + b_k z) f_k\|_p + \|1 - a_k f_k\|_p.$$

The last expression on the right side is uniformly bounded as k varies through $\mathbb{N}$, and hence $\{b_k\}$ is a bounded sequence. Once again we may draw a convergent subsequence, and relabel it so that

$$Q_k = a_k + b_k z$$

converges uniformly to some $R(z) = a + bz$.

It needs to be shown that $R = Q := q_{1,p}[f]$. For this we rely on the elementary result that if $v_k \to v$ in a Banach space and $\lambda_k \to \lambda$ in its dual space, then $\lambda_k(v_k) \to \lambda(v)$.

We apply this, identifying

$$\begin{aligned} v_k &= f_k, \\ v &= f, \\ \lambda_k(\,\cdot\,) &= \int_{\mathbb{T}} (1-[a_k+b_k z]f_k)^{\langle p-1\rangle}\,(\,\cdot\,)\,dm, \\ \lambda(\,\cdot\,) &= \int_{\mathbb{T}} (1-[a+bz]f)^{\langle p-1\rangle}\,(\,\cdot\,)\,dm. \end{aligned}$$

Then Lemma 3.1.2 ensures that $\lambda_k \to \lambda$, as needed.

The conclusion is that $\lambda(v) = \lim_{k\to\infty} \lambda_k(v_k) = \lim_{k\to\infty} 0 = 0$, or

$$1-(a+bz)f \perp_p f.$$

Repeat this argument with the choices

$$v_k = zf_k \quad \text{and} \quad v = zf$$

to see that

$$1-(a+bz)f \perp_p zf$$

as well. This forces $R(z) = Q(z) = a+bz = q_{1,p}[f](z)$, as claimed.

So far, we only know that there is a subsequence that satisfies the claim. However, we see that every subsequence of the original sequence $\{f_k\}$ has a further subsequence for which the linear OPAs tend *to the same limit* $a+bz$, the linear OPA from f being unique. This proves that in fact the full sequence $\{a_k+b_k z\}$ converges to $a+bz$.

This verifies the claim when $n=1$.

More generally, for arbitrary $n \in \mathbb{N}$, let

$$Q_k(z) = q_{n,p}[f](z) = a_0^{(k)} + a_1^{(k)} z + \cdots + a_n^{(k)} z^n.$$

From Lemma 3.1.3, we get

$$\begin{aligned} 1 &\geqslant \|1-Q_k f_k\|_p^r \\ &\geqslant |1-a_0^{(k)} f_0|^r + \sum_{m=1}^{\infty} K^m\, |a_0^{(k)} f_m + a_1^{(k)} f_{m-1} + \cdots + a_m^{(k)} f_0|^r, \end{aligned}$$

which implies

$$\frac{1}{K^{m/r}} \geqslant |a_0^{(k)} f_m + a_1^{(k)} f_{m-1} + \cdots + a_m^{(k)} f_0|$$

for all m.

We know that $a_0^{(k)}$ is bounded in k. It is also easy to see that $|f_j| \leqslant \|f\|_p$ for all j. If $a_0^{(k)}, a_1^{(k)}, \ldots, a_j^{(k)}$ are also bounded in k, then the relation

$$|a_{j+1}^{(k)}| \leqslant \frac{1}{K^{m/r}|f_0|} + \left| \frac{a_0^{(k)} f_{j+1}}{f_0} + \frac{a_1^{(k)} f_j}{f_0} + \cdots + \frac{a_j^{(k)} f_1}{f_0} \right|$$

ensures that $a_{j+1}^{(k)}$ is bounded as well. This proves that all of the coefficients of Q_k are uniformly bounded in k.

Arguing as before, we may find a subsequence from $\{Q_k\}$ that converges uniformly, and the limit must be $q_{n,p}[f]$. In fact, this must be the limit of the original sequence. $\square$

3.2. *Roots of OPAs.* As a corollary to the last continuity theorem, we begin this subsection with two results concerning the set of possible OPA roots. Let us first establish some notation.

Definition 3.2.1. For $1 < p < \infty$ and $n \geq 0$, we define the set of possible roots of OPAs of degree n in H^p as

$$\Omega_{n,p} := \{w \in \mathbb{C} : \text{there exists } f \in H^p \text{ such that } f(0) \neq 0 \text{ with } q_{n,p}[f](w) = 0\},$$

and let

$$\Omega_p := \bigcup_{n \geq 0} \Omega_{n,p}.$$

Note that $\Omega_{0,p} = \varnothing$ for all $p \in (1, \infty)$. We have an immediate proposition concerning these sets.

Proposition 3.2.2. *For $1 < p < \infty$ and each $n \geq 1$, we have $\Omega_{n,p} \subseteq \Omega_{1,p}$, and therefore*

$$\Omega_p = \Omega_{1,p}.$$

Proof. Suppose $w \in \Omega_{n,p}$ with $q_{n,p}[f](w) = 0$. Put $q_{n,p}[f] = (z - w)\tilde{q}$. Then, by optimality, we have

$$\begin{aligned} \|q_{1,p}[\tilde{q} f]\tilde{q} f - 1\|_p &\leq \|(z - w)\tilde{q} f - 1\|_p \\ &= \|q_{n,p}[f] f - 1\|_p \\ &\leq \|q_{1,p}[\tilde{q} f]\tilde{q} f - 1\|_p, \end{aligned}$$

and we deduce that $q_{1,p}[\tilde{q} f] = q_{n,p}[f]/\tilde{q} = z - w$, which implies that $w \in \Omega_{1,p}$. $\square$

Presently, we see that the set of OPA roots must contain the set $\mathbb{C} \setminus \overline{\mathbb{D}}$.

Proposition 3.2.3. *Let $1 < p < \infty$. If $w \in \mathbb{C} \setminus \overline{\mathbb{D}}$, then there exists $f \in H^p$ such that $q_{1,p}[f]$ has the root w, and so*

$$\mathbb{C} \setminus \overline{\mathbb{D}} \subseteq \Omega_p.$$

Proof. Let $w \in \mathbb{C} \setminus \overline{\mathbb{D}}$ and let

$$f(z) := \frac{1}{z - w},$$

which belongs to H^p for all $p \in (1, \infty)$. Further,

$$\|1 - Qf\|_p = 0$$

when

$$Q(z) = z - w.$$

Therefore, it must be that $q_{n,p}[f](z) = z - w$ for all $n \geqslant 1$. Hence, $w \in \Omega_p$. $\square$

We now show that this set is connected and symmetric under rotation.

Proposition 3.2.4. *For* $1 < p < \infty$, *the set* Ω_p *is rotationally symmetric and connected.*

Proof. We begin by establishing rotational symmetry. Let $f \in H^p$, and suppose $q_{1,p}[f] = a(z - w)$ (by Proposition 3.2.2, it suffices to take the linear OPA). Then, for any γ with $|\gamma| = 1$,

$$\begin{aligned}
\|1 - a(z - w)\|_p^p &= \int_0^{2\pi} |1 - a(z - w) f(z)|^p \, dm(z) \\
&= \int_0^{2\pi} |1 - a(\gamma\zeta - w) f(\gamma\zeta)|^p \, |\gamma|^p \, dm(\zeta) \\
&= \int_0^{2\pi} |1 - a(\gamma\zeta - w) f(\gamma\zeta)|^p \, dm(\zeta) \\
&= \int_0^{2\pi} |1 - (a\gamma)(\zeta - \bar{\gamma} w) f(\gamma\zeta)|^p \, dm(\zeta).
\end{aligned}$$

It must be that $(a\gamma)(z - \bar{\gamma} w)$ is the linear OPA for $f(\gamma z)$, for otherwise, by reversing these steps from

$$\|1 - q_{1,p}[f(\gamma z)] f(\gamma z)\|_p^p$$

we obtain a contradiction.

This shows that if w is an OPA root, then so is $\bar{\gamma} w$ for all γ, $|\gamma| = 1$. That is, the set Ω_p is rotationally symmetric.

Next, suppose that f and g belong to H^p, with real coefficients, and with $f(0) > 0$ and $g(0) > 0$. Let their linear OPA roots be r and R, respectively, where $0 < r < R$. By the continuity of the map $F \mapsto q_{1,p}[F]$, we see that the set of linear OPA roots of the collection of functions $tf + (1 - t)g$, $0 \leqslant t \leqslant 1$, must be an interval containing $[r, R]$; this is because the collection of functions is connected, and continuous maps preserve connectivity. Note that $tf(0) + (1 - t)g(0) > 0$ for all t, as required for the linear OPA to be nontrivial. Consequently, Ω_p is path connected, and hence connected. $\square$

3.3. ***Fixed-point approach.*** Again, we mention that computing OPAs when $p \neq 2$ is a challenging task. Here, we explore the idea of OPAs being fixed points of an iterative process. We begin with the degree-zero case and then move to the degree-one case.

Theorem 3.3.1. *Let $2 < p < \infty$, and let $f \in H^p$ be a nonconstant function. Then the degree-zero OPA $q_{0,p}[f]$ is the unique solution to the fixed-point equation*

$$\zeta = \Phi(\zeta),$$

where $\Phi : \mathbb{C} \mapsto \mathbb{C}$ is given by

$$\Phi(\zeta) := \left(\int_{\mathbb{T}} |1 - \zeta f|^{p-2} \bar{f}\, dm \right) \left(\int_{\mathbb{T}} |1 - \zeta f|^{p-2} |f|^2\, dm \right)^{-1}.$$

For any $\lambda_1 \in \mathbb{C}$, the sequence $\{\lambda_k\}$ given by $\lambda_{k+1} = \Phi(\lambda_k)$ converges to $q_{0,p}[f]$.

Proof. Since $p > 2$, and

$$1 = \frac{2}{p} + \frac{p-2}{p},$$

the parameters $\frac{p}{2}$ and $\frac{p}{p-2}$ are Hölder conjugates of each other. Hence Hölder's inequality gives

$$\begin{aligned} &\int_{\mathbb{T}} |1 - \zeta f|^{p-2} |f|^2\, dm \\ &\qquad \leqslant \left(\int_{\mathbb{T}} |1 - \zeta f|^{(p-2)p/(p-2)}\, dm \right)^{(p-2)/p} \left(\int_{\mathbb{T}} |f|^{2(p/2)}\, dm \right)^{2/p} < \infty. \end{aligned}$$

Furthermore, since f is nonconstant, the integral in the denominator of Φ is nonzero for any value of ζ.

When $\zeta \neq 0$, we have

$$\begin{aligned} |\zeta| \int_{\mathbb{T}} |1 - \zeta f|^{p-2} |f|\, dm &\leqslant \int_{\mathbb{T}} |1 - \zeta f|^{p-2} \left(|1 - \zeta f| + 1\right)\, dm \\ &= \int_{\mathbb{T}} |1 - \zeta f|^{p-1}\, dm + \int_{\mathbb{T}} |1 - \zeta f|^{p-2}\, dm \\ &< \infty; \end{aligned}$$

and, when $\zeta = 0$,

$$\int_{\mathbb{T}} |1 - \zeta f|^{p-2} |f|\, dm = \int_{\mathbb{T}} |1 - 0 \cdot f|^{p-2} |f|\, dm = \int_{\mathbb{T}} |f|\, dm < \infty.$$

This verifies that Φ is well defined for all $\zeta \in \mathbb{C}$.

In fact, Φ is continuous and bounded. Continuity of the numerator and denominator of Φ at any point ζ_0 can be established by a dominated convergence argument, with respective dominating functions

$$2^{p-2}(1+C|f|^{p-2})|f| \quad \text{and} \quad 2^{p-2}(1+C|f|^{p-2})|f|^2,$$

where $C > |\zeta_0|^{p-2}$. Continuity at infinity is established by

$$\begin{aligned}\lim_{\zeta\to\infty} \frac{\int_{\mathbb{T}} |1-\zeta f|^{p-2}\bar{f}\,dm}{\int_{\mathbb{T}} |1-\zeta f|^{p-2}|f|^2\,dm} &= \lim_{\zeta\to\infty} \frac{\int_{\mathbb{T}} |1/\zeta - f|^{p-2}\bar{f}\,dm}{\int_{\mathbb{T}} |1/\zeta - f|^{p-2}|f|^2\,dm} \\ &= \lim_{\zeta\to\infty} \frac{\int_{\mathbb{T}} |f|^{p-2}\bar{f}\,dm}{\int_{\mathbb{T}} |f|^{p}\,dm}.\end{aligned}$$

Consequently, Φ is a bounded function. For any choice of $\lambda_1 \in \mathbb{C}$, define $\lambda_{k+1} = \Phi(\lambda_k)$ for all $k = 1, 2, 3, \ldots$. The resulting sequence $\{\lambda_k\}$ is a bounded sequence, and must contain a convergent subsequence, $\{\lambda_{n_k}\}$, with $\lambda_k \to \lambda \in \mathbb{C}$. Continuity ensures that

$$\lambda = \Phi(\lambda),$$

which is to say that

$$\begin{aligned}\lambda \int_{\mathbb{T}} |1-\lambda f|^{p-2} f\bar{f}\,dm &= \int_{\mathbb{T}} |1-\lambda f|^{p-2}\bar{f}\,dm \\ 0 &= \int_{\mathbb{T}} |1-\lambda f|^{p-2}(1-\lambda f)\bar{f}\,dm,\end{aligned}$$

or $1-\lambda f \perp_p f$. This shows that $\lambda = q_{0,p}[f]$.

But any subsequence of $\{\lambda_k\}$ must have a further subsequence that converges to the same limit. Thus the sequence $\{\lambda_k\}$ itself must converge to $\lambda = q_{0,p}[f]$. □

We now discuss the linear case, first recording some notation.

Let a linear polynomial $Q_1(z) = a_1 + b_1 z$ be given and, for $k \geqslant 1$, let

$$\begin{aligned}\begin{bmatrix} a_{k+1} \\ b_{k+1}\end{bmatrix} &= \begin{bmatrix} C_k & \overline{D_k} \\ D_k & C_k\end{bmatrix}^{-1} \begin{bmatrix} A_k \\ B_k\end{bmatrix} \\ &= \frac{1}{|C_k|^2+|D_k|^2}\begin{bmatrix} C_k & -\overline{D_k} \\ -D_k & C_k\end{bmatrix}\begin{bmatrix} A_k \\ B_k\end{bmatrix} \\ &= \frac{1}{|C_k|^2+|D_k|^2}\begin{bmatrix} cA_kC_k - B_k\overline{D_k} \\ B_kC_k - A_kD_k\end{bmatrix},\end{aligned} \tag{3.3.2}$$

where

$$\begin{aligned} A_k &= \int_{\mathbb{T}} |1-Q_k f|^{p-2}\bar{f}\,dm, & B_k &= \int_{\mathbb{T}} |1-Q_k f|^{p-2}\overline{zf}\,dm, \\ C_k &= \int_{\mathbb{T}} |1-Q_k f|^{p-2}|f|^2\,dm, & D_k &= \int_{\mathbb{T}} |1-Q_k f|^{p-2}\bar{z}|f|^2\,dm,\end{aligned} \tag{3.3.3}$$

and $Q_k(z) = a_k + b_k z$. This determines a sequence of linear polynomials.

Theorem 3.3.4. *Let $2 < p < \infty$, and suppose that $f \in H^p$ is a nonconstant polynomial with $f(0) \neq 0$. If $Q_k(z) = a_k + b_k z$, $k \geqslant 0$, is the sequence of linear polynomials arising from* (3.3.2), *then Q_k converges to $q_{1,p}[f]$.*

Proof. If Q_1 is identically zero, then by inspection we see that Q_2 is not the zero polynomial. Thus, by relabeling if necessary, let us assume Q_1 is not identically zero.

By the hypotheses on f, the expression $w - Qf$ is a nonconstant polynomial for any complex number w and linear polynomial Q; hence $|w - Qf|^{p-2}$ will be integrable on the unit circle.

Consider the expression, with integrals being taken over the circle,

$$\Phi(Q) := \left| \frac{\begin{array}{r} \int |Qf|^{p-2} \bar{f}\, dm \int |Qf|^{p-2} |f|^2\, dm \qquad\qquad\qquad \\ - \int |Qf|^{p-2} \overline{zf}\, dm \int |Qf|^{p-2} \bar{z} |f|^2\, dm \end{array}}{\left|\int |Qf|^{p-2} |f|^2\, dm\right|^2 + \left|\int |Qf|^{p-2} \bar{z} |f|^2\, dm\right|^2} \right|,$$

as Q varies over set

$$\mathscr{Q} := \{a + bz \in \mathscr{P}_1 : \max\{|a|, |b|\} = 1\}.$$

Under the assumptions on f, the denominator is bounded away from zero. Thus $\Phi(Q)$ is a continuous function on a compact set, and achieves its maximum. In fact, the value of $\Phi(Q)$ is indifferent to rescaling Q, except for multiplying it by zero.

From this we can further deduce that the values of

$$\Psi(Q, w) := \left| \frac{\begin{array}{r} \int |w - Qf|^{p-2} \bar{f}\, dm \int |w - Qf|^{p-2} |f|^2\, dm \qquad\qquad\qquad \\ - \int |w - Qf|^{p-2} \overline{zf}\, dm \int |w - Qf|^{p-2} \bar{z} |f|^2\, dm \end{array}}{\left|\int |w - Qf|^{p-2} |f|^2\, dm\right|^2 + \left|\int |w - Qf|^{p-2} \bar{z} |f|^2\, dm\right|^2} \right|$$

are uniformly bounded for $Q \in \mathscr{Q}$ and $|w| \leqslant 1$.

Next, notice that for any nonzero linear polynomial $Q(z) = a + bz$ we have

$$\begin{aligned} \int_{\mathbb{T}} |1 - Qf|^{p-2} \bar{f}\, dm &= \int_{\mathbb{T}} |1 - (a + bz) f|^{p-2} \bar{f}\, dm \\ &= |c|^{p-2} \int_{\mathbb{T}} \left| \frac{1}{c} - \left(\frac{a}{c} + \left[\frac{b}{c} \right] z \right) f \right|^{p-2} \bar{f}\, dm, \end{aligned}$$

where $c := \max\{|a|, |b|\}$. This is to say that the value of A_k in (3.3.3) scales in a simple way with c, with the result that f is multiplied by a member of $\mathscr{Q}$, and the 1 inside the integrand is replaced by $\frac{1}{c}$. Similar remarks apply to the formulas for B_k, C_k, and D_k.

Consequently, when A_k, B_k, C_k and D_k are assembled together to yield a_{k+1} and b_{k+1}, the scaling factors $|c|^{p-2}$ attached to each integral cancel.

Let us write $c_k := \max\{|a_k|, |b_k|\}$. The above observations establish that $|c_{k+1}|$ is uniformly bounded as k varies over such indices that $|c_k| \geqslant 1$.

For the other values of k, for which $|c_k| < 1$, the corresponding expressions for $|1 - Q_k f|^{p-2}$ are again uniformly bounded in the obvious way, implying that the resulting c_{k+1} are also uniformly bounded.

This shows that $\{Q_k\}$ is a bounded sequence of linear polynomials, which must therefore have a convergent subsequence. The limit is a linear polynomial Q_∞, which satisfies the orthogonality conditions for $q_{1,p}[f]$, and hence must be the OPA. Uniqueness of the OPA ensures that, in fact, every subsequence of $\{Q_k\}$ has a further subsequence that converges to $q_{1,p}[f]$. In conclusion, we have

$$\lim_{k\to\infty} Q_k = q_{1,p}[f]. \qquad \square$$

We end this section by noting that Theorems 3.3.1 and 3.3.4 are only established for $2 < p < \infty$, and, in the degree-one case, only for polynomials. It is currently unclear if these results extend to $1 < p < 2$, or if analogous results hold for higher-degree OPAs.

4. Error bounds and duality arguments

The present section is concerned with estimating (both above and below) the quantity $\|q_{n,p}[f]f - 1\|_p$, i.e., the "error" in the optimal polynomial approximation algorithm. We begin by employing some duality methods, first recalling a fundamental result from classical functional analysis, tailored to our setting.

Lemma 4.0.1. *Let $1 < p < \infty$ and $f \in H^p$. For any $n \in \mathbb{N}$, we have*

$$\|q_{n,p}[f]f - 1\|_p = \Big[\inf_{\psi\in L^q}\{\|\psi\|_{L^q} : \psi_0 = 1,\ \langle z^k f, \psi\rangle = 0 \text{ for all } 0 \leqslant k \leqslant n\}\Big]^{-1}.$$

Proof. By an elementary duality theorem of functional analysis, with respect to the pairing

$$\langle f, g\rangle = \int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})}\,\frac{d\theta}{2\pi},$$

we have

$$\begin{aligned}
&\|q_{n,p}[f]f - 1\|_p \\
&\quad = \inf\{\|Qf - 1\|_p : Q \in \mathscr{P}_n\} \\
&\quad = \mathrm{dist}_{H^p}(1, \mathscr{P}_n f) \\
&\quad = \mathrm{dist}_{L^p}(1, \mathscr{P}_n f) \\
&\quad = \|1\|_{[(\mathscr{P}_n f)^\perp]^*} \\
&\quad = \sup\left\{\frac{|\langle \psi, 1\rangle|}{\|\psi\|_{L^q}} : \psi \in (\mathscr{P}_n f)^\perp \setminus \{0\}\right\}
\end{aligned}$$

$$= \sup\left\{ \frac{|\psi_0|}{\|\psi\|_{L^q}} : \psi \in (\mathscr{P}_n f)^\perp \setminus \{0\} \right\}$$
$$= \left[\inf\{\|\psi/\psi_0\|_{L^q} : \psi \in (\mathscr{P}_n f)^\perp \setminus \{0\}\}\right]^{-1}$$
$$= \left[\inf\{\|\psi\|_{L^q} : \psi \in (\mathscr{P}_n f)^\perp,\ \psi_0 = 1\}\right]^{-1}$$
$$= \left[\inf\{\|\psi\|_{L^q} : \psi \in L^q,\ \psi_0 = 1,\ \langle z^k f, \psi\rangle = 0 \text{ for all } 0 \leqslant k \leqslant n\}\right]^{-1}. \qquad \square$$

Remark 4.0.2. The reason to move to L^p in the third equality is that the dual of L^p is L^q. If we stick with the norm in H^p, then (caution!) the relevant dual space is the quotient space L^q/H^q, rather than H^q. These spaces are isometrically isomorphic only when $p = 2$.

In the fourth equality, we mean "the norm of the unit constant function, viewed as a bounded linear functional on the annihilator of the subspace spanned by $f\mathscr{P}_n$".

We first apply this duality to provide a lower bound for the OPA error in the case that we are approximating a polynomial with zeros in the disk.

Proposition 4.0.3. *Suppose f is a polynomial*

$$f(z) = (z - w_1)(z - w_2)\cdots(z - w_d), \tag{4.0.4}$$

with the roots being distinct, nonzero, and contained inside $\mathbb{D}$. Then, for $1 < p < \infty$ and $\lambda := q_{0,p}[f]$, we have

$$\|1 - \lambda f\|_p \geqslant 1 - |w_1 w_2 \cdots w_d|. \tag{4.0.5}$$

Proof. The space of functions in L^q which annihilate f contains functions of the form

$$c_1\Lambda_{w_1} + c_2\Lambda_{w_2} + \cdots + c_d\Lambda_{w_d},$$

where Λ_{w_j} denotes the point evaluation functional (or Szegö kernel) at the point $w_j \in \mathbb{D}$. Thus, by the Lemma 4.0.1, we have

$$\|1 - \lambda f\|_p = [\inf \|\psi\|_q]^{-1},$$

where the infimum is over $\psi \in L^q$ satisfying $\langle f, \psi\rangle = 0$ and $\psi_0 = 1$.

Let

$$B(z) = a \prod_{k=1}^{d} \frac{w_k - z}{1 - \overline{w}_k z},$$

a constant multiple of the Blaschke product with the same zeros as f. Its numerator has leading term $\pm a z^d$, while the denominator has leading term $\pm\overline{w_1 w_2 \cdots w_d} z^d$ (with matching signs). Thus long division followed by partial fractions expansion results in an expression of the form

$$B(z) = \frac{a}{\overline{w_1 w_2 \cdots w_d}} + c_1 \Lambda_{w_1}(z) + c_2 \Lambda_{w_2}(z) + \cdots + c_d \Lambda_{w_d}(z).$$

Evaluating this equation at $z = 0$ tells us that

$$a w_1 w_2 \cdots w_d = \frac{a}{\overline{w_1 w_2 \cdots w_d}} + c_1 + c_2 + \cdots + c_d.$$

This suggests making the specific choice of

$$\psi(z) = c_1 \Lambda_{w_1}(z) + c_2 \Lambda_{w_2}(z) + \cdots + c_d \Lambda_{w_d}(c)$$

with the coefficients determined above. The requirement of $\psi(0) = 1$ therefore gives

$$a w_1 w_2 \cdots w_d = \frac{a}{\overline{w_1 w_2 \cdots w_d}} + 1,$$

which will furnish the value of a, namely,

$$a = \left[w_1 w_2 \cdots w_d - \frac{1}{\overline{w_1 w_2 \cdots w_d}} \right]^{-1}.$$

Finally, an application of the triangle inequality yields

$$\begin{aligned} \|1 - \lambda f\|_p \geqslant \|\psi\|_q^{-1} \geqslant \left\| B - \frac{a}{\overline{w_1 w_2 \cdots w_d}} \right\|_q^{-1} \\ \geqslant \left[\|B\|_q + \left\| \frac{a}{\overline{w_1 w_2 \cdots w_d}} \right\|_q \right]^{-1} = \left[|a| + \left| \frac{a}{w_1 w_2 \cdots w_d} \right| \right]^{-1} \\ = \frac{1 - |f(0)|^2}{1 + |f(0)|} = 1 - |f(0)| = 1 - |w_1 w_2 \cdots w_d|. \end{aligned}$$ □

Remark 4.0.6. The above result holds for any function f vanishing at the points $w_1, \ldots, w_d$. Further, by Theorem 3.1.4, the result also extends to any infinite Blaschke product.

Let us now use duality to further investigate OPA errors for more general functions.

Proposition 4.0.7. *Let* $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, $n \in \mathbb{N}$, *and* $f \in H^p$ *with* $f(0) = 1$. *Then*

$$\|q_{n-1,p}[f] f - 1\|_p \geqslant \frac{1}{\|1 + \psi_1 z + \psi_2 z^2 + \cdots + \psi_n z^n\|_q},$$

where the coefficients ψ_k, $1 \leqslant k \leqslant n$, *satisfy the matrix equation*

$$\begin{bmatrix} f_1 & f_2 & f_3 & \cdots & f_n \\ f_0 & f_1 & f_2 & \cdots & f_{n-1} \\ 0 & f_0 & f_1 & \cdots & f_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_1 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_n \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Proof. It suffices to check that the function

$$\psi(z) = 1 + \psi_1 z + \psi_2 z^2 + \cdots + \psi_n z^n$$

satisfies the hypotheses of Lemma 4.0.1. That is, for $0 \leq k \leq n$, that

$$\langle z^k f, \psi \rangle = 0.$$

This is ensured precisely by the linear system in the statement of the proposition. □

With further calculation, the approach in the previous proposition can be used to show the following:

Proposition 4.0.8. *Let* $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, $n \in \mathbb{N}$, *and* $f \in H^p$ *with* $f(0) = 1$. *Let*

$$\frac{1}{f(z)} = 1 + g_1 z + g_2 z^2 + \cdots$$

be the power series of $\frac{1}{f}$ *about the origin. Then*

$$\|q_{n-1,p}[f]f - 1\|_p \geqslant \frac{|g_n|}{\|1 + g_1 z + g_2 z^2 + \cdots + g_n z^n\|_q}.$$

Proof. Let us begin with, from Proposition 4.0.7, the matrix equation

$$\begin{bmatrix} f_1 & f_2 & f_3 & \cdots & f_n \\ f_0 & f_1 & f_2 & \cdots & f_{n-1} \\ 0 & f_0 & f_1 & \cdots & f_{n-2} \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_1 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_n \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

There is no harm in multiplying both sides of the equation on the left by the elementary permutation matrix

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

(the next to last entry of the bottom row is 1), which has the effect of changing the equation to

$$\begin{bmatrix} f_0 & f_1 & f_2 & \cdots & f_{n-1} \\ 0 & f_0 & f_1 & \cdots & f_{n-2} \\ 0 & 0 & f_0 & \cdots & f_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_1 \\ f_1 & f_2 & f_3 & \cdots & f_n \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \vdots \\ \psi_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ -1 \end{bmatrix}.$$

By successively subtracting multiples of the other rows, the bottom row can be placed in the form

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & C \end{bmatrix}$$

for some constant C, which could be zero. In fact, recalling that $f_0 = 1$, we see that C must be given by

$$C = \det \begin{bmatrix} f_1 & f_2 & f_3 & \cdots & f_n \\ f_0 & f_1 & f_2 & \cdots & f_{n-1} \\ 0 & f_0 & f_1 & \cdots & f_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_1 \end{bmatrix}.$$

Furthermore, the sequence of row operations to diagonalize the matrix leaves the right side unchanged as

$$\begin{bmatrix} 0 & 0 & 0 & \cdots & -1 \end{bmatrix}^T.$$

Assuming that $C \neq 0$, and again recalling that $f_0 = 1$, our matrix equation can be written as

$$\begin{bmatrix} f_0 & f_1 & f_2 & \cdots & f_{n-1} \\ 0 & f_0 & f_1 & \cdots & f_{n-2} \\ 0 & 0 & f_0 & \cdots & f_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & f_0 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \vdots \\ \psi_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ -1/C \end{bmatrix}.$$

The inverse of the transposed (Toeplitz) matrix on the left is simply

$$\begin{bmatrix} g_0 & g_1 & g_2 & \cdots & g_{n-1} \\ 0 & g_0 & g_1 & \cdots & g_{n-2} \\ 0 & 0 & g_0 & \cdots & g_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & g_0 \end{bmatrix},$$

where $g(0)=1$ and $g(z)=g_0+g_1z+g_2z^2+\cdots$ is the Taylor expansion of $1/f(z)$, valid for some disk centered at the origin. The conclusion is that

$$\psi_k=-\frac{g_{n-k}}{C}\quad\text{for all } 0\leqslant k\leqslant n-1.$$

Our next challenge is to find an analytical meaning for the constant C. But notice that the row operations needed to clear entries from the bottom row of

$$\begin{bmatrix} f_0 & f_1 & f_2 & \cdots & f_{n-1}\\ 0 & f_0 & f_1 & \cdots & f_{n-2}\\ 0 & 0 & f_0 & \cdots & f_{n-3}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & f_1\\ f_1 & f_2 & f_3 & \cdots & f_n \end{bmatrix} \tag{4.0.9}$$

would (suitably modified) similarly clear the second through the last entries from the top row. Performing all of these (suitably modified) row operations on the identity matrix would have to result in

$$\begin{bmatrix} 1 & g_1 & g_2 & \cdots & g_{n-1}\\ 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Then following carefully what operations are correspondingly performed on the last column in (4.0.9), we conclude that

$$C=f_n+g_1f_{n-1}+g_2f_{n-2}+\cdots+g_{n-1}f_1=-g_n.$$

Finally, note that

$$\begin{aligned}\left\|1+\frac{g_{n-1}}{g_n}z+\frac{g_{n-2}}{g_n}z^2+\cdots+\frac{1}{g_n}z^n\right\|_q &= \frac{1}{|g_n|}\|g_n+g_{n-1}z+g_{n-2}z^2+\cdots+z^n\|_q\\ &= \frac{1}{|g_n|}\|z^{-n}(g_n+g_{n-1}z+g_{n-2}z^2+\cdots+z^n)\|_q\\ &= \frac{1}{|g_n|}\|1+g_1z^{-1}+g_2z^{-2}+\cdots+g_nz^{-n}\|_q\\ &= \frac{1}{|g_n|}\|1+g_1z+g_2z^2+\cdots+g_nz^n\|_q,\end{aligned}$$

where in the last step, the change of variable $\theta\mapsto-\theta$, for $z=e^{i\theta}$, leaves the norm integral unchanged. □

We will provide an improvement to the above proposition, but we must first consider the problem of finding $G \in zH^p$ such that

$$\|\overline{G} + F\|_p$$

is minimized. Duality tells us that (as we continue to mark extremal functions with *)

$$\begin{aligned}\|\overline{G^*} + F\|_p &= \sup\left\{\frac{|\langle F, \psi\rangle|}{\|\psi\|_q} : \psi \in H^q\right\}\\ &= \frac{|\langle F, \psi^*\rangle|}{\|\psi^*\|_q}\\ &= \frac{|\langle F, \psi^*\rangle|}{\inf\{\|\psi^* + K\|_q : K \in \overline{zH^q}\}}.\end{aligned}$$

Once again, we are up against the dual of H^p being isomorphic to H^q, but not isometrically.

Nonetheless, we must consider the metric projection of F onto the subspace $\overline{zH^p}$. Notice that $(\overline{G^*} + F)^{\langle p-1\rangle}$ annihilates any negative frequencies. Therefore, there exists $h \in H^q$ such that

$$(\overline{G^*} + F)^{\langle p-1\rangle} = h,$$

and thus, taking a $\langle q-1\rangle$ power, we have

$$|h|^{q-2}h = \overline{G^*} + F.$$

In turn, we see that finding G^* amounts to solving the above highly unpleasant functional equation.

Let us record this in the following result, where we write P_+ for the Riesz projection, given by

$$\sum_{k=-\infty}^{\infty} c_k z^k \mapsto \sum_{k=0}^{\infty} c_k z^k,$$

which is bounded from $L^p \to H^p$.

Proposition 4.0.10. *Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Suppose $h \in H^q$, and define $F := P_+\overline{h^{\langle q-1\rangle}}$. Then*

$$\inf\{\|F + \overline{G}\|_p : G \in zH^p\}$$

is attained by taking $F + \overline{G} = \overline{h^{\langle q-1\rangle}}$. In this case, the value of the infimum is given by

$$\inf\{\|F + \overline{G}\|_p : G \in zH^p\} = \|h\|_q^{q-1}.$$

This warrants the following observation:

Proposition 4.0.11. *For $1 < p < \infty$, the set of images $P_+\,\overline{(H^q)^{\langle q-1\rangle}}$ is dense in H^p.*

Proof. Suppose $g \in H^q$ has the property that

$$\langle P_+ \overline{h^{\langle q-1 \rangle}}, g \rangle = 0$$

for all $h \in H^q$. Then

$$0 = \langle P_+ \overline{h^{\langle q-1 \rangle}}, g \rangle = \int_{\mathbb{T}} P_+ \overline{h^{\langle q-1 \rangle}}\, \bar{g}\, dm = \int_{\mathbb{T}} \overline{h^{\langle q-1 \rangle}}\, \bar{g}\, dm.$$

We are able to drop the projection in the last line since integration against $\bar{g}$ will annihilate the negative frequencies of $\overline{h^{\langle q-1 \rangle}}$. In particular, this must hold for $h = g$, and hence $0 = \|g\|_q^q$. This forces g to be identically zero. □

In turn, we can make the following improvement to Proposition 4.0.8.

Proposition 4.0.12. *Let* $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, $n \in \mathbb{N}$, *and* $f \in H^p$ *with* $f(0) = 1$. *Let*

$$\frac{1}{f(z)} = 1 + g_1 z + g_2 z^2 + \cdots$$

be the power series of $\frac{1}{f}$ *about the origin. If* $g_n \neq 0$, *then*

$$\|q_{n-1,p}[f]f - 1\|_p \geqslant \frac{1}{\|h^{\langle p-1 \rangle}\|_q},$$

where $h \in H^p$ *satisfies*

$$P_+ \overline{h^{\langle p-1 \rangle}} = 1 + \frac{g_{n-1}}{g_n} z + \frac{g_{n-2}}{g_n} z^2 + \cdots + \frac{g_0}{g_n} z^n.$$

Let us now consider the case where $n \to \infty$. Then, by writing $f = JG$ for J inner and G outer, we see, for all $k \geq 0$, that $\psi \in L^q$ satisfies

$$0 = \langle z^k f, \psi \rangle = \langle z^k JG, \psi \rangle = \langle z^k G, \bar{J}\psi \rangle.$$

As $\{z^k G : k \geq 0\}$ is dense in H^p, we have, for any $n \geq 0$,

$$0 = \langle z^n, \bar{J}\psi \rangle = \int_0^{2\pi} J(e^{i\theta})\overline{\psi(e^{i\theta})} e^{in\theta}\, \frac{d\theta}{2\pi}.$$

From this, it follows that $K(z) := J(z)\overline{\psi(z)}/z$ is an element of H^q. We further divine that ψ must be determined by

$$\psi(z) = J(z)\overline{zK(z)}, \quad z \in \mathbb{T},$$

for some $K \in H^q$. The condition $\psi(0) = 1$ takes the form

$$1 = J_1\overline{K_0} + J_2\overline{K_1} + J_3\overline{K_2} + \cdots .$$

We now must minimize $\|J\overline{zK}\|_q$ subject to $K \in H^q$ satisfying the above constraint. It is tempting to try $K = \overline{J}$, but this will not work, since

$$0 = J_1\overline{J_0} + J_2\overline{J_1} + J_3\overline{J_2} + \cdots .$$

Instead, take $K(z) = c[J(z) - J(0)]/z$, where $c^{-1} = |J_1|^2 + |J_2|^2 + |J_3|^2 + \cdots$. Then

$$J_1\overline{K_0} + J_2\overline{K_1} + J_3\overline{K_2} + \cdots = c(|J_1|^2 + |J_2|^2 + |J_3|^2 + \cdots) = 1,$$

as needed. Using this choice of K to compute ψ, we obtain

$$\psi(z) = J(z)\overline{zK(z)} = cJ(z)\bar{z}\ \frac{\overline{J(z)} - \overline{J(0)}}{\bar{z}} = c(1 - J(z)\overline{J(0)}).$$

Since $\|q_{n,p}[f]f - 1\|_p \geqslant 1/\|\psi\|_q$, this furnishes the following bound.

Proposition 4.0.13. *Let* $1 < p < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$, $n \in \mathbb{N}$, *and* $f \in H^p$ *with* $f(0) = 1$. *Then*

$$\|q_{n,p}[f]f - 1\|_p \geqslant \frac{|J_1|^2 + |J_2|^2 + |J_3|^2 + \cdots}{\|1 - \overline{J(0)}J(z)\|_q},$$

where J *is the inner part of* f.

Incidentally, $c^{-1} = (|J_1|^2 + |J_2|^2 + |J_3|^2 + \cdots) = \|J\|_2^2 - |J(0)|^2 = 1 - |J(0)|^2$, so the lower bound above could be written equivalently as

$$\frac{1 - |J(0)|^2}{\|(1 - |J(0)|^2) - \overline{J(0)}(J_1 z + J_2 z^2 + \cdots)\|_q},$$

which is obviously no greater than 1, as needed.

We now step away from duality. Our final results concern OPA errors, but are proven with H^2 methods. The following proposition should be compared with Proposition 4.0.3; although the result below provides a better bound, it holds only for $p > 2$.

Proposition 4.0.14. *Let* $2 < p < \infty$, *and suppose* $f \in H^p$ *has a factorization* $f = JG$, *where* J *is inner and* G *is outer. If* $f(0) \neq 0$, *then, for any* $n \in \mathbb{N}$,

$$\|q_{n,p}[f]f - 1\|_p \geqslant \sqrt{1 - |J(0)|^2}.$$

Proof. Let $\mathscr{P}$ be the collection of all polynomials. Then

$$\begin{aligned}\|q_{n,p}[f]f - 1\|_p &\geqslant \inf_{Q \in \mathscr{P}} \|Qf - 1\|_p \geqslant \inf_{Q \in \mathscr{P}} \|Qf - 1\|_2 \\ &= \inf_{Q \in \mathscr{P}} \|QJG - 1\|_2 = \inf_{Q \in \mathscr{P}} \|QG - \overline{J}\|_2 = \|\overline{J(0)} - \overline{J}\|_2,\end{aligned}$$

with the last equality following from the fact that since G is outer, $\{QG : Q \in \mathscr{P}\}$ is dense in H^2. Now use

$$1 = \|J\|_2^2 = |J(0)|^2 + \|J - J(0)\|_2^2. \qquad \square$$

Note further that if $J = B$ is a Blaschke product, then this implies that

$$\|q_{n,p}[f]f - 1\|_p \geqslant \|J(0) - J\|_2 = \sqrt{1 - |B(0)|^2} = \sqrt{1 - |w_1 w_2 w_3 \cdots|^2},$$

where $w_1, w_2, w_3, \ldots$ are the zeros of B.

We end by providing a related result when $1 < p < 2$.

Proposition 4.0.15. *Let $1 < p < 2$ and suppose $f(0) \neq 0$. Then, for any $n \in \mathbb{N}$,*

$$\|q_{n,p}[f]f - 1\|_p \leqslant \sqrt{1 - (q_{n,2}[f]f)(0)}.$$

Proof. Routine bounds yield

$$\|q_{n,p}[f]f - 1\|_p \leqslant \|q_{n,2}[f]f - 1\|_p \leqslant \|q_{n,2}[f]f - 1\|_2 = \sqrt{1 - (q_{n,2}[f]f)(0)},$$

where the last equality is a consequence of the linear system described in (1.0.1). $\square$

Noting that $\|q_{n,2}[f]f - 1\|_2 \leqslant \|q_{0,2}[f]f - 1\|_2$, we can also establish the simple bound

$$\|q_{n,p}[f]f - 1\|_p \leqslant \left(1 - \frac{|f(0)|^2}{\|f\|_2^2}\right)^{1/2}.$$

References

[1] A. Acuaviva and D. Seco, "Equidistribution of zeros of some polynomials related to cyclic functions", *Anal. Math. Phys.* **12**:2 (2022), art. id. 53. MR Zbl

[2] J. Alonso, H. Martini, and S. Wu, "On Birkhoff orthogonality and isosceles orthogonality in normed linear spaces", *Aequationes Math.* **83**:1-2 (2012), 153–189. MR Zbl

[3] P. Arora, M. Augat, M. T. Jury, and M. Sargent, "An optimal approximation problem for free polynomials", *Proc. Amer. Math. Soc.* **152**:2 (2024), 455–470. MR Zbl

[4] C. Bénéteau and R. Centner, "A survey of optimal polynomial approximants, applications to digital filter design, and related open problems", *Complex Anal. Synerg.* **7**:2 (2021), art. id. 16. MR Zbl

[5] C. Bénéteau, A. A. Condori, C. Liaw, D. Seco, and A. A. Sola, "Cyclicity in Dirichlet-type spaces and extremal polynomials", *J. Anal. Math.* **126** (2015), 259–286. MR Zbl

[6] C. Bénéteau, A. A. Condori, C. Liaw, D. Seco, and A. A. Sola, "Cyclicity in Dirichlet-type spaces and extremal polynomials, II: Functions on the bidisk", *Pacific J. Math.* **276**:1 (2015), 35–58. MR Zbl

[7] C. Bénéteau, G. Knese, L. u. Kosiński, C. Liaw, D. Seco, and A. Sola, "Cyclic polynomials in two variables", *Trans. Amer. Math. Soc.* **368**:12 (2016), 8737–8754. MR Zbl

[8] C. Bénéteau, M. C. Fleeman, D. S. Khavinson, D. Seco, and A. A. Sola, "Remarks on inner functions and optimal approximants", *Canad. Math. Bull.* **61**:4 (2018), 704–716. MR Zbl

[9] C. Bénéteau, D. Khavinson, C. Liaw, D. Seco, and B. Simanek, "Zeros of optimal polynomial approximants: Jacobi matrices and Jentzsch-type theorems", *Rev. Mat. Iberoam.* **35**:2 (2019), 607–642. MR Zbl

[10] R. Centner, "Optimal polynomial approximants in L^p", *Concr. Oper.* **9**:1 (2022), 96–113. MR Zbl

[11] R. Centner, R. Cheng, and C. Felder, "Optimal polynomial approximants in H^p", preprint, 2023. arXiv 2305.16068

[12] R. Cheng and W. T. Ross, "Weak parallelogram laws on Banach spaces and applications to prediction", *Period. Math. Hungar.* **71**:1 (2015), 45–58. MR Zbl

[13] R. Cheng, J. Mashreghi, and W. T. Ross, *Function theory and ℓ^p spaces*, University Lecture Series **75**, American Mathematical Society, Providence, RI, 2020. MR Zbl

[14] R. Cheng, W. T. Ross, and D. Seco, "Zeros of optimal polynomial approximants in ℓ^p_A", *Adv. Math.* **404**:part A (2022), art. id. 108396. MR Zbl

[15] E. Fricain, J. Mashreghi, and D. Seco, "Cyclicity in reproducing kernel Hilbert spaces of analytic functions", *Comput. Methods Funct. Theory* **14**:4 (2014), 665–680. MR Zbl

[16] R. C. James, "Orthogonality and linear functionals in normed linear spaces", *Trans. Amer. Math. Soc.* **61** (1947), 265–292. MR Zbl

[17] M. Sargent and A. A. Sola, "Optimal approximants and orthogonal polynomials in several variables, II: Families of polynomials in the unit ball", *Proc. Amer. Math. Soc.* **149**:12 (2021), 5321–5330. MR Zbl

[18] M. Sargent and A. A. Sola, "Optimal approximants and orthogonal polynomials in several variables", *Canad. J. Math.* **74**:2 (2022), 428–456. MR Zbl

Received November 25, 2023. Revised February 4, 2024.

RAYMOND CHENG
DEPARTMENT OF MATHEMATICS AND STATISTICS
OLD DOMINION UNIVERSITY
NORFOLK, VA
UNITED STATES
rcheng@odu.edu

CHRISTOPHER FELDER
DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY BLOOMINGTON
BLOOMINGTON, IN
UNITED STATES
cfelder@iu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 327, No. 2, 2023
https://doi.org/10.2140/pjm.2023.327.297

THE GRIFFITHS DOUBLE CONE GROUP IS ISOMORPHIC TO THE TRIPLE

SAMUEL M. CORSON

It is shown that the fundamental group of the Griffiths double cone space is isomorphic to that of the triple cone. More generally if κ is a cardinal such that $2 \le \kappa \le 2^{\aleph_0}$ then the κ-fold cone has the same fundamental group as the double cone. The isomorphisms produced are nonconstructive, and no isomorphism between the fundamental group of the 2- and of the κ-fold cones, with $2 < \kappa$, can be realized via continuous mappings.

1. Introduction

The Griffiths double cone over the Hawaiian earring, which we denote $\mathbb{GS}_2$, was introduced by H. B. Griffiths [1954] and has long stood as an interesting example in topology (Figure 1). Although $\mathbb{GS}_2$ is a path connected, locally path connected compact metric space (a *Peano continuum*) which embeds as a subspace of $\mathbb{R}^3$, it has some subtle properties. Despite being a wedge of two contractible spaces, $\mathbb{GS}_2$ is not itself contractible, and more surprisingly the fundamental group of $\mathbb{GS}_2$ is uncountable. The fundamental group is freely indecomposable and includes a copy of the additive group of the rationals and of the fundamental group of the Hawaiian earring. This group has found use in defining cotorsion-free groups in the nonabelian setting [Eda and Fischer 2016] and continues to serve as a counterexample [Zastrow 1994] and as a test model for notions of infinitary abelianization [Brazas and Gillespie 2022].

It is easy to see that analogous behavior is exhibited when one uses more cones in the wedge, as in the triple wedge $\mathbb{GS}_3$ of cones over the Hawaiian earring or more generally in the κ-fold wedge $\mathbb{GS}_\kappa$ (the one-point union of cones, indexed by κ, with the natural metric topology). A natural question is whether the isomorphism type of the fundamental group changes with this change in subscript. In light of the

This work was supported by European Research Council grant PCG-336983, the Severo Ochoa Programme for Centres of Excellence in R&D SEV-20150554, and by the Heilbronn Institute for Mathematical Research, Bristol, UK.

MSC2020: primary 03E75, 20A15, 55Q52; secondary 20F10, 20F34.

Keywords: fundamental group, Griffiths space, Hawaiian earring, topological cone.

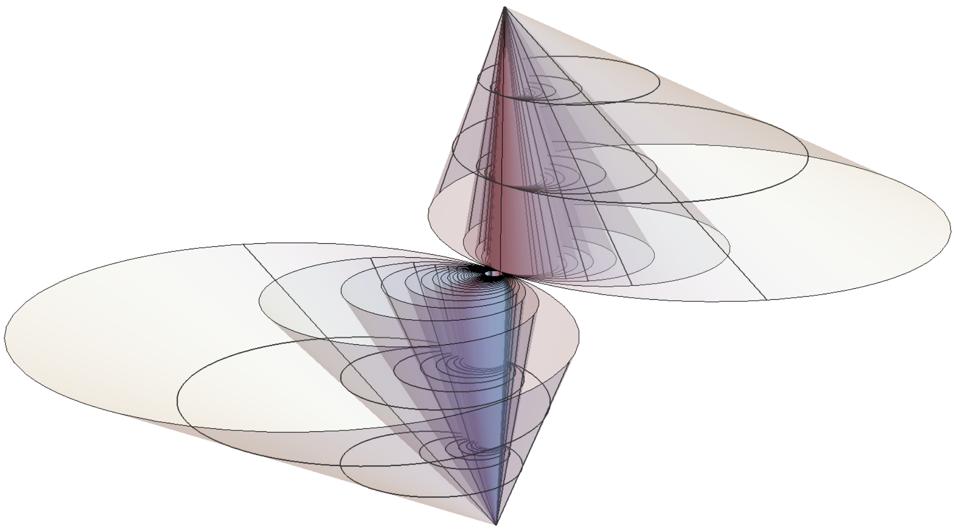

Figure 1. The Griffiths double cone $\mathbb{GS}_2$.

intuitive fact that no spatial isomorphism can be defined the following answer is surprising.

Theorem A. *If κ is a cardinal such that $2 \leq \kappa \leq 2^{\aleph_0}$ then $\pi_1(\mathbb{GS}_2) \simeq \pi_1(\mathbb{GS}_\kappa)$.*

The bounds on κ in the statement of Theorem A are the best possible. The spaces $\mathbb{GS}_0$ and $\mathbb{GS}_1$ both strongly deformation retract to a point and therefore have trivial fundamental group, and when $\kappa > 2^{\aleph_0}$ one has $|\pi_1(\mathbb{GS}_\kappa)| > 2^{\aleph_0} = |\pi_1(\mathbb{GS}_2)|$ (Theorem 2.11). Using techniques of [Eda and Fischer 2016] or [Herfort and Hojka 2017] one can compute the abelianizations of $\pi_1(\mathbb{GS}_2)$ and $\pi_1(\mathbb{GS}_3)$ and see that these abelianizations are isomorphic.

A notable point of comparison is that the wedge of 2, 3, etc. Hawaiian earrings (without cones) is again homeomorphic to the Hawaiian earring, and so these spaces have isomorphic fundamental groups. However the fundamental group of a wedge of $\aleph_0$ Hawaiian earrings, under the topology that we are considering, will not have isomorphic fundamental group. This follows since the $\aleph_0$-wedge of Hawaiian earrings retracts to a subspace which is the $\aleph_0$-wedge of circles each having diameter 1, and this shows that the fundamental group of the $\aleph_0$-wedge homomorphically surjects onto an infinite rank free group, which the fundamental group of the Hawaiian earring cannot do [Higman 1952].

The isomorphism given in Theorem A is produced combinatorially by a back-and-forth argument, using the axiom of choice. It is intuitively clear that there is no continuous function from $\mathbb{GS}_2$ to $\mathbb{GS}_3$ or vice versa which can yield an isomorphism of fundamental groups. A comparable situation in the setting of topological groups is that $\mathbb{R}$ and $\mathbb{R}^2$ are isomorphic as abstract groups, since by picking a Hamel basis over $\mathbb{Q}$ one sees that both are isomorphic to $\bigoplus_{2^{\aleph_0}} \mathbb{Q}$. There is no continuous, or even Baire measurable, isomorphism between these topological groups. By contrast Theorem A does not seem to follow by producing isomorphisms to an easily understood third group like $\bigoplus_{2^{\aleph_0}} \mathbb{Q}$.

Another curiosity worth mentioning is that despite the necessary constraints on the cardinality of κ in Theorem A, the first-order logical theory of $\pi_1(\mathbb{GS}_2)$ and $\pi_1(\mathbb{GS}_\kappa)$ are the same whenever $\kappa \geq 2$.

Theorem B. *If $2 \leq \gamma \leq \kappa$ then $\pi_1(\mathbb{GS}_\gamma)$ elementarily embeds in $\pi_1(\mathbb{GS}_\kappa)$. Thus for $\kappa \geq 2$ the groups $\pi_1(\mathbb{GS}_2)$ and $\pi_1(\mathbb{GS}_\kappa)$ are elementarily equivalent.*

Of course when κ is 0 or 1 the fundamental group $\pi_1(\mathbb{GS}_\kappa)$ is trivial and therefore not elementarily equivalent to $\pi_1(\mathbb{GS}_2)$. The proof of Theorem B utilizes Theorem A and the action of the automorphism group, and no previous knowledge of first-order logic is required to understand the proof.

The ideas used in proving Theorem A seem to have very broad applications, and we state two now. Another space that is often mentioned along with the Griffiths space is the harmonic archipelago $\mathbb{HA}$ of Bogley and Sieradski [2000]. The spaces $\mathbb{GS}_2$ and $\mathbb{HA}$ share many common properties. Each embeds as a subspace of $\mathbb{R}^3$, both contain a distinguished point at which every loop can be homotoped to have arbitrarily small image, and both have uncountable fundamental group. Cannon and Conner have conjectured that the two spaces share a further property, namely that they have isomorphic fundamental group [Conner 2011], and in a forthcoming paper we will show that this is indeed the case (the reader can see a proof of this fact in [Corson 2021, Theorem D]). By further reworking these ideas one can produce a correct proof of the main theorem of [Conner et al. 2015] (some errors have been pointed out by K. Eda) as well as answer many of the questions of that paper in the affirmative (see [Corson 2023]).

We describe the layout of this paper. In Section 2 we give the formal definition of the Griffiths space and its κ-fold analogues. We also present some combinatorially defined groups $\mathcal{C}_\kappa$ and show them to be isomorphic to the fundamental groups $\pi_1(\mathbb{GS}_\kappa)$. In Section 3 we prove Theorems A and B.

2. The cone groups

We give a construction of $\mathbb{GS}_2$ and more generally of the κ-fold Griffiths space $\mathbb{GS}_\kappa$ for any cardinal κ. We consider each cardinal number κ as being the set of all ordinals below it in the standard way. Thus $0 = \varnothing$, $n = \{0, \dots, n-1\}$ for each $n \in \omega$, $\omega + 2 = \{0, 1, \dots, \omega, \omega + 1\}$, etc. Let $2^{\aleph_0}$ denote the cardinal of the continuum. Given a point $p \in \mathbb{R}^2$ and $r \in [0, \infty)$ we let $C(p, r)$ denote the circle centered at p of radius r (in case $r = 0$ we obtain the degenerate circle consisting only of the point p). The *Hawaiian earring* is the subspace $\mathbb{E} = \bigcup_{n \in \omega} C\big(\big(0, \frac{1}{n+3}\big), \frac{1}{n+3}\big)$

of $\mathbb{R}^2$. Let $\mathbb{GS}_1 \subseteq \mathbb{R}^3$ be the subspace $\bigcup_{r\in[0,1]}\big(\bigcup_{n\in\omega} C\big(\big(0, \frac{1-r}{n+3}\big), \frac{r}{n+3}\big) \times \{r\}\big)$. The space $\mathbb{GS}_1$ may also be viewed as the space obtained by first taking the Hawaiian earring sitting in the xy-plane $\mathbb{E} \times \{0\}$ and joining each point of $\mathbb{E} \times \{0\}$ to the point $(0, 0, 1)$ by a geodesic line segment. A third, topological way of viewing $\mathbb{GS}_1$ is by simply taking the topological cone over the Hawaiian earring. In other words, $\mathbb{GS}_1$ is homeomorphic to the quotient space obtained by beginning with $\mathbb{E} \times [0, 1]$ and identifying all points which have 1 in the last coordinate.

We define $\mathbb{GS}_0$ to be the metric space consisting of the single point $\circ_0$. Let $\kappa \geq 1$ be a cardinal. We take $\mathbb{GS}_\kappa$ to be the set obtained by taking κ-many disjoint isometric copies $\bigsqcup_{\alpha<\kappa} X_\alpha$ of $\mathbb{GS}_1$ and identifying all copies of $(0, 0, 0)$ to a single point $\circ_\kappa$. Thus we consider $\circ_\kappa \in X_\alpha$ for all $\alpha < \kappa$. Metrize $\mathbb{GS}_\kappa$ by letting D_α be the metric on X_α (making X_α an isometric copy of $\mathbb{GS}_1$) and

$$D(x, y) = \begin{cases} D_\alpha(x, y) & \text{if } x, y \in X_\alpha, \\ D_\alpha(x, \circ_\kappa) + d_{\alpha'}(\circ_\kappa, y) & \text{if } x \in X_\alpha \setminus \{\circ_\kappa\} \text{ and } y \in X_{\alpha'} \setminus \{\circ_\kappa\}, \alpha \neq \alpha'. \end{cases}$$

We note that this definition yields an isometric copy of $\mathbb{GS}_1$ when $\kappa = 1$ and so the definition is consistent. When κ is finite, the space $\mathbb{GS}_\kappa$ is a Peano continuum and $\mathbb{GS}_\kappa$ is homeomorphic to the topological wedge of κ-many copies of $\mathbb{GS}_1$ with the copies of the point $(0, 0, 0)$ identified. When $\kappa \geq \aleph_0$ the space $\mathbb{GS}_\kappa$ is neither compact nor homeomorphic to the quotient space obtained by identifying all copies of $(0, 0, 0)$ in the topological disjoint union of κ-many copies of $\mathbb{GS}_1$.

Next we give a description of what we call the *cone group* $\mathcal{C}_\kappa$ for each cardinal κ. The description involves infinitary word combinatorics. Fix a cardinal κ. We start with a set $\mathcal{A}_\kappa = \{a_{\alpha,n}^{\pm 1}\}_{\alpha<\kappa, n<\omega}$ equipped with formal inverses. We call the elements of $\mathcal{A}_\kappa$ *letters* and a letter is *positive* if it has superscript 1. For convenience we shall usually leave off the superscript 1 on positive letters. A letter which is not positive is *negative*. Let proj_0 and proj_1 be the functions defined on $\mathcal{A}_\kappa$ which project respectively the first and second subscript of a letter. Thus $\mathrm{proj}_0(a_{\alpha,n}^{-1}) = \alpha$ and $\mathrm{proj}_1(a_{\alpha,n}^{-1}) = n$.

A *word* in $\mathcal{A}_\kappa$ is a function $W : \overline{W} \to \mathcal{A}_\kappa$ such that $\overline{W}$ is a totally ordered set and for each $N \in \omega$ the set $\{i \in \overline{W} \mid \mathrm{proj}_1(W(i)) \leq N\}$ is finite. The domain of a word is necessarily countable. We write $W_0 \equiv W_1$ if there exists an order isomorphism $\iota : \overline{W_0} \to \overline{W_1}$ such that $W_1(\iota(i)) = W_0(i)$ for all $i \in \overline{W_0}$, and write $\iota : W_0 \equiv W_1$ in this case. Let E denote the word with empty domain.

Let $\mathcal{W}_\kappa$ denote the set of all $\equiv$ classes of words in $\mathcal{A}_\kappa$. For $W \in \mathcal{W}_\kappa$ we let $d(W) = \min\{\mathrm{proj}_1(W(i)) \mid i \in \overline{W}\}$ and $d(E) = \infty$. There is a natural associative binary operation on $\mathcal{W}_\kappa$ given by word concatenation, defined by letting W_0W_1 be the word W such that $\overline{W} = \overline{W_0} \sqcup \overline{W_1}$ has the ordering that extends the orders of $\overline{W_0}$ and $\overline{W_1}$, placing elements in $\overline{W_0}$ below those of $\overline{W_1}$, and

$$W(i) = \begin{cases} W_0(i) & \text{if } i \in \overline{W_0}, \\ W_1(i) & \text{if } i \in \overline{W_1}. \end{cases}$$

There is similarly a notion of infinite concatenation. If Λ is a totally ordered set and $\{W_\lambda\}_{\lambda\in\Lambda}$ is a collection of words such that for every $N \in \omega$ the set $\{\lambda \in \Lambda : d(W_\lambda) \le N\}$ is finite then we can take a concatenation $\prod_{\lambda\in\Lambda} W_\lambda$ whose domain is the disjoint union $\bigsqcup_{\lambda\in\Lambda} \overline{W_\lambda}$ ordered in the natural way and whose outputs are given by $\bigl(\prod_{\lambda\in\Lambda} W_\lambda\bigr)(i) = W_\lambda(i)$ where $i \in \overline{W_\lambda}$. We also use this notation for the concatenation of ordered sets. If $\{\Lambda_\lambda\}_{\lambda\in\Lambda}$ is a collection of ordered sets and Λ is itself ordered we let $\prod_{\lambda\in\Lambda} \Lambda_\lambda$ be the ordered set obtained by taking the disjoint union of the Λ_λ and ordering the elements in the obvious way. To further abuse notation we write $\Lambda \equiv \Theta$ if Λ is order isomorphic to Θ.

We also have an inversion operation on words given by letting W^{-1} have domain $\overline{W}$ under the reverse order and letting $W^{-1}(i) = (W(i))^{-1}$. For each $N \in \omega$ and word W we let $p_N(W)$ be the restriction $W \upharpoonright \{i \in \overline{W} \mid \mathrm{proj}_1(W(i)) \le N\}$. Thus $p_N(W)$ is a finite word in the alphabet $\mathcal{A}_\kappa$. We write $W_0 \sim W_1$ if for every $N \in \omega$ the words $p_N(W_0)$ and $p_N(W_1)$ are equal when considered as elements in the free group on positive elements of $\mathcal{A}_\kappa$. As an example, the word $W \equiv a_{0,0}a_{0,0}^{-1}a_{0,1}a_{0,1}^{-1}\cdots$ satisfies $W \sim E$ since $p_N(W) \equiv a_{0,0}a_{0,0}^{-1}a_{0,1}a_{0,1}^{-1}\cdots a_{0,N}a_{0,N}^{-1}$ is freely equal to E for each $N \in \omega$. Let $[W]$ denote the $\sim$ equivalence class of W. We obtain a group structure on $\mathcal{W}_\kappa/\!\sim$ by letting $[W_0][W_1] = [W_0W_1]$, from which one gets inverses defined by $[W]^{-1} = [W^{-1}]$ and $[E]$ as the identity element. Let H_κ denote this group. Define a word W to be *α-pure* if $\mathrm{proj}_0 \circ W(i) = \alpha$ for all $i \in \overline{W}$. More generally a word is *pure* if it is α-pure for some α. The empty word E is α-pure for every α. Define the group $\mathcal{C}_\kappa$ to be the quotient of H_κ by the smallest normal subgroup including the set of $\sim$ equivalence classes of pure words.

We work towards the proof that $\mathcal{C}_\kappa \simeq \pi_1(\mathbb{GS}_\kappa, \circ_\kappa)$. Recall that the Hawaiian earring $\mathbb{E} \times \{0\}$ is a subspace of $\mathbb{GS}_1$. Each copy X_α of $\mathbb{GS}_1$ which appears in the wedge $\mathbb{GS}_\kappa$ therefore has such a copy of the Hawaiian earring, which we denote E_α, at its "base". Let $\mathbb{E}_\kappa$ denote the union of all of these copies E_α of the Hawaiian earring.

In [Cannon and Conner 2000] is a description of an isomorphism of H_1 with the fundamental group of the Hawaiian earring $\pi_1(\mathbb{E}_1, \circ_1)$, which we give and generalize here. Let $\mathcal{I}$ denote the set of maximal open intervals in the closed interval $[0, 1]$ minus the Cantor ternary set. The natural ordering on $\mathcal{I}$ is order isomorphic to that of the rationals, and so every countable order type embeds in $\mathcal{I}$. For each $n \in \omega$ let L_n be a loop based at $\circ_1$ which passes exactly once around the circle $C\bigl(\bigl(0, \frac{1}{n+3}\bigr), \frac{1}{n+3}\bigr)$ and is injective except at 0 and 1. Given a word $W \in \mathcal{W}_1$ we let $\iota : \overline{W} \to \mathcal{I}$ be an order embedding. Let $\mathrm{R}_\iota(W) : [0, 1] \to \mathbb{E}_1$ be the loop given by

$$\mathrm{R}_\iota(W)(t) = \begin{cases} L_n\bigl(\frac{t-\inf I}{\sup I-\inf I}\bigr) & \text{if } W(i) = a_{0,n} \text{ and } t \in I = \iota(i), \\ L_n^{-1}\bigl(\frac{t-\inf I}{\sup I-\inf I}\bigr) & \text{if } W(i) = a_{0,n}^{-1} \text{ and } t \in I = \iota(i), \\ \circ_1 & \text{otherwise.} \end{cases}$$

If $\iota_0 : \overline{W} \to \mathcal{I}$ is a distinct order embedding, then $\mathrm{R}_\iota(W)$ and $\mathrm{R}_{\iota_0}(W)$ are homotopic via a straightforward homotopy whose image lies inside the common image $\mathrm{R}_\iota(W)([0,1]) = \mathrm{R}_{\iota_0}(W)([0,1])$. Thus we have a well-defined map $\mathrm{R} : \mathcal{W} \to \pi_1(\mathbb{E}_1, \circ_1)$. Less obvious is the fact that $W \sim U$ implies $\mathrm{R}(W) = \mathrm{R}(U)$, so that R descends to a map, which we also name R, from H_1 to $\pi_1(\mathbb{E}_1, \circ_1)$ which is in fact an isomorphism. Each loop at $\circ_1$, moreover, can be homotoped in its image to a loop which is precisely $\mathrm{R}_\iota(W)$ for some ι and W.

We'll use these facts to produce such a map R for larger values of κ. To simplify the work we introduce the notion of reduced words. As is the case with finitary words, there is a notion of reducedness for words in $\mathcal{W}_\kappa$. We say $W \in \mathcal{W}_\kappa$ is *reduced* if $W \equiv W_0 W_1 W_2$ and $W_1 \sim E$ implies $W_1 \equiv E$. We state the following, whose proof would follow in precisely the same way as that of [Eda 1992, Theorem 1.4, Corollary 1.7].

Lemma 2.1. *Given $W \in \mathcal{W}_\kappa$ there exists a reduced word $W_0 \in \mathcal{W}_\kappa$ such that $[W] = [W_0]$ and this W_0 is unique up to $\equiv$. Moreover, letting W and U be reduced, there exist unique words W_0, W_1, U_0, U_1 such that*:

(1) $W \equiv W_0 W_1$.

(2) $U \equiv U_0 U_1$.

(3) $W_1 \equiv U_0^{-1}$.

(4) $W_0 U_1$ *is reduced.*

Let Red_κ denote the set of reduced words in $\mathcal{W}_\kappa$ and for each $W \in \mathcal{W}_\kappa$ let $\mathrm{Red}(W)$ be the reduced word such that $W \sim \mathrm{Red}(W)$. The proof of the following is straightforward.

Lemma 2.2. *We have* $\mathrm{Red}(WU) \equiv \mathrm{Red}(\mathrm{Red}(W)\,\mathrm{Red}(U))$ *given $W \in \mathcal{W}_\kappa$ and $U \in \mathcal{W}_\kappa$. Similarly, given W_0, W_1, $W_2 \in \mathcal{W}_\kappa$ we have*

$$\mathrm{Red}(W_0 W_1 W_2) \equiv \mathrm{Red}(W_0\,\mathrm{Red}(W_1 W_2)) \equiv \mathrm{Red}(\mathrm{Red}(W_0 W_1) W_2).$$

Lemma 2.2 implies the group H_κ is isomorphic to the set Red_κ under the group operation $W * U = \mathrm{Red}(WU)$. We give the following definition (see [Cannon and Conner 2000, Definition 3.4]):

Definition 2.3. Given a word $W \in \mathcal{W}_\kappa$ we say $\mathcal{S} \subseteq \overline{W} \times \overline{W}$ is a *cancellation* provided the following:

(1) For $\langle i_0, i_1 \rangle \in \mathcal{S}$, we have $i_0 < i_1$.

(2) If $\langle i_0, i_1 \rangle \in \mathcal{S}$ and $\langle i_0, i_2 \rangle \in \mathcal{S}$, then $i_2 = i_1$.

(3) If $\langle i_0, i_1 \rangle \in \mathcal{S}$ and $\langle i_2, i_1 \rangle \in \mathcal{S}$, then $i_2 = i_0$.

(4) If $\langle i_0, i_1 \rangle \in \mathcal{S}$ and $i_2 \in (i_0, i_1) \subseteq \overline{W}$, there exists $i_3 \in (i_0, i_1)$ such that either $\langle i_2, i_3 \rangle \in \mathcal{S}$ or $\langle i_3, i_2 \rangle \in \mathcal{S}$.

(5) If $\langle i_0, i_1\rangle \in \mathcal{S}$, then $W(i_0) = (W(i_1))^{-1}$.

The $\langle\,\cdot\,,\,\cdot\,\rangle$ notation for ordered pairs is used here in order to avoid confusion with parenthetical notation $(\,\cdot\,,\,\cdot\,)$ which can be interpreted as an open interval. We shall use $\langle\,\cdot\,\rangle$ to denote a generated subgroup, and the lack of a comma makes this use unambiguous.

A cancellation may be understood as a transfinite strategy for freely reducing a word. Conditions (2) and (3) imply that a cancellation is a pairing of elements in a subset of elements of $\overline{W}$. Condition (5) says that the pairing requires the associated letters in W to be inverses of each other. Condition (4) requires the pairing to be complete in the sense that each element between paired elements must also be paired by $\mathcal{S}$. Condition (4) also requires that the pairing is noncrossing in the sense that if an element i lies between two paired elements i_0 and i_1, then the element with which i is paired must also be between i_0 and i_1.

Zorn's lemma implies that each cancellation $\mathcal{S}$ in a word W is included in a maximal cancellation $\mathcal{S}'$; that is, $\mathcal{S} \subseteq \mathcal{S}'$ and $\mathcal{S}'$ is not a proper subset of a cancellation in W. It turns out that a maximal cancellation reveals the reduced word representative, as happens with freely reducing a finitary word until free reductions are no longer possible. We omit the proof of the following, but it follows in precisely the same manner as [Cannon and Conner 2000, Theorem 3.9]:

Lemma 2.4. *If $\mathcal{S}$ is a maximal cancellation for $W \in \mathcal{W}_\kappa$ then*

$$W \restriction \{i \in \overline{W} \mid (\neg\exists i')\,(\langle i, i'\rangle \in \mathcal{S} \text{ or } \langle i', i\rangle \in \mathcal{S})\} \equiv \mathrm{Red}(W).$$

Thus a word has only trivial cancellation if and only if that word is reduced. As a consequence, if $W \in \mathcal{W}_\kappa$ with $W \equiv \prod_{\lambda\in\Lambda} W_\lambda$ then $\mathrm{Red}(W) \equiv \mathrm{Red}\big(\prod_{\lambda\in\Lambda} \mathrm{Red}(W_\lambda)\big)$.

Now we define our homomorphism from Red_κ to $\pi_1(\mathbb{E}_\kappa, \circ_\kappa)$. For each $\alpha < \kappa$ and $n < \omega$ we let $L_{\alpha,n}$ be a loop based at $\circ_\kappa$ which goes exactly once around the n-th circle of E_α and is injective except at 0, 1. One can use an isometry between $\mathbb{E}_1$ and E_α to define $L_{\alpha,n}$ from L_n if wished. Given a reduced word $W \in \mathrm{Red}_\kappa$ and an order embedding $\iota : \overline{W} \to \mathcal{I}$ we get a loop $\mathrm{R}_\iota(W)$ defined by

$$\mathrm{R}_\iota(W)(t) = \begin{cases} L_{\alpha,n}\big(\frac{t-\inf I}{\sup I-\inf I}\big) & \text{if } W(i) = a_{0,n} \text{ and } t \in I = \iota(i),\\ L_{\alpha,n}^{-1}\big(\frac{t-\inf I}{\sup I-\inf I}\big) & \text{if } W(i) = a_{0,n}^{-1} \text{ and } t \in I = \iota(i),\\ \circ_\kappa & \text{otherwise.}\end{cases}$$

The check that this function on $[0, 1]$ is continuous is straightforward. Given some other order embedding $\iota_0 : \overline{W} \to \mathcal{I}$ we obtain a different loop R_{ι_0} which is homotopic to R_ι via a homotopy which is a reparametrization.

In particular we have a well-defined map $\mathrm{R} : \mathrm{Red}_\kappa \to \pi_1(\mathbb{E}_\kappa, \circ_\kappa)$. To see that this is a homomorphism, we let $W, U \in \mathrm{Red}_\kappa$ and let W_0, W_1, U_0, U_1 be as in

Lemma 2.1. The loop $\mathrm{R}(W_1)$ is readily seen to be the inverse of $\mathrm{R}(U_0)$. The word W_0U_1 is reduced and therefore we have

$$\begin{aligned}\mathrm{R}(W*U) &= \mathrm{R}(\mathrm{Red}(WU))\\ &= \mathrm{R}(W_0U_1)\\ &\simeq \mathrm{R}(W_0)\,\mathrm{R}(U_0)^{-1}\,\mathrm{R}(U_0)\,\mathrm{R}(U_1)\\ &= \mathrm{R}(W_0)\,\mathrm{R}(W_1)\,\mathrm{R}(U_0)\,\mathrm{R}(U_1)\\ &= \mathrm{R}(W_0W_1)\,\mathrm{R}(U_0U_1)\\ &= \mathrm{R}(W)\,\mathrm{R}(U).\end{aligned}$$

Suppose now that $W \in \mathrm{Red}_\kappa$ is in the kernel of R. Suppose for contradiction that $W \not\equiv E$. We'll construct a cancellation $\mathcal{S}$ of W to obtain a contradiction. Fix an order embedding $\iota : \overline{W} \to \mathcal{I}$. Let $H : [0, 1] \times [0, 1] \to \mathbb{E}_\kappa$ be a nullhomotopy of $\mathrm{R}_\iota(W)$. That is, $H(t, 0) = \mathrm{R}_\iota(W)(t)$ and $H(0, s) = H(1, s) = H(t, 1)$ for all $t, s \in [0, 1]$. For each $I \in \mathcal{I}$ we let $m(I)$ signify the midpoint $m(I) = \frac{1}{2}(\sup I + \inf I)$. Consider the set of points $M = \{(m(\iota(i)), 0)\}_{i\in\overline{W}} \subseteq [0, 1] \times [0, 1]$. For each point $p \in M$ we consider its path component P_p in $[0, 1] \times [0, 1] \setminus H^{-1}(\circ_\kappa)$. Each $p \in M$ is associated with a unique interval $\iota(i_p)$ and therefore with a unique element $i_p \in \overline{W}$, and each $i \in \overline{W}$ is in turn associated with a unique point $p \in M$. Moreover, the natural order on points in M is isomorphic with the elements of $\overline{W}$ in this association.

Fixing $p \in M$ the set $P_p \cap M$ is necessarily finite, because each element of $P_p \cap M$ corresponds to exactly one occurrence of a loop $L_{\alpha,n}$ or of its inverse, for a fixed α and n, and there are only finitely many such occurrences since there are finitely many occurrences of $a_{\alpha,n}^{\pm 1}$ in W. Write $P_p \cap M = \{p_0, p_1, \dots, p_j\}$ listing elements in the natural order. By modifying H to have output $\circ_\kappa$ outside of P_p, we see that H witnesses a nullhomotopy of the loop $\mathrm{R}_i(W \restriction \{i_{p_0}, \dots, i_{p_j}\})$, which lies entirely in the n-th circle of E_α. Then there are exactly as many i_{p_k} for which $W(i_{p_k}) = a_{\alpha,n}$ as there are for which $W(i_{p_k}) = a_{\alpha,n}^{-1}$. Select neighboring points p_k, p_{k+1} which are of opposite parity and let $\langle i_{p_k}, i_{p_{k+1}}\rangle \in \mathcal{S}$. Among the remaining points $P_p \cap M \setminus \{p_k, p_{k+1}\}$ select two which are neighboring under the new order and add this ordered pair to $\mathcal{S}$. Continue in this way until all elements of $P_p \cap M$ are used. Perform this procedure on all path components P_p for $p \in M$. It is straightforward to check that $\mathcal{S}$ satisfies the rules of a cancellation. We have obtained our contradiction. Thus R is an injection.

We check that R is a surjection. Let $L : [0, 1] \to \mathbb{E}_\kappa$ be a loop at $\circ_\kappa$. Let $\mathcal{J}$ be the set of maximal open intervals in $[0, 1] \setminus L^{-1}(\circ_\kappa)$. This set is countable and has a natural ordering. For each restriction $L \restriction \bar{J}$, where $J \in \mathcal{J}$, there is a homotopy $H_J : \bar{J} \times [0, 1] \to L(\bar{J})$ to a loop $L_J : \bar{J} \to L(\bar{J})$ which is either constant, or

$L_{\alpha,n}\big(\frac{t-\inf J}{\sup J-\inf J}\big)$ or $L^{-1}_{\alpha,n}\big(\frac{t-\inf J}{\sup J-\inf J}\big)$. By gluing these homotopies together we get a homotopy of L to a loop whose restriction to each nonconstant interval $\bar{J}$ is of the form $L_{\alpha,n}\big(\frac{t-\inf J}{\sup J-\inf J}\big)$ or $L^{-1}_{\alpha,n}\big(\frac{t-\inf J}{\sup J-\inf J}\big)$.

Thus assuming L is of this form, we define a word $W : \mathcal{J} \to \mathcal{A}_\kappa$ by letting $W(J) = a^{\pm 1}_{\alpha,n}$ where the α, n and superscript are determined in the straightforward way. That the mapping W is indeed a word (no n in the subscript occurs infinitely often) follows from the fact that L is continuous. Let $\mathcal{S}$ be a maximal cancellation on W. This $\mathcal{S}$ can be used to homotope L so that the new associated word is $\mathrm{Red}(W)$. More explicitly, we define $H : [0, 1] \times [0, 1] \to \mathbb{E}_\kappa$ by having $H(t, s) = L(t)$ if t does not lie inside an interval $(\inf J_0, \sup J_1)$ where $\langle J_0, J_1\rangle \in \mathcal{S}$. If a point $(t, s) \in [0, 1] \times [0, 1]$ lies on the semicircle determined by points $(t_0, 0)$ and $(t_1, 0)$ which is perpendicular to $[0, 1] \times \{0\}$ where $t_0 \in J_0$, $t_1 \in J_1$ and $\langle J_0, J_1\rangle \in \mathcal{S}$ with $L(t_0) = L(t_1)$, we let $H(t, s) = L(t_0) = L(t_1)$. Give H output $\circ_\kappa$ everywhere else. That H is continuous and produces a loop $H(t, 1)$ as described is intuitive but tedious to check. Thus we may now assume that the associated word W is reduced. By reparametrizing L we may make it so that all the intervals in $\mathcal{J}$ are elements in $\mathcal{I}$, which immediately gives an order embedding ι of $\overline{W}$ to $\mathcal{I}$ for which $L = \mathrm{R}_\iota(W)$. We have shown surjectivity and finished the proof of the following:

Lemma 2.5. *The function* $\mathrm{R} : \mathrm{Red}_\kappa \to \pi_1(\mathbb{E}_\kappa, \circ_\kappa)$ *is an isomorphism.*

We now approach the isomorphism $\mathcal{C}_\kappa \simeq \pi_1(\mathbb{GS}_\kappa, \circ_\kappa)$. For finite values of κ this can be done by a straightforward argument in which van Kampen's Theorem is iterated finitely many times, as is done in [Eda and Fischer 2016, Section 4]. We present an argument which works for every cardinal κ.

Lemma 2.6. *Given* $\epsilon > 0$ *and a loop* $L : [0, 1] \to \mathbb{GS}_\kappa$ *based at* $\circ_\kappa$, *there is a loop homotopic to L whose image is of diameter at most* ϵ.

Proof. Let $\mathcal{J}$ be the set of maximal open intervals in $[0, 1] \setminus L^{-1}(\circ_\kappa)$. There are only finitely many intervals $J \in \mathcal{J}$ for which the diameter of the image $\mathrm{diam}(L \restriction J)$ is at least $\frac{1}{2}\epsilon$. But for every $J \in \mathcal{J}$ the loop $L \restriction \bar{J}$ lies entirely in a contractible space, a homeomorph of $\mathbb{GS}_1$. In particular each restriction $L \restriction \bar{J}$ is nullhomotopic. Thus letting $\mathcal{J}' \subseteq \mathcal{J}$ be the set of those intervals whose images are of diameter at least $\frac{1}{2}\epsilon$ we have L homotopic to the loop $L' : [0, 1] \to \mathbb{GS}_\kappa$ given by

$$L'(t) = \begin{cases} L(t) & \text{if } t \notin \bigcup \mathcal{J}', \\ \circ_\kappa & \text{if } t \in \bigcup \mathcal{J}', \end{cases}$$

which has diameter at most ϵ. □

Let each copy of $(0, 0, 1)$ in the copies of $\mathbb{GS}_1$ whose wedge forms $\mathbb{GS}_\kappa$ be called a "cone tip". Let $\mathbb{GS}'_\kappa$ denote the space $\mathbb{GS}_\kappa$ minus the set of cone tips. The following is easy to see.

Lemma 2.7. *The space $\mathbb{GS}'_\kappa$ strongly deformation retracts to $\mathbb{E}_\kappa$.*

Now we let $U \subseteq \mathbb{GS}_\kappa$ be the open set which is the union over all $\alpha < \kappa$ of images $E_\alpha \times \left[0, \frac{2}{3}\right)$ in the cone over E_α. For each $\alpha < \kappa$ we let V_α be the image of $E_\alpha \times \left(\frac{1}{3}, 1\right]$ in the cone over E_α. An application of van Kampen's theorem gives the following.

Theorem 2.8. *The isomorphism* $\mathrm{R} : \mathrm{Red}_\kappa \to \pi_1(\mathbb{E}_\kappa, \circ_\kappa)$ *descends to an isomorphism* $\mathrm{R}_{\mathcal{C}_\kappa} : \mathcal{C}_\kappa \to \pi_1(\mathbb{GS}_\kappa, \circ_\kappa)$.

We immediately obtain the following (cf. [Bogopolski and Zastrow 2012, Theorem 8.1]):

Corollary 2.9. *A reduced word W is in the kernel of the map* $\mathrm{Red}_\kappa \to \mathcal{C}_\kappa$ *if and only if there exist finitely many intervals $I_0, \ldots, I_p$ such that $W \restriction I_j$ is pure for each j and* $\mathrm{Red}\big(W \restriction \big(\overline{W} \setminus \bigcup_{j=0}^{p} I_j\big)\big) = E$.

Lemma 2.10. *Suppose that we have a word $V \equiv \prod_{n\in\omega} V_n$ with $V \in \mathrm{Red}_\kappa$, and that the following properties are verified*:

(1) *Any interval $I \subseteq \overline{V}$ such that $V \restriction I$ is pure is a subinterval of $\overline{\prod_{n=0}^{m} V_n}$ for some $m \in \omega$.*

(2) *For each $n \in \omega$ there exists $j_n \in \omega$ such that* $\left|\{i \in \overline{V_n} \mid \mathrm{proj}_1(V_n(i)) = j_n\}\right| > \sum_{m\neq n}\left|\{i \in \overline{V_m} \mid \mathrm{proj}_1(V_m(i)) = j_n\}\right|$.

Then $[[V]] \neq [[E]]$ in $\mathcal{C}_\kappa$.

Proof. Suppose for contradiction that $[[V]] = [[E]]$, so by Corollary 2.9 we obtain a finite collection of intervals $I_0, \ldots, I_p$ in $\overline{V}$ such that $V \restriction I_k$ is pure for each $0 \le k \le p$ and $\mathrm{Red}\big(V \restriction \big(\overline{V} \setminus \bigcup_{k=0}^{p} I_k\big)\big) = E$. Let $\mathcal{S}$ be a maximal cancellation of $V \restriction \big(\overline{V} \setminus \bigcup_{k=0}^{p} I_k\big)$. We know by (1) that $\bigcup_{k=0}^{p} I_k \subseteq \overline{\prod_{n=0}^{m} V_n}$ for some $m \in \omega$. All elements of $Z = \{i \in \overline{V_{m+1}} \mid \mathrm{proj}_1(V_{m+1}(i)) = j_{m+1}\}$ must participate in $\mathcal{S}$ since $\mathrm{Red}\big(V \restriction \big(\overline{V} \setminus \bigcup_{k=0}^{p} I_k\big)\big) = E$, but since V_{m+1} is reduced we know that the elements of Z are paired with elements of $\overline{V} \setminus \big(\overline{V_{m+1}} \cup \bigcup_{k=0}^{p} I_k\big)$, but this is impossible by condition (2). □

For a reduced word W we let $[[W]]$ denote the equivalence class of W in $\mathcal{C}_\kappa$ and if $[[W]] = [[U]]$ we write $W \approx U$.

Theorem 2.11. *For each cardinal κ we have*

$$|\mathcal{C}_\kappa| = \begin{cases} 1 & \textit{if } \kappa = 0, \\ \kappa^{\aleph_0} & \textit{if } \kappa \geq 1. \end{cases}$$

Proof. We have already seen that the formula holds in case $\kappa = 0, 1$. Suppose $\kappa \geq 2$. Notice that the space $\mathbb{GS}_\kappa$ has $2^{\aleph_0} \cdot \kappa = \max\{2^{\aleph_0}, \kappa\}$ points in it. Every continuous function from $[0, 1]$ to the metric space $\mathbb{GS}_\kappa$ is totally determined by the restriction to $[0, 1] \cap \mathbb{Q}$. Thus there are at most $(\max\{2^{\aleph_0}, \kappa\})^{\aleph_0} = \kappa^{\aleph_0}$ loops in the space, so in particular $|\mathcal{C}_\kappa| \leq \kappa^{\aleph_0}$. We must show $|\mathcal{C}_\kappa| \geq \kappa^{\aleph_0}$.

If $2 \leq \kappa \leq 2^{\aleph_0}$ then let Σ be a collection of infinite subsets of ω such that for distinct $X, Y \in \Sigma$ we have $X \cap Y$ finite and such that $|\Sigma| = 2^{\aleph_0}$. Such a construction is straightforward, see for example [Kunen 1980, Chapter II, Theorem 1.3]. For each $X \in \Sigma$ let $X = \{n_{0,X}, n_{1,X}, \ldots\}$ be the enumeration of X in the natural order. Let

$$W_X \equiv a_{0,n_{0,X}} a_{1,n_{1,X}} a_{0,n_{2,X}} a_{1,n_{3,X}} \cdots .$$

Since W_X uses only positive letters it is clear that W_X and also any deletion of finitely many letters of W_X is a reduced word. By the conditions on Σ it is clear that $[[W_X]] \neq [[W_Y]]$ if $X \neq Y$. Then $\kappa^{\aleph_0} \leq |\mathcal{C}_\kappa|$.

Suppose that $2^{\aleph_0} < \kappa$ and that $\kappa^{\aleph_0} = \kappa$. Let $f : \kappa \times \omega \to \kappa$ be an injection and for each $\alpha < \kappa$ we define $W_\alpha \equiv a_{f(\alpha,0),0} a_{f(\alpha,1),1} \cdots$. It is clear that $[[W_\alpha]] \neq [[W_\beta]]$ for distinct $\alpha, \beta < \kappa$.

Suppose finally that $2^{\aleph_0} < \kappa$ and that $\kappa^{\aleph_0} > \kappa$. Let X be the set of all functions from ω to κ and consider two functions $\sigma_0, \sigma_1 \in X$ to be equivalent if they are eventually identical: for some $m \in \omega$ we have $\sigma_0(m + n) = \sigma_1(m + n)$ for all $n \in \omega$. Each equivalence class is of cardinality κ, so there are exactly $\kappa^{\aleph_0}$ distinct equivalence classes. Letting $Y \subset X$ be a selection from each equivalence class we define a map $Y \to \mathcal{C}_\kappa$ by letting $\sigma \mapsto W_\sigma$ where $W_\sigma \equiv a_{f(\sigma(0),0),0} a_{f(\sigma(1),1),1} \cdots$ and again $f : \kappa \times \omega \to \kappa$ is an injection. It is easy to see that for distinct elements of Y the assigned words are not equivalent in $\mathcal{C}_\kappa$. $\square$

An interval I in a totally ordered set Λ is *initial* if it is a union of intervals of the form $(-\infty, i]$ and is *terminal* if a union of intervals of form $[i, \infty)$ (an initial or terminal interval may be empty). Given a nonempty word $W \in \mathrm{Red}_\kappa$ there exists a unique maximal initial interval I_0 of $\overline{W}$ for which there exists a terminal interval $I_1 \subseteq \overline{W}$ such that $W \upharpoonright I_0 \equiv (W \upharpoonright I_1)^{-1}$. By the proof of [Eda 1992, Corollary 1.6] the maximal such initial interval I_0 and the accompanying I_1 are disjoint and $\overline{W} \setminus (I_0 \cup I_1)$ is nonempty, and this set is clearly an interval, say I_2. Thus $W \equiv (W \upharpoonright I_0)(W \upharpoonright I_2)(W \upharpoonright I_0)^{-1}$ and we call the word $W \upharpoonright I_2$ the *cyclic reduction* of W. Clearly if U is the cyclic reduction of W then the cyclic reduction of U is again U, so cyclic reduction is an idempotent operation. A word whose cyclic reduction is itself is called *cyclically reduced*. It is clear from Lemma 2.4 that a word U is cyclically reduced if and only if the word U^n is reduced for all $n \geq 1$, thus if and only if U^2 is reduced.

3. Theorem A

We begin with a description of the overall strategy and then describe the structure of this section. An isomorphism between two cone groups $\mathcal{C}_{\kappa_0}$ and $\mathcal{C}_{\kappa_1}$ will be constructed by induction on specially defined subgroups. We cannot expect that such an isomorphism will be imposed by a homomorphism $\mathrm{Red}_{\kappa_0} \to \mathrm{Red}_{\kappa_1}$. However, the idea is that establishing careful correspondences between certain words in Red_{κ_0} and certain words in Red_{κ_1} will allow us to ultimately produce homomorphisms $\phi_0 : \mathrm{Red}_{\kappa_0} \to \mathcal{C}_{\kappa_1}$ and $\phi_1 : \mathrm{Red}_{\kappa_1} \to \mathcal{C}_{\kappa_0}$ which will descend to isomorphisms $\Phi_0 : \mathcal{C}_{\kappa_0} \to \mathcal{C}_{\kappa_1}$ and $\Phi_1 : \mathcal{C}_{\kappa_1} \to \mathcal{C}_{\kappa_0}$ with $\Phi_1 = \Phi_0^{-1}$.

What sort of correspondences between words should be produced? They should not be so rigid as to produce a homomorphism $\mathrm{Red}_{\kappa_0} \to \mathrm{Red}_{\kappa_1}$. Rather, they should be forgiving enough to produce the homomorphisms ϕ_0 and ϕ_1 described above. The correspondences should also agree with each other so that the ϕ_0 and ϕ_1 are well defined.

Each word in Red_{κ_0} and Red_{κ_1} may be decomposed in a natural way as a concatenation of maximal pure subwords (the index over which concatenation is written is unique up to order isomorphism and is called the *p-index*). Taking concatenations over subintervals of the p-index gives us words which are recognizable pieces of the original word (which we will call *p-chunks*). There is a natural way of comparing certain words $W \in \mathrm{Red}_{\kappa_0}$ with other words $U \in \mathrm{Red}_{\kappa_1}$ via an order isomorphism between a subset of the p-index of W and that of U. These subsets will be large enough to "capture" any interval of the p-index, up to deletion of finitely many elements, and there will be a correspondence between the p-chunks of W and those of U. The bijections between the subsets of the p-indices will honor word concatenation (up to finite deletion of pure subwords) and will allow us to define isomorphisms between the subgroups of $\mathcal{C}_{\kappa_0}$ and $\mathcal{C}_{\kappa_1}$ which are generated by the p-chunks of the words on which we have defined such bijections.

In order to have the isomorphisms be well defined, it is essential that the imposed correspondences between p-chunks are in agreement with each other. That is, suppose that $W_0, W_1 \in \mathrm{Red}_{\kappa_0}$ and $U_0, U_1 \in \mathrm{Red}_{\kappa_1}$ and W_i is made to correspond to U_i for $i = 0, 1$. If $W \in \mathrm{Red}_{\kappa_0}$ is a p-chunk of each of W_0 and W_1 then we should be able to make W correspond to a word $U \in \mathrm{Red}_{\kappa_0}$ in a way that honors the correspondences $W_i \leftrightarrow U_i$, so any choice of such a U should be independent of whether we are considering W as a p-chunk of W_0 or of W_1, up to the equivalence $\approx$.

It will be necessary to be able to define many such correspondences between words, so as to make the isomorphism between subgroups of $\mathcal{C}_{\kappa_0}$ and $\mathcal{C}_{\kappa_1}$ have larger and larger domain and range. Keeping such new correspondences in agreement with the previously defined ones requires us to consider concatenations of words on which such bijections have already been defined, concatenations of order type ω and of order type $\mathbb{Q}$ are of particular concern. If we can continue to do this for sufficiently many steps ($2^{\aleph_0}$ steps will suffice) then we can succeed in the construction.

This section is organized into subsections for the sake of clarity. We introduce and prove some basic properties of p-chunks in Section 3A. In Section 3B we will make precise the concept of a "sufficiently large" subset of an ordered set. In Section 3C we define what it means for bijections between sufficiently large subsets of p-indices to honor word concatenation (up to deletion of finitely many pure subwords). In Section 3D we give some baby steps towards defining such bijections on more words, and in Sections 3E and 3F we show how to extend such notions for ω- and $\mathbb{Q}$-type concatenations, respectively. Finally in Section 3G we combine all the previous ideas to prove Theorems A and B.

3A. ***P-chunks.*** Let κ be a cardinal. For each word $W \in \mathrm{Red}_\kappa$ we have a decomposition of the domain $\overline{W} \equiv \prod_{\lambda\in\Lambda} \Lambda_\lambda$ such that each Λ_λ is a nonempty maximal interval with $W \upharpoonright \Lambda_\lambda$ pure. We'll call this decomposition the *pure decomposition of the domain of* W. Write $W \equiv_{\mathrm{p}} \prod_{\lambda\in\Lambda} W_\lambda$ to express that $\overline{W} \equiv \prod_{\lambda\in\Lambda} \overline{W_\lambda}$ is the p-decomposition of the domain of W, and call this writing $W \equiv_{\mathrm{p}} \prod_{\lambda\in\Lambda} W_\lambda$ the *p-decomposition of* W and Λ the *p-index*, denoted $\mathrm{p}{*}(W)$. By definition we therefore have $E \equiv_{\mathrm{p}} \prod_{\lambda\in\Lambda} W_\lambda$ with $\Lambda = \varnothing$. If $W \equiv_{\mathrm{p}} \prod_{\lambda\in \mathrm{p}*(W)} W_\lambda$ and I is an interval in $\mathrm{p}{*}(W)$ then let $W \upharpoonright_{\mathrm{p}} I$ denote the word $\prod_{\lambda\in I} W_\lambda$. Call a word W' a *p-chunk* of W if for some interval $I \subseteq \mathrm{p}{*}(W)$ we have $W' \equiv W \upharpoonright_{\mathrm{p}} I$. For a given $W \in \mathrm{Red}_\kappa$ we let $\text{p-chunk}(W)$ denote the set of p-chunks of W. A pure p-chunk of a word $W \equiv_{\mathrm{p}} \prod_{\lambda\in\Lambda} W_\lambda$ will, of course, either be empty or one of the W_λ. Notice as well that an equivalence $W \equiv U$ immediately gives an order isomorphism from $\mathrm{p}{*}(W)$ to $\mathrm{p}{*}(U)$.

Lemma 3.1. *Suppose that $W \equiv_{\mathrm{p}} \prod_{\lambda\in\Lambda} W_\lambda$ and $U \equiv_{\mathrm{p}} \prod_{\lambda'\in\Lambda'} U_{\lambda'}$. Then there exists a (possibly empty) initial interval $I \subseteq \Lambda$ and a (possibly empty) terminal interval $I' \subseteq \Lambda'$ such that either:*

(i) $\mathrm{Red}(WU) \equiv_{\mathrm{p}} \prod_{\lambda\in I} W_\lambda \prod_{\lambda'\in I'} U_{\lambda'}$; *or*

(ii) *there exist $\lambda_0 \in \Lambda$ which is the least element strictly above all elements in I, $\lambda_1 \in \Lambda'$ which is the greatest element strictly below all elements of I' and*

$$\mathrm{Red}(WU) \equiv_{\mathrm{p}} \Big(\prod_{\lambda\in I} W_\lambda\Big) V \Big(\prod_{\lambda'\in I'} U_{\lambda'}\Big)$$

where $V \equiv \mathrm{Red}(W_{\lambda_0} U_{\lambda_1}) \not\equiv E$ is pure.

Proof. Since both W and U are reduced we have reduced words W_0, W_1, U_0, U_1 such that $W \equiv W_0 W_1$, $U \equiv U_0 U_1$, $W_1 \equiv U_0^{-1}$ and $W_0 U_1$ is reduced, by Lemma 2.1. Select $I_0 \subseteq \Lambda$ to be a maximal initial interval for which $\bigcup_{\lambda\in I_0} \overline{W_\lambda} \subseteq \overline{W_0}$. Select $I_1' \subseteq \Lambda'$ to be a maximal terminal interval such that $\bigcup_{\lambda'\in I_1'} \overline{U_{\lambda'}} \subseteq \overline{U_1}$.

Suppose $\prod_{\lambda\in I_0} W_\lambda \equiv W_0$ and $\prod_{\lambda'\in I_1'} U_{\lambda'} \equiv U_1$. If I_0 has a maximal element λ_0 and I_1' has a minimal element λ_1 such that the words W_{λ_0} and U_{λ_1} are both α-pure for some α, then we let $I = I_0 \setminus \{\lambda_0\}$ and $I' = I_1' \setminus \{\lambda_1\}$ and $V \equiv W_{\lambda_0} U_{\lambda_1}$ and obviously condition (ii) holds. If there are no such maximal and minimal elements then condition (i) holds.

Suppose that $\prod_{\lambda\in I_0} W_\lambda \not\equiv W_0$. Then there exists some λ_0 which is the least element strictly above all elements in I_0 and nonempty words $W_{\lambda_0,0}$ and $W_{\lambda_0,1}$ such that

$$W_{\lambda_0} \equiv W_{\lambda_0,0}W_{\lambda_0,1}; \quad W_0 \equiv_{\mathrm{p}} \Big(\prod_{\lambda\in I_0} W_\lambda\Big)W_{\lambda_0,0}; \quad W_1 \equiv_{\mathrm{p}} W_{\lambda_0,1}\Big(\prod_{\lambda\in\Lambda\setminus(I_0\cup\{\lambda_0\})} W_\lambda\Big).$$

If in addition $\prod_{\lambda'\in I_1} U_{\lambda'} \equiv U_1$ then $\Lambda'\setminus I_1$ has a maximum element λ_1 which satisfies $U_{\lambda_1} \equiv W_{\lambda_0,1}^{-1}$. Thus we let $I = I_0 \setminus \{\lambda_0\}$ and $I' = I_1$ and $V \equiv W_{\lambda_0,0} \equiv \mathrm{Red}(W_{\lambda_0}U_{\lambda_1})$ and we have condition (ii). On the other hand, if in addition we have $\prod_{\lambda'\in I_1} U_{\lambda'} \not\equiv U_1$ then $\Lambda' \setminus I_1$ has a maximum element λ_1 and there exist nonempty words $U_{\lambda_1,0}$ and $U_{\lambda_1,1}$ for which

$$U_{\lambda_1} \equiv U_{\lambda_1,0}U_{\lambda_1,1}; \quad U_0 \equiv_{\mathrm{p}} \Big(\prod_{\lambda'\in\Lambda'\setminus I_1} U_{\lambda'}\Big)U_{\lambda_1,0}; \quad U_1 \equiv_{\mathrm{p}} U_{\lambda_1,1}\Big(\prod_{\lambda'\in I_1} U_{\lambda'}\Big).$$

Then we let $V \equiv W_{\lambda_0,0}V_{\lambda_1,1} \equiv \mathrm{Red}(W_{\lambda_0}U_{\lambda_1})$ and $I = I_0$ and $I' = I_1$ and condition (ii) holds.

The case where $\prod_{\lambda\in I_0} W_\lambda \equiv W_0$ and $\prod_{\lambda'\in I_1'} U_{\lambda'} \not\equiv U_1$ follows from dualizing the proof of an earlier case. □

Lemma 3.2. *Suppose that $X \subseteq \mathrm{Red}_\kappa$. For each nonempty element W of the subgroup $\langle\bigcup_{U\in X} \text{p-chunk}(U)\rangle \le \mathrm{Red}_\kappa$, if $W \equiv_{\mathrm{p}} \prod_{\lambda\in\Lambda} W_\lambda$, then there exist nonempty intervals $I_0, \dots, I_n$ in Λ such that*:

(i) $\Lambda \equiv \prod_{i=0}^n I_i$.

(ii) *For each $0 \le i \le n$, at least one of the following holds*:

 (a) *I_i is a singleton $\{\lambda\}$ such that W_λ is the reduction of a finite concatenation of pure p-chunks of elements in $X^{\pm 1}$.*

 (b) *$\prod_{\lambda\in I_i} W_\lambda$ is a p-chunk of some element in $X^{\pm 1}$.*

Proof. The elements of $\langle\bigcup_{U\in X} \text{p-chunk}(U)\rangle$ are of form $\mathrm{Red}(U_0\cdots U_l)$ where each U_i is a p-chunk of an element of $X^{\pm 1}$. The claim will follow by an induction on the number l. If $l = 0$ or $l = 1$ then we are already done. Supposing that the claim holds for l, we suppose $W \equiv \mathrm{Red}(U_0\cdots U_{l+1}) \equiv \mathrm{Red}(\mathrm{Red}(U_0\cdots U_l)U_{l+1})$ and let $W' \equiv \mathrm{Red}(U_0\cdots U_l)$ and $U \equiv U_{l+1}$. Let $W' \equiv W_0W_1$ and $U \equiv U_0U_1$ as in Lemma 2.1 for performing the reduction $\mathrm{Red}(W'U)$. Let $W' \equiv_{\mathrm{p}} \prod_{\lambda\in\Lambda} W_\lambda$ and $U \equiv_{\mathrm{p}} \prod_{\lambda'\in\Lambda'} U_\lambda$. By induction we have for the word W' a decomposition $I_0, \dots, I_{n'}$ as in the conclusion of this lemma. We can select an initial interval $I \subseteq \Lambda$ and a terminal interval $I' \subseteq \Lambda'$ as in the conclusion of Lemma 3.1. Consider the two possible cases in Lemma 3.1 for the word $W \equiv \mathrm{Red}(W'U)$. If case (i) of Lemma 3.1 holds then we can decompose the p-chunk total order for W into at most $n' + 1$

intervals as in (i) and (ii) of the statement of the lemma that we are proving. If case (ii) of Lemma 3.1 holds then we can decompose the p-chunk total order for W into at most $n' + 2$ intervals, at least one of which will be a singleton. □

We say a subgroup G of Red_κ is *p-fine* if each p-chunk U of each $W \in G$ is also in G (cf. [Eda 1999, page 600]).

Lemma 3.3. *If* $X \subseteq \mathrm{Red}_\kappa$ *then the subgroup* $\left\langle \bigcup_{U \in X} \text{p-chunk}(U) \right\rangle \leq \mathrm{Red}_\kappa$ *is p-fine. This is the smallest p-fine subgroup including the set* X.

Proof. This follows immediately from the characterization in Lemma 3.2. □

Given a set $X \subseteq \mathrm{Red}_\kappa$ we'll denote the subgroup

$$\left\langle \bigcup_{U \in X} \text{p-chunk}(U) \right\rangle \leq \mathrm{Red}_\kappa$$

by $\mathrm{Pfine}(X)$.

Lemma 3.4. *If* $X \subseteq \mathrm{Red}_\kappa$ *then there are at most* $(|X| + 1) \cdot \aleph_0$ *pure p-chunks of elements in* $\mathrm{Pfine}(X)$.

Proof. If X is empty then $\mathrm{Pfine}(X)$ has only the empty word and so there is one pure p-chunk of elements in $\mathrm{Pfine}(X)$ and the claim is true. If X is not empty then there are there are at most $|X| \cdot \aleph_0$ pure p-chunks of elements in X (since a p-index is at most countable), and therefore we have at most $|X| \cdot \aleph_0 \cdot \aleph_0 = |X| \cdot \aleph_0$ finite products of p-chunks, or their inverses, of elements in X. By Lemma 3.2 we know all pure p-chunks of elements in $\mathrm{Pfine}(X)$ arise in this way and so we are also done in this case. □

3B. *Close subsets.* We take a diversion through a concept which will be useful in later subsections.

Definition 3.5. Let Λ be a totally ordered set. We say $\Lambda_0 \subseteq \Lambda$ is *close in* Λ, and write $\mathrm{Close}(\Lambda_0, \Lambda)$, if every infinite interval in Λ has nonempty intersection with Λ_0.

The idea of a close subset Λ_0 in Λ is that there are no infinite gaps in Λ which miss elements in Λ_0. We give some elementary examples. If Λ_0 is cofinite in Λ then $\mathrm{Close}(\Lambda_0, \Lambda)$. Any infinite subset of the ordered set ω of natural numbers is close. A subset of $\mathbb{Z}$ is close precisely when it contains numbers of arbitrarily large positive numbers and arbitrarily large negative numbers. A subset of $\mathbb{Q}$ is close when it is dense.

Lemma 3.6. *The following hold*:

(i) *If* $\mathrm{Close}(\Lambda_0, \Lambda)$ *then for any infinite interval* $I \subseteq \Lambda$ *the set* $I \cap \Lambda_0$ *is infinite.*

(ii) *If* $\Lambda_2 \subseteq \Lambda_1 \subseteq \Lambda_0$ *with* $\mathrm{Close}(\Lambda_{i+1}, \Lambda_i)$ *for* $i = 0, 1$, *then* $\mathrm{Close}(\Lambda_2, \Lambda_0)$.

(iii) *If we have that* $\Lambda \equiv \prod_{\theta\in\Theta} \Lambda_\theta$, $\mathrm{Close}(\Lambda_{\theta,0}, \Lambda_\theta)$ *for each* $\theta \in \Theta$, *and also* $\mathrm{Close}(\{\theta \in \Theta \mid \Lambda_{\theta,0} \neq \varnothing\}, \Theta)$, *then* $\mathrm{Close}\big(\bigcup_{\theta\in\Theta} \Lambda_{\theta,0}, \Lambda\big)$.

(iv) *If* I_0 *is an interval in* Λ *and* $\mathrm{Close}(\Lambda_0, \Lambda)$, *then* $\mathrm{Close}(\Lambda_0 \cap I_0, I_0)$.

Proof. (i) If instead $I \cap \Lambda_0 = \{\lambda_0, \lambda_1, \ldots, \lambda_n\}$ with $\lambda_i < \lambda_{i+1}$ then at least one of the intervals $I \cap (-\infty, \lambda_0), (\lambda_0, \lambda_1), \ldots, (\lambda_{n-1}, \lambda_n), I \cap (\lambda_n, \infty)$ in Λ is infinite, but each has empty intersection with Λ_0 and this is a contradiction.

(ii) Let $I \subseteq \Lambda_0$ be an infinite interval. Notice that $I \cap \Lambda_1$ is infinite by (i) and so $I \cap \Lambda_1$ is an infinite interval in Λ_1, so $I \cap \Lambda_2 = (I \cap \Lambda_1) \cap \Lambda_2 \neq \varnothing$.

(iii) Let $I \subseteq \Lambda$ be an infinite interval. The set $\underline{I} = \{\theta \in \Theta \mid I \cap \Lambda_\theta \neq \varnothing\}$ is an interval in Θ. If $\underline{I}$ is finite then as $I = \bigsqcup_{\theta\in\underline{I}} (I \cap \Lambda_\theta)$ there is some $\theta_0 \in \underline{I}$ for which $|I \cap \Lambda_{\theta_0}| = \infty$, and as $I \cap \Lambda_{\theta_0}$ is an infinite interval in Λ_{θ_0} we see that $I \cap \Lambda_{\theta_0,0} \neq \varnothing$, so $I \cap \bigcup_{\theta\in\Theta} \Lambda_{\theta,0} \neq \varnothing$. If $\underline{I}$ is infinite then $\underline{I} \cap \{\theta \in \Theta \mid \Lambda_{\theta,0} \neq \varnothing\}$ is infinite by (i), as we are assuming $\mathrm{Close}(\{\theta \in \Theta \mid \Lambda_{\theta,0} \neq \varnothing\}, \Theta)$. Then there exists some $\theta_0 \in \underline{I} \cap \{\theta \in \Theta \mid \Lambda_{\theta,0} \neq \varnothing\}$ for which $I \supseteq \Lambda_{\theta_0}$. Thus $I \cap \Lambda_{\theta_0,0} \neq \varnothing$.

(iv) This is obvious. □

If $\mathrm{Close}(\Lambda_0, \Lambda)$ then for each interval $I \subseteq \Lambda$ we let $\propto(I, \Lambda_0)$ denote the smallest interval in Λ which includes the set $I \cap \Lambda_0$. In other words $\propto(I, \Lambda_0) = \bigcup_{\lambda_0,\lambda_1\in I\cap\Lambda_0,\lambda_0\leq\lambda_1} [\lambda_0, \lambda_1]$ where the intervals $[\lambda_0, \lambda_1]$ are being considered in Λ.

Lemma 3.7. *Let* $\mathrm{Close}(\Lambda_0, \Lambda)$ *and* $I \subseteq \Lambda$ *be an interval.*

(i) *The inclusion* $I \supseteq \propto(I, \Lambda_0)$ *holds and* $\propto(I, \Lambda_0) = \propto(\propto(I, \Lambda_0), \Lambda_0)$.

(ii) *The set* $I \setminus \propto(I, \Lambda_0)$ *is the disjoint union of an initial and terminal subinterval* $I_0, I_1 \subseteq I$ (*either subinterval could be empty*) *with* $|I_0|, |I_1| < \infty$.

Proof. (i) The claimed inclusion is obvious. For the claimed equality it is therefore sufficient to prove that $\propto(I, \Lambda_0) \subseteq \propto(\propto(I, \Lambda_0), \Lambda_0)$. We let $\lambda \in \propto(I, \Lambda_0)$ be given. Select $\lambda_0, \lambda_1 \in I \cap \Lambda_0$ such that $\lambda_0 \leq \lambda \leq \lambda_1$. Then $\lambda_0, \lambda_1 \in \propto(I, \Lambda_0) \cap \Lambda_0$ and $\lambda_0 \leq \lambda \leq \lambda_1$, so $\lambda \in \propto(\propto(I, \Lambda_0), \Lambda_0)$.

(ii) If $I \cap \Lambda_0 = \varnothing$ then I is finite (since $\mathrm{Close}(\lambda_0, \Lambda)$) and we can let $I_0 = \varnothing$ and $I_1 = I$. If $I \cap \Lambda_0 \neq \varnothing$ then we let $I_0 = \{\lambda \in I \mid (\forall \lambda_0 \in I \cap \Lambda_0)\ \lambda < \lambda_0\}$ and $I_1 = \{\lambda \in I \mid (\forall \lambda_0 \in I \cap \Lambda_0)\ \lambda > \lambda_0\}$. Clearly $I \equiv I_0 \propto(I, \Lambda_0) I_1$. Each of I_0 and I_1 is a subinterval of I and therefore a subinterval of Λ as well. If, say, I_0 is infinite then $I_0 \cap \Lambda_0 \neq \varnothing$ but this is an obvious contradiction. □

We will say that two totally ordered sets Λ and Θ are *close-isomorphic* if there exist $\Lambda_0 \subseteq \Lambda$ and $\Theta_0 \subseteq \Theta$ with $\mathrm{Close}(\Lambda_0, \Lambda)$, $\mathrm{Close}(\Theta_0, \Theta)$ and Λ_0 order isomorphic to Θ_0; and if ι is an order isomorphism between such a Λ_0 and Θ_0 then we will call ι a *close order isomorphism from* Λ *to* Θ. It is obvious that the inverse of a close order isomorphism from Λ to Θ is a close order isomorphism from Θ to Λ.

From a close order isomorphism (abbreviated *coi*) between totally ordered sets one obtains a reasonable way of identifying intervals in one totally ordered set with intervals in the other, which we now describe. Given coi ι between Λ and Θ, with Λ_0 and Θ_0 being the respective domain and range of ι, and an interval $I \subseteq \Lambda$ we let $\propto(I, \iota)$ denote the smallest interval in Θ which includes the set $\iota(I \cap \Lambda_0)$. Thus $\propto(I, \iota) = \bigcup_{\theta_0, \theta_1 \in \iota(I \cap \Lambda_0), \theta_0 \leq \theta_1} [\theta_0, \theta_1]$, where each interval $[\theta_0, \theta_1]$ is being considered in Θ.

Lemma 3.8. *If $\iota : \Lambda_0 \to \Theta_0$ is a coi between Λ and Θ and $I \subseteq \Lambda$ is an interval then $\propto(\propto(I, \iota), \iota^{-1}) = \propto(I, \Lambda_0)$.* □

We point out that a coi ι between Λ and Θ also induces a coi between the reversed orders Λ^{-1} and Θ^{-1} in the obvious way.

Lemma 3.9. *Let $\Lambda \equiv I_0 \cdots I_n$ and $\iota : \Lambda_0 \to \Theta_0$ a coi from Λ to Θ. Then there exist (possibly empty) finite subintervals $I'_0, \dots, I'_{n+1}$ of $\propto(\Lambda, \iota)$ such that*

$$\propto(\Lambda, \iota) \equiv I'_0 \propto(I_0, \iota) I'_1 \propto(I_1, \iota) I'_2 \cdots \propto(I_n, \iota) I'_{n+1}.$$

Proof. Assume the hypotheses. Clearly each $\propto(I_j, \iota)$ is a subinterval of $\propto(\Lambda, \iota)$, and it is easy to see that all elements of $\propto(I_j, \iota)$ are strictly below all elements of $\propto(I_{j+1}, \iota)$ for $0 \leq j < n$. Thus we may indeed write

$$\propto(\Lambda, \iota) \equiv I'_0 \propto(I_0, \iota) I'_1 \propto(I_1, \iota) I'_2 \cdots \propto(I_n, \iota) I'_{n+1}$$

and we conclude by pointing out that $I'_l \cap \Theta_0 = I'_l \cap \iota(\Lambda_0) = I'_l \cap \left(\bigcup_{j=0}^{n} \iota(I_j \cap \Lambda_0)\right) \subseteq \bigcup_{j=0}^{n} (I'_l \cap \propto(I_j, \iota)) = \varnothing$ for each $0 \leq l \leq n+1$, and since $\mathrm{Close}(\Theta_0, \Theta)$ we have I'_l finite. □

Lemma 3.10. *Let $\iota : \Lambda_0 \to \Theta_0$ be a coi from Λ to Λ_0. If $I \subseteq \Lambda$ is finite then $\propto(I, \iota)$ is finite.*

Proof. Since I is finite, we know $I \cap \Lambda_0$ is finite. Clearly we have $\propto(I, \iota) \cap \Theta_0 = \iota(I \cap \Lambda_0)$, so $\propto(I, \iota)$ is an interval in Θ having finite intersection with Θ_0. Thus $\propto(I, \iota)$ is finite by Lemma 3.6 (i). □

3C. *Coherent coi triples.* Suppose that κ_0 and κ_1 are cardinal numbers greater than or equal to 2. For words $W \in \mathrm{Red}_{\kappa_0}$ and $U \in \mathrm{Red}_{\kappa_1}$ we'll write $\mathrm{coi}(W, \iota, U)$ to denote that ι is a coi between $\mathrm{p}_*(W)$ and $\mathrm{p}_*(U)$ and say that $\mathrm{coi}(W, \iota, U)$ is a *coi triple from* Red_{κ_0} *to* Red_{κ_1}. We will often abuse language and say that ι is a coi from W to U when really ι is a coi from $\mathrm{p}_*(W)$ to $\mathrm{p}_*(U)$.

Definition 3.11. A collection $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x \in X}$ of coi triples from Red_{κ_0} to Red_{κ_1} is *coherent* if for any choice of $x_0, x_1 \in X$, intervals $I_0 \subseteq \mathrm{p}_*(W_{x_0})$ and $I_1 \subseteq \mathrm{p}_*(W_{x_1})$ and $i \in \{-1, 1\}$ such that $W_{x_0} \upharpoonright_{\mathrm{p}} I_0 \equiv (W_{x_1} \upharpoonright_{\mathrm{p}} I_1)^i$ we get

$$[[U_{x_0} \upharpoonright_{\mathrm{p}} \propto(I_0, \iota_{x_0})]] = [[(U_{x_1} \upharpoonright_{\mathrm{p}} \propto(I_1, \iota_{x_1}))^i]],$$

and similarly for any choice of $x_2, x_3 \in X$, intervals $I_2 \subseteq$ p-chunk(U_{x_2}) and $I_3 \subseteq$ p-chunk(U_{x_3}) and $j \in \{-1, 1\}$ such that $U_{x_2} \restriction_{\mathrm{p}} I_2 \equiv (U_{x_3} \restriction_{\mathrm{p}} I_3)^j$ we get

$$[[W_{x_2} \restriction_{\mathrm{p}} \propto(I_2, \iota_{x_2}^{-1})]] = [[(W_{x_3} \restriction_{\mathrm{p}} \propto(I_3, \iota_{x_3}^{-1}))^j]].$$

It is clear from the symmetric nature of this definition that if the collection of coi triples $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ from Red_{κ_0} to Red_{κ_1} is coherent then so is the collection of coi triples $\{\mathrm{coi}(U_x, \iota_x^{-1} W_x)\}_{x\in X}$ from Red_{κ_1} to Red_{κ_0}. We emphasize that a word can appear multiple times in a coherent collection. For example, if each element of $\{W_x\}_{x\in X}$ is pure then the collection $\{(W_x, \iota_x, E)\}_{x\in X}$ is obviously coherent (each ι_x is the empty function).

Lemma 3.12. *Suppose that Θ is a totally ordered set and that $\{\mathcal{T}_\theta\}_{\theta\in\Theta}$ is a collection of coherent collections of coi triples from* Red_{κ_0} *to* Red_{κ_1} *such that $\theta \le \theta'$ implies $\mathcal{T}_\theta \subseteq \mathcal{T}_{\theta'}$. Then $\bigcup_{\theta\in\Theta} \mathcal{T}_\theta$ is coherent.*

Proof. Supposing that $\mathrm{coi}(W_{x_0}, \iota_{x_0}, U_{x_0}), \mathrm{coi}(W_{x_1}, \iota_{x_1}, U_{x_1}) \in \bigcup_{\theta\in\Theta} \mathcal{T}_\theta$ and intervals $I_0 \subseteq \mathrm{p}*(W_{x_0})$ and $I_1 \subseteq \mathrm{p}*(W_{x_1})$ and $i \in \{-1, 1\}$ are such that $W_{x_0} \restriction_{\mathrm{p}} I_0 \equiv (W_{x_1} \restriction_{\mathrm{p}} I_1)^i$, we select $\theta \in \Theta$ such that $\mathrm{coi}(W_{x_0}, \iota_{x_0}, U_{x_0}), \mathrm{coi}(W_{x_1}, \iota_{x_1}, U_{x_1}) \in \mathcal{T}_\theta$. As $\mathcal{T}_\theta$ is coherent we get

$$[[U_{x_0} \restriction_{\mathrm{p}} \propto(I_0, \iota_{x_0})]] = [[(U_{x_1} \restriction_{\mathrm{p}} \propto(I_1, \iota_{x_1}))^i]].$$

The comparable check for words $U_{x_2}, U_{x_3} \in \mathrm{Red}_{\kappa_1}$ is analogous. □

Lemma 3.13. *Suppose* $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ *is coherent*, $x \in X$, $I \subseteq \mathrm{p}*(W_x)$ *is an interval*, $I \equiv I_0 I_1 \cdots I_n$. *Suppose also that for each* $0 \le j \le n$ *we have an* $x_j \in X$, *an interval* I'_j *in* $\mathrm{p}*(W_{x_j})$ *and* $i_j \in \{-1, 1\}$ *such that* $W_x \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} I'_j)^{i_j}$. *Then*

$$[[U_x \restriction_{\mathrm{p}} \propto(I, \iota_x)]] = \prod_{j=0}^{n} [[(U_{x_j} \restriction_{\mathrm{p}} \propto(I'_j, \iota_{x_j}))^{i_j}]].$$

Furthermore, *if* $L = \{0 \le j \le n \mid |I_j| > 1\}$ *we have*

$$[[U_x \restriction_{\mathrm{p}} \propto(I, \iota_x)]] = \prod_{j\in L} [[(U_{x_j} \restriction_{\mathrm{p}} \propto(I'_j, \iota_{x_j}))^{i_j}]].$$

Proof. For each $0 \le j \le n$ we have $W_x \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} I'_j)^{i_j}$, so that by the fact that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ is coherent we see that

$$[[U_x \restriction_{\mathrm{p}} \propto(I_j, \iota_x)]] = [[(U_{x_j} \restriction_{\mathrm{p}} \propto(I'_j, \iota_{x_j}))^{i_j}]]$$

for all $0 \le j \le n$. In particular we have

$$\prod_{j=0}^{n} [[U_x \restriction_{\mathrm{p}} \propto(I'_j, \iota_x)]] = \prod_{j=0}^{n} [[(U_{x_j} \restriction_{\mathrm{p}} \propto(I'_j, \iota_{x_j}))^{i_j}]]$$

and so we will be done with the first claim if we show that $[[U_x \restriction_{\mathrm{p}} \propto(I, \iota_x)]] = \prod_{j=0}^{n}[[U_x \restriction_{\mathrm{p}} \propto(I_j, \iota_x)]]$. But this is true since by Lemma 3.9 the (possibly unreduced) word $\prod_{j=0}^{n} U_x \restriction_{\mathrm{p}} \propto(I_j, \iota_x)$ is obtained from $U_x \restriction_{\mathrm{p}} \propto(I, \iota_x)$ by deleting finitely many pure subwords.

Next we let L be as in the statement of the lemma. Notice that for each $0 \leq j \leq n$ with $j \notin L$ we have $|I_j| = |I'_j| \leq 1$ and so $\propto(I'_j, \iota_{x_j})$ is a finite interval, by Lemma 3.10. Thus for each such j we have $[[(U_{x_j} \restriction_{\mathrm{p}} \propto(I'_j, \iota_{x_j}))^{i_j}]] = [[E]]$ since $U_{x_j} \restriction_{\mathrm{p}} \propto(I'_j, \iota_{x_j})$ is a finite concatenation of pure words. Thus removing all such j from the multiplication expression $\prod_{j=0}^{n}[[(U_{x_j} \restriction_{\mathrm{p}} \propto(I'_j, \iota_{x_j}))^{i_j}]]$ will not change the value in the group, and so we are done with the second claim. $\square$

What follows is a rather technical result that will allow us to conclude that certain natural maps are well defined despite certain choices that are made.

Lemma 3.14. *Let the collection* $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ *be coherent and let* W *be in* $\mathrm{Pfine}(\{W_x\}_{x\in X})$. *Let* $I_0, \dots, I_n$ *be a finite set of subintervals of* $\mathrm{p}*(W)$ *as in the conclusion of Lemma 3.2 and let* $J = \{0 \leq j \leq n \mid |I_j| > 1\}$. *For each* $j \in J$ *select* $x_j \in X$, $i_j \in \{-1, 1\}$, *and an interval* $\Lambda_j \subseteq \mathrm{p}*(W_{x_j})$ *such that* $W \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda_j)^{i_j}$. *Again, let* $I'_0, \dots, I'_{n'}$ *be a finite set of subintervals of* $\mathrm{p}*(W)$ *as in the conclusion of Lemma 3.2 and let* $J' = \{0 \leq j' \leq n' \mid |I'_{j'}| > 1\}$. *For each* $j' \in J'$ *select* $y_{j'} \in X$, $m_{j'} \in \{-1, 1\}$, *and an interval* $\Lambda'_{j'} \subseteq \mathrm{p}*(W_{y_{j'}})$ *such that* $W \restriction_{\mathrm{p}} I'_{j'} \equiv (W_{y_{j'}} \restriction_{\mathrm{p}} \Lambda'_{j'})^{m_{j'}}$. *Then*

$$\prod_{j\in J}[[(U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{i_j}]] = \prod_{j'\in J'}[[(U_{y_{j'}} \restriction_{\mathrm{p}} \propto(\Lambda'_{j'}, \iota_{y_{j'}}))^{m_{j'}}]].$$

Proof. Assume the hypotheses. Take $\mathbb{I}$ to be the set of nonempty intervals obtained by intersecting an I_j with an $I'_{j'}$. For each $0 \leq j \leq n$ we can write $I_j \equiv I_{(j,0)} I_{(j,1)} \cdots I_{(j,n_j)}$ where each $I_{(j,q)}$ is an element of $\mathbb{I}$. Similarly for each $0 \leq j' \leq n'$ we write $I'_{j'} \equiv I'_{(j',0)} \cdots I'_{(j',n'_{j'})}$ where each $I'_{(j',r)}$ is an element of $\mathbb{I}$.

We have $\mathbb{I} = \{I_{(j,q)}\}_{0\leq j\leq n, 0\leq q\leq n_j} = \{I'_{(j',r)}\}_{0\leq j'\leq n', 0\leq r\leq n'_{j'}}$. Let $F : \mathbb{I} \to \{(j, q) \mid 0 \leq j \leq n, 0 \leq q \leq n_j\}$ be the unique order isomorphism between the domain and codomain where the codomain is given the lexicographic order, comparing the leftmost coordinate first and define $F' : \mathbb{I} \to \{(j', r) \mid 0 \leq j' \leq n', 0 \leq r \leq n'_{j'}\}$ similarly. Let $\mathrm{h} : \{(j, q) \mid 0 \leq j \leq n, 0 \leq q \leq n_j\} \to \{0, \dots, n\}$ denote projection to the first coordinate, and similarly define $\mathrm{h}' : \{(j', r) \mid 0 \leq j' \leq n', 0 \leq r \leq n'_{j'}\} \to \{0, \dots, n'\}$. Let $\mathbb{J} \subseteq \mathbb{I}$ denote the set of intervals in $\mathbb{I}$ which are of cardinality at least 2; that is,

$$\mathbb{J} = \{I_{(j,q)} \mid 0 \leq j \leq n, 0 \leq q \leq n_j, |I_{(j,q)}| \geq 2\}.$$

For each $j \in J$ and each $I_{(j,q)} \in \mathbb{J}$ we know that $W \restriction_{\mathrm{p}} I_{(j,q)} \in \text{p-chunk}(W_{x_j}^{i_j})$, so select an interval $\Lambda_{(j,q)} \subseteq \mathrm{p}*(W_{x_j})$ such that $W \restriction_{\mathrm{p}} I_{(j,q)} \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda_{(j,q)})^{i_j}$. Now we have that

$$\begin{aligned}
&\prod_{j\in J}[[(U_{x_j}\restriction_{\mathrm{p}} \propto(\Lambda_j,\iota_{x_j}))^{i_j}]]\\
&\quad=\prod_{j\in J}\prod_{\substack{0\le q\le n_j\\ I_{(j,q)}\in\mathbb{J}}}[[(U_{x_j}\restriction_{\mathrm{p}} \propto(\Lambda_{(j,q)},\iota_{x_j}))^{i_j}]]\\
&\quad=\prod_{I\in\mathbb{J}}[[(U_{x_{\mathrm{h}\circ F(I)}}\restriction_{\mathrm{p}} \propto(\Lambda_{F(I)},\iota_{x_{\mathrm{h}\circ F(I)}}))^{i_{\mathrm{h}\circ F(I)}}]]\\
&\quad=\prod_{j'\in J'}\prod_{\substack{0\le r\le n'_{j'}\\ I_{(j',r)}\in\mathbb{J}}}[[(U_{x_{\mathrm{h}\circ F\circ(F')^{-1}(j',r))}}\restriction_{\mathrm{p}} \propto(\Lambda_{F\circ(F')^{-1}(j',r)},\iota_{x_{\mathrm{h}\circ F\circ(F')^{-1}(j',r)}}))^{i_{\mathrm{h}\circ F\circ(F')^{-1}(j',r)}}]]\\
&\quad=\prod_{j'\in J'}[[(U_{y_{j'}}\restriction_{\mathrm{p}} \propto(\Lambda'_{j'},\iota_{y_{j'}}))^{m_{j'}}]]
\end{aligned}$$

where the first equality holds by Lemma 3.13, the second and third equalities are simply a rewriting of the order index, and the last equality holds by another application of Lemma 3.13. This completes the proof. □

Now we may conclude that a coherent collection of cois produces well-defined homomorphisms. For each $i \in \{0, 1\}$ we let $\sqsupset_{\kappa_i} : \mathrm{Red}_{\kappa_i} \to \mathcal{C}_{\kappa_i}$ denote the surjection given by $W \mapsto [[W]]$.

Proposition 3.15. *Let* $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ *be coherent. By selecting for each* $W \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ *a finite set of subintervals* $I_0, \dots, I_n$ *of* $\mathrm{p}_*(W)$ *as in the conclusion of Lemma 3.2, letting* $J = \{0 \le j \le n \mid |I_j| > 1\}$, *selecting for each* $j \in J$ *an element* $x_j \in X$, $i_j \in \{-1, 1\}$, *and an interval* $\Lambda_j \subseteq \mathrm{p}_*(W_{x_j})$ *such that* $W \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda_j)^{i_j}$ *we obtain a homomorphism*

$$\phi_0 : \mathrm{Pfine}(\{W_x\}_{x\in X}) \to \sqsupset_{\kappa_1}(\mathrm{Pfine}(\{U_x\}_{x\in X}))$$

given by $\phi_0(W) = \prod_{j\in J}[[(U_{x_j}\restriction_{\mathrm{p}} \propto(\Lambda_j,\iota_{x_j}))^{i_j}]]$, *whose definition is independent of the choices made of the set of subintervals* $I_0, \dots, I_n$, *elements* $x_j \in X$ *and* $i_j \in \{-1, 1\}$, *and intervals* $\Lambda_j \subseteq \mathrm{p}_*(W_{x_j})$. *The comparable map*

$$\phi_1 : \mathrm{Pfine}(\{U_x\}_{x\in X}) \to \sqsupset_{\kappa_0}(\mathrm{Pfine}(\{W_x\}_{x\in X}))$$

similarly is a homomorphism whose definition is independent of the various selections made.

Proof. From Lemma 3.14 we see that the described function ϕ_0 is well defined and independent of the numerous choices made. We must check that ϕ_0 is a homomorphism.

We note first that if $W \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ and $\mathrm{p}_*(W)$ has a first or last element, say $\lambda = \max(\mathrm{p}_*(W))$, then $\phi_0(W) = \phi_0(W \restriction_{\mathrm{p}} \mathrm{p}_*(W) \setminus \{\lambda\})$. This is easily seen by selecting the set of intervals $I_0, \ldots, I_n$ for W to be such that $I_n = \{\lambda\}$. The fact that $|I_n| = 1$ and therefore $I_n \notin J$ completes the argument.

Suppose that $W \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ and $W \equiv W_0 W_1$ where also both $W_0, W_1 \in \mathrm{Pfine}(\{W_x\}_{x\in X})$. Choose subintervals $I_0, \ldots, I_n$ in $\mathrm{p}_*(W_0)$ as in Lemma 3.2, let $J = \{0 \le j \le n \mid |I_j| > 1\}$, select $x_j \in X$ and $i_j \in \{-1, 1\}$ and intervals $\Lambda_j \subseteq \mathrm{p}_*(W_{x_j})$ with $W \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda_j)^{i_j}$. Similarly choose intervals $I'_0, \ldots, I'_{n'}$ in $\mathrm{p}_*(W_1)$ and define J' and choose $y_{j'} \in X$, $m_{j'} \in \{-1, 1\}$ and $\Lambda'_{j'} \subseteq \mathrm{p}_*(W_{y_{j'}})$ for each $j' \in J'$. Notice that $\mathrm{p}_*(W) \equiv I_0 \cdots I_n I'_0 \cdots I'_{n'}$ is a decomposition as in Lemma 3.2 and $J \cup J'$ is precisely the set of indices whose accompanying interval is of cardinality at least two. Then

$$\begin{aligned}\phi_0(W) &= \left(\prod_{j\in J} [[(U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{i_j}]]\right)\left(\prod_{j'\in J'} [[(U_{y_{j'}} \restriction_{\mathrm{p}} \propto(\Lambda_{j'}, \iota_{x_j}))^{m_j}]]\right)\\ &= \phi_0(W_0)\phi_0(W_1).\end{aligned}$$

Next we suppose that $W \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ and let subintervals $I_0, \ldots, I_n$ in $\mathrm{p}_*(W_0)$ be as in Lemma 3.2, let $J = \{0 \le j \le n \mid |I_j| > 1\}$, select $x_j \in X$ and $i_j \in \{-1, 1\}$ and intervals $\Lambda_j \subseteq \mathrm{p}_*(W_{x_j})$ with $W \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda_j)^{i_j}$. Notice that $\mathrm{p}_*(W^{-1})$ may be written as $\mathrm{p}_*(W^{-1}) \equiv I'_n \cdots I'_0$ as in Lemma 3.2, where I'_j is order isomorphic to the ordered set $(I_j)^{-1}$, and $W \restriction_{\mathrm{p}} I_j \equiv (W^{-1} \restriction_{\mathrm{p}} I'_j)^{-1}$. Also, $\{0 \le j \le n \mid |I'_j| > 1\}$ is equal to the set J. Then

$$\begin{aligned}\phi_0(W) &= \prod_{j\in J} [[(U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{i_j}]] = \left(\prod_{j\in J^{-1}} [[(U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{-i_j}]]\right)^{-1}\\ &= (\phi_0(W^{-1}))^{-1},\end{aligned}$$

where J^{-1} denotes the set J under the reverse order. Thus $\phi_0(W^{-1}) = (\phi_0(W))^{-1}$.

Finally we let $W_0, W_1 \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ be given. As in Lemma 2.1 we write $W_0 \equiv W_{00}W_{01}$ and $W_1 \equiv W_{10}W_{11}$ with $W_{01} \equiv W_{10}^{-1}$ and the word $W_{00}W_{11}$ reduced. We will give the argument in the most difficult case and sketch how the argument goes in the less difficult ones. Suppose that W_{00} ends with a nonempty α-pure word and W_{11} begins with a nonempty α-pure word, and also that W_{01} begins with a nonempty α-pure word. From this last assumption we know that W_{10} ends with a nonempty α-pure word.

We have $W_{00}W_{11} \equiv W'_{00}W_a W'_{11}$ where we denote $\lambda_0 = \max(\mathrm{p}_*(W_{00}))$, $\lambda_1 = \min(\mathrm{p}_*(W_{11}))$ and

$$W'_{00} \equiv W_{00} \restriction_{\mathrm{p}} \{\lambda \in \mathrm{p}_*(W_{00}) \mid \lambda < \lambda_0\}, \quad W'_{11} \equiv W_1 \restriction_{\mathrm{p}} \{\lambda \in \mathrm{p}_*(W_{11}) \mid \lambda > \lambda_1\},$$
$$W_a \equiv (W_{00} \restriction_{\mathrm{p}} \{\lambda_0\})(W_{11} \restriction_{\mathrm{p}} \{\lambda_1\}).$$

Note that $W'_{00}, W_a, W'_{11} \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ since the concatenation $W_{00}W_{11}$ is in $\mathrm{Pfine}(\{W_x\}_{x\in X})$ and each of W'_{00}, W_a, and W'_{11} are p-chunks of this word, whereas for example $W_{00} \restriction_{\mathrm{p}} \{\lambda_0\}$ might not be in $\mathrm{Pfine}(\{W_x\}_{x\in X})$. Furthermore suppose $\lambda_2 = \min(\mathrm{p}_*(W_{01}))$ and $\lambda_3 = \max(\mathrm{p}_*(W_{10}))$ and define

$$W'_{01} \equiv W_{01} \restriction_{\mathrm{p}} (\mathrm{p}_*(W_{01}) \setminus \{\lambda_2\}), \quad W_b \equiv (W_{00} \restriction_{\mathrm{p}} \{\lambda_0\})(W_{01} \restriction_{\mathrm{p}} \{\lambda_2\}),$$
$$W'_{10} \equiv W_{10} \restriction_{\mathrm{p}} (\mathrm{p}_*(W_{10}) \setminus \{\lambda_3\}), \quad W_c \equiv (W_{10} \restriction_{\mathrm{p}} \{\lambda_3\})(W_{11} \restriction_{\mathrm{p}} \{\lambda_1\}).$$

Notice that $W'_{01} \equiv (W'_{10})^{-1}$ and that each of the words $W'_{01}, W_b, W'_{10}, W_c$ is in $\mathrm{Pfine}(\{W_x\}_{x\in X})$.

By our work so far we get

$$\begin{aligned}
\phi_0(W_{00}W_{11}) &= \phi_0(W'_{00}W_aW'_{11}) \\
&= \phi_0(W'_{00})\phi_0(W_a)\phi_0(W'_{11}) \\
&= \phi_0(W'_{00})\phi_0(W'_{11}) \\
&= \phi_0(W'_{00})\phi_0(W'_{01})\phi_0(W'_{10})\phi_0(W'_{11}) \\
&= \phi_0(W'_{00})\phi_0(W_b)\phi_0(W'_{01})\phi_0(W'_{10})\phi_0(W_c)\phi_0(W'_{11}) \\
&= \phi_0(W_0)\phi_0(W_1).
\end{aligned}$$

In the simpler case where W_{01} does not begin with an α-pure word (hence W_{10} does not end with an α-pure word) we let $W'_{01} = W_{01}$, $W'_{10} = W_{10}$ and both W_b and W_c be the empty word and the equalities above will all hold. In the case there does not exist $\alpha < \kappa_0$ such that both W_{00} ends with a nonempty α-pure word and W_{11} begins with an α-pure word we let $W'_{00} \equiv W_{00}$, $W'_{11} \equiv W_{11}$ and $W_a \equiv E$. It may still be the case that W_{00} ends with a nonempty β-pure word and W_{01} begins with a nonempty β-pure word, $\beta < \kappa_0$, and for this we define

$$W'_{01} \equiv W_{01} \restriction_{\mathrm{p}} (\mathrm{p}_*(W_{01}) \setminus \{\lambda_2\}), \quad W_b \equiv (W_{00} \restriction_{\mathrm{p}} \{\lambda_0\})(W_{01} \restriction_{\mathrm{p}} \{\lambda_2\}),$$
$$W'_{10} \equiv W_{10} \restriction_{\mathrm{p}} (\mathrm{p}_*(W_{10}) \setminus \{\lambda_3\}),$$

and let W_c be given by

$$\begin{cases} (W_{10} \restriction_{\mathrm{p}} \{\lambda_3\})(W_{11} \restriction_{\mathrm{p}} \{\lambda_1\}) & \text{in case } W_{11} \text{ begins with a nonempty} \\ & \beta\text{-pure word and } \lambda_3 = \min \mathrm{p}_*(W_{11}); \\ W_{10} \restriction_{\mathrm{p}} \{\lambda_3\} & \text{otherwise.} \end{cases}$$

The case where W_{11} and W_{10} respectively begin and end with a β-pure word, for some $\beta < \kappa_0$, is analogous. If none of these cases holds then we simply let $W'_{00} \equiv W_{00}$, $W'_{01} \equiv W_{01}$, $W'_{10} \equiv W_{10}$, $W'_{11} \equiv W_{11}$ and $W_a \equiv W_b \equiv W_c \equiv E$. This exhausts all possibilities and the proof is complete (the arguments for ϕ_1 are made in the analogous way). □

Proposition 3.16. *The homomorphisms ϕ_0 and ϕ_1 descend respectively to isomorphisms*

$$\Phi_0 : \beth_0(\mathrm{Pfine}(\{W_x\}_{x\in X})) \to \beth_1(\mathrm{Pfine}(\{U_x\}_{x\in X})),$$
$$\Phi_1 : \beth_1(\mathrm{Pfine}(\{U_x\}_{x\in X})) \to \beth_0(\mathrm{Pfine}(\{W_x\}_{x\in X})),$$

with $\Phi_0 = \Phi_1^{-1}$.

Proof. If $W \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ is a pure word the set $\mathrm{p}*(W)$ is a singleton and for any decomposition of $\mathrm{p}*(W)$ by Lemma 3.2 the accompanying set J will necessarily be empty. Thus all pure words in $\mathrm{Pfine}(\{W_x\}_{x\in X})$ are in $\ker(\phi_0)$ and so we get the induced Φ_0, and similarly we obtain an induced Φ_0.

By Lemma 3.2 each element of the group $\beth_0(\mathrm{Pfine}(\{W_x\}_{x\in X}))$ may be written as a product $[[W_0]][[W_1]]\cdots[[W_n]]$ where each W_i is in $\big(\bigcup_{x\in X} \text{p-chunk}(W_x)\big)^{\pm 1}$. For each $0 \leq j \leq n$ we select x_j and i_j and an interval $\Lambda_j \subseteq \mathrm{p}*(W_{x_j})$ such that $W_j \equiv (W_{x_j} \upharpoonright_{\mathrm{p}} \Lambda_j)^{i_j}$. Now

$$\begin{aligned}\Phi_1 \circ \Phi_0([[W_0]]\cdots[[W_n]]) &= \prod_{j=0}^{n} \Phi_1[[(U_{x_j} \upharpoonright_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{i_j}]]\\ &= \prod_{j=0}^{n} (\Phi_1([[U_{x_j} \upharpoonright_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j})]]))^{i_j}\\ &= \prod_{j=0}^{n} [[W_{x_j} \upharpoonright_{\mathrm{p}} \propto(\propto(\Lambda_j, \iota_{x_j}), \iota_{x_j}^{-1})]]^{i_j}\\ &= \prod_{j=0}^{n} [[W_{x_j} \upharpoonright_{\mathrm{p}} \Lambda_j]]^{i_j}\\ &= \prod_{j=0}^{n} [[W_j]],\end{aligned}$$

where the fourth equality holds by Lemma 3.8 — the word $W_{x_j} \upharpoonright_{\mathrm{p}} \propto(\propto(\Lambda_j, \iota_{x_j}), \iota_{x_j}^{-1})$ is obtained from the word $W_{x_j} \upharpoonright_{\mathrm{p}} \Lambda_j$ by deleting finitely many pure subwords, namely those associated with the set $\Lambda_j \setminus \propto(\propto(\Lambda_j, \iota_{x_j}), \iota_{x_j}^{-1})$. Thus $\Phi_1 \circ \Phi_0$ is the identity map, and that $\Phi_0 \circ \Phi_1$ is also the identity map follows from the same reasoning. □

3D. *Extensions of coherent collections.* By Proposition 3.16, the problem of finding an isomorphism between cone groups is reduced to that of finding a coherent collection of coi triples $\{\mathrm{coi}(W_x, U_x, \iota_x)\}_{x\in X}$ such that $\beth_0(\mathrm{Pfine}(\{W_x\}_{x\in X})) = \mathcal{C}_{\kappa_0}$ and $\beth_1(\mathrm{Pfine}(\{U_x\}_{x\in X})) = \mathcal{C}_{\kappa_1}$. Thus, in this and all remaining subsections we approach the problem of extending collections of coi triples. We still assume that $\kappa_0, \kappa_1 \geq 2$ and that the coi collections are from Red_{κ_0} to Red_{κ_1}.

Lemma 3.17. *Let* $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ *be coherent. If* W *is in* $\mathrm{Pfine}(\{W_x\}_{x\in X})$ *then there exists* $U \in \mathrm{Red}_{\kappa_1}$ *and a coi* ι *from* W *to* U *such that* $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{(W, \iota, U)\}$ *is coherent. Moreover if* W *is nonempty the domain* (*and range*) *of* ι *can be made to be nonempty.*

Proof. If W is empty then we let U and ι be empty. Else we choose subintervals $I_0, \ldots, I_n$ in $\mathrm{p}*(W)$ as in Lemma 3.2, let $J = \{0 \leq j \leq n \mid |I_j| > 1\}$, select $x_j \in X$ and

$i_j \in \{-1, 1\}$ and intervals $\Lambda_j \subseteq \mathrm{p}_*(W_{x_j})$ with $W \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda_j)^{i_j}$. Let $J' \subseteq J$ be given by

$$J' = \{j \in J \mid (U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{i_j} \not\equiv E\}.$$

For each $j \in J'$ let $U'_j \equiv (U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{i_j}$. For every $0 \leq j \leq n$ with $j \notin J'$ we let $U'_j \equiv a_{0,0}$.

The word $\prod_{j=0}^{n} U'_j$ is probably not reduced, and so we will make slight modifications in order to obtain a reduced word. We know that each subword U'_j is reduced and nonempty. Let $U_n \equiv U'_n$. Let $0 \leq j < n$ be given. There are a few possibilities:

- $\mathrm{p}_*(U'_j)$ has a maximal element and $\mathrm{p}_*(U'_{j+1})$ has a minimal element and both $U'_j \restriction_{\mathrm{p}} \{\max \mathrm{p}_*(U'_j)\}$ and $U'_{j+1} \restriction_{\mathrm{p}} \{\min \mathrm{p}_*(U'_{j+1})\}$ are α-pure for some $\alpha < \kappa_1$.
- $\mathrm{p}_*(U'_j)$ has a maximal element and $\mathrm{p}_*(U'_{j+1})$ has a minimal element and both $U'_j \restriction_{\mathrm{p}} \{\max \mathrm{p}_*(U'_j)\}$ and $U'_{j+1} \restriction_{\mathrm{p}} \{\min \mathrm{p}_*(U'_{j+1})\}$ are not α-pure for some $\alpha < \kappa_1$.
- $\mathrm{p}_*(U'_j)$ does not have a maximal element or $\mathrm{p}_*(U'_{j+1})$ does not have a minimal element.

In the middle case we let $U_j \equiv U'_j$. In the first or last case we choose $\alpha'_j < \kappa_1$ such that U'_j does not end with an α'_j-pure word (here we are using the fact that $\kappa_1 \geq 2$) and let $U_j \equiv U'_j a_{\alpha',0}$. The word $U_j U'_{j+1}$ is reduced, and so the word $U_j U_{j+1}$ is reduced (since U_{j+1} is nonempty), and so the word $U \equiv \prod_{j=0}^{n} U_j$ is reduced.

We now define the coi ι from W to U in a very natural way. If $j \in J'$ then we let the domain of ι_{x_j} be Λ'_j, and so $\mathrm{Close}(\Lambda'_j, \mathrm{p}_*(W_{x_j}))$. Let $\Lambda''_j \subseteq I_j$ be the image of $\Lambda'_j \cap \Lambda_j$ under the order isomorphism given by $W \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda_j)^{i_j}$. Similarly we let $\Theta''_j \subseteq \mathrm{p}_*(U'_j) \subseteq \mathrm{p}_*(U_j)$ be the image of $\iota_{x_j}(\Lambda_j \cap \Lambda'_j)$ under the order isomorphism given by $U'_j \equiv (U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{i_j}$. Define $\iota_j : \Lambda''_j \to \Theta''_j$ to be the order isomorphism given by the restriction to Λ''_j of the composition of the order isomorphism given by $W \restriction_{\mathrm{p}} I_j \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda_j)^{i_j}$ with ι_{x_j} with the order isomorphism given by $(U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j, \iota_{x_j}))^{i_j} \equiv U'_j$. It is easy to check that $\mathrm{Close}(\Lambda''_j, I_j)$ and $\mathrm{Close}(\Theta''_j, \mathrm{p}_*(U_j))$, since for $0 \leq j \leq n$ either $U_j \equiv U'_j$ or U_j is obtained from U'_j by appending a word of length one on the right.

If $0 \leq j \leq n$ and $j \notin J'$ then I_j is finite and nonempty, as is $\mathrm{p}_*(U_j)$, and we simply select elements $\lambda \in I_j$ and $\lambda' \in \mathrm{p}_*(U_j)$ and let $\Lambda''_j = \{\lambda\}$, $\Theta''_j = \{\lambda'_j\}$ and $\iota_j : \Lambda''_j \to \Theta''_j$ be the unique function. Clearly $\mathrm{Close}(\Lambda''_j, I_j)$ and $\mathrm{Close}(\Theta''_j, \mathrm{p}_*(U_j))$.

Let $\Lambda'' = \bigcup_{j=0}^{n} \Lambda''_j$ and $\Theta'' = \bigcup_{j=0}^{n} \Theta''_j$, and note that $\mathrm{Close}(\Lambda'', \mathrm{p}_*(W))$ and $\mathrm{Close}(\Theta'', \mathrm{p}_*(U))$ by Lemma 3.6 (iii). Let $\iota : \Lambda'' \to \Theta''$ be the unique extension of the ι_j. Now $\mathrm{coi}(W, \iota, U)$.

We check that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x \in X} \cup \{\mathrm{coi}(W, \iota, U)\}$ is coherent. Suppose that $y \in X$ and intervals $I \subseteq \text{p-chunk}(W)$ and $I' \subseteq \text{p-chunk}(W_y)$ and $i \in \{-1, 1\}$ are such that $W \restriction_{\mathrm{p}} I \equiv (W_y \restriction_{\mathrm{p}} I')^i$. Let $L \subseteq \{0, \ldots, n\}$ denote the set of those j such that $I_j \cap I \neq \varnothing$. For each $j \in L \cap J$ we have $W \restriction_{\mathrm{p}} (I_j \cap I) \equiv (W_{x_j} \restriction_{\mathrm{p}} \Lambda^*_j)^{i_j}$ for the obvious choice of interval $\Lambda_j^* \subseteq \Lambda_j \subseteq \text{p-chunk}(W_{x_j})$. Thus $(W_{x_j} \restriction_{\mathrm{p}} \Lambda_j^*)^{i \cdot i_j} \equiv W_y \restriction_{\mathrm{p}} I_j'$ for the obvious choice of interval $I_j' \subseteq I'$. By the coherence of $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x \in X}$ we therefore have

$$\begin{aligned}
[[U \restriction_{\mathrm{p}} \propto(I, \iota)]] &= \prod_{j \in L} [[U \restriction_{\mathrm{p}} \propto(I_j \cap I, \iota)]] \\
&= \prod_{j \in L \cap J'} [[U \restriction_{\mathrm{p}} \propto(I_j \cap I, \iota)]] \\
&= \prod_{j \in L \cap J'} [[U_{x_j} \restriction_{\mathrm{p}} \propto(\Lambda_j^*, \iota_{x_j})]]^{i_j} \\
&= \prod_{j \in (L \cap J')^i} [[U_y \restriction_{\mathrm{p}} \propto(I_j', \iota_y)]]^i \\
&= [[(U_y \restriction_{\mathrm{p}} \propto(I', \iota_y))^i]].
\end{aligned}$$

If we select intervals $I, I' \subseteq \mathrm{p}_*(W)$ and $i \in \{-1, 1\}$ such that $W \restriction_{\mathrm{p}} I \equiv (W \restriction_{\mathrm{p}} I')^i$ then a similar strategy of finitely decomposing I and I' is employed to show $[[U \restriction \propto(I, \iota)]] = [[(U \restriction_{\mathrm{p}} \propto(I', \iota))^i]]$.

With slight modifications, we check in a similar way that if $U \restriction_{\mathrm{p}} Q \equiv (U_z \restriction_{\mathrm{p}} Q')^i$, with $z \in X$, then the appropriate elements of $\mathcal{C}_{\kappa_0}$ are equal. Suppose $z \in X$, $i \in \{-1, 1\}$, and intervals $Q \subseteq \mathrm{p}_*(U)$ and $Q' \subseteq \mathrm{p}_*(U_z)$ are such that $U \restriction_{\mathrm{p}} Q \equiv (U_z \restriction_{\mathrm{p}} Q')^i$. By construction we know that $\mathrm{p}_*(U_j')$ is an initial interval in $\mathrm{p}_*(U_j)$, with $\mathrm{p}_*(U_j) \setminus \mathrm{p}_*(U_j')$ being of cardinality at most 1. Also, $\mathrm{p}_*(U) \equiv \mathrm{p}_*(U_0) \cdots \mathrm{p}_*(U_n)$. Let $T \subseteq \{0, \ldots, n\}$ be the set of those j such that $\mathrm{p}_*(U_j) \cap Q \neq \varnothing$. For each $j \in T \cap J'$ we have $U \restriction_{\mathrm{p}} \mathrm{p}_*(U_j') \cap Q \equiv (U_{x_j} \restriction_{\mathrm{p}} \Theta_j^*)^{i_j}$ for the obvious interval $\Theta_j^* \subseteq \text{p-chunk}(U_{x_j})$, and $(U_{x_j} \restriction_{\mathrm{p}} \Theta_j^*)^{i \cdot i_j} \equiv U_z \restriction_{\mathrm{p}} Q_j'$ for an appropriate $Q_j' \subseteq \mathrm{p}_*(U_z)$. We see that

$$\begin{aligned}
[[W \restriction_{\mathrm{p}} \propto(Q, \iota^{-1})]] &= \prod_{j \in T} [[W \restriction_{\mathrm{p}} \propto(\mathrm{p}_*(U_j) \cap Q, \iota^{-1})]] \\
&= \prod_{j \in L \cap J'} [[W \restriction_{\mathrm{p}} \propto(\mathrm{p}_*(U_j') \cap Q, \iota)]] \\
&= \prod_{j \in L \cap J'} [[W_{x_j} \restriction_{\mathrm{p}} \propto(\Theta_j^*, (\iota_{x_j})^{-1})]]^{i_j} \\
&= \prod_{j \in (L \cap J')^i} [[W_z \restriction_{\mathrm{p}} \propto(Q_j', \iota_y)]]^i \\
&= [[(W_z \restriction_{\mathrm{p}} \propto(Q', \iota_y))^i]].
\end{aligned}$$

Similar modifications are enacted if $Q, Q' \subseteq \mathrm{p}_*(U)$, and the proof is complete. □

We introduce some extra notation for convenience. For a not necessarily reduced word W we let

$$\|W\| = \sup\left\{\frac{1}{n+1} \mid n = \mathrm{proj}_1(W(i)) \text{ for some } i \in \overline{W}\right\}$$

where the supremum is considered in the set of nonnegative reals. As examples we have $\|E\| = 0$ and $\|a_{\alpha,5}^{-1}a_{\alpha',10}\| = \frac{1}{6}$. By comparison to earlier notation, we have $d(W) = 1/\|W\| - 1$.

Lemma 3.18. *Suppose that κ_0 and κ_1 are cardinal numbers greater than or equal to 2. Suppose that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ is coherent, $z \in X$ and that $\epsilon > 0$ is a real number. Then there exists $U \in \mathrm{Red}_{\kappa_1}$ with $\|U\| < \epsilon$ and a coi ι from W_z to U such that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_z, \iota, U)\}$ is coherent. Moreover the domain (and codomain) of ι may be chosen to be nonempty provided ι_z satisfies this property.*

Proof. If W_z is empty then let U be empty and $\iota = \varnothing$. Otherwise let $U_z \equiv_{\mathrm{p}} \prod_{\lambda\in \mathrm{p}_*(U_z)} U_\lambda$ and $J = \{\lambda \in \mathrm{p}_*(U_z) \mid \|U_\lambda\| \geq \epsilon\}$. Since U_z is a word, we know that J is finite. Let $N \in \omega$ be large enough that $\frac{1}{N+1} < \epsilon$. For each $\lambda \in \mathrm{p}_*(U_z)$ we let

$$U'_\lambda \equiv \begin{cases} U_\lambda & \text{if } \lambda \notin J, \\ a_{\alpha,N} & \text{if } \lambda \in J \text{ and } U_\lambda \text{ is } \alpha\text{-pure.} \end{cases}$$

We let $U \equiv \prod_{\lambda\in \mathrm{p}_*(U_x)} U'_\lambda$. It is easy to see that U is reduced (a cancellation in U would necessarily include the pairing of a letter $a_{\alpha,N} \equiv U_\lambda$, with $\lambda \in J$, with a letter in $U'_{\lambda'}$ where λ' is the immediate successor or immediate predecessor of λ in $\mathrm{p}_*(U_x)$, and thus U'_λ and $U'_{\lambda'}$ are both α-pure, so U_λ and $U_{\lambda'}$ are as well, a contradiction). Moreover $U \equiv_{\mathrm{p}} \prod_{\lambda\in \mathrm{p}_*(U_x)} U'_\lambda$ and clearly $\|U\| < \epsilon$. Letting $\iota = \iota_z$ it is immediate that ι is a coi from W_z to U. The rather intuitive fact that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_z, \iota, U)\}$ is coherent is proved along similar lines used in earlier proofs. □

Lemma 3.19. *Suppose that $\kappa_1 \geq 2$ and that $|X| < 2^{\aleph_0}$. Given $N \in \omega \setminus \{0\}$ and an ordinal $\alpha < \kappa_1$ there exists an α-pure word $U \in \mathrm{Red}_{\kappa_1}$ using only positive letters such that $\|U\| = \frac{1}{N}$, and $U(\max(\overline{U})) = a_{\alpha,N-1} = U(\min(\overline{U}))$, and $U \notin \mathrm{Pfine}(\{U_x\}_{x\in X})$.*

Proof. Assume the hypotheses. We will let $\overline{U} = [0, 1] \cap \mathbb{Q}$. It is easy to see that the set of all functions $f : ([0, 1] \cap \mathbb{Q}) \to \{a_{\alpha,n}\}_{n\geq N-1}$ such that $f(0) = f(1) = a_{\alpha,N-1}$ and the restriction $f \upharpoonright (0, 1) \cap \mathbb{Q}$ is injective is of cardinality $2^{\aleph_0}$, and each such function is an element of Red_{κ_1} since there are no inverse letters with which to perform a cancellation. On the other hand we have by Lemma 3.4 that there are less than $2^{\aleph_0}$ pure elements in $\mathrm{Pfine}(\{U_x\}_{x\in X})$. The lemma follows immediately. □

3E. *ω-type concatenations.* In this subsection we prove the following:

Proposition 3.20. *Suppose that κ_0 and κ_1 are cardinal numbers greater than or equal to 2. Suppose that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ is coherent, that W is reduced, that $\mathrm{p}_*(W) \equiv \prod_{n\in\omega} I_n$ with each $I_n \neq \varnothing$, $W \upharpoonright_{\mathrm{p}} I_n \in \mathrm{Pfine}(\{W_x\}_{x\in X})$, and $W \notin \mathrm{Pfine}(\{W_x\}_{x\in X})$. Suppose also that $|X| < 2^{\aleph_0}$. Then there exists $U \in \mathrm{Red}_{\kappa_1}$ and a coi ι from W to U such that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W, \iota, U)\}$ is coherent.*

Proof. For each $n \in \omega$ write $W_n \equiv W \upharpoonright_p I_n$. As $W_0 \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ is nontrivial we select a word $U_0 \in \mathrm{Red}_{\kappa_1}$ and coi ι_0 from W_0 to U_0 such that the domain of ι_0 is nonempty and such that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_0, \iota_0, U_0)\}$ is coherent, by Lemma 3.17. Assuming that the elements of $\{\mathrm{coi}(W_i, \iota_j, U_j)\}_{j\leq m}$ have already been chosen such that $\|U_j\| < \frac{1}{2}\|U_{j-1}\|$, each ι_j has nonempty domain and also that the union of collections $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_i, \iota_j, U_j)\}_{j\leq m}$ is coherent, we use Lemmas 3.17 and 3.18 to select $U_{m+1} \in \mathrm{Red}_{\kappa_1}$ and coi ι_{m+1} from W_{m+1} to U_{m+1} so that ι_{m+1} has nonempty domain, $\|U_{m+1}\| < \frac{1}{2}\|U_m\|$ and $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_i, \iota_j, U_j)\}_{j\leq m+1}$ is coherent.

By Lemma 3.12, the collection $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_i, \iota_j, U_j)\}_{j\in\omega}$ is coherent. For each $j \in \omega$ we will construct a word $V_j \in \mathrm{Red}_{\kappa_1}$ with $1 \leq |\mathrm{p}_*(V_j)| \leq 2$. Select $\alpha_j < \kappa_1$ such that the word U_j does not end with an α_j-pure subword. This is possible since $\kappa_1 \geq 2$ and U_j can end in at most one pure subword (and might possibly not end in a pure subword). By Lemma 3.19 we select an α_j-pure word $V_j' \in \mathrm{Red}_{\kappa_1} \setminus \mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{U_i\}_{i\in\omega})$ which uses only positive letters such that $\|V_j'\| = \|U_j\|$ and $\overline{V_j'}$ has maximum and minimum elements and $V_j'(\max(\overline{V_j'})) = a_{\alpha_j, d(U_j)+1} = V_j'(\min(\overline{V_j'}))$. If U_{j+1} begins with an α_j-pure subword, then select $\alpha_j'' \in \kappa_1 \setminus \{\alpha_j\}$, and again by Lemma 3.19, select $V_j'' \in \mathrm{Red}_{\kappa_1} \setminus \mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{U_i\}_{i\in\omega})$ which uses only positive letters such that $\|V_j''\| = \|U_j\|$ and $\overline{V_j''}$ has maximum and minimum elements and $V_j''(\max(\overline{V_j''})) = a_{\alpha_j'', d(U_j)-1} = V_j''(\min(\overline{V_j''}))$ and V_j'' is α_j''-pure. If U_{j+1} does not begin with an α_j-pure subword then let $V_j'' = E$. Let $V_j = V_j' V_j''$.

We know for each $n \in \omega$ that U_n, V_n' and V_n'' are each reduced. By how V_n' was selected, we know that $U_n V_n'$ is reduced since any cancellation would need to pair letters in V_n' with those in U_n, and U_n does not end in an α_j-pure word. Similarly, $U_n V_n' V_n'' \equiv U_n V_n$ is reduced.

As $\|U_n V_n\| \leq \frac{1}{2^n}$ we know the expression $\prod_{n\in\omega} U_n V_n \equiv U_0 V_0 U_1 V_1 \cdots$ is a word. By construction each of the words $\prod_{n=0}^{m} U_n V_n$ is reduced, and therefore the word $U \equiv \prod_{n\in\omega} U_n V_n$ is reduced. We note as well that by how V_n' and V_n'' were chosen we can write $\mathrm{p}_*(U) \equiv \prod_{n\in\omega} \mathrm{p}_*(U_n)\, \mathrm{p}_*(V_n)$, and $1 \leq |\mathrm{p}_*(V_n)| \leq 2$. Let ι be the function $\iota = \bigcup_{j\in\omega} \iota_j$. By Lemma 3.6 (iii) the domain of ι is close in $\mathrm{p}_*(W)$ and the range of ι is close in U, and thus we may write $\mathrm{coi}(W, \iota, U)$. We will show that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_i, \iota_j, U_j)\}_{j\in\omega} \cup \{\mathrm{coi}(W, \iota, U)\}$ is coherent, from which it will immediately follow that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W, \iota, U)\}$ is coherent.

Suppose that $y \in X \cup \omega$, $\Lambda_0 \subseteq \mathrm{p}_*(W)$ and $\Lambda_1 \subseteq \mathrm{p}_*(W_y)$ are intervals and $i \in \{-1, 1\}$ are such that $W \upharpoonright_p \Lambda_0 \equiv (W_y \upharpoonright_p \Lambda_1)^i$. If the set $\{n \in \omega \mid I_n \cap \Lambda_0 \neq \varnothing\}$ is infinite, then by the fact that Λ_0 is an interval there exist $m \in \omega$ and intervals $I_m', I_m'' \subseteq I_m$, with I_m' possibly empty, such that $I_m \equiv I_m' I_m''$ and $\Lambda_0 \equiv I_m'' \prod_{n=m+1}^{\infty} I_n$. Certainly $(W_y \upharpoonright_p \Lambda_1)^i \in \mathrm{Pfine}(\{W_x\}_{x\in X} \cup \{W_n\}_{n\in\omega})$, and since $W_n \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ for each n

we have in fact that $\mathrm{Pfine}(\{W_x\}_{x\in X}\cup\{W_n\}_{n\in\omega})=\mathrm{Pfine}(\{W_x\}_{x\in X})$. Therefore we have $W\restriction_{\mathrm{p}}\Lambda_0\equiv(W_y\restriction_{\mathrm{p}}\Lambda_1)^i\in\mathrm{Pfine}(\{W_x\}_{x\in X})$. But also $\big(\prod_{n=0}^{m-1}W_n\big)W\restriction_{\mathrm{p}}I'_m\in\mathrm{Pfine}(\{W_x\}_{x\in X})$. Thus $W\equiv\big(\big(\prod_{n=0}^{m-1}W_n\big)W\restriction_{\mathrm{p}}I'_m\big)(W\restriction_{\mathrm{p}}\Lambda_0)\in\mathrm{Pfine}(\{W_x\}_{x\in X})$, contrary to the assumptions of our lemma.

Thus we suppose that $y\in X\cup\omega$, $\Lambda_0\subseteq\mathrm{p}_*(W)$ and $\Lambda_1\subseteq\mathrm{p}_*(W_y)$ are intervals and $i\in\{-1,1\}$ are such that $W\restriction_{\mathrm{p}}\Lambda_0\equiv(W_y\restriction_{\mathrm{p}}\Lambda_1)^i$ and know from this that the set $K=\{n\in\omega\mid I_n\cap\Lambda_0\neq\varnothing\}$ is finite. If $K=\varnothing$ then $\Lambda_0=\varnothing=\Lambda_1$ and $[[U\restriction_{\mathrm{p}}\propto(\Lambda_0,\iota)]]=[[E]]=[[(U_y\restriction_{\mathrm{p}}\propto(\Lambda_1,\iota_y))^i]]$. If K has cardinality 1 then we let $K=\{m\}$ and we can write $I_m\equiv I'_m\Lambda_0I''_m$ where either or both of I'_m and I''_m may be empty. Since $\{\mathrm{coi}(W_x,\iota_x,U_x)\}_{x\in X}\cup\{\mathrm{coi}(W_j,\iota_j,U_j)\}_{j\in\omega}$ is coherent, we have

$$[[U\restriction_{\mathrm{p}}\propto(\Lambda_0,\iota)]]=[[U_m\restriction_{\mathrm{p}}\propto(\Lambda,\iota_m)]]=[[(U_y\restriction_{\mathrm{p}}\propto(\Lambda_1,\iota_y))^i]].$$

If K is of cardinality at least 2 then we let m_a and m_b be respectively the minimal and maximal elements and write $I_{m_a}\equiv I'_{m_a}I''_{m_a}$, $I_{m_b}\equiv I'_{m_b}I''_{m_b}$ (where either or both of I'_{m_a} and I''_{m_b} may be empty) and $\Lambda_0\equiv I''_{m_a}I_{m_a+1}\cdots I_{m_b-1}I'_{m_b}$. As $W\restriction_{\mathrm{p}}\Lambda_0\equiv(W_y\restriction_{\mathrm{p}}\Lambda_1)^i$, there exist subintervals $J_0,\dots,J_{m_b-m_a}$ of Λ_1 such that $W\restriction_{\mathrm{p}}I_j\equiv(W_y\restriction_{\mathrm{p}}J_j-m_a)^i$ for $m_a<j<m_b$ and $W\restriction_{\mathrm{p}}I''_{m_a}\equiv(W_y\restriction_{\mathrm{p}}J_0)^i$ and $W\restriction_{\mathrm{p}}I'_{m_b}\equiv(W_y\restriction_{\mathrm{p}}J_{m_b-m_a})^i$. Since $\{\mathrm{coi}(W_x,\iota_x,U_x)\}_{x\in X}\cup\{\mathrm{coi}(W_j,\iota_j,U_j)\}_{j\in\omega}$ is coherent, we have

$$\begin{aligned}[[U\restriction_{\mathrm{p}}\propto(\Lambda_0,\iota)]]&=[[U_{m_a}\restriction_{\mathrm{p}}\propto(I''_{m_a},\iota_{m_a})]][[U_{m_a+1}\restriction_{\mathrm{p}}\propto(I_{m_a+1},\iota_{m_a+1})]]\\&\qquad\cdots[[U_{m_b-1}\restriction_{\mathrm{p}}\propto(I_{m_b-1},\iota_{m_b-1})]][[U\restriction_{\mathrm{p}}\propto(I'_{m_b},\iota_{m_b})]]\\&=\prod_{j\in\{0,\dots,m_b-m_a\}^i}[[(U_y\restriction_{\mathrm{p}}\propto(J_j,\iota_y))^i]]\\&=[[(U_y\propto(\Lambda_1,\iota_y))^i]].\end{aligned}$$

Suppose now that $\Lambda_0,\Lambda_1\subseteq\mathrm{p}_*(W)$ are intervals and $i\in\{-1,1\}$ are such that $W\restriction_{\mathrm{p}}\Lambda_0\equiv(W\restriction_{\mathrm{p}}\Lambda_1)^i$. Let $K_0=\{n\in\omega\mid I_n\cap\Lambda_0\neq\varnothing\}$ and $K_1=\{n\in\omega\mid I_n\cap\Lambda_1\neq\varnothing\}$.

Case 1. K_0 is infinite. In this case, if K_1 is finite then $W\restriction_{\mathrm{p}}\Lambda_0\in\mathrm{Pfine}(\{W_x\}_{x\in X})$, and we have already seen that this implies $W\in\mathrm{Pfine}(\{W_x\}_{x\in X})$ since K_0 is infinite, and this is a contradiction. Thus K_1 must be infinite in this case. If $i=-1$ then $W\restriction_{\mathrm{p}}\Lambda_0\equiv(W\restriction_{\mathrm{p}}\Lambda_1)^{-1}$. As Λ_0 and Λ_1 are terminal intervals in $\mathrm{p}_*(W)$, let without loss of generality $\Lambda_0\subseteq\Lambda_1$ and set $Q\equiv W\restriction_{\mathrm{p}}\Lambda_0$. Then $W\restriction_{\mathrm{p}}\Lambda_1\equiv Q^{-1}\equiv PQ$ for nonempty Q and some possibly empty P. Then $Q\equiv Q^{-1}P^{-1}\equiv PQP^{-1}$, so $P\equiv E$, forcing $\Lambda_0=\Lambda_1$. Then $W\equiv AQ\equiv AQ^{-1}$ for a possibly empty A. Write $Q\equiv BCB^{-1}$ for nonempty cyclically reduced C. Then WW^{-1} has nonempty reduced representative $ABCCB^{-1}A^{-1}$, contradiction.

Therefore $i=1$ and $W\restriction_{\mathrm{p}}\Lambda_0\equiv W\restriction_{\mathrm{p}}\Lambda_1$, and both Λ_0 and Λ_1 are infinite terminal intervals in $\mathrm{p}_*(W)$. If without loss of generality Λ_1 is a proper subinterval of Λ_0, then since $W\restriction_{\mathrm{p}}\Lambda_0\equiv W\restriction_{\mathrm{p}}\Lambda_1$ we can select a proper terminal subinterval $\Lambda_2\subseteq\Lambda_1$ such that $W\restriction_{\mathrm{p}}\Lambda_1\equiv W\restriction_{\mathrm{p}}\Lambda_2$, and inductively we select proper terminal subintervals

$\Lambda_{i+1} \subseteq \Lambda_i$ with $W \upharpoonright_{\mathrm{p}} \Lambda_i \equiv W \upharpoonright_{\mathrm{p}} \Lambda_{i+1}$. Thus, letting $\lambda \in \Lambda_0 \setminus \Lambda_1$ we see that the nonempty $W \upharpoonright_{\mathrm{p}} \{\lambda\}$ occurs infinitely often as a subword of W, so that W is not a word, a contradiction. Thus $\Lambda_0 = \Lambda_1$ and $[[U \upharpoonright_{\mathrm{p}} \propto(\Lambda_0, \iota)]] = [[(U \upharpoonright_{\mathrm{p}} \propto(\Lambda_1, \iota)^i)]]$.

Case 2. K_0 is finite. In this case we know that K_1 is also finite (by applying the argument in Case 1, since $W \upharpoonright_{\mathrm{p}} \Lambda_1 \equiv (W \upharpoonright_{\mathrm{p}} \Lambda_0)^i$). Thus $W \upharpoonright_{\mathrm{p}} \Lambda_0 \in \mathrm{Pfine}(\{W_n\}_{n\in\omega})$. If $K_0 = \varnothing$ then so also $K_1 = \varnothing = \Lambda_0 = \Lambda_1$ and it is easy to see that $[[U \upharpoonright_{\mathrm{p}} \propto(\Lambda_0, \iota)]] = [[E]] = [[(U \upharpoonright_{\mathrm{p}} \propto(\Lambda_1, \iota))^i]]$. In case $K_0 \neq \varnothing$, from the correspondence $W \upharpoonright_{\mathrm{p}} \Lambda_0 \equiv (W \upharpoonright_{\mathrm{p}} \Lambda_1)^i$ we decompose $\Lambda_0 \equiv \Theta_0\Theta_1 \cdots \Theta_m$ and $\Lambda_1 \equiv \Theta'_0\Theta'_1 \cdots \Theta'_m$ so that $W \upharpoonright_{\mathrm{p}} \Theta_j \equiv (W \upharpoonright_{\mathrm{p}} \Theta'_{f(j)})^i$ where

$$f(j) = \begin{cases} j & \text{if } i = 1, \\ m - j & \text{if } i = -1, \end{cases}$$

and each Θ_j is a subinterval of one of $I_{\min(K_0)}, \ldots, I_{\max(K_0)}$ and each Θ'_j is a subinterval of one of $I_{\min(K_1)}, \ldots, I_{\max(K_1)}$. Let $f_0 : \{0, \ldots, m\} \to \{\min(K_0), \ldots, \max(K_0)\}$ be the nondecreasing surjective function given by $\Theta_j \subseteq I_{f_0(j)}$, and also let $f_1 : \{0, \ldots, m\} \to \{\min(K_1), \ldots, \max(K_1)\}$ be given by $\Theta'_j \subseteq I_{f_1(j)}$. We have

$$\begin{aligned}
[[U \upharpoonright_{\mathrm{p}} \propto(\Lambda_0, \iota)]] &= \prod_{j=0}^{m} [[U_{f_0(j)} \upharpoonright_{\mathrm{p}} \propto(\Theta_j, \iota_{f_0(j)})]] \\
&= \prod_{j=0}^{m} [[(U_{f_1(f(j))} \upharpoonright_{\mathrm{p}} \propto(\Theta_{f(j)}, \iota_{f_1(f(j))}))^i]] \\
&= [[(U \upharpoonright_{\mathrm{p}} \propto(\Lambda_1, \iota))^i]]
\end{aligned}$$

where the first and third equalities hold by performing a deletion of finitely many pure words in Red_{κ_1} (Lemma 3.13) and the second equality holds by the coherence of the collection $\{\mathrm{coi}(W_n, \iota_n, U_n)\}_{n\in\omega}$. This completes case 2 and this part of the argument.

Suppose $y \in X \cup \omega$, $\Lambda_0 \subseteq \mathrm{p}_*(U)$ and $\Lambda_1 \subseteq \mathrm{p}_*(U_y)$ are intervals and $i \in \{-1, 1\}$ are such that $U \upharpoonright_{\mathrm{p}} \Lambda_0 \equiv (U_y \upharpoonright_{\mathrm{p}} \Lambda_1)^i$. Recalling that $U \equiv \prod_{n\in\omega}(U_n V_n)$ and none of the nonempty p-chunks of V_n are in $\mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{U_n\}_{n\in\omega})$ we see that $\Lambda_0 \subseteq \mathrm{p}_*(U_n)$ for some $n \in \omega$. From the coherence of $\{\mathrm{coi}(W_n, \iota_n, U_n)\}_{n\in\omega} \cup \{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ it is easy to see that $[[W \upharpoonright_{\mathrm{p}} \propto(\Lambda_0, \iota^{-1})]] = [[(W_y \upharpoonright_{\mathrm{p}} \propto(\Lambda_1, \iota_y^{-1}))^i]]$.

Finally suppose intervals $\Lambda_0, \Lambda_1 \subseteq \mathrm{p}_*(U)$ and $i \in \{-1, 1\}$ are such that $U \upharpoonright_{\mathrm{p}} \Lambda_0 \equiv (U \upharpoonright_{\mathrm{p}} \Lambda_1)^i$. Recall that $U \equiv \prod_{n\in\omega} U_n V_n$ with

$$\mathrm{p}_*(U) \equiv \prod_{n\in\omega} \mathrm{p}_*(U_n)\, \mathrm{p}_*(V_n)$$

and for all $n \in \omega$ we have $\|U_n\| = \|V_n\| \geq 2\|U_{n+1}\|$ and V_n uses only positive letters, satisfies $1 \leq |\mathrm{p}_*(V_n)| \leq 2$ and every nonempty p-chunk of V_n is not an element of $\mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{U_n\}_{n\in\omega})$.

If there exists $\lambda \in \Lambda_0$ and $n \in \omega$ such that $\lambda \in \mathrm{p}*(V_n)$ then $i = 1$ since every pure p-chunk of U which is not in $\mathrm{Pfine}(\{U_x\}_{x\in X}\{U_n\}_{n\in\omega})$ is a p-chunk in some V_m and therefore has positive letters only. Furthermore the order isomorphism $h : \Lambda_0 \to \Lambda_1$ induced by the word equivalence $U \restriction_{\mathrm{p}} \Lambda_0 \equiv U \restriction_{\mathrm{p}} \Lambda_1$ must have $h(\lambda) = \lambda$, for if $U \restriction_{\mathrm{p}} \{\lambda\}$ is, say, α-pure, then $U \restriction_{\mathrm{p}} \{\lambda\}$ is the unique α-pure p-chunk of U which has value $\|U \restriction_{\mathrm{p}} \{\lambda\}\|$ under the function $\|\cdot\|$. But this implies that h is the identity function since if, say, $\lambda' < \lambda$ and $h(\lambda') < \lambda'$, then $\lambda' < h^{-1}(\lambda') < h^{-2}(\lambda') < \cdots < \lambda$ and so the word $U \restriction_{\mathrm{p}} \Lambda_0$ has infinitely many disjoint occurrences of subwords equivalent to $U \restriction_{\mathrm{p}} \{\lambda'\}$, which contradicts the fact that U is a word. Thus $\Lambda_0 = \Lambda_1$ and obviously $[[W \restriction_{\mathrm{p}} \propto(\Lambda_0, \iota^{-1})]] = [[W \restriction_{\mathrm{p}} \propto(\Lambda_1, \iota^{-1})]]$.

On the other hand if $\Lambda_0 \cap \mathrm{p}*(V_n) = \varnothing$ for all $n \in \omega$ then $\Lambda_0 \subseteq \mathrm{p}*(U_m)$ for some $m \in \omega$. Thus $U \restriction_{\mathrm{p}} \Lambda_0 \in \mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{U_n\}_{n\in\omega})$, so $\Lambda_1 \cap \mathrm{p}*(V_n) = \varnothing$ for all $n \in \omega$ as well. Thus $\Lambda_1 \subseteq \mathrm{p}*(U_{m'})$ for some $m' \in \omega$. Then

$$\begin{aligned}[[W \restriction_{\mathrm{p}} \propto(\Lambda_0, \iota^{-1})]] &= [[W_m \restriction_{\mathrm{p}} \propto(\Lambda_0, \iota_m^{-1})]]\\ &= [[(W_{m'} \restriction_{\mathrm{p}} \propto(\Lambda_1, \iota_{m'}^{-1}))^i]]\\ &= [[(W \restriction_{\mathrm{p}} \propto(\Lambda_1, \iota^{-1}))^i]]\end{aligned}$$

since $U_m \restriction_{\mathrm{p}} \Lambda_0 \equiv U_{m'} \restriction_{\mathrm{p}} \Lambda_1$ and $\{\mathrm{coi}(W_n, \iota_n, U_n)\}_{n\in\omega}$ is coherent. □

3F. $\mathbb{Q}$-*type concatenations.* In this subsection we will devote our attention to proving the following:

Proposition 3.21. *Suppose that κ_0 and κ_1 are cardinal numbers greater than or equal to* 2. *Suppose that* $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ *is coherent, that* $\mathrm{p}*(W) \equiv \prod_{q\in\mathbb{Q}} I_q$ *with each $I_q \neq \varnothing$, $W \restriction_{\mathrm{p}} I_q \in \mathrm{Pfine}(\{W_x\}_{x\in X})$ for each $q \in \mathbb{Q}$, and $W \restriction_{\mathrm{p}} \bigcup \Lambda \notin$* $\mathrm{Pfine}(\{W_x\}_{x\in X})$ *for each interval $\Lambda \subseteq \mathbb{Q}$ with more than one point. Suppose also that $|X| < 2^{\aleph_0}$. Then there exists $U \in \mathrm{Red}_{\kappa_1}$ and a coi ι from W to U such that* $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W, \iota, U)\}$ *is coherent.*

Proof. Let $\{W_n\}_{n\in\omega}$ be a list such that for each $q \in \mathbb{Q}$ we have some $n \in \omega$ for which either $W \restriction_{\mathrm{p}} I_q \equiv W_n$ or $W \restriction_{\mathrm{p}} I_q \equiv W_n^{-1}$, and $n \neq n'$ implies $W_n \not\equiv W_{n'} \not\equiv W_n^{-1}$. Notice that indeed such a list must be infinite, for otherwise there is some $q' \in \mathbb{Q}$ such that $\{q \in \mathbb{Q} \mid W \restriction_{\mathrm{p}} I_q \equiv W \restriction_{\mathrm{p}} I_{q'}\}$ is infinite, which contradicts the fact that W is a word. By assumption, $\{W_n\}_{n\in\omega} \subseteq \mathrm{Pfine}(\{W_x\}_{x\in X})$. Select $P_0 \in \mathrm{Red}_{\kappa_1}$ and a coi ι_0 from W_0 to P_0 with nonempty domain such that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_0, \iota_0, P_0)\}$ is coherent by Lemma 3.17. Assuming we have chosen P_n and ι_n we select $P_{n+1} \in \mathrm{Red}_{\kappa_1}$ and a coi ι_{n+1} from W_{n+1} to P_{n+1} such that $\|P_{n+1}\| \leq \frac{1}{2}\|P_n\|$, the domain of ι_{n+1} is nonempty, and $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_j, \iota_j, P_j)\}_{j=0}^{n+1}$ is coherent by Lemmas 3.17 and 3.18. The collection $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_n, \iota_n, P_n)\}_{n\in\omega}$ is coherent by Lemma 3.12.

For each $m \in \omega$ select ordinals $\alpha_{m,b}, \alpha_{m,c} < \kappa_1$ such that P_m does not begin with an initial subword which is $\alpha_{m,b}$-pure and P_m does not end with a terminal subword which is $\alpha_{m,c}$-pure. By Lemma 3.19 we select an $\alpha_{m,b}$-pure word $V_{m,b}$ which uses only positive letters such that $\|V_{m,b}\| = \|P_m\|$, and $V_{m,b}(\max(\overline{V_{m,b}})) = a_{\alpha_{m,b},d(P_m)+1} = V_{m,b}(\min(\overline{V_{m,b}}))$ and $V_{m,b} \notin \mathrm{Pfine}(\{U_{x\in X}\}_{x\in X} \cup \{P_n\}_{n\in\omega})$. Similarly select an $\alpha_{m,c}$-pure word $V_{m,c}$ which uses only positive letters such that $\|V_{m,c}\| = \|P_m\|$, and $V_{m,c}(\max(\overline{V_{m,c}})) = a_{\alpha_{m,c},d(P_m)+1} = V_{m,c}(\min(\overline{V_{m,c}}))$ and $V_{m,c} \notin \mathrm{Pfine}(\{U_{x\in X}\}_{x\in X} \cup \{P_n\}_{n\in\omega})$.

Define functions $f_0 : \mathbb{Q} \to \omega$ and $f_1 : \mathbb{Q} \to \{\pm 1\}$ by $W \restriction_{\mathrm{p}} I_q \equiv W_{f_0(q)}^{f_1(q)}$. For each $m \in \omega$ the preimage $f_0^{-1}(m)$ is nonempty (by how the list $\{W_n\}_{n\in\omega}$ was chosen) and finite (since W is a word). For each $q \in \mathbb{Q}$ let $U_q \equiv (V_{f_0(q),b} P_{f_0(q)} V_{f_0(q),c})^{f_1(q)}$ and $U \equiv \prod_{q\in\mathbb{Q}} U_q$. Notice that this is a word since for each real number $\epsilon > 0$ the set $\{q \in \mathbb{Q} \mid \|U_q\| \geq \epsilon\}$ is finite. It is easy to see that each U_q is reduced and that moreover $\mathrm{p}_*(P_{f_0(q)}^{f_1(q)})$ is a subinterval of $\mathrm{p}_*(U_q)$ and $|\mathrm{p}_*(U_q) \setminus \mathrm{p}_*(P_{f_0(q)}^{f_1(q)})| = 2$.

Lemma 3.22. *U is reduced.*

Proof. For each $n \in \omega$ we let $J_n = \left\{q \in \mathbb{Q} \mid \|U_q\| = \frac{1}{n+1}\right\}$. We see that each J_n is finite since U is a word. For any cancellation $\mathcal{S}$ on U we define $L_n(\mathcal{S})$ to be the set of those $q \in J_n$ for which there exists $i \in \overline{U_q}$ which occurs in some ordered pair in $\mathcal{S}$. Define $L_n'(\mathcal{S}) \subseteq L_n(\mathcal{S})$ to be the set of all $q \in L_n(\mathcal{S})$ for which there exists a unique $q' \in L_n(\mathcal{S})$ such that $\mathcal{S}$ pairs each element in $\overline{U_q}$ with an element in $\overline{U_{q'}}$ and each element in $\overline{U_{q'}}$ with an element in $\overline{U_q}$. Our strategy will be to assume for contradiction that a nonempty cancellation over U exists and then to inductively modify the cancellation into a cancellation which witnesses a cancellation over W, contradicting the reducedness of W.

Suppose that $\mathcal{S}_0$ is a nonempty cancellation over U and let n_0 be minimal such that $L_{n_0}(\mathcal{S}) \neq \varnothing$. If $L_{n_0}(\mathcal{S}_0) = L_{n_0}'(\mathcal{S}_0)$ then we write $\mathcal{S}_1 = \mathcal{S}_0$ and move on to the next step of our induction. If $L_{n_0}(\mathcal{S}_0) \neq L_{n_0}'(\mathcal{S}_0)$ then we write $L_{n_0}(\mathcal{S}_0) \setminus L_{n_0}'(\mathcal{S}_0) = \{q_0, \ldots, q_k\}$ with $q_r < q_{r+1}$ under the ordering on $\mathbb{Q}$. Define a relation E on $L_{n_0}(\mathcal{S}_0) \setminus L_{n_0}'(\mathcal{S}_0)$ by writing $E(q_{r_0}, q_{r_1})$, where $q_{r_0}, q_{r_1} \in L_{n_0}(\mathcal{S}_0) \setminus L_{n_0}'(\mathcal{S}_0)$, if there exist $i_0 \in \overline{U_{q_{r_0}}}$ and $i_1 \in \overline{U_{q_{r_1}}}$ such that $\langle i_0, i_1\rangle \in \mathcal{S}_0$. Since each U_q is reduced we see that $E(q_r, q_r)$ is false for all $0 \leq r \leq k$. Also, $E(q_{r_0}, q_{r_1})$ implies that $q_{r_0} < q_{r_1}$ since $\langle i_0, i_1\rangle \in \mathcal{S}_0$ implies $i_0 < i_1$ in $\overline{U}$. By how each U_q is defined, we see that $U_q(\min(\overline{U_q})) = U_q(\max(\overline{U_q})) \in \{a_{\alpha_{n_0},n_0}^{\pm 1}\}$ for each $q \in L_{n_0}(\mathcal{S}_0)$. For $q' \in \bigcup_{n>n_0} L_n(\mathcal{S}_0)$ we have $\|U_{q'}\| < 1/(n_0+1)$. Since U_q is reduced for each $q \in L_{n_0}(\mathcal{S}_0)$, we see that for each $q \in L_{n_0}(\mathcal{S}_0)$ at least one of $\max(\overline{U_q})$ or $\min(\overline{U_q})$ must appear in some element of $\mathcal{S}_0$. Moreover, by how $L_n'(\mathcal{S}_0)$ is defined, for each $q \in L_{n_0}(\mathcal{S}_0) \setminus L_{n_0}'(\mathcal{S}_0)$ at least one of $\max(\overline{U_q})$ or $\min(\overline{U_q})$ must appear in $\mathcal{S}_0$ and be paired with some element in $\overline{U_{q'}}$ for some $q' \in L_{n_0}(\mathcal{S}_0) \setminus (L_{n_0}'(\mathcal{S}_0) \cup \{q\})$.

Thus we see that each $q \in L_{n_0}(\mathcal{S}_0) \setminus L_{n_1}'(\mathcal{S}_0)$ must appear as a first or second coordinate in the relation E. Notice as well that if $E(q_{r_0}, q_{r_1})$ and $E(q_{r_2}, q_{r_3})$ where $q_{r_0} < q_{r_2} \leq q_{r_1}$ then $q_{r_0} < q_{r_2} < q_{r_3} \leq q_{r_1}$ by property (4) of cancellations (see

Definition 2.3). Similarly if $E(q_{r_0}, q_{r_1})$ and $E(q_{r_2}, q_{r_3})$ hold and $q_{r_0} \le q_{r_3} < q_{r_1}$ then we have $q_{r_0} \le q_{r_2} < q_{r_3} < q_{r_1}$. Since the set $L_{n_0}(\mathcal{S}_0) \setminus L'_{n_1}(\mathcal{S}_0)$ is finite, we therefore have some $0 \le r < k$ such that $E(q_r, q_{r+1})$. Again, since U_{q_r} and $U_{q_{r+1}}$ are each reduced we must have $\langle \max(\overline{U_{q_r}}), \min(\overline{U_{q_{r+1}}})\rangle \in \mathcal{S}_0$. Thus $U_{q_r} \equiv (U_{q_{r+1}})^{-1}$ and we let $f : \overline{U_{q_r}} \to \overline{U_{q_{r+1}}}$ be an order reversing bijection with $U_{q_{r+1}}(f(i)) = (U_{q_r}(i))^{-1}$ witnessing this equivalence.

We let $\mathcal{S}_0^{(1)}$ be given by

$$\begin{aligned}\mathcal{S}_0^{(1)} = \{\langle i_0, i_1\rangle \in \mathcal{S}_0 \mid i_0, i_1 \notin \overline{U_{q_r}} \cup \overline{U_{q_{r+1}}}\} \\ \cup \{\langle i_0, f(i_0)\rangle \mid i_0 \in \overline{U_{q_r}}\} \\ \cup \{\langle i_0, i_1\rangle \in \overline{U} \times \overline{U} \mid (\exists i_2 \in \overline{U_{q_r}})\ \langle i_0, i_2\rangle, \langle f(i_2), i_1\rangle \in \mathcal{S}_0\} \\ \cup \{\langle i_0, i_1\rangle \in \overline{U} \times \overline{U} \mid (\exists i_2 \in \overline{U_{q_r}})\ \langle i_1, i_2\rangle, \langle i_0, f(i_2)\rangle \in \mathcal{S}_0\} \\ \cup \{\langle i_0, i_1\rangle \in \overline{U} \times \overline{U} \mid (\exists i_2 \in \overline{U_{q_r}})\ \langle i_2, i_1\rangle, \langle f(i_2), i_0\rangle \in \mathcal{S}_0\}.\end{aligned}$$

It is straightforward to see that $\mathcal{S}_0^{(1)}$ is a cancellation and $L_n(\mathcal{S}_0^{(1)}) \subseteq L_n(\mathcal{S}_0)$ for all $n \in \omega$. But also $L'_{n_0}(\mathcal{S}_0^{(1)}) = L'_{n_0}(\mathcal{S}_0) \sqcup \{q_r, q_{r+1}\}$. Iterating the argument to produce $\mathcal{S}_0^{(2)}$, $\mathcal{S}_0^{(3)}$, etc. so as to make $L'_{n_0}(\mathcal{S}_0^{(j+1)})$ strictly include $L'_{n_0}(\mathcal{S}_0^{(j)})$ and have $L_{n_0}(\mathcal{S}_0^{(j+1)}) \subseteq L_{n_0}(\mathcal{S}_0^{(j)})$, we see, since $L_{n_0}(\mathcal{S}_0)$ is finite, that eventually $L'_{n_0}(\mathcal{S}_0^{(j)}) = L_{n_0}(\mathcal{S}_0^{(j)})$. Set $\mathcal{S}_1 = \mathcal{S}_0^{(j)}$ for sufficiently large j.

Notice that $\mathcal{S}_1$ does not pair any element of $\overline{U_q}$ with $\overline{U_{q'}}$ when $q \in L_{n_0}(\mathcal{S}_1)$ and $q' \notin L_{n_0}(\mathcal{S}_1)$. Letting $n_1 \in \omega$ be minimal such that $n_1 > n_0$ and $L_{n_1}(\mathcal{S}_1) \neq \varnothing$ (an $n > n_0$ with $L_n(\mathcal{S}_1) \neq \varnothing$ must exist since $\mathbb{Q}$ is order dense), we may thus repeat the arguments as before to create $\mathcal{S}_2$ such that $L_{n_1}(\mathcal{S}_2) = L'_{n_1}(\mathcal{S}_2)$ and also $\mathcal{S}_2$ agrees with $\mathcal{S}_1$ on $L_{n_0}(\mathcal{S}_1) = L_{n_0}(\mathcal{S}_2)$. Select $n_2 > n_1$ which is minimal such that $L_{n_2}(\mathcal{S}_2) \neq \varnothing$, produce $\mathcal{S}_3$, and continue this process inductively. Let $\mathcal{S}_\infty$ equal $\big\{\langle i_0, i_1\rangle \mid (\exists p \in \omega)\ i_0, i_1 \in \bigcup_{q \in L_{n_p}} U_q \text{ and } \langle i_0, i_1\rangle \in \mathcal{S}_{p+1}\big\}$ and we have that $\mathcal{S}_\infty$ is a cancellation such that $L_n(\mathcal{S}_\infty) = L'_n(\mathcal{S}_\infty)$ for all $n \in \omega$ and $\mathcal{S}_\infty \neq \varnothing$.

But now let $\mathcal{S}' = \{\langle q_0, q_1\rangle \mid \exists (i_0 \in \overline{U_{q_0}}, i_1 \in \overline{U_{q_1}})\ \langle i_0, i_1\rangle \in \mathcal{S}_\infty\}$ and notice that $\mathcal{S}'$ is a pairing of a subset of elements in $\mathbb{Q}$ that satisfies the comparable properties (1)–(4) of Definition 2.3, and $\langle q_0, q_1\rangle \in \mathcal{S}'$ implies that $U_{q_0} \equiv (U_{q_1})^{-1}$. Then $W_{q_0} \equiv (W_{q_1})^{-1}$ for $\langle q_0, q_1\rangle \in \mathcal{S}'$ and it is easy to use $\mathcal{S}'$ to define a nonempty cancellation $\mathcal{S}$ on W, and we have a contradiction. □

Now that we know that U is reduced, it is easy to see that

$$\mathrm{p}_*(U) \equiv \prod_{q \in \mathbb{Q}} \mathrm{p}_*(U_q) \equiv \prod_{q \in \mathbb{Q}} (\mathrm{p}_*(V_{f_0(q),b})\, \mathrm{p}_*(P_{f_0(q)})\, \mathrm{p}_*(V_{f_0(q),c}))^{f_1(q)}.$$

Using the collection $\{\mathrm{coi}(W_n, \iota_n, P_n)\}_{n\in\omega}$ we define the coi ι from W to U in the natural way. Namely, let T_q denote the subword $W \restriction_{\mathrm{p}} I_q$, and recall that $W_{f_0(q)}^{f_1(q)} \equiv T_q$ and $U_q \equiv (V_{f_0(q)} P_{f_0(q)} V_{f_0(q)})^{f_1(q)}$. Let $g : \mathrm{p}_*(P_{f_0(q)}^{f_1(q)}) \to \mathrm{p}_*(U_q)$ denote the order embedding given by this last equivalence, and ι_q be the function whose domain

$\mathrm{dom}(\iota_q)$ is the image of $\mathrm{dom}(\iota_{f_0(q)})$ under the order isomorphism $f : \mathrm{p}_*(W_{f_0(q)}^{f_1(q)}) \to \mathrm{p}_*(W_q)$, whose image lies in $\mathrm{p}_*(U_q)$ and such that $\iota_q(i) = g \circ \iota_{f_0(q)} \circ f^{-1}(i)$.

Notice that ι_q is an order isomorphism between its domain and image since $\iota_{f_0(q)}$ is order preserving and exactly one of the following holds:

- f is an order isomorphism between $\mathrm{p}_*(T_q)$ and $\mathrm{p}_*(W_{f_0(q)})$ and g is an order embedding from $\mathrm{p}_*(P_{f_0(q)})$ to $\mathrm{p}_*(U_q)$.
- f gives an order reversing bijection between $\mathrm{p}_*(T_q)$ and $\mathrm{p}_*(W_{f_0(q)})$ and g gives an order reversing embedding from $\mathrm{p}_*(P_{f_0(q)})$ to $\mathrm{p}_*(U_q)$.

Since $\mathrm{Close}(\mathrm{dom}(\iota_n), \mathrm{p}_*(W_n))$, the relation $\mathrm{Close}(\mathrm{dom}(\iota_q), \mathrm{p}_*(T_q))$ is easily seen to hold. Also, since $|\mathrm{p}_*(V_{f_0(q),b})| = 1 = |\mathrm{p}_*(V_{f_0(q),c})|$, we easily see that $\mathrm{Close}(\mathrm{im}(\iota_q), \mathrm{p}_*(U_q))$. Let ι be the order isomorphism given by $\iota = \bigcup_{q\in\mathbb{Q}} \iota_q$. By Lemma 3.6 (iii) we have $\mathrm{Close}(\mathrm{dom}(\iota), \mathrm{p}_*(W))$ and $\mathrm{Close}(\mathrm{im}(\iota), \mathrm{p}_*(U_q))$, so ι is a coi from W to U. We check the coherence of

$$\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_n, \iota_n, P_n)\}_{n\in\omega} \cup \{\mathrm{coi}(W, \iota, U)\},$$

which will imply the coherence of $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W, \iota, U)\}$.

Suppose that $x_0 \in X \cup \omega$, $\Lambda_0 \subseteq \mathrm{p}_*(W)$ and $\Lambda_1 \subseteq \mathrm{p}_*(W_{x_0})$ are intervals, and $i \in \{-1, 1\}$ are such that $W \restriction_{\mathrm{p}} \Lambda_0 \equiv (W_{x_0} \restriction_{\mathrm{p}} \Lambda_1)^i$. Notice that Λ_0 must be a subinterval of some $\mathrm{p}_*(T_q)$ since $\mathbb{Q}$ is order dense, $W \restriction_{\mathrm{p}} \Lambda \notin \mathrm{Pfine}(\{W_x\}_{x\in X})$ for each interval $\Lambda \subseteq \mathbb{Q}$ with more than one point and $(W_{x_0} \restriction_{\mathrm{p}} \Lambda_1)^i \in \mathrm{Pfine}(\{W_x\}_{x\in X} \cup \{W_n\}_{n\in\omega}) = \mathrm{Pfine}(\{W_x\}_{x\in X})$. But letting $f : \mathrm{p}_*(W_{f_0(q)}^{f_1(q)}) \to \mathrm{p}_*(T_q)$ be the natural order isomorphism and $\Lambda_0' \subseteq \mathrm{p}_*(W_{f_0(q)}^{f_1(q)})$ be the interval given by $f^{-1}(\Lambda_0)$, it is easy to see that

$$[[U \restriction_{\mathrm{p}} \propto(\Lambda_0, \iota)]] = [[(P_{f_0(q)} \restriction_{\mathrm{p}} \propto(\Lambda_0', \iota_{f_0(q)}))^{f_1(q)}]] = [[(U_{x_0} \restriction_{\mathrm{p}} \propto(\Lambda_1, \iota_{x_0}))^i]]$$

by how the function ι_q was defined (for the first equality) and the coherence of $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_n, \iota_n, P_n)\}_{n\in\omega}$ (for the second equality).

Next, suppose that $\Lambda_0, \Lambda_1 \subseteq \mathrm{p}_*(W)$ are intervals and $i \in \{-1, 1\}$ are such that $W \restriction_{\mathrm{p}} \Lambda_0 \equiv (W \restriction_{\mathrm{p}} \Lambda_1)^i$. Let $J_0 = \{q \in \mathbb{Q} \mid \mathrm{p}_*(T_q) \cap \Lambda_0 \neq \varnothing\}$ and $J_1 = \{q \in \mathbb{Q} \mid \mathrm{p}_*(T_q) \cap \Lambda_1 \neq \varnothing\}$. Clearly each of J_0 and J_1 are intervals in $\mathbb{Q}$. If, say, J_0 is empty or a singleton then $W \restriction_{\mathrm{p}} \Lambda_0 \in \mathrm{Pfine}(\{W_x\}_{x\in X})$, and so J_1 is not infinite (since we are assuming $W \restriction_{\mathrm{p}} \Lambda \notin \mathrm{Pfine}(\{W_x\}_{x\in X})$ for each interval $\Lambda \subseteq \mathbb{Q}$ with more than one point). Similarly if J_1 is empty or a singleton then J_0 is finite (hence a singleton or empty). In case J_0 is finite we can argue as before, using the coherence of the collection $\{\mathrm{coi}(W_n, \iota_n, P_n)\}_{n\in\omega}$ to obtain $[[U \restriction_{\mathrm{p}} \propto(\Lambda_0, \iota)]] = [[(U \restriction_{\mathrm{p}} \propto(\Lambda_1, \iota))^i]]$.

Suppose now that J_0 (and therefore also J_1) is infinite. Since J_0 is order dense and $W \restriction_{\mathrm{p}} \Lambda \notin \mathrm{Pfine}(\{W_x\}_{x\in X})$ for each interval $\Lambda \subseteq \mathbb{Q}$ with more than one point, we notice that J_0 has a minimum if and only if the word $W \restriction_{\mathrm{p}} \Lambda_0$ has a nonempty initial subword which is an element of $\mathrm{Pfine}(\{W_x\}_{x\in X})$. Also, if J_0 has minimum q

then $W \restriction_p (p_*(W_q) \cap \Lambda_0)$ is the maximal initial subword of $W \restriction_p \Lambda_0$ which is an element in $\mathrm{Pfine}(\{W_x\}_{x\in X})$. Similarly J_0 has a maximum if and only if the word $W \restriction_p \Lambda_0$ has a nonempty terminal subword which is an element of $\mathrm{Pfine}(\{W_x\}_{x\in X})$, and if J_0 has maximum q then $W \restriction_p (p_*(W_q) \cap \Lambda_0)$ is the maximal terminal subword of $W \restriction_p \Lambda_0$ which is an element in $\mathrm{Pfine}(\{W_x\}_{x\in X})$. Let $J_0' \subseteq J_0$ be the subinterval which consists of J_0 minus any maximum or minimum that J_0 might have. By similar reasoning, we see that for each $q \in J_0'$ the subword T_q is a maximal subword of $W \restriction_p \Lambda_0$ which is an element of $\mathrm{Pfine}(\{W_x\}_{x\in X})$.

The comparable claims hold for J_1; for example J_1 has a minimum if and only if the word $W \restriction_p \Lambda_1$ has a nonempty initial subword which is an element of $\mathrm{Pfine}(\{W_x\}_{x\in X})$, and if $q \in J_1$ is minimal then $W \restriction_p (p_*(T_q) \cap \Lambda_1)$ is the maximal initial subword of $W \restriction_p \Lambda_1$ which is an element in $\mathrm{Pfine}(\{W_x\}_{x\in X})$. Define the interval $J_1' \subseteq J_1$ similarly. As $W \restriction_p \Lambda_0 \equiv (W \restriction_p \Lambda_1)^i$, we see that if $i = 1$:

- J_0 has a minimum if and only if J_1 has one.
- J_0 has a maximum if and only if J_1 has one.
- If $q_0 = \min(J_0)$ and $q_1 = \min(J_1)$, $W \restriction_p (\Lambda_0 \cap p_*(T_{q_0})) \equiv W \restriction_p (\Lambda_1 \cap p_*(T_{q_1}))$.
- If $q_0 = \max(J_0)$ and $q_1 = \max(J_1)$, $W \restriction_p (\Lambda_0 \cap p_*(T_{q_0})) \equiv W \restriction_p (\Lambda_1 \cap p_*(T_{q_1}))$.
- There is an order isomorphism $h : J_0' \to J_1'$ such that $W_{h(q)} \equiv W_g$.

Now if $i = -1$:

- J_0 has a minimum if and only if J_1 has a maximum.
- J_0 has a maximum if and only if J_1 has a minimum.
- If $q_0 = \min(J_0)$ and $q_1 = \max(J_1)$, $W \restriction_p (\Lambda_0 \cap p_*(T_{q_0})) \equiv \big(W \restriction_p (\Lambda_1 \cap p_*(T_{q_1}))\big)^{-1}$.
- If $q_0 = \max(J_0)$ and $q_1 = \min(J_1)$, $W \restriction_p (\Lambda_0 \cap p_*(T_{q_0})) \equiv \big(W \restriction_p (\Lambda_1 \cap p_*(T_{q_1}))\big)^{-1}$.
- There is an order reversing bijection $h : J_0' \to J_1'$ such that $T_{h(q)} \equiv (T_q)^{-1}$.

From this and how the ι_q were defined it is clear that

$$U \restriction_p \propto\Bigg(\bigcup_{q\in J_0'} p_*(T_q), \iota\Bigg) \equiv \Bigg(U \restriction_p \propto\Bigg(\bigcup_{q\in J_1'} p_*(T_q), \iota\Bigg)\Bigg)^i.$$

Now suppose, for example, $i = -1$ and J_0 has maximum and minimum. Let $K \equiv U \restriction_p \propto\big(\bigcup_{q\in J_0'} p_*(T_q), \iota\big)$. By Lemma 3.13 we have that $[[U \restriction_p \propto(\Lambda_0, \iota)]]$ is equal to

$$[[U \restriction_p \propto(\Lambda_0 \cap p_*(T_{\min(J_0)}), \iota)]][[K]][[U \restriction_p \propto(\Lambda_0 \cap p_*(T_{\max(J_0)}), \iota)]]$$

and that $[[(U \restriction_p \propto(\Lambda_1, \iota))^{-1}]]$ is equal to

$$[[(U \restriction_p (\Lambda_1 \cap p_*(T_{\max(J_1)}), \iota))^{-1}]][[K]][[(U \restriction_p \propto(\Lambda_1 \cap p_*(T_{\min(J_1)}), \iota))^{-1}]].$$

Since the coi collection $\{\mathrm{coi}(W_n, \iota_n, P_n)\}_{n\in\omega}$ is coherent we also have the equalities

$$[[U \restriction_{\mathrm{p}} \propto(\Lambda_0 \cap \mathrm{p}_*(T_{\min(J_0)}), \iota)]] = [[(U \restriction_{\mathrm{p}} \propto(\Lambda_1 \cap \mathrm{p}_*(T_{\max(J_1)}), \iota))^{-1}]]$$

and

$$[[U \restriction_{\mathrm{p}} \propto(\Lambda_0 \cap \mathrm{p}_*(T_{\max(J_0)}), \iota)]] = [[(U \restriction_{\mathrm{p}} \propto(\Lambda_1 \cap \mathrm{p}_*(T_{\min(J_1)}), \iota))^{-1}]].$$

Thus $[[U \restriction_{\mathrm{p}} \propto(\Lambda_0, \iota)]] = [[(U \restriction_{\mathrm{p}} \propto(\Lambda_1, \iota))^{-1}]]$ by direct substitution. All other possibilities can be similarly argued.

Suppose that $x_0 \in X$ and $\Lambda_0 \subseteq \mathrm{p}_*(U)$, $\Lambda_1 \subseteq \mathrm{p}_*(U_{x_0})$ are intervals and $i \in \{-1, 1\}$ are such that $U \restriction_{\mathrm{p}} \Lambda_0 \equiv (U_{x_0} \restriction_{\mathrm{p}} \Lambda_1)^i$. As $(U_{x_0} \restriction_{\mathrm{p}} \Lambda_1)^i \in \mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{P_n\}_{n\in\omega})$, and $V_{m,b}, V_{m,c} \notin \mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{P_n\}_{n\in\omega})$ for all $m \in \omega$ we see that Λ_0 must be a subinterval of some $\mathrm{p}_*(U_q)$, and more particularly a subinterval of $\mathrm{p}_*(P_{f_0(q)}^{f_1(q)})$. By how ι_q was defined, and since $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W_n, \iota_n, P_n)\}_{n\in\omega}$ is coherent it follows that

$$[[W \restriction_{\mathrm{p}} \propto(\Lambda_0, \iota^{-1})]] = [[(W_{x_0} \restriction_{\mathrm{p}} \propto(\Lambda_1, \iota_{x_0}^{-1}))^i]].$$

If $n_0 \in \omega$ and $\Lambda_0 \subseteq \mathrm{p}_*(U)$, $\Lambda_1 \subseteq \mathrm{p}_*(P_{n_0})$ are intervals and $i \in \{-1, 1\}$ are such that $U \restriction_{\mathrm{p}} \Lambda_0 \equiv (P_{n_0} \restriction_{\mathrm{p}} \Lambda_1)^i$ then the same argument shows that

$$[[W \restriction_{\mathrm{p}} \propto(\Lambda_0, \iota^{-1})]] = [[(W_{n_0} \restriction_{\mathrm{p}} \propto(\Lambda_1, \iota_{n_0}^{-1}))^i]].$$

Finally, suppose that intervals $\Lambda_0, \Lambda_1 \subseteq \mathrm{p}_*(U)$ and $i \in \{-1, 1\}$ are such that $U \restriction_{\mathrm{p}} \Lambda_0 \equiv (U \restriction_{\mathrm{p}} \Lambda_1)^i$. As before we define

$$J_0 = \{q \in \mathbb{Q} \mid \mathrm{p}_*(U_q) \cap \Lambda_0 \neq \varnothing\}, \quad J_1 = \{q \in \mathbb{Q} \mid \mathrm{p}_*(U_q) \cap \Lambda_1 \neq \varnothing\}.$$

Once again, the cases where J_0, hence also J_1, is empty or a singleton are treated the same. We therefore assume that both J_0 and J_1 are infinite. One sees that J_0 has a minimum if and only if $U \restriction_{\mathrm{p}} \Lambda_0$ has a nonempty initial subword which is a pure p-chunk (i.e., a word $V_{m,b}^{\pm 1}$ or $V_{m,c}^{\pm 1}$ for some $m \in \omega$) or which is in $\mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{P_n\}_{n\in\omega})$, and not both since the words $V_{m,b}$ and $V_{m,c}$ were not in $\mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{P_n\}_{n\in\omega})$. In either case, J_0 has a minimum if and only if there is an element $\lambda \in \Lambda_0$ for which $U \restriction_{\mathrm{p}} \{\lambda\} \notin \mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{P_n\}_{n\in\omega})$ which is minimal. Similar such statements for maxima and J_1 apply. Thus we see that when $i = 1$, J_0 has a minimum if and only if J_1 has one, and J_0 has a maximum if and only if J_1 has one. When $i = -1$ the comparable dual statements hold. Let J_0' be the set J_0 minus any maximal or minimal element and define J_1' analogously. For each $q \in J_0'$ (or $q \in J_1'$) we have that $U_{f_0(q)}^{f_1(q)}$ is a maximal subword of U which is in $\mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{P_n\}_{n\in\omega})$, and each of $V_{f_0(q),b}^{f_1(q)}$ and $V_{f_0(q),c}^{f_1(q)}$ is a maximal p-chunk of U all of whose nonempty p-chunks are not in $\mathrm{Pfine}(\{U_x\}_{x\in X} \cup \{P_n\}_{n\in\omega})$.

In particular, $U \restriction_{\mathrm{p}} \bigcup_{q\in J_0'} \mathrm{p}_*(U_q)$ is the word obtained from U by removing an initial pure p-chunk (if it exists) and then removing an initial nonempty p-chunk which is an element of $\mathrm{Pfine}(\{U_x\}_{x\in X}\cup\{P_n\}_{n\in\omega})$ (if it exists) and then removing an initial pure p-chunk (if step two applies) and doing the similar three-step process to the terminal part of the word U. Hence it is clear that $U \upharpoonright_{\mathrm{p}} \bigcup_{q\in J_0'} \mathrm{p}_*(U_q) \equiv (U \upharpoonright_{\mathrm{p}} \bigcup_{q\in J_1'} \mathrm{p}_*(U_q))^i$. Moreover this word equality will pair maximal intervals $\Lambda \subseteq \bigcup_{q\in J_0'} \mathrm{p}_*(U_q)$ for which $U \upharpoonright_{\mathrm{p}} \Lambda \in \mathrm{Pfine}(\{U_x\}_{x\in X}\cup\{P_n\}_{n\in\omega})$ with such intervals in $\bigcup_{q\in J_1'} \mathrm{p}_*(U_q)$, and for such a Λ we'll have $U \upharpoonright_{\mathrm{p}} \Lambda \equiv P_n^{\pm 1}$ for some $n\in\omega$. As $P_n \not\equiv P_{n'} \not\equiv P_n^{-1}$ when $n\neq n'$ we have a bijection $h: J_0' \to J_1'$ which is an order isomorphism in case $i=1$, or an order reversal in case $i=-1$, such that $U_{h(q)} \equiv (U_q)^i$ once again. Thus we get

$$W \upharpoonright_{\mathrm{p}} \propto\Bigg(\bigcup_{q\in J_0'} \mathrm{p}_*(U_q),\, \iota^{-1}\Bigg) \equiv \Bigg(W \upharpoonright_{\mathrm{p}} \propto\Bigg(\bigcup_{q\in J_1'} \mathrm{p}_*(U_q),\, \iota^{-1}\Bigg)\Bigg)^i.$$

Thus for example, if $i=-1$ and J_0 has maximum and minimum then we let $K \equiv W \upharpoonright_{\mathrm{p}} \propto\big(\bigcup_{q\in J_0'} \mathrm{p}_*(U_q),\, \iota^{-1}\big)$. Then $[[W \upharpoonright_{\mathrm{p}} \propto(\Lambda_0, \iota^{-1})]]$ is equal to the product

$$[[W \upharpoonright_{\mathrm{p}} \propto(\Lambda_0 \cap \mathrm{p}_*(U_{\min(J_0)}), \iota^{-1})]][[K]][[W \upharpoonright_{\mathrm{p}} \propto(\Lambda_0 \cap \mathrm{p}_*(U_{\max(J_0)}), \iota^{-1})]]$$

by Lemma 3.13. By the same reasoning we have that $[[(W \upharpoonright_{\mathrm{p}} \propto(\Lambda_1, \iota^{-1}))^{-1}]]$ is equal to

$$[[(W \upharpoonright_{\mathrm{p}} \propto(\Lambda_1 \cap \mathrm{p}_*(U_{\max(J_1)}), \iota^{-1}))^{-1}]][[K]][[(W \upharpoonright_{\mathrm{p}} \propto(\Lambda_1 \cap \mathrm{p}_*(U_{\min(J_1)}), \iota^{-1}))^{-1}]].$$

By coherence we get that

$$[[W \upharpoonright_{\mathrm{p}} \propto(\Lambda_0 \cap \mathrm{p}_*(U_{\min(J_0)}), \iota^{-1})]] = [[(W \upharpoonright_{\mathrm{p}} \propto(\Lambda_1 \cap \mathrm{p}_*(U_{\max(J_1)}), \iota^{-1}))^{-1}]],$$

and similarly

$$[[W \upharpoonright_{\mathrm{p}} \propto(\Lambda_0 \cap \mathrm{p}_*(U_{\max(J_0)}), \iota^{-1})]] = [[(W \upharpoonright_{\mathrm{p}} \propto(\Lambda_1 \cap \mathrm{p}_*(U_{\min(J_1)}), \iota^{-1}))^{-1}]],$$

and so the equality $[[W \upharpoonright_{\mathrm{p}} \propto(\Lambda_0, \iota^{-1})]] = [[(W \upharpoonright_{\mathrm{p}} \propto(\Lambda_1, \iota^{-1}))^{-1}]]$ is immediate. □

3G. ***Arbitrary extensions.*** In this subsection we will prove the following proposition and then complete the proof of Theorem A as well as prove Theorem B.

Proposition 3.23. *Suppose that κ_0 and κ_1 are cardinal numbers greater than or equal to 2. Suppose that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X}$ is coherent and that $|X| < 2^{\aleph_0}$. Then given $W \in \mathrm{Red}_{\kappa_0}$ there exists $U \in \mathrm{Red}_{\kappa_1}$ and a coi ι from W to U such that $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W, \iota, U)\}$ is coherent.*

Proof. Assume the hypotheses. If W is the empty word E then we let $U \equiv E$ and ι be the empty function. This clearly satisfies the conclusion of the proposition. Thus we may now assume that W is not E and so $\mathrm{p}_*(W)$ is nonempty. For each $\lambda \in \mathrm{p}_*(W)$ we let ι_λ be the empty function, so ι_λ is a coi from $W \upharpoonright_{\mathrm{p}} \{\lambda\}$ to E. It is quite trivial to see that $\mathcal{T}_0 = \{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x\in X} \cup \{\mathrm{coi}(W \upharpoonright_{\mathrm{p}} \{\lambda\}, \iota_\lambda, E)\}_{\lambda\in \mathrm{p}_*(W)}$

is coherent. Let $\prec$ be a well-order on the set $\mathrm{p}_*(W)$ and if $\mathcal{T}$ is a collection of cois then we let $h(\mathcal{T})$ denote the set of first words listed in the ordered triples (for example $h(\mathcal{T}_0) = \{W_x\}_{x\in X} \cup \{W \upharpoonright_{\mathrm{p}} \{\lambda\}\}_{\lambda\in \mathrm{p}_*(W)})$.

Step 1. Define a function f_0 from an initial subset of the set $\aleph_1$ of countable ordinals to $\mathrm{p}_*(W)$, as well as a function f_1 with the same domain as f_0 and with codomain the set of two letters $\{L, R\}$ and f_2 a function with the same domain as f_0 and with codomain the set of intervals in $\mathrm{p}_*(W)$. We shall also extend the coi collection. If each $\lambda \in \mathrm{p}_*(W)$ is contained in a maximal interval $I \subseteq \mathrm{p}_*(W)$ such that $W \upharpoonright_{\mathrm{p}} I \in h(\mathcal{T}_\zeta)$ then we cease our construction of step 1 and proceed to step 2. If it is not the case that each $\lambda \in \mathrm{p}_*(W)$ is contained in a maximal interval $I \subseteq \mathrm{p}_*(W)$ such that $W \upharpoonright_{\mathrm{p}} I \in h(\mathcal{T}_\zeta)$ then we select a minimal such λ under the well-ordering $\prec$ and let $f_0(\zeta) = \lambda$. Note that it is possible that each singleton $\{\lambda\}$ is already maximal such that $W \upharpoonright_{\mathrm{p}} \{\lambda\} \in h(\mathcal{T}_0)$. At least one of two possibilities holds:

Case i. If there is a sequence $\{I_m\}_{m\in\omega}$ such that $\lambda = \min(I_m)$ and I_m is strictly included in I_{m+1} for all $m \in \omega$ with $W \upharpoonright_{\mathrm{p}} I_m \in \mathrm{Pfine}(h(\mathcal{T}_\zeta))$ but $W \upharpoonright_{\mathrm{p}} \bigcup_{m\in\omega} I_m \notin \mathrm{Pfine}(h(\mathcal{T}_\zeta))$, then we let $f_1(\zeta) = L$ (for Left endpoint) and $f_2(\zeta) = \bigcup_{m\in\omega} I_m$. By Proposition 3.20 we select $U_\zeta \in \mathrm{Red}_{\kappa_1}$ and a coi ι_ζ from $W \upharpoonright_{\mathrm{p}} f_2(\zeta)$ to U_ζ such that $\mathcal{T}_{\zeta+1} = \mathcal{T}_\zeta \cup \{\mathrm{coi}(W \upharpoonright_{\mathrm{p}} f_2(\zeta), \iota_\zeta, U_\zeta)\}$ is coherent.

Case ii. If such a sequence as in case i does not exist then there exists a sequence $\{I_m\}_{m\in\omega}$ such that $\lambda = \max(I_m)$ and I_m is strictly included in I_{m+1} for all $m \in \omega$ with $W \upharpoonright_{\mathrm{p}} I_m \in \mathrm{Pfine}(h(\mathcal{T}_\zeta))$, but $W \upharpoonright_{\mathrm{p}} \bigcup_{m\in\omega} I_m \notin \mathrm{Pfine}(h(\mathcal{T}_\zeta))$. In this case we let $f_1(\zeta) = R$ (for Right endpoint) and $f_2(\zeta) = \bigcup_{m\in\omega} I_m$. By Proposition 3.20 applied to the word W^{-1} we select $U_\zeta \in \mathrm{Red}_{\kappa_1}$ and a coi ι_ζ from $W \upharpoonright_{\mathrm{p}} f_2(\zeta)$ to U_ζ such that $\mathcal{T}_{\zeta+1} = \mathcal{T}_\zeta \cup \{\mathrm{coi}(W \upharpoonright_{\mathrm{p}} f_2(\zeta), \iota_\zeta, U_\zeta)\}$ is coherent.

Iterating this recursion and letting $\mathcal{T}_\zeta = \bigcup_{\zeta_0<\zeta} \mathcal{T}_{\zeta_0}$ when ζ is a limit ordinal, we define the functions f_0, f_1, f_2 over an increasingly large initial segment of $\aleph_1$. We claim, however, that this recursion must terminate at some stage, and thus move us into step 2. If, otherwise, the recursion does not terminate, then the functions f_0, f_1, f_2 are defined on all of $\aleph_1$. Since the codomains, $\mathrm{p}_*(W)$ and $\{L, R\}$, of f_0 and f_1 are countable, there exists some $\lambda \in \mathrm{p}_*(W)$ and, say, $R \in \{L, R\}$, and uncountable $J \subseteq \aleph_1$ such that $f_0(J) = \{\lambda\}$ and $f_1(J) = \{R\}$. Suppose that $\zeta_0, \zeta_1 \in J$ are such that $\zeta_0 < \zeta_1$. Then by construction, at step ζ_1 we see that $f_2(\zeta_1)$ is an interval in $\mathrm{p}_*(W)$ with right endpoint λ which is larger than any interval I in $\mathrm{p}_*(W)$ with $\lambda = \max(I)$ and $W \upharpoonright_{\mathrm{p}} I \in \mathrm{Pfine}(h(\mathcal{T}_{\zeta_1}))$. As $W \upharpoonright_{\mathrm{p}} f_2(\zeta_0) \in \mathrm{Pfine}(\mathcal{T}_{\zeta_0+1}) \subseteq \mathrm{Pfine}(\mathcal{T}_{\zeta_1})$ we get that $f_2(\zeta_0)$ is strictly included into $f_2(\zeta_1)$. But as J is well ordered under the restriction of the order on $\aleph_1$ we let $s(\zeta)$ denote the successor of $\zeta \in J$ in J and select $\lambda_\zeta \in f_2(s(\zeta)) \setminus f_2(\zeta)$, giving us an injection from the uncountable set J to the countable set $\mathrm{p}_*(W)$, contradiction.

Step 2. From step 1 we obtain a coherent collection $\mathcal{T}_\zeta$ of cois, with $|\mathcal{T}_\zeta| < 2^{\aleph_0}$, and each $\lambda \in \mathrm{p}_*(W)$ includes into a maximal interval $I_\lambda \subseteq \mathrm{p}_*(W)$ with respect to the property that $W \restriction_{\mathrm{p}} I_\lambda \in \mathrm{Pfine}(h(\mathcal{T}_\zeta))$. Note that it is possible that $I_\lambda = \{\lambda\}$ for each $\lambda \in \mathrm{p}_*(W)$. The collection Λ of all such maximal intervals has a natural induced ordering and is necessarily order dense, for if there existed distinct I_λ and $I_{\lambda'}$ between which there are no elements in Λ then the word $W \restriction_{\mathrm{p}} I_\lambda \cup I_{\lambda'}$ would be in $\mathrm{Pfine}(h(\mathcal{T}_\zeta))$, contradicting maximality. As W is not the empty word we know that $\Lambda \neq \varnothing$. If Λ is a singleton then $\Lambda = \{\mathrm{p}_*(W)\}$, so $W \in \mathrm{Pfine}(\mathcal{T}_\zeta)$, so by Lemma 3.17 select $U \in \mathrm{Red}_{\kappa_1}$ and ι such that $\mathcal{T}_\zeta \cup \{\mathrm{coi}(W, \iota, U)\}$ is coherent.

If Λ is not a singleton let Λ' be the interval in Λ which excludes $\min(\Lambda)$ and $\max(\Lambda)$ if either or both exist. If Λ' is not empty then it is order isomorphic to $\mathbb{Q}$, and in either case by Proposition 3.21 we may add, if necessary, a single coi triple to $\mathcal{T}_\zeta$ to obtain a coherent collection $\mathcal{T}'_\zeta$ such that $W \restriction_{\mathrm{p}} \left(\bigcup \Lambda'\right) \in \mathrm{Pfine}(h(\mathcal{T}'_\zeta))$. Next, since $W \restriction_{\mathrm{p}} \min(\Lambda)$, $W \restriction_{\mathrm{p}} \max(\Lambda) \in \mathrm{Pfine}(h(\mathcal{T}'_\zeta))$ if either of $\min(\Lambda)$ or $\max(\Lambda)$ exists, we have that $W \in \mathrm{Pfine}(h(\mathcal{T}'_\zeta))$ as W is the concatenation of one or two or three words in $\mathrm{Pfine}(h(\mathcal{T}'_\zeta))$. By Lemma 3.17 we select $U \in \mathrm{Red}_{\kappa_1}$ and a coi ι such that $\mathcal{T}'_\zeta \cup \{\mathrm{coi}(W, \iota, U)\}$ is coherent. Then $\{\mathrm{coi}(W_x, \iota_x, U_x)\}_{x \in X} \cup \{\mathrm{coi}(W, \iota, U)\}$ is coherent and our proposition is proved. □

Proof of Theorem A. Let κ be a cardinal such that $2 \leq \kappa \leq 2^{\aleph_0}$. It is easy to see from Theorem 2.11 that $|\mathrm{Red}_2| = |\mathrm{Red}_\kappa| = 2^{\aleph_0}$. Thus we let $\prec$ well-order Red_2 in such a way that each element has fewer than $2^{\aleph_0}$ predecessors. Similarly let $\prec'$ well-order Red_κ in such a way that each element has fewer than $2^{\aleph_0}$ predecessors. We inductively define a coherent collection $\{\mathrm{coi}(W_\zeta, \iota_\zeta, U_\zeta)\}_{\zeta < 2^{\aleph_0}}$ of coi triples from Red_2 to Red_κ.

Recall that each ordinal ζ may be written uniquely as an ordinal sum $\zeta = \beta + m$ where β is either 0 or a limit ordinal and $m \in \omega$, and so ζ can be considered even or odd depending on the parity of m. Select a word $W_0 \in \mathrm{Red}_2$ minimal under $\prec$ and by Proposition 3.23 select $U_0 \in \mathrm{Red}_\kappa$ and a coi ι_0 such that $\{\mathrm{coi}(W_0, \iota_0, U_0)\}$ is coherent. Suppose that we have defined coherent $\{\mathrm{coi}(W_\zeta, \iota_\zeta, U_\zeta)\}_{\zeta < \mu}$ for all $\mu < \nu < 2^{\aleph_0}$. By Lemma 3.12 we know $\{\mathrm{coi}(W_\zeta, \iota_\zeta, U_\zeta)\}_{\zeta < \nu}$ is coherent. If ν is even then by Lemma 3.19 we select a word $W_\nu \notin \mathrm{Pfine}(\{W_\zeta\}_{\zeta < \nu})$ which is minimal under $\prec$ and by Proposition 3.23 select $U_\nu \in \mathrm{Red}_\kappa$ and a coi ι_ν such that $\{\mathrm{coi}(W_\zeta, \iota_\zeta, U_\zeta)\}_{\zeta < \nu+1}$ is coherent (using $\kappa_0 = 2$ and $\kappa_1 = \kappa$). Similarly if ν is odd then by Lemma 3.19 we select a word $U_\nu \notin \mathrm{Pfine}(\{U_\zeta\}_{\zeta < \nu})$ which is minimal under $\prec'$ and by Proposition 3.23 select $W_\nu \in \mathrm{Red}_\kappa$ and a coi ι_ν such that $\{\mathrm{coi}(W_\zeta, \iota_\zeta, U_\zeta)\}_{\zeta < \nu+1}$ is coherent (using $\kappa_0 = \kappa$ and $\kappa_1 = 2$).

Notice that $\mathrm{Pfine}(\{W_\zeta\}_{\zeta < 2^{\aleph_0}}) = \mathrm{Red}_2$ and $\mathrm{Pfine}(\{U_\zeta\}_{\zeta < 2^{\aleph_0}}) = \mathrm{Red}_\kappa$. Thus by Proposition 3.16 we have an isomorphism $\Phi : \mathcal{C}_2 \to \mathcal{C}_\kappa$. □

We will derive Theorem B as a consequence of Theorem A. Instead of defining the notions of elementary equivalence and elementary subsumption, we will trust the reader to know these concepts or to look them up. We will rely on the following classical result.

Lemma 3.24. *Suppose $\mathcal{U}_0$ is a submodel of $\mathcal{U}_1$ such that for every $a_0, \dots, a_{n-1} \in U_0$ and $a \in U_1$ there exists an automorphism $\phi : \mathcal{U}_1 \to \mathcal{U}_1$ such that $\phi(a_i) = a_i$ for all $i < n$ and $\phi(a) \in U_0$. Then $\mathcal{U}_0$ is an elementary submodel of $\mathcal{U}_1$.*

Proof of Theorem B. Certainly if $\gamma = \kappa$ or if $2 \leq \gamma \leq \kappa \leq 2^{\aleph_0}$ then we have $\mathcal{C}_\gamma \simeq \mathcal{C}_\kappa$ (using Theorem A in the second case) and the isomorphism is an elementary embedding. We may therefore assume that $2^{\aleph_0} \leq \gamma < \kappa$, for the result will follow for $2 \leq \gamma < 2^{\aleph_0} < \kappa$ as well by the fact that $\mathcal{C}_\gamma \simeq \mathcal{C}_{2^{\aleph_0}}$ in this case.

The map $\psi_{\gamma,\kappa} : \mathcal{C}_\gamma \to \mathcal{C}_\kappa$ given by $[[W]] \mapsto [[W]]$ is easily seen to be an injection and we consider $\mathcal{C}_\gamma$ as the substructure of $\mathcal{C}_\kappa$ consisting of those $[[W]]$ which have a representative utilizing only letters with first coordinate less than γ. Any bijection $f : \kappa \to \kappa$ induces a bijection $F_f : \mathcal{A}_\kappa \to \mathcal{A}_\kappa$ given by $a_{\alpha,n}^{\pm 1} \mapsto a_{f(\alpha),n}^{\pm 1}$ which induces a bijection $\mathcal{F}_f : \mathcal{W}_\kappa \to \mathcal{W}_\kappa$ given by $W \mapsto \prod_{i \in \overline{W}} F_f(W(i))$. This $\mathcal{F}_f$ induces an automorphism $\theta_f : \mathrm{Red}_\kappa \to \mathrm{Red}_\kappa$ given by $W \mapsto \mathcal{F}_f(W)$ which descends to an automorphism $\overline{\theta_f} : \mathcal{C}_\kappa \to \mathcal{C}_\kappa$.

Lemma 3.25. *Suppose $\gamma \leq \kappa$ with γ uncountable. If $X \subseteq \mathcal{C}_\gamma$ and $Y \subseteq \mathcal{C}_\kappa$ with $|X|, |Y| < \gamma$ there exists a bijection $f : \kappa \to \kappa$ such that $\overline{\theta_f}(x) = x$ for all $x \in X$ and $\overline{\theta_f}(Y) \subseteq \mathcal{C}_\gamma$.*

Proof. Assume the hypotheses. For each $x \in X$ fix a representative $W_x \in x$ such that $\mathrm{proj}_0(W) \subseteq \gamma$. For each $y \in Y$ fix a representative W_y. Since each set $\mathrm{proj}_0(W_x)$ is at most countable, the set $\bigcup_{x \in X} \mathrm{proj}_0(W_x)$ is of cardinality at most $\aleph_0 \cdot |X|$. Similarly the set $\bigcup_{y \in Y} \mathrm{proj}_0(W_y)$ is of cardinality at most $\aleph_0 \cdot |Y|$.

Since γ is uncountable, $\bigcup_{x \in X} \mathrm{proj}_0(W_x) \subseteq \gamma$ is of cardinality less than γ and $\bigcup_{y \in Y} \mathrm{proj}_0(W_y) \subseteq \kappa$ is also of cardinality less than γ, we can easily select a bijection $f : \kappa \to \kappa$ which fixes the elements in $\bigcup_{x \in X} \mathrm{proj}_0(W_x)$ and such that $f\big(\bigcup_{y \in Y} \mathrm{proj}_0(W_y)\big) \subseteq \gamma$. The automorphism $\overline{\theta_f}$ satisfies the desired properties. □

The proof of Theorem B is now complete by appealing to Lemma 3.24. □

Note that the map $f \mapsto \overline{\theta_f}$ gives a homomorphic injection from the full symmetric group on the set κ, S_κ, to the automorphism group $\mathrm{Aut}(\pi_1(\mathbb{GS}_\kappa))$. Since $\pi_1(\mathbb{GS}_2) \simeq \pi_1(\mathbb{GS}_{2^{\aleph_0}})$ we immediately get the following, which is not obvious a priori:

Corollary 3.26. *The group* $\mathrm{Aut}(\pi_1(\mathbb{GS}_2))$ *includes a subgroup isomorphic to the full symmetric group $S_{2^{\aleph_0}}$ on a set of size continuum.*

Acknowledgement

The author thanks Jeremy Brazas for the beautiful picture used in this article. Deep gratitude is also due to the referees whose careful reading and comments made this paper much more understandable.

References

[Bogley and Sieradski 2000] W. A. Bogley and A. J. Sieradski, "Weighted combinatorial group theory and wild metric complexes", pp. 53–80 in *Groups — Korea '98* (Pusan, 1998), edited by Y. G. Baik et al., de Gruyter, Berlin, 2000. MR Zbl

[Bogopolski and Zastrow 2012] O. Bogopolski and A. Zastrow, "The word problem for some uncountable groups given by countable words", *Topology Appl.* **159**:3 (2012), 569–586. MR Zbl

[Brazas and Gillespie 2022] J. Brazas and P. Gillespie, "Infinitary commutativity and abelianization in fundamental groups", *J. Aust. Math. Soc.* **112**:3 (2022), 289–311. MR Zbl

[Cannon and Conner 2000] J. W. Cannon and G. R. Conner, "The combinatorial structure of the Hawaiian earring group", *Topology Appl.* **106**:3 (2000), 225–271. MR Zbl

[Conner 2011] G. R. Conner, "Some interesting open problems in low-dimensional wild topology", conference talk, Workshop on Topology of Wild Spaces and Fractals, 2011, available at http://dmg.tuwien.ac.at/dorfer/wild_topology/abstracts.pdf.

[Conner et al. 2015] G. R. Conner, W. Hojka, and M. Meilstrup, "Archipelago groups", *Proc. Amer. Math. Soc.* **143**:11 (2015), 4973–4988. MR Zbl

[Corson 2021] S. M. Corson, "The Griffiths double cone group is isomorphic to the triple", preprint article, 2021. arXiv 2012.06794

[Corson 2023] S. M. Corson, "The nonabelian product modulo sum", preprint article, 2023. arXiv 2206.02682

[Eda 1992] K. Eda, "Free σ-products and noncommutatively slender groups", *J. Algebra* **148**:1 (1992), 243–263. MR Zbl

[Eda 1999] K. Eda, "Free subgroups of the fundamental group of the Hawaiian earring", *J. Algebra* **219**:2 (1999), 598–605. MR Zbl

[Eda and Fischer 2016] K. Eda and H. Fischer, "Cotorsion-free groups from a topological viewpoint", *Topology Appl.* **214** (2016), 21–34. MR Zbl

[Griffiths 1954] H. B. Griffiths, "The fundamental group of two spaces with a common point", *Quart. J. Math. Oxford Ser.* (2) **5** (1954), 175–190. MR Zbl

[Herfort and Hojka 2017] W. Herfort and W. Hojka, "Cotorsion and wild homology", *Israel J. Math.* **221**:1 (2017), 275–290. MR Zbl

[Higman 1952] G. Higman, "Unrestricted free products, and varieties of topological groups", *J. London Math. Soc.* **27** (1952), 73–81. MR Zbl

[Kunen 1980] K. Kunen, *Set theory*: *An introduction to independence proofs*, Studies in Logic and the Foundations of Mathematics **102**, North-Holland, Amsterdam, 1980. MR Zbl

[Zastrow 1994] A. Zastrow, "The second van-Kampen theorem for topological spaces", *Topology Appl.* **59**:3 (1994), 201–232. MR Zbl

Received April 29, 2021. Revised January 19, 2024.

SAMUEL M. CORSON
SCHOOL OF MATHEMATICS
UNIVERSIDAD DEL PAÍS VASCO
LEIOA
SPAIN

sammyc973@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 327, No. 2, 2023
https://doi.org/10.2140/pjm.2023.327.337

ARITHMETIC MODULAR LINKS

TALI PINSKY, JESSICA S. PURCELL AND
JOSÉ ANDRÉS RODRÍGUEZ-MIGUELES

We construct a sequence of geodesics on the modular surface such that the complements of the canonical lifts to the unit tangent bundle are arithmetic 3-manifolds.

1. Introduction

The modular group $\mathrm{PSL}(2, \mathbb{Z})$ is one of the simplest examples of an arithmetic group. The quotient of the upper half plane by the modular group is called the modular surface Σ_{Mod}; it is an arithmetic hyperbolic 2-dimensional orbifold.

One dimension higher, arithmetic hyperbolic 3-manifolds and 3-orbifolds form families of manifolds with very rich structure. They are also quite special. For example, among knot complements, only the figure-8 knot is arithmetic [22], and there exist closed orientable 3-manifolds that do not contain a simple closed curve with arithmetic complement [2]. However, every closed orientable 3-manifold contains an arithmetic link [15].

Associated with each oriented closed geodesic γ on the modular surface is a 3-manifold. This is obtained by lifting the geodesic γ into the unit tangent bundle over the modular surface $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$ to obtain a corresponding periodic orbit of the geodesic flow $\widehat{\gamma}$ called the *canonical lift*. The 3-manifold is the complement of $\widehat{\gamma}$ in the unit tangent bundle.

By Thurston's hyperbolisation theorem, the complement of a canonical lift of a closed modular geodesic will always be hyperbolic; see Foulon and Hasselblatt [12]. What is unknown in general is whether it will be arithmetic for some cases, and if so, what topological, geometric, and algebraic properties of the geodesic yield arithmeticity.

In this paper, we find an explicit family of canonical lift complements that are arithmetic.

Theorem 1.1. *There exists a sequence $\{\gamma_n\}_{n\in\mathbb{N}}$ of distinct closed geodesics on the modular surface such that for each n, the union of the first n canonical lifts $\bigcup_{j=1}^{n} \widehat{\gamma}_j$*

MSC2020: primary 57K32; secondary 11S20.
Keywords: arithmetic 3-manifolds, modular surface, closed geodesics.

has complement in the unit tangent bundle $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$ *that is an arithmetic hyperbolic* 3*-manifold* $\mathrm{UT}(\Sigma_{\mathrm{Mod}}) \setminus \bigcup_{j=1}^{n} \widehat{\gamma_j}$*, obtained by gluing regular ideal octahedra.*

Note that for $n > 1$, the manifolds of Theorem 1.1 are complements of more than one geodesic. When $n = 1$, the theorem produces a 3-manifold homeomorphic to the Whitehead link complement, which is well known to be arithmetic [18, § 4.5]. This corresponds to $\mathrm{UT}(\Sigma_{\mathrm{Mod}}) \setminus \widehat{\gamma_0}$ for γ_0 the shortest geodesic on the modular surface. It is an open question as to whether this is the only arithmetic canonical lift complement of a single geodesic on the modular surface.

The theorem is proved by considering canonical lifts of geodesics on a once-punctured torus, which is a six-fold cover of the modular surface. In Theorem 4.2 below, we build an explicit family of geodesics on the punctured torus and we prove that their canonical lifts are built of regular ideal octahedra. Such manifolds are always arithmetic, and the main theorem follows as arithmeticity is invariant under finite covers.

Because of the explicit nature of the construction, we are further able to obtain geometric information on these manifolds. For example, their volumes are given explicitly, and can be related to the lengths of the geodesics.

Corollary 1.2. *There exists a sequence* $\{\gamma_k\}_{k\in\mathbb{N}}$ *of closed geodesics on the modular surface with length* $\ell(\gamma_k) \nearrow \infty$ *such that, for* $\Gamma_n := \bigcup_{k=1}^{n} \gamma_k$,

(1) $\mathrm{UT}(\Sigma_{\mathrm{Mod}}) \setminus \widehat{\Gamma}_n$ *is arithmetic,*
(2) $\mathrm{Vol}(\mathrm{UT}(\Sigma_{\mathrm{Mod}}) \setminus \widehat{\Gamma_n}) = n\, v_{\mathrm{oct}}/2$*, and*
(3) $\mathrm{Vol}(\mathrm{UT}(\Sigma_{\mathrm{Mod}}) \setminus \widehat{\Gamma_n}) \asymp \sqrt{\ell(\Gamma_n)}$.

Here v_{oct} *is the volume of a regular ideal octahedron.*

In Corollary 1.2, $\asymp$ means *coarsely equivalent*: there are constants A, B, C, and D such that

$$A\sqrt{\ell(\Gamma_n)} + B \leq \mathrm{Vol}(\mathrm{UT}(\Sigma_{\mathrm{Mod}}) \setminus \widehat{\Gamma_n}) \leq C\sqrt{\ell(\Gamma_n)} + D.$$

Note that others have related volume to length of geodesics. Bergeron, Pinsky, and Silberman showed that the volume is bounded by a constant times the length [5]. Rodríguez-Migueles showed that there is a sequence of geodesics such that the volume grows linearly in the length of the geodesics up to a logarithmic factor [23]. Upper and lower bounds were extended by Cremaschi and Rodríguez-Migueles [8]. Cremaschi, Rodríguez-Migueles and Yarmola related volumes of the canonical lifts of a pair of simple closed curves to the Weil–Petersson distance in Teichmüller space [9].

More generally, by taking finite covers, we obtain:

Corollary 1.3. *Let* $\Sigma_{g,r}$ *be an orientable punctured surface with any hyperbolic metric. Then there exists a sequence* $\{\Gamma_k\}_{k\in\mathbb{N}}$ *of filling finite sets of closed geodesics* *on* $\Sigma_{g,r}$ *with lengths* $\ell(\Gamma_k) \nearrow \infty$*, such that* $\mathrm{UT}(\Sigma_{g,r}) \smallsetminus \widehat{\Gamma_k}$ *is arithmetic for each* $k \in \mathbb{N}$ *and*

$$\mathrm{Vol}(\mathrm{UT}(\Sigma_{g,r}) \smallsetminus \widehat{\Gamma_k}) \asymp \sqrt{\ell(\Gamma_k)}.$$

2. Surfaces and unit tangent bundles

Let Σ be a hyperbolic surface or orbifold. The unit tangent bundle $\mathrm{UT}(\Sigma)$ consists of points of the form (x, v), where x lies on Σ, and v is a unit vector tangent to Σ at x. Given a smooth oriented curve γ on Σ, any point $x \in \gamma$ determines a point (x, v) in the unit tangent vector, by letting v be the unit vector at x pointing in the direction of γ. Then γ lifts to a embedded closed curve $\widehat{\gamma}$ in $\mathrm{UT}(\Sigma)$.

The modular surface. The *modular surface* is the quotient of $\mathbb{H}^2$ by the modular group $\mathrm{PSL}(2, \mathbb{Z})$. Background on the modular group can be found in many places, for example in the work of Series [26]; see also the work of Brandts, Pinsky, and Silberman [7]. We review a few relevant facts here.

Consider the upper half plane $\mathbb{H}^2$ with its hyperbolic metric. Let U be a rotation of π about the point i and let V be a rotation of $2\pi/3$ about the point $\frac{1}{2} + i\frac{\sqrt{3}}{2}$, permuting points $\infty, 1, 0$. These two rotations generate the modular group $\mathrm{PSL}(2, \mathbb{Z})$. As elements of $\mathrm{PSL}(2, \mathbb{Z})$, U and V have the form

$$U = \pm\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad V = \pm\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}.$$

The rotation V fixes the hyperbolic ideal triangle in $\mathbb{H}^2$ with vertices $0, 1, \infty$, while U maps it to an adjacent ideal triangle. Thus the orbit of this ideal triangle under $\mathrm{PSL}(2, \mathbb{Z})$ is an invariant tessellation of $\mathbb{H}^2$ by ideal triangles called the Farey tessellation. It has an ideal vertex at each point of $\mathbb{Q} \cup \infty$ on $\partial\mathbb{H}^2$.

The quotient of $\mathbb{H}^2$ by the modular group $\mathrm{PSL}(2, \mathbb{Z})$ is an orbifold that is a sphere with a cusp, a cone point of order three, and a cone point of order two. This is called the *modular surface* and denoted by Σ_{Mod}. A fundamental domain for Σ_{Mod} is given by one third of the $0, 1, \infty$ ideal triangle.

Elements of finite order in $\mathrm{PSL}(2, \mathbb{Z})$ are exactly the conjugates of $1, U, V, V^2$. Every element of infinite order is a finite word in U, V and $V^{-1} = V^2$, involving both letters. Conjugating, one may always obtain a word beginning with V or V^2 and ending with U. Thus, up to conjugation, any infinite-order element can be written in positive powers of $L = V^2U$ and $R = VU$ [13], where

$$L = \pm\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad R = \pm\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \tag{2.1}$$

A closed geodesic on the modular surface Σ_{Mod} is called a *modular geodesic*. Modular geodesics are in one-to-one correspondence with conjugacy classes of

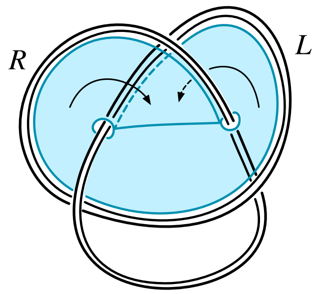

Figure 1. The branched surface inside the complement of the trefoil, with the direction of the semiflow indicated, pointing downwards from the branchline.

hyperbolic elements in $\mathrm{PSL}(2, \mathbb{Z})$, i.e., those with trace more than two. Note that R and L are parabolic elements, with trace two, but any word in positive powers in R and L involving both letters is hyperbolic.

A modular geodesic lifts to $\mathbb{H}^2$, tiled by the Farey tessellation. Series observed that such lifts cut out a sequence of triangles [26]. Within a given triangle an oriented geodesic enters through one side and then either exits through the side on its left (cutting off a single ideal vertex on its left side) or exits to its right. The sequence of rights and lefts determines a word in positive powers of R and L up to cyclic order called the *cutting sequence*. This agrees with the matrix product corresponding to the geodesic.

Now consider the unit tangent bundle of the modular surface, $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$. This is a Seifert fibred space whose base orbifold is Σ_{Mod}, with cone points of orders two and three and a cusp. In [19], Milnor proves that $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$ is homeomorphic to the complement of the trefoil in S^3, which proof he credits to Quillen. A neighbourhood of the cusp point of Σ_{Mod} lifts to give a neighbourhood of the trefoil. By the work of Ghys [13], for any finite collection of closed geodesics on the modular surface, their canonical lifts can be jointly isotoped in $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$ to lie on the branched surface shown in Figure 1. These are called *modular links*.

A modular link follows two lobes of the branched surface, one on the right and one on the left, and it is determined up to cyclic permutation by the word in the letters L and R. Thus the complement of a modular link corresponding to an n-component geodesic on the modular surface will be homeomorphic to the complement of a link in S^3 with $n+1$ components, with the additional component corresponding to the trefoil. Examples are shown at the end of Section 6.

The once-punctured torus. Begin with the closed torus with no punctures, which we will denote by $\Sigma_{1,0}$: the surface of genus one with zero punctures. Once we fix a choice of generators $\frac{1}{0}$ and $\frac{0}{1}$ for $\pi_1(\Sigma_{1,0})$, any simple closed curve on the torus is determined by an element of $\mathbb{Q} \cup \left\{\frac{1}{0}\right\}$. A geodesic representative of p/q has constant tangent vector; the curve lifts to a line of constant slope p/q in the universal cover $\mathbb{R}^2$.

The unit tangent bundle $\mathrm{UT}(\Sigma_{1,0})$ in this case is homeomorphic to $\Sigma_{1,0} \times S^1$. For ease of notation, we will write a point e^{it} in S^1 simply as t; in this form, two points in S^1 are equivalent if they differ by addition of an integer multiple of 2π. Then the canonical lift of a curve γ of slope p/q is a curve $\gamma \times \{\arctan(p/q)\}$ when oriented with tangent vector pointing towards $e^{i \arctan(p/q)}$ in $\mathbb{C}$. The curve has two orientations; when oriented in the opposite direction the canonical lift becomes $\gamma \times \{\arctan(p/q) + \pi\}$. Note that in either case, it has constant second coordinate. (This discussion needs to be modified for $p/q = \frac{1}{0}$; we leave that to the reader.)

Now consider the once-punctured torus, which we denote by $\Sigma_{1,1}$: the genus one surface with one puncture. Consider the abelian cover of the punctured torus; for now we view this as the plane $\mathbb{R}^2$ with integer lattice points removed. The line $y = 0$ in $\mathbb{R}^2$ projects to an arc μ on $\Sigma_{1,1}$ with both endpoints on the puncture. Similarly, the line $x = 0$ projects to an arc λ.

Consider those simple closed curves on the punctured torus that are parallel to lines in $\mathbb{R}^2$ of rational slope p/q, but disjoint from points on the integer lattice. These lines of rational slope project to closed curves in $\Sigma_{1,1}$ meeting μ a total of $|p|$ times, and meeting λ a total of $|q|$ times. We let p/q denote the closed curve. In particular, a closed curve parallel to μ is $\frac{0}{1}$, and one parallel to λ is $\frac{1}{0}$. Note these are not all the closed curves in $\Sigma_{1,1}$; we are omitting curves that wrap around the puncture in more complicated ways. However, these are the closed curves we will encounter in this paper.

Now consider the canonical lifts of such curves. The unit tangent bundle of the punctured torus is homeomorphic to the product $\Sigma_{1,1} \times S^1$. Just as for the closed torus, up to homeomorphism, the canonical lift of a curve of slope p/q in $\mathrm{UT}(\Sigma_{1,1})$ has the form $\gamma \times \{\arctan(p/q)\}$ oriented in one direction, or $\gamma \times \{\arctan(p/q) + \pi\}$ oriented in the other direction. That is, in either case we may isotope p/q in $\Sigma_{1,1}$ to have constant tangent vector.

In addition to the unit tangent bundle one may consider the projective tangent bundle $\mathrm{PT}(\Sigma_{1,1})$, where one quotients out by the action of ± 1 on S^1, i.e., antipodal points are identified. The unit tangent bundle is a degree-two cover of the projective tangent bundle. The two lifts of any fixed geodesic are identified in the quotient, and hence an unoriented closed geodesic has a unique lift to the projective tangent bundle. Its complement in the projective tangent bundle is covered via a degree-two covering map by the complement of both its lifts in the unit tangent bundle.

Lemma 2.2. *The punctured torus forms a* 6*-fold cover of the modular surface. The group of covering transformations is generated by a rotation of order three and a rotation of order two.*

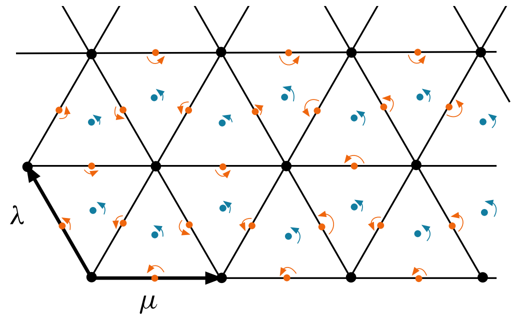

Figure 2. Taking the quotient of $\mathbb{R}^2 \smallsetminus \Lambda$ by translations gives $\Sigma_{1,1}$. Quotient further by $2\pi/3$ rotations about centres of triangles and π rotations about centres of edges to obtain Σ_{Mod}.

Similarly, the unit tangent bundle of the punctured torus forms a 6*-fold cover of the unit tangent bundle of the modular surface. The group of covering transformations is generated by two glide rotations of orders three and two.*

Proof. We will study the cover $\Sigma_{1,1} \to \Sigma_{\mathrm{Mod}}$ by considering first the abelian cover $\mathbb{R}^2 \smallsetminus \Lambda \to \Sigma_{1,1}$, where Λ is a lattice, and showing that Σ_{Mod} is obtained as a further quotient of this space.

Triangulate $\Sigma_{1,1}$ by adding the edges λ parallel to $\frac{1}{0}$ and μ parallel to $\frac{0}{1}$ as above, and an arc parallel to the slope $\frac{1}{1}$. This subdivides $\Sigma_{1,1}$ into two triangles, which we view as equilateral triangles. The abelian cover of $\Sigma_{1,1}$ can then be viewed as obtained by tiling $\mathbb{R}^2$ by these equilateral triangles, and removing all vertices to form the lattice Λ. We obtain $\Sigma_{1,1}$ by taking the quotient of $\mathbb{R}^2 \smallsetminus \Lambda$ by covering transformations that translate in the direction of μ and λ.

To obtain Σ_{Mod}, we quotient further, first by a rotation by $2\pi/3$, fixing the centre of one of the equilateral triangles and rotating its three vertices (the second triangle will also be rotated around its centre as a result), and then by a rotation by π, fixing the centre of an edge of an equilateral triangle and rotating that edge back to itself, swapping its endpoints and swapping the two triangles (note this will rotate simultaneously the other two edges about their centres). These two rotations generate a group of order 6, and the quotient is Σ_{Mod}. See Figure 2.

Now consider the unit tangent bundles. The unit tangent bundle $\mathrm{UT}(\Sigma_{1,1})$ is a trivial product, so it is covered by $(\mathbb{R}^2 \smallsetminus \Lambda) \times \mathbb{R}$. We obtain $\mathrm{UT}(\Sigma_{1,1})$ by taking the quotient by translations on $\mathbb{R}^2$ in the directions of μ and λ, and by a translation $(x, y, 0) \mapsto (x, y, 2\pi)$ in the $\mathbb{R}$ direction.

To obtain $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$, further quotient by a covering transformation of order three, and one of order two. The first is the glide rotation V that rotates an equilateral triangle in $\mathbb{R}^2$ by $2\pi/3$ about its centre, and translates it in the $\mathbb{R}$ direction by $2\pi/3$. Then V has order three in $\Sigma_{1,1} \times S^1 = \mathrm{UT}(\Sigma_{1,1})$. The second is the glide rotation U

that rotates $\mathbb{R}^2$ by π in the centre of an edge of an equilateral triangle, and shifts in the $\mathbb{R}$ direction by π. This has order two in $\mathrm{UT}(\Sigma_{1,1})$. Observe it takes the canonical lift of an oriented curve in $\Sigma_{1,1}$ to the canonical lift of the oppositely oriented curve.

We claim that the quotient of $\mathrm{UT}(\Sigma_{1,1})$ by U and V is $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$. To see this, note that the quotient is Seifert fibred, with base orbifold a sphere with one cusp, one cone point of order two, and one cone point of order three. This is homeomorphic to $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$. $\square$

Remark 2.3. More generally, any orientable hyperbolic surface with at least one puncture can be tiled by ideal triangles. There is then a hyperbolic structure that allows us to identify its fundamental domain with a finite portion of the Farey tessellation of $\mathbb{H}^2$. Since the modular group $\mathrm{PSL}_2(\mathbb{R})$ is the full symmetry group of the tessellation, this yields a representation of the surface's fundamental group as a subgroup of the modular group of finite index, and the surface is therefore a branched cover of Σ_{Mod}. Thus one can consider lifts of modular geodesics to any such surface and, as the unit tangent bundle is always trivial in this case, if the lift is simple the situation will be similar.

Curves on the once-punctured torus and the Farey tessellation. Isotopy classes of simple closed curves on the punctured torus are organised by the same Farey tessellation. Recall that the Farey complex can be considered as $\mathbb{H}^2$ with boundary $\mathbb{R} \cup \left\{\frac{1}{0}\right\}$. Isotopy classes of simple closed curves on $\Sigma_{1,1}$ correspond to points in $\mathbb{Q} \cup \left\{\frac{1}{0}\right\}$. The geometric intersection number of curves a/b and c/d is given by $|ad - bc|$. When a/b and c/d intersect exactly once, they correspond to an edge in the Farey complex: a hyperbolic geodesic running from a/b in $\mathbb{Q} \cup \left\{\frac{1}{0}\right\}$ to c/d in $\mathbb{Q} \cup \left\{\frac{1}{0}\right\}$. We say such curves are *Farey neighbours*. The matrix $\left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right)$ in $\mathrm{PSL}(2, \mathbb{Z})$ takes the edge between $\frac{1}{0}$ and $\frac{0}{1}$ to the edge between a/b and c/d in $\mathbb{H}^2$.

Definition 2.4. We say an ordered collection of simple closed curves $\alpha_1, \ldots, \alpha_n$ in $\Sigma_{1,1}$ are *Farey neighbours* if each α_j and α_{j+1} are connected by an edge of the Farey triangulation, for $j = 1, \ldots, n-1$, and if α_n and α_1 are also connected by an edge of the Farey triangulation.

3. Arithmetic Kleinian groups

Let K be a link in a compact 3-manifold with torus boundary. Suppose that the interior of the complement has a complete hyperbolic structure, meaning it is isometric to $\mathbb{H}^3/G$, where $\mathbb{H}^3$ is the hyperbolic 3-space and G is a torsion-free, noncocompact Kleinian group of finite covolume. The following definition of arithmeticity is a consequence of [18, Theorem 9.2.2].

Definition 3.1. A noncocompact Kleinian group is *arithmetic* if it is conjugate in $\mathrm{PSL}(2, \mathbb{C})$ to a group commensurable with $\mathrm{PSL}_2(O_d)$, where O_d is the ring of integers in the imaginary quadratic number field $\mathbb{Q}(\sqrt{-d})$, with d a positive integer. Such a group $\mathrm{PSL}(2, O_d)$ is called a *Bianchi group*. We say that a hyperbolic 3-manifold is *arithmetic* if the corresponding Kleinian group is arithmetic. Similarly, a knot or link with arithmetic complement is said to be arithmetic.

An example of an Bianchi group is the group $\mathrm{PSL}(2, \mathbb{Z}[i])$, called the Picard group. The Picard group is generated by face pairings of a fundamental region

$$\mathcal{F} = \left\{(x, y, t) \in \mathbb{H}^3 \mid x^2 + y^2 + t^2 \geq 1,\ -\tfrac{1}{2} \leq x \leq \tfrac{1}{2},\ 0 \leq y \leq \tfrac{1}{2}\right\};$$

see [18, § 1.4.1]. This is a quotient of a regular ideal octahedron. In fact, analogous to the two-dimensional case, $\mathbb{H}^3$ is tessellated by regular ideal octahedra, with ideal vertices at all points of $\mathbb{Q}[i]$. The Picard group $\mathrm{PSL}(2, \mathbb{Z}[i])$ is a subgroup of index two of the full symmetry group of this tessellation. Thus we have the following well-known result; see [1; 18, § 9.4; 20].

Lemma 3.2. *Any finite-volume hyperbolic* 3*-manifold obtained by gluing regular ideal octahedra is arithmetic.*

Note that arithmeticity is preserved by taking finite covers or quotients; any space that is finitely covered by such a space is also arithmetic.

4. Regular octahedra for neighbouring slopes

We now return to curves on the punctured torus $\Sigma_{1,1}$, and build arithmetic links in $\mathrm{UT}(\Sigma_{1,1})$.

Lemma 4.1. *Suppose α and β are two simple closed curves on the punctured torus $\Sigma_{1,1}$ that share an edge in the Farey triangulation. Let $N_{\alpha,\beta}$ denote the space obtained from $\Sigma_{1,1} \times [0, 1]$ by removing α from $\Sigma_{1,1} \times \{0\}$ and removing β from $\Sigma_{1,1} \times \{1\}$. Then $N_{\alpha,\beta}$ admits a complete hyperbolic structure obtained by gluing in pairs the eight faces of a regular ideal octahedron.*

Proof. When $\alpha = \frac{1}{0}$ and $\beta = \frac{0}{1}$, this is well known and is illustrated in Figure 3; see, for example, [16, Lemma 2.4]. On the left of that figure, $\Sigma_{1,1} \times [0, 1]$ is obtained by gluing the front face to the back, and the left face to the right.

On the right of the figure, observe that this gluing now identifies the front and back triangles opposite each other across the ideal vertex at the top of the octahedron, and the left and right triangles opposite each other across the ideal vertex at the bottom of the octahedron. If we give the ideal octahedron the hyperbolic geometry of a regular ideal octahedron, then each edge is identified to two edges of the ideal octahedron. The remaining unglued top and bottom faces become totally geodesic once-punctured annuli.

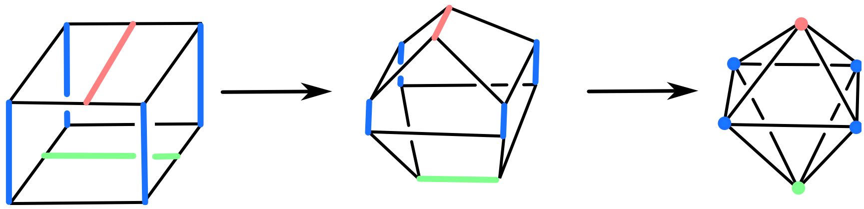

Figure 3. Starting on the left with $\Sigma_{1,1} \times [0, 1]$ with $\alpha = \frac{1}{0}$ drilled from $\Sigma_{1,1} \times \{0\}$ and $\beta = \frac{0}{1}$ drilled from $\Sigma_{1,1} \times \{1\}$, we obtain a regular ideal octahedron on the right.

For general $\alpha = p/q$ and $\beta = r/s$, α and β are Farey neighbours if $|ps - qr| = 1$. In this case there exists a homeomorphism from $N_{0,\infty=1/0}$ to $N_{\alpha,\beta}$ induced by the action of the linear automorphism $\left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right)$ taking $\Sigma_{1,1} \times \{t\}$ to $\Sigma_{1,1} \times \{t\}$ for all t, and taking $(\Sigma_{1,1} \times \{0\}) \smallsetminus \left\{\frac{1}{0}\right\}$ to $(\Sigma_{1,1} \times \{0\}) \smallsetminus \alpha$ and $(\Sigma_{1,1} \times \{1\}) \smallsetminus \left\{\frac{0}{1}\right\}$ to $(\Sigma_{1,1} \times \{1\}) \smallsetminus \beta$. This can be realised by a hyperbolic isometry. $\square$

Theorem 4.2. *Let $\alpha_1, \ldots, \alpha_n$ be simple closed curves in $\Sigma_{1,1}$ that are Farey neighbours. Drill $\Sigma_{1,1} \times S^1$ by removing α_j from $\Sigma_{1,1} \times \{j/n\}$. The resulting manifold has a complete hyperbolic structure obtained by gluing n regular ideal octahedra.*

Proof. Cut the drilled manifold along each surface $\Sigma_{1,1} \times \{j/n\}$. Obtain blocks of the form $N_{\alpha_j,\alpha_{j+1}}$. By Lemma 4.1, each of these can be given the hyperbolic structure of a regular ideal octahedron, with two top faces unglued and two bottom faces unglued.

Glue the top faces of $N_{\alpha_j,\alpha_{j+1}}$ to the bottom faces of $N_{\alpha_{j+1},\alpha_{j+2}}$ for $j = 1, \ldots, n$ modulo n. The gluing will be by the identity, along totally geodesic once-punctured annuli. These have a unique hyperbolic structure, hence the gluing is by isometry.

We claim this gives a complete hyperbolic structure on the original drilled manifold. The proof is by the Poincaré polyhedron theorem; see the work of Epstein and Petronio [11] for a careful exposition. The gluing identifies blocks $N_{\alpha_j,\alpha_{j+1}}$ top to bottom, yielding a manifold homeomorphic to the desired manifold. Under the gluing, each edge is 4-valent. Thus when edges are glued, the monodromy around any edge is the identity: formed by gluing four right-dihedral angles. This is sufficient to ensure that the manifold has a (possibly incomplete) hyperbolic structure. For completeness, notice that in the boundary of a horoball neighbourhood of any cusp, we identify a sequence of truncated neighbourhoods of the ideal vertices; these are squares. The squares are glued to obtain a tiling of the horospherical torus. Thus the regular ideal octahedra induce a Euclidean structure on each cusp. It follows that the hyperbolic metric obtained from the octahedra is a complete metric on the drilled manifold; see also [21, Theorem 4.10]. $\square$

We wish to apply Theorem 4.2 to a result about canonical lifts of Farey neighbours in the unit tangent bundle $\mathrm{UT}(\Sigma_{1,1})$. However, we need to take some care in orienting the curves. As noted above, each curve $\gamma = p/q$ has two orientations. For one orientation, the canonical lift $\widehat{\gamma}$ will lie in $\Sigma_{1,1} \times \{\arctan(p/q)\}$ and the other will lie in $\Sigma_{1,1} \times \{\arctan(p/q) + \pi\}$. The canonical lift $\widetilde{\gamma}$ to the projective tangent bundle $\mathrm{PT}\,\Sigma_{1,1}$ (which is the same trivial bundle $\Sigma_{1,1} \times S^1$) is well defined.

Theorem 4.3. *Let $\Gamma := \{\gamma_j = a_j/b_j\}_{j=1}^n$ be a collection of simple closed geodesics on the punctured torus made of Farey neighbours, with each γ_j oriented in the direction of* $\exp(i \arctan(a_j/b_j))$. *Let $\overline{\Gamma} := \{\overline{\gamma_j}\}_{j=1}^n$ be the same collection, with each curve oriented in the opposite direction. Then*:

(1) $\mathrm{UT}(\Sigma_{1,1}) \smallsetminus \widehat{\Gamma} \cong \mathrm{UT}(\Sigma_{1,1}) \smallsetminus \widehat{\overline{\Gamma}} \cong \mathrm{PT}(\Sigma_{1,1}) \smallsetminus \widetilde{\Gamma}$ *is arithmetic, obtained by gluing n regular ideal octahedra.*

(2) $\mathrm{UT}(\Sigma_{1,1}) \smallsetminus (\widehat{\Gamma} \cup \widehat{\overline{\Gamma}})$ *is arithmetic, obtained by gluing $2n$ regular ideal octahedra.*

Proof. Each $\gamma_j = a_j/b_j$ corresponds to a distinct slope in $\mathbb{Q} \cup \left\{\frac{1}{0}\right\}$. We may assume the b_j are nonnegative integers. By our orientation convention, each curve $\widehat{\gamma_j}$ will be drilled from $\Sigma_{1,1} \times \{\arctan(a_j/b_j)\} \subset \Sigma_{1,1} \times S^1$. Because Γ is a collection of Farey neighbours, there is some minimal slope in $\mathbb{Q}$, which we may relabel to be $\gamma_1 = a_1/b_1$, and then up to relabelling, the slopes satisfy $a_1/b_1 < a_2/b_2 < \cdots < a_n/b_n$. Then when we drill, the curves are drilled in cyclic order γ_1, γ_2, up to γ_n in the S^1 factor of $\Sigma_{1,1} \times S^1$. The drilling is therefore homeomorphic to the drilling of Theorem 4.2. Then the fact that $M_{\widehat{\Gamma}}$ is obtained by gluing n regular ideal octahedra follows from Theorem 4.2, and the fact that it is arithmetic follows from Lemma 3.2. An identical argument holds for $\overline{\Gamma}$.

For the union of $\widehat{\Gamma}$ and $\widehat{\overline{\Gamma}}$, the arithmeticity follows from the fact it is a double cover of $\mathrm{PT}(\Sigma_{1,1}) \smallsetminus \widehat{\Gamma}$. Furthermore, the first n canonical lifts will be at heights $\arctan(a_1/b_1) < \cdots < \arctan(a_n/b_n)$, and the next n at $\arctan(a_1/b_1) + \pi$ through $\arctan(a_n/b_n) + \pi$. Thus again we drill the Farey neighbours in an order homeomorphic to that of Theorem 4.2, and so that theorem implies that the complement is built of $2n$ regular ideal octahedra. ☐

5. Projecting and lifting on the modular surface

Lemma 5.1. *Let γ be an oriented geodesic on the modular surface Σ_{Mod}, obtained by projecting the simple closed curve $p/q \subset \Sigma_{1,1}$ via the covering map of Lemma 2.2. Then under the covering map, γ has six lifts in $\Sigma_{1,1}$. These are p/q, $q/(q-p)$, $(p-q)/p$, and each of these three curves oriented in the opposite direction: $\overline{p/q}$, $\overline{q/(q-p)}$, and $\overline{(p-q)/p}$.*

Proof. We consider the images of p/q under the rotations of order two and three of Lemma 2.2. As in the proof of that lemma, we will view $\Sigma_{1,1}$ as a quotient of the tiling of $\mathbb{R}^2$ by equilateral triangles with vertices removed.

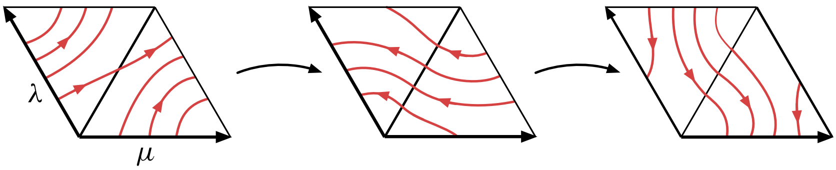

Figure 4. A rotation by $2\pi/3$ about the centre of each equilateral triangle takes the curve p/q to the curve $\overline{(p-q)/p}$, and a further rotation takes it to $\overline{q/(q-p)}$. Shown is the case $q > p > 0$. Similar pictures give other cases.

Recall that the rotation of order three rotates an ideal triangle, permuting its vertices. Consider its effect on the curve p/q. We may assume without loss of generality that $q \geq 0$. If $p \geq 0$, then the curve p/q meets the side μ of an equilateral triangle in the fundamental domain for $\Sigma_{1,1}$ a total of p times. It meets λ a total of q times, and meets the diagonal $|q-p|$ times. See Figure 4, which shows the case $q > p > 0$.

Rotating by $2\pi/3$ takes the curve to one meeting μ a total of $|q-p|$ times, meeting λ a total of $|p|$ times, and meeting the diagonal q times. In case $q > p > 0$, as shown in Figure 4, the resulting slope is negative, of value $\overline{(p-q)/p}$, and a further rotation gives the curve of slope $\overline{q/(q-p)}$. Our convention is to take an overline if the curve crosses lambda from right to left; we will see that in all cases we obtain each curve in both directions so this convention will not matter.

If $p > q > 0$, the result of the rotation is positive, of slope $\overline{(p-q)/p}$, and a further rotation results in a curve of slope $q/(q-p)$.

If $p < 0$ then the curve p/q meets μ a total of $|p|$ times, meets λ a total of q times, and meets the diagonal $|p|+q = q-p$ times. The resulting slopes after rotating are $(q-p)/p$ and $\overline{(p-q)/p}$.

Finally if one of p or q is zero, or $p = q = 1$, the three slopes up to rotation are $\frac{0}{1}$, $\frac{1}{0}$, and $\frac{1}{1}$, and the lemma holds for these.

Now consider the rotation of order two, with fixed point on an edge of the triangle. This takes the p/q curve back to itself, but it gives it the opposite orientation. This will give us the curve $\overline{p/q}$. Similarly it gives the other two curves with opposite orientations. Thus in all cases we obtain the set of both orientations of each of the slopes $\{p/q, q/(q-p), (p-q)/p\}$, as required. □

The following lemma shows that in lieu of rotating the closed geodesics and then considering the resulting slopes as above, one may instead directly rotate the slopes along the circle at infinity.

Lemma 5.2. *For any* $p/q \in \mathbb{Q} \cup \left\{\frac{1}{0}\right\}$, $V(p/q) = q/(q-p)$ *and* $V^2(p/q) = (p-q)/p$.

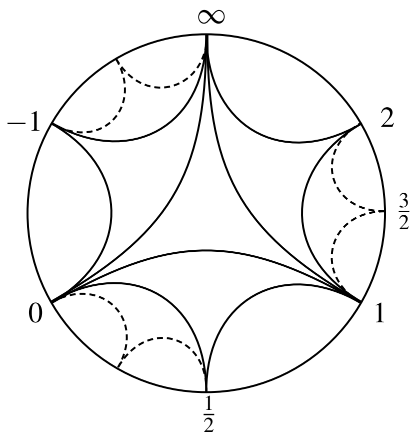

Figure 5. The Farey graph of rational slopes.

Proof. Recall that V has the form $\pm\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$. Then $V(p/q) = q/(q-p)$, and $V^2(p/q) = (p-q)/p$. $\square$

Observe that the first rotation shown in Figure 4 is V^2.

We will now turn a sequence of geodesics in the punctured torus into a sequence of geodesics on the modular surface. We start with an example, shown in Figure 5. Consider the $\frac{3}{2}$ curve. There is a shortest path from the Farey triangle with vertices $\left(\frac{1}{1}, \frac{0}{1}, \frac{1}{0}\right)$ to a Farey triangle with vertex $\frac{3}{2}$. The path meets three Farey triangles, with vertices $\left(\frac{1}{1}, \frac{0}{1}, \frac{1}{0}\right)$, $\left(\frac{1}{1}, \frac{2}{1}, \frac{1}{0}\right)$, and $\left(\frac{1}{1}, \frac{2}{1}, \frac{3}{2}\right)$. Form a collection of curves Γ by adding all the distinct slopes in all these triangles to Γ.

Thus Γ consists of $\frac{0}{1}$, $\frac{1}{0}$, $\frac{1}{1}$, $\frac{2}{1}$, and $\frac{3}{2}$. Note these are Farey neighbours, so Theorem 4.3 implies that the complement of their canonical lifts (oriented both ways) is an arithmetic manifold.

We wish to apply the covering projection from $\mathrm{UT}(\Sigma_{1,1})$ to $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$. However, note that the canonical lift of Γ does not cover any link complement in the unit tangent bundle of the modular surface, because Γ does not contain all the preimages of its projections to the modular surface. Thus we extend Γ, by including all images of Γ under the rotations V and V^2. Thus in the example of Figure 5, we would add $-\frac{1}{1} = V\left(\frac{2}{1}\right)$, $\frac{1}{2} = V^2\left(\frac{2}{1}\right)$, $-\frac{2}{1} = V\left(\frac{3}{2}\right)$ and $\frac{1}{3} = V^2\left(\frac{3}{2}\right)$. The result is a again a collection of Farey neighbours, and now the complement of all canonical lifts is a cover of the complement of a modular link. We generalise this example.

Theorem 5.3. *Any modular geodesic that lifts to a simple closed curve α on the once-punctured torus is part of an arithmetic link in* $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$ *with all components being modular geodesics. Suppose the shortest path in the Farey triangulation between the triangle* $(0, 1, \infty)$ *and any triangle with vertex α passes through x Farey triangles. Then the complement of the lift can be decomposed into x regular ideal octahedra.*

Proof. Given any slope p/q, there is a shortest path in the Farey triangulation from the centre of the triangle with vertices $\left(\frac{0}{1}, \frac{1}{1}, \frac{1}{0}\right)$ to a triangle with a vertex p/q. This will pass through some number of Farey triangles. Build a collection of curves Γ by adding all the slopes corresponding to all the vertices of the Farey triangles in the path. Thus Γ will contain $\frac{0}{1}, \frac{1}{1}, \frac{1}{0}$ and p/q, as well as additional curves at vertices of Farey triangles. At this step, Γ will contain a total of $2+x$ slopes: three corresponding to the first triangle $\left(\frac{0}{1}, \frac{1}{1}, \frac{1}{0}\right)$, and $x-1$ additional slopes, one for each new triangle in the path.

Next, expand Γ by adding all images of Γ under the rotations V and V^2 of Lemma 5.2. Note this adds $2(x-1)$ additional slopes to Γ, so that in total, Γ now contains $3x$ slopes.

Observe that the collection Γ can now be ordered in $\mathbb{Q} \cup \left\{\frac{1}{0}\right\}$ to give a set of Farey neighbours, invariant under the action of V. Theorem 4.3 then implies $\mathrm{UT}(\Sigma_{1,1}) \smallsetminus (\widehat{\Gamma} \cup \widehat{\overline{\Gamma}})$ is arithmetic, obtained by gluing $6x$ regular ideal octahedra. By Lemma 5.1, the drilled curves are exactly the canonical lifts of all curves projecting to a collection of x simple closed curves on the modular surface.

Now consider the action of the covering transformations of Lemma 2.2 from $\mathrm{UT}(\Sigma_{1,1})$ to $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$. By construction, the order-two transformation will take the canonical lift of p/q to that of $\overline{p/q}$. The order-three transformation will take the canonical lift of p/q to $V(p/q)$ and $V^2(p/q)$. Thus $\mathrm{UT}(\Sigma_{1,1}) \smallsetminus (\widehat{\Gamma} \cup \widehat{\overline{\Gamma}})$ is a six-fold cover of the complement of a collection of canonical lifts in $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$.

Finally, observe that each of the covering transformations maps a regular ideal octahedron to a distinct regular ideal octahedron. By the construction of Theorem 4.2, the regular ideal octahedra lie between canonical lifts that share an edge in the Farey triangulation. The covering transformation of degree three takes the octahedron between a/b and c/d to that between $V(a/b)$ and $V(c/d)$, and then again to that between $V^2(a/b)$ and $V^2(c/d)$; these are all distinct edges of the Farey triangulation. The covering transformation of degree two takes the octahedron between a/b and c/d to that between $\overline{a/b}$ and $\overline{c/d}$; this octahedron differs from the original by a rotation by π in the S^1 factor of $\mathrm{UT}(\Sigma_{1,1}) \cong \Sigma_{1,1} \times S^1$.

Then when we take the quotient by covering transformations, we obtain an arithmetic canonical link complement in $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$, with the link containing the original curve, and built from $6x/6 = x$ regular ideal octahedra. □

Theorem 1.1 from the introduction is an immediate consequence.

Corollary 5.4. *There are infinitely many arithmetic modular links.* □

6. Cutting sequences

As explained in Section 2, canonical lifts of geodesics in $\mathrm{UT}(\Sigma_{\mathrm{Mod}})$ can be viewed as links in $S^3 \smallsetminus K$ where K is the trefoil knot. In the previous section, we found infinitely many arithmetic canonical link complements. We wish to identify these links as the complement of links in the 3-sphere. To do so, we will find cutting sequences for the links, enabling us to identify them in the branched surface of Figure 1 following [13]. That is the main goal of this section.

Definition 6.1. Let α be a closed geodesic in the modular surface Σ_{Mod}. The *LR-cutting sequence* of α is the bi-infinite sequence of instances of L and R obtained as follows. Recall that a fundamental domain for Σ_{Mod} is the quotient of an ideal triangle by an order-three and an order-two rotation. As in the proof of Lemma 2.2, we take a cover of Σ_{Mod} that tiles $\mathbb{R}^2$ by equilateral triangles, and remove the lattice Λ consisting of the vertices of these triangles. Lift α to this cover. Consider a point of intersection of α with an edge of a triangle. Then in the adjacent triangle, α either runs next to the edge to the left or to the right. If it runs to the left, take the letter L. If it runs right, take the letter R. Now repeat for the next triangle, and so on. Because α is a closed geodesic, eventually α returns to an edge identified with the original edge of intersection, and the sequence will repeat.

Remark 6.2. Since different lifts of the geodesic α differ by an element of $\mathrm{PSL}(2, \mathbb{Z})$ which preserves the Farey tessellation by ideal triangles, the cutting sequence remains the same up to cyclic order no matter which lift of α we start with. By reversing the orientation of α if necessary we may always assume its cutting sequence begins with an L. We may always assume it enters the $0, 1, \infty$ triangle through the imaginary axis (oriented to the right) by using the rotation about i given by U above.

We can similarly define a cutting sequence for simple closed curves in $\Sigma_{1,1}$. Take a curve p/q with p/q positive, and lift to the abelian cover of $\Sigma_{1,1}$ that we build by tiling $\mathbb{R}^2$ with equilateral triangles, again as in the proof of Lemma 2.2. Lift p/q to this cover. The lift will intersect lifts of the arcs μ and λ. If it intersects μ, assign an instance of A. If it intersects λ, assign an instance of B. This gives an AB-cutting sequence for geodesics on $\Sigma_{1,1}$.

The following algorithm, from Series [25] and Davis [10, Algorithm 7.6], gives the AB-cutting sequence in terms of the continued fraction expansion of p/q.

Algorithm 6.3. (1) Start with an infinite string consisting of incidences of the letter A. This corresponds to a lift of a geodesic of slope 0. If $p/q \neq 0$, take a continued fraction expansion of the slope p/q of the form $[a_1, a_2, \dots, a_k]$ where all the a_j are positive.

(2) Insert a_k instances of the letter B between each pair of letters A. The corresponding trajectory now has slope a_k. If $p/q = a_k$ we are done.

(3) Else swap every A to B and vice-versa. The corresponding trajectory now has slope $1/a_k$. If $p/q = 1/a_k$ we are done.

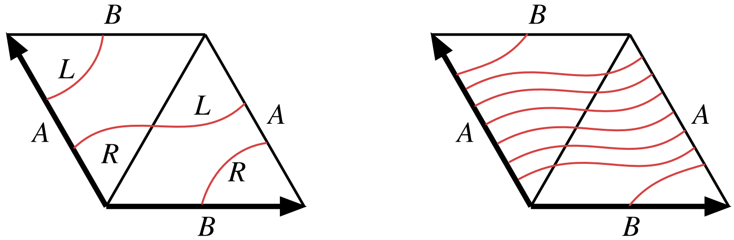

Figure 6. On the left is the general rule for determining the cutting sequence of a positive slope, on the right is the cutting sequence $LR(RL)^6$ corresponding to the projection of the geodesic of slope $\frac{1}{7}$.

(4) Else insert a_{k-1} instances of the letter B between each pair of letters A. The corresponding trajectory now has slope $a_{k-1} + 1/a_k$. If we have reached p/q we are done.

(5) Else reverse B and A. The corresponding trajectory now has slope

$$\frac{1}{a_{k-1} + \frac{1}{a_k}}.$$

If we have reached p/q we are done.

(6) Else continue this process, ending by inserting a_1 instances of the letter B between each pair of letters A. This yields the AB-cutting sequence corresponding to the fractional slope $[a_1, a_2, \dots, a_k]$.

We wish to find the LR-cutting sequence corresponding to a modular geodesic, and the AB-cutting sequence of Algorithm 6.3 for its lift to the once-punctured torus. By Remark 6.2 the lift we choose does not change the LR-cutting sequence and thus we may choose the lift to be a curve of slope p/q on $\Sigma_{1,1}$ where p and q are nonnegative. We can obtain the LR-cutting sequence corresponding to its projection as follows.

Algorithm 6.4. Let p/q be a slope, where p and q are both positive. Then for $j = 1, \dots, n-1$:

(1) If the j-th letter is A and the next letter is B, add L.

(2) If the j-th letter is B followed by A, add R.

(3) If the j-th letter is A followed by A, add RL.

(4) If the j-th letter is B followed by B, add LR.

If the slope is $\frac{0}{1}$ or $\frac{1}{0}$ (these are both lifts of the same modular geodesic) the cutting sequence is LR.

See Figure 6.

Example 6.5. Given a straight line of slope $1/n$, its AB-cutting sequence is the bi-infinite sequence given by concatenating copies of BA^n. Its LR-cutting sequence is the bi-infinite sequence given by concatenating $LR(RL)^{n-1}$.

Modular links. Now return to the arithmetic modular links of Theorem 5.3. We will construct examples of such links in the trefoil complement in the 3-sphere.

From the proof of that theorem, the links are obtained by adding curves from the Farey triangulation that are invariant under the rotation W that rotates $\frac{0}{1}$ to $\frac{1}{1}$, $\frac{1}{1}$ to $\frac{1}{0}$, and $\frac{1}{0}$ to $\frac{0}{1}$. The smallest collection of curves comes from the initial triangle $\frac{0}{1}$, $\frac{1}{1}$, and $\frac{1}{0}$. All three curves at the vertices of this triangle are identified when we project to Σ_{Mod}. Hence we may use any of the three curves to determine the modular link. We take $p/q = \frac{1}{1}$.

Then observe that the AB-cutting sequence in this case is simply obtained by concatenating copies of BA. By Algorithm 6.4, the LR-cutting sequence is then obtained by concatenating copies of LR (or equivalently RL). Therefore the modular geodesic corresponds to RL. In Figure 7, shown are the three distinct lifts of this geodesic in the parallelogram that is a fundamental domain for $\Sigma_{1,1}$. There are six lifts in total. As discussed above, the other three lifts traverse these curves in opposite directions. Note all six curves determine a cutting sequence RL or LR, which gives the same bi-infinite sequence.

Thus we have proved:

Lemma 6.6. *The modular geodesic RL is arithmetic.* □

The corresponding curve in the trefoil complement is obtained by drawing a closed curve on the branched surface of Figure 1. The cutting sequence LR instructs us that this curve must first run over the L lobe of the branched surface, then the R lobe, then close. This is shown on the left of Figure 8. Note that Lemma 6.6 is easily proved directly by the fact that its complement is homeomorphic to the Whitehead link complement as shown by the deformations of Figure 8.

Now consider the next simplest arithmetic modular link arising from the construction in the proof of Theorem 5.3. This is obtained by adding a single additional curve, coming from a new vertex of a Farey triangle of distance one from that with vertices $\frac{1}{0}$, $\frac{1}{1}$, and $\frac{0}{1}$, and then taking the image of this curve under the degree-three

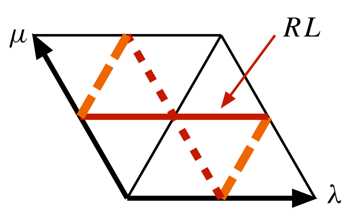

Figure 7. A fundamental domain for the two-dimensional torus, and three different lifts corresponding to the modular geodesic RL.

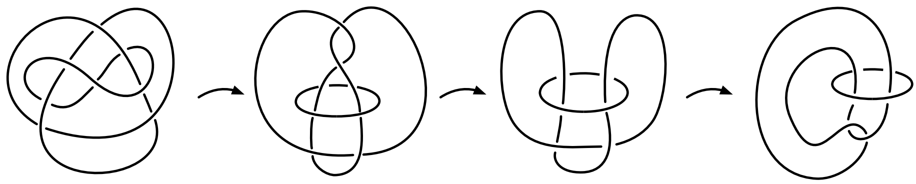

Figure 8. The homeomorphism between the complement of the RL geodesic and the Whitehead link complement.

rotation. We see from Figure 5 that the only possibility is to next include $\frac{2}{1}$, $-\frac{2}{1}$, and $\frac{1}{2}$, which are all identified in Σ_{Mod}.

In particular, the curve $\frac{1}{2}$ has AB-cutting sequence BAA, and LR-cutting sequence obtained by concatenating copies of $LRRL$, which is equivalent to L^2R^2. Thus in the trefoil complement, it runs twice over the L lobe of the branched surface, then twice over the R lobe, before closing up.

The link given by the union of LR and L^2R^2 is also arithmetic, by Theorem 5.3. It is shown on the left of Figure 9. This is a three-component link in S^3. As mentioned, any finite union of modular geodesics has an embedding as orbits on the template. We remark this embedding is unique, and can be found in general using an algorithm, for example, as in Birman and Williams [6, Algorithm 2.4.3]; see also Hui and Rodríguez-Migueles [17].

There are two possibilities for a four-component link in S^3 that arises from Theorem 5.3. One choice is to add slopes $\frac{3}{2}$, $-\frac{2}{1}$, and $\frac{1}{3}$, which are identified to a modular curve with LR-cutting sequence with repeating portion LRL^2R^2. Thus the four-component arithmetic link in S^3 consists of the trefoil and the geodesics LR, L^2R^2, and LRL^2R^2. This link is shown in the middle of Figure 9.

The other option is to add slopes $\frac{3}{1}$, $-\frac{1}{2}$, and $\frac{2}{3}$, which are identified to a modular curve with LR-cutting sequence with repeating portion LR^2L^2R. Thus another four-component arithmetic link in S^3 consists of the trefoil, the link LR, L^2R^2, and LR^2L^2R.

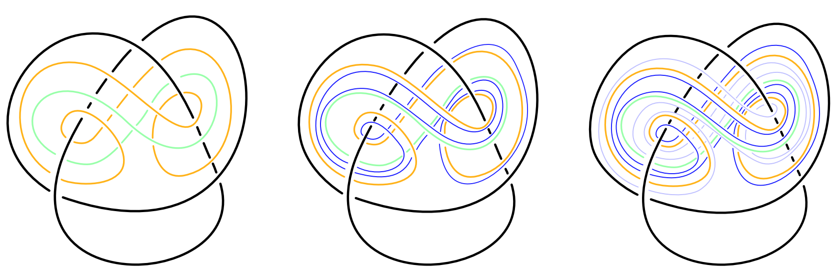

Figure 9. After the Whitehead link, the next three simplest arithmetic links from Theorem 5.3 are shown (note that in our conventions the symbol R corresponds to the left side of the figure).

Note that the five-component link consisting of the trefoil and the geodesics LR, L^2R^2, LR^2L^2R and LRL^2R^2 is also arithmetic by Theorem 5.3. This link is shown on the right of Figure 9.

7. Volume versus hyperbolic length

Our goal is to make explicit the relationship between volume of the canonical lift complement and geometric length of the original geodesic, for some sequence of geodesics in some surfaces.

Remark 7.1. Recall that for $A \in \mathrm{PSL}(2, \mathbb{R})$ a hyperbolic element of trace t, the eigenvalues of A are $(-t \pm \sqrt{t^2-4})/2$. Let λ_A be the eigenvalue satisfying $|\lambda_A| > 1$. Then the length of the closed geodesic determined by A is $2\ln|\lambda_A|$.

Lemma 7.2. *Let γ_n be the unique closed geodesic on the modular surface lifting to the geodesic $1/n$ on $\Sigma_{1,1}$. For $\Gamma_n := \{\gamma_i\}_{i=1}^n$, the length $\ell(\Gamma_n)$ satisfies*

$$\ell(\Gamma_n) \asymp n^2.$$

Proof. The matrix representative corresponding to $1/n$ is $A_n := LR(RL)^{n-1}$; see Example 6.5. Let

$$\begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix} := (RL)^{n-1}, \quad \text{so} \quad \begin{pmatrix} a_{n+1} & b_{n+1} \\ c_{n+1} & d_{n+1} \end{pmatrix} = \begin{pmatrix} a_n+c_n & b_n+d_n \\ a_n+2c_n & b_n+2d_n \end{pmatrix},$$

and

$$A_{n+1} = \begin{pmatrix} 3a_n+4c_n & 3b_n+4d_n \\ 2a_n+3c_n & 2b_n+3d_n \end{pmatrix}.$$

Then $\frac{3}{2}\,\mathrm{Trace}\,A_{n-1} \le \mathrm{Trace}\,A_n \le 4\,\mathrm{Trace}\,A_{n-1}$. As $\mathrm{Trace}\,A_1 = 3$, by induction

$$\left(\tfrac{3}{2}\right)^n \le \mathrm{Trace}\,A_n \le 4^n.$$

The eigenvalue λ_n of A_n with $|\lambda_n| > 1$ is bounded by

$$\tfrac{1}{2}|\lambda_n| \le \tfrac{1}{2}\,\mathrm{Trace}\,A_n \le |\lambda_n|.$$

Thus the length of γ_n satisfies

$$n \ln \tfrac{3}{2} \le \ell(\gamma_n) \le 2n \ln 4,$$

and thus

$$n^2 \ln \tfrac{3}{2} \le \ell(\Gamma_n) \le 2n^2 \ln 4. \qquad \square$$

Corollary 7.3. *Let $\Gamma_k := \{\gamma_{1,n} = 1/n,\ \gamma_{2,n} = n/(n-1),\ \gamma_{3,n} = (1-n)/1\}_{n=1}^k$ be a collection of oriented simple closed geodesics on the once-punctured torus with a hyperbolic metric ρ. Then, for $\widehat{\Gamma_k}$, the canonical lifts of Γ_k,*

(1) $\mathrm{UT}(\Sigma_{1,1}) \smallsetminus \widehat{\Gamma_k}$ *is arithmetic,*

(2) $\mathrm{Vol}(\mathrm{UT}(\Sigma_{1,1})\smallsetminus\widehat{\Gamma_k}) = 3k\, v_{\mathrm{oct}}$, *and*

(3) $\mathrm{Vol}(\mathrm{UT}(\Sigma_{1,1})\smallsetminus\widehat{\Gamma_k}) \asymp \sqrt{\ell_\rho(\Gamma_k)}$.

Proof. Notice that Γ_k are Farey neighbours, so by Theorem 4.3, $\mathrm{UT}(\Sigma_{1,1})\smallsetminus\widehat{\Gamma_k}$ is arithmetic and

$$\mathrm{Vol}(\mathrm{UT}(\Sigma_{1,1})\smallsetminus\widehat{\Gamma_k}) = 3k\, v_{\mathrm{oct}}.$$

Observe that the geodesics $1/n$, $n/(n-1)$, $(1-n)/1$ project under the 6-fold cover of the modular surface to $LR(RL)^{n-1}$; see Example 6.5. Then by Lemma 7.2, the length of the projection of Γ_k to Σ_{Mod} is coarsely equivalent to k^2. Thus in the 6-fold cover $\Sigma_{1,1}$, the lengths satisfy

$$\ell_{\rho_{1,1}}(\Gamma_k) \asymp 6k^2,$$

where $\rho_{1,1}$ is the pullback metric induced on $\Sigma_{1,1}$ by the metric on the modular surface Σ_{Mod}. Then

$$\mathrm{Vol}(\mathrm{UT}(\Sigma_{1,1})\smallsetminus\widehat{\Gamma_k}) \asymp v_{\mathrm{oct}}\sqrt{3/2}\sqrt{\ell_{\rho_{1,1}}(\Gamma_k)}.$$

The proof of this result for any hyperbolic metric on the once-punctured torus follows from the fact that any pair of hyperbolic metrics on a hyperbolic surface are bilipschitz; see, for example, [5, Lemma 4.1]. □

By projecting the geodesics in Corollary 7.3 under the 6-fold cover to the modular surface we obtain the following result from the introduction.

Corollary 1.2. *There exists a sequence $\{\gamma_k\}_{k\in\mathbb{N}}$ of closed geodesics on the modular surface with length $\ell(\gamma_k) \nearrow \infty$ such that, for $\Gamma_n := \bigcup_{k=1}^n \gamma_k$,*

(1) $\mathrm{UT}(\Sigma_{\mathrm{Mod}})\smallsetminus\widehat{\Gamma}_n$ *is arithmetic,*

(2) $\mathrm{Vol}(\mathrm{UT}(\Sigma_{\mathrm{Mod}})\smallsetminus\widehat{\Gamma_n}) = n\, v_{\mathrm{oct}}/2$, *and*

(3) $\mathrm{Vol}(\mathrm{UT}(\Sigma_{\mathrm{Mod}})\smallsetminus\widehat{\Gamma_n}) \asymp \sqrt{\ell(\Gamma_n)}$.

Here v_{oct} is the volume of a regular ideal octahedron.

Corollary 1.3. *Let $\Sigma_{g,r}$ be an orientable punctured surface with any hyperbolic metric. Then there exists a sequence $\{\Gamma_k\}_{k\in\mathbb{N}}$ of filling finite sets of closed geodesics on $\Sigma_{g,r}$ with lengths $\ell(\Gamma_k) \nearrow \infty$, such that* $\mathrm{UT}(\Sigma_{g,r})\smallsetminus\widehat{\Gamma_k}$ *is arithmetic for each $k \in \mathbb{N}$ and*

$$\mathrm{Vol}(\mathrm{UT}(\Sigma_{g,r})\smallsetminus\widehat{\Gamma_k}) \asymp \sqrt{\ell(\Gamma_k)}.$$

Proof. By Remark 2.3 we can construct a finite (branched) covering map p from any orientable punctured hyperbolic surface $\Sigma_{g,r}$ of genus g with r punctures to the modular surface Σ_{Mod}.

Let $\widetilde{\Gamma}_k$ be the finite set of closed geodesics on Σ obtained as the preimage under p of the closed geodesics $\{\gamma_n\}_{n=1}^k$ of Lemma 7.2. By Lemma 3.2, $\mathrm{UT}(\Sigma_{g,r})\smallsetminus\widehat{\Gamma}_k$ is arithmetic. A similar estimation of the volume and lengths as in Corollary 7.3 gives

$$\mathrm{Vol}(\mathrm{UT}(\Sigma_{g,r})\smallsetminus\widehat{\Gamma}_k) \asymp \sqrt{\ell_\rho(\Gamma_k)},$$

where the length $\ell_\rho(\Gamma_k)$ is measured in the pullback metric $\Sigma_{g,r}$ induced by the metric on Σ_{Mod}. Again the proof of this result for any hyperbolic metric on $\Sigma_{g,r}$ follows from the fact that any pair of hyperbolic metrics on a hyperbolic surface are bilipschitz; see, for example, [5, Lemma 4.1]. □

8. Further questions

There is only one arithmetic knot complement in the 3-sphere, namely the figure-8 knot, due to Reid [22]. Is the modular geodesic LR the only modular geodesic with arithmetic complement of its canonical lift? Notice that the question has a negative answer in the general context of any knot in the complement of the trefoil. Hatcher found an example of an arithmetic two-component link, where one component is the trefoil knot, and the trace field is $\mathbb{Q}(\sqrt{-2})$; see Figure 17 in [14]. However, the unknotted component in Hatcher's example is not a canonical lift of a closed geodesic in the modular surface.

All arithmetic modular links produced in this paper are conjugate in $\mathrm{PSL}(2,\mathbb{C})$ to a group commensurable with $\mathrm{PSL}(2,\mathbb{Z}(\sqrt{-1}))$. Are there examples of arithmetic modular links conjugate to groups commensurable with $\mathrm{PSL}(2, O_d)$ for O_d a ring of integers in a different quadratic number field $\mathbb{Q}(\sqrt{-d})$? More generally, is some classification possible? For example, in the 3-sphere, there are infinitely many arithmetic links. However, Baker and Reid showed that there are only finitely many principal congruence link complements in the 3-sphere [3], where a noncompact finite-volume hyperbolic 3-manifold is principal congruence if it is isometric to $\mathbb{H}^3/\Gamma(I)$ where $\Gamma(I)=\ker\{\mathrm{PSL}(2, O_d)\to\mathrm{PSL}(2, O_d/I)\}$ for some ideal I in O_d. Baker, Goerner, and Reid have now enumerated all principal congruence link complements in the 3-sphere [4]. Is a similar classification possible for modular links?

Any closed geodesic on the modular surface naturally corresponds to a real quadratic extension of $\mathbb{Q}$ [24]. Does the arithmeticity of the complement of the corresponding canonical lift relate to this? For the examples in this paper, the quadratic field corresponding to the LR geodesic is $\mathbb{Q}(\sqrt{5})$. The geodesic L^2R^2 has quadratic field $\mathbb{Q}(\sqrt{2})$. The geodesics LR^2L^2L and LRL^2R^2 have the same length, and both correspond to the same quadratic field $\mathbb{Q}(\sqrt{221})$. In general, geodesics corresponding to different maximal ideals in the same quadratic field will have the same length.

Acknowledgements

The authors thank the anonymous referee for helpful suggestions that have improved the paper. They also thank Neil Hoffman for pointing them towards helpful references. This work began during a workshop at the Matrix Centre in Australia; the authors thank the organisers and the centre for hosting them. Purcell was partially supported by the Australian Research Council, grant DP210103136. Rodríguez-Migueles is also very grateful for discussions with Juan Souto and Connie On Yu Hui on related topics, and recognises the support by the Special Priority Programme SPP 2026 Geometry at Infinity funded by the DFG.

References

[1] R. Abe and I. R. Aitchison, "Geometry and Markoff's spectrum for $\mathbb{Q}(i)$, I", *Trans. Amer. Math. Soc.* **365**:11 (2013), 6065–6102. MR Zbl

[2] M. D. Baker and A. W. Reid, "Arithmetic knots in closed 3-manifolds", *J. Knot Theory Ramifications* **11**:6 (2002), 903–920. MR Zbl

[3] M. D. Baker and A. W. Reid, "Principal congruence link complements", *Ann. Fac. Sci. Toulouse Math.* (6) **23**:5 (2014), 1063–1092. MR Zbl

[4] M. D. Baker, M. Goerner, and A. W. Reid, "All principal congruence link groups", *J. Algebra* **528** (2019), 497–504. MR Zbl

[5] M. Bergeron, T. Pinsky, and L. Silberman, "An upper bound for the volumes of complements of periodic geodesics", *Int. Math. Res. Not.* **2019**:15 (2019), 4707–4729. MR Zbl

[6] J. S. Birman and R. F. Williams, "Knotted periodic orbits in dynamical systems, I: Lorenz's equations", *Topology* **22**:1 (1983), 47–82. MR Zbl

[7] A. Brandts, T. Pinsky, and L. Silberman, "Volumes of hyperbolic three-manifolds associated with modular links", *Symmetry* **11**:10 (2019), art. id. 1206.

[8] T. Cremaschi and J. A. Rodríguez-Migueles, "Hyperbolicity of link complements in Seifert-fibered spaces", *Algebr. Geom. Topol.* **20**:7 (2020), 3561–3588. MR Zbl

[9] T. Cremaschi, J. A. Rodríguez-Migueles, and A. Yarmola, "On volumes and filling collections of multicurves", *J. Topol.* **15**:3 (2022), 1107–1153. MR Zbl

[10] D. Davis, "Lines in positive genus: an introduction to flat surfaces", pp. 1–55 in *Dynamics done with your bare hands*, edited by F. Dal'Bo et al., Eur. Math. Soc., Zürich, 2016. MR Zbl

[11] D. B. A. Epstein and C. Petronio, "An exposition of Poincaré's polyhedron theorem", *Enseign. Math.* (2) **40**:1-2 (1994), 113–170. MR Zbl

[12] P. Foulon and B. Hasselblatt, "Contact Anosov flows on hyperbolic 3-manifolds", *Geom. Topol.* **17**:2 (2013), 1225–1252. MR Zbl

[13] E. Ghys, "Knots and dynamics", pp. 247–277 in *International Congress of Mathematicians*, *I*, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2007. MR Zbl

[14] A. Hatcher, "Hyperbolic structures of arithmetic type on some link complements", *J. London Math. Soc.* (2) **27**:2 (1983), 345–355. MR Zbl

[15] H. M. Hilden, M. T. Lozano, and J. M. Montesinos, "On knots that are universal", *Topology* **24**:4 (1985), 499–504. MR Zbl

[16] C. O. Y. Hui and J. S. Purcell, "On the geometry of rod packings in the 3-torus", preprint, 2022. arXiv 2212.04662

[17] C. O. Y. Hui and J. A. Rodríguez-Migueles, "Modular links: bunch algorithm and upper volume bounds", preprint, 2023. arXiv 2308.12847

[18] C. Maclachlan and A. W. Reid, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Mathematics **219**, Springer, 2003. MR Zbl

[19] J. Milnor, *Introduction to algebraic K-theory*, Annals of Mathematics Studies **72**, Princeton University Press, 1971. MR Zbl

[20] W. D. Neumann and A. W. Reid, "Notes on Adams' small volume orbifolds", pp. 311–314 in *Topology '90* (Columbus, OH, 1990), edited by B. Apanasov et al., Ohio State Univ. Math. Res. Inst. Publ. **1**, de Gruyter, Berlin, 1992. MR Zbl

[21] J. S. Purcell, *Hyperbolic knot theory*, Graduate Studies in Mathematics **209**, American Mathematical Society, Providence, RI, 2020. MR Zbl

[22] A. W. Reid, "Arithmeticity of knot complements", *J. London Math. Soc.* (2) **43**:1 (1991), 171–184. MR Zbl

[23] J. A. Rodríguez-Migueles, "Periods of continued fractions and volumes of modular knots complements", *J. Knot Theory Ramifications* **32**:10 (2023), art. id. 2350063. MR Zbl

[24] P. Sarnak, "Class numbers of indefinite binary quadratic forms", *J. Number Theory* **15**:2 (1982), 229–247. MR Zbl

[25] C. Series, "The geometry of Markoff numbers", *Math. Intelligencer* **7**:3 (1985), 20–29. MR Zbl

[26] C. Series, "The modular surface and continued fractions", *J. London Math. Soc.* (2) **31**:1 (1985), 69–80. MR Zbl

Received September 5, 2023. Revised January 26, 2024.

TALI PINSKY
DEPARTMENT OF MATHEMATICS
TECHNION – ISRAEL INSTITUTE OF TECHNOLOGY
HAIFA
ISRAEL
talipi@technion.ac.il

JESSICA S. PURCELL
SCHOOL OF MATHEMATICS
MONASH UNIVERSITY
CLAYTON VIC
AUSTRALIA
jessica.purcell@monash.edu

JOSÉ ANDRÉS RODRÍGUEZ-MIGUELES
MATHEMATISCHES INSTITUT
UNIVERSITÄT MÜNCHEN
MÜNCHEN
GERMANY
migueles@math.lmu.de

PACIFIC JOURNAL OF MATHEMATICS
Vol. 327, No. 2, 2023
https://doi.org/10.2140/pjm.2023.327.359

TENSOR TRIANGULATED CATEGORY STRUCTURES IN THE DERIVED CATEGORY OF A VARIETY WITH BIG (ANTI)CANONICAL BUNDLE

ANGEL TOLEDO

Let X be a smooth projective variety over $\mathbb{C}$ with big (anti)canonical bundle. It is known that for such X the Balmer spectrum of the tensor triangulated category of perfect complexes $\mathrm{Perf}(X)$, equipped with the derived tensor product $\otimes_X^{\mathbb{L}}$, recovers the space X. We study the possible tensor triangulated category structures one can put on $\mathrm{Perf}(X)$. As an application, we prove a monoidal version of the well-known Bondal–Orlov reconstruction theorem.

1. Introduction

Bondal and Orlov [2001] showed that if X is a smooth projective variety over $\mathbb{C}$ with ample (anti)canonical bundle then its bounded derived category $D^b(X)$ completely recovers the space. More precisely, they showed that:

Theorem 1.1 [Bondal and Orlov 2001, Theorem 2.5]. *Let X be an irreducible smooth projective variety with ample* (*anti*)*canonical bundle. If $D^b(X) \simeq D^b(Y)$ for some other smooth algebraic variety Y, then $X \cong Y$.*

This theorem came in contrast with the discovery by Mukai [1987] that for an abelian variety A, there exists an equivalence $D^b(A) \simeq D^b(\hat{A})$, as triangulated categories, between the bounded derived category of A and the bounded derived category of its dual $\hat{A}$.

This observation sparked the study of what are now called Fourier–Mukai partners of a given variety X, that is, those varieties which are triangulated equivalent to the bounded derived category of X.

Bondal and Orlov's reconstruction pointed out that a (birational) geometric condition on the variety can introduce some control on these derived equivalences, and with this in mind, Kawamata generalized this theorem for varieties with big (anti)canonical bundle, clarifying from a geometric point of view the role of this condition on the possible equivalence of derived categories. Namely he showed:

MSC2020: 14F08, 18G80.

Keywords: tensor triangulated category, Balmer spectrum, Bondal–Orlov reconstruction, derived categories.

Theorem 1.2 [Kawamata 2002, Theorem 1.4]. *Let X, Y be smooth projective varieties such that there is an equivalence*

$$\mathcal{F} : D^b(X) \xrightarrow{\simeq} D^b(Y),$$

as triangulated categories. Then:

(1) $\dim X = \dim Y$.

(2) *If the canonical divisor K_X is nef, so is K_Y and there is an equality in the numerical Kodaira dimensions $\nu(X)$ and $\nu(Y)$.*

(3) *If X is of general type, then X and Y are birational, and furthermore, there is a smooth projective variety $p : Z \to X$, $q : Z \to Y$ such that $p^* K_X \simeq q^* K_Y$.*

This theorem should be understood as a strong indication of a relationship between the birational geometry of a variety and its derived category.

On the other hand, Balmer showed [Balmer 2002; Bondal and Orlov 2001] that when equipped with the derived tensor product $\otimes_X^{\mathbb{L}}$, the derived category of perfect complexes $\mathrm{Perf}(X)$ of any coherent scheme X can recover the space X by what is now known as the Balmer spectrum $\mathrm{Spc}(\mathrm{Perf}(X), \otimes_X^{\mathbb{L}})$. The Balmer spectrum can be constructed for a general tensor triangulated category, that is, a triangulated category equipped with a compatible monoidal structure, and produce a locally ringed space.

The existence of nonisomorphic Fourier–Mukai partners Y for a smooth variety X implies, using the Balmer spectrum construction, that the bounded derived category $D^b(X)$ can be equipped with at least as many tensor triangulated category structures as nonisomorphic Fourier–Mukai partners, up to monoidal equivalence.

In other words, if $\mathrm{FM}(X)$ is the set of isomorphism classes of Fourier–Mukai partners of X and $\mathrm{TTS}(X)$ is the set of equivalence classes of tensor triangulated category structures on the bounded derived category $D^b(X)$, there exists an injection

$$\mathrm{FM}(X) \to \mathrm{TTS}(X), \quad Y \mapsto (\otimes_Y^{\mathbb{L}}, \mathscr{O}_Y),$$

where the pair $(\otimes_Y^{\mathbb{L}}, \mathscr{O}_Y)$ denotes the tensor triangulated category structure given by the derived tensor product $\otimes_Y^{\mathbb{L}}$ with unit $\mathscr{O}_Y$.

Our main interest in this work is the study of this function, its surjectivity and the properties that one can deduce about possible tensor triangulated category structures outside of the image of this injection, all under the condition that the (anti)canonical bundle of X is big.

In Section 2 we give a brief general overview of the results we will need about general derived categories of quasicoherent sheaves on a smooth projective variety, together with a reminder of the Balmer spectrum construction through Thomason's classification theorem.

In Section 3, given a tensor triangulated category structure $(D^b(X), \boxtimes, \mathbb{1})$ with unit $\mathbb{1}$ on a bounded derived category $D^b(X)$, we introduce the notion of almost spanning class with respect to a thick subcategory I (Definition 3.9) and we show (Theorem 3.10) that if X is a smooth projective variety of general type then there exists a proper tensor ideal I_{X^*} of $(D^b(X), \otimes^{\mathbb{L}}_X, \mathscr{O}_X)$ such that the set of tensor powers of ω_X forms an almost spanning sequence with respect to this ideal I_{X^*}. This result is meant to highlight the more general behavior of almost spanning classes through the use of Thomason's classification theorem and properties of the Balmer spectrum. We see that this collection of objects can be used to prove the following result:

Lemma 1.3 (Lemma 3.12). *Suppose X is a smooth projective variety of general type. If $\boxtimes$ is a tensor triangulated structure on $D^b(X)$ with unit $\mathscr{O}_X$, and U is a $\boxtimes$-invertible object such that $U \boxtimes I_{X^*} \subseteq I_{X^*}$, then there is a natural equivalence between the functors induced by $U \boxtimes \_$ and $U \otimes^{\mathbb{L}}_X \_$ in $D^b(X)/I_{X^*}$.*

When the $\otimes^{\mathbb{L}}_X$-tensor ideal I_{X^*} is also a $\boxtimes$-tensor ideal for a tensor triangulated category structure as described in the previous lemma, then we obtain that the Picard group of $\boxtimes$-invertible objects is a subgroup of the Picard group of $\otimes^{\mathbb{L}}_X$-invertible objects (Corollary 3.15). This hypothesis holds true in particular when the (anti)canonical bundle of X is ample.

With this observation, our main corollary is the following monoidal version of the Bondal–Orlov reconstruction theorem:

Corollary 1.4 (Corollary 3.18). *Let X be a smooth projective variety with ample (anti)canonical bundle. If $\omega_X[n]$ is an invertible object for a tensor triangulated structure $\boxtimes$ on $D^b(X)$ with unit $\mathscr{O}_X$, then $\boxtimes$ and $\otimes^{\mathbb{L}}_X$ coincide on objects.*

2. Derived categories and the Balmer reconstruction

Throughout the rest of this work we will be working exclusively with smooth projective varieties over $\mathbb{C}$. We will omit the mention of the base field. We recommend [Huybrechts 2006] as a good reference for the material concerning derived categories in this section.

The goal of this section is to introduce the basic results and notions we will be using for our results. Let us start by recalling that if X is a smooth projective variety then there exists an equivalence as triangulated categories between the derived category $\mathrm{Perf}(X)$ of perfect complexes on X and the bounded derived category $D^b(X)$. As a consequence of this, whenever we work with such a variety we will at times make no distinction between these two categories.

One important feature of these categories is the existence of Serre functors.

Definition 2.1. Let $\mathscr{T}$ be a triangulated category. An autoequivalence $S : \mathscr{T} \to \mathscr{T}$ satisfying $\mathrm{Hom}(A, B) \cong \mathrm{Hom}(B, S(A))^*$ for all objects $A, B \in \mathscr{T}$ is called a Serre functor.

Example 2.2. If the triangulated category is a derived category of a smooth projective scheme of dimension n, we have Grothendieck–Verdier duality, which implies that for every pair of objects $M, N \in D^b(X)$, $\mathrm{Hom}(M, N) = \mathrm{Hom}(N, M \otimes \omega_X[n])^*$, where ω_X is the canonical bundle of X.

This notion was first defined by Bondal and Kapranov [1989]. The following two properties of the Serre functor are essential to our work:

Lemma 2.3 [Bondal and Orlov 2001, Proposition 1.3]. *Let $\mathscr{T}$ be a triangulated category with Serre functor S, and let $\psi : \mathscr{T} \to \mathscr{T}$ be any autoequivalence. Then $\psi \circ S \cong S \circ \psi$.*

Proposition 2.4 [Bondal and Kapranov 1989, Proposition 3.4]. *Let $\mathscr{T}$ be a triangulated category and let S be a Serre functor on $\mathscr{T}$. Then S is unique up to graded isomorphism.*

This latter proposition implies that whenever the Serre functor exists, it is intrinsic to the given category. In our case of interest, because one can write this functor using the derived tensor product $\otimes_X^{\mathbb{L}}$, we have now some possible control on the monoidal structure $\otimes_X^{\mathbb{L}}$ directly from the category without knowledge of X.

Another crucial notion we will use is that of spanning classes.

Definition 2.5. A collection of objects $\{X_i\} \subseteq \mathscr{T}$ of a triangulated category is called a spanning class if the following hold:

(1) If $\mathrm{Hom}(X_i, D[j]) = 0$ for all i and j, then $D \simeq 0$.

(2) If $\mathrm{Hom}(D[j], X_i) = 0$ for all i and j, then $D \simeq 0$.

However, whenever the Serre functor exists in the triangulated category we see that only one of the conditions is necessary and the other will be automatically satisfied by use of the Serre functor isomorphism.

A general way to produce spanning classes in derived categories of abelian categories is from ample sequences:

Definition 2.6. Let $\mathcal{A}$ be an abelian category. A collection of objects $\{L_i\} \subset \mathcal{A}$ is called an ample sequence if for $i \ll 0$ and all $A \in \mathcal{A}$,

(1) $\mathrm{Hom}(L_i, A) \otimes_k L_i \to A$ is surjective,

(2) $\mathrm{Hom}(A, L_i) = 0$, and

(3) $\mathrm{Ext}^j(L_i, A) = 0$ for $j \neq 0$.

As the name suggests, an important example of such sequences comes from collections of tensor powers of ample line bundles. The relation between the two notions of spanning class and ample sequence was shown by Orlov [1997]:

Lemma 2.7. *Let $\mathcal{A}$ be an abelian category of finite homological dimension, and let $\{L_i\}$ be an ample sequence. Then the collection $\{L_i\}$, seen as objects of $\mathcal{D}(\mathcal{A})$, form a spanning class.*

The next example shows how we should exploit the existence of ample sequences.

Example 2.8. Let X be a smooth projective variety with ample canonical bundle. Then the set $\{\omega_X^{\otimes x}\}$ forms an ample sequence, so by the previous lemma it forms a spanning class in the derived category $D^b(X)$.

As a consequence, we see that any bounded complex $\mathscr{F}$ of coherent sheaves can be resolved by tensor powers of the canonical bundle ω_X. In other words, there exists an exact sequence

$$0 \to \bigoplus_{j_0}(\omega_X^{\otimes i_0}) \to \cdots \to \bigoplus_{j_k}(\omega_X^{\otimes i_k}) \to \mathscr{F} \to 0.$$

Remark 2.9. In general, for a triangulated category $\mathscr{T}$ with a spanning class $\Omega \subset \mathscr{T}$, if $\phi : \mathscr{T} \to \mathscr{T}$ is an autoequivalence then the set $\phi(\Omega)$ is a spanning class.

This implies that in the example above one can resolve any bounded complex $\mathscr{F}$ by tensor powers of sheaves of the form $\omega_X(i)[j]$ for a fixed $i, j \in \mathbb{Z}$.

Tensor triangulated geometry. When dealing with derived categories of coherent sheaves on a variety one can equip this category with a monoidal structure given by the derived tensor product. One can axiomatize this sort of structure in what is known as a tensor triangulated category.

In this subsection we recall Balmer's spectrum construction, which inputs a tensor triangulated category and outputs a locally ringed space, which as we will see recovers a variety whenever we work with the derived category of perfect complexes on said variety.

Definition 2.10. An essentially small tensor triangulated category (TTC) $\mathscr{T}$ is a triangulated category together with the following data:

(1) A closed symmetric monoidal structure given by a functor $\otimes : \mathscr{T} \times \mathscr{T} \to \mathscr{T}$ that is additive and exact (with respect to the k-linear structure) on both entries.

(2) The internal Hom functor $\underline{\mathrm{hom}} : \mathscr{T} \times \mathscr{T} \to \mathscr{T}$ sending triangles to triangles (up to a sign).

(3) Coherent natural isomorphisms for each n and m in $\mathbb{Z}$, $r : x \otimes (y[n]) \cong (x \otimes y)[n]$ and $l : (x[n]) \otimes y \cong (x \otimes y)[n]$, compatible with the symmetry, associative and unit coherence morphisms from the symmetric monoidal category structure (see for example [Dell'Ambrogio 2016, Section 2.1.1] for the explicit diagrams).

We will refer to a TTC by the triple $(\mathscr{T}, \otimes, \mathbb{1}_{\mathscr{T}})$, where $\otimes$ refers to the monoidal structure and $\mathbb{1}_{\mathscr{T}}$ to the unit object. Often if there is no confusion or the unit plays no role we will omit it and write $(\mathscr{T}, \otimes)$ instead.

At times when we deal with a fixed underlying triangulated category $\mathscr{T}$ we will write $\otimes$ or $(\otimes, \mathbb{1})$ to refer to a tensor triangulated category structure on $\mathscr{T}$.

Let us remark however that the functor $\otimes$ and unit $\mathbb{1}_{\mathscr{T}}$ do not completely determine a tensor triangulated category since the compatibility conditions in the symmetric monoidal category structure can in principle change while maintaining the functor $\otimes$ and unit $\mathbb{1}$. As we will explain in the following, this does not represent a problem for our purposes.

We proceed with a number of definitions.

Definition 2.11. Let $\mathscr{T}$ be a triangulated category, and $\mathscr{I} \subseteq \mathscr{T}$ a full triangulated subcategory. We say that $\mathscr{I}$ is thick if it is closed under direct summands, so that if $A \oplus B \in \mathscr{I}$ then $A, B \in \mathscr{I}$.

Definition 2.12. Let $(\mathscr{T}, \otimes)$ be a TTC. We will say that a thick subcategory $\mathscr{I} \subset \mathscr{T}$ is a $\otimes$-ideal if for every $A \in \mathscr{T}$ we have $A \otimes \mathscr{I} \subset \mathscr{I}$

Definition 2.13. Let $(\mathscr{T}, \otimes)$ be a tensor triangulated category. Let $\mathscr{I}$ be a $\otimes$-ideal. We will say that $\mathscr{I}$ is prime if when $A, B \in \mathscr{T}$ with $A \otimes B \in \mathscr{I}$, it follows that $A \in \mathscr{I}$ or $B \in \mathscr{I}$.

As in affine algebraic geometry we can define the spectrum of a tensor triangulated category.

Definition 2.14. Let $(\mathscr{T}, \otimes, \mathbb{1})$ be an essentially small tensor triangulated category. The set of all prime $\otimes$-ideals will be denoted by $\mathrm{Spc}(\mathscr{T}, \otimes, \mathbb{1})$ (alternatively $\mathrm{Spc}(\mathscr{T})$, $\mathrm{Spc}(\otimes, \mathbb{1})$ or $\mathrm{Spc}(\otimes)$, depending on which information is clear from context).

Importantly, whenever the triangulated category $\mathscr{T}$ is nonzero we have that $\mathrm{Spec}(\mathscr{T}, \otimes) \neq \varnothing$ for any tensor triangulated category structure $\otimes$ we can put on $\mathscr{T}$ (see [Balmer 2005, Proposition 2.3]).

On this set we will put a topological structure.

Definition 2.15. Let $(\mathscr{T}, \otimes, \mathbb{1})$ be a TTC. The support of an object $A \in \mathscr{T}$, denoted by $\mathrm{supp}(A)$, is the set $\{\mathfrak{p} \in \mathrm{Spc}(\mathscr{T}) \mid A \notin \mathfrak{p}\}$.

Lemma 2.16 [Balmer 2005, Lemma 2.6]. *Let $S \subset \mathscr{T}$ be a family of objects. The sets of the form $\mathcal{Z}(S) := \bigcap_{A \in S} \mathrm{supp}(A)$ form a basis for a topology on $\mathrm{Spc}(\mathscr{T})$.*

An important result regarding this topology is the following, which restricts the kind of spaces we should be expecting from the construction.

Theorem 2.17 [Balmer 2005, Propositions 2.15, 2.18]. *For any TTC $(\mathscr{T}, \otimes, \mathbb{1})$, the space $\mathrm{Spc}(\mathscr{T})$ is a spectral space in the sense of Hochster, meaning it is sober and has a basis of quasicompact open subsets.*

Now that the topology on $\mathrm{Spc}(\mathscr{T})$ has been chosen, the next step is to equip this space with sheaf of rings which will act as the structure sheaf.

To a subset $Y \subset \mathrm{Spc}(\mathscr{T})$ we can assign a thick $\otimes$-ideal denoted by $\mathscr{I}_Y$ and defined as the subcategory supported on Y, meaning $\mathscr{I}_Y := \{A \in \mathscr{T} \mid \mathrm{supp}(A) \subset Y\}$.

Finally, with Y as above, we denote by $\mathbb{1}_{T_Y}$ the image of the unit $\mathbb{1}$ of $\mathscr{T}$ under the localization functor $\pi : \mathscr{T} \to \mathscr{T}/\mathscr{I}_Y$.

Definition 2.18. Let $\mathscr{T}$ be a nonzero TTC. We define a structure sheaf $\mathcal{O}_{\mathrm{Spc}(\mathscr{T})}$ over $\mathrm{Spc}(\mathscr{T})$ as the sheaffification of the assignment $U \mapsto \mathrm{End}(\mathbb{1}_{T_Z})$, where $Z := \mathrm{Spc}(\mathscr{T}) \setminus U$, for an open subset $U \subset \mathrm{Spc}(\mathscr{T})$.

It is not hard to see the assignment $\mathrm{Spc}(F)$ respects composition of exact monoidal functors, so if $F : \mathscr{T} \to \mathscr{T}'$ is such a functor, we get a morphism of ringed spaces, since for a closed $Z = \mathrm{Spc}(\mathscr{T}) \setminus U$ we have $F(\mathscr{I}_Z) \subset \mathscr{I}_{Z'}$, where $Z' = \mathrm{Spc}(\mathscr{T}') \setminus \mathrm{Spc}(F)^{-1}(U)$, which implies there is a morphism $\mathcal{O}_{\mathscr{T}} \to \mathrm{Spc}(F)_*\mathcal{O}_{\mathcal{L}}$, so $\mathrm{Spc} : \mathbb{TTC} \to \mathrm{RS}$ is a functor, where $\mathbb{TTC}$ and RS denote the categories of essentially small tensor triangulated categories and of ringed spaces, respectively.

Under nice conditions (for example, $\mathscr{T}$ rigid), this can be shown to be a functor $\mathrm{Spc} : \mathbb{TTC} \to \mathrm{LRS}$, where LRS denotes the category of locally ringed spaces.

With this construction in mind we can now describe the anticipated reconstruction theorem as described by Balmer.

Theorem 2.19 [Balmer 2005, Corollary 5.6]. *Let X be a quasicompact and quasiseparated scheme. There is a homeomorphism*

$$f : X \xrightarrow{\cong} \mathrm{Spc}(\mathrm{Perf}(X), \otimes_X^{\mathbb{L}}).$$

This homeomorphism follows from Thomason's classification theorem [1997, Theorem 3.15], which establishes a correspondence between certain subsets of a quasicompact and quasiseparated scheme X and $\otimes_X^{\mathbb{L}}$-ideals of $\mathrm{Perf}(X)$. The following is a general version of this classification for tensor triangulated categories as presented by Balmer [2005, Theorem 4.10]

Theorem 2.20. *Let $(\mathscr{T}, \otimes, \mathbb{1})$ be a TTC. Let $\mathscr{S}$ be those subsets $Y \subset \mathrm{Spc}(\mathscr{T})$ which are unions $Y = \bigcup_{i \in I} Y_i$, where Y_i is a closed subset with quasicompact complement for all $i \in I$. Let $\mathscr{R}$ be the set of radical $\otimes$-ideals of $\mathscr{T}$. Then there is an order-preserving bijection $\mathscr{S} \to \mathscr{R}$ given by the assignment which sends Y to the subcategory $\mathscr{T}_Y := \{A \in \mathscr{T} \mid \mathrm{supp}(A) \subset Y\}$ and with inverse sending a radical $\otimes$-ideal $\mathscr{I}$ to the subset $S_{\mathscr{I}} := \bigcup_{A \in \mathscr{I}} \mathrm{supp}(A)$.*

Here by radical $\otimes$-ideal we mean a $\otimes$-ideal $\mathscr{I}$ such that whenever $A^{\otimes n}$ is in $\mathscr{I}$ then A is in $\mathscr{I}$.

In practice every $\otimes$-ideal is automatically a radical $\otimes$-ideal and it certainly depends on the monoidal structure one can put on the triangulated category $\mathscr{T}$. As pointed out by Balmer [2005] this condition is satisfied as soon as the tensor triangulated category is rigid, meaning that every object is dualizable.

When X is a variety the classification theorem can be specialized to a very simple form, as pointed out by Rouquier [2003].

Theorem 2.21. *Let X be a variety. There is a correspondence between the set of closed subsets of X and $\otimes_X^{\mathbb{L}}$-ideals of finite type, that is, those ideals generated by a single object.*

Using the homeomorphism from Theorem 2.19 and the construction of the structure sheaf on $\mathrm{Spc}(\mathrm{Perf}(X), \otimes_X^{\mathbb{L}})$ from Definition 2.18 we only need the following theorem to complete the reconstruction theorem of Balmer.

Theorem 2.22 [Balmer 2002]. *Let X be scheme with underlying noetherian topological space. There is an isomorphism $\mathscr{O}_X \cong \mathscr{O}_{\mathrm{Spc}(\mathrm{Perf}(X), \mathscr{O}_X)}$.*

The next proposition should inform us how localizations behave under taking Spc.

Proposition 2.23 [Balmer 2005, Proposition 3.11]. *Let $\mathscr{I} \subset \mathscr{T}$ be a thick $\otimes$-ideal. Then the localization functor $\pi : \mathscr{T} \to \mathscr{T}/\mathscr{I}$ is an exact monoidal functor and induces an homeomorphism* $\mathrm{Spc}(\mathscr{T}/\mathscr{I}) \cong \{\mathfrak{p} \in \mathrm{Spc}(\mathscr{T}) \mid \mathscr{I} \subset \mathfrak{p}\}$.

In particular, when this is combined with the classification theorem in the form of Theorem 2.21, we see that open subvarieties U of a variety X are isomorphic to $\mathrm{Spc}(\mathrm{Perf}(X)/\mathscr{I}_Z)$, where Z is the complement of U in X.

Remark 2.24. So far we have been dealing with tensor triangulated categories as described in Definition 2.10, meaning we require there to be a closed symmetric monoidal category structure on $\mathscr{T}$. However, under closer inspection, one sees that nowhere in the classification theorem nor in Balmer's construction does one need the full monoidal structure.

In fact so far we really only need the data of a functor $\otimes : \mathscr{T} \times \mathscr{T} \to \mathscr{T}$, covariant and exact in each variable, together with a unit object $\mathbb{1}$ and isomorphisms corresponding to the symmetric, associative and unit conditions. In other words, if $(\mathscr{T}, \otimes, \mathbb{1})$ and $(\mathscr{T}, \boxtimes, \mathbb{1}')$ are two tensor triangulated categories with underlying triangulated category $\mathscr{T}$ such that $\otimes \simeq \boxtimes$ for every pair of objects in $\mathscr{T}$, and $\mathbb{1} \simeq \mathbb{1}'$, then the Balmer spectra satisfy $\mathrm{Spc}(\mathscr{T}, \otimes, \mathbb{1}) \cong \mathrm{Spc}(\mathscr{T}, \boxtimes, \mathbb{1}')$ as locally ringed spaces. The associators, unitors and braidings of the monoidal categories have no influence in the resulting space.

It is this that justifies our notation $(\mathscr{T}, \otimes, \mathbb{1})$, as we have mentioned before. In the following we shall keep referring to tensor triangulated categories although our results apply for slightly more general but more awkward structures.

3. Tensor triangulated categories and Picard groups

While the Bondal–Orlov reconstruction (Theorem 1.1) tells us that one can directly recover a smooth projective variety X with ample (anti)canonical bundle from the derived category $D^b(X) \simeq \mathrm{Perf}(X)$, there are plenty of smooth projective varieties which have nonisomorphic Fourier–Mukai partners, varieties Y such that $D^b(X) \simeq D^b(Y)$, which implies that on a given derived category $D^b(X)$ there might be many nonequivalent tensor triangulated category structures.

However, even in the case where our variety X has ample (anti)canonical bundle, as in the hypothesis of the Bondal–Orlov reconstruction theorem, it is not immediate that there is only one possible tensor triangulated category structure. It is, in principle, possible that there might be one such structure $(D^b(X), \boxtimes, \mathbb{1})$ such that $D^b(\mathrm{Spc}(\boxtimes, \mathbb{1})) \not\simeq D^b(X)$ and so Bondal–Orlov does not apply.

In some sense our motivating question is whether Balmer's reconstruction implies Bondal–Orlov. In this section we will be looking into this and related ideas by exploring the possible tensor triangulated categories one can equip on $D^b(X)$ under the slightly more general hypothesis of X having a big (anti)canonical bundle.

We start by mentioning a result by Liu and Sierra [2013] that shows, in particular, that there are smooth projective varieties X with ample anticanonical bundle — thus under the hypothesis of Bondal–Orlov — for which the derived category $D^b(X)$ admits a tensor triangulated category structure $(\boxtimes, \mathbb{1})$ such that $\mathrm{Spc}(\boxtimes, \mathbb{1}) \not\cong X$.

Recall that there are varieties X that are known to have derived categories equivalent to the derived category of representations on a quiver (possibly with relations). For example, in the presence of a full strong exceptional collection $\{E_i\}$, we have that $D^b(X)$ is equivalent to $D^b\big(\text{mod-End}\big(\bigoplus E_i\big)\big)$, the derived category of finitely generated modules over the algebra $\mathrm{End}\big(\bigoplus E_i\big)$. This latter algebra on the other hand is equivalent to the path algebra of a quiver, and so we obtain an equivalence between the derived category of X and the derived category of finite-dimensional representations of a quiver $Q = (P_n, E_{ij})$.

The important point here is that this derived category of representations of a quiver comes with a tensor triangulated category structure induced by the tensor product of representations. Recall that for two such representations (V_i, p_{ik}) and (W_j, q_{js}) the tensor product is given entrywise: $(V_i, p_{ik}) \otimes_{\mathrm{rep}} (W_j, q_{js}) := (V_i \otimes W_j, p_{ik} \otimes q_{js})$.

We write rep Q for the abelian k-linear category of finite-dimensional quiver representations on a quiver Q. Let us denote by $(D^b(\mathrm{rep}\, Q), \otimes^{\mathbb{L}}_{\mathrm{rep}}, \mathbb{1}_{\mathrm{rep}})$ the resulting tensor triangulated category structure on $D^b(\mathrm{rep}\, Q)$, where $\otimes^{\mathbb{L}}_{\mathrm{rep}}$ is the derived tensor product coming from $\otimes_{\mathrm{rep}}$, and $\mathbb{1}_{\mathrm{rep}} := (k_i, Id_{ij})$ is the representation given by putting k on every vertex and the identity morphism in each edge of the quiver.

Liu and Sierra [2013, Definition 1.2.5] consider quivers with relations satisfying a compatibility condition with the tensor product and say that in this case the quiver has tensor relations.

Theorem 3.1 [Liu and Sierra 2013, Theorem 2.1.5.1]. *Let Q be a finite ordered quiver with tensor relations. Then* $\mathrm{Spc}((D^b(\mathrm{rep}\, Q), \otimes^{\mathbb{L}}_{\mathrm{rep}}))$ *is the discrete space* $\{P_n\}$.

They also describe completely the structure sheaf in this case.

Theorem 3.2 [Liu and Sierra 2013, Theorem 2.2.4.1]. *Let Q be a finite ordered quiver with tensor relations. Then $\mathscr{O}_Q := \mathscr{O}_{\mathrm{Spc}(\otimes^{\mathbb{L}}_{\mathrm{rep}})}$ is the constant sheaf of algebras k, so that for any open $W \subset \mathrm{Spc}(\otimes^{\mathbb{L}}_{\mathrm{rep}})$ we have $\mathscr{O}_Q(W) = k^{\oplus W}$.*

In particular, for $X = \mathbb{P}^n$, we have by a well-known result of Beilinson [1978] that $D^b(X)$ is equivalent to the category of representations of a quiver with $n+1$ vertices. Thus the derived category $D^b(X)$ has a tensor triangulated category structure $(\otimes^{\mathbb{L}}_{\mathrm{rep}}, \mathbb{1}_{\mathrm{rep}})$ such that $\mathrm{Spc}((\otimes^{\mathbb{L}}_{\mathrm{rep}}, \mathbb{1}_{\mathrm{rep}})) \ncong X = \mathbb{P}^n$. As $\mathbb{P}^n$ is a smooth projective variety with ample anticanonical bundle, this previous result implies that the study of tensor triangulated category structures on $D^b(X)$ is not trivial even in the cases falling under the hypothesis of the Bondal–Orlov reconstruction theorem and might shed some light on the internal structure of the derived category in itself.

In general the behavior of the dynamics of the Balmer spectrum and taking derived categories can be complex. Since the Balmer spectrum is a locally ringed space, it has an abelian category of sheaves of modules which admits a tensor product, and we can derive this category as usual, but the category of sheaves of modules is in general much more complicated than a category of coherent or even quasicoherent sheaves.

Having said that, let us put ourselves in the slightly more general situation of derived categories of varieties of general type. Recall a variety is of general type if its canonical bundle is big. In particular, varieties with ample canonical bundle are of general type.

One alternative characterization of bigness for a variety is the following:

Theorem 3.3 [Lazarsfeld 2004, Example 2.2.9]. *A smooth projective variety is of general type if and only if, for any sheaf $\mathscr{F} \in \mathrm{Coh}(X)$, there exists an integer i_0 depending on $\mathscr{F}$ such that the sheaf $\mathscr{F} \otimes_X \omega_X^i$ is generically globally generated for $i \gg i_0$.*

As a consequence of the Kodaira lemma (see [Lazarsfeld 2004, Proposition 2.2.6]), we have the corollary:

Corollary 3.4. *Let X be a smooth projective variety of general type. Then there exists an open subvariety X^* such that for any $\mathscr{F} \in \mathrm{Coh}(X)$, there exists a positive integer i_0 for which, for any $i \gg i_0$, the sheaf $\mathscr{F}|_{X^*} \otimes_X \omega^i_{X^*}$ on X^* is globally generated.*

Let us explain the previous corollary and the nature of the open subvariety X^*. We recall some basic definitions.

Definition 3.5. Let X be a projective variety, and let $\mathscr{L}$ be a line bundle on X. The augmented base locus is the Zariski closed set

$$B_+(\mathscr{L}) := \bigcap_{m \in \mathbb{N}} B(m\mathscr{L} - A),$$

where A is any ample line bundle, and for any line bundle $\mathscr{L}'$ the set $B(\mathscr{L}')$ is defined as the intersection of the base loci of multiples of the line bundle, that is,

$$B(\mathscr{L}') := \bigcap_{m \in \mathbb{N}} Bs(m\mathscr{L}').$$

In [Boucksom et al. 2014] the following theorem characterizing the complement of the augmented base locus is proven:

Theorem 3.6. *Let $\mathscr{L}$ be a big line bundle on a normal projective variety X over an algebraically closed field. Then the complement $X \setminus B_+(\mathscr{L})$ of the augmented base locus is the largest Zariski open subset $U \subseteq X \setminus B(\mathscr{L})$ such that for all large and divisible $m(\mathscr{L}) \in \mathbb{Z}$ the restriction of the morphism*

$$\phi_m : X \setminus B(\mathscr{L}) \dashrightarrow \mathbb{P}H^0(X, m\mathscr{L})$$

to U is an isomorphism onto its image.

Two important observations follow immediately from the definition, the fact that the augmented base locus is independent of the choice of ample line bundle, and Kodaira's decomposition of big line bundles.

Remark 3.7. (1) $B_+(\mathscr{L}) = \varnothing$ if and only if $\mathscr{L}$ is ample.

(2) $B_+(\mathscr{L}) \neq X$ if and only if $\mathscr{L}$ is big.

From the remarks above and Thomason's classification theorem, Theorem 2.20, we know that since there exists a correspondence between closed subsets of the Balmer spectrum and radical tensor ideals in the tensor triangulated category, there exists a radical tensor ideal corresponding to the augmented base locus $B_+(\mathscr{L})$ for any given line bundle $\mathscr{L}$. In particular, the open subvariety X^* from Corollary 3.4 is the complement of the augmented base locus, $X \setminus B_+(\omega_X)$, and corresponds to a $\otimes^{\mathbb{L}}_X$-ideal generated by a single object (using Theorem 2.21) whose homological support gives back the closed subset $B_+(\omega_X)$.

Remark 3.8. Denote by I_{X^*} the $\otimes^{\mathbb{L}}_X$-ideal corresponding to the open subvariety X^*. By Remark 3.7, this ideal must be a proper $\otimes^{\mathbb{L}}_X$-ideal of $D^b(X)$ and is the ideal 0 precisely when the (anti)canonical bundle is ample.

We would like to understand the effect of the positivity of the canonical bundle (in this case the fact that the variety is of general type) on the tensor triangulated structure of the category. We know from Proposition 2.4 that the Serre functor in a triangulated category is unique up to degree whenever it exists and so it is intrinsic to the category. In our concrete case we know furthermore that the Serre functor is isomorphic to $\_ \otimes^{\mathbb{L}}_X \omega_X[n]$, where $n \in \mathbb{N}$ is the dimension of the variety and ω_X is the dualizing sheaf of X.

Let us start with a definition mimicking that of spanning class:

Definition 3.9. Let $(\mathscr{T}, \otimes)$ be a tensor triangulated category, let $\mathscr{I} \subseteq \mathscr{T}$ be a thick subcategory, and let us denote by $\pi : \mathscr{T} \to \mathscr{T}/\mathscr{I}$ the localization functor. We say that a collection of objects $\Omega \subset \mathscr{T}$ is an almost spanning class with respect to $\mathscr{I}$ if the following two conditions hold:

(1) If $X \in \mathscr{T}/\mathscr{I}$ is such that $\mathrm{Hom}_{\mathscr{T}/\mathscr{I}}(\pi(B), X[j]) = 0$ for all $B \in \Omega$ and $j \in \mathbb{Z}$, then $X \cong 0$.

(2) If $X \in \mathscr{T}/\mathscr{I}$ is such that $\mathrm{Hom}_{\mathscr{T}/I}(X[j], \pi(B)) = 0$ for all $B \in \Omega$ and $j \in \mathbb{Z}$, then $X \cong 0$.

It is immediate to see that the previous definition is equivalent to asking that the collection Ω maps through π to a spanning class on the quotient $\mathscr{T}/\mathscr{I}$. When the thick subcategory in question is the 0 subcategory then the definition reduces to that of a spanning class as in Definition 2.5.

Additionally when the triangulated category $\mathscr{T}/\mathscr{I}$ has a Serre functor, only one of the conditions in the definition is necessary as the Serre duality implies the other automatically.

We would like to generalize Lemma 2.7 but for a big canonical bundle instead of an ample one and see that a big bundle induces an almost spanning class in the derived category with respect to a $\otimes_X^{\mathbb{L}}$-ideal $\mathscr{I}$.

Theorem 3.10. *Let X be a smooth projective variety of general type. Then the collection of tensor powers $(\omega_X^{\otimes i})_{i \in \mathbb{Z}}$ forms an almost spanning class with respect to the tensor ideal I_{X^*} in the tensor triangulated category $(D^b(X), \otimes_X^{\mathbb{L}})$.*

Proof. We need to show that $\pi(\{\omega_X^{\otimes i}\})$ forms a spanning class in the quotient $D^b(X)/I_{X^*}$. As I_{X^*} is the ideal corresponding to the open smooth subvariety X^*, from Corollary 3.4 we know that there is an isomorphism $\mathrm{Spc}(D^b(X)/I_{X^*}) \cong X^*$. Since ω_X restricted to X^* is ample by the characterization of Theorem 3.6, we get that $\{\omega_X^{\otimes^{\mathbb{L}} i}|_{X^*}\}$ forms a spanning class, by Lemma 2.7, of the derived category of X^* which coincides with the quotient category $D^b(X)/I_{X^*}$ by Proposition 2.23. □

The main key in our arguments is the fact that one can construct, as in the ample case, a resolution for any bounded complex of coherent sheaves on X^* in terms of tensor powers of the canonical bundle ω_{X^*} of X^*, with the advantage that one is able to have a concrete description of the derived category of this space in terms of a quotient of the derived category of the larger variety X.

Explicitly, for any bounded complex A of coherent sheaves over X^*, there is a resolution

$$\cdots \to 0 \to \bigoplus_{j_0} (\omega_{X^*}^{\otimes i_0}) \to \cdots \to \bigoplus_{j_k} (\omega_{X^*}^{\otimes i_k}) \to A \to 0.$$

Another thing to notice is that in the example given above for the nonequivalent tensor triangulated category structures on $D^b(\mathbb{P}^n)$, one immediate issue with the two given such structures was that the units were nonisomorphic. So we should proceed to work with tensor triangulated categories with a fixed unit isomorphic to $\mathscr{O}_X$.

Definition 3.11. Let $(\mathscr{T}, \otimes, \mathbb{1})$ be a TTC. An object $X \in \mathscr{T}$ is $\otimes$-invertible if there exists $X^{-1} \in \mathscr{T}$ such that $X \otimes X^{-1} \cong \mathbb{1}$. We will denote by $\mathrm{Pic}(D^b(X), \otimes)$ the group of isomorphism classes of $\otimes$-invertible objects.

We will make use of the following lemma:

Lemma 3.12. *Suppose X is a smooth projective variety of general type of dimension n. If $\boxtimes$ is a tensor triangulated structure on $D^b(X)$ with unit $\mathscr{O}_X$, and U is a $\boxtimes$-invertible object such that $U \boxtimes I_{X^*} \subseteq I_{X^*}$, then there is a natural equivalence between the functors induced by $U \boxtimes \_$ and $U \otimes^{\mathbb{L}}_X \_$ in $D^b(X)/I_{X^*}$.*

Proof. By our previous discussion we know that any bounded complex can be resolved in $D^b(X)/I_{X^*}$ by a resolution

$$\cdots \to 0 \to \bigoplus_{j_0} (\omega_{X^*}^{\otimes i_0}) \to \cdots \to \bigoplus_{j_k} (\omega_{X^*}^{\otimes i_k}) \to A \to 0.$$

Because the Serre functor in $D^b(X^*)$ is given by $\_ \otimes^{\mathbb{L}}_X \omega_{X^*}[n']$, where n' is the dimension of X^*, and we know any exact equivalence must commute with it, if we let $U\widehat{\boxtimes}$ and $U\widehat{\otimes^{\mathbb{L}}_X}$ denote the endofunctors of $D^b(X)/I_{X^*}$ induced respectively by $U\boxtimes$ and $U\otimes^{\mathbb{L}}$, then, since $U\widehat{\boxtimes}$ is an autoequivalence,

$$(U \mathbin{\widehat{\boxtimes}} \hat{A}) \mathbin{\widehat{\otimes^{\mathbb{L}}_X}} \omega_{X^*}[n'] \cong U \mathbin{\widehat{\boxtimes}} (\hat{A} \mathbin{\widehat{\otimes^{\mathbb{L}}_X}} \omega_{X^*}[n']).$$

As $\mathscr{O}_X$ is a unit for both $\otimes_X$ and $\boxtimes$, and after shifting by $[-n']$, we deduce

$$U \mathbin{\widehat{\otimes^{\mathbb{L}}_X}} \omega_{X^*} \cong U \mathbin{\widehat{\boxtimes}} \omega_{X^*}.$$

From this, the exactness of $\otimes^{\mathbb{L}}$ and $\boxtimes$, and the resolutions in terms of $\omega^i_{X^*}$, we obtain the isomorphisms

$$U \mathbin{\widehat{\otimes^{\mathbb{L}}}} A \cong U \mathbin{\widehat{\boxtimes}} A. \qquad \square$$

Remark 3.13. Let us point out the slight abuse of notation of the functor $U\widehat{\otimes^{\mathbb{L}}}$. This functor would formally be denoted by $\widehat{U}\otimes^{\mathbb{L}}_{D^b(X)/I_{X^*}}$, as it is induced by the object $\widehat{U}$ in the tensor triangulated category $(D^b(X)/I_{X^*}, \otimes^{\mathbb{L}}_{D^b(X)/I_{X^*}})$, but as the only tensor ideal we are taking a quotient by in this section is I_{X^*}, we believe our notation is lighter without losing sight of which functors they represent.

We have the following corollary:

Corollary 3.14. *Let X be a variety of general type, and let $\boxtimes$ be a tensor triangulated category structure on $D^b(X)$ with unit $\mathscr{O}_X$. Then, for any $\boxtimes$-invertible object U such that $U \boxtimes I_{X^*} \subseteq I_{X^*}$, the equivalence $U\widehat{\boxtimes} : D^b(X)/I_{X^*} \to D^b(X)/I_{X^*}$ induced by $U \boxtimes \_$ is equivalent to an equivalence given by objects in the group* $\mathrm{Pic}(D^b(X)/I_{X^*}, \widehat{\otimes^{\mathbb{L}}})$ *of invertible $\widehat{\otimes^{\mathbb{L}}}$-objects.*

Proof. From Lemma 3.12 we have that if U^{-1} is such that $U \boxtimes U^{-1} \cong \mathscr{O}_X$ then in the quotient $D^b(X)/I_{X^*}$,

$$U \mathbin{\widehat{\otimes^{\mathbb{L}}}} \widehat{U^{-1}} \cong U \mathbin{\widehat{\boxtimes}} \widehat{U^{-1}} \cong \mathscr{O}_{X^*}.$$

As $(D^b(X)/I_{X^*}, \widehat{\otimes^{\mathbb{L}}})$ is a tensor triangulated category, we have that $\widehat{U} \in D^b(X)/I_{X^*}$ is a $\widehat{\otimes^{\mathbb{L}}}$-invertible object. □

In Lemma 3.12 and Corollary 3.14, the ideal I_{X^*} might not be a $\boxtimes$-tensor ideal and thus the quotient $D^b(X)/I_{X^*}$ does not necessarily carry a tensor triangulated category structure induced by $\boxtimes$. However, our result guarantees that after passing to the quotient, the equivalences induced by the functors $U \boxtimes \_$ are equivalent to equivalences given by invertible objects in $(D^b(X)/I_{X^*}, \otimes^{\mathbb{L}}_X)$ induced by the same object, under the condition that I_{X^*} is stable by $U\boxtimes$.

In particular, we have:

Corollary 3.15. *Let X be a variety of general type and $\boxtimes$ a tensor triangulated structure on $D^b(X)$ with unit $\mathscr{O}_X$. If I_{X^*} is a $\boxtimes$-ideal then the Picard group* $\operatorname{Pic}(D^b(X)/I_{X^*}, \widehat{\boxtimes})$ *is a subgroup of the Picard group* $\operatorname{Pic}(D^b(X)/I_{X^*}, \widehat{\otimes^{\mathbb{L}}_X})$.

Proof. The proof is as in the previous two. If U is in $\operatorname{Pic}(D^b(X)/I_{X^*}, \widehat{\boxtimes})$ then it induces an autoequivalence of $D^b(X)/I_{X^*}$ and so it commutes with the Serre functor on $D^b(X^*) \simeq D^b(X)/I_{X^*}$. By writing a resolution for any bounded complex A in terms of direct sums of derived tensor powers of ω_{X^*} we can use the same argument as in the proof of Lemma 3.12 and we arrive at the isomorphisms

$$U \mathbin{\widehat{\otimes^{\mathbb{L}}}} A \cong U \mathbin{\widehat{\boxtimes}} A.$$ □

Remark 3.16. In the results above we have chosen to work with varieties of general type, but the same argument applies to varieties with big anticanonical bundle.

The case when our variety has an ample (anti)canonical bundle allows us to relate the Picard group of the full derived category to that of any other tensor triangulated category structure on it.

The following result follows from the previous argument.

Corollary 3.17. *Let X be a variety with ample* (*anti*)*canonical bundle. Then if $\boxtimes$ is a tensor triangulated category structure on $D^b(X)$ with unit $\mathscr{O}_X$, the Picard group* $\operatorname{Pic}(D^b(X), \boxtimes)$ *is isomorphic to a subgroup of* $\operatorname{Pic}(D^b(X), \otimes_X)$.

Proof. We just need to notice that in this case the $\otimes_X$-ideal from Corollary 3.4 is the 0 ideal and thus we can resolve any object $A \in D^b(X)$ by a sequence of powers of the Serre functor. By the same reasoning as above we see that

$$U \otimes^{\mathbb{L}} A \cong U \boxtimes A.$$ □

Although Bondal and Orlov had already classified the group of autoequivalences of a derived category of a variety with ample (anti)canonical bundle, we are working without the condition of an equivalence between the derived category of the Balmer spectrum of $\boxtimes$ and the derived category $D^b(X)$, and as such it is not immediate from their result that the Picard group of $\boxtimes$ must involve invertible sheaves over X.

In other words, since $\mathrm{Spc}(\boxtimes)$ is not necessarily isomorphic to X, understanding the autoequivalences of $D^b(X)$ alone does not give us an immediate relationship to the Picard group of $\boxtimes$.

We should think of the following corollary as a monoidal version of the Bondal–Orlov reconstruction theorem:

Corollary 3.18. *Let X be as above. If $\omega_X[n]$ is an invertible object for a tensor triangulated structure $\boxtimes$ on $D^b(X)$ with unit $\mathcal{O}_X$ then $\boxtimes$ and $\otimes_X^{\mathbb{L}}$ coincide on objects.*

Proof. As ω_X is $\boxtimes$-invertible, Corollary 3.17 tells us that for any $A \in D^b(X)$,

$$\omega_X \otimes_X^{\mathbb{L}} A \cong \omega_X \boxtimes A.$$

But we can resolve any other bounded complex B in terms of derived powers of the canonical sheaf, so by the exactness of $\boxtimes$, we have

$$B \otimes_X^{\mathbb{L}} A \cong B \boxtimes A. \qquad \square$$

The nature of this result comes precisely from the fact that the tensor triangulated category structure $(\boxtimes, \mathbb{1})$ does not necessarily come from a derived equivalence $D^b(X) \simeq D^b(Y)$, and although the extra assumption on the unit is required, the result is in this direction slightly more general than the original theorem.

Corollary 3.19. *Let X be a variety with ample (anti)canonical bundle. If $(\mathcal{O}_X, \boxtimes)$ is a tensor triangulated structure on $D^b(X)$ such that $\mathrm{Pic}(\boxtimes) \cong \mathrm{Pic}(D^b(X))$ via the assignment $U \mapsto U$ then $\boxtimes$ coincides with $\otimes_X^{\mathbb{L}}$ on objects.*

Proof. In this case if this morphism is an isomorphism, then ω_X is $\boxtimes$-invertible and the result follows from the previous corollary. $\square$

In fact if we are under the same hypothesis for X then as soon as we are able to show that the generators of $\mathrm{Pic}(D^b(X), \otimes^{\mathbb{L}})$ are $\boxtimes$-invertible then by the previous corollary there must be an equivalence between $\boxtimes$ and $\otimes_X^{\mathbb{L}}$.

Example 3.20. Let $X = \mathbb{P}^n$ be the projective space. In this case we know that $\mathrm{Pic}(D^b(X)) = \mathbb{Z} \oplus \mathbb{Z}$, corresponding to the line bundles plus their shifts. The result above then says that whenever there is a tensor triangulated structure $\boxtimes$ on $D^b(X)$ with unit $\mathcal{O}_X$ then the Picard group of this tensor structure must necessarily be a subgroup of $\mathbb{Z} \oplus \mathbb{Z}$.

If $\omega_X = \mathcal{O}_X(-n-1)$ is $\boxtimes$-invertible then we get that $\boxtimes$ coincides with $\otimes_X^{\mathbb{L}}$. Similarly if $\mathcal{O}_X(-1)$ is $\boxtimes$-invertible.

One natural question to ask when working with Picard groups of tensor triangulated category structures is what the relationship with line bundles on the associated space is. Balmer and Favi [2007, Proposition 4.4] proved the following result:

Proposition 3.21. *Let X be a scheme and consider* $\mathrm{Perf}(X)$, *its derived category of perfect complexes. Then there is a split short exact sequence of abelian groups*

$$0 \to \mathrm{Pic}(X) \to \mathrm{Pic}(\mathrm{Perf}(X), \otimes^{\mathbb{L}}_X) \to C(X;\mathbb{Z}) \to 0,$$

where $C(X;\mathbb{Z})$ stands for the group of locally constant functions from X to $\mathbb{Z}$.

Again, under the hypothesis of X having an ample (anti)canonical bundle, by using Proposition 3.21 we see that for a TTC $(\boxtimes, \mathscr{O}_X)$, if $\mathrm{Spc}(\boxtimes)$ is a scheme then the Picard group of $\mathrm{Spc}(\boxtimes)$ must be a subgroup of the Picard group of $\mathrm{Pic}(D^b(X), \boxtimes) \leq \mathrm{Pic}(D^b(X), \otimes^{\mathbb{L}}_X)$. So a line bundle in $\mathrm{Spc}(\boxtimes)$ has to be $\otimes^{\mathbb{L}}_X$-invertible.

Remark 3.22. From the original proof of the Bondal–Orlov reconstruction theorem, we know that it is actually possible to fully characterize line bundles up to a shift from purely categorical properties. Given the importance of the Picard group of the variety, we can ask whether it is possible to reconstruct the derived tensor product in $D^b(X)$ without having to pass through a reconstruction theorem.

Antieau [2013] sketched a construction in which by considering invertible objects (in the sense of Bondal and Orlov) one can define a collection of tensor products $\otimes^{\mathbb{L}}_U$ by exploiting the resolution by derived tensor powers of ω_X.

The idea is to pick an invertible object U, shown in [Bondal and Orlov 2001] to be isomorphic to a shift of a line bundle in X, and then by use of the resolution we only need to define the products $\omega_X^{\otimes^{\mathbb{L}}_X i}[ni] \otimes^{\mathbb{L}}_U A^\bullet$ for any object $A^\bullet$. As the Serre functor $S \simeq \_\otimes^{\mathbb{L}}_X \omega_X[n]$ comes from the categorical structure alone then we can set these products to be simply $S^i(A^\bullet)$.

These tensor products $\otimes^{\mathbb{L}}_U$ have U as unit, and they all have X as Balmer spectrum.

In general for a triangulated category $\mathscr{T}$ we have an action by $\mathrm{Aut}(\mathscr{T})$ on the collection $\mathrm{TTS}(\mathscr{T})$.

If $(\otimes, \mathbb{1}) \in \mathrm{TTS}(\mathscr{T})$ and $\phi \in \mathrm{Aut}(\mathscr{T})$ we have a tensor structure defined by

$$X \otimes_\phi Y := \phi^{-1}\big(\phi(X) \otimes \phi(Y)\big),$$

with unit given by $\phi^{-1}(\mathbb{1})$.

We have now justified enough the following definition:

Definition 3.23. Let $\mathscr{T}$ be a triangulated category. Denote by $\mathrm{TTS}(\mathscr{T})$ the collection of equivalence classes of tensor triangulated category structures on $\mathscr{T}$, where we consider two tensor triangulated category structures to be equivalent if there is a monoidal equivalence between the two of them.

To keep some control and avoid counting structures coming from autoequivalences as we saw, we should at least fix the unit object.

Definition 3.24. Let $\mathscr{T}$ be a triangulated category and $U \in \mathscr{T}$ an object. Then the set $\mathrm{TTS}_U(\mathscr{T})$ is the set of equivalence classes of tensor triangulated structures on $\mathscr{T}$, where U is the unit.

This is the set we are mainly interested in classifying.

Let us finish by discussing the original Bondal–Orlov reconstruction theorem in terms of the results we have shown so far.

Theorem 3.25. *Let X be a variety with ample (anti)canonical divisor, and let $\boxtimes$ be a tensor triangulated structure on $D^b(X)$ with unit $\mathscr{O}_X$. If $\mathrm{Spc}(\boxtimes)$ is a smooth projective space with ample (anti)canonical bundle and there is an equivalence $D^b(X) \simeq D^b(\mathrm{Spc}(\boxtimes))$, then $X \cong \mathrm{Spc}(\boxtimes)$.*

Proof. The only thing to note here is that since $\mathrm{Spc}(\boxtimes)$ has ample (anti)canonical bundle, ω_X has to be $\boxtimes$-invertible. Indeed, we recall that one can pick the equivalence $D^b(X) \simeq D^b(\mathrm{Spc}(\boxtimes))$ to send ω_X to $\omega_{\mathrm{Spc}(\boxtimes)}$ and then the assertion follows by applying Corollary 3.17 to $\mathrm{Spc}(\boxtimes)$ so that we obtain that $\mathrm{Pic}(D^b(X), \otimes^{\mathbb{L}}_X)$ is isomorphic via the assignment $\mathscr{L} \mapsto \mathscr{L}$ to a subgroup of $\mathrm{Pic}(D^b(X), \boxtimes)$. Since ω_X is $\boxtimes$-invertible, by Corollary 3.18 we obtain our result. $\square$

Remark 3.26. We need to explain our choice of hypotheses. First, the assumption that $\mathrm{Spc}(\boxtimes)$ is a smooth projective variety is necessary just as in the original Bondal–Orlov theorem formulation. We have added a couple more assumptions, however. We suppose that the (anti)canonical bundle of $\mathrm{Spc}(\boxtimes)$ is also ample to highlight the use of the monoidal structures in the theorem. This hypothesis is however not necessary as it can be directly deduced from the derived equivalence between the two spaces, just as in the original proof of Bondal and Orlov. Alternatively, we can formulate the theorem as follows:

Theorem 3.27. *Let X be a variety with ample (anti)canonical divisor, and let $\boxtimes$ be a tensor triangulated structure on $D^b(X)$ with unit $\mathscr{O}_X$. Suppose that $\mathrm{Spc}(\boxtimes)$ is a smooth projective space and that we have an equivalence $D^b(X) \simeq D^b(\mathrm{Spc}(\boxtimes))$. Then $X \cong \mathrm{Spc}(\boxtimes)$.*

Of more importance is perhaps the choice of unit, as we have seen that there are tensor triangulated category structures on the derived category of such a variety which will produce very different spaces under the Balmer reconstruction. This choice of unit allows us to keep some control in the classification of structures producing the same space.

A natural next step for future work would be to deal with the possible sort of objects which can be units for such a structure.

Remark 3.28. There is some nuance in the way in which Bondal–Orlov follows from our results as we make use of some important technical results from the original proof. We expect however that the discussion in this work has provided enough of a justification and motivation for looking at this problem in terms of monoidal structures.

We can close our discussion with the following theorem:

Theorem 3.29. *Let X be a smooth projective variety with ample* (*anti*)*canonical bundle. If $\boxtimes$ is a tensor triangulated category structure on $D^b(X)$ such that $\mathscr{O}_X$ is its unit and* $\mathrm{Spc}(\boxtimes)$ *is isomorphic to X, then $\boxtimes$ and $\otimes_X^{\mathbb{L}}$ coincide on objects.*

This however does not fully classify $\mathrm{TTS}_{\mathscr{O}_X}(D^b(X))$, since we require Balmer's spectrum to be a Fourier–Mukai partner, but there is no reason to expect in general a relationship between the derived category of the Balmer spectrum and the original triangulated category.

The lack of morphisms between a space X and the Balmer spectrum $\mathrm{Spc}(\boxtimes)$ for some tensor triangulated structure, and thus of functors between the derived categories of these two spaces, is one of the obstacles to being able to understand the possible structures $\boxtimes$.

Acknowledgements

The results in this work were obtained as part of the author's PhD thesis at the Laboratoire J.A. Dieudonné at the Université Côte d'Azur. The author would like to thank his advisor Carlos Simpson for his constant support and many discussions, and to Ivo Dell'Ambrogio and Bertrand Toën for their careful and valuable comments on the thesis manuscript. The PhD thesis was partially financed by the CONACyT-Gobierno Francés 2018 doctoral scholarship.

References

[Antieau 2013] B. Antieau, Reply to "Can we define the tensor product in the derived category $D^b_{\mathrm{coh}}(x)$ just from $D^b_{\mathrm{coh}}(x)$ in certain cases?", October 2013, available at https://mathoverflow.net/q/143661.

[Balmer 2002] P. Balmer, "Presheaves of triangulated categories and reconstruction of schemes", *Math. Ann.* **324**:3 (2002), 557–580. MR Zbl

[Balmer 2005] P. Balmer, "The spectrum of prime ideals in tensor triangulated categories", *J. Reine Angew. Math.* **588** (2005), 149–168. MR Zbl

[Balmer and Favi 2007] P. Balmer and G. Favi, "Gluing techniques in triangular geometry", *Q. J. Math.* **58**:4 (2007), 415–441. MR Zbl

[Beĭlinson 1978] A. A. Beĭlinson, "Coherent sheaves on $\mathbb{P}^n$ and problems in linear algebra", *Funktsional. Anal. i Prilozhen.* **12**:3 (1978), 68–69. In Russian; translated in *Funct. Anal. Appl.* **12**:3 (1978), 214–216. MR Zbl

[Bondal and Kapranov 1989] A. I. Bondal and M. M. Kapranov, "Representable functors, Serre functors, and reconstructions", *Izv. Akad. Nauk SSSR Ser. Mat.* **53**:6 (1989), 1183–1205. In Russian; translated in *Math. USSR-Izv.* **35**:3 (1990), 519–541. MR Zbl

[Bondal and Orlov 2001] A. Bondal and D. Orlov, "Reconstruction of a variety from the derived category and groups of autoequivalences", *Compositio Math.* **125**:3 (2001), 327–344. MR Zbl

[Boucksom et al. 2014] S. Boucksom, S. Cacciola, and A. F. Lopez, "Augmented base loci and restricted volumes on normal varieties", *Math. Z.* **278**:3-4 (2014), 979–985. MR Zbl

[Dell'Ambrogio 2016] I. Dell'Ambrogio, *Triangulated categories and applications*, habilitation, Université de Lille 1, 2016, available at https://idellambrogio.github.io/HDR-DellAmbrogio.pdf.

[Huybrechts 2006] D. Huybrechts, *Fourier–Mukai transforms in algebraic geometry*, Clarendon, 2006. MR Zbl

[Kawamata 2002] Y. Kawamata, "D-equivalence and K-equivalence", *J. Differential Geom.* **61**:1 (2002), 147–171. MR Zbl

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, I: Classical setting: line bundles and linear series*, Ergeb. der Math. (3) **48**, Springer, 2004. MR Zbl

[Liu and Sierra 2013] Y.-H. Liu and S. J. Sierra, "Recovering quivers from derived quiver representations", *Comm. Algebra* **41**:8 (2013), 3013–3031. MR Zbl

[Mukai 1987] S. Mukai, "Fourier functor and its application to the moduli of bundles on an abelian variety", pp. 515–550 in *Algebraic geometry* (Sendai, 1985), edited by T. Oda, Adv. Stud. Pure Math. **10**, North-Holland, 1987. MR Zbl

[Orlov 1997] D. O. Orlov, "Equivalences of derived categories and $K3$ surfaces", *J. Math. Sci.* **84**:5 (1997), 1361–1381. MR Zbl

[Rouquier 2003] R. Rouquier, "Catégories dérivées et géométrie algébrique", from the conference "Exposés à la semaine Géométrie algébrique complexe" (Luminy, 2003).

[Thomason 1997] R. W. Thomason, "The classification of triangulated subcategories", *Compositio Math.* **105**:1 (1997), 1–27. MR Zbl

Received September 23, 2023. Revised January 28, 2024.

ANGEL TOLEDO
LABORATOIRE JEAN ALEXANDRE DIEUDONNÉ
UNIVERSITÉ CÔTE D'AZUR
NICE
FRANCE

toledo@unice.fr

Guidelines for Authors

Authors may submit articles at msp.org/pjm/about/journal/submissions.html and choose an editor at that time. Exceptionally, a paper may be submitted in hard copy to one of the editors; authors should keep a copy.

By submitting a manuscript you assert that it is original and is not under consideration for publication elsewhere. Instructions on manuscript preparation are provided below. For further information, visit the web address above or write to pacific@math.berkeley.edu or to Pacific Journal of Mathematics, University of California, Los Angeles, CA 90095–1555. Correspondence by email is requested for convenience and speed.

Manuscripts must be in English, French or German. A brief abstract of about 150 words or less in English must be included. The abstract should be self-contained and not make any reference to the bibliography. Also required are keywords and subject classification for the article, and, for each author, postal address, affiliation (if appropriate) and email address if available. A home-page URL is optional.

Authors are encouraged to use LaTeX, but papers in other varieties of TeX, and exceptionally in other formats, are acceptable. At submission time only a PDF file is required; follow the instructions at the web address above. Carefully preserve all relevant files, such as LaTeX sources and individual files for each figure; you will be asked to submit them upon acceptance of the paper.

Bibliographical references should be listed alphabetically at the end of the paper. All references in the bibliography should be cited in the text. Use of BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 327 No. 2 December 2023